PERGAMON

Vision Research 38 (1998) 3569–3582

Vision Research

Position displacement, not velocity, is the cue to motion detection of second-order stimuli

Adriane E. Seiffert *, Patrick Cavanagh

Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA

Received 8 August 1997; received in revised form 29 December 1997

Abstract

Motion detection can be achieved either with mechanisms sensitive to a target's velocity, or sensitive to change in a target's position. Using a procedure to dissociate these two provided by Nakayama and Tyler (Vis Res 1981;21:427–433), we explored detection of first-order (luminance-based) and various second-order (texture-based and stereo-based) motion. In the first experiment, observers viewed annular gratings oscillating in rotational motion at various rates. For each oscillation temporal frequency, we determined the minimum displacement of the pattern for which observers could reliably see motion. For first-order motion, these motion detection thresholds decreased with increasing temporal frequency, and thus were determined by a minimum velocity. In contrast, motion detection thresholds for second-order motion remained roughly constant across temporal frequency, and thus were determined by a minimum displacement. In Experiment 2, luminance-based gratings of different contrasts were tested to show that the velocity-dependence was not an artifact of pattern visibility. In the remaining experiments, results similar to Experiment 1 were obtained with a central presentation of a linear grating, instead of an annular grating (Experiment 3), and with a motion discrimination (phase discrimination) rather than motion detection task (Experiment 4). We conclude that, within the ranges tested here, second-order motion is more readily detected with a mechanism which tracks the change of position of features over time.

1. Introduction

One can infer that motion has occurred in a visual scene by detecting the change in the position of elements in the scene. For example, one can be aware of the slow continuous movement of the setting sun simply by watching it (with dark sunglasses) over a long period of time. Perhaps because of the importance of information about object motion, for navigation and for image segmentation, a more direct derivation of motion is also performed by the visual system. Studies have shown that motion is computed directly, by a 'low-level' motion system (see Nakayama [2] for review) based on the direction selectivity of neurons in the primary [3] and supplementary visual areas [4,5]. Perceptual studies have supported this distinction between low-level motion analysis and 'high-level' position tracking in many ways. For example, Nakayama and Tyler [1] demonstrated that the low-level system is sensitive to the velocity of motion, as opposed to the distance travelled. Others have revealed that the perception of a change in position is not even necessary for motion perception to occur; the 'fine grain motion illusion' shows that motion can be perceived between two points which are not spatially resolvable [6,7]. In addition, it has been shown that motion can be seen in a two-frame sequence where one of the individual frames is below pattern detection threshold [8]. This dissociation is also clear to an observer of the motion after-effect because the perception of visual motion is not accompanied by a perceived translation of the features in the inducing pattern [9].

Although this evidence has established the existence of low-level motion analysis, it does not rule out motion perception based on noticing changes in position. Observers can localize objects with high degree accuracy in very brief presentations [10,11]. It is reasonable that changes in position over short time intervals could be tracked by a process independently of the low-level motion system. This alternative motion system may be especially important if low-level analysis is unable to derive a reliable motion signal from a given stimulus.

* Corresponding author. E-mail: aeseiffe@wjh.harvard.edu.

0042-6989/98/$19.00
PII: S0042-6989(98)00035-2

Many researchers in the field of motion perception have proposed that two independent systems exist. Anstis [12] distinguished two motion mechanisms most like those outlined in the previous paragraph. His 'system 1' is a network of hard-wired motion detectors that are based on correlational mechanisms, similar to the low-level motion mechanism. 'System 2' or the 'cognitive' motion system codes motion after the extraction of edges or forms in a scene, thereby solving the motion correspondence problem by tracking the position of visible features over time. He suggested that system 2 was responsible for detecting motion of edges defined by texture and/or binocular depth. These stimuli have been called 'second-order' motion stimuli, because motion is produced by two areas with the same mean luminance, but with different spatial, temporal or ocular distributions of luminance [13]. The name 'second-order' refers to the fact that, first, two samples need to be taken to determine the structure of the stimulus. The luminance of two points of a pattern are needed to define its contrast, and two ocular images define disparity. Then, these can to be tracked over time for motion to be identified. Correspondingly, for 'first-order' stimuli, only one sample (e.g. luminance) needs to be tracked over time [13]. In more recent years, the distinction of two motion analyzers has revolved around these two stimulus categories.

Most models, however, posit that second-order motion is perceived via a system very similar to, or identical to, the low-level motion system [14–17]. Evidence supporting these models comes from a variety of comparisons between motion perception of first-order and second-order stimuli, showing that the two have similar profiles along many dimensions [13,18–21]. A parallel line of research, however, has shown that motion perception produced by second-order stimuli differs from that produced by first-order stimuli [22–24]. Much of this research has concluded that two low-level motion systems exist: one for detecting luminance-based motion, and another for detecting second-order motion. However, none of these studies has directly tested whether the second-order motion mechanism shows velocity-based or position-based characteristics. In this paper, we show that second-order motion is not detected with a velocity-sensitive mechanism, under conditions where first-order motion is. The results from these experiments suggest that second-order motion is based on detecting position change, consistent with a feature-tracking mechanism.

1.1. The paradigm

In this study, velocity and position-change sensitivity were dissociated with a paradigm provided by Nakayama and Tyler [1]. A stimulus oscillated sinusoidally with two independent parameters: temporal frequency, and displacement. In this paper, the term 'displacement' will be used to denote the total spatial shift of the pattern as shown in Fig. 1A; Space-time plots of oscillating motion in Fig. 1B; show how peak velocity and displacement can be independently varied. Notice that two patterns with the same total displacement, but different frequencies have different peak velocities (B1 and B2). Two patterns with the same velocity, but different frequencies have different displacements (B2 and B3). Thus one can determine if motion detection is sensitive to velocity or displacement by measuring displacement thresholds over a range of temporal frequencies. If motion is detected with a position-based system, then displacement thresholds should remain constant across temporal frequency; that is, B1 and B2 would be above threshold, and B3 would be below threshold. However, if motion is detected with a velocity-based system, then displacement thresholds should fall with increasing temporal frequency; that is, B1 would be above threshold, and B2 and B3 would be below threshold.

Using this procedure, we examined the detection of first and second-order motion. Consistent with results of Nakayama and Tyler [1], we expect that first-order

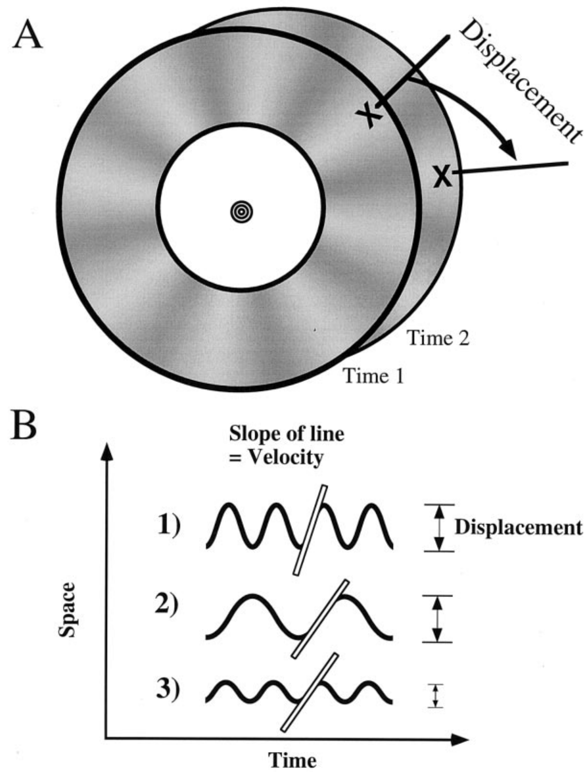

Fig. 1. (A) Radial luminance grating similar to that used in this experiment, at two points in time. As the pattern oscillates, points on the pattern (such as that marked by X's) transverse the path shown. 'Displacement' refers to the total extent of spatial shift. (B) Space-time plots of a point on the pattern such as that marked in (A). The slope of the white, oblique lines indicate the peak velocity of the motion. See text.

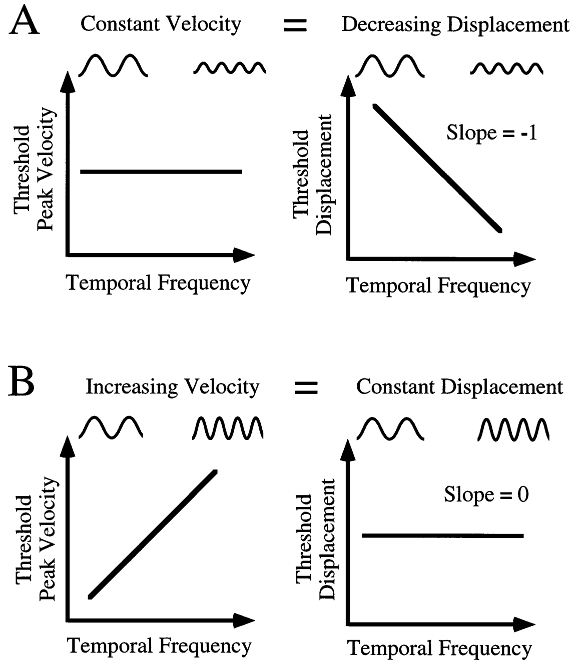

Fig. 2. (A) The thresholds across temporal frequency of a system sensitive to a minimum peak velocity. Thresholds trace a curve of constant peak velocity (left), therefore the corresponding displacements fall with a slope of − 1.00 (right) across temporal frequency. Wavy lines above each graph are space-time plots as in Fig. 1B. (B) If thresholds are determined by a minimum displacement, then the peak velocity of the motion rises (left) with temporal frequency to maintain this constant displacement (right).

motion is detected by a low-level motion system sensitive to a minimum velocity (Fig. 2A, left). Thus, displacement thresholds should decrease with a slope of − 1.00 as the temporal frequency increases (Fig. 2A, right). In contrast, if second-order motion is detected with a mechanism sensitive to changes in position, not velocity, it should show no variation in displacement thresholds over temporal frequency (Fig. 2B, right). Regardless of the speed, a minimum displacement will be necessary to perceive motion.

2. Experiment 1: detection of first-order versus second-order motion

2.1. Subjects

The first author (AES) and five other observers, naive to the purpose of the experiment, served as subjects for this study. All had normal or corrected to normal acuity and were experienced psychophysical observers.

2.2. Stimuli

Stimuli were radial gratings (similar to Fig. 1A) displayed in an annulus (3.6–8.1° of eccentricity) about a fixation bullseye (0.5° diameter) on a black (0.51 cd/m^2) background. Mean luminance of the display annulus was 39.8 cd/m^2, and the viewing distance was 57 cm, set by a headrest. Eight cycles of the grating appeared over the annulus, so the spatial frequency of the pattern varied from approximately 0.37 cycles/degree (cpd) at the inner edge to 0.164 cpd at the outer edge. Each grating oscillated by rotating about the central fixation point with a sinusoidal temporal course as depicted in Fig. 1B.

Five conditions in this experiment included one first-order and four second-order motion stimuli. The first-order motion stimulus (luminance condition) was a luminance sine wave grating (Fig. 3A). The reported contrast of this stimulus was the Michelson contrast. The four second-order motion stimuli were as follows.

The CM rings texture was a sinusoidal contrast modulation of a pattern of 30 concentric circles (Fig. 3B) which alternated light and dark. Each circle was approximately 0.15° of visual angle in width, thereby matching luminance changes in the direction of motion (circumferentially) in every 0.3° of visual angle. Notice that the contrast between the light and dark circles of the pattern was modulated over the annulus between 0% contrast and some maximum value. This maximum contrast defined the contrast of the CM rings texture.

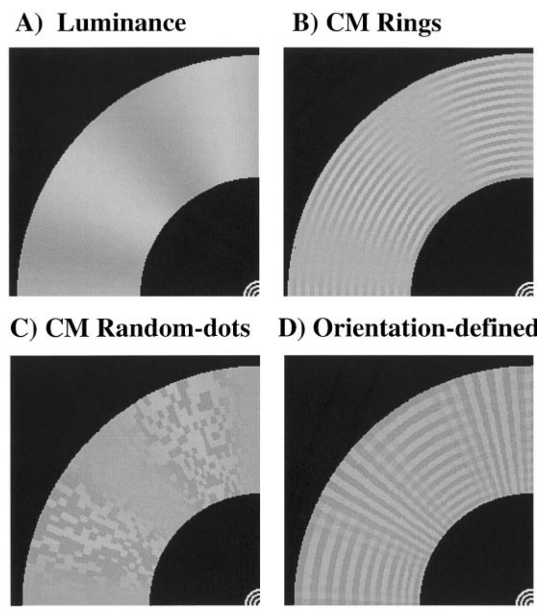

Fig. 3. Upper right quadrants of the luminance grating and the three textured gratings used in Experiment 1. The contrast of each of the patterns shown is much higher than the actual stimuli used. (A) Luminance condition: a luminance-based sine wave radial grating. (B) Contrast-modulated rings texture: a sinusoidal contrast modulation of a set of concentric rings. (C) Contrast-modulated random-dots texture: a sinusoidal contrast modulation of a pattern of dynamic random dots, replaced pseudo-randomly at a rate of 33 Hz. The size of the random dots is enlarged here for demonstration purposes. See text for actual size. (D) Orientation-defined texture: a sinusoidal variation in contrast of the concentric rings pattern (as in B) 180° out of phase with a sinusoidal variation in contrast of a pattern of radial lines. Note: that the contrast envelopes moved while the lines making up the pattern remained stationary.

The CM random-dot texture was a sinusoidal contrast modulation of a pattern of random dots, depicted in Fig. 3C. The radial grating was divided into arc segments which were randomly assigned to light or dark. Each segment varied in actual size depending on its eccentricity; the smallest segments were approximately $0.0565 \times 0.0555°$ and the largest were approximately $0.127 \times 0.122°$ of visual angle. The contrast of this pattern was defined in the same way as the CM rings texture. Each segment was replaced pseudo-randomly at a rate of 33 Hz making the display resemble a wheel of made out of eight twinkling spokes.

The orientation-defined texture was a sinusoidal variation of the contrast of a pattern of rings varied 180° out of phase with the contrast of a pattern of radial lines (Fig. 3D). To create the second-order motion of the orientation-defined texture, while keeping the first-order signals balanced, the pattern of dark and light lines defining the orientation never moved. Only the contrast of both patterns changed over space and time. During rotation, a particular segment of the image would change from the rings pattern to the crossed pattern to the radial pattern through a constant gradual change in contrast. Thus, it was the sections of the same orientation that defined the motion, rather than any of the luminance edges. The contrast of this orientation-defined grating was defined as the maximum contrast of each pattern.

Finally, the stereo-defined grating was a sinusoidal corrugation in depth of a field of dynamic random dots ($0.006 \times 0.006°$ square patches) flickering at 30 Hz, presented on a background of static random dots with intermediate depth. Note that the corrugated surface rotated in the plane perpendicular to the line of sight and, therefore, was not moving in depth. The contrast of the stereo-defined grating was defined as the disparity difference between the most crossed and most uncrossed (closest and farthest) points in the sinusoidal pattern. Contrast of 100% was arbitrarily set to represent a disparity difference of about 0.1°.

2.3. Apparatus

There were two sets of apparatus used in these experiments. On the first apparatus set, stimuli were generated on a Power Macintosh 7500/100 and displayed on an Apple High-Resolution Monochrome monitor. The 640×480 pixel video signal with the help of an ISR Video Attenuator [25], had 12 bits of intensity resolution calibrated for linearity, with a refresh rate of 67 Hz. The large contrast range of the monitor allowed for a close approximation of a sine wave, and thus small displacements, to be displayed.

The second apparatus generated the stereo-depth display with a Macintosh IIci computer and a Datacube image processor, on a Mitsubishi Diamond Scan color screen. The image was drawn interlaced at 60 Hz. A Tektronix Stereoscopic modulator polarizing screen created the illusion of stereo-depth by presenting alternate fields of the display to the left and right eyes of the observers. Subpixel resolution was attained by modulating the contrast of pixels at light–dark borders, so as to modulate the perceived position of each border [26,27]. All observers reported a clear sense of depth from the image.

Table 1
Contrast threshold levels for each observer for each stimulus type

Condition	Observer	Contrast threshold (%)	S.D.
Annulus			
Luminance	AES	0.67	0.09
	CMH	0.60	0.12
	CDD	1.08	0.33
	DVW	1.03	0.22
	MW	0.82	0.21
	FT	0.75	0.15
	JDW	0.80	0.09
CM rings	AES	1.58	0.37
	CMH	1.13	0.08
	CDD	1.30	0.35
	DVW	1.25	0.22
CM random-dots	AES	1.98	0.28
	CMH	2.05	0.37
Orientation-defined	AES	1.30	0.29
	CMH	1.90	0.28
	MW	2.05	0.21
Stereo-defined	AES	3.50[a]	0.56
	CMH	3.50[a]	0.56
	FT	1.50[b]	0.58
Linear			
Luminance	AES	1.94	0.62
	CDD	1.34	0.50
	DVW	0.68	0.26
CM rings	AES	1.53	0.37
	CDD	1.35	0.48
	DVW	0.99	0.15
CM random-dots	AES	1.55	0.13
	DVW	1.48	0.20

All stimuli were presented at 10 times these threshold values. [a] Approximately 12 arc s disparity difference. [b] Approximately 6 arc s disparity difference.

2.4. Procedure

An experimental session consisted of two data-collecting procedures. First, at the beginning of each session, contrast thresholds for pattern detection were estimated with the method of adjustment. The contrast for each different type of grating was defined as described above. Each grating was shown oscillating at 2 Hz with a displacement of 180°. After a button press, observers moved a mouse from a starting point (either zero or 25% contrast) to their threshold level. Four settings were made for each display type, for each observer. These threshold levels are listed in Table 1 for all conditions and all observers. The contrast of the stimulus was then fixed at ten times this threshold level.

In the second part of the session, displacement thresholds were estimated with the yes–no method of constant stimuli for a wide range of temporal frequencies. Each trial began with the presentation of the blank display annulus (with no pattern) and the fixation point. After a 1.5 s interval, the sinusoidal pattern appeared and immediately began to oscillate. The spatial phase of the pattern at onset was chosen randomly to decrease inter-trial effects of contrast adaptation. Observers were instructed to judge whether or not the display was moving, and were asked to keep a high criterion for the procedure. The experimenter told subjects to only say 'yes' if they were sure that the stimulus was moving. Observers were asked to maintain fixation on the central bullseye throughout a trial, and all reported doing so without difficulty for all conditions. Each observer completed two blocks of trials consisting of 64 random presentations of eight different displacements of the same display pattern at the same temporal frequency. The range of displacements tested within a block varied in eight equal steps from zero (the no motion, 'catch' trials) to a 'high-displacement' value, which was chosen based on pilot data for each subject. If the results from a single block were uninterpretable because the 'high-displacement' value was too high or too low, or if subjects reported seeing motion during the catch trials, the block was repeated in a more appropriate range.

At the end of the experiment, if the subjects did not spontaneously report a description of their subjective experience in the two conditions, the experimenter asked them to report their experience or strategy. The experimenter also asked them whether they experienced 'seeing motion' or 'noticing a position change'.

2.5. Results

For each observer, each pattern and each temporal frequency, the number of positive responses (indicating motion was detected) were plotted as a function of the displacements tested. Cumulative normal functions were fitted to these data as shown in Fig. 4. Displacement threshold values were defined as the displacement corresponding to the point on the curve at the 0.50 proportion of motion detections. Because spatial frequency of the pattern depended on eccentricity, displacement thresholds were reported in terms of the degrees of phase shift of the grating. A displacement of 360° denotes that the grating moved exactly one spatial cycle. Error bars shown in all the following data graphs are 95% confidence intervals based on these curve fits.

Displacement thresholds as a function of temporal frequency for each observer are shown in Figs. 5–8. In every panel, the solid circles denote the first-order condition, the open squares denote the second-order condition and the dotted line shows the slope of − 1.00 which indicates the constant peak velocity line on these plots. For all observers, the displacement thresholds found for the luminance-based motion stimulus decreased with increasing temporal frequency at a rate comparable to that predicted by sensitivity to velocity. Because thresholds fell with a slope of approximately − 1.00, the stimuli represented by each point on the curve had about the same peak velocity, which varied from observer to observer. As shown in Table 2, these peak velocities averaged across observers were around 0.015° of visual angle per second for the fastest moving part of the rotating grating. These velocities were estimated by determining the speed at which the outermost part of the grating was moving at displacement threshold.

In contrast, the displacement thresholds for the second-order motion stimuli did not vary systematically as a function of temporal frequency. In fact, thresholds from all three texture-defined patterns and the stereo-defined pattern were roughly constant over the same temporal frequency range that the first-order patterns dropped significantly. Second-order displacement thresholds hovered around 10% of a cycle, which is a displacement of about 16 arc min at the inner edge and 37 arc min at the outer edge. For all second-order motion stimuli tested, then, motion detection thresholds were limited by a minimum displacement rather than a minimum velocity.

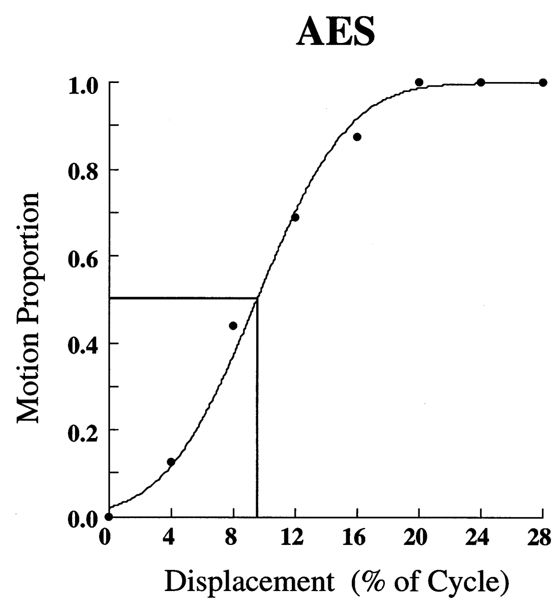

Fig. 4. Proportion of trials for which subject AES responded that the stimulus was moving, plotted as a function of the amount of displacement of the grating. The stimulus in this case was the Contrast-modulated rings texture display, moving at 0.1 Hz. The 0.5 proportion threshold was estimated at 9.53% of a cycle, or 34.3° of phase.

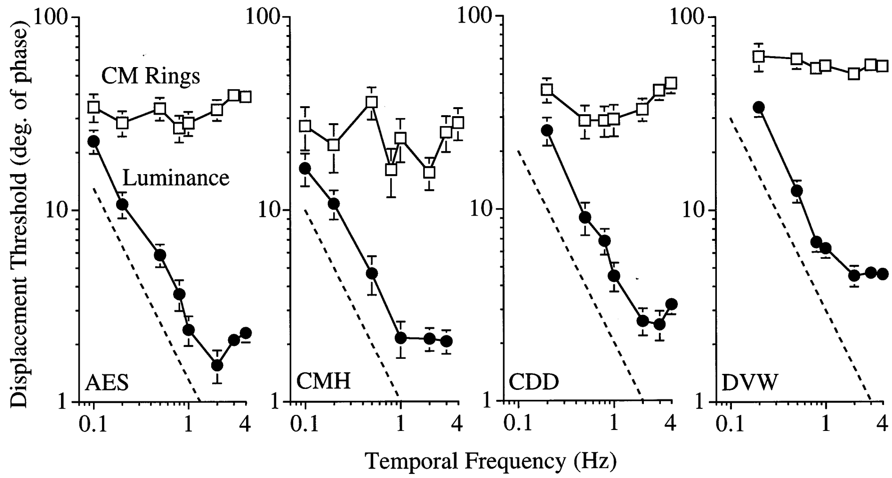

Fig. 5. Experiment 1: displacement thresholds as a function of temporal frequency for four observers. The contrast-modulated rings condition thresholds (□) did not vary systematically with temporal frequency, but the luminance condition thresholds (●) decreased as temporal frequency increased. The dashed line in each figure shows the slope of − 1.00 that represents stimuli with the same peak velocity across temporal frequencies. Notice that the luminance thresholds fall with a slope close to this predicted value until about 2 Hz, and then fall higher than the prediction. This discrepancy is discussed further in Experiment 4.

Although subjects indicated on which trials they detected the motion of the stimulus, this detection did not require that subjects experience visual motion. In fact, two of the subjects spontaneously mentioned that, for some of the blocks of trials, they were not 'seeing motion' as much as they felt they were 'noticing a change' in the position of the features in the stimulus. As it turned out, the conditions that were reported as 'noticing a change' were the very slow oscillation temporal frequencies (e.g. 0.1 and 0.2 Hz) for both types of stimuli (first and second-order patterns). For the higher rates (1.0 and 2.0 Hz), these subjects reported 'seeing motion' for both types of stimuli. All of the other of the subjects gave similar descriptions of 'seeing motion' and 'noticing a position change' across conditions.

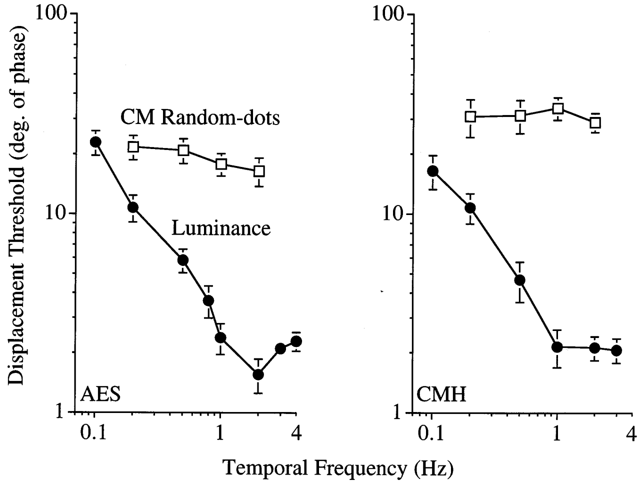

Fig. 6. Experiment 1: contrast-modulated random-dot condition (□) plotted with the luminance condition (●) replotted from Fig. 5.

2.6. Discussion

Similar to Nakayama and Tyler [1], we find detection of luminance-based motion was determined by a minimum velocity. Replotted in Fig. 9 (the dashed line) are the results Nakayama and Tyler's Experiment 1 (Fig. 3) from testing the displacement thresholds for an oscillating, random dot display. Although the scales are very different, the subject averages from the luminance condition of the present study (Fig. 9, solid line, circles) follow a pattern over temporal frequency very similar to the Nakayama and Tyler results. However, the subject averages from each of the second-order conditions (Fig. 9, solid line, squares) remain roughly constant over temporal frequency, showing sensitivity to a minimum displacement, not velocity.

Nakayama and Tyler [1] concluded that velocity sensitivity was higher than position sensitivity for their random dot display because position cues were reduced by making the adjacent sections of random dots move differentially. In the present work, the stimuli were sine waves, so both position cues and velocity cues were readily available. In fact, because the spatial and temporal frequencies were equal, the first and second-order stimuli had equivalent position and velocity information. Despite this, observers were sensitive to different aspects of these two stimuli. Observers were sensitive to the velocity of the first-order motion, and the position-change of the second-order motion. Interestingly, the displacement thresholds in the second-order conditions were always higher than the thresholds in the first-order condition. This lack of sensitivity to velocity indicates that low-level motion systems could not produce reliable motion signals from the shifts in these second-order patterns.

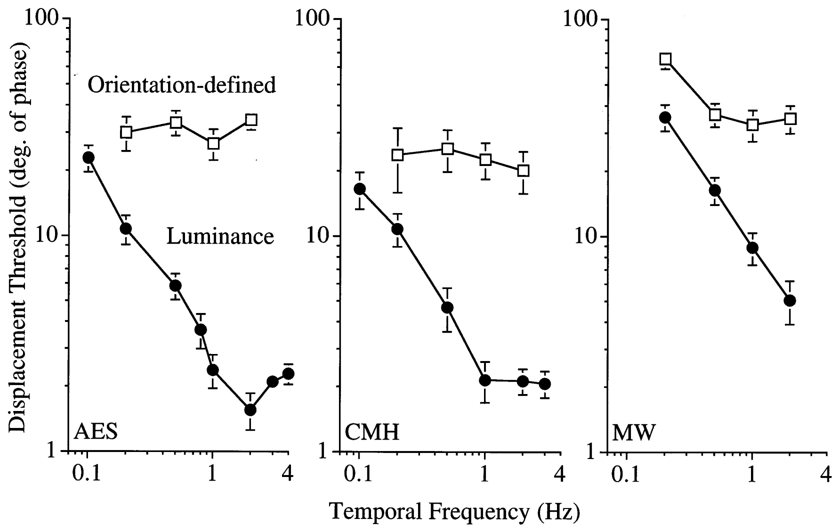

Fig. 7. Experiment 1: orientation-defined texture condition (□) plotted with the luminance condition (●) as in Fig. 5 for AES and CMH.

Although thresholds were determined by different factors, subjects reported having the same (or at least very similar) experience of 'seeing motion' in the second-order conditions as they did in the first-order conditions. Note that the inference that observers are detecting motion with a position-based mec is not inconsistent with the statement that they are experiencing visual motion. Although detecting position-change can be a consciously chosen strategy, it may also be a more automatized mechanism. Because the output of such a system would be functionally equivalent from the output of a velocity-tuned motion system, the two may produce the same mental state within the observer.

The conclusion that second-order motion is detected with a position-sensitive mechanism is contrary to the notion that contrast-modulated motion is detected via a system similar to luminance-based motion after a contrast full-wave or half-wave rectifying first stage [14]. Regardless of the details of the signal pre-processing stage, a model of motion processing based on this type of low-level motion analysis would predict velocity sensitivity. Rather, these results are consistent with the idea that second-order motion is detected by a mechanism similar to the feature matching or feature tracking mechanisms proposed previously [12,28,29]. Before making this conclusion, however, some other issues regarding the nature of the perception of motion in this experiment need to be addressed.

Due to our choice of stimuli and task, the decision of whether or not the stimulus moved on each trial could have been based, at times, on factors other then the perception of motion. For example, in the luminance condition at the very low temporal frequencies, observers reported that the task was especially difficult because the stimuli seemed to fade from view. Therefore, the degree of apparent contrast, rather than motion, may have determined observers' responses. For temporal frequencies higher than 2 Hz, observers reported seeing little rotational motion, but rather reported detecting flicker. Two additional experiments described below evaluate the role of these alternative cues.

3. Experiment 2: luminance at different contrasts

All subjects in Experiment 1 reported difficulty in performing the motion detection task for the luminance-based motion display, because the luminance grating slowly disappeared. Because the display was a low spatial frequency, sinusoidal luminance modulation at a moderately low contrast, adaptation to the stimulus occurred relatively quickly after display onset, especially for the lower temporal frequencies. It is possible then, that the increase in displacement threshold at lower temporal frequencies was due to a reduction in visibility of the stimulus pattern, rather than a reduc-

Table 2
Peak velocities estimated from the luminance condition for each observer in Experiment 1

Observer	Mean peak velocity (°/s)	S.D.
AES	0.0099	0.0015
CMH	0.0094	0.0038
CDD	0.0186	0.0016
DVW	0.0254	0.0051
Average	0.0154	0.0033

Values were obtained for each point on the curve between 0.2 and 2.0 Hz with the following equation: Peak velocity = displacement/2 × temporal frequency where the displacement value was determined in degrees of visual angle for the maximum distance the pattern traversed.

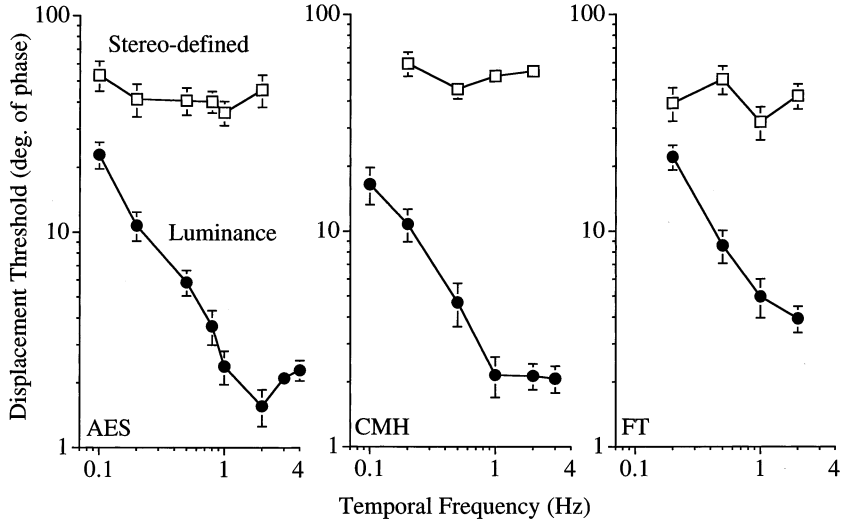

Fig. 8. Experiment 1: stereo-defined texture condition (□) plotted with the luminance condition (●) as in Fig. 5 for AES and CMH.

tion in motion detection per se. Experiment 2 tested displacement thresholds for the luminance-based motion stimulus with a much higher contrast to eliminate visibility difficulties.

3.1. Method

Two of the five subjects that participated in Experiment 1 also participated in Experiment 2. In addition, one other observer, also naive to the purposes of the experiment, was recruited for this experiment. The stimuli, and procedure used for this experiment was identical to the luminance condition of Experiment 1, except the grating was displayed at two different contrast levels.

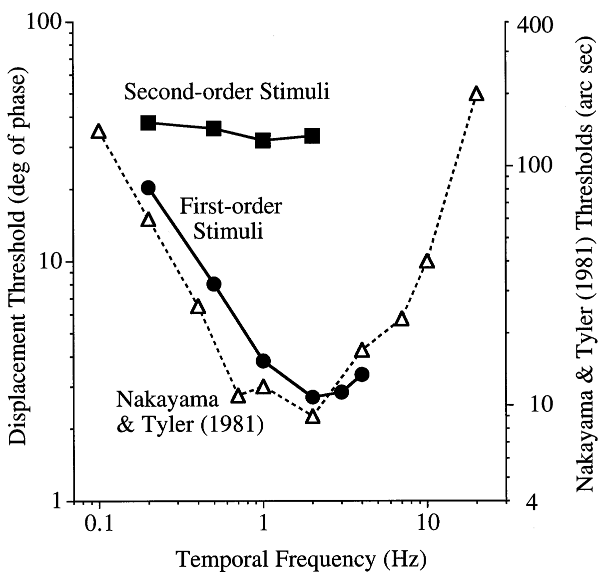

Fig. 9. Summary of present results and comparison with Nakayama and Tyler [1]. Results of Experiment 1 averaged across subjects are plotted with solid lines with respect to the left Y-axis. Only those points common to all subjects who participated in that condition are shown. With respect to the right Y-axis, open triangles and dashed line replot the results from one observer (CWT) from Nakayama and Tyler's Experiment 1 (Fig. 3).

3.2. Results and discussion

Displacement thresholds as a function of temporal frequency for luminance-based motion stimulus at different contrasts are shown for each observer in Fig. 10. Controlling for pattern visibility did not change the pattern of results for displacement thresholds. Whether the contrast was approximately ten or forty times detection threshold, a linear decrease of displacement threshold occurred over increasing temporal frequency. In all cases, the slopes of the displacement thresholds were significantly below zero, and were good approximations of the expected slope of -1.00. The main effect of increasing the contrast was to lower all the displacement threshold levels. Therefore, increasing visibility did have a marked effect on performance, but did not alter the dependence of motion detection thresholds on a minimum velocity.

4. Experiment 3: discrimination task

All of the previous studies reported here have relied upon observers to judge whether or not motion is present in each display. One potential problem with this procedure is that subjects may be responding positively on the basis of some other cue besides velocity or position-change, such as flicker. It is important, then, that the same type of experiment be performed with a test that isolates motion sensitivity from flicker sensitiv-

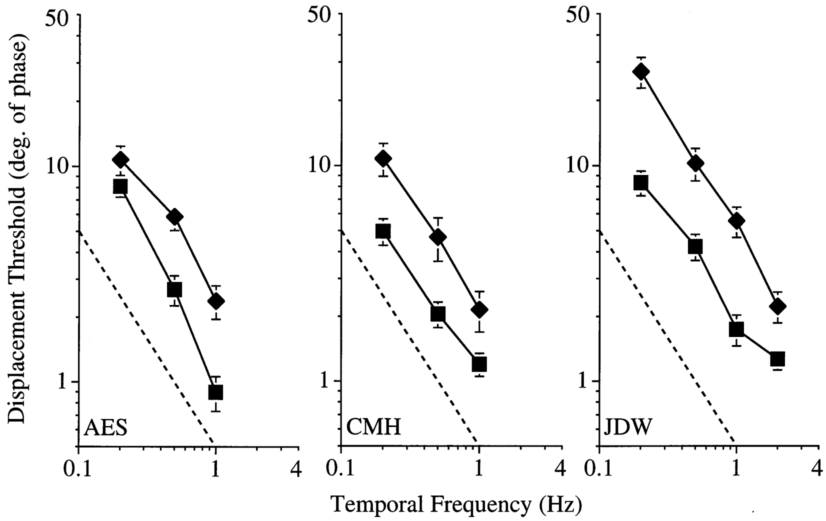

Fig. 10. Experiment 2: displacement thresholds for oscillating luminance gratings of two different contrasts. Gratings of contrast ten times detection threshold (◆) and gratings of contrast 40 times threshold (■) both produce thresholds with a slope close to the predicted − 1.00 value. Dashed lines show the slope of − 1.00 for each plot.

ity. To accomplish this, we divided the display in half which allowed the two parts of the stimulus to oscillate independently. Observers were tested on their ability to discriminate whether the two halves were oscillating temporally in phase or out of phase.[1] Detection of flicker would not allow observers to perform this task, as all phase information would be lost; subjects can only respond correctly if they detect the direction of motion of both halves of the display. Thus, the results from this experiment isolate motion sensitivity from sensitivity to flicker.

4.1. Methods

The displays used in this experiment were very similar to those used in Experiment 1. The display annulus was drawn with a 1.8° wide strip taken out of the top and bottom, so that each side of the display contained a semi-annulus section. Each of the two sections contained gratings with randomly selected spatial phase. On half the trials, the two gratings moved in phase, such that the whole display looked like a rotating grating with parts occluded, and on the other half of the trials, the gratings moved out of phase to look more like the motion of flapping wings. The gratings were in all other ways identical to those used in the luminance and CM rings conditions of Experiment 1. The procedure for this experiment was identical to the previous studies, except observers responded with a button press to indicate whether the display was moving in phase or out of phase.

[1] 1 We thank Edward Adelson for suggesting this discrimination task.

4.2. Results and discussion

Displacement thresholds as a function of temporal frequency are shown for each observer in Fig. 11. Notice, first, that the pattern of results is consistent with that obtained in the previous experiments. Namely, the luminance grating elicited thresholds that fell with increasing temporal frequency, whereas the CM rings grating thresholds were roughly constant across temporal frequency. Second, note that the CM rings thresholds for the highest temporal frequencies tested (4 Hz and even 3 Hz for subject CDD) are much higher than for the other temporal frequencies. This deviation was not found in the previous experiments, indicating that for these rates, it was flicker detection, rather than motion detection, which determined the thresholds in previous studies. In addition, in the luminance condition, notice that for these rates the data points deviate from the predicted slope of − 1.00 in all of the experiments reported here. The deviation from the prediction in the luminance case is most likely also due to degradation of motion perception at these rates

The last two experiments buttress the conclusions from the first experiment and run contrary to existing literature regarding the nature of the process detecting second-order motion. There are two main differences between this study and previous work that might have created this discrepancy. One difference is in location of the stimuli. Many studies of second-order motion have used central presentation rather than perifoveal presentation, because previous reports have found that second-order stimuli presented outside the fovea are more difficult to detect [13,24]. In fact, Pantle [24] has suggested that second-order stimuli are perceived qualitatively different from first-order stimuli by peripheral

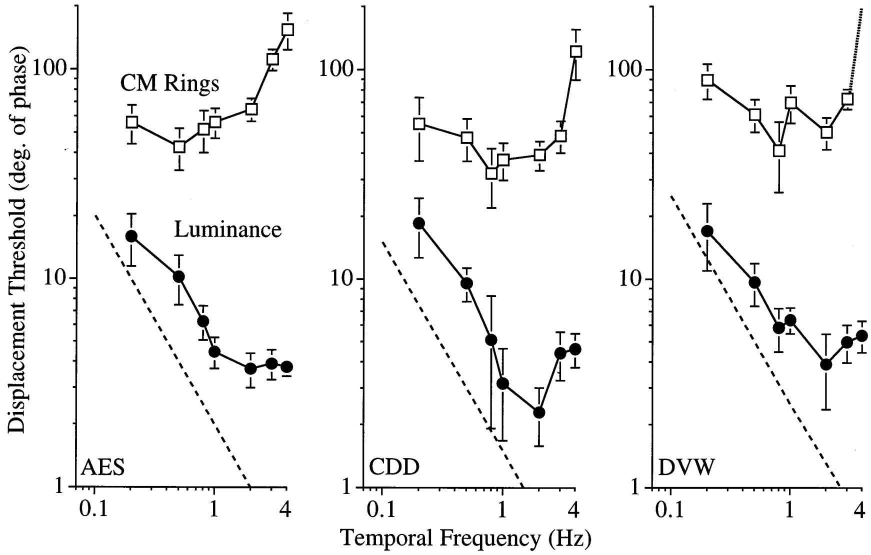

Fig. 11. Experiment 3: displacement thresholds determined with the motion discrimination task. As in previous experiments, thresholds for luminance gratings (●) fell with a slope of about − 1.00 until around 3 Hz. Similarly, thresholds for CM rings gratings (□) were roughly constant, until about 3 or 4 Hz. Dashed lines show the slope of − 1.00 in each graph. Dotted line in DVW graph indicates that the threshold for 4 Hz was too high to be measured.

vision. Therefore, different results may be obtained from this paradigm if the stimuli were centrally presented. The second difference is that the present work used radial gratings, rather than linear gratings. Although it is unclear at this time why this difference in presentation would produce such a difference in the results, it remains a discrepancy to be addressed.

5. Experiment 4: foveal presentation of linear gratings

Using the same paradigm as in previous experiments, this experiment tested foveated linear gratings with first or second-order characteristics. Efforts were made to produce stimuli in this experiment similar in all other ways to the those used in Experiment 1. Gratings filled a $12 \times 17.5°$ field of view, and no fixation marker was presented. This was done to decrease the position cues between the bars of the stimulus and the ends of the display or the edges of the fixation mark, because the previous experiments had no such cues available. The spatial frequency of the gratings were within the range tested in Experiment 1, but the texture elements making up the second-order displays were much smaller, in order to attempt to control for the increase in visual acuity in the fovea relative to the periphery. In this way, the stimuli were constructed so that the same cues were available to the observer in this experiment as in Experiment 1, and only the retinal location of stimulation was different.

5.1. Method

A $12 \times 17.5°$, vertical sine wave grating of 0.23 cpd oscillated horizontally in the center of the screen. No fixation bullseye or point was present in the displays, and eye movements were not constrained. Observers were asked to keep their eyes roughly in the center of the display so that accretion and deletion cues of the pattern at the display edge were not foveated.

Three conditions were tested: (1) The luminance condition was a simple luminance-based sine wave grating. (2) The CM lines texture condition was analogous to the CM rings condition of Experiment 1, such that a pattern of horizontal lines of alternating luminance was sinusoidally contrast modulated to produce a vertical second-order grating. Because of the central presentation, the spatial frequency of the pattern of horizontal lines was increased to make the luminance changes more closely balanced. Lines in this pattern were approximately 0.068° wide, so that luminance was balanced within each 0.136° vertically. (3) The CM random-dots condition was analogous to that from Experiment 1, again, with smaller sized elements ($0.070 \times 0.073°$). The procedure was identical to Experiment 1. Three subjects from Experiment 1. participated in this experiment; two (CDD and DVW) participated in Experiment 3 before doing Experiment 1. Observer CDD did not participate in the CM random-dots condition.

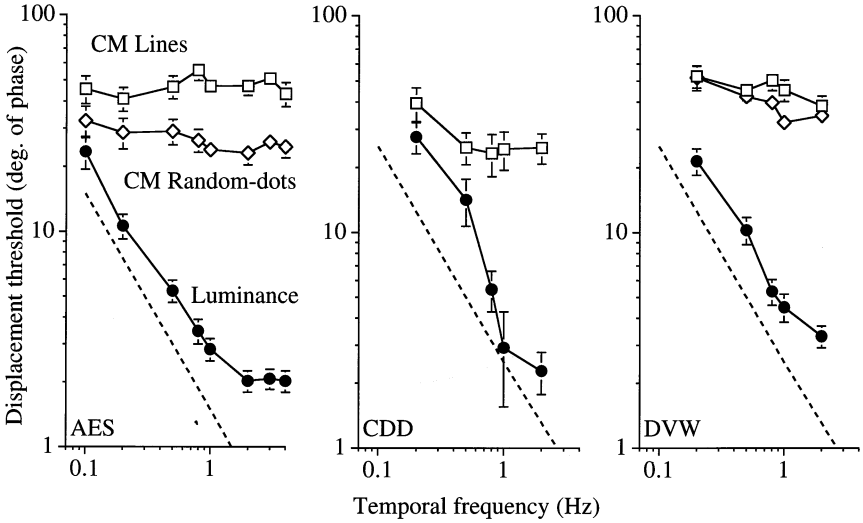

Fig. 12. Experiment 4: displacement thresholds for linear gratings presented foveally. The pattern of results is identical to that obtained with peripheral, radial gratings in Experiment 1. luminance grating thresholds (●) decrease with increasing temporal frequency, but the second-order gratings, CM lines (□) and CM random-dots (◇), remain roughly constant. Dashed lines show the slope of − 1.00. Thresholds for the CM random-dots condition were not obtained for subject CDD.

5.2. Results and discussion

Displacement thresholds as a function of temporal frequency are shown in Fig. 12, separately for each observer. Replicating the first experiment, the luminance condition thresholds fell with a slope of approximately − 1.00 between a temporal frequency of 0.2 and 2.0 Hz, indicating that a velocity-sensitive mechanism was employed. In addition, both second-order patterns elicited displacement thresholds that did not vary systematically as a function of temporal frequency. A position-based mechanism was used to detect the second-order motion with this centrally presented linear grating, as with the radial grating used in Experiment 1. These results show that the position-dependence of second-order motion detection was not obtained because of the peripheral presentation, but instead seems to be a more general trait of the perception of second-order motion.

6. General discussion

Throughout the experiments reported here, performance for gratings defined by first-order characteristics (luminance) were compared to performance for gratings defined by second-order characteristics (contrast-modulation, texture and depth) with identical spatial and temporal frequency profiles. Studying motion in oscillation allowed us to distinguish between sensitivity to the velocity and sensitivity to the change in the position of these patterns. In every case, we have found that the motion of first-order stimuli was detected by a mechanism sensitive to velocity, where the motion of second-order stimuli was detected by a mechanism sensitive to position-change. In the following section, these results are discussed with regard to existing models of second-order motion perception, and the results supporting these models. Alternative accounts are offered for these previous results, and the importance of position information for the perception of second-order motion is discussed.

6.1. Models of second-order motion perception

Several investigations involving a comparison of first-order and second-order stimuli have come to the conclusion that both types of stimuli are detected by the same, or very similar motion mechanisms [13,30,14]. Various authors have proposed that specific detectors exist which extract information from the second-order display, transform it in some way, such as full or half-wave rectification in the case of contrast-modulated textures, and send output to a correlational motion energy analysis system similar to that used for luminance motion [14–16]. Results from the present experiment suggest that this model of second-order motion does not always hold. The correlational motion energy analysis system is based on Reichardt [31] delay-and-compare structure, which proposes motion detection units tuned to a specific velocities of motion. These motion detectors do not code position. However, in the present experiments, second-order motion is detected by a mechanism sensitive to changes in position, not velocity. Second-order motion detection via correlational motion energy analysis cannot account for these results.

Yet, there is a great deal of evidence supporting the notion that second-order motion is detected with a system similar to first-order motion. Here we consider two possible ways in which these results can arise. For the contrast-modulated stimuli (but not the stereo-defined or orientation-modulated ones) motion energy analysis could operate on the distortion products produced by any input nonlinearity. This is the essence of several proposals involving half or full-wave rectification [14]. Alternatively, there may be specialized motion detectors which respond directly to the spatial variation of the second-order property, whether texture or stereo. We consider the possibility of nonlinear transformations first. Brown [32], for example, has argued that texture patterns with equal increments and decrements of luminance may not be balanced in the visual system, because of non-linear coding of luminance prior to motion detecting units. A compressive nonlinearity, in particular, would produce increasing strengths of luminance signals as the contrast of the contrast-modulated pattern increased. This relationship between luminance distortion products and signal strength could explain the absence of low-level motion response in our study (with a maximum at about 20% contrast) as compared to its presence in others using contrasts as high as 100% [14,20,30,33,34]. Moreover, signal strength increases with the speed of the stimulus as well as its contrast, and our stimuli were also very slow, between about 0.01 and 0.15° of visual angle per second, compared to many studies reporting low-level response to texture gratings. Smith et al. [35], for example, tested speeds no slower than 0.25°/s. Two studies [20,21] have addressed whether the components produced by input nonlinearities in response to textures are analyzed by the same mechanisms which analyze the motion of the luminance patterns or by a separate set of mechanisms. They reported results favoring distinct, independent processes, at least if the only distortion products at the fundamental frequency (or second-harmonic) of the stimuli were considered. Whatever the outcome of this controversy concerning the possible role of input nonlinearities, it is clear that this class of model cannot recover motion from either the stereo-defined stimuli or the orientation-modulation textures used here as no pointwise nonlinearity will produce any modulated signal at the fundamental frequency for either of these two stimulus types.

Alternatively, then, there may be specialized, second-order motion detectors that are just much less sensitive to the stimuli used in this experiment than is the available position-based mechanism. If such a system exists, our data place some bounds on its sensitivity. Specifically, threshold level velocity must be high enough that the corresponding displacement thresholds would all be higher than those found in our experiments. This is depicted in Fig. 13, where for the temporal frequency of about 1.0 Hz, this displacement threshold could be around 36° of phase, increasing to approximately 360° at 0.1 Hz. This iso-velocity line represents a velocity which is approximately ten times higher than that obtained from the first-order condition in this experiment. The minimum velocities were estimated for first-order stimuli in Experiment 1 by determining the speed at which the outer-most part of the grating was moving for stimuli at the displacement thresholds. The average value was 0.015° of visual angle per second. So, the minimum velocity to which a second-order velocity-based system must respond would be one-tenth of that or about 0.15°/s.

Johns [36] found just this pattern of results by testing motion in oscillation of two forms of dynamic random dot stereograms. Stereograms formed either a depth-defined bar or a random field of depth-defined patches. They found that displacement thresholds across temporal frequency were constant at approximately 6 arc min for the single bar condition, but fell from about 40–10 arc min for the random patches condition. Thus, when position cues are reduced by creating motion with a field of elements, a second-order stimulus in motion is detected with a velocity-sensitive mechanism. These results suggest that a velocity-dependent mechanism is available to detect stereo-depth defined motion and it is much less sensitive than the position-change detection mechanism. The same may hold for contrast-modulated and other texture-based motion.

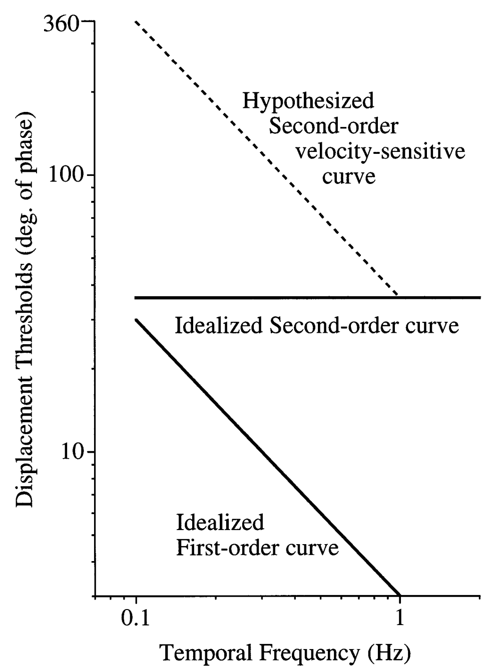

Fig. 13. Solid lines show the idealized curves from the results of the present study for the first-order (luminance) and second-order conditions. Dashed line shows the hypothesized thresholds for the velocity-sensitive mechanism for second-order stimuli.

6.2. The importance of position information

Regardless of whether or not specialized velocity-sensitive mechanisms exist for second-order stimuli, the present study underlines the importance and sensitivity of position tracking mechanisms for the perception of second-order motion. We find that when both position and velocity information is available, first-order stimuli are detected via velocity signals and second-order stimuli are detected via position signals. Given that whenever something moves it also must change its position, any motion display will have information available for both the velocity-sensitive and position-sensitive mechanisms. The important factor, then, is the relative strength of these cues for determining motion thresholds. We find that the velocity-sensitive mechanism is more sensitive to motion of luminance-defined patterns, and the position-based system is more sensitive to motion of texture and depth-defined patterns.

Many of the studies comparing first and second-order stimuli incorporate some procedure, such as using random patches (as in [36]) or short durations [37,22], to bypass a position-sensitive mechanism. The assumption in these procedures is that position tracking must be ruled out before the analysis of second-order motion perception can begin. However, our results indicate that a position-sensitive mechanism is a key element in the perception of second-order motion, in that position tracking may be the more sensitive, ‘default’ mechanism used for detecting and discriminating this motion, at least at slower speeds. The importance of position information in detection of second-order motion is also emphasized by the finding that inter-attribute apparent motion (first to second-order and vice versa) is not readily detectable once position information is removed [38]. To fully account for observers’ ability to discriminate motion of second-order patterns, it seems crucial that position-sensitive motion mechanisms be taken into account.

What kind of motion mechanism is sensitive to position-change, but not velocity? Several possibilities exist, such as the pattern matching process described by Julesz [39], feature tracking system by Ullman [40] and Anstis [12], feature-salience (or ‘third-order’) motion detection system by Lu and Sperling [21], and the attention-based process described by Cavanagh [29]. These position-based systems are an obvious choice for the second-order patterns studied here, because of their clearly delineated features, such as patches of texture, or close points in depth.

Acknowledgements

This work was supported by Grant EY09258 from the National Eye Institute to Patrick Cavanagh, and Graduate Fellowship from the National Science Foundation awarded to Adriane Seiffert. We thank Doug Willen, Robert O’Shea, Martin Wainwright, Frank Tong, and Ted Adelson for giving advice on the work and the manuscript.

References

[1] Nakayama K, Tyler CW. Psychophysical isolation of movement sensitivity by removal of familiar position cues. Vis Res 1981;21:427–33.
[2] Nakayama K. Biological image motion processing: A review. Vis Res 1985;25(5):625–60.
[3] Barlow HB, Levick WR. The mechanism of directionally selective units in the rabbit’s retina. J Physiol 1965;178:477–504.
[4] Zeki SM. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. J Physiol (Lond) 1974;236:549–73.
[5] Maunsell JHR, Van Essen DC. Response properties of single units in middle temporal visual area of the macaque. J Neurosci 1983;3:2563–86.
[6] Exner S. Experimentelle Untersuchung der einfachsten psychischen Processe. Pflügers Arch für Gesamte Physiologie 1875;11:403–32.
[7] Thorson J, Lange GD, Biederman-Thorson M. Objective measure of the dynamics of a visual movement illusion. Science 1969;164:1087–8.
[8] Morgan MJ, Cleary R. Effects of contrast substitutions upon motion detection in spatially random patterns. Vis Res 1992;32(4):639–43.
[9] Wohlgemuth A. Monograph Supplement, I, On the after-effect of seen movement. Br J Psychol, 1911.
[10] Smythe L, Finkel DL. Masking of spatial and identity information from geometric forms by a visual noise field. Can J Psychol 1974;28:399–408.
[11] Breitmeyer BG. Visual Masking: An Integrative Approach. New York: Oxford University Press, 1984:133.
[12] Anstis SM. The perception of apparent motion. Philos Trans R Soc Lond Ser B 1980;290:153–68.
[13] Cavanagh P, Mather G. Motion: The long and short of it. Spat Vis 1989;4:103–29.
[14] Chubb C, Sperling G. Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. J Opt Soc Am A 1988;5:1986–2007.
[15] Johnston A, McOwen PW, Buxton H. A computational model of the analysis of some first-order and second-order motion patterns by simple and complex cells. Proc R Soc Lond Ser B 1992;250:297–306.
[16] Grzywacz NM. One-path model for contrast-independent perception of Fourier and non-Fourier motions. Investig Ophthalmol Vis Sci 1992;33:954.
[17] Wilson HR, Ferrera VP, Yo C. A psychophysically motivated model for two-dimensional motion perception. Vis Neurosci 1992;9:79–97.
[18] Turano K, Pantle A. On the mechanism that encodes the movement of contrast variations: Velocity discrimination. Vis Res 1989;29(2):207–21.
[19] Cavanagh P, Arguin M, von Grünau M. Inter-attribute apparent motion. Vis Res 1989;29:1197–204.
[20] Solomon JA, Sperling G. First and second-order motion and texture resolution in central and peripheral vision. Vis Res 1995;35(1):59–64.
[21] Lu Z, Sperling G. The functional architecture of human visual motion perception. Vis Res 1995;35(19):2697–722.

[22] Derrington AM, Badcock DR, Henning GB. Discriminating the direction of second-order motion at short stimulus durations. Vis Res 1993;33:1785–94.
[23] Pantle A, Turano K. Visual resolution of motion ambiguity with periodic luminance and contrast-domain stimuli. Vis Res 1992;32(11):2093–106.
[24] Pantle A. Immobility of some second-order stimuli in human peripheral vision. J Opt Soc Am A 1992;9(6):863–7.
[25] Pelli DG, Zhang L. Accurate control of contrast on microcomputer displays. Vis Res 1991;31(7/8):1337–50.
[26] Mather G, Morgan M. Irradiation: Implications for theories of edge localization. Vis Res 1986;26(6):1007–15.
[27] Morgan MJ, Mather G, Moulden B, Watt RJ. Intensity-response nonlinearities and the theory of edge localization. Vis Res 1984;24(7):713–9.
[28] Georgeson MA, Shackleton TM. Monocular motion sensing, binocular motion perception. Vis Res 1989;29:1511–23.
[29] Cavanagh P. Attention-based motion perception. Science 1992;257:1563–5.
[30] Smith AT. Correspondence-based and energy-based detection of second-order motion in human vision. J Opt Soc Am A 1994;11(7):1940–8.
[31] Reichardt W. Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In: Rosenblith's WA, editor. Sensory Communication. New York: Wiley, 1961.
[32] Brown RO. The effective luminance of complex stimuli in motion: First-order signals from second-order stimuli. Soc Neurosci Abstr 1994;20:1666.
[33] Nishida S. Spatio-temporal properties of motion perception for random check contrast modulations. Vis Res 1993;33:633–645.
[34] Ledgeway T, Smith AT. The perceived speed of second-order motion and its dependence on stimulus contrast. Vis Res 1995;35(10):1421–34.
[35] Smith AT, Hess RF, Baker CL Jr. Direction identification thresholds for second-order motion in central and peripheral vision. J Opt Soc Am A 1994;11(2):506–14.
[36] Johns AM, Rogers BJ, Eagle RA. Thresholds for detecting moving cyclopean forms are limited by velocity not displacement. Investig Ophthalmol Vis Sci 1996;37(3):285.
[37] Cropper SJ, Derrington AM. Motion of chromatic stimuli: First or second-order. Vis Res 1994;34:49–58.
[38] Mather G, West S. Evidence for second-order motion detectors. Vis Res 1993;33(8):1109–12.
[39] Julesz B. Foundations of Cyctopian Perception. Chicago, Ill: University of Chicago Press, 1971.
[40] Ullman S. The Interpretation of Visual Motion. Cambridge, MA: MIT Press, 1979.