Depth perception from second-order-motion stimuli yoked to head movement

Makoto Ichikawa [a,*], Shin'ya Nishida [b], Hiroshi Ono [c,d]

[a] *Department of Perceptual Sciences and Design Engineering, Yamaguchi University, 2-16-1 Tokiwadai, Ube, Yamaguchi 755-8611, Japan*
[b] *NTT Communication Science Laboratories, NTT Corporation, Kanagawa 243-0198, Japan*
[c] *Department of Psychology, York University, Ont., Canada M3J 1P3*
[d] *Department 2, ATR Human Information Science Laboratories, Kyoto 619-0288, Japan*

Abstract

We examined whether depth perception was produced by the parallax of second-order motion (i.e., movement of non-luminance features, such as flicker, texture size modulation, or contrast modulation that moved in synchrony with lateral head movement). The results, obtained with second-order motion from a simple grating stimuli, showed that depth order was judged correctly with probabilities well above chance, but the reported depth magnitude did not co-vary with parallax magnitude. When we used a complex spatial pattern for which feature tracking was difficult, the accuracy of depth-order judgments descended to chance level. Our results suggest that the visual system (a) can detect the correct depth order by tracking a relative shift in the salient features of a stimulus pattern, but (b) cannot determine depth magnitude from a velocity field given by second-order-motion stimuli.
© 2004 Elsevier Ltd. All rights reserved.

Keywords: Second-order motion; Motion parallax; Depth order; Depth magnitude; Feature tracking

1. Introduction

Retinal image motion produced by an observer's head movement is a reliable and unambiguous depth cue (e.g., Heine, 1905; Ono, Rivest, & Ono, 1986; Rogers & Graham, 1979). For instance, Rogers and Graham demonstrated that depth order, depth magnitude, and the profile of the surface undulation could be derived from the relative retinal image motion among a random dot pattern yoked to a head movement (i.e., the spatial position of each dot was shifted in accordance with the lateral position of the observer's head). They called this depth cue *motion parallax*. (They also used the term motion parallax to refer to the relative retinal image motion produced by a moving surface, but in this paper we restrict our consideration of motion parallax to the relative retinal image motion produced by a head movement.) Using luminance-defined gratings (first-order-motion stimuli), Ono et al. (1986) found that when the magnitude of parallax was small, observers saw depth without seeing motion. [1] For larger parallax, on the other hand, observers saw the movement of the pattern in addition to the depth. These results suggest that retinal motion information is used exclusively for

---

[*] Corresponding author. Tel.: +81 836 85 9724; fax: +81 836 85 9701.

E-mail address: ichikawa@yamaguchi-u.ac.jp (M. Ichikawa).

[1] When the stimulus movement on a screen is yoked to the head movement, it simulates a situation in which the stimulus is stationary and the head is moving. When the extent of motion parallax is small in this situation, observers see depth but no motion (Ono et al., 1986).

generating parallactic depth until the extent of the relative retinal image motion exceeds some critical level. It is still not known, however, whether motion parallax produced by the movement of non-luminance features (i.e., second-order motion) yoked to an observer's head movement is an effective depth cue. Non-luminance features to be considered here are: flicker, texture difference, and contrast modulation (see e.g., Cavanagh & Mather, 1989; Chubb & Sperling, 1988).

Previous studies have examined the effectiveness of second-order motion in determining depth perception from kinetic depth cues. Kinetic depth cues are produced by the projected image of a rotating object independent of an observer's head movement. It is a reliable depth cue, but it contains ambiguous information about the depth order (Wallach & O'Connell, 1953). Prazdny (1986) reported that the image of a rotating wire frame was seen in 3D when depicted by a disparity difference or by dynamic noise. However, other studies using more complex second-order-motion stimuli reported that observers did not perceive surface depth (Dosher, Landy, & Sperling, 1989; Hess & Ziegler, 2000; Landy, Dosher, Sperling, & Perkins, 1991). Thus, from studies using kinetic depth cues, it is still not known what aspect of second-order motion provides information for depth perception.

The studies discussed above indicate that the contribution of second-order motion to depth perception differs from that of first-order motion. This difference is reasonable in that the processing of second-order-motion stimuli differs from that of first-order-motion stimuli for motion perception (e.g., Derrington & Badcock, 1985; Harris & Smith, 1992; Nishida, Edwards, & Sato, 1997; Nishida & Sato, 1992). Perception of motion for second-order stimuli is likely achieved by detecting second-order motion energy by means of specialized low-level sensors, by attentionally tracking the shift in position of salient features in a high-level processing stage, or by both (e.g., Cavanagh, 1992; Lu & Sperling, 1995). Moreover, the position tracking mechanism is said to be different from the velocity-sensitive motion energy detectors (McKay, 1976; Nakayama & Tyler, 1981; Regan & Beverley, 1984). In light of what is known about the perception of motion for second-order stimuli, we investigated how depth from motion parallax is created by second-order-motion stimuli. We conducted three experiments in which observers were asked to judge the order and magnitude of parallactic depth when second-order- or first-order-motion stimuli were yoked to their lateral head movements. The first two experiments examined whether observers perceive consistent depth-order and depth-magnitude for a second-order stimulus. The third experiment examined how depth-order perception for second-order-motion stimuli depends on trackable features of the stimulus pattern.
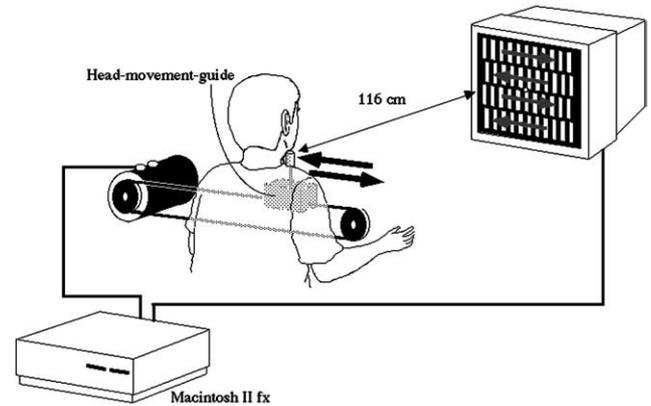

Fig. 1. Apparatus for the experiments. The movement of the head-movement-guide was synchronized with the relative movement of the four bands of gratings on the monitor.

2. Experiment 1

We presented the same magnitudes of motion parallax for the second- and first-order-motion stimuli, and asked observers to make depth order judgments and depth magnitude judgments. The parallax magnitude was specified by the retinal image motion relative to the head movement. We used a simple stimulus configuration consisting of four horizontal bands of vertical gratings that were specified by first- or second-order features. Each band moved horizontally in a direction opposite to its vertical neighbors (see Fig. 1). Relative stimulus movement was yoked to the observer's head movement by manipulating the phase of gratings in accordance with the lateral position of the chin rest.

2.1. Methods

2.1.1. Observers

Five observers, including one of the authors (MI), took part. All were experienced in viewing motion parallax displays with first-order motion, but not with second-order motion. [2] Except for MI, all were naive as to the purpose of the experiment.

2.1.2. Stimulus and apparatus

We presented the stimulus on an Apple Color 13″ High Resolution Monitor using a Macintosh II fx computer with 8•24 Apple video card. The viewing distance was 116 cm (Fig. 1).

The movement of the gratings was yoked to a flexible head-movement-guide under the observer's chin. The head-movement-guide moved sinusoidally and had an amplitude of 15 cm. The observer's chin rested on the

---

[2] During the series of experiments, there was no indication that prolonged experience with second-order displays might change the pattern of reported results.

guide and followed its movement. Since the guide was bendable, the observer had to actively move their head, rather than resting their chin on the guide. (Before the experimental sessions, all of the observers had training sessions in order to teach them to follow the head-movement-guide beneath their chin.) To move the gratings on the monitor in synchrony with the head movement, a computer changed the spatial phase of gratings in accordance with the lateral position of the chin rest with a negligible amount of delay. Given that motion parallax is an effective depth cue, the gratings that move in the opposite direction to the head movement should be seen in front of the fixation plane, while those moving in the same direction should be seen behind of the fixation plane. The magnitude of parallax depth should be determined by the speed of the grating motion relative to that of the head motion. For a more detail description of parallax presentation, see Ujike and Ono (2001).

Three types of second-order-motion stimuli were used: flicker gratings, size modulation gratings, and contrast modulation gratings (Fig. 2a–c). The stimulus consisted of four horizontal bands of 1.3 cpd square-wave gratings, each subtending 2.8×9.8 arc deg. Each band was separated by a vertical gap of 15 arc min, and adjacent bands moved in opposite directions. This made the first and third bands appear either nearer or further than the second and fourth bands. There was a red fixation point (diameter of 10.0 arc min) between the second and third bands. Parts of the stimuli consisted of binary random dot elements (50% white and 50% black). The flicker grating had fixed areas and dynamic areas in which contrast polarity of each element was reversed every 67 ms. The element size was 1.1×1.1 arc min. (This was the minimum step of the display.) The size modulation grating had small element areas (1.1×5.4 arc min for each element) and large element areas (2.2×10.8 arc min). The element arrangement was refreshed every 67 ms. The contrast modulation grating had high contrast areas made of dynamic random dot

elements (1.1×1.1 arc min each, refreshed every 67 ms) and zero contrast areas (uniform gray areas) which had the same mean luminance as the high contrast areas. In these second-order-motion stimuli, the boundaries of the areas, which were defined by flicker, size modulation, or contrast modulation, moved. The luminance of the white elements was 22.2 cd/m² while the luminance of the black elements, as well as that of the background, was 0.1 cd/m². The effects of luminance artifacts in the second-order-motion stimuli were negligibly small, and we confirmed that no coherent motion was seen when the second-order-motion grating and the first-order-motion grating were alternately presented with a phase shift of 90° (see Ledgeway & Smith, 1994).

The first-order-motion stimulus consisted of black and white areas (Fig. 2d). The luminance of the white areas was 22.2 cd/m² while the luminance of the black areas, as well as that of the background, was 0.1 cd/m².

There were two head speed conditions: 0.1 Hz oscillation with a peak velocity of 6 cm/s, and 0.4 Hz oscillation with a peak velocity of 24 cm/s. For each head speed, we used two magnitudes of parallax; 9.3 and 18.5 arc min in terms of equivalent disparity. Equivalent disparity, the unit used to describe the magnitude of motion parallax, was introduced by Rogers and Graham (1982): It provides a unit that is comparable to the value of binocular retinal disparity. It is defined by the difference in the extent of the retinal image motion (in terms of visual angle) caused by a head movement of 6.2 cm (i.e., the average interocular distance) (Rogers & Graham, 1982), or, equivalently, by the ratio of relative retinal image velocity of the two images to head velocity multiplied by 6.2 cm (Ujike & Ono, 2001). For the condition having equivalent disparity of 9.3 arc min, for instance, each grating shifted 11.2 arc min as the head moved 15 cm (i.e., $11.2 \times 2/15 \times 6.2 = 9.3$). The peak velocities of retinal image motion of each grating were 4.5 arc min/s (for equivalent disparity of 9.3 arc min and peak head velocity of 6 cm/s), 9.0 arc min/s (for
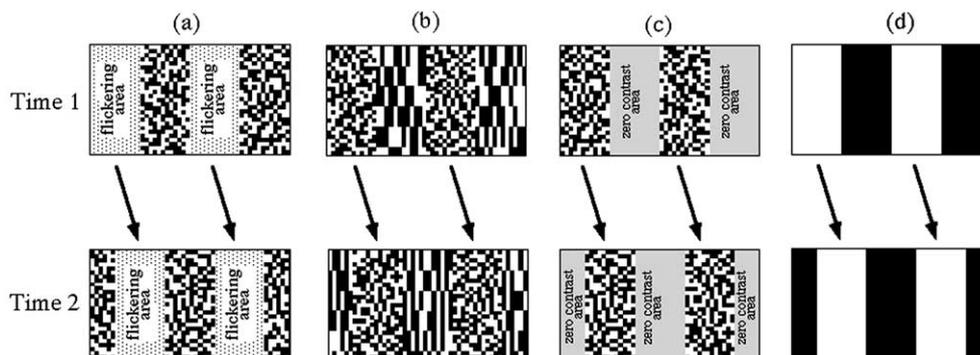

Fig. 2. Diagrams of the four types of motion used in Experiment 1: (a) Motion of flickering grating, (b) motion of texture size modulation grating, (c) motion of contrast modulation grating, and (d) motion of luminance grating. From Time 1 to Time 2, the area specified by a non-luminous feature (a–c), or a luminous feature (d) shift their horizontal positions. Each example shows two cycles from a single band of the stimulus, and each panel shows only two cycles of the 13 cycles in a band.

equivalent disparity of 18.5 arc min and peak head velocity of 6 cm/s), 18.0 arc min/s (for equivalent disparity of 9.3 arc min and peak head velocity of 24 cm/s), and 36.0 arc min/s (for equivalent disparity of 18.5 arc min and peak head velocity of 24 cm/s). We confirmed in a preliminary test that even at the slowest velocity (4.5 arc min/s), all the observers could see all the types of second-order motion. Using a stimulus similar to our first-order-motion stimulus, Ono and Ujike (1993a, 1993b) found that observers see motion as well as depth when the equivalent disparity is above 4 arc min/s. Therefore, under the conditions we used, observers would see both depth and motion—at least for the first-order-motion stimulus. Moreover, one would not expect the apparent depth magnitude to increase proportionally with an increase in the parallax magnitude, when both motion and depth are seen (Ono et al., 1986). We used relatively large motion parallaxes, because the results of preliminary tests indicated that either depth or motion was hard to see with a second-order-motion stimulus with an equivalent disparity of 4.7 arc min.

2.1.3. Procedure

There were 32 conditions (four motion types, two parallax magnitudes, two depth orders, and two peak velocities of the head movements). The stimuli were presented in four blocks, one for each type of motion (flicker, size modulation, contrast modulation,[3] and luminance). In each block of trials, eight different conditions (two parallax magnitudes, two depth orders, and two head movement velocities) were presented 10 times each, in random order. In each trial, observers viewed the stimuli monocularly with their preferred eye. During the observation, they fixated the center red point. They reported verbally which gratings appeared closer. Then, by pulling a tape measure out of its case, without seeing the scale, they reported the magnitude of the apparent depth between adjacent gratings. The viewing time was unlimited.

2.2. Results and discussion

The correct depth order was defined according to the geometry of the motion parallax; it consisted of seeing those bands that moved in the same direction as the head as farther than those that moved in the opposite direction. The reported magnitude of depth was analyzed only when the depth order was judged correctly.

The two head movement velocity conditions produced no consistent differences in the frequency of correct depth order or in the apparent depth magnitude.

Table 1
The percent correct depth order judgements for each motion type in Experiment 1

Conditions	Observers				
	MI	PG	MS	SS	GW
First-order motion					
Luminance	100.0	100.0	80.0	97.5	97.5
Second-order motion					
Flicker	95.0	100.0	80.0	100.0	100.0
Size modulation	92.5	100.0	75.0	92.5	100.0
Contrast modulation	100.0	100.0		100.0	

Also, the two depth-order conditions produced no consistent difference in the apparent depth magnitude. Therefore, we present the results averaged over the two conditions of head movement velocities and depth orders.

Table 1 shows the percentage of correct apparent depth-order judgments for each stimulus condition. The performance was similar across stimulus condition within each observer. These data indicate that the observers reliably reported the depth order specified by the geometry of parallax for all types of motion.

Fig. 3 shows the mean and standard error (SE) of the apparent depth magnitudes obtained in the trials in which the reported depth order was consistent with that specified by motion parallax. The mean magnitudes for the three second-order-motion stimuli (flicker, size modulation, and contrast modulation) were consistently smaller than those for the first-order-motion stimulus (luminance). Moreover, they were about the same magnitude for the two different parallax magnitudes. The apparent depth magnitudes for the first-order-motion stimuli, on the other hand, increased with the increase in the parallax magnitude.
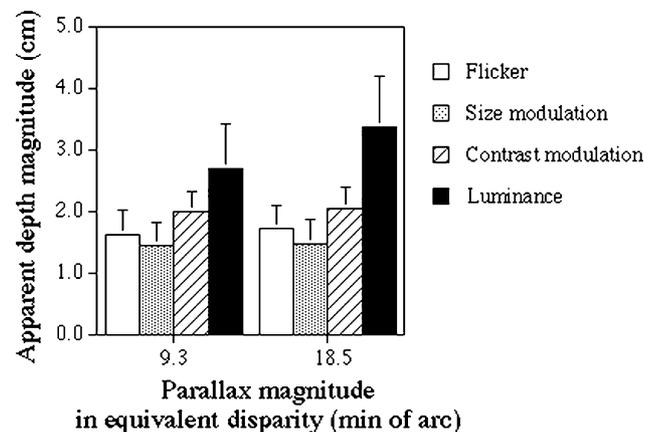

Fig. 3. The means and SE of the apparent depth magnitude for each condition in Experiment 1. Data are from five observers except for the data of the contrast-modulation condition, which is from three observers.

---

[3] The contrast-modulation condition was suggested to us when we presented the initial results at ARVO 1996, and was presented only to observers SS, MI, and PG. The other observers were no longer available to serve in this part of the experiment.

We conducted a 3×2 analysis of variance for repeated measures. The factors were motion type (flicker, size modulation, and luminance) and parallax magnitudes (9.3, and 18.5 arc min). (The data for the contrast-modulation condition were not included in the analysis because this condition was only presented to three observers. Their results in this condition were very similar to their results in the other conditions.) The interaction of the two factors was significant ($F(2,8) = 12.047$, $p < 0.01$), as were the two main effects; motion type ($F(2,8) = 4.758$, $p < 0.05$) and the parallax magnitude ($F(1,4) = 14.123$, $p < 0.05$). To analyze the interaction, Tukey's post hoc *HSD* test was performed. For the luminance condition, the mean apparent depth for the 18.5 arc min condition was larger than that of the 9.3 arc min condition ($p < 0.05$). For all of the second-order-motion stimuli, there was no significant difference between 9.3 and 18.5 arc min conditions. Moreover, for the 18.5 arc min condition, the mean apparent depth for the luminance condition was larger than those for both flicker and size modulation conditions ($p < 0.05$). For the main effect of motion type, Tukey's post hoc *HSD* test showed that the mean apparent depth magnitude of the luminance condition was larger than those for flicker and size modulation conditions ($p < 0.05$).

The apparent depth magnitudes for the second-order-motion conditions (shown in Fig. 3) are larger than zero, but there are no differences across these conditions. These results suggest that second-order-motion stimuli provide information for depth-order perception but do not provide reliable information for depth-magnitude perception.

3. Experiment 2

In Experiment 1, apparent depth magnitude did not covary with parallax magnitude for the second-order-motion stimuli. Since the different types of motion were not equally visible, the observers' difficulty in discriminating the different depth magnitudes with the second-order-motion stimuli may be due to their difficulty in extracting motion signals from the noisy display of the second-order-motion stimuli. (The observers in Experiment 1 reported that the visibility of motion was lower for second-order-motion stimuli than for first-order-motion stimuli.) In Experiment 2, we investigated whether the lack of correlation between the predicted and obtained apparent depth magnitudes in the second-order-motion conditions was due to the second-order-motion stimulus itself, or due to the poor visibility of motion. We equated the visibility of motion for the first- and second-order-motion stimuli by adjusting the modulation amplitude of the stimuli for each individual.

3.1. Method

3.1.1. Observers

Five observers took part. Two of them (MI, PG) participated in Experiment 1. The other three were inexperienced in viewing motion parallax displays. Except for MI, all were naive as to the purpose of the experiment.

3.1.2. Stimulus and apparatus

The same apparatus and viewing distance as in Experiment 1 were used. The stimulus consisted of four horizontal bands of 0.9 cpd sinusoidal gratings defined by contrast modulation (second-order-motion stimulus) or luminance modulation (first-order-motion stimulus). (In order to present smooth movement of the gratings, we used sinusoidal gratings, which could be shifted by a sub-pixel amount by changing the drawing pattern, instead of the rectangle gratings used in Experiment 1.) Each band subtended 2.8×9.8 arc deg, and was separated by a vertical gap of 15 arc min. There was a red fixation point (diameter of 10.0 arc min) between the second and third band. Adjacent bands moved in opposite directions. There were three parallax magnitudes (7.6, 15.2, and 22.8 arc min in terms of equivalent disparity), and two depth orders (the first and third bands of gratings appeared nearer or further than the other two). These parallax magnitudes were selected because in preliminary tests, observers could see depth. The peak retinal velocities were 2.1, 4.2, and 6.4 arc min/s for the three parallax magnitude conditions, respectively. As in Experiment 1, we expected that observers would see both motion and depth—at least with the first-order-motion stimuli in all three parallax-magnitude conditions.

Both the first- and second-order-motion stimuli contained random noise (Fig. 4). Each noise element subtended 1.1×1.1 arc min. In the second-order-motion stimulus, the contrast of dynamic noise was modulated sinusoidally to give vertically oriented second-order-motion gratings. To adjust the motion strength, we manipulated the carrier (noise) contrast, which is known to vary the strength of second-order motion, just as effectively as modulation contrast (Cropper, 1998; Nishida, 1993). At the trough of the sinusoidal modulation, the noise contrast was zero; at the peak, it was at the level determined by the preliminary test described below. In the first-order-motion stimulus, the local mean luminance of dynamic noise was modulated sinusoidally and the luminance modulation amplitude (which was always the same as the noise contrast) was at the level determined by the preliminary test. There were two amplitude conditions for both the first- and second-order-motion stimuli. In the small amplitude condition, the amplitude of the contrast/luminance modulation was equal to the threshold for motion perception; the movement of the contrast/luminance modulation is
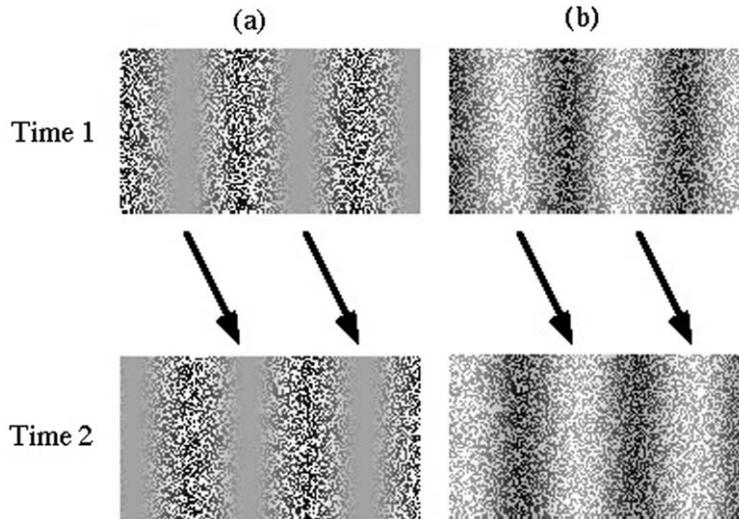
Fig. 4. Diagrams of the stimuli used in Experiment 2: (a) second-order-motion stimuli: the grating was defined by contrast modulation, and (b) first-order-motion stimuli: the grating was defined by luminance modulation. From Time 1 to Time 2, the area specified by a non-luminous feature (a), or a luminous feature (b) shift their horizontal positions. Each panel shows only two cycles of the 8.8 cycles in a band.

expected to be seen 75% of the time for this condition. In the large amplitude condition, the amplitude was twice the threshold value. For both types of stimuli, the mean luminance, as well as the background luminance was 11.1 cd/m^2. The random-dot pattern was updated every 67 ms.

The preliminary test, using the method of constant stimuli, determined the motion threshold for each observer for both the first- and second-order-motion stimuli with a velocity of 5.3 min/s. Observers judged the direction of the motion 32 times for each of eight amplitudes of the contrast/luminance modulation (from 100% to 0.8% of 8 bits depth of modulation amplitude) presented in random order. The direction discrimination performance improved when the amplitude of the contrast/luminance modulation was increased, and probit analysis determined the amplitude that allowed correct perception in 75% of the trials. It ranged from 1.1% to 2.1% for the first-order stimuli, and from 10.2% to 13.9% for the second-order stimuli.

3.1.3. Procedure

There were 24 conditions (two motion types, two modulation amplitudes, three parallax magnitudes, and two depth-order conditions). Each condition was presented six times in random order. During the observation, observers fixated the center red point. In each trial, they viewed the stimuli monocularly (with their preferred eye) while their head moved sinusoidally with an amplitude of 6.5 cm and a peak velocity of 12 cm/s (frequency of 0.46 Hz). We used a smaller amplitude than the one in Experiment 1 so that the observers could follow the head-movement-guide more easily. This amplitude, as well as the amplitude used in Experiment 1, was within the range (from 5 to 30 cm) for which

Ujike and Ono (2001) found that the peak velocity on the retina determined the sensitivity of depth perception regardless of the amplitude. The observers reported verbally which gratings appeared closer. Then, by pulling a tape measure out of its case without seeing the scale, they reported the magnitude of the apparent depth between adjacent gratings. The viewing time was unlimited.

3.2. Results and discussion

Table 2 shows the percentage of correct apparent depth order responses, for each motion type condition and modulation amplitude condition. One of the naive observers (SM) did not perceive a consistent depth order in any of the conditions, including the large amplitude condition for the first-order-motion stimulus. In the following analyses, we used the data of the other four observers.

As we found in Experiment 1, for the large amplitude conditions of both the first- and second-order-motion

Table 2
The percent correct depth order judgements for each motion type in Experiment 2

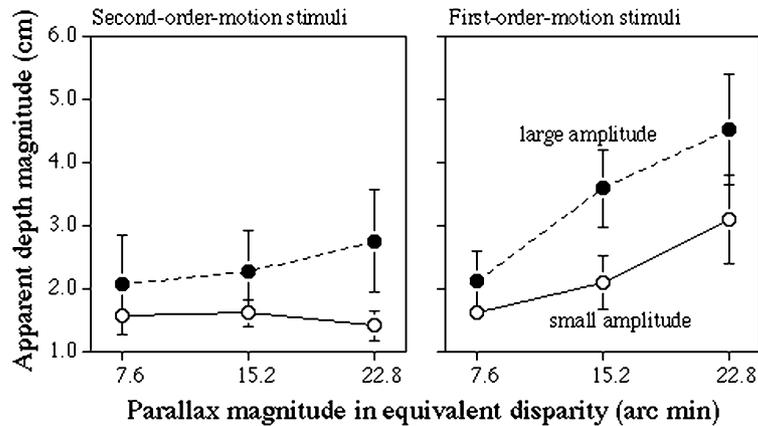| Conditions	Observers				
	MI	PG	JZ	RK	SM
First-order motion					
Small amplitude	97.2	97.2	72.2	77.8	80.6
Large amplitude	91.7	100.0	97.2	83.3	69.4
Second-order motion					
Small amplitude	94.4	97.2	75.0	55.6	50.0
Large amplitude	94.4	97.2	80.6	88.9	52.8

Fig. 5. The means and SE of the apparent depth magnitude from four observers for each condition in Experiment 2. Each data point includes the results of the trials in which the apparent depth order was consistent with the depth order specified by motion parallax.

stimuli, the four observers reported correct depth order with a probability higher than the 75% level. There was little difference in the performance among the different parallax-magnitude conditions.

Fig. 5 shows the mean and SE of the apparent depth magnitudes for each parallax-magnitude condition with the two motion types and the two modulation amplitudes. We performed a three-way repeated measures analysis of variance with motion type, modulation amplitude, and parallax magnitude as factors for the apparent depth magnitude. The interaction of the motion type and parallax magnitude was significant ($F(2,6) = 6.25$, $p < 0.05$), as was the main effect of parallax magnitude ($F(2,6) = 10.79$, $p < 0.05$). There were no other significant interactions (motion type and amplitude, $F(1,3) = 0.71$, $p > 0.40$; amplitude and parallax magnitude, $F(2,6) = 2.98$, $p > 0.10$; motion type, amplitude, and parallax magnitude, $F(2,6) = 4.20$, $p > 0.05$) nor main effects (motion type, $F(1,3) = 7.31$, $p > 0.05$; amplitude, $F(1,3) = 3.59$, $p > 0.15$). These indicate that the difference in amplitude, which was expected to be related to the visibility of the motion, did not contribute to any significant interaction and main effect. For the interaction of the motion type and parallax magnitude, Tukey's post hoc *HSD* test showed that, only for the first-order-motion stimulus, the apparent depth magnitude for the 22.8 arc min condition was larger than that of the 7.6 arc min condition ($p < 0.05$). The same comparisons for the second-order-motion conditions did not indicate significant differences ($p > 0.10$). (Similar statistical significances were obtained for the data analysis that included the response of the fifth observer (SM) who did not report consistent depth order in any of the conditions.) These results indicate that the lack of a consistent relationship between apparent depth magnitude and parallax magnitude is not a consequence of the poor visibility of the motion and it is likely a consequence of the second-order-motion stimulus itself.

4. Experiment 3

In Experiments 1 and 2, observers reported correct depth order with a probability higher than chance level for second-order-motion stimuli. It is suggested that the visual system detects second-order motion either by sensing second-order motion energy, as it does for first-order motion, or by tracking the position change of salient features (e.g., Cavanagh, 1992; Lu & Sperling, 1995). Failure to see depth magnitude may indicate that second-order motion energy does not contribute to the depth percept. Instead, position tracking of multiple features, or sequential comparison of configuration change, may enable observers to infer depth order. Indeed, all the second-order-motion stimuli we used contained trackable features. In Experiment 3, we investigated whether the trackable features in the stimulus were necessary for the consistent depth-order perception from second-order-motion stimuli. We tested depth-order perception for a stimulus in which multiple-feature tracking and sequential configuration change were made difficult.

4.1. Methods

4.1.1. Observers

Two observers from Experiment 2 (MI, PG) took part. Observer PG was naive as to the purpose of the experiment but experienced in psychophysical experiments.

4.1.2. Stimulus and apparatus

The same apparatus and viewing distance as in Experiments 1 and 2 were used. The magnitude of parallax was fixed at an equivalent disparity of 15.2 arc min (peak retinal velocity of 4.2 arc min/s). There were two phase conditions for the sinusoidal function that specified the modulation of the gratings; the consistent-phase condition and the random-phase condition. The stimuli
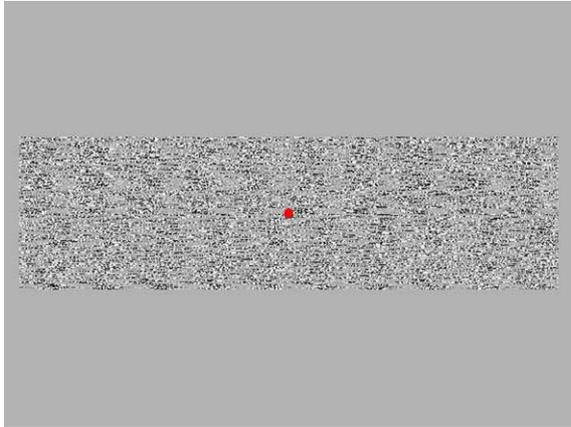
Fig. 6. Example of the stimuli used as the random-phase condition in Experiment 3. The small point at the center depicts the red fixation point.

of the consistent-phase condition were first- and second-order-motion sinusoidal gratings similar to those used for the large amplitude conditions in Experiment 2, but differed in two ways, in order to make it more difficult for the observers to track features at motion boundaries. First, there was no gap between the bands. Second, there were only two bands of sinusoidal modulation that presented the parallactic depth cue. For the random-phase condition, we used the same sinusoidal gratings as in the consistent-phase condition, except that the phase of the sinusoidal modulation was randomly varied independently at each vertical position (1.1 arc-min) within the bands (Fig. 6). For both consistent- and random-phase conditions, each grating subtended 2.8×9.8 arc deg. When we asked the observers to report the difference in the direction of the motion between the gratings above or below the fixation point, they were able to correctly do so for all the stimulus conditions.

4.1.3. Procedure

There were eight conditions (two motion types, two phases, and two depth orders). Each stimulus condition was presented 32 times in random order. During the observation, the observers fixated the center red point. In each trial, they moved their heads back-and-forth twice (for about 4.3 s) by following the head-movement-guide. The guide moved sinusoidally with an amplitude of 6.5 cm and a peak velocity of 12 cm/s. While moving their head, the observers viewed the stimulus monocularly, and then reported the apparent depth order by pressing a computer key.

4.2. Results and discussion

Table 3 shows the percentage of correct responses for each condition. For both observers, the depth order judgment was at the chance level for the random-phase condition of the second-order-motion stimuli, while it

Table 3
The percent correct depth order judgements for each motion type in Experiment 3

Conditions	Observers	
	MI	PG
First-order motion		
Consistent-phase	100.0	100.0
Random-phase	96.9	96.9
Second-order motion		
Consistent-phase	100.0	100.0
Random-phase	53.1	43.8

was nearly perfect for the other three conditions. As mentioned above, in the preliminary test, we confirmed that direction perception per se was not impaired in the random-phase condition. These results are consistent with the hypothesis that depth-order perception from a second-order-motion stimulus is not supported by motion energy involved in second-order-motion stimuli, but by relative position shifts in salient stimulus features, or configuration changes combined with information about the direction of the head movement. Additionally, failure to judge veridical depth order from random-phase second-order-motion stimuli rejects an alternative interpretation that the observers cognitively inferred depth order of second-order stimuli from the perceived direction of a grating band in relation to his/her head movement.

The procedures of Experiment 3 and that of the other two experiments were different with regard to the number of tasks and the stimulus viewing duration. In order to eliminate the concern that these differences are responsible for observers' inability to report the depth order in some of the conditions of Experiment 3, we conducted an additional experiment with five new naive observers. For the stimuli used in Experiment 3, the observers reported both the order and magnitude of the apparent depth with no restriction in the viewing time. The head-movement-guide used in Experiments 1 and 2 was not available for the new experiments. Therefore, we asked the observers to move their head from side to side in synchrony with the sinusoidal stimulus movement. They were to change the direction of head movement when the computer made a beep sound (i.e., at every directional change of the stimulus movement). As found in the main experiment, all five observers correctly reported the depth order specified by parallax except when they viewed the random-phase second-order stimuli. The apparent depth magnitudes obtained with first-order stimuli was significantly larger than those obtained with second-order stimuli, and the two means of the phase conditions (consistent vs. randomized) were not statistically different. While obtained with slightly different apparatus than what was used in the original experiments, these results do not support

the idea that the procedural differences among the experiments had any significant effects on the pattern of the results.

Because Experiment 3 used the same physical magnitude of equivalent disparity for all the different stimulus conditions, one of the anonymous reviewers raised the possibility that our results depended on differences in the sensitivity of relative motion for each different condition. However, in our new experiment, in which we used the stimulus from Experiment 3 and a procedure similar to that of Experiments 1 and 2, we obtained similar results. This was so even when the magnitudes of equivalent disparity were adjusted to equate the visibility of the minimum relative motion. The equivalent disparity of 15.2 arc min (peak retinal velocity of 4.2 arc min/s) was found to correspond to 6.4 times the threshold of the first-order/consistent-phase condition. In the new experiment, the equivalent disparities of the other three conditions were also 6.4 times their thresholds. They were 18.4, 23.2 and 51.2 arc min for the first-order/inconsistent-phase, second-order/consistent-phase and second-order/inconsistent-phase conditions, respectively.

5. General discussion

Experiments 1 and 2 show that the visual system can specify depth order, but not depth magnitude by second-order motion parallax. This is consistently found regardless of the type of second-order motion. According to the results of Experiment 3, the depth order perception with the second-order-motion stimuli used in Experiments 1 and 2 is based on the detection of position shifts of trackable features. The same pattern of results was robustly obtained despite a number of differences in conditions among different experiments (observers, type of second-order motion, stimulus parameters such as spatial frequency of the gratings, waveforms of the modulation, and spatial extent of the stimuli).

Although this study shows that consistent depth-order perception requires trackable features for second-order-motion stimuli, the existence of trackable features is not a necessary condition for parallactic depth-order perception from a first-order-motion stimulus. That is, even when there is no trackable shift in a first-order-motion stimulus, consistent depth-order perception is established by coupling a first-order-motion signal with the signal of head movement direction. For example, observers see the correct parallactic depth order without any perception of position shift when the retinal image motion is slow (Ono et al., 1986; Ono & Ujike, 1994), and without any retinal image shift when the motion signal is derived from a motion aftereffect (Ono & Ujike, 1994). Moreover, for first-order-motion stimuli consisting of multiple elements, in which it is unlikely that observers can attentionally track all the shift of the ele-

ments (note that spatial attention cannot be directed simultaneously to many locations (e.g., Eriksen & Yeh, 1985; Posner & Snyder, 1980)), observers can perceive the depth order specified by motion parallax (e.g., present Experiment 3; Ono et al., 1986; Rogers & Graham, 1979). In these cases, a relative shift was not trackable. These findings suggest that consistent depth-order perception is achieved by coupling the directional signal of a head movement with the directional signal of stimulus motion, which could be derived from either pre-attentive detection of motion energy or attentive position tracking of relative shift (or recognition of configuration change).

The present results suggest that a first-order-motion component that is not available in the second-order motion is required for consistent depth-magnitude perception. For the first-order-motion stimuli, there is a correlation between apparent depth magnitude and parallax magnitude (Experiments 1 and 2), and this correlation did not disappear even when the amplitude of the luminance modulation was reduced to threshold levels of motion perception (Experiment 2). In contrast, there was no such correlation for the second-order-motion stimuli (Experiments 1 and 2). The correlation we found with the first-order-motion stimuli is consistent with previous studies (Ono & Ujike, 1994; Ujike & Ono, 2001), which demonstrate that the apparent depth magnitude produced by the retinal image motion, or motion after effect, accompanied with lateral head movement, increases with an increase in gain (relative velocity signal/head movement velocity). These studies, together with the results of this study, suggest that the first-order-motion component that is not available in the second-order-motion stimuli is the relative velocity signal, and that the visual system cannot extract it from a second-order-motion stimulus. The later suggestion is consistent with the claim of Nishida, Ledgeway, and Edwards (1997) that the processing of second-order motion does not provide effective input to relative motion processing. [4]

Finally, we point out that previous findings about kinetic depth cues are compatible with the present finding that depth-order perception requires information about a trackable position shift when viewing second-order-motion stimuli. Prazdny (1986) demonstrated that second-order-motion stimuli can produce a reliable

---

[4] By using the contrast modulation gratings (the same stimuli as used in Experiment 2), we conducted informal observations in which all the bands moved in the same direction, but the speed of the first and third bands was different from that of the second and fourth bands. For those stimuli, observers tended to see the depth order specified by relative velocity; they saw the bands whose retinal velocity was higher as nearer. The apparent depth magnitude, however, was constant even when the extent of the relative velocity increased. These results further support our conclusion that, when determining the depth order for a second-order-motion stimulus, the visual system depends only upon the sign in the relative motion velocity, but not upon the relative motion velocity itself.

perception of rotating 3D objects. In his study, an object profile specified by the second-order-motion properties was always clearly visible on a stable background, and therefore, it was trackable. In contrast, several studies reported that observers failed to see 3D shapes defined by a kinetic depth cue when viewing second-order-motion stimuli (Dosher et al., 1989; Hess & Ziegler, 2000; Landy et al., 1991). In these studies, the stimulus gave a kinetic depth cue by means of a motion vector field carried by randomly distributed multiple elements without providing trackable features of the objects. Furthermore, Hess and Ziegler (2000) found that kinetic depth was seen for a second-order-motion stimulus consisting of only two elements, although it was not seen for those consisting of multiple elements (more than 60). They interpreted this result as indicating that the failure to see kinetic depth from a second-order-motion stimulus is due to difficulties in the coherence or binding of second-order-motion signals across space. Their finding and interpretation are compatible with our proposal that the visual system requires a trackable position shift for consistent depth-order perception when viewing second-order-motion stimuli. The present study and these previous studies indicate that the visual system requires trackable features to create depth perception from cues from the retinal image motion (including motion parallax and kinetic depth cues) defined by second-order motion.

Acknowledgments

References

Cavanagh, P. (1992). Attention-based motion perception. *Science, 257*, 1563–1565.

Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision, 4*, 103–129.

Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A, 5*, 1986–2007.

Cropper, S. J. (1998). Detection of chromatic and luminance contrast modulation by the visual system. *Journal of the Optical Society America A, 15*, 1969–1986.

Derrington, A. M., & Badcock, D. R. (1985). Separate detectors for simple and complex grating patterns? *Vision Research, 25*, 1869–1878.

Dosher, B. A., Landy, M. S., & Sperling, G. (1989). Kinetic depth effect and optic flow-I. 3D shape from Fourier motion. *Vision Research, 29*, 1789–1813.

Eriksen, C. W., & Yeh, Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 583–597.

Harris, L. R., & Smith, A. T. (1992). Motion defined exclusively by second-order characteristics does not evoke optokinetic nystagmus. *Visual Neuroscience, 9*, 565–570.

Heine, L. (1905). Über wahrehmung und vorstellung von entfernungsunterschieden. *Albrecht von Graefes Archiv für Klinsche und Experimentelle Ophthalmologie, 61*, 484–498.

Hess, R. F., & Ziegler, L. R. (2000). What limits the contribution of second-order motion to the perception of surface shape? *Vision Research, 40*, 2125–2133.

Landy, M. S., Dosher, B. A., Sperling, G., & Perkins, M. E. (1991). Kinetic depth effect and optic flow-II. First- and second-order motion. *Vision Research, 31*, 859–876.

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.

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

McKay, D. M. (1976). Perceptual conflict between visual motion and change of location. *Vision Research, 16*, 557–558.

Nakayama, K., & Tyler, C. W. (1981). Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Research, 21*, 427–433.

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

Nishida, S., Edwards, M., & Sato, T. (1997). Simultaneous motion contrast across space: involvement of second-order motion. *Vision Research, 37*, 199–214.

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

Nishida, S., & Sato, T. (1992). Positive motion after-effect induced by bandpass-filtered random-dot cinematograms. *Vision Research, 32*, 1635–1646.

Ono, M. E., Rivest, J., & Ono, H. (1986). Depth perception as a functional of motion parallax and absolute distance information. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 331–337.

Ono, H., & Ujike, H. (1993a). Zone in which motion parallax is completely effective. *Investigative Ophthalmology and Visual Science, 34*, 1052 (supplement).

Ono, H., & Ujike, H. (1993b). Equal depth contours as a function of head velocity. *Perception, 22*, 81 (supplement).

Ono, H., & Ujike, H. (1994). Apparent depth with motion aftereffect and head movement. *Perception, 23*, 1241–1248.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160–174.

Prazdny, K. (1986). Three-dimensional structure from long-range apparent motion. *Perception, 15*, 619–625.

Regan, D., & Beverley, K. I. (1984). Figure-ground segregation by motion contrast and by luminance contrast. *Journal of Optical Society of America A, 1*, 433–442.

Rogers, B. J., & Graham, M. E. (1979). Motion parallax as an independent cue for depth perception. *Perception, 8*, 125–134.

Rogers, B. J., & Graham, M. E. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research, 22*, 261–270.

Ujike, H., & Ono, H. (2001). Depth thresholds of motion parallax as a function of head movement velocity. *Vision Research, 41*, 2835–2843.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45*, 205–217.