Perception, 1998, volume 27, pages 937–949

Apparent depth with retinal image motion of expansion and contraction yoked to head movement

Takanao Yajima, Hiroyasu Ujike¶, Keiji Uchikawa
Department of Image Science and Engineering Laboratory, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, 226-8503 Japan; ¶ Human Informatics Department, National Institute of Bioscience and Human-Technology, 1-1 Higashi, Tsukuba, 305-8566 Japan; e-mail: ujike@nibh.go.jp
Received 18 December 1996, in revised form 22 July 1998

Abstract. The two main questions addressed in this study were (a) what effect does yoking the relative expansion and contraction (EC) of retinal images to forward and backward head movements have on the resultant magnitude and stability of perceived depth, and (b) how does this relative EC image motion interact with the depth cues of motion parallax? Relative EC image motion was produced by moving a small CCD camera toward and away from the stimulus, two random-dot surfaces separated in depth, in synchrony with the observers' forward and backward head movements. Observers viewed the stimuli monocularly, on a helmet-mounted display, while moving their heads at various velocities, including zero velocity. The results showed that (a) the magnitude of perceived depth was smaller with smaller head velocities (<10 cm s^{-1}), including the zero-head-velocity condition, than with a larger velocity (10 cm s^{-1}), and (b) perceived depth, when motion parallax and the EC image motion cues were simultaneously presented, is equal to the greater of the two possible perceived depths produced from either of these two cues alone. The results suggested the role of nonvisual information of self-motion on perceiving depth.

1 Introduction

As we move about in the environment the retinal images of stationary objects in our visual world are constantly changing. They sweep across our retinae and change size at varying rates, yet we rarely perceive these changes as motion or changes in size. One type of retinal image motion, namely, relative image motion during lateral or vertical head movements (ie motion parallax), has been studied extensively (Ferris 1972; Johansson 1973; Eriksson 1974; Rogers and Graham 1979; Ono et al 1986; Ono and Steinbach 1990; Steinbach et al 1991) and has been shown to be a principal cue for depth perception under monocular viewing (Rogers and Graham 1979; Steinbach et al 1991). Another type of image motion, relative expansion and contraction (hereafter EC) during forward and backward head movements, however, has been rarely focused on as a depth cue (however, for a relation with size constancy see Gregory and Ross 1964a, 1964b; and for a relation with motion perception, Wallach and Flaherty 1975; Harris et al 1981). The purpose of this study is to examine the effectiveness, as a depth cue, of relative EC image motion during forward and backward head movements. To explore the cue effectiveness, we started to examine the effectiveness of forward and backward head movements on the perception of depth.

Motion parallax has been shown to be effective both (a) to produce veridical depth and also larger depth than that from relative image motion without head movement (Rogers and Graham 1979; Ono and Steinbach 1990), and (b) to produce rigid depth (Ono and Steinbach 1990; Rogers and Rogers 1992). Ono and Steinbach (1990) presented the same relative image motion with and without observers' head movement, and found that the magnitude of perceived depth was larger with head movement than without head movement. Moreover, Rogers and Graham (1979) compared relative retinal image motion yoked to movement of the observer's head with that yoked to

¶ Author to whom correspondence should be addressed.

movement of the stimulus display, and found that the magnitude of perceived depth in the head-movement condition was veridical and also larger than that in the display-movement condition. Both Ono and Steinbach (1990) and Rogers and Rogers (1992) showed that relative retinal image motion yoked to observers' head movements produced a rigid perception of depth, while the same motion viewed without head movements produced an unstable perception in which the direction of depth often reversed.

Unlike motion parallax, relative EC image motion may not need to be yoked to head movements to produce a stable perception of depth. For example, the expansion of retinal images has been discussed in terms of a cue for motion in depth (eg Lee 1976), and it was shown that time-to-contact of two expanding images which were simultaneously presented to an observer was discriminated with accuracy of 90% while the difference of time-to-contact was 100 ms (Todd 1981). Moreover, time-to-contact can be discriminated by using only the ratio of the angular size of objects to the rate of increase of angular size (Regan and Hamstra 1993). These findings indicate that the stable depth between two objects can be perceived solely from relative EC retinal image motion without head movements.

Yoking relative EC image motion to forward and backward head movements, however, may increase the effectiveness of this depth cue. Using a motion-adaptation paradigm, Wallach and Flaherty (1975), and Harris et al (1981) indicated that image motion signal obtained during observers' movement can be used for perception of 3-D structure, but not necessarily for motion perception. Wallach and Flaherty (1975) compared velocity aftereffects of adaptation to expansion of image motion while observers were moving forward, moving backward, or stationary. They found that the aftereffects were decreased for the adaptation during observers' movement. Moreover, Wallach and Flaherty (1975) and Harris et al (1981) compared motion aftereffects of adaptation to image motion of either expansion or contraction only while observers were moving forward, moving backward, or stationary. They both found that the motion aftereffect decreased for the adaptation to image expansion during observers' forward movement. These results showing decrement of motion-adaptation effects imply that the motion signal yoked to head movement was used to perceive 3-D structure, which is consistent with Ono and Steinbach's (1990) motion-parallax argument of a trade-off between motion perception and depth perception.

To explore the cue effectiveness of relative EC image motion yoked to forward and backward head movements, we also examined the interaction of relative EC image motion with motion parallax. The types of the interaction that most of the literature focused on were (a) weighted linear averaging (Dosher et al 1986; Rogers and Collett 1989; Johnston et al 1993; Ichikawa and Saida 1996), in which the amplitudes of depth calculated from each depth cue are linearly combined, with weights assigned to each cue according to its effectiveness, and (b) cue dominating (Bülthoff and Mallot 1988, Buckley and Frisby 1993; Norman and Todd 1995), in which the amplitude of depth is determined by an effective cue with the other, ineffective cues being inhibited. Therefore, the type of the interaction and the weight for averaging reveal the effectiveness of a cue relative to that of the others combined. Because motion parallax is an effective depth cue, the interaction observed between these monocular depth cues would reveal the cue effectiveness of relative EC image motion yoked to head movements, and would also suggest the visual strategy that extracts the 3-D environment from retinal image motion caused by head movements.

The two main questions we address in this study are (a) what effect does yoking relative EC image motion to forward and backward head movements have on the perception of depth, and (b) how does relative EC image motion yoked to forward and backward head movements interact with motion parallax? To answer these questions, we performed the following three experiments. In experiment 1, we measured the magnitude of perceived depth with and without head movements. In experiment 2, we measured the magnitude of perceived depth as a function of head velocity. In experiment 3, we measured the magnitude of perceived depth when relative EC image motion and motion parallax were presented together, and compared it with that when each cue was presented separately.

2 Experiment 1

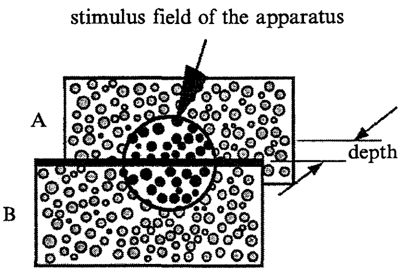
Is the relative EC of retinal images a more effective depth cue when coupled to forward and backward head movements than when not? Although Todd (1981) and Regan and Hamstra (1993) reported that time-to-contact of two objects can be discriminated, whether the depth between them is veridically perceived is unknown. Moreover, both Wallach and Flaherty's (1975) and Harris et al's (1981) reports implied that head-movement information is used to interpret 3-D structure from 2-D image motion of expansion. Given these, perceived depth from the relative EC image motion may depend on whether or not it is coupled to forward and backward head movements. To answer the above question, we measured the magnitude of perceived depth, between two textured surfaces, produced by the relative EC motion of their retinal images in the following two conditions: (a) head-movement condition, in which the retinal image motion was yoked to head movement, and (b) no-head-movement condition, in which the retinal image motion was reproduced without head movement.

2.1 *Method*

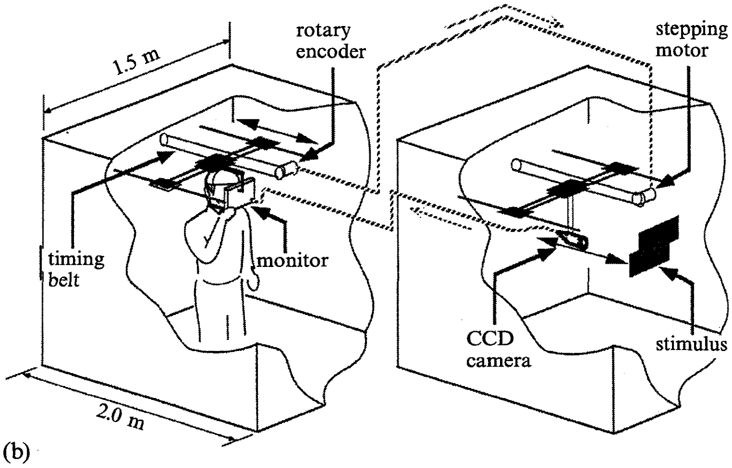
Two male observers, 23 and 24 years of age, participated in this experiment. One of the observers (KM) was naive as to the purpose of the experiment while the other (TY) was one of the authors.

The stimulus consisted of the videoed image (Toshiba CCD camera, IK-UM42) of two random-dot patterns separated in depth by 2 to 20 cm, and was presented on a helmet-mounted, 6-inch NTSC monitor (Epson, ET-S6Z). The monitor was positioned 15 cm in front of the observer's preferred eye, and was viewed monocularly through a 20 mm aperture cut into a piece of black cardboard. A 7 D convex lens was positioned in front of the observer's eye to allow for comfortable accommodation of the monitor and minimize the difference between the accommodation distance and the camera-to-stimulus distance. The random dots, which varied in size from 1.5 to 18 mm in diameter, had a luminance of 3.6 cd m^{-2} and the background luminance was 47.9 cd m^{-2}. The stimulus field subtended 53 deg in diameter at the eye (see figure 1a).

The relative EC motion of the image of the stimulus was produced as follows. The observer and the stimulus were located in two separate dark booths, of approximately 1.5 m $\times$ 2.0 m in size (figure 1b). In the observer booth, the helmet housing the monitor was attached to a guide rail via a timing belt, which was in turn attached to a rotary encoder (Japan Servo, MXE302P). Thus, when the observer's head was strapped into the helmet, its movements were restricted to the forward and backward directions and were monitored by the rotary encoder. In the stimulus booth, the CCD camera videoed the stimulus at a rate of 60 frames s^{-1} and was made to move towards and away from the stimulus, in synchrony with the observer's forward and backward head movements, via a stepping motor (Oriental Motor, PK268-02A) which received its pulse signal from the rotary encoder in the observer booth. The delay of the movement of the camera from the movement of the observer's head was well within the frame period (17 ms) of the videoed image when the velocity of the head movement was less than 30 cm s^{-1}. Amplitude of possible excursion of both the camera and the helmet was 90 cm. In the experiment, the camera was positioned a mean distance of 30.5 cm from the center of the stimulus.



(a)



(b)

Figure 1. (a) The stimulus consisted of two textured surfaces (A and B) separated in depth. Only the area of the stimulus within the circle (diameter 53 deg) was imaged by the CCD camera; thus when viewing the stimulus without head movements, observers saw two random-dot surfaces separated by a black horizontal line. (b) A schematic diagram of the apparatus. Observers' movements were transmitted to the stepping motor to which controlled the movement of the CCD camera. The image from the CCD camera was viewed by the observers on the helmet-mounted display.

The observer's task was to report the magnitude and the direction of perceived depth while either (a) moving his head forward and backward through an extent of 25.0 cm by swaying his body back and forth at the same position (head-movement condition) or (b) keeping his head stationary (no-head-movement condition). In the head-movement condition, the observer viewed the stimulus in real time while moving his head in synchrony with an electric metronome which beeped at a frequency of 0.8 Hz. Thus, the head moved at a velocity of approximately 20 cm s^{-1}. To keep this head movement control accurate and precise, we installed head-movement stoppers at both ends of the 25.0 cm excursion, and also had observers practice the head movement sufficiently before experimental sessions. The practice for each observer was continued until both the experimenter and the observers felt that the extent of head movement and its synchronization with the metronome was precise and accurate. The CCD camera moved through the same extent and at the same velocity as the observer's head. In the no-head-movement condition, the observer viewed a video of the stimulus taken during the head-movement condition. Throughout the no-head-movement condition the observer's head was held stationary. If, after viewing the stimulus, the observer reported seeing depth he reported which surface appeared in front and matched the magnitude of the perceived depth by manually adjusting the distance between two pieces of cardboard which were arranged parallel to each other and attached to a metal rod orthogonally.

Each observer completed a total of 160 trials, performed in sixteen blocks of 10 trials each. Within each block the stimulus was presented once at each of ten different magnitudes of physical depths (2 to 20 cm in steps of 2 cm) whose direction was such that the lower surface was presented either in front of or behind the upper surface. The directions of the physical depths were balanced between blocks. The presentation order of the different magnitudes of physical depths was randomized within blocks and the directions of the depths and the head-movement/no-head-movement conditions varied randomly between blocks.

2.2 *Results and discussion*

The results from each observer (magnitude of perceived depth, for each condition, as a function of the magnitude of physical depth) are presented in figure 2. The results from trials of different directions but the same magnitude of physical depths were averaged. It can be seen clearly from the figure that (a) the magnitude of perceived depth was consistently greater in the head-movement condition than in the no-head-movement condition, (b) the magnitude of perceived depth increased with that of the physical depth, and (c) depth was not perceived when the physical depth was 6 cm or less in the no-head-movement condition, or when it was 2 cm in the head-movement condition. We performed a two-way repeated-measures analysis of variance on the data from each observer using head-movement condition (head-movement/no-head-movement) and physical depth (2 to 20 cm in steps of 2 cm) as factors. The analysis showed a significant effect of head-movement condition (for KM, $F_{1,140} = 140.3$, $p < 0.001$; and for TY, $F_{1,140} = 490.5$, $p < 0.001$), a significant effect of physical depth (for KM, $F_{9,140} = 174.8$, $p < 0.001$; and for TY, $F_{9,140} = 134.5$, $p < 0.001$) and a significant (head-movement condition) $\times$ (physical depth) interaction (for KM, $F_{9,140} = 6.72$, $p < 0.001$; and for TY, $F_{9,140} = 10.67$, $p < 0.001$).

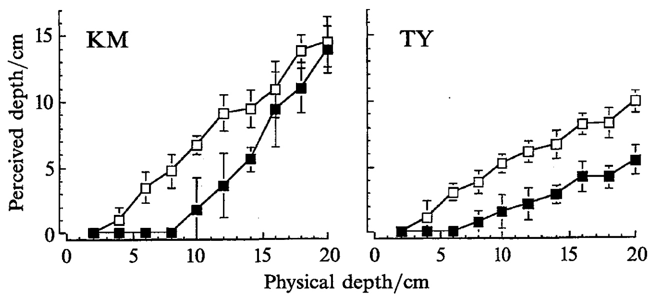


Figure 2. The magnitude of perceived depth in the head-movement condition (open squares) and no-head-movement condition (filled squares) as a function of the magnitude of physical depth for observers KM and TY. In both conditions, perceived depth increased with physical depth. The magnitude of perceived depth was larger, however, in the head-movement condition than in the no-head-movement condition.

Our finding that the magnitude of perceived depth produced by relative EC image motion was greater in the head-movement condition than in the no-head-movement condition indicates that coupling this type of motion to forward and backward head movements increases its effectiveness as a depth cue. Moreover, observers reported that they perceived the depth more easily in the head-movement condition than in the no-head-movement condition. These findings are consistent with those of Rogers and Graham (1979) and Ono and Steinbach (1990), who reported that the magnitude of perceived depth produced by lateral shearing motion (ie motion parallax) was greater when the image motion was yoked to leftward and rightward head movements than when it was presented without head movements.

Given that retinal image motion per se was identical in the two head movement conditions, the results showing that head movements increased the effectiveness of perceiving depth from relative EC motion indicate that nonvisual information such as the vestibular and proprioceptive signals were used for perceiving depth. This indication is consistent with Wallach and Flaherty's (1975) and Harris et al's (1981) reports suggesting expanding image motion that was perceived as motion when observers were stationary was used for 3-D spatial information during observers' movement. Moreover, the indication is consistent with Gregory and Ross's (1964b) report implying that proprioceptive information was effective for size constancy of the EC image.

Whether nonvisual information improves the sensitivity of 3-D perception has been discussed in the literature (eg Bingham and Stassen 1994). Cornilleau-Pérès and Droulez (1994) compared the detection of 3-D curvature from lateral motion of the retinal image in three conditions: self-motion, object translation, and object rotation. They found that the performance was the best in the object-rotation condition and the poorest in the object-translation condition. From this result, they suggested that the nonvisual information of self-motion indirectly improves the performance by stabilizing the retinal image through the vestibulo-ocular reflex. However, van Damme and van de Grind (1996) suggested that nonvisual information is used to improve 3-D perception from their finding that the improvement of integration time of 3-D curvature in the head-movement condition could not be explained by motion-detection performance. Moreover, Heidenreich and Turano (1996) reported that stabilizing the retinal image could not improve speed discrimination, which is contrary to the suggestion of Cornilleau-Pérès and Droulez (1994). In our experiment, the type of retinal image was expansion and contraction and not lateral motion; thus the retinal image stability through some reflexes (eg the vestibulo-ocular reflex) can not explain the different amount of perceived depth in the two head-movement conditions.

Although the magnitude of perceived depth more closely approximated the physical depth in the head-movement condition than in the no-head-movement condition, stable depth was perceived by all observers in the no-head-movement condition. Moreover, the magnitude of perceived depth in the no-head-movement condition increased as a function of the magnitude of the physical depth. This result is not surprising, however, because the relative EC image motion per se (ie image motion not yoked to head movements) contains inherent information about the direction of depth. For example, images of near surfaces expand faster than those of more distant surfaces. Moreover, it has been reported (Todd 1981; Regan and Hamstra 1993) that the time-to-contact of approaching objects can be discriminated on the basis solely of the relative image motion of expansion.

Our finding that depth was not perceived when the physical depth was 6 cm or less in the no-head-movement condition or when it was 2 cm in the head-movement condition indicates that the depth threshold for depth produced by relative EC motion is higher than that produced by lateral shearing motion (ie motion parallax). Rogers and Graham (1982) measured parallactic depth thresholds with lateral head movements for different spatial frequencies (0.05 to 1.6 cycles deg^{-1}) of sinusoidally corrugated depth among random-dot patterns and reported thresholds, in terms of equivalent disparity, ranging from 30 to 150 s of arc; these thresholds correspond to 0.73 to 3.6 mm at their stimulus distance of 57 cm. Moreover, Gonzalez et al (1989) who measured the parallactic depth thresholds of children with one eye enucleated, and Steinbach et al (1991) who measured parallactic depth thresholds for both horizontal and vertical head movements, both reported thresholds of less than 1 mm. Given that depth thresholds from relative EC image motion increase as the physical absolute distance to the stimulus increase and the physical absolute distance was smaller (30.5 cm) in the present study than in the cited motion-parallax studies, clearly depth thresholds for depth produced by relative EC motion is much higher than the reported parallactic thresholds of less than a few millimeters.

3 Experiment 2

The results of the first experiment show that relative EC image motion is a more effective depth cue when it is coupled to forward and backward head movements than when it is presented without head movements. In this experiment we examine the relationship between head velocity and the magnitude of perceived depth produced by relative EC image motion yoked to forward and backward head movements.

Recently, Ono and Ujike (1993) reported that the magnitude of parallactic depth is head-velocity dependent. Specifically, they reported that (a) the magnitude of perceived depth is constant for constant simulated depths when the head moves faster than about 10 cm s^{-1} and (b) the magnitude of perceived depth decreases for constant simulated depths, when the head moves more slowly than about 10 cm s^{-1}. Although they did not discuss the underlying mechanism responsible for this difference in the magnitude of parallactic depth, it is likely that the differing vestibular signals associated with the different head velocities played some role. If this is so, then the magnitude of perceived depth from relative EC image motion may also vary as a function of the velocity of the forward and backward head movements. To examine the possibility of the head-velocity dependency of the depth from relative EC image motion, we measured the magnitude of perceived depth, using the same stimulus as in experiment 1, under four different head-velocity conditions.

3.1 *Method*

Three male observers, ranging in age from 22 to 24 years, participated in this experiment. Two of the observers (KM and TY) had participated previously in experiment 1 and all but one observer (TY) were naive as to the purpose of the experiment. These same three observers also participated subsequently in experiment 3.

The apparatus and the stimulus pattern were identical to those used in experiment 1. The experimental conditions included (a) five magnitudes of physical depth ranging from 8 to 24 cm in steps of 4 cm, (b) four head velocities ranging from 5 to 20 cm s^{-1} in steps of 5 cm s^{-1}, and (c) one direction of physical depth; the lower surface was always presented in front of the upper surface. We used only one direction of physical depth in the experiment, because no different tendency of the results from different directions of physical depths was observed in experiment 1.

The observers' task was similar to that in experiment 1. They viewed the stimulus in real time while moving their head forward and backward, through an extent of 25 cm by swaying the body back and forth at the same position, in synchrony with the beeps of an electric metronome. The frequency of the beeps was set at 0.2, 0.4, 0.6 and 0.8 Hz corresponding to head velocities of 5, 10, 15 and 20 cm s^{-1}, respectively. To keep this head-movement control accurate and precise, we installed head-movement stoppers at both ends of the 25.0 cm excursion, and also had observers practice the head movement sufficiently before experimental sessions. The practice for each observer was continued until both the experimenter and the observers felt that the extent of head movement and its synchronization with the metronome was precise and accurate. The CCD camera, positioned a mean distance of 34.5 cm from the center of the stimulus, moved through the same extent and at the same velocity as the observer's head. As in experiment 1, the observer reported which surface appeared in front and matched the magnitude of the perceived depth by manually adjusting the distance between two pieces of cardboard which were arranged parallel to each other and attached to a metal rod orthogonally.

Each observer completed a total of 100 trials, performed in five blocks of 20 trials each. Within each block, each of the 20 possible (physical depth magnitudes) × (head-velocity conditions) was presented once. The presentation order was randomized both within and between blocks.

3.2 *Results and discussion*

The results from each observer (magnitude of perceived depth, for each head velocity, as a function of the magnitude of physical depth) are presented in figure 3. The results show that (a) the magnitude of perceived depth was always the smallest in the lowest-head-velocity (5 cm s^{-1}) condition, except at 20 cm of physical depth for observer KM, and (b) the magnitude of perceived depth increased with that of the physical depth as shown in experiment 1. As described above, the perceived depth is relatively small when the head moves slowly (5 cm s^{-1}) and relatively large when the head moves quickly (10 to 20 cm s^{-1}), which is clear for observers TY and AK, and somewhat ambiguous for KM. We performed a two-way repeated-measures analysis of variance on the data from each observer using head velocity (5, 10, 15, and 20 cm s^{-1}) and physical depth (8, 12, 16, 20, and 24 cm) as factors. The analysis showed a significant effect of head-velocity condition for observers TY ($F_{3,80} = 117$, $p < 0.001$) and AK ($F_{3,80} = 8.49$, $p < 0.001$), but not for KM ($F_{3,80} = 2.02$, $p > 0.10$). The effect of physical depth was significant for all observers (for TY, $F_{4,80} = 677$, $p < 0.001$; for AK, $F_{4,80} = 216$, $p < 0.001$; and for KM, $F_{4,80} = 108$, $p < 0.001$). The interaction was significant for TY ($F_{12,80} = 6.73$, $p < 0.001$), but not for AK ($F_{12,80} = 1.72$, $p > 0.05$) and KM ($F_{12,80} = 0.323$, $p > 0.20$).

To illustrate more clearly the effect of head velocity on the magnitude of perceived depth, we recalculated the data in terms of perceived-depth ratio. The perceived-depth ratio is the ratio of magnitude of perceived depth to that of physical depth. The ratio was averaged (by geometrical mean because of averaging ratios) for each head velocity for each observer. The calculated ratios are plotted as a function of head velocity for each observer (see figure 4). It can be seen that magnitude of perceived depth is clearly smaller in the 5 cm s^{-1} head-velocity condition than in the other head-velocity conditions for TY and AK, while the magnitude was almost the same for all the head-velocity conditions (although the value at 5 cm s^{-1} is the smallest) for KM.

We speculate that the difference, for the lowest head velocity and the others, in the magnitude of perceived depth as a function of head velocity, clearly shown for the two observers, is the result of the differing vestibular signals associated with the acceleration component of the different head velocities. Although the observers were asked to move their heads at as constant a speed as possible, it is probable that their head

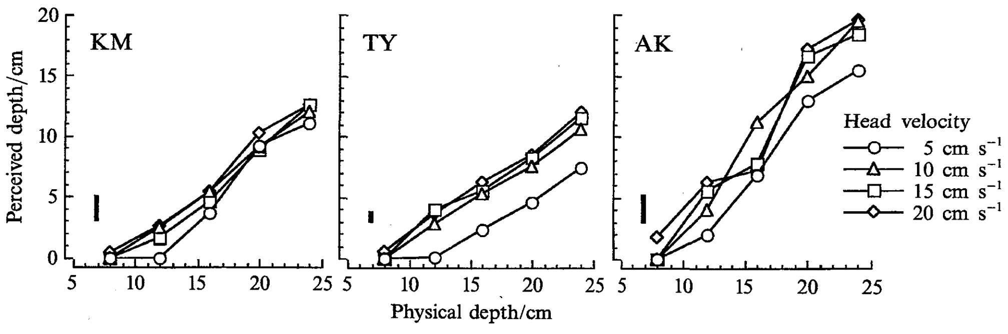


Figure 3. The magnitude of perceived depth, with the four different head velocities, as a function of the magnitude of physical depth. The magnitude of perceived depth, for observers TY and AK, was clearly smaller in the 5 cm s^{-1} velocity condition than in the other head-velocity conditions. The vertical bars represent the average SD.

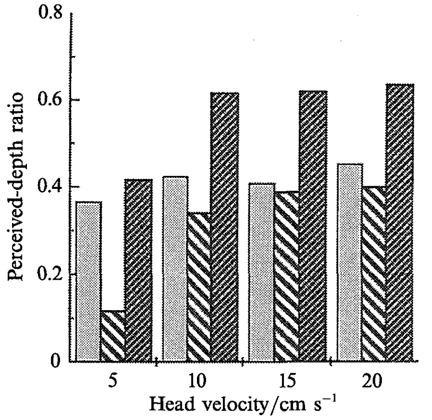


Figure 4. The mean ratio of perceived depth to physical depth with different head velocities. For observers TY and AK, the magnitude of perceived depth was clearly smaller in the 5 cm s^{-1} head-velocity condition than in the other velocity conditions, which confirmed the finding shown in figure 3.

movements were sinusoidal. If this is so, then the values of the peak acceleration would be approximately 20 cm s^{-2} for the slowest head velocity (5 cm s^{-1}) and 79 cm s^{-2} for the second-slowest one (10 cm s^{-1}). Given that the former value is close to the 10 cm s^{-2} threshold value of the vestibular system (see Howard 1982), this may be an indication that as one approaches the limits of the vestibular system the perception of depth produced by head movements is affected. This same implication may also apply to the data of Ujike and Ono (1992), and Ono and Ujike (1993), who measured parallactic depth thresholds and equal depth contours of motion parallax with different head velocities.

If we accept the above speculation, then the individual difference for the head-velocity effect on the magnitude of perceived depth can be explained by individual difference of acceleration threshold for the vestibular system. As mentioned above, the peak acceleration value for the lowest-head-velocity condition would be approximately 20 cm s^{-2}, while the threshold value of the vestibular system was reported as about 10 cm s^{-2}. Thus, the lowest-head-velocity condition may not be enough to show the effect with all the observers if there are individual variations, albeit small, of vestibular threshold. Another possible reason explaining the individual difference may be the failure of head-movement control; we, however, do not think this could be the main reason. Although a little variation of head velocity may have existed, all observers had practiced well moving their heads in synchrony with the metronome, which is supposed to reduce the variation as far as possible.

An alternative possible cause of the reduced extent of perceived depth in the slowest-head-velocity condition may be that the relative-velocity gradient across the depth boundary was near or below the relative-motion threshold and thus too weak to produce depth perception effectively. We, however, do not think this plausible, because of the following two reasons: (a) the relative-velocity gradient across the depth boundary could be detected well even for the smallest physical depth (12 cm) in the slowest-head-velocity condition, and (b) the fact that the extent of perceived depth increased with physical depth even in the slowest-head-velocity condition reveals that the relative-velocity gradient for larger physical depth was also well detected. To verify that the relative-velocity gradient was well detected, we calculated the relative velocity at several boundary positions between the two stimulus surfaces separated by the smallest physical depth (12 cm) in the slowest-head-velocity (5 cm s^{-1}) condition. When the observer was at the middle of the excursion, the relative velocity was approximately 0.26 deg s^{-1} at 5 deg from the center of the stimulus, 0.52 deg s^{-1} at 10 deg, and 1.04 deg s^{-1} at 20 deg; the velocity ratio across the boundary was more than 1.4 at any boundary position except the center of the stimulus. The values of the relative velocity are well above the relative-motion threshold, reported to be much less than 0.1 deg s^{-1} (eg Johnston and Wright 1985).

4 Experiment 3

In experiment 3, we examined the interaction between relative EC image motion yoked to forward and backward head movements and another monocular depth cue, motion parallax. We measured the stability and magnitude of perceived depth when the cues of relative EC image motion and motion parallax were presented simultaneously.

In experiments 1 and 2, both the lower and the upper thresholds for depth produced by relative EC image motion were greater than those for depth produced by motion parallax. Specifically, larger extents of physical depth were required to perceive depth from relative EC image motion than from motion parallax, and the upper limit of depth from relative EC image motion, more than 10 cm in our experimental conditions, was greater than that for motion parallax. Given these differences, experiment 3 was designed to determine how the magnitude of perceived depth is determined when the depth cues of both relative EC image motion and lateral shearing motion yoked to observers' head movements are presented simultaneously. To achieve this, we measured the stability and magnitude of perceived depth in the following three conditions: (a) a motion-parallax condition in which image motion was yoked to the observers' leftward and rightward head movements, (b) an EC-motion condition in which image motion was yoked to the observers' forward and backward head movements, and (c) a combination condition in which image motion was yoked to the observers' circular head movements.

4.1 *Method*

Three observers (KM, TY, and AK) participated in this experiment. They had all participated previously in experiment 2, and all but TY were naive as to the purpose of the experiment.

The apparatus and the stimuli were identical to those used in experiment 1, with the following exception. To allow for lateral head movements for the motion-parallax cue, a second timing belt and rotary encoder were attached to the helmet housing the monitor in the observer booth. Also, a second stepping motor was attached to the CCD camera in the stimulus booth to yoke the motion of the camera to the observers' lateral head movements.

The observers' task was as follows. In the motion-parallax condition, they were asked to move their head laterally within 80 cm (because of a mechanical limit) and at a velocity that felt comfortable. During this head movement, the CCD camera, which was positioned 26 cm from the center of the stimulus, moved in synchrony with the observer's head. In the EC-motion condition, observers were asked to move their head forward and backward within 82 cm (because of a mechanical limit), and at a velocity that felt comfortable. During this head movement the CCD camera moved in synchrony with the observer's head through the range of distances from 26 to 108 cm from the center of the stimulus. In the combination condition, observers were asked to move their head in a circular pattern within a diameter of 80 cm (because of a mechanical limit) and at a circular velocity that felt comfortable. During this head movement, the CCD camera moved in synchrony with the observer's head and the smallest distance between the camera and the center of the stimulus was 26 cm. As in the previous two experiments, the observers reported which surface appeared in front and matched the magnitude of the perceived depth by manually adjusting the distance between two pieces of cardboard which were arranged parallel to each other and attached to a metal rod orthogonally.

In the three experimental conditions observers were asked to move their head to a comfortable extent and at a comfortable velocity. The reason was that (a) if head velocity is not slow (ie is > 10 cm s^{-1} for both relative EC motion and motion parallax; see Ono and Ujike 1993 for the latter), magnitude of perceived depth does not depend on head velocity, (b) when the observers were asked to move their head comfortably, they moved their head clearly faster than 10 cm s^{-1}, and (c) head-movement control for a circular pattern was rather difficult when active movement was also needed.

Each observer completed a total of 150 trials, performed in fifteen blocks of 10 trials each. Within each block the stimulus was presented once at each of ten different physical depths, from 4 to 22 cm in steps of 2 cm, in one of the three head-movement conditions; there were five blocks for each head-movement condition. The presentation order was randomized both within and between blocks.

4.2 *Results and discussion*

The results from each observer (magnitude of perceived depth, for each condition, as a function of the magnitude of physical depth) are presented in figure 5. It can clearly be seen that (a) the magnitude of perceived depth in the motion-parallax condition increased until approximately 16 cm of physical depth, and then decreased toward zero at approximately 22 cm of physical depth, (b) no depth was perceived in the EC-motion condition when the magnitude of the physical depth was less than about 8 cm, and then the magnitude of perceived depth increased monotonically with physical depth, (c) the magnitude of perceived depth was larger in the EC-motion condition than in the motion-parallax condition for physical depths greater than about 16 cm, while the reverse was true for physical depths less than about 16 cm, and (d) the magnitude of perceived depth in the combination condition was approximately equal to the greater of the two perceived depths in the motion-parallax and the EC-motion conditions.

We confirmed statistically that the magnitude of perceived depth for physical depths smaller than 18 cm in the combination condition was equal to that in the motion-parallax condition, but not to that in the EC-motion condition. We performed a two-way repeated-measures analysis of variance on the data from each observer using head-movement condition and physical depths less than 18 cm as factors. This analysis was done for each possible pair of the three head-movement conditions. For all three observers the main effect of head-movement condition was not significant for the pair consisting of the motion-parallax and the combination conditions

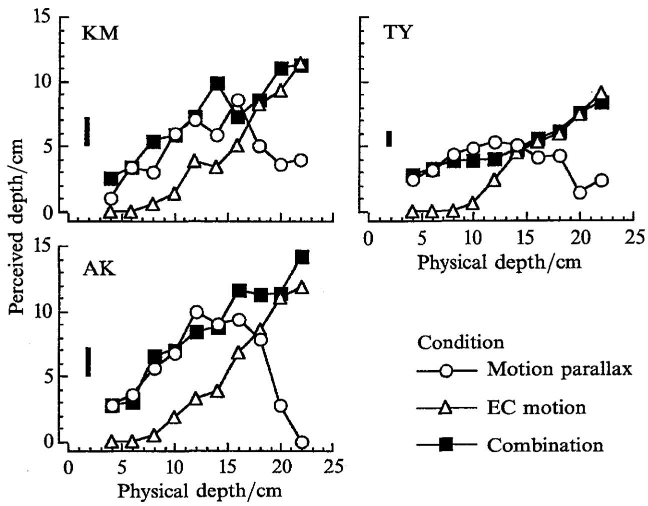


Figure 5. The magnitude of perceived depth in the motion-parallax (open circles), the EC-motion (open triangles), and the combination (filled squares) conditions as a function of the magnitude of physical depth. For all three observers (KM, TY, AK), the magnitudes of perceived depth in the combination and the motion-parallax conditions were approximately equal with slower head movements (<16 cm s^{-1}), while those in the combination and the EC-motion conditions were approximately equal with faster head movements. The vertical bars represent the average SD.

($F_{1,56} = 0.0988$, $p > 0.20$ for AK; $F_{1,56} = 3.57$, $p > 0.05$ for KM; and $F_{1,56} = 0.655$, $p > 0.20$ for TY), but was significant for the other pairs of conditions ($p < 0.001$).

These findings indicate that the visual system adopts the depth cue which yields the largest or veridical depth, and yields a stable perception of depth through a wider range of physical depth. This type of integration can not be classified as either weighted linear averaging or a simple cue domination of one cue over the other. Moreover, this integration seems to occur to supplement a weak output based on one cue with another output based on the other cue. A similar type of interaction, that motion parallax becomes relatively ineffective with larger simulated depth, was reported by Ono et al (1988), who combined motion parallax and dynamic occlusion. They made the directions of the simulated depth from each cue in conflict and measured the percentages of depth direction consistent with occlusion cue. They found that perceived depth order was determined by motion parallax when the simulated depth was smaller (equivalent disparity < 25 min of arc) and the perceived depth order was determined by dynamic occlusion when the simulated depth was larger (equivalent disparity > 25 min of arc).

The type of interaction found in the experiment is interesting because it seems to represent a strategy to perceive a spatial layout in everyday situations. There are reports in the literature that magnitudes of perceived depth from binocular disparity and motion parallax, for example, have some maximum values that are not so large (Tyler 1983; Ono and Ujike 1993); the values are not enough to explain the perception of spatial layout we experience in everyday situations. Even if other cues, such as relative EC motion and dynamic occlusion, for perceiving larger depth exist, cue averaging and simple cue domination will yield perception either of smaller depth or of larger depth. In that sense, the interaction found in the experiment is reasonable and appropriate to recognize any spatial layouts, although we do not suggest how the visual system learns and realizes the strategy.

5 Summary

The main results of our study are the following two points. First, magnitudes of perceived depth produced from relative EC image motion were greater when the motion was yoked to forward and backward head movements than when it was not, and magnitudes of perceived depth decreased with head movements of less than 10 cm s^{-1}. These different magnitudes of perceived depth with and without head movement or with different head velocities seem to be led by nonvisual information of self-motion, because visual information per se did not differ with and without head movement, and head movement with a velocity of 5 cm s^{-1}, which led to a smaller amount of perceived depth, may produce a weak vestibular signal. Second, when depth cues of relative EC image motion yoked to head movement and of motion parallax were simultaneously presented, the resultant magnitude of perceived depth was equal to the greater of two perceived depths, each of which was produced from either motion parallax or relative EC image motion alone; the cue interaction was neither weighted linear averaging nor simple cue domination.

Acknowledgements. We appreciate very helpful comments by Hiroshi Ono on the earlier version of the manuscript and improvements to its English by Alistair Mapp.

References

Bingham G P, Stassen M G, 1994 "Monocular egocentric distance information generated by head movement" *Ecological Psychology* **6** 219 – 238

Buckley D, Frisby J P, 1993 "Interaction of stereo, texture and outline cues in the sharp perception of three-dimensional ridges" *Vision Research* **33** 919 – 933

Bülthoff H H, Mallot H A, 1988 "Integration of depth modules: stereo and shading" *Journal of the Optical Society of America A* **10** 1749 – 1758

Cornilleau-Pérès V, Droulez J, 1994 "The visual perception of three-dimensional shape from self-motion and object-motion" *Vision Research* **34** 2331–2336
Damme W J M van, Grind W A van de, 1996 "Non-visual information in structure-from-motion" *Vision Research* **36** 3119–3127
Dosher B A, Sperling G, Wurst S A, 1986 "Tradeoffs between stereopsis and proximity luminance covariance as determents of perceived 3D structure" *Vision Research* **26** 973–990
Eriksson E S, 1974 "Movement parallax during locomotion" *Perception & Psychophysics* **16** 197–200
Ferris S H, 1972 "Motion parallax and absolute distance" *Journal of Experimental Psychology* **95** 258–263
Gonzalez E G, Steinbach M J, Ono H, 1989 "Depth perception in children enucleated at an early age" *Clinical Vision Sciences* **4** 173–177
Gregory R L, Ross H E, 1964a "Visual constancy during movement I" *Perceptual and Motor Skills* **18** 3–8
Gregory R L, Ross H E, 1964b "Visual constancy during movement II" *Perceptual and Motor Skills* **18** 23–26
Harris L R, Morgan M J, Still A W, 1981 "Moving and the motion after-effect" *Nature (London)* **293** 139–141
Heidenreich S M, Turano K A, 1996 "Speed discrimination under stabilized and normal viewing conditions" *Vision Research* **36** 1819–1825
Howard I P, 1982 *Human Visual Orientation* (New York: John Wiley)
Ichikawa M, Saida S, 1996 "How is motion disparity integrated with binocular disparity in depth perception?" *Perception & Psychophysics* **58** 271–282
Johansson G, 1973 "Monocular movement parallax and near-space perception" *Perception* **2** 135–146
Johnston A, Wright M J, 1985 "Lower thresholds of motion for gratings as a function of eccentricity and contrast" *Vision Research* **25** 179–185
Johnston E B, Cumming B G, Parker A J, 1993 "Integration of depth modules: stereopsis and texture" *Vision Research* **33** 813–826
Lee D N, 1976 "A theory of visual control of braking based on information about time-to-collision" *Perception* **5** 437–459
Norman J F, Todd J T, 1995 "The perception of 3-D structure from contradictory optical patterns" *Perception & Psychophysics* **57** 826–834
Ono M E, Rivest J, Ono H, 1986 "Depth perception as a function of motion parallax and absolute-distance information" *Journal of Experimental Psychology: Human Perception and Performance* **12** 331–337
Ono H, Rogers B J, Ohmi M, Ono M E, 1988 "Dynamic occlusion and motion parallax" *Perception* **17** 255–266
Ono H, Steinbach M J, 1990 "Monocular stereopsis with and without head movement" *Perception & Psychophysics* **48** 179–187
Ono H, Ujike H, 1993 "Equal-depth contours as a function of different velocities of head movement" *Perception* **22** Supplement, 81
Regan D, Hamstra S J, 1993 "Dissociation of discrimination thresholds for time to contact and for rate of angular expansion" *Vision Research* **33** 447–462
Rogers B J, Collett T S, 1989 "The appearance of surfaces specified by motion parallax and binocular disparity" *Quarterly Journal of Experimental Psychology* **41A** 697–717
Rogers B, Graham M, 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
Rogers S, Rogers B J, 1992 "Visual and nonvisual information disambiguate surfaces specified by motion parallax" *Perception & Psychophysics* **52** 446–452
Steinbach M J, Ono H, Wolf M E, 1991 "Motion parallax judgments of depth as a function of the direction and type of head movement" *Canadian Journal of Psychology* **45** 92–98
Todd J, 1981 "Visual information about moving object" *Journal of Experimental Psychology: Human Perception and Performance* **7** 795–810
Tyler C W, 1983 "Sensory processing of binocular disparity", in *Vergence Eye Movements: Basic and Clinical Aspects* Eds C M Schor, K J Ciuffreda (Boston, MA: Butterworth) pp 199–296
Ujike H, Ono H, 1992 "Two different limits of parallactic depth" *Perception* **21** Supplement 2, 86
Wallach H, Flaherty E W, 1975 "A compensation for field expansion caused by moving forward" *Perception & Psychophysics* **17** 445–449