Asymmetry in the perception of motion in depth induced by moving cast shadows

Tomoko Imura
Language and Intelligence Section, Primate Research Institute, Kyoto University, Inuyama, Aichi, Japan, & Japan Society for the Promotion of Science, Japan

Nobu Shirai
Department of Psychology, Tokyo Metropolitan University, Hachioji, Tokyo, Japan, & Japan Society for the Promotion of Science, Japan

Masaki Tomonaga
Language and Intelligence Section, Primate Research Institute, Kyoto University, Inuyama, Aichi, Japan

Masami K. Yamaguchi
Department of Integrated Psychological Science, Kwansei Gakuin University, Nishinomiya, Hyogo, Japan, & Language and Intelligence Section, Primate Research Institute, Kyoto University, Inuyama, Aichi, Japan

Akihiro Yagi

An expanding object, which may represent an approaching motion, is easier to detect than a contracting one, which may represent a receding object. To confirm the generality of asymmetry in the detection of approaching and receding motions, we focused on the perception of apparent motion in depth created by moving cast shadows. The visual search for an approaching target among receding distractors was more efficient than for the opposite condition (Experiment 1). However, this asymmetry disappeared when a light shadow was added (Experiments 2 and 3). This suggests that the visual system is specialized to detect approaching motion defined by cast shadows, as well as other three-dimensional cues such as expanding motion and shading.

Keywords: asymmetry, cast shadow, motion in depth, visual search


Introduction

One of the most important functions of the visual system is to avoid dangerous situations such as a collision. Humans use various cues, including expanding motion, binocular disparity, and pictorial cues, to detect approaching objects toward the eyes and our visual system integrates information from these multiple depth cues. The expansion and contraction of the retinal image are useful motion-in-depth cues (Beverley & Regan, 1979; De Bruyn & Orban, 1990; Gibson, 1979; Johansson, 1964; Regan & Beverley, 1978; Regan & Hamstra, 1993; Schiff, 1965; Schiff, Caviness, & Gibson, 1962). Furthermore, some studies have indicated that the human visual system has a specific sensitivity to expanding motion (Takeuchi, 1997) and that this expansion-detection bias emerges very early in life (Shirai, Kanazawa, & Yamaguchi, 2004).

A bias toward expansion was originally reported by Schiff and his colleagues. They demonstrated that defensive responses to expansion of a form contrasted with responses to contraction. In their study, monkeys (Schiff et al., 1962) and other animals (Schiff, 1965) generated avoidant body movements, only to radial expanding motion when the dark form underwent a symmetrical hyperbolic expansion beyond 30 degrees of visual angle. Later developmental studies also showed that 1-month-old or even younger human infants made defensive motor responses such as blinking or head movements to a large field radial expansion pattern (e.g., Ball & Tronick, 1971; Ball, Ballot, & Dibble, 1983; Bower, Broughton, & Moore, 1970; Nanez, 1988; Nanez & Yonas, 1994; Yonas,
Pettersen, & Lockman, 1979), and these defensive responses to apparent collision events increase with age in the first few months of life (Nanez & Yonas, 1994). Takeuchi (1997) presented an example of the expansion-detection bias in search asymmetry for expansion and contraction. He showed that the detection of an expanding object among contracting distractors was easier than detection of a contracting target among expanding distractors in visual search tasks. This suggests that the visual system is more sensitive to expansion, which can represent an approaching object, than to contraction, which can represent a receding object. In addition, three-dimensional shapes defined by shading cues affect the perception of expanding and contracting motion patterns (Shirai & Yamaguchi, 2004). Visual search involving expanding and contracting disks with shading showed an anisotropic sensitivity between the detection of expanding convex circles and the detection of expanding concave ones, whereas there was no anisotropy between the detection of contracting convex and contracting concave disks. Shirai and Yamaguchi (2004) suggested that the concave shape impaired the impression of approaching motion, causing this anisotropy. They also suggested that the human visual system is specialized to detect approaching motion defined by multiple depth cues.

However, there is a possibility that the reported asymmetry in the detection of motion in depth was specific to the stimuli, involving expansion and contraction and shading cues. Therefore, it is necessary to explore whether asymmetry of perception of motion in depth occurs regardless of the type of visual information.

We explored the perception of apparent motion in depth induced by moving cast shadows. A cast shadow is an effective cue for the spatial layout of objects relative to the background (Yonas, Goldsmith, & Hallstrom, 1978). In particular, the motion of cast shadows provides useful information about the relative motion of objects in a scene (Imura et al., 2006; Kersten, Knill, Mamassian, & Bülthoff, 1996; Kersten, Mamassian, & Knill, 1997; Kim & Grocki, 2006). Kersten et al. (1996) provided evidence that the motion of a cast shadow creates the impression of an object moving in depth, even if the object itself does not change in size and position. An object appears to approach observers when a cast shadow detaches from the object, whereas it appears to recede when the shadow comes close to the object, even when the size of the object is fixed. These findings suggest that moving cast shadows, as well as the relative size of the retinal image and binocular disparity cue (Regan & Beverley, 1979), influence the perception of motion in depth. In addition, moving white shadows do not create a three-dimensional impression by the manipulation of their luminance, suggesting that the shadow should be darker than the background to obtain this effect (Kersten et al., 1997).

Using the motion of dark cast shadows, we created stimuli that were perceived as apparently approaching or receding from observers. In Experiment 1, we examined the anisotropy of perceived approaching and receding motions defined by the moving cast shadows. In Experiments 2 and 3, we manipulated the luminance of shadows to confirm whether the findings of Experiment 1 were caused by the three-dimensional impression created by cast shadows. Finally, we discuss whether the human visual system is specialized to detect approaching objects.

### Experiment 1

#### Methods

**Participants**

Seven undergraduate students (average age = 22.2 years) participated in the experiment. All had normal or corrected-to-normal visual acuity and color vision, and all were naïve as to the purpose of the experiment before participating.

**Apparatus**

Stimuli were presented on a 21-inch color CRT monitor (Nanao Eizo FlexScan T966) controlled by a personal computer (Dell, Dimension, XPST600). The resolution of the CRT monitor was set at 1024 × 768 pixels with an 8-bit color mode. The refresh rate of the CRT monitor was 75 Hz. The response was recorded through a keyboard. Participants wore an eye patch on the left eye to maintain a monocular viewing condition. Chin and head rests maintained the viewing distance from the CRT monitor to a participant at 57 cm.

**Stimuli**

The stimuli consisted of static blue squares (8.7 cd/m²) placed on moving gray squares (12.7 cd/m²) with blurred outlines that could be interpreted as shadows cast by lighting from the top-left side (Figure 1). Shadow motions can create a strong perception of approaching or receding.
of objects (Kersten et al., 1996); in our experiment, each shadow moved from the top left to bottom right of the monitor. The area of each shadow was increased or decreased by 37% per 60 ms. The sizes of the squares were randomly varied in each trial from 1.37 to 1.72. The initial areas of the shadows were also randomized to prevent participants from searching based on a particular stimulus size (Figure 1). The background was a uniform gray surface (82.7 cd/m²), and the Michelson contrast of shadows for the background was 73.4%. Shadows were generated using Shade 6.0 and retouched using Adobe Photoshop.

Displays consisted of 12, 24, or 36 items, which were arranged randomly on the 23.4° x 23.4° display area, with possible locations selected such that items never overlapped. The position of each item was “jittered” by ±1.95° to minimize possible effects of item arrangement.

**Procedure**

A trial began with the presentation of a fixation point (black cross subtending 0.84°). After participants pressed the space key, the fixation point disappeared and the stimulus display was presented. The presentation of the stimulus display was fixed for 600 ms. The subjects were asked to search for the unique motion pattern and to respond to the presence or absence of the target by pressing one of two keys. Participants were required to respond as quickly as possible while avoiding errors. The error rate was measured, and the response time was defined as the time from the onset of the stimulus display until a response was made.

Each participant was tested under two conditions: the approaching condition, in which a square with an expanding shadow was contained among different sizes of squares with contracting shadows; and the receding condition, in which a square with a contracting shadow was contained among different sizes of squares with expanding shadows.

**Results**

For target-present trials, the mean error rate as a function of the number of items for each condition was lower in the approaching condition than the receding condition, regardless of the number of items (Figure 2). Two-way repeated-measures analyses of variance (ANOVAs) were performed on the error rates, with condition (approaching or receding) and number of items as main factors. There was a significant main effect of condition ($F(1, 6) = 10.86, p < 0.05$).

The mean correct response times (RTs) for target-present trials as a function of the number of items for each condition in Experiment 1. Error bars indicate mean standard errors.

A session consisted of two blocks: one for the approaching condition and one for the receding condition. Each block contained 120 trials, half of which had a target. The number of items and the presence or absence of the target varied in each trial. Each participant was given one practice session followed by two test sessions. The order of these blocks was counterbalanced across participants.

The error rates in the approaching condition were lower than those in the receding condition, and the RTs were faster in the approaching than in the receding condition. These results are consistent with the asymmetry of the perception of motion in depth defined by radial expansion and contraction and shading cues representing 3-D shapes (Shirai & Yamaguchi, 2004). Our results extend the findings of anisotropy in the human visual system for the
perception of approaching objects defined by moving cast shadows and suggest that this asymmetry occurs across different motion-in-depth cues. However, there are two other possible interpretations of these findings. One is that an expanding object, like a shadow, is easier to detect than a contracting object, regardless of the impression of approach. The other is that the impression of approach defined by the relationship between shadows and objects causes the asymmetry in detection. To examine these possibilities, in Experiment 2, we changed the luminance of the shadows to manipulate the effectiveness of shadows as depth cues. If the findings of Experiment 1 depended on an impression of approach based on the relationship between shadows and objects, expanding but light-colored shadows would not produce detection asymmetry. That is, in Experiment 2, we examined whether the findings of Experiment 1 were caused by the three-dimensional impression created by cast shadows.

**Experiment 2**

**Methods**

**Participants**

Seven undergraduate students (average age = 22.2 years) participated in the experiment. All had normal or corrected-to-normal visual acuity and color vision, and all participants also took part in Experiment 1.

**Apparatus**

The same apparatus used in Experiment 1 was used in Experiment 2.

**Stimuli**

The stimuli were identical to those used in Experiment 1, except that the luminance of the shadows was lighter than the background (114.7 cd/m²; Figure 4). Although the light shadows moved away from the squares, the blue squares could not be perceived as approaching or receding. Previous studies have shown that white shadows are less effective than dark shadows in creating the perception of motion in depth (Elder, Trithart, Pintilie, & MacLean, 2004; Kersten et al., 1996; Rensink & Cavanagh, 2004). The Michelson contrast of the shadows in relation to the background was 16.2%. Shadows were generated using three-dimensional modeling and animation software (Shade 6.0) and retouched using Adobe Photoshop.

**Procedure**

The procedure was identical to that used in Experiment 1. Each participant was tested under two conditions: the approaching condition, in which a square with an expanding white shadow was contained among different sizes of squares with contracting white shadows; and the receding condition, in which a square with a contracting white shadow was contained among different sizes of squares with expanding white shadows. The design was identical to that of Experiment 1.

**Results**

For target-present trials, there were no differences in mean error rates as a function of the number of items for each condition between the approaching condition and the receding condition (Figure 5). Two-way repeated-measures ANOVAs were performed on the error rates, with condition and number of items as main factors. There were no significant main effects or interactions. In addition, we made direct comparisons of advantages of “approaching” over “receding” between Experiments 1 and 2 to clarify the effects of dark shadows. We calculated
the average of differences in error rates between “approaching” and “receding” conditions regardless of display size in every participant and conducted non-paired t-test. There were significant differences in advantages of “approaching” over “receding” between experiments, suggesting that differences in error rates between “approaching” and “receding” conditions were larger in Experiment 1 than in Experiment 2 (t(9) = 2.68, p < 0.05, two-tailed).

The mean correct RTs for target-present trials as a function of the number of items for each condition are shown (Figure 6). Two-way repeated-measures ANOVAs were performed on the correct RTs, with condition and number of items as main factors. There were no significant main effects or interactions. The mean standard error suggests that these results were caused by large variance among participants. We also made between-experiments comparisons of advantages of “approaching” over “receding” using differences in RTs between both conditions. A non-paired t-test with experiments revealed significant differences in advantages of “approaching” over “receding” (t(9) = 3.74, p < 0.01, two-tailed).

There were no significant differences between the approaching and receding conditions when the shadows had lighter luminance than the background. This suggests that the asymmetry observed in Experiment 1 is not explained by an advantage in the detection of expanding objects. Rather, the spatial relationship between dark shadows and objects provides depth information regarding the objects. Together, the experiments show anisotropic sensitivity in the detection of approaching and receding objects defined by moving cast shadows. However, because the Michelson contrast of the shadows in Experiment 2 was lower than that in Experiment 1, the contrast may not have been sufficient for target detection. To exclude this possibility, in Experiment 3, we used a display that reversed the tones in Experiment 1.

### Experiment 3

#### Methods

**Participants**

Four undergraduate students (average age = 26.5 years) participated in the experiment. All had normal or corrected-to-normal visual acuity and color vision, and all were naive as to the purpose of the experiment before participating.

**Apparatus**

The same apparatus used in Experiments 1 and 2 was used in Experiment 3.

**Stimuli**

The stimuli were identical to those of Experiment 1, except that the tones of the stimuli were reversed (Figure 7). Although the contrasts among the background, squares, and shadows were identical to those of Experiment 1, light shadows were moved away from the square, and blue squares were not perceived as approaching or receding. The Michelson contrast of shadows against the background was 86.1%, which was almost equal to that of Experiment 1. Stimuli were retouched using Adobe Photoshop.

**Procedure**

The procedure and design were identical to those used in Experiments 1 and 2.

**Results**

For target-present trials, there were no differences in mean error rates as a function of the number of items for each condition between the approaching condition
and the receding condition (Figure 8). Two-way repeated-measures ANOVAs were performed on the error rates, with condition and number of items as main factors. There were no significant main effects or interactions. In addition, we made direct comparison of the advantage of “approaching” over “receding” between Experiments 1 and 3 using average of differences in error rates between “approaching” and “receding” conditions. A non-paired $t$-test revealed significant differences between experiments ($t(9) = 2.52, p < 0.05$, two-tailed).

The mean correct RTs for target-present trials as a function of the number of items for each condition are shown (Figure 9). The search for an approaching square was faster than that for a receding square. Two-way repeated-measures ANOVAs were performed on the correct RTs, with condition and number of items as main factors. There were no significant main effects or interactions. We also compared the advantage of “approaching” over “receding” in Experiment 3 with that in Experiment 1 using average of differences in RTs between “approaching” and “receding” conditions. A non-paired $t$-test showed significant difference between experiments ($t(9) = 2.32, p < 0.05$, two-tailed).

Although the Michelson contrast of shadows in relation to the background was identical to that in Experiment 1, there were no differences in error rates or RTs between approaching and receding conditions in Experiment 3. These results indicate that the anisotropy disappears when the shadow luminance is lighter than the background.

**Discussion**

Our results indicate that the detection of an “approaching” target defined by moving cast shadows was faster than the detection of a “receding” target (Experiment 1). In addition, there was no difference between the detection of an object with an expanding light shadow and one with a contracting light shadow (Experiments 2 and 3).

Our findings are consistent with previous findings that objects moving toward an observer, represented by the interaction between radial motion and shading cues, are detected faster than those moving away from an observer (Shirai & Yamaguchi, 2004). The motion of cast shadows created the impression of apparent motion in depth by the object in our study, and the search for approaching objects was faster than that for receding objects. Our results provide more evidence that the visual system may be specifically sensitive to approaching motion.

The search performance observed differed slightly from that reported by Takeuchi (1997), who also found search asymmetry for expansion and contraction, the basic visual information of motion-in-depth perception. Takeuchi (1997) found that the search time for an expanding target did not depend on the number of items, whereas that for a contracting target became longer as the number of distractors increased. He found search asymmetry between the detection of expanding objects and the detection of contracting objects. In contrast, we found that the search time for approaching and receding targets did not increase independent of the number of items, but the search time for an approaching target was shorter than that for a receding one. The process of detecting expansion and contraction may differ from that of detecting motion in depth defined by moving cast shadows. Studies exploring the effects of static cast shadows have shown conflicting findings regarding search slopes (Elder et al., 2004; Rensink & Cavanagh, 2004). Further systematic studies are needed to better understand the process of perceiving moving cast shadows.

In Experiments 2 and 3, we examined whether the asymmetry in detection was caused by depth impressions defined by moving cast shadows or simply by an advantage in the detection of expanding shadows. In these experiments, we used light shadows, which impair the
motion-in-depth impression (Elder et al., 2004; Kersten et al., 1996). The expanding light shadows did not facilitate the detection of a target. This evidence suggests that the asymmetry was caused not by sensitivity to expanding objects, but by the impression of motion in depth of objects caused by the moving cast shadows.

Previous studies have also found that depth impressions obtained from multiple depth cues, including motion and shading, influence the detection of “approaching” motion. Shirai and Yamaguchi (2004) showed the effects of shading information on the detection of expanding or contracting objects. They proposed that concavity defined by a shading cue impaired the detection of an expanding target. Taken together, these findings suggest that the visual system is specialized to detect “approaching” motion and that asymmetry of motion-in-depth perception might be caused by the higher visual area, which requires a combination of multiple depth cues.

Recent neurophysiological studies have examined neural mechanisms that mediate the processing of combinations of multiple depth cues: perspective and disparity (Liu, Vogels, & Orban, 2004; Tsutsui, Jiang, Yara, Sakata, & Taira, 2001; Welchman, Deubelius, Conrad, Bültthoff, & Kourtzi, 2005) and shading and disparity (Taira, Nose, Inoue, & Tsutsui, 2001). Such studies have provided evidence that an object’s three-dimensional shape may be represented independent of depth cues in macaque inferior temporal cortex (Liu et al., 2004) and humans lateral occipital complex (Kourtzi & Kanwisher, 2001). Functional magnetic resonance imaging studies in humans (Shikata et al., 2001; Taira et al., 2001) and monkeys suggest that a group of neurons in the occipitotemporal and parietal regions is involved in the detection of three-dimensional objects defined by multiple depth cues. In addition, well-known motion-sensitive cortical areas, i.e., MT/MST, may compute the surface based on shading cues (Kourtzi, Bültthoff, Erb, & Grodd, 2002). Considering these findings, it is plausible that there are some groups of neurons that respond to apparent object motion, as defined by motion and cast shadows.

On the other hand, there have been arguments that motion in depth defined by cast shadows can compute in different reference frame from other depth cues (Schrater & Kersten, 2000). Schrater and Kersten (2000) pointed out that while the image size of expanding or contracting objects is determined by the depth of square from the observer and the physical size of the square, the cast shadow position is determined by the allocentric distance of the square from the background and the position of the light source. To combine the cast shadow with image size data, it might require a different type of mechanism from other depth cues. The present findings suggest an advantage to the detection of “approaching” objects defined by cast shadows as well as other depth cues, it is still unclear, however, whether depth from the multiple cues is processed in individual module or common module.

Our results suggest that the human visual system is specialized to detect an “approaching” object among “receding” objects defined by moving cast shadows. This specialization would be ecologically valid because avoiding dangerous situations such as collision with approaching objects is one of the most important functions for various species. Hence, the detection bias for approach observed here and in previous research might be an ecological function of the visual system.

Acknowledgments

This study and the preparation of this manuscript were financially supported by Grants-in-Aid for Scientific Research from the Japanese Ministry of Education, Culture, Sports, Science, and Technology (MEXT), Japan Society for the Promotion of Science (JSPS; Grants 13610086, 16300084, and 1704414), and MEXT Grants-in-Aid for the 21st Century COE Program (A14 and D10).

Commercial relationships: none.
Corresponding author: Tomoko Imura.
Email: imura@pri.kyoto-u.ac.jp.
Address: Language and Intelligence Section, Primate Research Institute, Kyoto University, Aichi 484-8506, Japan.

References


