PHENOMENAL REPORTS

Motion-driven enhancement of a lower region cue in depth perception

Yuki Kubota[1,2]*, Ryota Mima[1], Takahiro Kawabe[2], Taiki Fukiage[2] and Masahiko Inami[1]

[1]Information Somatics Laboratory, RCAST, The University of Tokyo, Tokyo, Japan; [2]NTT Communication Science Laboratory, Kanagawa, Japan

Abstract

The authors report a demonstration in which a motion-defined boundary enhances the effects of a lower region cue in depth perception. Although the lower region cue has been proposed as a potential depth cue, its effect is weak in the static image. Their demonstration reveals that the lower region is almost unambiguously perceived as being in front when defined by horizontal motion mimicking motion parallax. The authors further investigated phenomenological aspects of the lower region cue by combining it with other depth cues.

Keywords: *depth illusion*; *depth order*; *lower region cue*; *motion-defined boundary*

> To access the movies for this article, please visit the article where all movies are embedded.

Edited by:
Kohske Takahashi
Ritsumeikan University, Japan

Reviewed by:
Satoshi Shioiri, Tohoku University, Japan

Yuki Kobayashi
Ritsumeikan University, Japan

A lower region cue has been proposed as one of the figure-ground cues (Hulleman & Humphreys, 2004; Vecera & Palmer, 2006; Vecera et al., 2002). The lower region of images is more likely to be recognized as a figure than the upper region, which is independent of observer's eye movements, contrast, or voluntary spatial attention. However, when the image is rotated at 90° and divided on the left and right, this asymmetry disappears (Vecera et al., 2002). Subsequent studies have shown that the stimuli with a wider base and a narrow top were more frequently recognized as a figure (Hulleman & Humphreys, 2004). Furthermore, the regions attached to the walls with texture gradient or linear perspectives tended to be recognized as a figure (Vecera & Palmer, 2006).

Figure-ground assignments are closely related to depth perception (Grossberg, 1997; Vecera et al., 2002), suggesting that the lower region cue is a potential monocular depth cue. However, its effect is weak in the static image. For example, when the still image has a straight horizontal boundary, the lower region cue disappears or becomes weaker. In Fig. 1a with a non-straight boundary, the lower region appears to be a figure or is perceived as the front, whereas in Fig. 1b, it is challenging to identify which region is perceived as the front.

In this report, we present a demonstration in which the lower region cue can provide a significant effect in determining depth relationships. In our stimuli, the boundary between upper and lower regions is defined by horizontal motion mimicking motion parallax. Contrary to the static image cases in which figure-ground relationships can sometimes be reversed, in our demonstration the lower region is almost unambiguously perceived as being in front. The authors further investigate phenomenological aspects of the lower region cue by combining it with other monocular depth cues.

Basic finding

In Movie 1, two binary noise regions oscillate sinusoidally at 0.5 Hz in opposite phase with the same amplitude, 10% of the image width. The two regions consist of top/bottom regions with horizontal motion (Movie 1a), left/right regions with vertical motion (Movie 1b), and top/bottom regions with vertical motion (Movie 1c). A schematic diagram of the depth perception is shown on the right of each movie.

The main finding in this report is that in Movie 1a; the lower region of the movie is almost always perceived to be in front of the upper region. To confirm their phenomenal observations, the authors collected informal observations

---

1
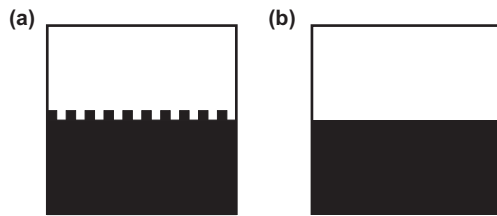
Fig. 1. Sample images related to the lower region cue with (a) a non-straight boundary and (b) a straight boundary. The former image was created with reference to the one in the previous study (Vecera et al., 2002).
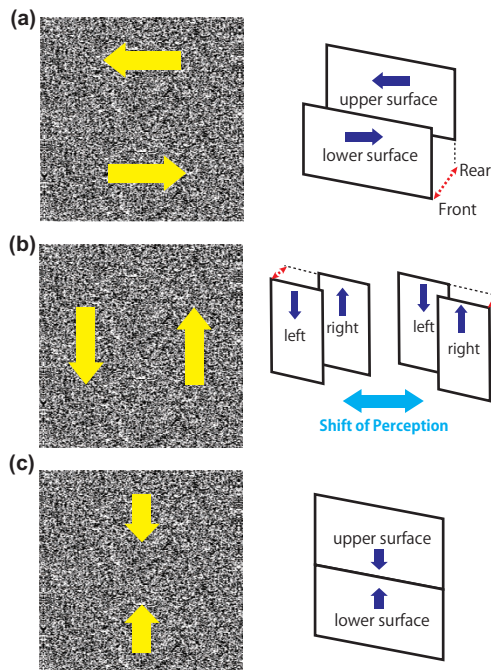
Movie 1. Stimulus movies with a motion-defined boundary. The movies consist of two regions, (a) top/bottom regions with horizontal motion, (b) left/right regions with vertical motion, and (c) top/bottom regions with vertical motion, oscillating at 0.5 Hz with the same amplitude (10% of the image width) and opposite phase. A schematic diagram of depth perception is shown on the right of each movie.

from 18 observers (Refer to Appendix for details about the experiment.). The results of the informal observations confirmed that most observers (97.2%) perceived the lower region as being in front.

However, in Movie 1b, which is a 90° rotation of Movie 1a (in this article, all angles are rotated in counter-clockwise rotation), the depth order of the left and right regions are bistable similar to Necker's cube (Necker, 1832). Specifically, in certain times, the left region appears to be in front of the right, while in other times, the right region appears to be in front of the left.

These observations are consistent with the characteristics of a lower region cue in figure-ground segmentation: (1) asymmetry in a figure-ground assignment in the upper and lower regions and (2) disappearance of this asymmetry when the movie is rotated at 90°. The demonstrations suggest that a lower region cue can contribute to depth identification in horizontal-motion-defined regions.

To make the lower region cue effective, it may be essential to move the two regions with horizontal motion mimicking motion parallax. In Movie 1c, which has vertical motion with a horizontal boundary, the depth order is more ambiguous than in Movie 1a. The two regions are sometimes perceived to be almost flat.

Furthermore, when the movie is observed from above by leaning forward, the same perception can be obtained: the region closer to the observer's body is perceived as being in front. Even when the movie is observed with monocular vision, the depth order is almost stable. The perceived depth disappears once the horizontal motion of the movie is stopped. Note that the movie size, the speed of the dots, and the gaze movement may affect the depth perception of Movies 1a–c, and one can find a specific condition in which the upper region appears to be in front. We leave these questions for a future study.

Dynamic versus static

The velocity cue is known as one of the depth cues (Braunstein & Andersen, 1981; Kaneko & Uchikawa, 1993); fast-moving regions are more likely to appear closer than the slow-moving regions. In the present study, the authors investigate whether the lower region cue or the velocity cue is dominant in determining depth perception of the illusion. Stimulus movies with one region static from Movie 1 are shown in Movie 2.

In Movies 2a and b, the majority of observers reported that the lower region was perceived to be in front of the upper region, even when the lower region was static (Movie 2a: 77.8%, 2b: 94.4%). When the movie was rotated at 90°, the dynamic region tended to appear in front of the static region (Movie 2c: 55.6%, 2d: 58.3%). The authors further confirmed this tendency in our observation. When the regions were separated with the horizontal boundary, the lower region was predominantly perceived as being in front and the reversal of depth order was less frequent compared with the 90°-rotated version of the stimulus.

These observations suggest that the lower region cue is more dominant than the velocity cue in Movies 2a and b. Specifically, the lower region appears to be in front of the upper region in Movie 2a, even when the lower region cue and the velocity cue conflict with each other. However, the authors believe that this conflict reduced the frequency of perceiving the lower region to be 'in front' in Movie 2a, compared with Movie 2b in our results. When the movie
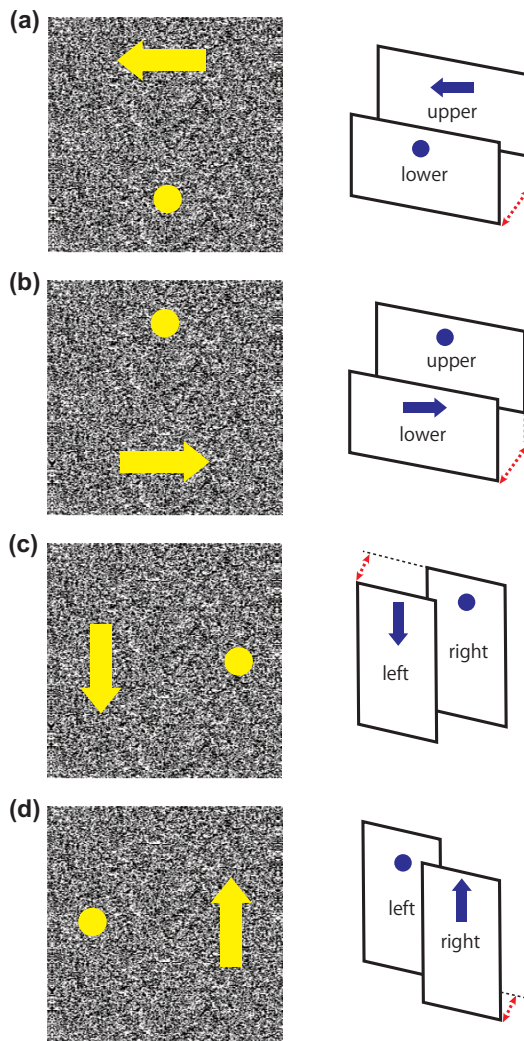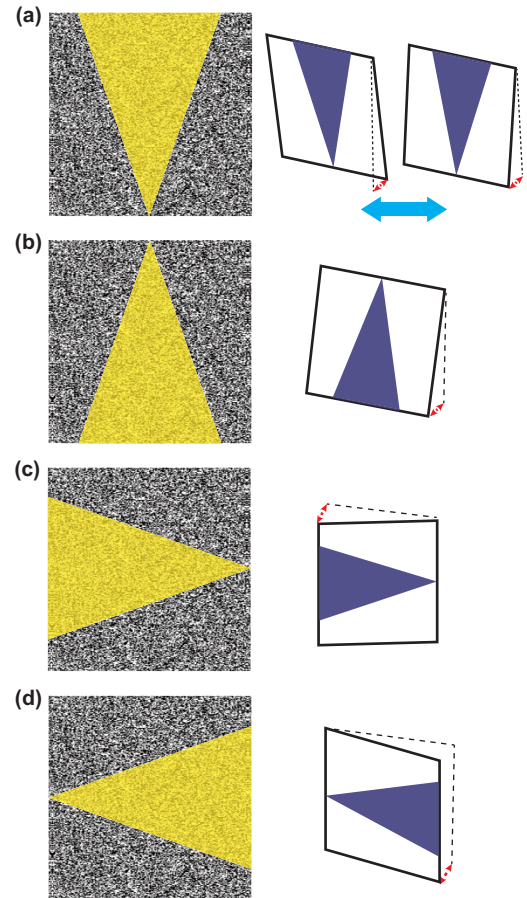
Movie 2. Stimulus movies with one region static from Movie 1. Each movie is rotated at (a) 0°, (b) 180°, (c) 90°, (d) 270° from Movie 2a. A schematic diagram of depth perception is shown on the right of each movie.

Movie 3. Continuous variation of Movie 2. The motion amplitude at the top and bottom of the movie is the same as that in Movie 2 and the motion between them alters continuously. The movie is rotated at (a) 0°, (b) 180°, (c) 90°, (d) 270°. A schematic diagram of depth perception is shown on the right of each movie.

is tilted at 90°, (Movies 2c and d), the lower region cue disappears; hence the dynamic region should appear in front more often. However, the rate of the dynamic region perceived as being 'in front' was close to the chance rate (50%), suggesting that the effect of the velocity cue was not strong in these demonstrations.

Discrete versus smooth

Next, let us explore whether the presence of motion continuity preserves the illusion. Movie 3 presents a continuous variation of Movie 2. The i-th pixel from the bottom of Movie 3a with N-pixel square ($N = 256$) moves at a rate of $i = 0.1 \times \dfrac{i}{N}$ to the image width as the oscillation amplitude. Specifically, the binary noise moves at 10% of

the image width at the very top of Movie 3a and static at the very bottom. Furthermore, the amplitude of its motion decays continuously and linearly from top to bottom.

According to the informal study, in Movies 3a–d, the majority of the observers reported that the region with faster motion appeared to be in front (Movie 3a: 75.0%, 3b: 86.1%, 3c: 80.6%, 3d: 77.8%). Accordingly, in the movie with continuous motion, the velocity cue is more dominant in determining the depth order than the lower region cue. Notably, this contradicts the results obtained with discrete motion (Movie 2).

Although this discrepancy cannot be explained conclusively, a previous study that investigated depth perception using stimulus configuration similar to ours reported that the perceived depth order became less stable when discontinuity of the velocity existed (Kitazaki & Shimojo, 1998). However, the given study did not report the lower region

bias found in our study, possibly because of the difference in the experimental protocol such as the presence of self-motion in their study.

In addition, the depth order was relatively unstable when the upper side moved faster (Movie 3a). According to our observation, the upper part of the movie was often perceived as being in front. However, at some points, the lower part appeared to be in front, and at other times, the movie appeared to be almost flat. This ambiguity in perception may be caused by the conflict between the lower region cue and the velocity cue.
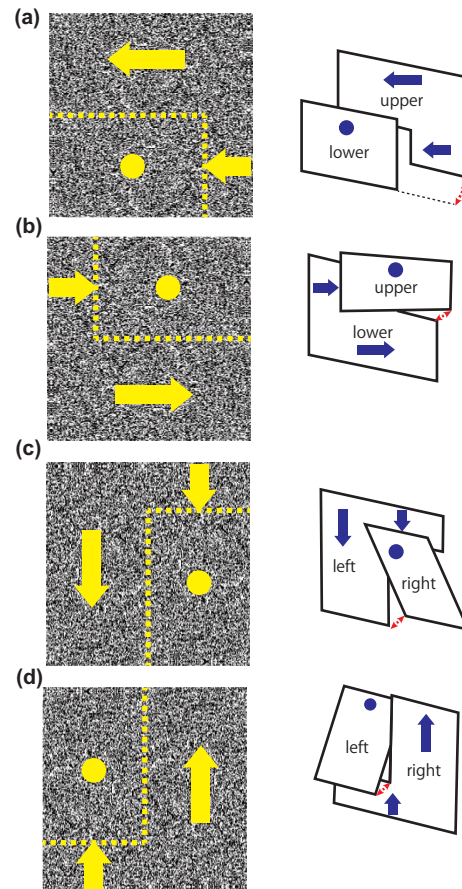
Presence of an occlusion cue

A previous study reported that a region with faster accretion/deletion was more likely to be perceived as being behind another region (Kaplan, 1969). This dynamic occlusion cue with accretion/deletion is known as a monocular depth cue. In this section, we examine whether the occlusion cue or the lower region cue is dominant in depth identification.

Movie 4 shows stimulus movies in the presence of an occlusion cue with a static region. The movies are created by changing 10% width of the static region in Movie 2 to a dynamic region that moves with the same phase and amplitude as the residual dynamic region.

In the informal experiment, observers more often (63.9%) perceived the occluding region as being in front, even when the occlusion cue and the lower region cue conflicted with each other (Movie 4b). However, the probability that the occluding region appeared in front was slightly lower than the other conditions (Movie 4a: 83.3%, 4c: 83.3%, 4d: 72.2%). These results suggest that the occlusion cue is more dominant than the lower region cue.

The demonstrations in Movies 4a–d also contain the velocity cue. The occlusion cue conflicts with the velocity cue and the lower region cue in Movie 4b and conflicts with the velocity cue in all the demonstrations in Movie 4. This may explain why the occluding region is not always perceived as being in front.

Regarding the qualitative appearance, it is observed that the entire area of the lower occluding region is perceived as being in front of the dynamic region in Movie 4a. However, in Movie 4b, the depth order appears to be reversed from the left to the right side of the occluding region such that the static region is perceived as being in front near the boundary with the dynamic occlusion. Simultaneously, the dynamic region appears to be in front away from the boundary. Qualitatively, similar depth perception can be observed in Movies 4c and d. This may be the result of the interpretation of the visual system, such that the local depth cues do not conflict with each other as much as possible.

Movie 4. Stimulus movies in the presence of an occlusion cue with a static region. The stimuli are created by replacing 10% of static region in Movie 2 into the dynamic region. The movie is rotated at (a) 0°, (b) 180°, (c) 90°, and (d) 270° from Movie 4a. A schematic diagram of depth perception is shown on the right of each movie.

Conclusion

In this study, we first demonstrated that the horizontal motion mimicking motion parallax can enhance the lower region cue as a depth cue such that the lower half region of the movie is perceived almost unambiguously as being in front of the upper region. In a series of demonstrations, we showed that the lower region cue dominated the velocity cue in the stimuli with motion discontinuity. However, the relationship could be reversed in the stimulus with a continuous velocity change from the upper to lower region. Furthermore, in the movies with occlusion cue, the occluding region appeared as being in front, suggesting that the occlusion cue is more dominant than the lower region cue. Although these results help to estimate the strength of the lower region cue with respect to the other depth cues, there is no guarantee that the same pattern would hold for other stimulus configurations with different velocity, texture patterns, shapes, and visual angle. We leave this issue for future study.

Table 1. Summary of results obtained in the informal observation

Movie	Lower region appears in front	Right region appears in front
Movie 1a	97.2% (35/36)	-
Movie 1b	-	80.6% (29/36)
Movie 2a	77.8% (28/36)	-
Movie 2b	94.4% (34/36)	-
Movie 2c	-	44.4% (16/36)
Movie 2d	-	58.3% (21/36)
Movie 3a	25.0% (9/36)	-
Movie 3b	86.1% (31/36)	-
Movie 3c	-	19.4% (7/36)
Movie 3d	-	77.8% (28/36)
Movie 4a	83.3% (30/36)	-
Movie 4b	36.1% (13/36)	-
Movie 4c	-	83.3% (30/36)
Movie 4d	-	27.8% (10/36)

Appendix

The informal observations were collected from 18 participants aged between 19 and 59 years with normal or corrected-to-normal vision (the authors did not take part as participants). The 14 movies presented in this study, except for Movie 1c, were tested. Each stimulus was presented twice in a randomized order. The movies were displayed as a 30 Hz image sequence on a web browser. A credit card was used as a reference of known size to adjust the image scale such that the stimulus size was 5 cm square at each observer's screen. The participants observed the movies from approximately 60 cm apart and chose whether the upper or lower (left or right) side of each movie appeared in front. The participants answered their choice by means of a horizontally arranged two buttons corresponding to the upper and lower (left and right) region. There was no time limit and the participants could observe the stimuli for as long as they needed.

The summary of the results is shown in Table 1. The columns of Table 1 show the percentage in which the lower (right) side for each movie appeared as being in front. One condition (i.e. Movie 1b) was not found to be consistent with the observation of the authors. In Movie 1b, the depth order was expected to be the most ambiguous because there was no depth cue that could provide a biased response. However, in 29 of 36 trials, the participants answered that the right side appeared to be in front. Currently, there is no definitive answer to explain this bias, although it may be accounted for to some extent by the hysteresis effect (Bonaiuto et al., 2016). Specifically, it can be assumed that the participants pressed the same response button as in the previous trial because the decisions were particularly challenging in

this condition[1]. Furthermore, the result might be related to the visual system's left-right asymmetry in the depth interpretation (Sun & Perona, 1998). This issue will be clarified in future study.

Conflict of interest and funding

References

Bonaiuto, J. J., de Berker, A., & Bestmann, S. (2016). Response repetition biases in human perceptual decisions are explained by activity decay in competitive attractor models. *Elife*, *5*, e20047. doi: 10.7554/eLife.20047.025

Braunstein, M. L., & Andersen, G. J. (1981). Velocity gradients and relative depth perception. *Perception & Psychophysics*, *29*(2), 145–155. doi: 10.3758/BF03207278

Grossberg, S. (1997). Cortical dynamics of three-dimensional figure–ground perception of two- dimensional pictures. *Psychological Review*, *104*(3), 618–658. doi: 10.1037/0033-295X.104.3.618

Hulleman, J., & Humphreys, G. W. (2004). A new cue to figure–ground coding: Top–bottom polarity. *Vision Research*, *44*(24), 2779–2791. doi: 10.1016/j.visres.2004.06.012

Kaneko, H., & Uchikawa, K. (1993). Apparent relative size and depth of moving objects. *Perception*, *22*(5), 537–547. doi: 10.1068/p220537

Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Perception & Psychophysics*, *6*(4), 193–198. doi: 10.3758/BF03207015

Kitazaki, M., & Shimojo, S. (1998). Surface discontinuity is critical in a moving observer's perception of objects depth order and relative motion from retinal image motion. *Perception*, *27*(10), 1153–1176. doi: 10.1068/p271153

Necker, L. A. (1832). Lxi. observations on some remarkable optical phenomena seen in Switzerland; and on an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *1*(5), 329–337. doi: 10.1080/14786443208647909

Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience*, *1*(3), 183–184. doi: 10.1038/630

Vecera, S. P., & Palmer, S. E. (2006). Grounding the figure: Surface attachment influences figure-ground organization. *Psychonomic Bulletin & Review*, *13*(4), 563–569. doi: 10.3758/BF03193963

Vecera, S. P., P., Vogel, E. K., & Woodman, G. F. (2002). Lower region: A new cue for figure-ground assignment. *Journal of Experimental Psychology: General, 131*(2), 194–205. doi: 10.1037/0096-3445.131.2.194

---

[1]Out of 36 participants for Movie 1b, 23 clicked on the same button as in the previous trial. In particular, of these 23 participants, only two responses corresponded to the left-hand button, whereas the remaining 21 responses corresponded to the right-hand button. The bias toward the right button could be due to the lower region bias. Of the 21 trials, the stimuli in the preceding 16 trials were divided into the upper and lower regions and the participants answered 'lower region' by pressing the right-hand button.