Reviews (Round 1): Motion-driven enhancement of a lower region cue in depth perception (#8028)

Edited by: Kohske Takahashi, Ritsumeikan University, Japan

Review by: Satoshi Shioiri, Tohoku University, Japan

For author and editor

This study reports a phenomenon of lower region being perceived closer in depth in motion defined figures as a novel visual illusion. Although perception of lower region appears being closer is a known phenomenon, no previous report used motion defined figures. By the reason, authors claim that their display is new. The demonstrations are nice and manuscript is well written. However, it is a critical to ask whether the phenomenon reported is novel. My understanding of the phenomenon is that the well-known bias of the lower region seen closer is found for the shape defined by motion differences. While I wouldn't claim that this is new if I were an author because the difference from known phenomenon is how the shape is defined and I see no reason there may be no effect. But, this is my personal opinion. While objective estimation of originality is not easy in this case, and I agree that the stimuli are new.

Most of the phenomena found in the demonstration movies can be explained basically by known factors as depth cues: lower region effect, motion effect, and dynamic occlusion. One exception may be the change in surface perception between Figs 3 (a) and (b): perceiving concave and convex surfaces. Although adding lower region effect to the flat oriented surface could explain the surface orientation, perhaps nonlinear effect is required to explain the change between concave and convex surfaces. This is not the effect for figure–ground organization but the effect of depth interactions and may not be the effect what the author want to report. The perception of Figs 4 (a) and (b) may also indicate nonlinear effect of depth and/or depth interactions, which may also be novel although more detailed description is necessary.

The author may want to make the novel points clearer, rather than stating that the lower region effect in motion figure is new.

Other points.

Title: using motion figures is a unique point of the study. The title should indicate the point.

Movie 1: region. The lower region is perceived stably in the foreground even when the eye moves over this movie.

I have tried to see if there are conditions, where the upper region can be seen in front. To me, tracking the upper area biases to the direction of the upper to be closer in general but not see as in front in general. However, I found that the upper region appears to be in front right after the start of the fixation at the area occasionally. Also, I found perception changes depends on the movie size or speed of dot motion. The readers may want to know the meaning of "stably" clearer, limiting the conditions should be satisfied.

1.2 2nd paragraph: Furthermore, ,,,

It is not clear the condition. I think the authors mean that observing from above by leaning forward.

1.2.2 Dynamic vs. Static, 1st paragraph: The qualitative tendency of depth perception does not change in relation to where the observer looks or spatial scale of the movie.

I think the effect becomes weaker with moving field at the upper region due to the bias of moving object to be seen closer or as a figure. The location to fixate altered the depth relationship to me.

Movie 3: The motion between them alters continuously

Any reason to use the cosine function? I am not sure if I agree with the description.

and it has a smaller depth gap between the top and bottom than Movie 2(a), which have a discrete motion profile.

1.4 line 7: the unevenness of the perceived surface is also reversed.What does "unevenness" mean?

1.4 line 8: surface appear to vibrate

Not clear how the surface vibrate. Is the surface changing depth?

Recommendation: Resubmit for review

Completed: 2021-08-12

Review by: Yuki Kobayashi, Ritsumeikan University, Japan

For author and editor

The authors report that a lower region of a movie tends to be perceived as the foreground even when the boundary between upper and lower regions is defined by motion. The authors further examined whether this effect is still present in different conditions in the following movies. I agree that the first demonstration (Movie 1a and 1b) is so convincing to show the existence of the effect. However, I have three major comments about the interpretation, citations, and the observations for Movie 2-4. Some minor comments follow the major ones.

Major comment 1 - Interpretation of the main finding

The authors argue that their demonstration shows a novel phenomenon, but I am not completely convinced of this claim. It is true that the phenomenon shown in this manuscript is driven by the two different motion patterns while Vecera et al. (2002) used a static image with a boundary defined by luminance difference. However, the difference of how the central boundary is defined (by motion or by luminance) does not seem to be so critical to let the authors' demonstration be "a novel phenomenon" and different from Vecera et al.'s finding that a lower region tends to appear as figure and closer. I considered the present demonstration as a new expression form of a known phenomenon. I would appreciate clearer explanation regarding why the difference of how the central boundary is defined is such a large leap from Vecera et al. that the present demonstration is essentially different from it.

However, it is notable that a motion-defined boundary seems to enhance the effect of the lower region cue. This effect is absent or very weak with a straight horizontal boundary in a static image (for example, the white half of Indonesia's national flag does not appear in front, at least to me), while Vecera et al.'s stimuli that were halved by a sawtooth or wave-like edge showed the strong lower region preference. The difference between their stimuli and Indonesia's flag suggests that the straightness of the central boundary weakens the lower region cue in static images. The present demonstration shows that the lower region does serve as a depth cue even with a straight horizontal boundary when it is defined by motion. I think this finding is novel. If the authors agree with my opinion, I recommend that the authors put focus on the motion-driven enhancement of the lower region cue, rather than argue a novelty of the phenomenon per se. Journal of Illusion accepts "Demonstration, Design, Artwork or Application cases of known illusion". I consider this submission meets this scope.

Major comment 2 - Relevant previous studies

I recommend that the authors cite some relevant studies. Some studies have showed that faster objects, without observers' head movement, tend to appear closer to observers especially when they have velocity gradient (Braunstein & Andersen, 1981; Kaneko & Uchikawa, 1993). These studies are closely related to the discussion for Movie 3, where regions with high velocity tend to appear in front. Moreover, Kaneko and Uchikawa (1993) suggested the lower region preference before Vecera et al. (2002) (though Vecera did not cite this study).

Kaplan (1969) also seems to be relevant to the discussion for Movie 4. The effect of accretion/deletion defined by texture motion on depth perception is discussed in this paper. Citing these previous studies will enrich the discussions.

Braunstein, M. L., & Andersen, G. J. (1981). Velocity gradients and relative depth perception. Perception & Psychophysics, 29(2), 145-155.

Kaneko, H., & Uchikawa, K. (1993). Apparent relative size and depth of moving objects. Perception, 22(5), 537-547.

Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. Perception & Psychophysics, 6(4), 193-198.

Major comment 3 - The observations for Movie 2-4

Although the first demonstration (Movie 1a and 1b) was convincing, the authors' observations for the following movies (Movie 2-4) were not necessarily agreeable for me. I might be an atypical observer, but assertive expressions do not seem to be appropriate for some observations which will not necessarily obtain wide agreement. It may be also a possible choice to collect data to justify the authors' observations.

- Movie 2

The authors wrote "The depth gap is larger when the stationary region is located in the lower than in the upper." (p.2) I do not completely agree with this observation (I do not see the clear difference of the apparent depth gap between in Movie 2a and in Movie 2b) although I agree that the lower regions consistently appear in front. Moreover, this observation seems contradictory to the observation for Movie 2c and 2d, in which the authors argue that the dynamic region appears to be in front. Considering the observations for Movie 2c and 2d, it is expectable that the gap of apparent depth would be larger in Movie 2b than in 2a, because both the effects of a lower region and motion work in the same region of Movie 2b. More detailed discussion on the relationship between these two

effects would help readers' better understanding.

- Movie 3 and 4

The observations for Movie 3 and 4 were not acceptable to me, either. For Movie 3a, the authors claim that the lower part appears in front, but this appearance is not stable for me (sometimes it appears flat) although I perceive the depth in Movie 3b, 3c, and 3d as the authors claim. As mentioned above, previous studies explain the perceived depth in Movie 3b, 3c, and 3d. I suppose that the depth cues by the velocity gradient and the lower region are conflicting in Movie 3a. The fact that the dynamic region does not always appear in front in Movie 3a while it does in Movie 3b, 3c, and 3d sufficiently shows the effect of the lower region cue. It would be safer to avoid asserting that Movie 3a's lower region appear in front. Moreover, the appearances of Movies 4b, 4c, and 4d are not so stable to me, so I also recommend concessive expressions for the discussions for them.

Minor comments

I am not an English native so my sense is not reliable, but some expressions in the manuscript did not sound natural. Please consider modifying the sentences if you agree with my following comments.

- p.1 left column "…tends to be recognized as a figure more easily": "is more likely to be recognized as a figure" might be more natural.

- p.2 left column "in the right": "on the right" seems more natural.

- p.2 right column "... regions affect the illusion.": "affects" should be grammatically correct.

- p.2 right column Section 2: Here, "static" and "stationary" are mixed. If the authors intentionally used these two similar words for different meanings, please clarify their difference. If not, it would improve readability to adopt either of them.

- p.4 left column: Here, "unevenness" seems to mean the apparent curve of the image surface, but the first "unevenness" (p4. line 7) appeared to mean the apparent depth gap between the top and bottom regions. I recommend that the authors choose a different word here.

- p.5 left column "However, Movie 4(b),…": "in" should be inserted before "Movie 4(b)".

- p.5 left column "a 90° and 270° rotation of": "90° and 270° rotations of" may be correct.

- I found some "rotated 90° " and "tilted 90° ", but "rotated by 90° " and "tilted by 90° " should be correct.

I have some more minor comments. Please consider modifying these points if you agree with me.

- p.2 right column: I guess that "Movie 3(a) with N-pixel square" means a movie whose one side is N-pixel long. If I understanding this part correctly, floor function did not seem to be necessary for i/N. I would appreciate an explanation for this formula.

- p.4 right column Movie 4 caption: “(b) 90° , (c) 180° ” should be “(b) 180° , (c) 90° ”.

- p.5 left column “In addition, by hiding the dynamic region providing an occlusion cue, the lower region becomes in front as in the original movie.”: Here, does “the original movie” mean Movie 1a? I am afraid I did not understand why Movie 4c and 4d can be perceived as Movie 1a by hiding the dynamic and occlusion-providing region. I would appreciate if the authors could make the meaning of this sentence clearer.

- p.5 References: The last pages’ numbers of Grossberg (1997) and Vecera et al. (2002) are missing.

- For each rotation angle, it should be clarified that the angles are defined by counterclockwise rotation

Recommendation: Revisions required

Completed: 2021-07-22
