Motion from occlusion

Stephen A. Engel Department of Psychology, UCLA, Los Angeles, CA, USA

David A. Remus Department of Psychology, Stanford University, Stanford, CA, USA

Rajiv Sainath Institute of Neurological Sciences, University of Pennsylvania, Philadelphia, PA, USA

Vision uses specific image features or cues to infer physical properties of the world. Here, we use a novel illusion to show that occlusion, traditionally thought of as a cue to depth, is also a powerful cue to motion. A display of stacking disks that contains only occlusion as a cue to depth generates a vivid sense of movement that is likely computed in early or middle levels of visual processing.

Keywords: depth perception, illusory motion, illusion, third-order motion

Introduction

When one object blocks another from view, as might two papers lying upon a desk, observers immediately infer which object is on top. Such occlusion relations are well-known cues to depth in static images (Ratoosh, 1949), but in dynamic displays, they have historically been considered as a source of error or ambiguity (Wallach, 1953). When occlusion is not properly inferred in displays of contour-defined motion, erroneous or ambiguous perception can result, but when occlusion is properly inferred, it disambiguates a variety of motion displays (e.g., McDermott & Adelson, 2004; Shimojo, Silverman, & Nakayama, 1989).

No study to date, however, has investigated whether occlusion, by itself, can produce a percept of motion. Below, we describe an illusion that strongly suggests that it can.

Methods

Participants were recruited from within the UCLA community. Each gave informed consent, as approved by the UCLA Office for the Protection of Research Subjects.

Our basic display presented identical disks drawn in rapid succession, where new disks occluded older ones (Movie 1; animations can also be seen on the web at http://engellab.psych.ucla.edu/DiskDemos.html). This display generates a vivid impression of motion as the disks stack and appear to move toward the viewer. The display contains no previously identified cue to motion in depth.

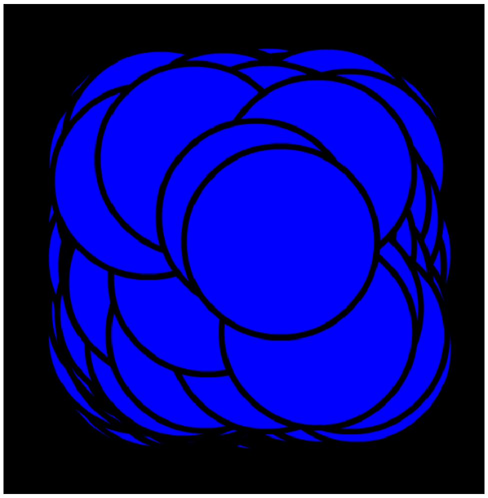
To measure this percept, we had five participants compare the motion of stacking disks to that of a single disk whose motion was defined by a smooth change in diameter and a moving cast shadow (Figure 1). Participants monocularly viewed both displays side by side, fixating on a mark between them, and indicated whether the stacking disks or the single disk was "moving toward you more quickly." The rate at which stacking disks appeared was varied across trials (2.4, 4, or 5 disks/s), as was the single disk's simulated velocity (0.067, 0.089, 0.111, 0.133, or 0.156 radii/s; 300 total trials per participant). The initial background for stacking disks was identical to the background of the single-disk display.

Participants were seated in a dimly lit room and their heads were stabilized using a chin rest. Each stacking disk subtended 7.9 deg of visual angle, as did the single moving disk at the beginning of its motion. The entire display subtended approximately 60 deg. Disks were blue with a black outline on a black background.

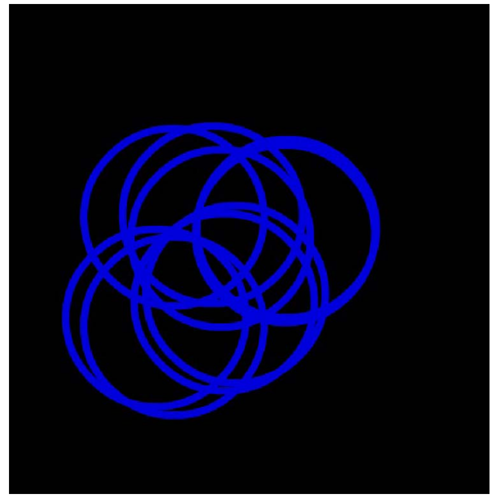
In a second experiment, we tested whether adaptation to the stacking disks affected the apparent motion of a display of rings that appeared in rapid succession (Movie 2). The ring's transparent centers provided no occlusion cues, causing them to generate an inconsistent or ambiguous sense of motion when display parameters were otherwise identical to the stacking disk display. We expected perception of this neutral display to be altered by adaptation.

Six participants viewed 1-s displays of these rings (appearing at a rate of 10 rings/s; initial rings subtended 13 deg of visual angle) and judged whether the rings generally appeared to be moving toward or away from them. Participants performed this task both before and after 3 min of viewing the stacking disks (disks stacked at a rate of 10/s; each disk subtended 13 deg of visual angle). In the latter case, ring displays were interleaved with 15-s "top–up" displays of the stacking disks that were intended to keep participants in a consistent state of adaptation. Each successive ring could be either smaller or larger than the previous one to vary motion bias in the ring displays. The magnitude

doi: 10.1167/6.5.9 Received December 20, 2005; published May 10, 2006 ISSN 1534-7362 © ARVO



Movie 1. Basic motion from occlusion display. Blue disks are drawn on a computer screen, with new disks partially blocking older ones from view. The disks appear to stack, and the stack appears to move toward the viewer. Click on the image to view the movie.



Movie 2. Experiment 2: rings display. Blue rings are drawn without cues to occlusion; motion is ambiguous. Click on the image to view the movie.

and direction of this size change were varied across trials (0-, 1-, or 3-pixel increments per frame; 200 total trials per participant).

Results

Responses from our first task showed that the stacking disks generated a consistent motion percept (Figure 2). As the stacking disks' presentation rate increased, participants reported more often that the stacking disks moved in depth more quickly than the single disk, $F(2,32) = 70.5$, $p < .001$ (ANOVA, main effect of stacking speed). Likewise, as the single disk's simulated velocity increased, participants reported more often that the single disk moved more quickly, $F(4,32) = 25.9$, $p < .001$.

Responses from our second task showed a strong motion aftereffect (Figure 3). After viewing the stacking disks, the rings appeared to move away from observers with reliably greater frequency, $F(1,20) = 7.0$, $p < .05$ (ANOVA, main effect of adaptation), than they did prior to viewing the stacking disks. Trends in this direction were seen at every level of the ring size change.

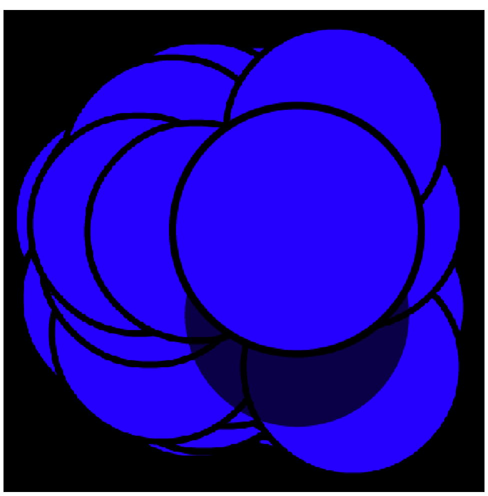


Figure 1. Experiment 1: comparison display. A single disk's size, edge locations, and shadows changed over time, providing several cues to motion in depth.

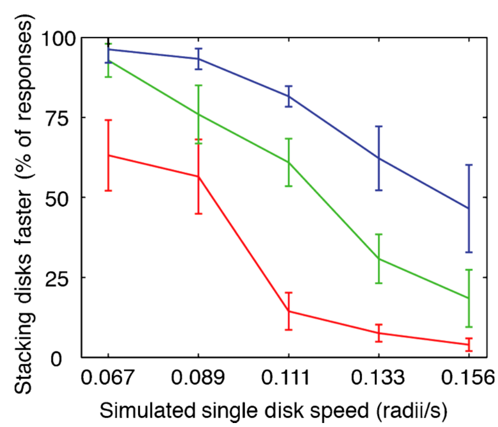


Figure 2. Results of Experiment 1. Participants compared the speed of motion of the stacking disks with the single moving disk. The disks were stacked at three different rates: 2.4 (red), 4 (green), or 5 (blue) disks/s. Error bars: ±1 between-subjects standard error.

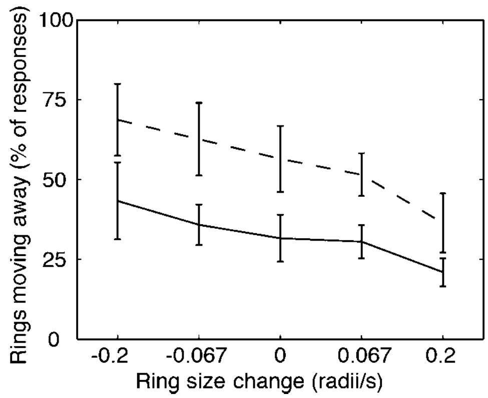
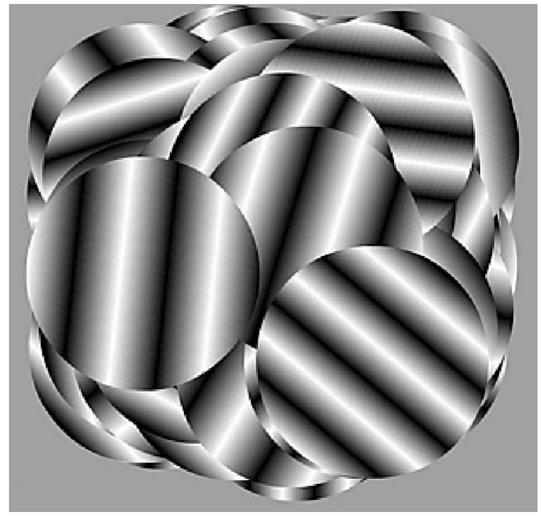


Figure 3. Results of Experiment 2. Participants viewed rings before and after 3 min of exposure to stacking disks (before, solid line; after, dotted line). To measure the magnitude of the motion aftereffect, we varied whether successive rings became larger or smaller, which biased them to appear to move toward or away from the observer. Viewing the stacking disks increased how frequently the rings appeared to move away from the viewer.



Movie 3. Stacking illusion with disks defined by contrast rather than luminance. Click on the image to view the movie.

Discussion

Experiment 1 suggests that information from occlusion alone can support motion perception. Because visual aftereffects are generally believed to arise relatively early in visual processing, Experiment 2 reveals that occlusion enters the motion computation at an early or perhaps midlevel stage.

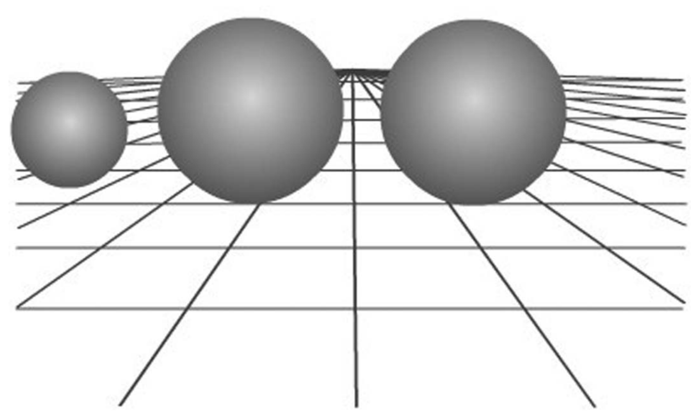
Can existing models of motion perception explain the stacking disks' illusion? While our basic stimulus was intended to contain no traditional cues to motion, the appearance of each disk is a change in image structure over time that can serve as input to computational models of motion perception. These models could possibly infer motion in depth from our displays; for example, the appearance of each disk could be interpreted as an instantaneous expansion from a point, which is consistent with rapid movement toward an object. Indeed, isolated disks appearing on a uniform background can generate percepts of expansion (Bartely, 1941).

The simplest, "first-order" models of motion perception use changes in image luminance as a cue. Movie 3 shows a version of the illusion where the disks are defined by spatially oriented contrast patterns rather than simple luminance and, hence, cannot be explained by such models.

Models of "second-order" motion use changes in quantities such as orientation or contrast as a cue to motion and, thus, could possibly account for the movement perceived in Movie 3. Although it seems unlikely, we cannot, in principle, rule out the possibility that these or more complex models could infer motion in depth in our displays without using information provided by occlusion.

Movie 4 further suggests that traditional models may be unable to explain motion in depth from occlusion. In this movie, occlusion strongly affects the interpretation of motion signaled by a traditional cue, local motion energy. The moving ball's horizontal motion is identical in both sequences, but whether it occludes or is occluded by the other balls greatly affects its perceived trajectory (the demonstration is closely related to one in Ringach, 1996). This demonstration also shows the effectiveness of occlusion cues to motion in slightly more natural displays than the stacking disks.

The neural mechanisms underlying motion from occlusion remain unknown. One possibility is that visual attention tracks the stack's depth and that attention's location feeds into the same mechanisms that process traditional motion cues (Lu & Sperling, 2001). Alternatively, representations of depth from occlusion, perhaps computed preattentively (Nakayama & Shimojo, 1992; Rensink & Enns, 1998),



Movie 4. Occlusion and translation. The moving ball's trajectory is greatly influenced by whether it occludes or is occluded by the other balls. Click on the image to view the movie.

could feed more directly into motion processing. Future work will distinguish between these and other possible mechanisms.

Observers can also see motion when stereo disparity (Patterson, 1999), shadows (Kersten, Knill, Mamassian, & Bulthoff, 1996), and spatial scale (Schrater, Knill, & Simoncelli, 2001) change over time. Our results suggest a common hypothesis for interpreting these effects: Each of the numerous static cues to depth may support a robust percept of motion.

Conclusions

Occlusion information alone can produce a strong percept of motion in depth. The visual system can integrate a large number of relatively complex cues to infer visual motion in the world.

Acknowledgments

We thank Dario Ringach, Zili Liu, and Phil Kellman for helpful discussion. This work was supported by NIH EY11862.

Commercial relationships: none.
Corresponding author: Stephen A. Engel.
Email: engel@psych.ucla.edu.
Address: UCLA Psychology, Franz Hall 1282a, Los Angeles, CA 90095-1563, USA.

References

Bartley, S. H. (1941). *Vision, a study of its basis*. New York: Van Nostrand Reinhold.

Kersten, D., Knill, D. C., Mamassian, P., & Bulthoff, I. (1996). Illusory motion from shadows. *Nature, 379,* 31. [PubMed]

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 18,* 2331–2370. [PubMed]

McDermott, J., & Adelson, E. H. (2004). The geometry of the occluding contour and its effect on motion interpretation. *Journal of Vision, 4*(10), 944–954, http://journalofvision.org/4/10/9/, doi:10.1167/4.10.9. [PubMed] [Article]

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science, 257,* 1357–1363. [PubMed]

Patterson, R. (1999). Stereoscopic (cyclopean) motion sensing. *Vision Research, 39,* 3329–3345. [PubMed]

Ratoosh, P. (1949). On interposition as a cue for the perception of distance. *Proceedings of the National Academy of Sciences of the United States of America, 35,* 257–259. [PubMed] [Article]

Rensink, R. A., & Enns, J. T. (1998). Early completion of occluded objects. *Vision Research, 38,* 2489–2505. [PubMed]

Ringach, D. L. (1996). *Binocular eye movements caused by the perception of three-dimensional structure from motion*. Unpublished doctoral dissertation, New York University.

Schrater, P. R., Knill, D. C., & Simoncelli, E. P. (2001). Perceiving visual expansion without optic flow. *Nature, 410,* 816–819. [PubMed]

Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research, 29,* 619–626. [PubMed]

Wallach, H. (1953). Ueber visuell wahrgenommene bewegungrichtung. *Psychologie Forschung, 20,* 325–380.