

Second-order motion conveys depth-order information

Jay Hegdé

The Salk Institute for Biological Studies, La Jolla, CA, USA



Thomas D. Albright

The Salk Institute for Biological Studies, La Jolla, CA, USA



Gene R. Stoner

The Salk Institute for Biological Studies, La Jolla, CA, USA



Psychophysical and neurophysiological studies have revealed that the visual system is sensitive to both “first-order” motion, in which moving features are defined by luminance cues, and “second-order” motion, in which motion is defined by nonluminance cues, such as contrast or flicker. Here we show psychophysically that common types of second-order stimuli provide potent cues to depth order. Although motion defined exclusively by nonluminance cues may be relatively rare in natural scenes, the depth-order cues offered by second-order stimuli arise ubiquitously as a result of occlusion of one moving object by another. Our results thus shed new light on the ecological importance of second-order motion. Furthermore, our results imply that visual cortical areas that have been shown to be responsive to second-order motion may be extracting information not just about object motion as has been assumed, but also about the relative depth of objects.

Keywords: depth, 3D surface and shape perception, motion-2D

Introduction

Imagine a window shade being drawn shut. As the shade is drawn, it progressively occludes scene elements outside the window. Conversely, when the shade is pulled open, those same scene elements are progressively disoccluded. This dynamic occlusion/dissocclusion has long been recognized as a potential cue for depth order and has been termed accretion-deletion (AD) (Kaplan, 1969; Gibson, Kaplan, Reynolds, & Wheeler, 1969; Thompson, Mutch, & Berzins, 1985).

More recently, it has been recognized that depth order can also be inferred by identifying a moving boundary, and then contrasting the velocity of the boundary with the velocity of features adjacent to the boundary (Yonas, Craton, & Thompson, 1987; Mutch & Thompson, 1988; Craton & Yonas, 1990). Suppose, for example, that the window shade in our example were textured. Under such conditions, the shade can be identified as the occluding foreground by virtue of the fact that its textured surface moves together with its edge (i.e., the surface and the edge are in common motion [CM]). The background side of the boundary, conversely, can be identified by the differential motion (DM) between it and the moving edge.

Stimuli used in previous investigations of dynamic depth-ordering cues (Kaplan, 1969; Gibson et al., 1969; Thompson, et al., 1985; Yonas et al., 1987; Mutch & Thompson, 1988; Craton & Yonas, 1990) included moving luminance-defined features, which activate first-order motion detectors. We observed, however, that this first-order motion component could be removed, thereby creating second-order motion stimuli that offer the AD and DM

depth-ordering cues. We hypothesized that these cues, and the second-order stimuli that provide them, were sufficient to support perceptual depth ordering. Using human subjects, we assayed the depth-order information contained in commonly used types of second-order stimuli. We found that, for all subjects, second-order motion stimuli provided cues for depth order that were at least as effective as stimuli that possessed both first- and second-order motion components.

Methods

Subjects

Thirteen naïve volunteer human subjects (six male and seven female) participated in this study. All had normal or corrected-to-normal vision. All experimental procedures were carried out in accordance with National Institutes of Health and Salk Institute guidelines for the use of human subjects.

Stimuli

In this work, we represent motion stimuli by their space-time plots, which permit explicit identification of depth-order cues (Figure 1). Our stimuli were generated using the Matlab utility (Mathworks Inc., Natick, MA). We constructed six stimuli, each of which consisted of two abutting panels with a moving boundary dividing them (Figure 2). A given panel in each stimulus was either uniform gray, or consisted of flickering or nonflickering random dots. For the random dot panels, dots were either

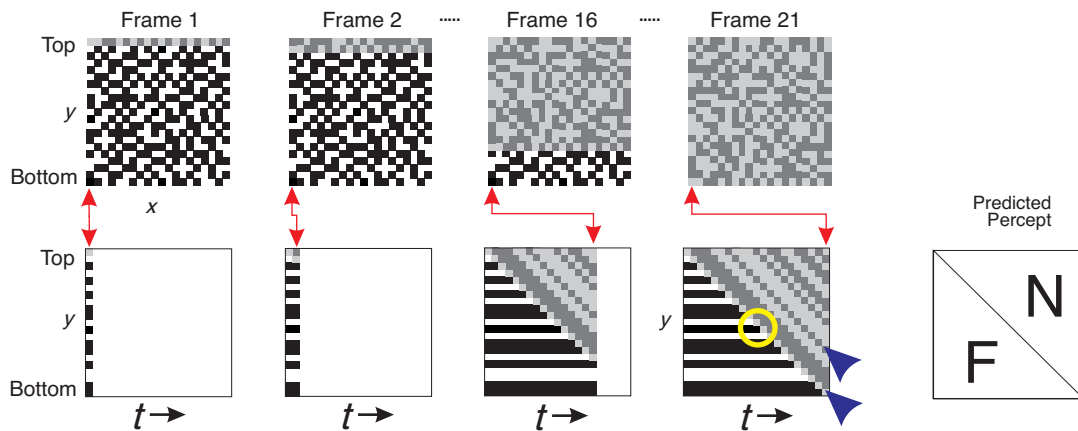


Figure 1. Construction of space-time plot (bottom row) of a motion stimulus (top row). A space-time plot is constructed by concatenating a designated column of pixels from successive frames of the motion stimulus, so that successive columns of the space-time plot represent the temporal progression of a given set of pixels in the movie (click on the figure to view). For the stimulus shown, the top panel (highlighted in gray) slides over the stationary bottom panel in successive frames. The finished space-time plot represents the temporal progression of the far left column of pixels of the movie. The horizontal and oblique lines in this plot correspond to stationary and moving textures, respectively. The AD cue is indicated by the termination of the horizontal lines (highlighted by the circle) at the boundary (i.e., where the two sets of lines meet) between the two panels. The CM cue is reflected by the match between the orientation (corresponding to velocity) of the boundary and that of the highlighted texture (arrowheads). Conversely, the DM cue is represented by the disparate orientation of the motion boundary relative to the lines corresponding to the static texture. All cues predict the same depth ordering: the highlighted panel is near (N) relative to the other “Far” (F) panel.

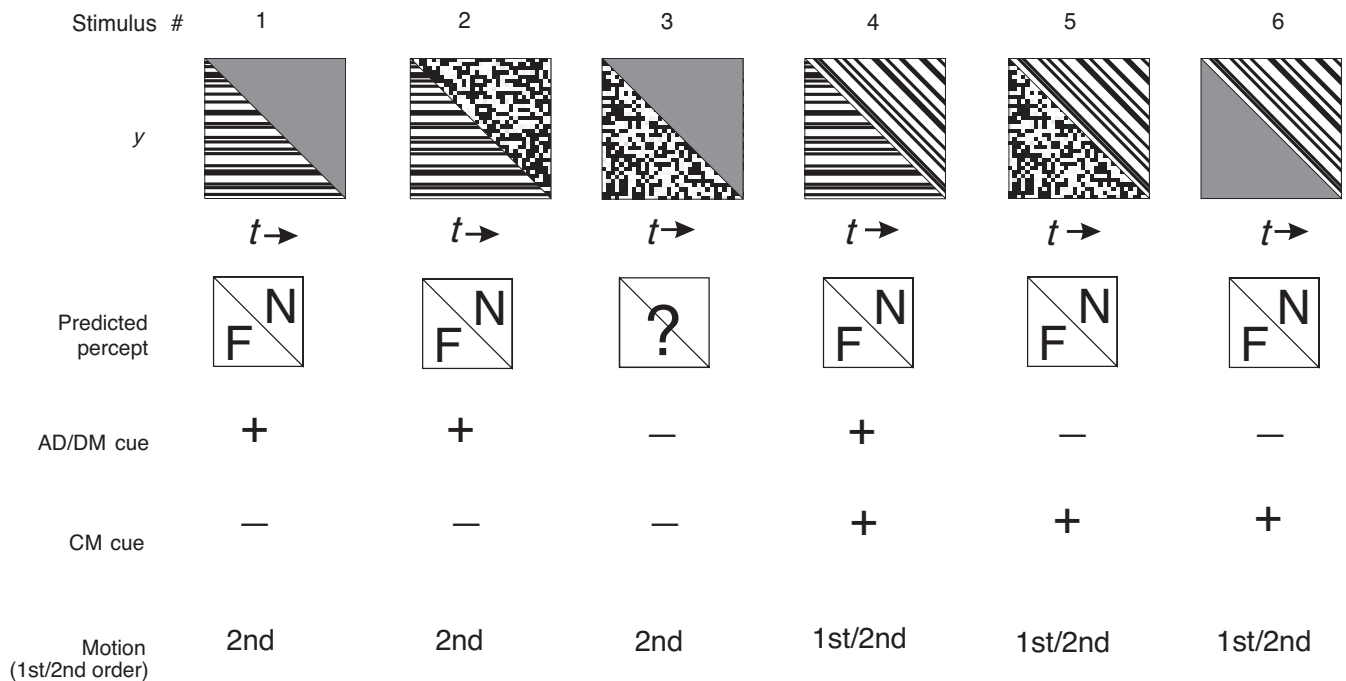


Figure 2. The stimuli used in the experiments, shown in space-time format. For each stimulus, the predicted depth order (N = near; F = far; ? = not defined), and the depth-order and motion cues present in the stimulus are indicated. Click on the stimuli to view.

black or white and subtended one pixel. Dot density was 50%. For flickering panels, each dot was independently redrawn at every refresh, with 50% probability, as either black or white. Each stimulus was 199 x 199 pixels (5.8° x 5.8°), with one panel at the top and the other at the bottom, so that the motion boundary between them was horizontal.

Stimuli were presented on a 19" flat screen Sony Trinitron Multiscan E500 monitor (frame rate, 75 Hz) at an eccentricity of 8°. All panels in all stimuli were equiluminant with each other at 29.4 cd/m² (as measured by the Photo-Research PR650 photometer). Nonetheless, it is possible that there were residual luminance differences between the panels. Note, however, that any such residual luminance differences are not likely to have affected perceived depth because, as we note in the "Discussion," a luminance difference between the two panels of our stimuli would not constitute a depth-ordering cue.

All stimuli were presented against a neutral gray background, which had a luminance of 15.1 cd/m². The horizontal border between the two panels of the stimulus moved smoothly up and down the middle one-third of the stimulus at a rate of 6.6°/s.

The trials were carried out and the data collected using the CORTEX program for experimental control and stimulus presentation (<http://www.cortex.salk.edu>) running on a Pentium-class computer. Each trial began with the onset of the fixation spot, which the subjects were asked to fixate throughout the trial. After establishing fixation, the subject indicated readiness for the trial using a key press. Although the subjects were told that fixation was necessary, the depth-order percepts elicited by our stimuli were not dependent on fixation (unpublished observations). After a 500-ms delay, the moving stimulus was presented for 4 s. After stimulus offset, the subject was required to report which side of the motion boundary appeared as the near surface by pressing an appropriate key. Each stimulus was tested with a given panel at top or at bottom for 20 trials each, and the data were pooled across the two mirror-image configurations. All trials were randomly interleaved.

Results

We tested perceived depth order for six motion stimuli, which are depicted in space-time format in Figure 2. Each stimulus consisted of two abutting panels with a moving boundary dividing them (see "Methods"; for demo, see http://www.vcls.salk.edu/Demos/depth_order_movies). By independently varying the panel types (uniform gray, static texture, coherently moving texture, or dynamic random texture), we created stimuli with the AD and DM cues, the CM cue, both sets of cues, or neither set of cues. Simple rules determined the relationship between panel type and depth cue:

- i. Stimuli 1 and 2 (numbered from left in Figure 2) contained a panel of static texture. This texture was occluded or disoccluded by the moving boundary, thus presenting the AD cue. Because the motion boundary moved relative to the stationary texture, the DM cue was also provided. We refer to these stimuli as AD/DM-cue stimuli.
- ii. Stimuli 5 and 6 contained a panel of coherently moving texture. The motion of this texture was the same as the moving boundary, thereby providing the CM cue. We refer to these as CM-cue stimuli.
- iii. Stimulus 4 possessed both static and coherently moving texture panels, and thus presented all three dynamic depth cues. We refer to this as the AD/DM/CM-cue stimulus.
- iv. Stimulus 3 lacked panels with either static or coherently moving texture and thus presented none of the three depth cues. We refer to this as the no-cue stimulus.

All stimuli containing the CM cue (stimuli 4, 5, and 6) possessed both first- and second-order motion components, whereas the remaining stimuli (stimuli 1, 2, and 3) were purely second-order. The second-order stimuli were simple variants of those used in previous investigations of second-order motion (Chubb & Sperling, 1988; Cavanagh & Mather, 1989; Albright, 1992; Smith, Greenlee, Singh, Kraemer, & Hennig, 1998; Baker, & Mareschal, 2001; Seiffert, Somers, Dale, & Tootell, 2003; Dumoulin, Baker, Hess, & Evans, 2003).

Naïve human subjects ($n = 13$) viewed stimuli and reported which side of the motion boundary appeared as the near surface (two-alternative forced-choice; see "Methods") via a key press. Figure 3 shows the reported depth percepts of two individual subjects. Figure 4 illustrates the responses averaged across all 13 subjects. When presented with the two second-order AD/DM-cue stimuli (stimuli 1 and 2), all subjects reported the depth order consistent with those cues (Binomial Sign Test, $p < .05$), thus revealing that the depth cues provided by second-order motion are sufficient to support perceptual depth order. Importantly, the effectiveness of the first-order CM-cue stimulus (stimuli 5 and 6) was less, on average, than the AD/DM cue combination. In addition, the effectiveness of the second-order AD/DM-cue stimuli was no less than that of the AD/DM/CM-cue stimulus, which contained both first- and second-order motion (Fisher Test of Independence, $p > .05$). Finally, the sole second-order stimulus that offered no depth-order cues (stimulus 3) did not elicit a consistent depth-order percept. Taken together, these results reveal that depth-order cues conveyed exclusively by second-order motion are as effective as those conveyed by stimuli possessing first-order motion components. An unexpected finding was that the effective-

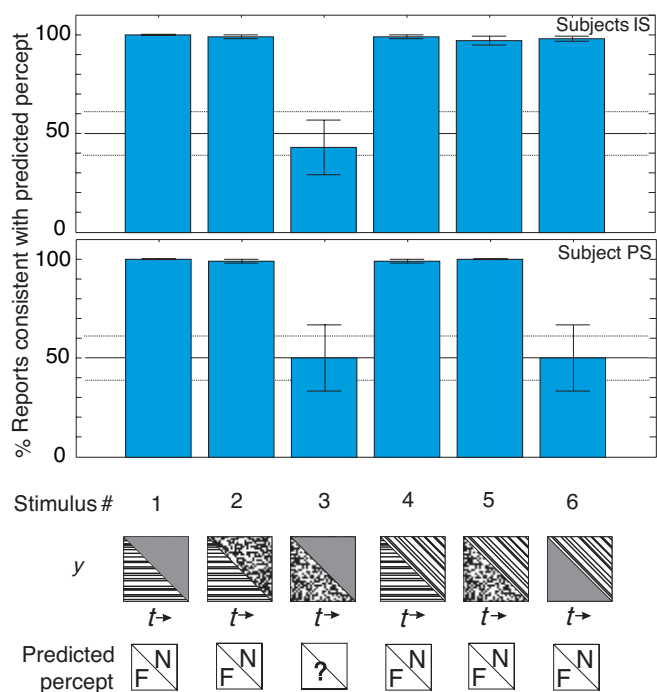


Figure 3. Depth-order percepts elicited by the motion stimuli for two individual subjects. Stimuli were presented in two mirror-image spatial orientations (not shown) and data were pooled across orientations (see “Methods”). The bars denote percentage of trials in which each subject reported the bottom surface to be nearer. The two panels show the results from two individual subjects (\pm across-trial SEM). Chance performance (solid line) and $\pm 95\%$ confidence intervals (dotted lines) are indicated. Click on the stimuli to view.

ness of CM-cue stimulus 6 was more variable across subjects, and lower, on average, than that of CM-cue stimulus 5 (two-way ANOVA, subjects \times stimulus types, $p < .05$ for both factors and the factor interaction). We speculate on the significance of this finding below.

We obtained qualitatively similar results when the subjects reported the perceived depth order by adjusting the depth order of panels in a separate matching stereoscopic stimulus (data not shown).

Discussion

It has been argued previously that motion processing is “form-cue” invariant (Albright, 1992) in that motion sensitive neurons have the same directional preferences for features defined by first- and second-order image variation. In natural scenes, however, first- and second-order motions have different origins (Fleet & Langley, 1994), which suggests that they may convey different types of information (Stoner & Albright, 1996). In particular, first-order motion frequently originates from variation in reflectance on a single moving surface, whereas second-order motion typically arises at occlusion boundaries. Even when an occlusion

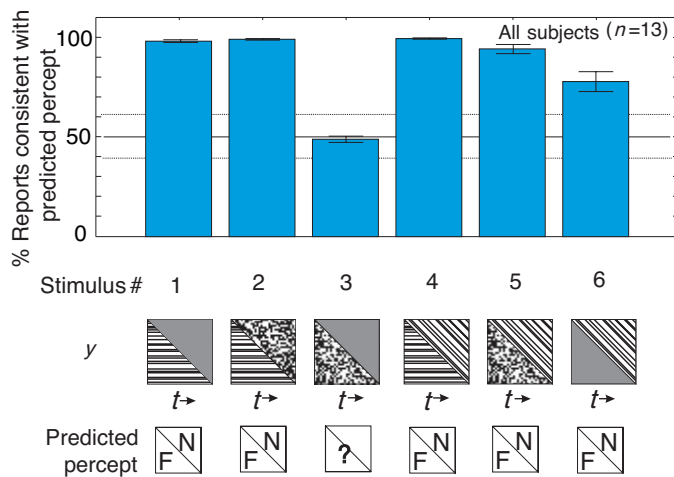


Figure 4. Average depth-order percepts of all subjects. The bars denote the average of all 13 subjects (\pm across-subject SEM) for each stimulus. The solid and the dotted lines denote chance performance and $\pm 95\%$ confidence intervals, respectively. Click on the stimuli to view.

boundary is defined by luminance variation, it does not, by itself, provide any of the depth-order cues discussed here. To illustrate, if the shade and the background scene of our example were both featureless and distinguished only by a luminance (i.e., first-order) difference, there would be no basis for distinguishing occluder from occluded background: A dark surface can occlude a lighter surface just as easily as the other way around. Indeed, first-order motion contributes depth-order information only to the extent that it provides comparative velocity information. For example, though the coherently moving dots in our CM-cue stimuli activate first-order motion mechanisms, extraction of the CM cue requires that the resultant velocity estimate be compared with the velocity of the moving boundary. This velocity comparison, by definition, constitutes a second-order operation.

The finding that our two CM-cue stimuli (stimuli 5 and 6) differed in their effectiveness (Figure 4) raises the important point that the mechanisms underlying perceptual depth ordering may not respect our formal cue descriptions. It is thus possible that these mechanisms interpret differently the flickering texture panel of stimulus 5 and the uniform gray panel of stimulus 6, even though neither panel offers unambiguous velocity or accretion-deletion information. This remains speculation, however, because the precise mechanisms underlying perceptual depth ordering are unknown. It is to be hoped that examination of neuronal responses to these different second-order depth cues will illuminate these mechanisms.

Our results confirm the speculation that depth-order information based on occlusion is conveyed by commonly used second-order motion stimuli (Stoner & Albright, 1996), and that first-order motion cues are unnecessary. Second-order stimuli have been shown to activate a subset

of neurons in a number of cortical areas (Albright, 1992; Smith et al., 1998; Seiffert et al., 2003; Dumoulin et al., 2003). But why, given that moving objects distinguished solely by second-order variation seem relatively rare, would we need this additional motion mechanism? One answer is that second-order mechanisms, unlike first-order mechanisms, can reliably process object motion when faced with variable illumination (see Baker & Mareschal, 2001). Our results suggest another answer, namely that neurons responsive to second-order motion function not merely to extract motion but to extract depth order from dynamic scenes. From this point of view, second-order mechanisms are functionally important whenever one object dynamically occludes another, regardless of whether those objects are distinguished from one another by luminance variation.

Acknowledgments

We thank Greg Horwitz, John Reynolds, and Geoff Boynton for helpful comments on the manuscript, and Laura Abavare for expert technical assistance. We thank William Thompson for helpful discussions. TDA is an Investigator of the Howard Hughes Medical Institute.

Commercial relationships: none.

Corresponding author: Gene R. Stoner.

Email: gene@salk.edu.

Address: Systems Neurobiology Lab, Salk Institute, P.O. Box 85800, San Diego, CA 92186-5800.

References

- Albright, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science*, 255, 1141-1143. [PubMed]
- Baker, C. L., Jr., & Mareschal, I. (2001). Processing of second-order stimuli in the visual cortex. *Progress in Brain Research*, 134, 171-191. [PubMed]
- Cavanagh, P., & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, 4, 103-129. [PubMed]
- Chubb, C., & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America*, 5, 1986-2007. [PubMed]
- Craton L. G., & Yonas, A. (1990). Kinetic occlusion: Further studies of the boundary-flow cue. *Perception and Psychophysics*, 47, 169-179. [PubMed]
- Dumoulin, S. O., Baker, C. L., Jr., Hess, R. F., & Evans, A. C. (2003). Cortical specialization for processing first- and second-order motion. *Cerebral Cortex*, 13, 1375-1385. [PubMed]
- Fleet, D. J., & Langley, K. (1994). Computational analysis of non-Fourier motion. *Vision Research*, 34, 3057-3079. [PubMed]
- Gibson, J. J., Kaplan, G. A., Reynolds, H. N., & Wheeler, K. (1969). The change from visible to invisible: A study of optical transitions. *Perception and Psychophysics*, 5, 113-116.
- Kaplan, G. A. (1969). Kinetic disruption of optical texture: The perception of depth at an edge. *Perception and Psychophysics*, 6, 193-198.
- Mutch, K. M., & Thompson, W. B. (1988). Analysis of accretion and deletion at boundaries in dynamic scenes. In W. Richards (Ed.), *Natural Computation* (pp. 44-54). Cambridge, MA: MIT Press.
- Seiffert, A. E., Somers, D. C., Dale, A. M., & Tootell, R. B. (2003). Functional MRI studies of human visual motion perception: Texture, luminance, attention and after-effects. *Cerebral Cortex*, 13, 340-349. [PubMed]
- Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M., & Hennig, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *Journal of Neuroscience*, 18, 3816-3830. [PubMed]
- Stoner, G. R., & Albright, T. D. (1996). The interpretation of visual motion: Evidence for surface segmentation mechanisms. *Vision Research*, 9, 1291-1310. [PubMed]
- Thompson, W. B., Mutch, K. M., & Berzins, V. A. (1985). Dynamic occlusion analysis in optical flow fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7, 374-383.
- Yonas, A., Craton, L. G., & Thompson, W. B. (1987). Relative motion: Kinetic information for the order of depth at an edge. *Perception and Psychophysics*, 41, 53-59. [PubMed]