

The minimal effect of occlusion on perceived depth from motion parallax

DAVID W. EBY and JACK M. LOOMIS
University of California, Santa Barbara, California

Two experiments tested the ability of subjects to perceive the depths of three-dimensional (3-D) objects oscillating behind a series of occluding bars. Occlusion was manipulated by the spatial period of the occluding bar series as well as by its duty cycle (the proportion of field of view that was not occluded). It was found that the estimates of perceived 3-D shape from motion parallax were remarkably unaffected when the projected image of a rotating 3-D object was partially occluded. Occlusion of 60% of the object had little effect on perceived shape, as judged by estimates of perceived depth, and even occlusion of 90% of the object produced only slight reductions in perceived depth in many cases. It was also found that even when there were no obvious occluding boundaries, judgments of depth were relatively unaffected.

In most natural scenes, near objects often partially occlude more distant objects, yet interpretation of the scene and identification of partially occluded objects seem to proceed effortlessly. How the visual system is able to interpolate the occluded portions of objects as part of the unit formation process and the subsequent process of identification has been the topic of growing interest in recent years (e.g., Kanizsa, 1979; Kellman & Shipley, 1991; Nakayama, Shimojo, & Silverman, 1990).

The concern of this article is with the extent to which occlusion interferes with the perception of shape from optical motion. Previous work (Day, 1989; Fujita, 1990) has demonstrated that the visual system is capable of perceiving three-dimensional (3-D) shape from motion under the most difficult of circumstances, those of anorthoscopic presentation (Rock, 1981)—that is, when an object translates behind a narrow stationary aperture. However, these studies reported only qualitative demonstrations that very simple 3-D shapes could be perceived during their simultaneous rotation and translation behind an aperture; no comparisons were made of the ease and accuracy of perception under these conditions with those of unoccluded presentation.

The present work investigated whether metric shape, as assessed by verbal estimates of object depth, was affected by partial occlusion of the simulated objects, when the only information available about the 3-D object structure was the 2-D projected motions of luminous texture elements randomly distributed on the surface of the object (i.e., motion parallax, MP, information). The visual system's use of MP information in perceiving 3-D objects (in the absence of occlusion) has been investigated extensively (e.g.,

see Braunstein, 1962; Loomis & Eby, 1989; Todd, 1985). The dependent measure in the present study, a verbal estimate of metric depth, has been shown to be quite sensitive to variations in simulated depth and to variables that affect perceived shape (Loomis & Eby, 1988, 1989); this measure allowed us to assess slight variations in perceived depth.

EXPERIMENT 1

Method

Subjects. Four graduate students from the University of California, Santa Barbara, participated for pay. All of the subjects had normal or corrected-to-normal visual acuity in their right eye. All of the subjects were experienced psychophysical observers.

Stimuli and Apparatus. Two simulated 3-D objects were used. The objects were half-ellipsoids that varied only in their base to apex distance (the projected base diameter in each object was 10 cm or 5.7° visual angle). The objects were defined by a set of 247 luminous points (each 0.4 arc min) distributed with equal probability on the surface. Thus, some 2-D static texture cues to depth were present in the imagery. Because we wanted subjects to track an identifiable feature on the object on half of the trials, each object had an additional nine-point fixation target (a plus sign 0.3° wide) plotted on its apex.

On selected trials during the experiment, the subjects were requested to hold their fixation stationary by directing their gaze to a stationary fixation target. This target was a luminous plus sign (0.3° wide) painted with phosphorescent paint on the display screen in the center of the display imagery. To maintain the visibility of the phosphorescent fixation target throughout the experiment, two flood lamps were illuminated for about 4 sec between trials.

On all trials, the half-ellipsoids were simulated as oscillating about a vertical axis within the plane of the ellipsoid base through a range of 30° (15° to either side of the orientation in which the base was parallel to the display screen). With the 60-Hz video refresh rate and the 120 distinct views of the object used in the experiment, one complete oscillation cycle (60°) lasted 2 sec. The animation sequence continued to cycle until the subject made a response.

The main interest of this experiment was the effect of partially occluding the half-ellipsoids with a series of vertically oriented bars. The primary manipulations were changes in bar width and bar spacing. However, we prefer to specify our manipulations with respect to two other variables, *spatial period* and *duty cycle*. Spatial period is the separation between the left (or right) vertical edges of adjacent occluding bars and is thus the sum of bar width and bar spacing; here, it is expressed in terms of number of horizontal pixels. Duty cycle is the ratio of bar spacing divided

This research was supported by Grant 15129 from NINCDS and by a grant from the Academic Senate of the University of California, Santa Barbara. A preliminary report of Experiment 1 was presented at the annual meeting for the Association for Research in Vision and Ophthalmology, Sarasota, FL, in 1989. Correspondence should be addressed to D. W. Eby, Department of Psychology, California State University, San Bernardino, CA 92407.

by spatial period. It specifies the proportion of the field of view that is unoccluded by the bars.

The MP sequences were created on an IBM PS/2 Model 80 micro-computer using software written for the 640×480 resolution mode of the IBM VGA video board. The imagery was displayed on a Zenith Flat Screen RGB monitor (ZTM 1490).

The display was directly viewed with only the right eye from a distance of 1 m. The subject's chin was positioned in an adjustable chinrest. Located beneath the display monitor were a ruler graded in centimeters and a small light source connected to a footpedal. The ruler was used to facilitate a subject's judgment of metric depth. During an experimental trial, the subject could depress the footpedal to make the ruler visible. A black baffle was placed between the monitor and the light source so that the light did not directly illuminate the display.

Design. Four factors were investigated in this experiment. As mentioned above, two of the factors were spatial period and duty cycle. Three spatial periods were investigated: 10, 20, and 40 pixels. These periods corresponded to visual angles of .06°, .11°, and .23°, respectively. Three duty cycles were studied: 0.1, 0.4, and 0.7. The various combinations of spatial period and duty cycle yielded bars that varied in width from 3 to 36 pixels and separations between the bars that varied from 1 to 28 pixels. To compare the occlusion conditions with a condition without occlusion, an additional stimulus with a 1.0 duty cycle was added.

The third factor we investigated was simulated depth of the half-ellipsoids. Two values were studied: 5 and 25 cm. These values produced imagery with perspective ratios (Braunstein, 1962) of 1.05 and 1.25, respectively.

The fourth factor we studied was the type of fixation instruction. We manipulated this factor by instructing the subject to fixate the phosphorescent target on the screen on half of the trials and to track the fixation target located on the back of the object during the other half of the trials.

All possible combinations of the first three factors (3 spatial periods × 3 duty cycles × 2 object depths) yielded 18 stimulus conditions. In addition, there was a stimulus condition in which the duty cycle was 1.0 (no occlusion) for both simulated object depths, yielding a total of 20 unique stimulus conditions in the experiment. Within an experiment, the subject viewed each block of conditions in random order six times. Three of the blocks were viewed with stationary fixation instructions and three were viewed with tracking fixation instructions. Each block alternated in the type of viewing instructions. Half of the subjects started with stationary fixation instructions.

Procedure. The subjects were run individually. Prior to the experiment, the subjects were shown, with the aid of a diagram, the dimension of the half-ellipsoid that they were being asked to judge. They were told to respond to the apparent depth extent of the half-ellipsoid between its base and its apex. Because perspective displays occasionally reverse in depth, and such reversals can affect estimates of apparent depth, the subjects were instructed to maintain a constant orientation of the object in depth. When a spontaneous reversal occurred, they were told to wait until the object reverted to its original configuration prior to making their judgment. No subject reported any difficulties in following these instructions; the subjects reported that they saw most of the displays as 3-D and based their judgment on this percept.

A trial proceeded as follows. The experimenter extinguished the floodlights and initiated the stimulus sequence. The subject viewed the receding stimulus sequence for as long as was necessary to make a judgment of perceived depth. When ready, the subject verbally reported his or her depth impression in centimeters, using the ruler to provide a scale for the response. The experimenter terminated the stimulus, recorded the depth value, turned on the floodlights, and timed the 4-sec interval for the floodlights.

Results

The average estimates of the 4 observers are shown in Figure 1. The left panel depicts the results for tracking fixation instructions; the right panel shows the results for the stationary fixation instructions. Within each panel are the results for the two different-sized objects and the three spatial periods as a function of duty cycle. A comparison of the results between the two panels shows that when

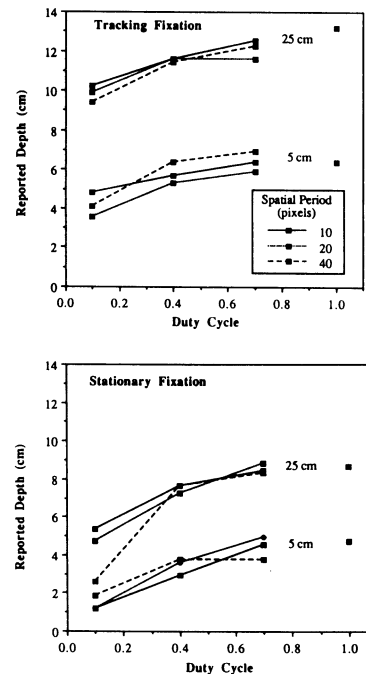


Figure 1. The results of Experiment 1. The unconnected points are the average estimates of depth when no occlusion was present. Note that depth estimates in the lower duty cycles are quite robust, including the 10-pixel spatial period, 0.1 duty-cycle condition in which the visible area between the occluding bars was only 1 pixel.

subjects are given instructions to fixate a stationary target while the stimulus moves behind the target, the average estimates of depth tended to be smaller in value than when they fixate a target that moves with the stimulus. A 2 (viewing instructions) × 2 (depth) × 3 (spatial period) × 3 (duty cycle) repeated measures analysis of variance (ANOVA) on the estimates of depth (excluding the 1.0 duty cycle) showed this effect to be statistically significant [$F(1,3) = 9.10, p < .05$].

As can be seen by the nonoverlap of the functions for the 25- and the 5-cm objects in Figure 1, the large object was judged to be consistently deeper than the small simulated object across all conditions. This main effect was statistically significant [$F(1,3) = 29.74, p < .02$] and is consistent with the results of several studies (e.g., Braunstein & Tittle, 1988; Loomis & Eby, 1988, 1989). More interestingly, there was no main effect of spatial period [$F(2,6) = 1.20$], but there was a significant interaction between the depth and the spatial-period variables [$F(2,6) = 11.70, p < .01$]. Estimates of depth were significantly affected by the duty cycle of the occluding bars [$F(2,6) = 10.40, p < .02$]. None of the other comparisons were statistically significant.

Discussion

The principal finding of this experiment is that placing a series of bars in front of a rotating object had relatively little effect on the perceived shape of the object, as judged by the subjects' verbal estimates of perceived depth. Although reducing the duty cycle led to reliable decreases in depth estimates for all conditions, occluding even 90% of the object points led to only modest reductions in many of the 0.1 duty-cycle conditions.

One of our manipulations was the mode of fixation used by the subject. When the subject maintained fixation on a target stationary on the computer display, estimates of perceived depth were consistently smaller than they were when the subject tracked a target positioned at the apex of the rotating object. This effect does not seem to result solely from the occlusion manipulations. If it did, we would expect to observe little difference in the conditions in which there was no occlusion (i.e., 1.0 duty cycles) across the different viewing conditions, for each simulated object depth. As shown in Figure 1, this expectation was not met; the average depth estimates for the 1.0 duty-cycle conditions were much smaller with stationary fixation than with moving fixation. Similar results have been previously reported with polar projections of translating dihedral angles (Braunstein & Andersen, 1981; Farber & McConkie, 1979).

These results suggest that fields of MP, even when disrupted both spatially and temporally, are readily interpretable in terms of depth variation. However, before concluding that it is MP that conveyed depth to the subjects, we need to consider three alternative possibilities. First, it might be argued that static variation of texture density underlies the perception of depth. This possibility is ruled out by the fact that little depth is perceived when the parallax is removed, leaving only static variations of texture density, as when the object is rotated about the line of sight (Loomis & Eby, 1988). With respect to the present stimuli, we have also observed that depth is not perceived with static images. Second, it is possible that the subjects in Experiment 1 based their estimates on the information provided by the changing projected contour of the half-ellipsoid as it oscillated. This type of information in isolation has been shown to provide information that is useful to an observer in judging object shape (e.g., see Loomis & Eby, 1989; Todd, 1985). A third possibility is that the subjects may have used information provided by the fixation target on the apex of each object. This target provided a feature that could have appeared unique between the different objects when they oscillated. This feature, then, could have been used by the subject as a means of object identification, independently of the perceived depth in the object. Experiment 2 was designed, in part, to test the latter two possibilities.

EXPERIMENT 2

Experiment 1 revealed a minimal effect of a series of occluding bars in front of a MP display; estimates of depth were diminished relatively little even when as few as 10% of the stimulus points were visible during any single view of the simulated object. One possible interpretation of this effect is suggested by the results of Nakayama, Shimojo, & Silverman (1990), who have shown that previously unrecognizable objects behind occluding elements become recognizable when the object is stereoscopically separated from the occluders. They suggest that the visual system uses an occlusion constraint in resolving ambiguity about a stimulus. The stereoscopic depth provides information that a patch is an occluder in front of an object and this information is used to "fill in" the information beneath the occluding patch, facilitating the object-recognition process. It is possible that in our stimuli a similar process was occurring. As the simulated object oscillated, its image points moved back and forth horizontally, appearing in one aperture, being occluded temporarily, and then reappearing in the adjacent aperture. With several points moving behind the dark occluding bars, a subjective contour is produced at the accretion and deletion boundary (e.g., Andersen & Braunstein, 1983; Andersen & Cortese, 1989; Yonas, Craton, & Thompson, 1987). When such a boundary is stationary, it is usually seen as being in front of the texture elements (Yonas et al., 1987). This ordinal depth information (called the boundary flow cue by Yonas et al., 1987), combined

with an occlusion constraint (Nakayama et al., 1990), might enable observers to more effectively interpolate a 3-D surface behind the occluding bars and to enhance their judgment of metric depth. We tested this possibility by comparing "correlated" bars of the type studied in Experiment 1 with "uncorrelated" bars in which the occluding bar series for each image point was vertically misaligned. If the robust performance observed at the small duty cycles is based on an interpretation that the object is oscillating behind a series of occluding bars, then depth estimates should dramatically decrease in the uncorrelated bar conditions in which no subjective occluding surfaces were apparent.

Method

Subjects. Five graduate students from the University of California, Santa Barbara, participated as paid subjects. All of the subjects had normal or corrected-to-normal visual acuity in the right eye.

Stimuli and Apparatus. The simulated objects were the same as those used in Experiment 1, except that no fixation target was present. The correlated bars were produced in the same way as in Experiment 1. The uncorrelated bars were produced in software by assigning each point a horizontal set of bars with fixed period, duty cycle, and horizontal phase (with respect to the edge of the display screen). The bars were decorrelated by assigning a fixed random phase for each point. Thus, both types of occluding bars were locally identical, but they differed globally with respect to the alignment of the bar series associated with different points. The apparatus was the same as in Experiment 1, except that a circular mask was centered around the display imagery.

Design and Procedure. Four factors were investigated: simulated object depth (5 and 10 cm), spatial period (10, 20, and 40 pixels), duty cycle (.1, .4, and .7), and type of occluding bar series (correlated and uncorrelated). The perspective ratios for the 5- and 10-cm objects were 1.05 and 1.10, respectively. Two stimulus conditions (one for each simulated object depth) in which no occlusion was present were also included.

The procedure was the same as in Experiment 1. The subject participated in a single session in which he or she judged all 38 stimulus conditions three times in random order.

Results

The average depth estimates as a function of type of occluding bars, simulated object depth, spatial period, and duty cycle are shown in Figure 2. A 2 (aperture type) \times 2 (depth) \times 2 (period) \times 4 (duty cycle) ANOVA was performed on the depth estimates (excluding the 1.0 duty cycle). As in Experiment 1, depth estimates increased with increasing simulated depth [$F(1,4) = 21.34, p < .01$]. There were also significant main effects of duty cycle [$F(2,8) = 7.16, p < .02$] and spatial period [$F(2,8) = 6.06, p < .05$].

Overall, the estimates of depth for the different types of occluding bars were essentially the same as those suggested by the nonsignificant main effect of type of occluding bar series [$F(1,4) = .02, p > .5$]. There was, however, a reliable interaction between this factor and duty cycle [$F(2,8) = 8.67, p < .01$].

Discussion

This experiment showed that an interpretation of the stimulus as an object oscillating behind a series of occluding bars slightly facilitates the depth processing of the MP information. Despite the fact that there was no overall main effect of occluding bar type, there was a significant interaction between this variable and duty cycle, suggesting that the duty cycle of the occluding bar series produced differential depth estimates depending on the type of occlusion that was present. This effect is seen in Figure 2, in

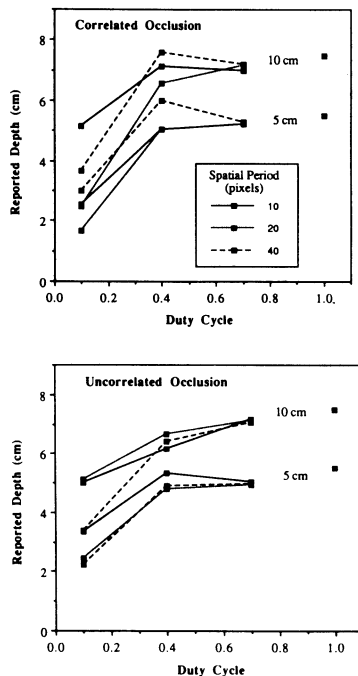


Figure 2. The results of Experiment 2. The unconnected points are the average estimates of depth when no occlusion was present; accordingly, these points are the same between the two panels. The depth estimates in the 0.1 duty cycle are quite robust.

which the depth estimates for the uncorrelated occluding bars began to decline as the duty cycle was decreased from 1.0 to 0.4; in the correlated occluding bar cases, this decline did not occur until the .1 duty cycle. This indicates that the disruption in the recovered depth from the MP processing mechanism occurred at a smaller duty cycle when the occluding bars were vertically misaligned than when they were vertically aligned. Because the only difference between the two stimulus conditions was this alignment, we conclude that the alignment facilitated depth estimates in the correlated occlusion conditions for the .4 duty cycle.

GENERAL DISCUSSION

These experiments demonstrate that the perception of 3-D shape from motion parallax is remarkably unperturbed when the projected image of a rotating 3-D object is partially occluded. Occlusion of 60% of the object has little effect on perceived shape, as judged by reports of perceived depth, and even occlusion of 90% of the object produces only slight reductions in perceived depth for many of the conditions.

A second finding was that perceptual organization of the stimulus with respect to the occluded object and the occluding surfaces is not required in order for observers to perceive the shape of the occluded objects, for even when there were no obvious occluding boundaries, judgments of depth were relatively unaffected. At the same time, however, organization with respect to occluding surfaces did have a slight effect for the 0.4 duty-cycle condition.

Perhaps the most surprising result of all is the amount of depth reported for conditions involving a spatial period of 10 pixels and a duty cycle of 0.1. In these conditions, the apertures through which the ob-

ject points were visible were only 1 pixel wide, and the occluded regions were 9 pixels wide. This means that once a moving object point appeared, it could only reappear, if at all, at the same location or at some other location a multiple of 9 pixels away. Not only is there no motion signal associated with any single object point on the scale of 10 pixels (corresponding to an angular separation of $.06^\circ$), but the problem of solving the correspondence of object points would seem to be formidable. Clearly, it must be that the motion signals and information for solving correspondence are conveyed by the overall patterning of the appearance and disappearance of illuminated pixels on the screen.

REFERENCES

- ANDERSEN, G. J., & BRAUNSTEIN, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, *34*, 356-362.
- ANDERSEN, G. J., & CORTESE, J. M. (1989). 2-D contour perception resulting from kinetic occlusion. *Perception & Psychophysics*, *46*, 49-55.
- BRAUNSTEIN, M. L. (1962). Depth perception in rotating dot patterns: Effects of numerosity and perspective. *Journal of Experimental Psychology*, *64*, 415-420.
- BRAUNSTEIN, M. L., & ANDERSEN, G. J. (1981). Velocity gradients and relative depth perception. *Perception & Psychophysics*, *29*, 145-155.
- BRAUNSTEIN, M. L., & TITTLE, J. S. (1988). The observer-relative velocity field as the basis for effective motion parallax. *Journal of Experimental Psychology: Human Perception & Performance*, *14*, 582-590.
- DAY, R. H. (1989). Apparent depth from progressive exposure of moving shadows: The kinetic depth effect in a narrow aperture. *Bulletin of the Psychonomic Society*, *27*, 320-322.
- FARBER, J. M., & MCCONKIE, A. B. (1979). Optical motions as information for unsigned depth. *Journal of Experimental Psychology: Human Perception & Performance*, *5*, 494-500.
- FUJITA, N. (1990). Three-dimensional anorthoscopic perception. *Perception*, *19*, 767-771.
- KANIZSA, G. (1979). *Organization in vision*. Chicago: University of Chicago.
- KELLMAN, P. J., & SHIPLEY, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, *23*, 141-221.
- LOOMIS, J. M., & EBY, D. W. (1988). Perceiving structure from motion: failure of shape constancy. In *Proceedings of the Second International Conference on Computer Vision* (pp. 383-391). Washington, DC: Computer Society Press.
- LOOMIS, J. M., & EBY, D. W. (1989). Relative motion parallax and the perception of structure from motion. In *Proceedings of the IEEE Workshop of Visual Motion* (pp. 204-211). Washington, DC: IEEE.
- NAKAYAMA, K., SHIMOJO, S., & SILVERMAN, G. H. (1990). Stereoscopic depth: Its relation to image segmentation, grouping and recognition of occluded objects. *Perception*, *18*, 55-68.
- ROCK, I. (1981). Anorthoscopic perception. *Scientific American*, *244*, 145-153.
- TODD, J. T. (1985). The perception of structure from motion: Is projective correspondence of moving elements a necessary condition? *Journal of Experimental Psychology: Human Perception & Performance*, *11*, 689-710.
- YONAS, A., CRATON, L. G., & THOMPSON, W. B. (1987). Relative motion: Kinetic information for the order of depth at an edge. *Perception & Psychophysics*, *41*, 53-59.