

Recovering viewer-centered depth from disparity, occlusion, and velocity gradients

MYRON L. BRAUNSTEIN

University of California, Irvine, California

GEORGE J. ANDERSEN

University of Illinois at Urbana-Champaign, Champaign, Illinois

MICHAEL W. ROUSE

Southern California College of Optometry, Fullerton, California

and

JAMES S. TITTLE

University of California, Irvine, California

Two experiments were conducted to assess the effects of corresponding and conflicting binocular and monocular information on the recovery of depth order (signed depth). Subjects viewed displays in which the same or opposite depth orders were indicated by disparity and occlusion, in one experiment, or by disparity and velocity gradients, in a second experiment. The same 36 subjects, 17 who had failed a Random Dot E test and 19 who had passed, were run in both experiments. When binocular and monocular information indicated conflicting depth orders, most subjects responded in accordance with the monocular information on some trials in both experiments. This was true even for a subgroup who always responded in accordance with the stereoscopic information on control trials that did not provide monocular information for depth order. For this subgroup, the impact of conflicting monocular information in the velocity gradient task correlated with performance on the uncrossed version of the Random Dot E test. We also found that some subjects who failed static tests of stereoscopic depth perception could respond accurately to continuously changing disparities.

The information that human observers recover about the three-dimensional environment from two-dimensional images normally includes both the shapes of objects and the relative distances of objects, or parts of objects, from the observer. Shape information can be represented in object-centered coordinates, without regard to the location of the viewer (Marr & Nishihara, 1978). Relative distances, on the other hand, are defined in relationship to a viewer and would be represented more appropriately in a viewer-centered (or viewer-relative) coordinate system. Relationships between object-centered and viewer-centered representations have been of interest in both the literature on depth perception (see, e.g., Braunstein, 1983, in press) and the literature on long-term memory (Jolicoeur & Kosslyn, 1983). (See also Sedgwick, 1983, for a discussion of environment-centered representations.) According to Marr (1982), a viewer-centered representation, the "2½-D sketch," is recovered from informa-

tion in the retinal image, and is used to build an object-centered representation, the "3-D model." These representations are thus based on the same distance information, and both representations could be derived from a single source of depth information. Consider, for example, a situation in which binocular disparity is the only source of information about the distances from the observer to feature points on the surface of a sphere. A viewer-centered representation would provide the relative distances of each point from the observer. This set of relative distances from the observer, in turn, could specify the spherical shape (provided that the depth dimension is scaled correctly relative to the horizontal and vertical dimensions). If human depth perception always followed this two-stage path, from a viewer-centered representation to an object-centered representation, any process that provided an unambiguous perception of object shape might be expected to provide an unambiguous perception of the relative distances of parts of the object from the observer. The recovery of three-dimensional structure from two-dimensional images, however, appears not always to fit this paradigm. Consider the information recovered from a parallel projection of an object rotating about an axis other than the line of sight, a perception

This research was supported by National Eye Institute Grant EY 04553. We thank Lionel Shapiro for assistance in various aspects of this research and a reviewer for helpful comments. Requests for reprints should be sent to Myron L. Braunstein, Cognitive Sciences Group, School of Social Sciences, University of California, Irvine, CA 92717.

to which the terms "kinetic depth effect" (Wallach & O'Connell, 1953) and "structure-from-motion" (Ullman, 1979) have been applied. Shape can be recovered in these displays in object-centered coordinates, but important viewer-centered information is absent: The recovered shape cannot be distinguished from its reflection along the line of sight (Braunstein, 1966; Ullman, 1979). Depth order, or signed depth, cannot be resolved unless other sources of information, such as polar perspective (Braunstein, 1966), occlusion (Braunstein, Andersen, & Riefer, 1982), or relative brightness (Schwartz & Sperling, 1983) are added.

These considerations suggest that there are two paths to the recovery of representations of the three-dimensional environment that include both shapes of objects and distances of objects relative to an observer: (1) Relative distances (viewer-centered information) may be recovered from information in the retinal projections, with shape (object-centered information) derived from these distances; or (2) shape (object-centered information) may be recovered directly from information in the retinal projections with viewer-centered information (depth order) added from additional sources.¹ Binocular disparity may be considered prototypical of the first path, as disparity appears to provide both shape information and unambiguous depth order. Structure-from-motion may be considered as prototypical of the second path, as the recovery of shape without depth order has been demonstrated for parallel projections of rotation in depth (Braunstein, 1966, 1977; Petersik, 1980).²

The ambiguity of depth order in a parallel projection simulating an object rotating in depth can be resolved by adding occlusion to the display (Braunstein et al., 1982). In that study, the displays simulated parallel projections of pentagonal texture elements on the surface of a rotating sphere. In one condition, the surface of the sphere was treated as opaque and the elements disappeared as they rounded the edge (edge occlusion). In another condition, the surface was treated as transparent, but the texture elements were treated as opaque, so that the projections of nearer elements covered and uncovered those of more distant elements as the sphere rotated (element occlusion). Direction-of-rotation judgments were used to determine the accuracy with which depth order was recovered in accordance with occlusion. Accuracy was determined primarily by the type of occlusion, with edge occlusion providing a generally higher level. For the element-occlusion displays, element size determined the level of accuracy. In a later experiment (Andersen & Braunstein, 1983), we showed that the presence of contours is not necessary for occlusion to be effective in resolving ambiguities in depth order.

These studies demonstrated that viewer-centered information (depth order) can be added to an object-centered representation (shape) by supplementary sources of information, such as occlusion. This may be contrasted to the paradigm proposed by Marr in which an object-centered representation is derived from viewer-centered

information. The present study was concerned with the relationship of these two paths to the recovery of viewer-centered depth, when information for both paths is available. Specifically, we considered whether information that is used to resolve ambiguities in depth order in dynamic monocular displays continues to affect perceived depth order when unambiguous information is available from binocular disparity.³ Experiment 1 considered the influence of occlusion on judgments based on depth order when conflicting information about depth order was provided by binocular disparity. Two additional questions were addressed: If occlusion continues to influence perceived depth order when conflicting information is provided by disparity, does the use of occlusion depend on the same variables (edge vs. element occlusion and element size) that determined the effectiveness of occlusion in the earlier, monocular research? Is the relative use of occlusion and disparity, by individual subjects, related to performance on tests of static stereoscopic depth perception?

We examined direction-of-rotation judgments for stereoscopic versions of displays from Braunstein et al. (1982). Each display was generated with disparity indicating a depth order corresponding to that indicated by occlusion or with disparity indicating a depth order in conflict with that indicated by occlusion. Control stimuli were added in which no occlusion information was available and depth order could be resolved only through disparity. Use of disparity in resolving depth order was compared with performance on two tests of static stereoscopic depth perception.

The second experiment addressed similar questions for displays in which depth order can be resolved from the polar perspective information available in a velocity gradient produced by translating slanted surfaces (Braunstein & Andersen, 1981). If the perceptual analysis of this type of display involves recovery of relative distances from polar perspective and the recovery of shape from these relative distances, it would fit the first proposed path (Marr's paradigm). However, for reasons discussed in the introduction to Experiment 2, we believe that the perceptual analysis of velocity gradients, unless linked to head movements (Rogers & Graham, 1979), follows the second path: Shape is initially recovered by means of a structure-from-motion analysis, and depth order is subsequently resolved through the use of supplementary information.

EXPERIMENT 1

Method

Subjects. The subjects were 36 students in lower division psychology courses at the University of California, Irvine, who received extra credit for their participation. In selecting the subjects, 235 students were screened by staff members of the Southern California College of Optometry. The screening began with the Random Dot E test (Reinecke & Simons, 1974). Both crossed and uncrossed disparity targets were presented twice to each subject. The targets were exposed until the subject responded, usually 5–10 sec,

up to a maximum of 30 sec. Seventeen subjects, who failed to respond correctly to either a crossed or uncrossed presentation, were selected for participation. Two of these subjects did not bring their corrective lenses to the screening session but wore them during the experimental session. Their data is excluded from the correlations, reported below, that involve the E test. Nineteen additional subjects were selected at random from among those who passed both versions of the E test. Three additional subjects were run, but their data were not used because of apparatus problems. Nineteen subjects, 11 from the first group and 8 from the second group, participated first in Experiment 2. The remaining subjects participated first in Experiment 1.

Design. Four stimulus variables were examined in the main design: type of occlusion (edge or element), size of the texture elements (the three largest sizes used by Braunstein et al., 1982), correspondence versus conflict of the stereoscopic and occlusion information for depth order, and direction of rotation (as indicated by the stereoscopic information). There were also six control stimuli with no occlusion displayed. These varied in size of the texture elements and direction of rotation. The classification of the subjects on the basis of the Random Dot E test was not treated as a variable in the design because of the low correlation between performance on this test and performance on the control trials, as discussed below.

Stimuli. The stimuli were 30 15-sec (360-frame) computer-generated 16-mm motion picture sequences. The sequences were generated one frame at a time on a Tektronix 4010 display scope, recorded on Kodak Plus-X film using an Automax camera, and printed on high-contrast film. The 30 sequences were randomly arranged on the film with a 10-sec (240-frame) ISI. Each sequence displayed two views, separated horizontally, of a parallel projection of texture elements (irregular pentagons) located on the surface of a sphere rotating 360° about a vertical axis. The resolution of each view was 420 × 420. The two views were from the same 360° rotation sequence, but were 5° out of phase. Examples of frames from the film are shown in Figure 1. The methods used to

generate the pentagons and to locate them on the surface of the sphere, and the distinction between edge and element occlusion, were the same as those used by Braunstein et al. (1982). For each edge- and element-occlusion display, there was a version produced in which the left and right positions of the two views on the film frame were reversed, producing a conflict between the direction of rotation indicated by occlusion and the direction indicated by disparity. In addition to the edge- and element-occlusion displays, there were control displays in which both the texture elements and the surface of the sphere were treated as transparent; thus, the direction of rotation was specified only by binocular disparity.

Three stationary stereo pairs were used for preliminary and checking purposes. Two—the “square targets”—displayed small squares inside larger squares, with the smaller squares displaced either temporally or nasally relative to the large squares. One pair—the “cross target”—displayed a circle enclosing a horizontal line in one member of the pair and a vertical line in the other.

Apparatus. A 16-mm projector (Bell and Howell, Model 2585) was used to present the film. The film was projected through a glass window onto a translucent screen (Polacoat) in a separate room. The pentagon and background luminances at the screen were approximately 1.8 and .03 cd/m², respectively. The projected image of a sphere on the screen was 5 cm in diameter. The center-to-center separation was 6 cm. The displays were viewed through a Brewster-type stereoscope (Keystone No. 50 Home Training Stereoscope), with the lenses located 15.2 cm from the screen. Black cloth around the stereoscope concealed the screen from the subject, except for the portion viewed through the lenses. The diameter of the sphere in visual angle, viewed through the stereoscope, was 20.5°. The maximum disparity between contours in the two views was 799" of arc. The convergence required to align the images was 4.8°. ⁴

The response device consisted of a 50-mm-diam spherical knob mounted on a vertical shaft. Turning the knob clockwise or counterclockwise triggered one of two microswitches which turned on one of two lights at the experimenter's console. A 15-cm-diam glass sphere with white dots unsystematically positioned on the surface served as a demonstration model, used only during the reading of the instructions. An 8-rpm reversible motor was attached to the vertical axis of the glass sphere.

Procedure. The subjects were run individually. The two square targets were used to acquaint the subject with the viewing apparatus. The subjects then were instructed to indicate, by turning the response device clockwise or counterclockwise, the direction of rotation of the sphere viewed through the stereoscope as soon as each display disappeared. The two directions of rotation were demonstrated using the glass display model, and the subject practiced the response procedure. The room was then made totally dark, and the film was started. Different orders of stimulus presentation were obtained by starting the film in different locations. Each order was used for either 1 or 2 subjects. The sequences were repeated twice for each subject, in the same order. The first five sequences, which served as practice trials, were repeated at the end of the session.

The cross target was presented twice during the session (between stimuli) as a check against possible suppression of the view in one eye. The subject was asked for a verbal description of what was seen. All subjects reported seeing the horizontal line and vertical line as crossed (except in two of the three cases of apparatus problems noted above). All subjects were tested on the RANDOT Circles scale following their completion of both Experiment 1 and Experiment 2.

Results

In considering the competition and cooperation of stereoscopic and monocular information in determining depth order, it would be useful to analyze the effects of the stimulus variables separately for subjects who perform

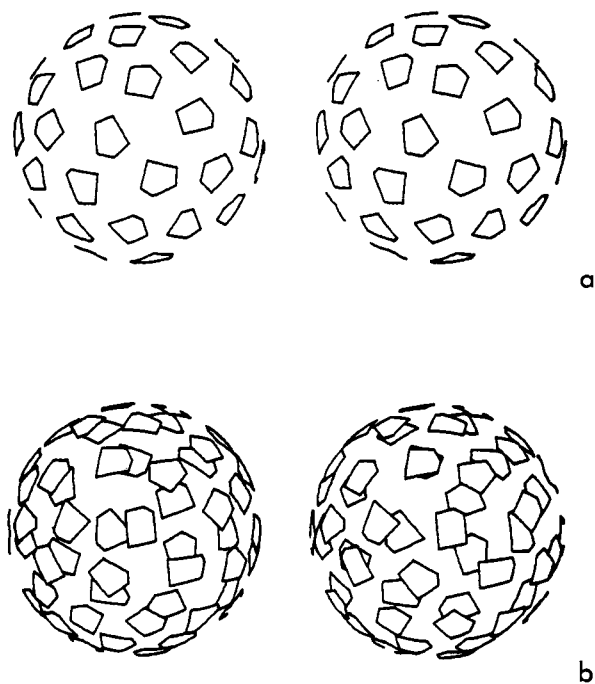


Figure 1. Stereo pairs illustrating conflicting depth order indicated by disparity and occlusion, for (a) edge occlusion and (b) element occlusion, for the largest texture-element size.

well on tests of stereoscopic vision and for those who perform poorly. However, only a small number of the people who fail particular measures of stereoscopic depth perception consistently fail all measures (see, e.g., Patterson & Fox, 1984). We considered three measures of stereoscopic depth perception: (1) failure on a Random Dot E test, defined as responding incorrectly on either of two crossed-disparity presentations or either of two uncrossed-disparity presentations; (2) the score on the RANDOT Circles test (0-10); and (3) the proportion of control stimuli, which contained no monocular information for depth order, on which the subject responded in accordance with the stereoscopic information. We found a significant correlation (point-biserial) of .594 ($p < .01$) between the first two criteria. The correlations between the first two criteria and the third criterion were .338 and .089, respectively, and were not statistically significant. The analysis of variance results for the four stimulus variables were essentially the same regardless of whether or not subjects who failed the E test or received low scores on the Circles test were eliminated. The third criterion, however, was used in the following manner: To assure that a group of subjects could be defined who were able to perform the direction-of-rotation task on the basis of stereoscopic information alone, the subjects were classified according to whether or not they responded in accordance with the stereoscopic information on all 12 control trials. A separate analysis of variance was conducted for the four stimulus variables for the 19 subjects who met this stringent criterion. Results will be reported for all subjects and also for this "special group" of 19 subjects.

For all subjects, the main effects of type of occlusion [$F(1,35) = 62.9$] and of correspondence of stereoscopic and monocular information [$F(1,35) = 65.8$], and the interaction of these variables [$F(1,35) = 26.4$], were sig-

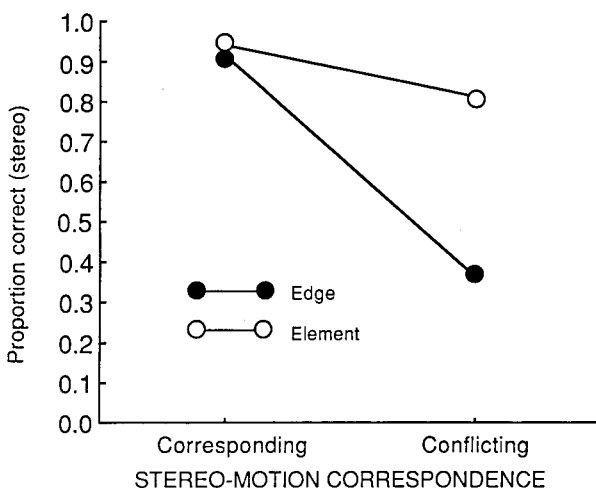


Figure 2. Effect of conflicting monocular information on the proportion of direction-of-rotation judgments made in accordance with disparity, for edge- and element-occlusion displays.

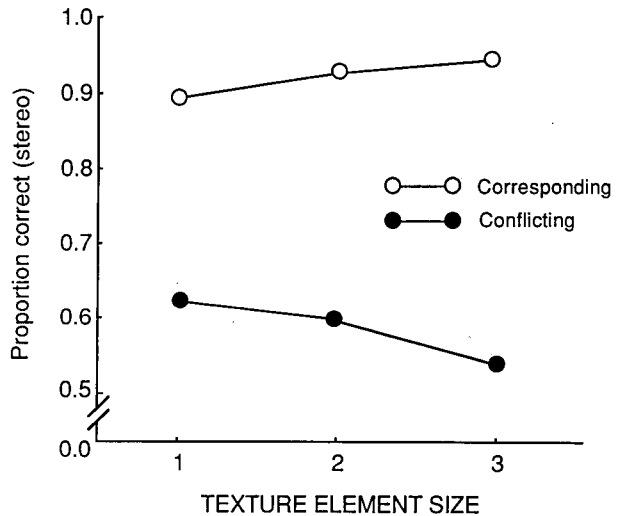


Figure 3. Effect of conflicting monocular information on the proportion of direction-of-rotation judgments made in accordance with disparity, as a function of texture-element size.

nificant ($p < .01$). As shown in Figure 2, subjects responded less frequently in accordance with the stereoscopic information, in the presence of conflicting occlusion, for the edge-occlusion displays. There are two possible reasons for this difference. First, edge occlusion is more effective as an indicator of depth order for monocular stimuli (Braunstein et al., 1982), and information that is more effective monocularly may simply be more effective in competing with stereoscopic depth. Alternatively, there may have been a difference in the reversibility of stereoscopic depth in the edge- and element-occlusion displays. In the element-occlusion displays, pentagons at the front will move across pentagons on the back of the sphere, providing a large difference in disparity for adjacent contours. For the edge-occlusion displays, only one hemisphere is visible, providing only a gradual variation in disparities across the display. Considering the importance of relative disparity in stereoscopic depth perception (Gillam, Flagg, & Finlay, 1984), the relative resistance to reversal of the element-occlusion displays is not surprising. The large difference between the use of stereoscopic information in judging depth order for the edge- and element-occlusion displays suggests that a combination of these two factors—greater monocular effectiveness of edge-occlusion displays and greater stereoscopic effectiveness of the element-occlusion displays—may have been responsible, but this will have to be determined in future research.

There was a significant interaction of texture-element size with correspondence of stereoscopic and monocular information [$F(2,70) = 11.8$]. As shown in Figure 3, the use of stereoscopic information decreased with increasing element size in the conflicting occlusion condition. This probably reflects the increased effectiveness of occlusion with increasing element size, as found by Braun-

stein et al. (1982). The remaining main effects for all subjects were not significant ($F < 1$), and there were no other significant interactions.

The main effects of type of occlusion and of correspondence of stereoscopic and monocular information, and the interaction of these two variables, were also significant for the special group [$F(1,18) = 31.4$, $F(1,18) = 34.1$, and $F(1,18) = 31.0$, respectively, $p < .01$]. A significant interaction of texture-element size with correspondence of monocular and stereoscopic information also was found for that group [$F(2,36) = 8.34$, $p < .01$]. The results for that group differed only in the finding of additional significant effects. Significant main effects were found for direction of rotation [$F(1,18) = 5.59$, $p < .05$] and for texture-element size [$F(2,36) = 8.35$, $p < .01$]. There was also a significant interaction of these two variables with type of occlusion [$F(2,36) = 3.67$, $p < .01$]. Direction-of-rotation judgments were more often in accord with the stereoscopic information when this information indicated a clockwise direction, especially for the edge-occlusion displays. This probably reflects a direction bias that, for this group of subjects, sometimes competed successfully with the information for direction of rotation. The significant effects of correspondence of stereoscopic and monocular information, and of the interaction of this variable with type of occlusion, for the special group demonstrates that these results were not due to any inability to perceive the stereoscopic depth relationships in the displays.

To determine whether the impact of conflicting monocular information on judgments of direction of rotation was related to performance on static tests of stereoscopic depth perception, a monocular impact score was computed for each subject. Impact was defined as (corresponding - conflicting)/corresponding, where the values for corresponding and conflicting were the number of responses made by a subject in accordance with the stereoscopic information in the presence of corresponding or conflicting monocular information, respectively. The impact scores did not correlate significantly with the E test or the Circles test, either for all subjects or for the special group.

EXPERIMENT 2

In the displays used in Experiment 1, monocular information about the spherical shape was available in the orthographic projection of the rotation (structure-from-motion) and monocular information about depth order was available from occlusion, when present. Stereoscopic information was available for both shape and depth order. In Experiment 2, we considered a second case in which both shape and depth order could be resolved from monocular information. The velocity gradient displays are different in several respects. First, they are not kinetic depth or structure-from-motion displays (at least in the more limited sense of those terms), in that shape cannot be recovered from orthographic projections. An ortho-

graphic projection of a translation along an axis perpendicular to the line of sight does not provide information about three-dimensional shape (Braunstein, 1966). Polar perspective must be used if shape is to be recovered (assuming that static depth information is controlled). The information in a horizontal translation, viewed with a polar projection, is sometimes classified as motion parallax. It is probably best, however, to restrict that expression to cases of head-movement parallax, and to refer to the information available in passively observed translations of this type as velocity gradients.

Although velocity gradients cannot be the result of orthographic projections of pure translations, it is possible that the visual system analyzes velocity-gradient displays in a manner similar to that in which it analyzes orthographic projections of rotational motion (Hoffman, personal communication, February 1984; see also Braunstein, in press). According to this suggestion, the velocity gradient is subjected to a local analysis appropriate to an orthographic projection of a rotation about a vertical axis. The differences in the projected velocities are attributed to differences in proximity of the points to the axis of rotation, rather than to the differences in distance from the observer that are being simulated. The faster moving points in the projection are processed as if they were more distant from the axis and the slower moving points are processed as if they were closer to the axis. This perceptual analysis would be approximately consistent with the display if applied locally in space and time, but of course would be incorrect globally. Evidence consistent with this analysis was found in the earlier research by Braunstein and Andersen (1981). First, latency was shorter for recovery of shape than for recovery of depth order. Second, there was a perception of rotational motion, especially when the depth order was reported incorrectly. Perceived rotation was also found by Rogers and Collett (1985), when motion parallax indicated depth variations but binocular disparity indicated constant depth. Regardless of whether this suggestion of initial orthographic analysis of velocity-gradient information is correct, the velocity-gradient displays represent a second case in which dynamic monocular information—in this case polar perspective—can be used to resolve depth order.

There is a second difference between the velocity-gradient displays and the rotating-sphere displays which may be important. Although both displays are dynamic, in the sense that almost all points are in continuous motion, the Z coordinates of points on the rotating sphere change continuously, whereas the points in the velocity-gradient displays maintain constant Z coordinates as they translate along the X axis. The disparity of each visible point in the sphere displays thus changes continuously over time, whereas each point in the velocity-gradient displays maintains a constant disparity.

Experiment 2 examined the cooperative and competitive effects of polar perspective and disparity in determining judged depth order. The relationship between the conditions under which the monocular information is most effective in isolation and conditions that determine its ef-

fectiveness in competition with stereoscopic information were also considered. Finally, the relationship between the use of stereoscopic depth information and performance on static tests of stereoscopic ability was measured.

Method

Subjects. The subjects were the same as in Experiment 1. Seventeen subjects had participated first in Experiment 1; the remaining subjects participated first in Experiment 2.

Design. There were four stimulus variables: corresponding versus conflicting monocular and stereoscopic information for depth order, translation speed (two levels), orientation of the dihedral angle (center-near or center-far), and direction of translation (left-to-right or right-to-left).

Stimuli. The stimuli were 16 15-sec computer-generated displays of 250 bright green dots against a dark green background. The positions of the dots on the display scope were randomly determined with a uniform distribution. The dots moved along horizontal linear trajectories. The velocities of the dots were computed as if the dots were projections of points on two slanted planes meeting at a horizontal line bisecting the field of view. The angle formed by the intersecting planes represented either the most distant portion of the display ("center-far") or the closest portion of the display ("center-near"). In the center-far displays, the planes intersected at infinity, so that a point exactly at the intersection would remain stationary on the scope. In the center-near displays, the angle of intersection was the same as the viewing angle, so that a point at the top or bottom of the viewing area would represent an infinite distance, and would remain stationary. (See Braunstein & Andersen, 1981.)

Apparatus. The stimuli were displayed on a Hewlett-Packard Model 1321B X-Y Display Scope with a P31 phosphor, under the control of a PDP-11/44 computer. Each view in the stereo pair was 5.5-cm². The center-to-center separation was 6.5 cm. The resolution of each view in a stereo pair was 1,800 × 1,800. The dot and background brightnesses were approximately .4 and .007 cd/m², respectively. A 1.0 neutral density filter was placed between the display scope and the stereoscope to eliminate any perceptible phosphor trace. The displays were viewed through the same model stereoscope as in Experiment 1, with a lens-to-screen distance of 15.2 cm. Black cloth concealed the face of the display scope, except for the portion visible through the stereoscope. The visual angle subtended by a side of the square viewed through the stereoscope was 22.4°. The fastest dot speed for both the center-near and center-far displays was either 4.35°/sec or 8.70°/sec, depending on the translation-speed condition. The maximum disparities were 241" and 482" of arc, respectively, for the two translation speeds. The required convergence was 3.9°. A sheet of black posterboard, bent to form a dihedral angle, served as a demonstration model during the reading of the instructions. White dots were unsystematically positioned on both sides of the posterboard. The square and cross targets were the same as in Experiment 1.

Procedure. The square targets were used to familiarize the subjects with the viewing apparatus, as in Experiment 1. Each subject was instructed to indicate, immediately after each display disappeared, whether the center dots appeared nearer or more distant by saying "center near" or "center far." The two conditions were demonstrated using the posterboard model. The subjects were run individually in a dark room. Thirty-seven displays (five practice trials and two replications of the 16 stimulus conditions) were presented to each subject. The order of the 16 conditions was randomized separately for each replication and for each subject. The stimulus conditions used in the practice trials for a given subject were the same as the conditions in the last five trials in the session, for that subject. As in Experiment 1, there were two presentations of the cross target. All subjects reported seeing the horizontal and vertical lines as crossed.

Results

The main effects of correspondence versus conflicting stereoscopic information [$F(1,35) = 44.6$] and of angle orientation [$F(1,35) = 20.5$] and the interaction of these two variables [$F(1,35) = 36.7$] were significant ($p < .01$) for all subjects. This interaction is illustrated in Figure 4. Subjects more frequently judged depth order in accordance with stereoscopic information in the conflicting displays when the angle orientation indicated by disparity was center-far. There were also significant interactions of direction of translation with correspondence versus conflicting information [$F(1,35) = 4.19$, $p < .05$] and of these two variables with angle orientation [$F(1,35) = 6.94$, $p < .05$] and with speed [$F(1,35) = 8.87$, $p < .01$].

There was a significant interaction of speed with correspondence for all subjects [$F(1,35) = 10.9$, $p < .01$]. This is illustrated in Figure 5. As expected, the faster speed, which was more effective in conveying depth order in monocular displays (Braunstein & Andersen, 1981), was also more effective in competing with stereoscopic information in the present experiment. The interaction of these two variables with angle orientation was also significant [$F(1,35) = 5.74$, $p < .05$].

A separate analysis was conducted for the special group of subjects who responded in accordance with stereoscopic information on all of the control trials in Experiment 1. The main effect of correspondence versus conflicting stereoscopic information and of angle orientation, and the interaction of these two variables, were also significant for the special group [$F_s(1,18) = 12.4$, 9.9, and 19.4, respectively, $p < .01$ in each case]. There were no other significant main effects or interactions for this group.

As in Experiment 1, the variables previously shown to be most important in determining the accuracy with which depth order was recovered from velocity gradients in monocular displays (Braunstein & Andersen, 1981) were most important in determining the effectiveness of monocular information for depth order when placed in

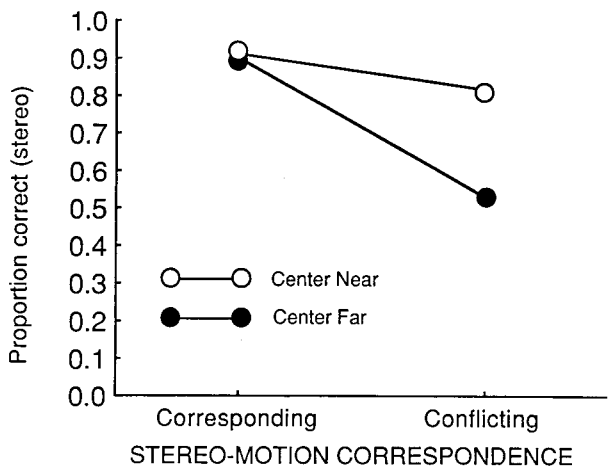


Figure 4. Effect of conflicting monocular information on the proportion of angle-orientation judgments made in accordance with disparity, for center-near and center-far displays.

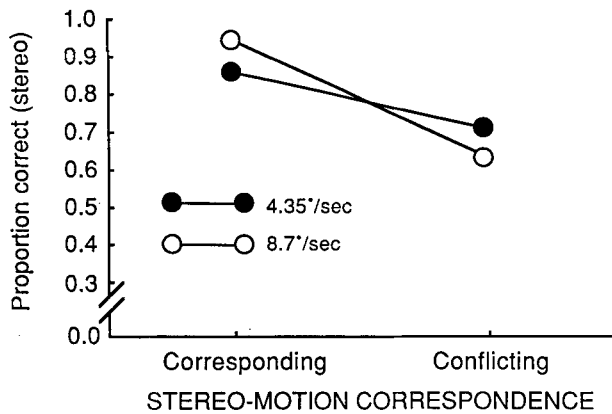


Figure 5. Effect of conflicting monocular information on the proportion of angle-orientation judgments made in accordance with disparity, for the two translation speeds.

conflict with disparity. In the earlier study, maximum dot speed, angle orientation, and the interactions of angle orientation with maximum dot speed and with direction of translation significantly affected depth order judgments. In the present experiment, there were significant interactions between correspondence of monocular and stereoscopic information and each of these effects and interactions, for all subjects. The interaction of information correspondence with angle orientation also was significant when the data were analyzed separately for subjects who responded correctly on all control trials in Experiment 1, but the remaining interactions, involving direction of translation and translation speed, were not significant for this group.

Monocular impact scores were computed for each subject, as in Experiment 1. There was a significant correlation between the impact of conflicting monocular information and performance on the uncrossed version of the Random Dot E test ($r = -.616, p < .01$) for the special group. (For all subjects, this correlation was $-.205, p > .05$.) This result indicates an inverse relationship between the use of conflicting monocular information and performance on a static test of stereoscopic depth perception, among subjects who had responded perfectly in accordance with stereoscopic information on a task that contained no monocular information for depth order. It is a limited result, however, as the correlations between the monocular impact score and two other measures, the crossed Random Dot E and the RANDOT Circles tests, were not significant for this group ($r = -.431$ and $r = -.296$, respectively, $ps > .05$).

Reports of perceived rotational motion, found by Braunstein and Andersen (1981), also were found in this experiment. When the debriefing question "Did you notice any type of motion other than horizontal?" was asked, 27 of the 36 subjects described rotations, generally indicating a vertical axis.

DISCUSSION

We propose that there are two types of processes that can lead to the recovery of shape and depth order. One type of process recovers both shape and depth order from the same information about distances of feature points from the observer, in the sequence proposed by Marr (1982). The other type of process can recover shape from information that leaves depth order ambiguous, unless supplementary information is used to resolve ambiguities in depth order. Disparity, for example, normally provides unambiguous depth order, whereas depth order is ambiguous when shape is recovered by structure-from-motion, unless supplemented by occlusion, polar perspective, or other information. The present experiments included stimuli that provided information for both types of processes. The results indicate that both types of processes remain effective when they provide conflicting information. Almost all subjects made some judgments in accordance with occlusion in Experiment 1 and in accordance with the velocity gradients in Experiment 2 when disparity indicated the opposite depth order. The use of monocular information to resolve depth order in the presence of conflicting disparity information is not contingent on inability to use the stereoscopic depth information. Most of the subjects who exhibited perfect performance on the control trials in Experiment 1, in which there was no monocular information for depth order, sometimes used monocular information in the presence of conflicting stereoscopic information. Although the subjects were not informed that conflicting information for direction of rotation was present in some displays, several reported that they sometimes could switch the perceived direction of rotation at will. All of this suggests that both types of representations are available—one in which viewer-centered depth is recovered from the same information used to recover structure and one in which structure is recovered independently of viewer-centered depth, but to which depth order can be added from other sources. This is compatible with the suggestion, from the study of coordinate systems in long-term memory (Jolicoeur & Kosslyn, 1983), that people develop both types of representations.

The same variables that determined the effectiveness of monocular information for depth order in isolation also determined the relative effectiveness of this information in determining judgments based on depth order with conflicting disparity information. This is not surprising; it is generally true for studies in which conflicting cues are paired (see Epstein, 1975, for a review). The marked difference between the effectiveness of edge and element occlusion when in conflict with disparity, however, may go beyond the matter of effectiveness of the monocular cues in isolation. The edge-occlusion displays exhibited a gradual change in disparity over the image, much like the inverted masks that have been shown to be perceptu-

ally reversible (Yellott & Kaiwi, 1979). On the other hand, the element-occlusion displays exhibited discontinuities in disparity, when near and far elements crossed in the image. Although the present study does not provide a separation of that effect from the greater effectiveness of edge occlusion in isolation, the reversibility of depth based on gradually changing disparity as opposed to discontinuities in disparity should be considered in future research.

Some subjects who performed poorly on static tests of stereoscopic vision were able to respond accurately to changing disparities. Most remarkably, 6 subjects who scored 0 on the RANDOT Circles test responded in accordance with the depth order indicated by disparity on over 90% of the control trials in Experiment 1. The impact of conflicting monocular information on direction-of-rotation judgments in Experiment 1 was not significantly correlated with performance on the static stereoscopic tests, but there was a relationship between performance on uncrossed presentations of the Random Dot E test and the impact of conflicting monocular information on the velocity gradient task in Experiment 2.⁵ It would be premature to reach any conclusions about differences in the relationships of the two tasks to static disparity tests on the basis of this one result, but differences between the effectiveness of continuously changing disparities and constant disparities in determining relative depth judgments should be investigated further.

The present study was concerned primarily with responses to conflicting stereoscopic and motion information. It is more likely, in natural settings, that both sources of information would provide corresponding information about three-dimensional relationships. There have been several computational studies of how disparity and motion may provide mutual constraints for the recovery of three-dimensional structure (e.g., Richards, 1985; Waxman & Duncan, 1985). This combining of sources of information may be especially important when the length of the motion sequence is minimal (Lappin, Doner, & Kottas, 1980) or the display contains noise (Todd, 1985). This would seem to be a worthwhile direction for continued empirical research on the relationships between disparity and structure-from-motion.

REFERENCES

- ANDERSEN, G. J., & BRAUNSTEIN, M. L. (1983). Dynamic occlusion in the perception of rotation in depth. *Perception & Psychophysics*, **34**, 356-362.
- BRAUNSTEIN, M. L. (1966). Sensitivity of the observer to transformations of the visual field. *Journal of Experimental Psychology*, **72**, 683-687.
- BRAUNSTEIN, M. L. (1977). Perceived direction of rotation of simulated three-dimensional patterns. *Perception & Psychophysics*, **21**, 553-557.
- BRAUNSTEIN, M. L. (1983). Perception of rotation in depth: The psychophysical evidence. In Association for Computing Machinery, *Motion: Representation and perception* (Order No. 439830). New York: ACM.
- BRAUNSTEIN, M. L. (in press). The empirical study of structure from motion. In W. N. Martin & J. K. Aggarwal (Eds.), *Motion under-*
- standing: Robot and human vision*. Hingham, MA: Kluwer Academic Publishers.
- BRAUNSTEIN, M. L., & ANDERSEN, G. J. (1981). Velocity gradients and relative depth perception. *Perception & Psychophysics*, **29**, 145-155.
- BRAUNSTEIN, M. L., ANDERSEN, G. J., & RIEFER, D. M. (1982). The use of occlusion to resolve ambiguity in parallel projections. *Perception & Psychophysics*, **31**, 261-267.
- CORMACK, R., & FOX, R. (1985). The computation of disparity and depth in stereograms. *Perception & Psychophysics*, **38**, 375-380.
- EPSTEIN, W. (1968). Modification of the disparity-depth relationship as a result of exposure to conflicting cues. *American Journal of Psychology*, **81**, 189-197.
- EPSTEIN, W. (1975). Recalibration by pairing: A process of perceptual learning. *Perception*, **4**, 59-72.
- ESKRIDGE, J. B. (1976). Accommodation and vergence with the Brewster stereoscope. *Journal of the American Optometric Association*, **47**, 919-924.
- GILLAM, B., FLAGG, T., & FINLAY, D. (1984). Evidence for disparity change as primary stimulus for stereoscopic processing. *Perception & Psychophysics*, **36**, 559-564.
- JOLICOEUR, P., & KOSSLYN, S. M. (1983). Coordinate systems in the long-term memory representation of three-dimensional shapes. *Cognitive Psychology*, **15**, 301-345.
- LAPPIN, J. S., DONER, J. F., & KOTTAS, B. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. *Science*, **209**, 717-719.
- MARR, D. (1982). *Vision*. San Francisco: Freeman.
- MARR, D., & NISHIHARA, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London*, **200**, 269-294.
- PATTERSON, R., & FOX, R. (1984). The effect of testing method on stereoaomaly. *Vision Research*, **24**, 403-408.
- PETERSIK, J. T. (1980). Rotation judgments and depth judgments: Separate or dependent processes. *Perception & Psychophysics*, **27**, 588-590.
- PRAZDNY, K. (1986). *Visual detection of three-dimensional structure from image movements without relevant luminance-domain motion signals*. Manuscript submitted for publication.
- REINECKE, R. D., & SIMONS, K. (1974). A new stereoscopic test for amblyopia screening. *American Journal of Ophthalmology*, **78**, 714-721.
- RICHARDS, W. (1985). Structure from stereo and motion. *Journal of the Optical Society of America*, **A2**, 343-349.
- RICHARDS, W., & LIEBERMAN, H. R. (1985). Correlation between stereo ability and the recovery of structure-from-motion. *American Journal of Optometry & Physiological Optics*, **62**, 111-118.
- ROGERS, B. J., & COLLETT, T. (1985). Rigidity and perceived depth in surfaces specified by disparity and parallax cues. *Investigative Ophthalmology & Visual Science*, **26**(3, Supplement), 242.
- ROGERS, B., & GRAHAM, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, **8**, 125-134.
- SCHWARTZ, B. J., & SPERLING, G. (1983). Nonrigid 3D percepts from 2D representations of rigid objects. *Investigative Ophthalmology & Visual Science*, **24**(3, Supplement), 239.
- SEDGWICK, H. A. (1983). Environment-centered representation of spatial layout: Available visual information from texture and perspective. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 425-458). New York: Academic Press.
- TODD, J. (1985). Perception of structure from motion: Is projective correspondence of moving elements a necessary condition? *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 689-710.
- ULLMAN, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- WALLACH, H., & KARSH, E. B. (1963). Why the modification of stereoscopic depth perception is so rapid. *American Journal of Psychology*, **76**, 413-420.
- WALLACH, H., MOORE, M. E., & DAVIDSON, L. (1963). Modification of stereoscopic depth-perception. *American Journal of Psychology*, **76**, 191-204.

- WALLACH, H., & O'CONNELL, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*, 205-217.
- WAXMAN, A. M., & DUNCAN, J. H. (1985). *Binocular image flows: Steps toward stereo-motion fusion* (Report CAR-TR-119). College Park: University of Maryland, Center for Automation Research.
- YELLOTT, J. I., JR., & KAIWI, J. L. (1979). Depth inversion despite stereopsis: The appearance of random-dot stereograms on surfaces seen in reverse perspective. *Perception*, *8*, 135-142.

NOTES

1. Depth information can be distinguished along other dimensions, in addition to object-centered versus viewer-centered depth. Depth relationships may be specified in relative units (up to a scale factor) or in absolute units. The relationships specified may be metric or nonmetric. (For example, occlusion may provide only ordinal information about viewer-centered depth.) Particular sources of visual information, such as shading, may be related to specific types of information about objects, such as surface orientation and curvature.

2. Prazdny (1986) makes a similar distinction, describing the binocular correlation mechanism as a depth measurement device and the structure-from-motion mechanism as a layout-sensing device.

3. Earlier research considering the effect of inconsistent stereoscopic and kinetic depth information on perceived shape (Epstein, 1968; Wallach & Karsh, 1963) found that stereoscopic depth was recalibrated in the direction indicated by structure-from-motion but that the reverse recalibration did not occur (Wallach, Moore, & Davidson, 1963). This is consistent with the interpretation of structure-from-motion as a direct source of information about object shape.

4. The disparity calculations assume midsagittal fixation and symmetric convergence (Cormack & Fox, 1985). The disparity and convergence calculations are based on Eskridge's (1976) tables for the Brewster-type stereoscope.

5. Richards and Lieberman (1985) have reported a relationship between performance on a structure-from-motion task and crossed disparity. Their results, however, were concerned with judgments of amount of depth and do not relate directly to the depth order tasks in the present study.

(Manuscript received March 31, 1986;
revision accepted for publication July 11, 1986.)