Recovery of structure from motion: Implications for a performance theory based on the structure-from-motion theorem

J. TIMOTHY PETERSIK
Ripon College, Ripon, Wisconsin

The structure-from-motion theorem asserts that three projective views of four elements in motion are sufficient to specify the structure of the object to which the elements belong. Seven experiments were conducted to determine whether the perceptual recovery of structure in human observers is favorable to an interpretation of the theorem as a *performance theory*. In six experiments, subjects either discriminated simulations of objects in rotation around the *y*-axis versus randomly perturbed counterparts or tried to identify them. As the rotation angle between frames of the simulations increased, performance accuracy decreased. Furthermore, performance accuracy did not depend upon the type of projection (parallel versus polar) or upon the coplanarity/noncoplanarity of the stimulus elements. A seventh experiment showed that the subjective appearance of structure breaks down when the angle of rotation between successive frames exceeds about 35.5°. Because the subjects' performance in some conditions exceeded expectations based upon constraints imposed by the structure-from-motion theorem, it is suggested that additional algorithms or heuristic rules might need to be considered when interpreting human recovery of structure in such displays.

Since the pioneering study of the kinetic depth effect by Wallach and O'Connell (1953), it has generally been accepted that under appropriate conditions motion in a two-dimensional (2-D) stimulus array can lead to the perception of structure and depth. Braunstein (1976) summarized much of the work in this area, and Ullman (1979) advanced a computational theory explaining how the visual system might use 2-D transformations to recover a scene's "3-D structure, motion, and change." The theory consists of two major components: the rigidity assumption and the structure-from-motion theorem (SFMT). Regarding the rigidity assumption, Ullman stated that "any set of elements undergoing a two-dimensional transformation which has a unique interpretation as a rigid body moving in space should be interpreted as such a body in motion." The SFMT states that three distinct (parallel-projection) views of four noncoplanar points in a rigid configuration are sufficient (although not always necessary) to determine the structure and motion of a 3-D object. Ullman also provided a mathematical proof of the SFMT, and considered the rigidity assumption and the SFMT together to account for the visual system's "remarkable capacity for correctly interpreting the chang-

ing projections of unrecognizable objects in motion" (p. 130).

As Ullman (1979) noted, considerable empirical support exists for the rigidity assumption (e.g., Johansson, 1978). However, despite the fact that the SFMT has obvious potential application to problems of human vision (as Ullman himself suggested in the text from which the above quotation was taken), virtually no studies to date have directly examined the utility of the SFMT as it applies to human perception (although many papers have referred to it in a post hoc manner). One reason for this lack of human laboratory investigation may be the compelling quality of Ullman's mathematical proof. Another may be that displays incorporating *fewer than* four points have already been shown to lead to percepts of structure in depth (Börjesson & von Hofsten, 1973). Still another reason may be that Ullman presented a number of caveats that appear to restrict the interpretation of potential laboratory tests. For example, Ullman (p. 151) noted that four elements are sufficient to recover structure only if "the input data are given with a high degree of accuracy, and the computation process is similarly accurate." Furthermore, Ullman suggested that the SFMT may not be suited to living organisms: "Other algorithms, that satisfy better the requirement for biological feasibility . . . also exist, but they will not be discussed here" (p. 198). Other limitations on human laboratory investigation are that the SFMT may not guarantee a unique solution for coplanar points "depending upon the initial orientation of the planar object and its subsequent rotations in space" and that the theorem does not apply to perspective transformations without additional theoretical assumptions.

I gratefully acknowledge the assistance of Ramzi Baydoun, Cathy Guarnieri, Lisa Handley, Lynn Kahney, Beth Laufenberg, Jeff Nomoto, and Sherry Schwandt in both discussion of and data collection for the present experiments. The distinction between the SFMT considered as a performance theory rather than as a competence theory was based on a suggestion made by an anonymous reviewer. I thank the reviewer for this helpful point. Requests for reprints should be sent to the author at the Department of Psychology, Ripon College, P.O. Box 248, Ripon, WI 54971.

Thus, the overall psychophysical plausibility of the SFMT as a statement of human perceptual ability remains in question.

Ullman (1979, p. 1) described the theoretical framework within which the SFMT is embedded as a *computational theory* of the process of interpretation of visual motion. As such, the theory is a formal statement of the goals of the computations that underlie the interpretation of visual motion. The computational theory itself is not committed to any specific algorithm or implementation (Marr, 1982). Because computational theories can be implemented by more than one algorithm (and by more than one hardware scheme), the algorithm rather than the computational theory best lends itself to psychophysical investigation (Marr, 1982). Therefore, it is not possible to conduct a test of the SFMT in the sense of potentially falsifying it. In light of this, the SFMT was considered, in the present studies, to be a potential statement of human perceptual abilities. That is, recognizing the validity of the SFMT as a competence theory, it was examined, in the present studies, as a performance theory. My goal was to begin an assessment of its psychophysical plausibility, especially in the interpretation of 2-D simulations of a 3-D object in rigid rotation around the y-axis.

Several considerations entered into the design of the experiments. First, I assumed that the "correspondence problem" (Ullman, 1980) must be solved correctly (i.e., corresponding stimulus elements must be matched correctly from frame to frame) before the correct interpretation of structure from motion could be expected to occur. The average displacement of corresponding elements in a rotation simulation depends to a great extent upon the angle of rotation depicted between successive frames. Given that the closer neighboring elements are in successive frames of an apparent-movement display the more likely they are to be matched, I expected observers to be more likely to recover 3-D structure in a four-element/three-frame display in which the angle of rotation between frames was small rather than large. In the present experiments, the angle of rotation between frames varied between 5° and 60°. A failure to correctly recover structure in displays with large angles of rotation could be due either to a failure in the correspondence process or to a failure in the recovery-of-structure process. In either event, recovery of structure that varied inversely with angle of rotation would be evidence of a limitation in the human perceiver's ability to implement the SFMT.

Other considerations in the design of the experiments derived from the fact that the SFMT is technically restricted to noncoplanar stimulus elements and orthographic (parallel) projections (although Ullman, 1979, noted that the SFMT might apply to coplanar elements under certain conditions and that it might form a foundation for the interpretation of perspective, or polar-projection, simulations). Thus, the present experiments included both planar and nonplanar simulations under conditions of both parallel and polar projection. Superior performance with

planar elements and polar projections would suggest a recovery-of-structure process extending beyond that expected by a performance theory based on the SFMT.

In six of the current experiments, simulations of rigid rotations around the y-axis were presented along with versions of the same displays in which the positions of the stimulus elements were randomly perturbed. The subjects were asked either to discriminate such displays or to identify them. Thus, in Experiments 1-6, if the SFMT is psychophysically plausible, observers in general would be expected to be better able to discriminate and to identify the displays containing rigid transformations. However, because of the limitations inherent in the SFMT or in the human perceiver's implementation of it, one or more of the following might also be expected: (1) the observers' discrimination and identification abilities might depend upon the angle of rotation between successive frames; (2) discrimination and identification might be better for parallel projections than for perspective displays; and (3) discrimination and identification might be better for noncoplanar than for coplanar elements.

EXPERIMENT 1

In Experiment 1 the observers were exposed to pairs of displays, each consisting of three frames containing four noncoplanar elements (luminous dots). One display depicted rigid rotation of the elements around the y-axis (with either parallel or polar projection); in the other display, the same elements were randomly perturbed in each frame by various amounts within a constant range. The subjects were asked to indicate which display contained the rigid rotation.

Method

Subjects. Four experienced psychophysical observers, 1 female and 3 males between 21 and 35 years old, served as subjects. All of the observers had normal or corrected-to-normal vision, and only 1 observer, the author, had prior knowledge of the rationale for the experiment.

In addition to the main observers described above, the experiment was replicated (over fewer trials) with 16 volunteers from an introductory course in psychology.

Apparatus. The stimuli were generated by an Apple II+ microcomputer and displayed on a monochrome Apple Monitor III whose intensity was adjusted to avoid visible persistence. The observers' viewing distance was 130 cm, although head movements were not restricted.¹ The stimuli were displayed within a rectangular area of the monitor screen that measured 280 pixels horizontally $\times$ 180 pixels vertically. The display parameters of the monitor were adjusted so that interpixel spacing in the vertical and horizontal dimensions was equal. Under these conditions, the display area subtended 9.3° $\times$ 5.8° of visual angle. Each pixel subtended approximately 2' vertically and horizontally. A hardware modification of the Apple II+ described by Cavanagh and Anstis (1980) was used to prevent "interleaving" of the stimulus frames that composed each rotation simulation on the raster-scanning monitor.

Stimuli. The observers were presented with rotation simulations generated by a modified version of a program previously described in Petersik (1979). Each rotation simulation began with the random placement of four luminous elements (pixels) in an area of

computer memory that corresponded to the monitor screen. The elements were made noncoplanar by randomly assigning them values in each of the x , y , and z dimensions. The random placement constituted the first frame of the three-frame rotation simulation. Next, the second and third frames of the simulation were created; the exact projections of their elements depended upon the parameters of rotation chosen for a particular experimental condition (see below). This sequence of frames corresponded to a rigid rotation of the initial four elements around the y -axis. The total extent of rotation depended upon the angle of rotation that intervened between successive frames. The routine used to generate the projections of the rotating elements guaranteed that both the "cosine factor" responsible for information regarding rotation and the "perspective factor" responsible for information regarding direction of rotation were represented in the polar-projection displays (Braunstein, 1976). In the parallel-projection displays, only the cosine factor was represented. In the polar-projection displays, the simulated projection distance was 32 cm. This yielded a perspective ratio of 2.00 (Braunstein, 1976).² In the parallel-projection displays, the simulated projection distance was 10,000 cm, yielding a perspective ratio of 1.00. In all other respects, the displays were created according to the descriptions provided by Braunstein.

After the simulation of a rigidly rotating object was created, a new display was produced by randomly perturbing the location of all of the elements in each of the original three frames by a non-zero value no greater than 5 pixels in either the horizontal or vertical dimension. Their perturbations could be positive (e.g., up or to the right) or negative (e.g., down or to the left), yielding a perturbation range in each dimension of 10 pixels. Thus, the second set of frames depicted motion of the elements that was the sum of the original transformations and the perturbations added to each frame. For each set of frames, a check was performed so that no element was ever transformed beyond the display boundary described above.

The rotation angle was also varied between successive frames of a display. Rotation angles were 5°, 10°, 20°, 40°, or 60°. When displayed, each element of a frame had a luminance of 2–9 cd/m². Each frame was presented for 128 msec with a 16-msec interval between frames.

Procedure. Each practiced observer served in 40 replications of a block of 10 trials determined by the factorial combination of the two projection types and the five angles of rotation. Each experimental session consisted of no more than two blocks of trials, and the sessions were separated by no less than 4 hr. Each inexperienced observer served in three blocks of trials in one experimental session.

The experiment was conducted in a room diffusely illuminated by a 40-W incandescent bulb placed 60 cm above and just behind the display screen. There was no visible glare on the screen. Prior to the beginning of a block of trials, the order of the 10 experimental conditions was randomly determined by and stored in the computer. Each experimental session began with the subject's instructions printed on the screen. After reading the instructions, each subject gazed directly at the center of the display area. A trial consisted of the successive presentation of two three-frame displays, one depicting a rigid rotation and the other its corresponding perturbed display. The order of these presentations was randomly determined by the computer. A warning buzzer preceded each set of frames by 2 sec, and there was a total of 4 sec between the two sets of a trial. Following the presentation of the two sets of frames, a message appeared on the screen asking the subject to indicate (by pressing an appropriate key) whether the first or the second display had shown the rigidly rotating "object." The subject's responses were automatically stored on disk. There was a 5–10 sec interval between trials while the computer prepared the next set of displays.

The practiced observers participated in one additional experimental

Figure 1. Mean correct discrimination of object rotation as a function of rotation angle. These data were obtained when one stimulus of a pair was perturbed within a constant range in Experiment 1. Circles represent data from practiced observers; squares, data from naive observers. Filled symbols represent parallel projections; open symbols, polar projections.

session in which, instead of discriminating the rigidly rotating object, they reported their subjective impressions of all of the displays.

Results and Discussion

The percentages of correct discriminations over all conditions of the experiment are shown in Figure 1 for both the practiced and the naive observers. Both groups of observers experienced a monotonic decrease in the ability to discriminate rigidly rotating collections of elements from perturbed collections as the angle of rotation depicted between frames increased, although even at the greatest angle of rotation observers performed better than chance. A completely repeated measures analysis of variance conducted on the data from the practiced observers revealed the effect of the angle of rotation to be significant [$F(4,12) = 30.60$, $p < .001$]. On the other hand, the observers' ability to correctly recover structure from motion was independent of the type of projection used [$F(1,4) = 0.21$, $p > .05$]. Finally, there was no interaction between the factors [$F(4,12) = .86$, $p > .05$]. Standard deviations for the practiced observers' data points ranged from 3.5% (parallel projection, 5° rotation angle) to 14.8% (polar projection, 20° rotation angle); the variation was not linearly related to rotation angle for polar projections, but was an increasing function of rotation angle for parallel projections.

Because of the small number of replications obtained with each naive subject, no statistical analyses were conducted on their responses. Their results are shown in Figure 1 for comparison. For present purposes their results are important in demonstrating a practice effect (relative to the primary observers) for the kinds of judgments required by this experiment, a finding similar to that of Petersik (1978).

Thus, the overall psychophysical plausibility of the SFMT as a statement of human perceptual ability remains in question.

Ullman (1979, p. 1) described the theoretical framework within which the SFMT is embedded as a *computational theory* of the process of interpretation of visual motion. As such, the theory is a formal statement of the goals of the computations that underlie the interpretation of visual motion. The computational theory itself is not committed to any specific algorithm or implementation (Marr, 1982). Because computational theories can be implemented by more than one algorithm (and by more than one hardware scheme), the algorithm rather than the computational theory best lends itself to psychophysical investigation (Marr, 1982). Therefore, it is not possible to conduct a test of the SFMT in the sense of potentially falsifying it. In light of this, the SFMT was considered, in the present studies, to be a potential statement of human perceptual abilities. That is, recognizing the validity of the SFMT as a competence theory, it was examined, in the present studies, as a performance theory. My goal was to begin an assessment of its psychophysical plausibility, especially in the interpretation of 2-D simulations of a 3-D object in rigid rotation around the y-axis.

Several considerations entered into the design of the experiments. First, I assumed that the "correspondence problem" (Ullman, 1980) must be solved correctly (i.e., corresponding stimulus elements must be matched correctly from frame to frame) before the correct interpretation of structure from motion could be expected to occur. The average displacement of corresponding elements in a rotation simulation depends to a great extent upon the angle of rotation depicted between successive frames. Given that the closer neighboring elements are in successive frames of an apparent-movement display the more likely they are to be matched, I expected observers to be more likely to recover 3-D structure in a four-element/three-frame display in which the angle of rotation between frames was small rather than large. In the present experiments, the angle of rotation between frames varied between 5° and 60°. A failure to correctly recover structure in displays with large angles of rotation could be due either to a failure in the correspondence process or to a failure in the recovery-of-structure process. In either event, recovery of structure that varied inversely with angle of rotation would be evidence of a limitation in the human perceiver's ability to implement the SFMT.

Other considerations in the design of the experiments derived from the fact that the SFMT is technically restricted to noncoplanar stimulus elements and orthographic (parallel) projections (although Ullman, 1979, noted that the SFMT might apply to coplanar elements under certain conditions and that it might form a foundation for the interpretation of perspective, or polar-projection, simulations). Thus, the present experiments included both planar and nonplanar simulations under conditions of both parallel and polar projection. Superior performance with

planar elements and polar projections would suggest a recovery-of-structure process extending beyond that expected by a performance theory based on the SFMT.

In six of the current experiments, simulations of rigid rotations around the y-axis were presented along with versions of the same displays in which the positions of the stimulus elements were randomly perturbed. The subjects were asked either to discriminate such displays or to identify them. Thus, in Experiments 1-6, if the SFMT is psychophysically plausible, observers in general would be expected to be better able to discriminate and to identify the displays containing rigid transformations. However, because of the limitations inherent in the SFMT or in the human perceiver's implementation of it, one or more of the following might also be expected: (1) the observers' discrimination and identification abilities might depend upon the angle of rotation between successive frames; (2) discrimination and identification might be better for parallel projections than for perspective displays; and (3) discrimination and identification might be better for noncoplanar than for coplanar elements.

EXPERIMENT 1

In Experiment 1 the observers were exposed to pairs of displays, each consisting of three frames containing four noncoplanar elements (luminous dots). One display depicted rigid rotation of the elements around the y-axis (with either parallel or polar projection); in the other display, the same elements were randomly perturbed in each frame by various amounts within a constant range. The subjects were asked to indicate which display contained the rigid rotation.

Method

Subjects. Four experienced psychophysical observers, 1 female and 3 males between 21 and 35 years old, served as subjects. All of the observers had normal or corrected-to-normal vision, and only 1 observer, the author, had prior knowledge of the rationale for the experiment.

In addition to the main observers described above, the experiment was replicated (over fewer trials) with 16 volunteers from an introductory course in psychology.

Apparatus. The stimuli were generated by an Apple II+ microcomputer and displayed on a monochrome Apple Monitor III whose intensity was adjusted to avoid visible persistence. The observers' viewing distance was 130 cm, although head movements were not restricted.¹ The stimuli were displayed within a rectangular area of the monitor screen that measured 280 pixels horizontally $\times$ 180 pixels vertically. The display parameters of the monitor were adjusted so that interpixel spacing in the vertical and horizontal dimensions was equal. Under these conditions, the display area subtended 9.3° $\times$ 5.8° of visual angle. Each pixel subtended approximately 2' vertically and horizontally. A hardware modification of the Apple II+ described by Cavanagh and Anstis (1980) was used to prevent "interleaving" of the stimulus frames that composed each rotation simulation on the raster-scanning monitor.

Stimuli. The observers were presented with rotation simulations generated by a modified version of a program previously described in Petersik (1979). Each rotation simulation began with the random placement of four luminous elements (pixels) in an area of

computer memory that corresponded to the monitor screen. The elements were made noncoplanar by randomly assigning them values in each of the x , y , and z dimensions. The random placement constituted the first frame of the three-frame rotation simulation. Next, the second and third frames of the simulation were created; the exact projections of their elements depended upon the parameters of rotation chosen for a particular experimental condition (see below). This sequence of frames corresponded to a rigid rotation of the initial four elements around the y -axis. The total extent of rotation depended upon the angle of rotation that intervened between successive frames. The routine used to generate the projections of the rotating elements guaranteed that both the "cosine factor" responsible for information regarding rotation and the "perspective factor" responsible for information regarding direction of rotation were represented in the polar-projection displays (Braunstein, 1976). In the parallel-projection displays, only the cosine factor was represented. In the polar-projection displays, the simulated projection distance was 32 cm.¹ This yielded a perspective ratio of 2.00 (Braunstein, 1976).² In the parallel-projection displays, the simulated projection distance was 10,000 cm, yielding a perspective ratio of 1.00. In all other respects, the displays were created according to the descriptions provided by Braunstein.

After the simulation of a rigidly rotating object was created, a new display was produced by randomly perturbing the location of all of the elements in each of the original three frames by a non-zero value no greater than 5 pixels in either the horizontal or vertical dimension. Their perturbations could be positive (e.g., up or to the right) or negative (e.g., down or to the left), yielding a perturbation range in each dimension of 10 pixels. Thus, the second set of frames depicted motion of the elements that was the sum of the original transformations and the perturbations added to each frame. For each set of frames, a check was performed so that no element was ever transformed beyond the display boundary described above.

The rotation angle was also varied between successive frames of a display. Rotation angles were 5°, 10°, 20°, 40°, or 60°. When displayed, each element of a frame had a luminance of 2–9 cd/m². Each frame was presented for 128 msec with a 16-msec interval between frames.

Procedure. Each practiced observer served in 40 replications of a block of 10 trials determined by the factorial combination of the two projection types and the five angles of rotation. Each experimental session consisted of no more than two blocks of trials, and the sessions were separated by no less than 4 hr. Each inexperienced observer served in three blocks of trials in one experimental session.

The experiment was conducted in a room diffusely illuminated by a 40-W incandescent bulb placed 60 cm above and just behind the display screen. There was no visible glare on the screen. Prior to the beginning of a block of trials, the order of the 10 experimental conditions was randomly determined by and stored in the computer. Each experimental session began with the subject's instructions printed on the screen. After reading the instructions, each subject gazed directly at the center of the display area. A trial consisted of the successive presentation of two three-frame displays, one depicting a rigid rotation and the other its corresponding perturbed display. The order of these presentations was randomly determined by the computer. A warning buzzer preceded each set of frames by 2 sec, and there was a total of 4 sec between the two sets of a trial. Following the presentation of the two sets of frames, a message appeared on the screen asking the subject to indicate (by pressing an appropriate key) whether the first or the second display had shown the rigidly rotating "object." The subject's responses were automatically stored on disk. There was a 5–10 sec interval between trials while the computer prepared the next set of displays.

The practiced observers participated in one additional experimental

Figure 1. Mean correct discrimination of object rotation as a function of rotation angle. These data were obtained when one stimulus of a pair was perturbed within a constant range in Experiment 1. Circles represent data from practiced observers; squares, data from naive observers. Filled symbols represent parallel projections; open symbols, polar projections.

session in which, instead of discriminating the rigidly rotating object, they reported their subjective impressions of all of the displays.

Results and Discussion

The percentages of correct discriminations over all conditions of the experiment are shown in Figure 1 for both the practiced and the naive observers. Both groups of observers experienced a monotonic decrease in the ability to discriminate rigidly rotating collections of elements from perturbed collections as the angle of rotation depicted between frames increased, although even at the greatest angle of rotation observers performed better than chance. A completely repeated measures analysis of variance conducted on the data from the practiced observers revealed the effect of the angle of rotation to be significant [$F(4,12) = 30.60$, $p < .001$]. On the other hand, the observers' ability to correctly recover structure from motion was independent of the type of projection used [$F(1,4) = 0.21$, $p > .05$]. Finally, there was no interaction between the factors [$F(4,12) = .86$, $p > .05$]. Standard deviations for the practiced observers' data points ranged from 3.5% (parallel projection, 5° rotation angle) to 14.8% (polar projection, 20° rotation angle); the variation was not linearly related to rotation angle for polar projections, but was an increasing function of rotation angle for parallel projections.

Because of the small number of replications obtained with each naive subject, no statistical analyses were conducted on their responses. Their results are shown in Figure 1 for comparison. For present purposes their results are important in demonstrating a practice effect (relative to the primary observers) for the kinds of judgments required by this experiment, a finding similar to that of Petersik (1978).

The decrease in recovery-of-structure ability with increasing rotation angle constrains the SFMT as a performance theory if the theory is interpreted to mean that *any* three views of rigid motion in depth should yield equivalent recovery of structure. The finding is consistent with the conjecture that the correspondence problem is not easily solved when elements in successive frames are separated by large 2-D distances and that therefore the recovery of structure would be less efficient. The absence of a difference in recovery of structure between parallel and polar projections is not consistent with a performance theory based on the SFMT alone (since the SFMT is limited to parallel projections). Ullman (1979) argued that the SFMT can be used to recover structure in perspective transformations if it is simultaneously applied to a number of local parts of the scene and the outputs of such local computations are then combined. However, this use of the SFMT requires that four points be visible in *each* of the local areas of the scene examined. Therefore, the good recovery of structure obtained with the polar projections in this experiment is unaccounted for by the SFMT alone. This finding does suggest the existence of a psychophysical interpretation of structure based upon three successive views of four points in perspective rotation. Such an extension is beyond the SFMT and may be accounted for by some other algorithm—for instance, one based upon the apparent-velocity field in these displays, or some heuristic rule (e.g., Braunstein, 1983).

A linear-regression analysis was performed on the practiced observers' data shown in Figure 1, averaged over projection type. The equation relating the percentage of correct responses to angle of rotation was $y' = 98.4\% - .60(x)$. The linear relationship was significant [$F(1,3) = 115.2, p < .05; r^2 = .98$]. In an attempt to account for the decrease in accuracy with greater rotation angles, the subjects were questioned carefully about their perceptual experiences while viewing several displays. A common theme that emerged was that coherent percepts of depth and structure were obtainable and useful in making judgments only at the smaller rotation angles (in the 5°–40° range). Previous experiments (Petersik, 1978, 1980a) had shown that rated depth judgments do not decrease with increasing rotation angles up to 20° or with increasing stimulus onset asynchronies up to about 100 msec. However, these earlier experiments included more stimulus elements per frame and rotation through a complete 360° circuit and are therefore not directly comparable to the conditions of the present experiment. The earlier experiments did show that when spatiotemporal conditions produced perceptually discontinuous ("choppy") displays, depth ratings tended toward two-dimensionality. This interpretation was consistent with that of the subjects in this experiment when they viewed the four-element/three-frame displays at the larger rotation angles.

The perceptual process that underlies the subjective experience of coherent structure in depth apparently was not able to operate with large rotation angles, and therefore large 2-D displacements, in the present experiment.

Nonetheless, the greater-than-chance accuracy shown here might suggest that even at large rotation angles structure from motion information was available at some level for the purpose of discrimination. The subjects reported that at the large rotation angles they frequently made negative judgments, that is, they ruled out the display that appeared to be "more random." Indeed, even with the small rotation angles, the subjects reported the subjective experience of structure in the perturbed displays; however, the structure of the perturbed displays was generally not as "good" as that of the unperturbed displays. Under any circumstance, the subjects were able to make discriminations by applying some rule regarding the appearance of the displays containing unperturbed structure.

The decrease in recovery of structure found with increasing rotation angles, along with the subjects' subjective reports, suggests that the constant perturbation used had unequal effects on the five rotation-angle conditions employed. That is, at small rotation angles the changes in the 2-D projection of elements from frame to frame are also small, and therefore a perturbation of ± 5 pixels is comparatively large and noticeable. However, the same perturbation seems to have been much less noticeable at the large rotation angles, in which the normal 2-D projections changed considerably more. In fact, the perturbation of ± 5 pixels is comparable to the position changes of elements in a three-frame simulation involving only 10°–11° of rotation per frame.³ If the notion that the recovery of structure is limited by the size of the 2-D projection changes from frame to frame is true, then eliminating the rotation-angle effect found in Experiment 1 ought to be possible by scaling the perturbations used in each simulation by the angle of rotation depicted. This logic formed the basis of Experiment 2.

EXPERIMENT 2

A number of different ways exist for scaling perturbations to the extent of change represented from frame to frame of a rotation simulation. Pilot work showed that a convenient scaling factor was simply to base the 2-D perturbation on a percentage of the angular change shown in any given display. Such a factor produces a relatively constant discrimination performance for all angles of rotation.

Method

Subjects. The 4 subjects who participated in Experiment 1 served in Experiment 2. Again, 3 of them were psychophysically experienced observers but naive as to the purpose and design of the experiment. As in Experiment 1, Experiment 2 was replicated with 16 naive subjects who participated in 3 trials for each condition.

Apparatus and Stimuli. The apparatus and viewing conditions were the same as those described for Experiment 1. The stimuli were generated and presented in the same way but with the following exceptions: (1) The 2-D perturbation range used for one member of each pair of displays was 20% of the angle of rotation depicted between frames in the nonperturbed display. For example, the perturbations corresponding to a display with 30° rotations were ± 6

Figure 2. Mean correct discrimination as a function of rotation angle. These data were obtained when one stimulus was perturbed by a constant percentage of its rotation angle in Experiment 2. Circles represent data from practiced observers; squares, data from naive observers. Filled symbols represent parallel projections; open symbols, polar projections.

pixels. Again, all perturbations were nonzero. (2) The rotation angles used in the present experiment were 10°, 20°, 30°, 40°, and 50°, because of the tendency for perturbations to be highly noticeable at small rotation angles and less noticeable at larger ones.

Procedure. The procedure employed was the same as for Experiment 1.

Results and Discussion

Figure 2 shows the percentages of correct discriminations of displays depicting rigidly rotating objects versus perturbed displays, both for practiced and for naive subjects. The major finding of this experiment was that the scaled perturbations eliminated the effect of the angle of rotation seen in Figure 1 [$F(4,12) = 1.49, p > .05$]. That is, displays showing successively larger rotations between frames, and therefore correspondingly larger 2-D transformations, required greater absolute levels of perturbation to maintain a relatively constant degree of performance. (Correct discrimination ranged from 67.8% to 81.6% with an overall mean of 72.8%.) These results are consistent with the notion that the correspondence process underlying the recovery of structure from motion is not as finely tuned, or at least is not as efficient, at the larger rotation angles.⁴ The results further suggest that the process that yields structure from motion takes, as its input, information derived about the motion of individual display elements from the correspondence process.

Although the data shown in Figure 2 suggest a superiority of parallel projections over polar projections, as might be consistent with a performance theory based on the SFMT, this effect was not statistically significant [$F(1,3) = 5.87, p > .05$]. Furthermore, there was no evidence of an interaction between rotation angle and projection type [$F(4,12) = .48, p > .05$].

Once again, during self-report, the subjects stated that both the perturbed and the rigid displays yielded kinetic

depth effects, which declined at the larger rotation angles. Furthermore, the subjects claimed to be able to "see" which of the two displays depicted a more rigidly rotating object. To determine the importance of having the perturbed and unperturbed displays appearing in close succession for comparison, Experiments 3 and 4 were designed so that only one display, either an unperturbed or a perturbed rotation simulation, was shown in each trial.

EXPERIMENTS 3 AND 4

Psychophysically, the SFMT could be interpreted to be an all-or-none process. That is, the output of the structure-from-motion computations could yield a binary "yes/no" judgment regarding whether any particular motion display represents a rigidly moving object. If this interpretation is true, the observers may not require the type of comparisons afforded by the procedure of Experiments 1 and 2. Therefore, Experiments 3 and 4 were simply replications of Experiments 1 and 2, except that an identification procedure was used. On any given trial, the computer chose randomly to present either an unperturbed or a perturbed simulation in accord with randomly selected angular-rotation and projection-type parameters. Following the single display, the subject was to report whether the object depicted was or was not rigidly rotating. If the two types of displays need not be compared to provide an answer regarding the occurrence of rigid structure (i.e., if the "negative judgments" discussed above are not necessary), we would expect the results of these experiments to be comparable to those of the first two.

Method

Subjects. The same 4 subjects who served in Experiments 1 and 2 also served in the present experiments. Again, 3 subjects were psychophysically experienced, but naive as to the rationale of the experiments. No naive subjects were tested.

Apparatus and Stimuli. The stimuli were generated with the same apparatus and in the same manner as that described for Experiments 1 and 2, except that on any given trial the computer chose randomly to display either the perturbed or the unperturbed version of a display. In Experiment 3, the parameters and degree of constant perturbation used in Experiment 1 were again employed. The parameters and scaled perturbations used in Experiment 2 were used in Experiment 4. The viewing conditions were the same as those described previously.

Procedure. The instructions displayed to the subjects were changed to reflect the new identification task. In the course of an experimental session, each display was followed by a request for the subject to indicate whether or not the display had depicted a rigidly rotating object by pressing the appropriate key on the computer keyboard. If uncertain, the subjects were to guess. Each subject served in 30 replications of displays depicting each combination of experimental conditions.

Results and Discussion

The upper panel of Figure 3 shows the results of Experiment 3; the lower panel shows the results from Experiment 4. As can be seen by comparing the upper panel of Figure 3 with Figure 1, the subjects showed the same

Figure 3. Mean correct identification of object motion and nonobject motion as a function of rotation angle. Upper panel: data from Experiment 3 (constant perturbation range); lower panel: data from Experiment 4 (scaled perturbation range). Filled symbols represent parallel projections; open symbols, polar projections.

pattern of results in Experiment 3 (constant range of perturbation) as they did in Experiment 1. However, in this case their identification performance dropped by about 10% as compared with their discrimination performance in Experiment 1. Nonetheless, only at a rotation angle of 60° did the subjects' average performance (± 1 SD) extend to the chance level (50%). Therefore, a safe conclusion seems to be that at the first four rotation angles the subjects were fairly accurate in recovering structure from motion, and that their overall performance is nearly as good for identification as for discrimination (when comparisons are available).

As in Experiment 1, a repeated measures analysis of variance of the data of Experiment 3 showed a significant effect of rotation angle [$F(4,12) = 17.44, p < .001$]. Again, there was no effect of projection type [$F(1,3) < 1.00$] or interaction of rotation angle with projection type [$F(4,12) = 1.83, p > .05$].

A linear-regression analysis of the average data shown in the top panel of Figure 3 showed a significant linear component [$F(1,3) = 17.54, p < .001; r^2 = .854$]. The regression equation was $y' = 87.8\% - 0.61(x)$. The slope of this function is virtually identical to that found in Experiment 1 (.60), further suggesting that a structure-from-motion process, rather than a form of comparison of the displays, accounts for the data in both experiments.

The bottom panel of Figure 3 shows the results of Experiment 4. As was the case in Experiment 2, no significant effect of rotation angle occurred when the perturbation factor was 20% of the rotation angle [$F(4,12) = 1.34, p > .05$]. Again, there was no effect of projection type or interaction [$F(1,3) < 1.00$ and $F(4,12) = 1.06, p < .05$, respectively]. Once again, the identification task

reduced the subjects' performance relative to the discrimination task by about 10%.

The similarity between the results of Experiments 3 and 4 and those of Experiments 1 and 2 suggests that a similar process was responsible for the results in all four. For the four-element/three-frame displays used here, the structure-from-motion process alone resulted in an average identification ability of 72.8% with the constant perturbation range; this dropped to 63.1% with the scaled perturbations. Because the corresponding averages for the discrimination tasks were 82% and 71%, respectively, the availability of comparison contributes only about 10% to the discrimination ability. Thus, for the conditions used in these experiments, the structure-from-motion process alone produces 63%–71% correct judgments. Although this seems to be a fairly low percentage of correct judgments, it may be expected under the SFMT, because Ullman (1979, p. 151) stated "a less accurate device (such as, perhaps, the human visual system) might require more elements (or more views) for a reliable and accurate interpretation."

EXPERIMENTS 5 AND 6

The SFMT guarantees the recovery of structure only from parallel projections of noncoplanar elements. Experiments 1–4 have shown that the human's ability to recover structure is not a function of polar versus parallel projection. Experiments 5 and 6 extend the analysis of a performance theory based upon the SFMT to coplanar elements. Despite the apparent limitation of the SFMT to noncoplanar elements, Ullman (1979, p. 173) noted that, for coplanar elements initial orientation and subsequent rotations about the vertical axis are ambiguous, although coplanarity itself can be established. Coplanarity itself, then, could possibly be used as information about structure. In Experiments 5 and 6, the subjects were once again asked to discriminate between rigidly rotating and perturbed displays containing coplanar elements. The subjects were not asked to indicate the orientation of a perceived plane or the direction of its rotation, but rather whether or not it was rotating rigidly.

Method

Subjects. The 4 subjects from Experiments 1–4 served in the present experiments. Once again, 3 subjects were naive as to the rationale of the experiment and the design of the displays.

Apparatus and Stimuli. The apparatus and viewing conditions were the same as previously described. The stimuli were generated in the same manner, except that the elements of all displays were initially assigned a constant value on the z-axis: They were projected to be in the plane of the display monitor. For the rigidly rotating displays, the elements were always maintained in coplanarity. For the perturbed displays, the projected positions of the elements were determined in the same way as previously (i.e., their initial projections, given coplanarity, were determined, and then both x and y perturbations were added). With the exception of coplanarity, Experiment 5 was a replication of Experiment 1: rotation-angle and projection parameters were the same and perturbations were ± 5 pixels. Experiment 6 was similarly a replica-

Figure 4. Mean correct discrimination of object rotation as a function of rotation angle. Data obtained when stimuli were planar. Upper panel: data from Experiment 5; lower panel: data from Experiment 6. Filled symbols represent parallel projections; open symbols, polar projections.

tion of Experiment 2: parameters were the same and perturbations were scaled at 20% of the rotation angle. Adjustments to displays were made whenever the projection of a display led to a frame in which all elements were colinear, because this would be a strong cue to the rigidly rotating plane. When such a display occurred, the starting point of the simulation was changed by 4° to eliminate the colinearity.

Procedure. The instructions to the subjects and the procedure were the same as for Experiment 1. The subjects were not told that the displays would depict planar objects or coplanar elements. Once again, their task was simply to discriminate the display containing a rigidly rotating object. The experiments differed only in having constant or proportional perturbations and in having slightly different angles of rotation.

Results and Discussion

The upper panel of Figure 4 shows the discrimination results for the constant range of perturbation (Experiment 5); the lower panel shows them for the proportional perturbations (Experiment 6). An examination of the figures shows first that the pattern of results is similar to that shown in Figures 1 and 2 (and also in Figure 3). Second, the subjects were apparently better able to recover structure from the coplanar elements than from the noncoplanar elements. The subjects viewing coplanar elements in Experiment 5 improved their discrimination performance an average of 6.7% as compared with discrimination performance with the noncoplanar displays of Experiment 1. In Experiment 6, coplanar elements improved performance an average of 11.6% as compared with performance in Experiment 2. Although the SFMT cannot provide for recovery of structure from coplanar elements, it does assure recovery of their coplanarity. This, along with a heuristic rule that asserts that coplanar elements should belong to the same object, could account for the performance of the subjects in Experiments 5 and 6.

As in previous experiments with a constant range of perturbation (i.e., Experiments 1 and 3), the data of Experiment 5 yielded a significant effect of rotation angle [$F(4,12) = 10.08, p < .001$]. Again, neither the projection type nor the interaction proved significant [$F(1,3) = 2.55, p > .05$, and $F(4,12) = 1.93, p > .05$, respectively]. The average data of the top panel of Figure 4 were subjected to a linear-regression analysis, which showed a significant linear trend over rotation angle [$F(1,3) = 178.46, p < .001; r^2 = .98$]. The regression equation was $y' = 102\% - 0.49(x)$. The slope in this case is quite a bit shallower than was found for the noncoplanar elements, which again suggests that the process of recovering structure from motion was more accurate with the coplanar displays.

As in previous experiments with a scaled perturbation range (i.e., Experiments 2 and 4), the data of Experiment 6 failed to show significance for either of the main effects or for the interaction (in each case, $F < 1.00$).

EXPERIMENT 7

The results of Experiments 1–6 suggest that the SFMT can serve well as a performance theory if certain additional heuristic rules or algorithms are assumed to account for recovery of structure with coplanar elements and with polar projections. However, a difficulty arises in accounting for the finding that discrimination and identification accuracy decrease with increasing rotation angle (see also, Donner, Lappin, & Perfetto, 1984). A failure to recover structure with a large rotation angle could be due to a failure in the correspondence process that logically precedes recovery of structure, in which case the failure would not be evidence against a performance theory based upon the SFMT. Alternatively, such a failure in recovering structure could be due to a breakdown in the process responsible for the recovery of structure itself. Such a case would be evidence against the applicability of the SFMT to perceptual performance, at least at large rotation angles.

To address this issue, the goal of Experiment 7 was to determine whether there was some consistent rotation angle (and therefore some corresponding 2-D displacement) beyond which subjects failed to acquire a compelling subjective impression of structure from motion. My reasoning was that if such a consistent displacement did exist, it might be related to what was already known about the spatial-integration limits of the correspondence process.

Method

Subjects. Five subjects, one of whom (the author) had served in Experiments 1–6, participated in Experiment 7. Four of the subjects were female, 1 was male; all were between 20 and 35 years old. All had normal or corrected-to-normal vision. Each of the 4 naive subjects participated in 10 practice sessions, using the program that had been used to run Experiment 1, prior to serving in the experimental sessions of this study. The practice sessions served to familiarize the subjects with the kinds of stimuli that would be observed as well as with the use of the computer.

Stimuli and Apparatus. The apparatus and the testing environment were the same as in all previous experiments. The program used to generate and display the stimuli was modified to present a series of displays (trials) that depicted unperturbed polar projections of a rotating object. The simulated projection distance was 32 cm. As in previous experiments, each display consisted of four elements shown in three frames. The initial position of the elements was randomly determined in each trial so that expectations regarding the appearance of the "rotating object" could not influence subsequent trials in a series. Display parameters were the same as in previous experiments, except that the program either incremented or decremented the angle of rotation between frames by 3° from trial to trial, depending upon whether an ascending or a descending series was being run.

Procedure. The method of ascending and descending limits was used to determine the rotation angle at which subjects experienced a transition between subjectively experiencing and not experiencing well-formed kinetic-depth effects. Each subject participated individually and in no more than 10 series in a single session. A session began with the subject specifying whether an ascending or descending series (randomly chosen) was to be run. For an ascending series, the first display observed had a rotation angle of 3° between successive frames; for a descending series the angle was 60°. Following a trial, the subject indicated by pressing the appropriate key on the computer whether or not the display gave the *appearance* of a coherently rotating object. The subject was instructed to base his/her answers on perceptual experience rather than on inference. Following the subject's answer, new locations for the four display elements were determined by the computer and the angle of rotation was incremented or decremented appropriately. A warning buzzer prepared the subject for the next trial in the series. The interval between the subject's answer and the subsequent display was approximately 10 sec. A series was terminated after the subject made a transition from the category of his or her initial judgment to three successive judgments of the opposite type. All responses were permanently recorded by the computer. Twenty ascending and 20 descending series were run for each subject. The series themselves were run in the *a-d-d-a* order.

Results and Discussion

All subjects reported seeing coherent object rotation in the initial display (3° rotation angle) of each ascending series. The subjects reported random or incoherent motion in the initial display (60° rotation angle) of each descending series in 95% of all trials. (They always reported incoherent motion on one of the first *two* displays.) Table 1 shows the ascending, descending, and average transition angle, along with standard deviations,

for each subject. The bottom row of Table 1 shows the transition angles averaged over all subjects, along with the associated standard deviations. As can be seen in the table, the transition angles obtained from the 5 subjects were in good agreement for both ascending and descending series as well as for the average of both. The overall transition angle below which subjects tended to report seeing objects in rotation and above which they did not was 35.5°. This value corresponds to a horizontal 2-D motion of elements along the midline of the display of about 33' of visual angle. Because elements higher and lower than the midline and toward the edges of the display undergo less 2-D motion with a given angle of rotation, one can reasonably hypothesize that the matching of corresponding elements from frame to frame is mediated by the short-range process in apparent movement (Braddick, 1974; Petersik & Pantle, 1979). Although the short-range process is frequently said to operate over 2-D distances no greater than 15' of arc, recent evidence suggests that its limits might be much greater under certain conditions (Baker & Braddick, 1982; Petersik, Pufahl, & Krasnoff, 1983). Furthermore, there is evidence that local matches made within the displacement limit of the short-range process can influence neighboring matches over still greater distances (Petersik, 1987). Thus, matches accomplished by the short-range process at some locations of the 2-D projection of a rotating object may possibly determine matches of other elements over greater distances. The obtained transition angle of 35.5° (33') is consistent with this interpretation. The results of the present experiment thus suggest that the decrease in discrimination and identification accuracy obtained in Experiments 1-6 was due to a gradual decrease in the ability of the short-range process to solve the correspondence problem over increasing rotation angles rather than to an inherent limit in the structure-from-motion process over increasing rotation angles.

GENERAL DISCUSSION

The most general finding to emerge from the present series of experiments is that human observers can recover structure in displays consisting of three views of four elements in rotation with about 60%-80% accuracy, depending upon the type of simulation and the observer's task.⁵ Furthermore, the process of comparing a perturbed display to a rigid rotation increases the probability of correctly identifying a rigid rotation by about 10%. Although the overall accuracy might not be as high as one might expect under a performance theory based on the SFMT, it is reasonable given the "precision of computation" constraints determined by Ullman (1979). The finding that recovery of structure varies with the angle of rotation between frames can be accounted for by assuming a variation in the ability of the short-range apparent-movement process to solve the correspondence problem as rotation angle increases.

Table 1
Ascending, Descending, and Average Transition Points (Tr),
with Standard Deviations (SD), between Rotation
in Depth and Incoherent Motion for All Subjects

Subject	Series Type					
	Ascending		Descending		Average	
	Tr	SD	Tr	SD	Tr	SD
T.P.	34.8	9.7	32.8	9.7	33.8	9.5
S.S.	36.7	6.4	40.8	7.1	38.8	6.9
B.L.	34.6	10.8	28.1	11.5	31.3	11.4
L.H.	42.9	7.5	35.5	5.8	39.1	7.6
C.G.	28.5*	10.2	32.0	6.5	30.3	8.4
Average	35.5	5.2	33.8	4.7	34.7	4.1

*Based on 7 trials.

Two points that are seemingly at odds with predictions based upon the SFMT—that structure is recovered well in displays consisting of coplanar objects and in polar-projection displays (or both)—can be accounted for by assuming additional algorithms or heuristic rules. For example, the perspective factor identified by Braunstein (1976) for polar projections provides information regarding the direction of rotation of stimulus elements. Assuming the correct solution of the correspondence problem, the assignment of a direction of rotation to a collection of moving elements could serve as a heuristic rule for determining structure in polar-projection displays.

The SFMT and any performance theory based upon it are necessarily limited to predictions regarding the recovery of structure from motion and are essentially neutral regarding other aspects of the perceptual experience (e.g., the subjective experience of depth, the perception of rotation direction, etc.). Nonetheless, human subjects are clearly capable of basing their judgments regarding structure on a variety of criteria and heuristic rules (Braunstein, 1983; Cutting, 1986). In this respect, the subjective reports of the current subjects in conjunction with aspects of the present data are instructive. For example, despite the breakdown in the appearance of structure beyond rotation angles of about 35.5° , most subjects were able to recover structure above chance levels even with rotation angles of 50° and larger. This ability suggests that under some conditions the subjects made use of a decision process that went beyond a simple perceptual judgment. (For example, the comparison of Experiments 1 and 2 with Experiments 3 and 4 showed that the successive presentation of both perturbed and unperturbed displays contributed about 10% accuracy to judgments.) Also, with relatively small rotation angles, the observers reported the subjective experience of structure in many (if not most) perturbed displays, although such displays had no exact correspondence to rigidly rotating objects. In such cases, the judgment of structure seemed more to be relative, rather than absolute. As mentioned earlier, in trials in which both perturbed and unperturbed displays were presented, the subjects sometimes made their judgments by eliminating the display with the less compelling subjective experience of structure.

The possibility that the subjects based judgments of structure on a variety of criteria suggests that the human judgment of structure from motion is multidetermined. Human observers can make separate judgments of depth, rotation direction, and coherence in the kinds of displays described here (Braunstein, 1962; Petersik, 1980b). A relatively strong subjective response on any of these criteria, as well as on others discussed above, could lead a subject to report structure in a stimulus display, because ecologically such phenomena as depth and a constant rotation direction are correlated with structure. Furthermore, subjects seem to operate with a strong bias favoring the rigidity assumption because they frequently report structure in successive displays of randomly placed stimu-

lus elements⁶ and they locate structure relatively easily amidst a great deal of noise (Todd, 1985).

Because of the likely multiple determination of structure-from-motion judgments, a performance theory based upon the SFMT would probably be a theory of the ideal observer at best. Ullman (1984) developed an "incremental rigidity" scheme that accommodates many of the findings from studies of human perception. In this scheme, glimpses of an object undergoing a 3-D transformation lead to refinements of an internal model of the object's structure. A computer implementation of the incremental-rigidity scheme integrated views over time, preferred rigid interpretations over nonrigid interpretations, and tolerated stimulus nonrigidity well. In general, the incremental scheme converged on a correct interpretation of the 2-D data, although the correct interpretation frequently required more than one revolution (rotation) of the object. Furthermore, the incremental-rigidity scheme was reported not to recover structure well with only three views of four elements undergoing transformation. Because subjects in the present experiments were able to recover structure with only three views of four elements rotating less than one complete revolution, a performance theory based on the SFMT and incorporating the incremental-rigidity scheme would still seem to be incomplete.

In summary, a performance theory based upon the SFMT is psychophysically plausible if one allows for additional decision-making or heuristic rules. Incorporating the incremental-rigidity scheme (Ullman, 1984) would enhance the ability of the theory to predict many aspects of human perceptual performance but would still leave the theory incomplete.

REFERENCES

- BAKER, C. L., & BRADDICK, O. J. (1982). The basis of area and dot number effects in random-dot motion perception. *Vision Research*, **22**, 1253-1259.
- BÖRJESSON, E., & VON HOFSTEN, C. (1973). Visual perception of motion in depth: Application of a vector model to three-dot motion patterns. *Perception & Psychophysics*, **13**, 169-179.
- BRADDICK, O. J. (1974). A short-range process in apparent movement. *Vision Research*, **14**, 519-527.
- 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. (1976). *Depth perception through motion*. New York: Academic Press.
- BRAUNSTEIN, M. L. (1983). Contrasts between human and machine vision: Should technology recapitulate phylogeny? In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision*. New York: Academic Press.
- CAVANAGH, P., & ANSTIS, S. M. (1980). Visual psychophysics on the Apple II: Getting started. *Behavior Research Methods & Instrumentation*, **12**, 614-626.
- CUTTING, J. E. (1986). *Perception with an eye for motion*. Cambridge, MA: MIT Press.
- DONNER, J., LAPPIN, J. S., & PERFETTO, G. (1984). Detection of three-dimensional structure in moving optical patterns. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 1-11.
- HERSHBERGER, M., & URBAN, D. (1970). Depth perception from mo-

NOTES

- tion parallax in one-dimensional polar projections: Projection versus viewing distance. *Journal of Experimental Psychology*, **86**, 133-136.
- JOHANSSON, G. (1978). Visual event perception. In R. Held, H. W. Leibowitz, & H. L. Teuber (eds.), *Handbook of sensory physiology*. Berlin: Springer-Verlag.
- MARR, D. (1982). *Vision*. San Francisco: Freeman.
- PETERSIK, J. T. (1978). *Judgments of direction and depth in dynamic perspective transformations*. Unpublished doctoral dissertation, Miami University, Oxford, OH.
- PETERSIK, J. T. (1979). Three-dimensional object constancy: Coherence of a simulated rotating sphere in noise. *Perception & Psychophysics*, **25**, 328-337.
- PETERSIK, J. T. (1980a). The effects of spatial and temporal factors on the perception of stroboscopic rotation simulations. *Perception*, **9**, 271-283.
- PETERSIK, J. T. (1980b). Rotation judgments and depth judgments: Separate or dependent processes? *Perception & Psychophysics*, **27**, 588-590.
- PETERSIK, J. T. (1987). *Spread of effect involving short-range apparent motion*. Manuscript in preparation.
- PETERSIK, J. T., & PANTLE, A. J. (1979). Factors controlling the competing sensations produced by a bistable stroboscopic motion display. *Vision Research*, **19**, 143-154.
- PETERSIK, J. T., PUFALL, R., & KRASNOFF, E. (1983). Failure to find an absolute retinal limit of a putative short-range process in apparent motion. *Vision Research*, **23**, 1663-1670.
- TODD, J. T. (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.
- ULLMAN, S. (1980). The effect of similarity between line segments on the correspondence strength in apparent motion. *Perception*, **9**, 617-626.
- ULLMAN, S. (1984). Maximizing rigidity: The incremental recovery of 3-D structure from rigid and nonrigid motion. *Perception*, **13**, 255-274.
- WALLACH, H., & O'CONNELL, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, **45**, 205-217.

1. The viewing distance was not varied to match simulated projection distance because Hershberger and Urban (1970) had shown that the veridicality of judgments of rotation varies with *simulated* projection distance, not *actual* viewing distance.

2. A perspective ratio of 2.0 was chosen because it is very close to the value used by Braunstein (1962) at which depth judgments and coherence judgments of rotation simulations tend to be equivalently strong. It was hoped that such equivalence would maximize the determination of structure from motion.

3. This figure resulted from recording the 2-D projections of elements over successive frames in 100 different simulations of the type described here and subsequently averaging the position changes.

4. I do not intend to draw a strong conclusion from a failure to reject the null hypothesis alone. However, it does seem important that changing the computation of the extent of perturbation used in Experiment 2, relative to the constant perturbation used in Experiment 1, produced a change in the resulting functions. Had Experiments 1 and 2 been conducted as a single study, "perturbation type" would probably have emerged as a significant factor.

5. It could be argued that these experiments are confounded, because interstimulus interval (ISI) was not varied (in accord with Korte's third law) as rotation angle increased. To determine whether lengthening ISI would improve performance at greater rotation angles, I replicated Experiments 1 and 2 with ISIs that were proportional to the angle of rotation. The results were virtually identical to those obtained in the experiments reported above.

6. In demonstration displays that were recently developed here, three frames that consisted of four randomly placed elements were presented in the same manner as displays of the present experiments. With such random-element displays, observers occasionally have reported perceiving an object undergoing some (often complex) transformation. Sometimes the "object" is seen in a transversal motion while rotating, sometimes as an object stretching, sometimes as an object undergoing rotation and expansion, and so forth. At the same time, observers have usually been aware that no real object is represented in the display and that their experiences are perceptually constructive.