

Percepts of Rigid Motion Within and Across Apertures

Maggie Shiffrar
Stanford University

M. Pavel
Department of Psychology and Center for Neural Science
New York University

Humans consistently err in their percepts of rotational motion viewed through an aperture. Such errors provide insight into the constraints observers use to interpret retinal images. In the 1st of 2 experiments, Ss consistently perceived the fixed center of rotation for an unmarked line viewed through an aperture as located on the line, regardless of its actual location. Accuracy greatly improved with visible line endings. This finding was extended to explain why a square appears nonrigid when it rotates behind a partial occluder. This illusion may result from observers misperceiving the center of rotation of the unmarked square sides. In this situation, Ss seemed unable to apply an object rigidity constraint across apertures. These findings support a conceptualization of the visual system in which consistent local information must be clearly present before prior knowledge can be used to interpret retinal stimulation.

The perception and recognition of objects from two-dimensional images is frequently a difficult problem because many different objects may be consistent with any particular image. Yet, human observers often interpret ambiguous images in systematic ways. Constraints can aid in the interpretation of images by restricting the set of possible solutions. For example, rigidity may be a useful constraint to help people perceive objects in motion.

Researchers frequently have assumed that when a sequence of images represents a moving object, the visual system favors solutions corresponding to rigid objects (Ullman, 1979). Although rigidity is certainly an important constraint, there are circumstances in which a nonrigid percept is preferred by the visual system. For example, a nonrigid interpretation occurs when another constraint is in competition with and more salient than rigidity (Braunstein & Andersen, 1984; Nakayama & Silverman, 1988a, 1988b; Schwartz & Sperling, 1983; Wallach & O'Connell, 1953). One of the assumptions underlying our work is that the effectiveness of various constraints can be evaluated by examining the competition among constraints.

Most of the current models of visual information processing are based on the assumption that the visual system is composed of interconnected units forming a hierarchical structure (e.g., Hildreth & Koch, 1987; Marr & Ullman, 1981; Nakayama, 1985; van Santen & Sperling, 1984; Watson & Ahumada, 1985). In these models, which are consistent with much of the recent research in physiology and anatomy, units at lower levels are likely to process local information from

relatively smaller retinotopic neighborhoods. The results of these local analyses are then processed by higher-level units that can integrate information over larger spatial extents and consequently can perform more global analyses. The assumption of such a multiresolution, hierarchical structure gives rise to several questions: What analyses and decisions are made at the different levels of the hierarchy? What is the spatial extent over which information can be combined at different levels of analysis? In particular, how does information observed at one location affect the interpretation of observations at another?

One convenient way to interpret the process of combining information from different locations is in terms of constraint satisfaction. That is, observed motion at one point may eliminate certain interpretations of local observations at another point if the points are associated in the image. The visual system attempts to interpret the image such that all constraints are satisfied. If the different constraints cannot be satisfied simultaneously, the visual system interprets the spatially distinct observations as arising from different, or even nonrigid, objects.

For the purpose of this discussion, we distinguish global constraints from those that arise from local analyses. To clarify the distinction between local and higher-level, global constraints, we define local analyses in terms of the computations that can be performed on arbitrarily small neighborhoods of points. Local constraints therefore can be used to interpret the motion of a single, small segment of a contour of an object. Thus, edge detectors that approximate differentiation (e.g., Laplacian zero-crossings) are examples of local analyses that result in local constraints.

Higher-level constraints generally are based on prior knowledge and biases, that is, information about the environment that is available to the visual system in addition to an image. For example, suppose that when confronted with a new image, the visual system has a preference to interpret the image as containing a single, rigid object. The visual system can use this preference for rigid objects to constrain the interpretation of motion information at different points in the image. These

This work was supported by U.S. Air Force Office for Scientific Research Grant AFOSR-84-03-08, and by National Aeronautics and Space Administration Grant NCC 2-269 to Stanford University.

We thank Helen Cunningham, Ken Nakayama, and Roger Shepard for many valuable discussions. We thank Jim Todd for his comments on an earlier draft of this article.

Correspondence concerning this article should be addressed to M. Pavel, Department of Psychology, New York University, 6 Washington Place, New York, New York 10003.

assumptions about the nature of the world can be applied across disconnected contour segments. Examples of higher-level, or global, constraints include object rigidity, object constancy, and a constant or slowly varying illuminant over space and time. Whereas global constraints are likely to involve top-down processes, local constraints may reflect the behavior of lower-level, or bottom-up, processing.

Although the distinction between local and global analyses and constraints is relative (e.g., in the case of multiple objects), we use it as a convenient way to describe integration of information over images. In general, global constraints underlie rules that govern the process of information integration across space. We note that the distinction between local and global analyses in the context of the aperture problem has been used by other investigators to describe analysis of motion (e.g., Koenderink, 1986; Waxman, Kamgar-Parsi, & Subbarao, 1987). In fact, most solutions of the motion problem involve some form of global constraint.

In this article, we report the results of an investigation of a competition between local and global constraints by examining simple moving objects viewed through one or more disconnected windows or apertures. As we describe below, a task that involves viewing two-dimensional objects through relatively small apertures permitted us to design experiments in which the results of motion perception within each aperture contradict the global analysis across all apertures. This strategy allowed us to investigate the nature of constraint competition in the perception of motion. Before discussing the experiments, we consider the effects of viewing moving objects through small apertures.

Absence of Motion Information: The Aperture Problem

If an infinitely long (extending beyond the field of view) homogeneous translating line is viewed, only the component of motion perpendicular to the orientation of the line can be measured. Both human and ideal observers are unable to extract the parallel component of motion, as all points along the length of the line are identical. Because of this inability to perceive the parallel motion of a line having no visible endpoints or markers, all line motions consisting of the same perpendicular component but a different parallel component appear to move identically, as shown in Figure 1. In fact, such motion ambiguities arise whenever the intensity variations of a particular orientation extend beyond the area of measurement (Horn & Schunk, 1981). Thus, motion perception either of an infinitely long line or of a line viewed through a relatively small window cannot be determined without additional constraints. This ambiguity, known as *the aperture problem*, has received much attention (Adelson & Movshon, 1982; Burt & Sperling, 1981; Hildreth, 1984; Nakayama & Silverman, 1988a, 1988b; Poggio, Torre, & Koch, 1985; Rock, 1981; Wallach, 1935) because any conceivable motion detection mechanism, whether biological or computational, is likely to have receptive field(s) that are limited in size.

Perception of Translation in an Aperture

The lack of motion information leads to many interesting perceptual effects. Perhaps the best known illusion due to the

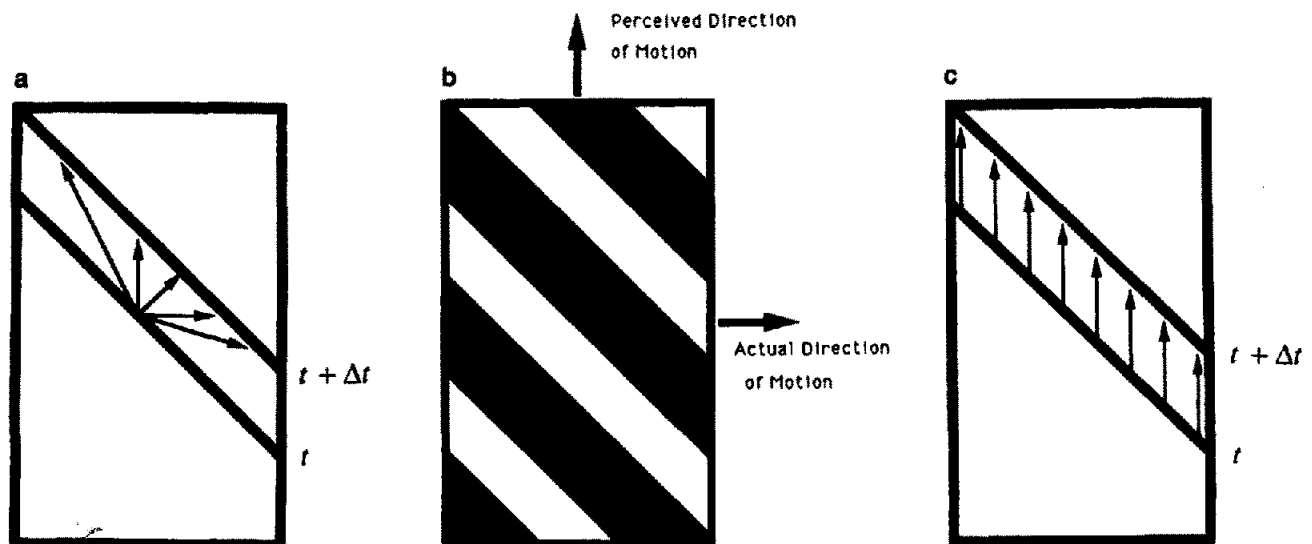


Figure 1. (a) A moving line seen at two different times behind a rectangular aperture. (Because the parallel component of motion is not visible, any motions with the same perpendicular motion but differing parallel motion appear identical.) (b) The barber pole illusion. (Observers report a vertical motion consistently, although the stimulus is actually consistent with an infinite number of real motions.) (c) The direction vectors resulting from the application of a smoothness constraint. (The vectors that were perpendicular to the orientation of the contour are now parallel with the edges of the aperture.)

aperture problem is the *barber pole illusion*, which can be described as a set of strips translating horizontally behind a rectangular aperture, as shown in Figure 1. Even though the strips move horizontally, observers see them as moving in a vertical direction. Because observers see this ambiguous stimulus as moving in only one direction, they must use additional information or constraints to eliminate all other possible interpretations. The specific constraints used to interpret images are the focus of a number of theoretical approaches involving various global constraints.

In one class of theories, the visual system resolves motion ambiguity by assuming that the luminance distribution of the underlying image (except for the motion to be determined) is constant and that the velocity varies smoothly over space (Horn & Schunk, 1981) or along contours (Hildreth, 1984). This approach to velocity estimation is framed as the optimization of an objective function. One of the two components of the objective function to be minimized by the visual system is the deviation from the observed perpendicular motion at each point along a contour (local analysis); the other component is the smoothness constraint. This smoothness constraint can be thought of as arising from local rigidity. The basic idea is that nearby points that belong to the same contour most likely will move with similar velocities. As a result, interpretations that minimize velocity differences between neighboring points are favored by this local rigidity, or smoothness constraint. An example of a particular objective function, J , defining such a constraint applied along a contour can be formally expressed as a sum of two terms integrated along the contour length S ,

$$J(\mathbf{V}) = \int_{s \in S} \left\{ \left| \frac{d\mathbf{V}(s)}{ds} \right|^2 + [v_L(s) - u_L(s)]^2 \right\} ds, \quad (1)$$

in which \mathbf{V} is the estimate of the velocity vectors (field) over the contour, $v_L(s)$ is the magnitude of the estimated perpendicular component at s , and $u_L(s)$ is the measured perpendicular component. The first term, $|d\mathbf{V}(s)/ds|^2$, is a measure of the roughness of the velocity field. The second term, $[v_L(s) - u_L(s)]^2$, represents a measure of the error of the estimated velocity field relative to the unconstrained, measured values of the velocity field. The estimate of the actual velocity is obtained by minimizing J over all feasible assignments for \mathbf{V} . This formalization of the smoothness constraint is one of several that have been proposed (e.g., Horn & Schunk, 1981; Yuille & Grzywacz, 1988). Moreover, other constraints imposed by the visual system can be included as additional terms in Equation (1). For example, Nakayama and Silverman (1988b) added a term to characterize the tendency of the visual system to favor velocities perpendicular to contours.

The constraint maximizing the smoothness of a velocity field could be used by the visual system to interpret the translation of a contour behind an aperture, as in the barber pole illusion. The velocity, $u_L(s)$, measured at the points of the contour located along an aperture's edge is parallel to the aperture. To minimize variation in velocity, or to maintain local rigidity, all intervening points of the contour must move in a parallel direction relative to these edge velocity vectors.

Thus, the stripes in the barber pole illusion are interpreted as moving vertically because the points along each strip are interpreted as moving in a direction parallel to the velocities at the aperture edge, as illustrated in Figure 1.

One implication of the class of theories involving the global smoothness constraints is that the perceived motion of a line moving in an aperture is, to a large extent, determined by the line's motion at the edges of the aperture. The predictions of this smoothness constraint theory are therefore the same as those of Wallach (1976), who proposed that lines will be seen as moving in the vertical direction if the predominant motion at an aperture boundary is vertical. In a similar manner, the perceived motion of a line moving in a circular aperture is determined by the direction of motion of the visible endpoints of the line. As recently proposed by Shimojo, Silverman, and Nakayama (1989), the critical role of the motion of endpoints may result from the classification of line terminators. If aperture boundaries could cause visible endpoints of a line to be interpreted as intrinsic terminators, then the visual system would be more likely to use the motion of these endpoints to interpret the motion of the entire line.

In another class of theories, proposed by Fennema and Thompson (1979) and extended by Adelson and Movshon (1982), motion is determined by intersections of constraints. According to these theories, the motion of an object is represented in a velocity space with coordinates corresponding to the velocities v_x and v_y in the x and y directions, respectively. A measurement of velocity, u_L , perpendicular to a contour is consistent with all velocities on a line in the velocity space. The slope of the velocity line is determined by the orientation of the contour. A unique velocity of an object can be determined by the intersection of different velocity lines that correspond to contours with different orientations within a single aperture. As stated, this theory is directly applicable only to the ambiguity problem arising from a single aperture. To extend this theory to multiple apertures, we must assume rigidity of the moving object. For example, if a polygon moved behind several apertures such that different edges were seen through different apertures, the visual system could interpret uniquely the translational motion, with the constraint proposed by Adelson and Movshon, by finding the point of intersection of the velocity vector from each aperture.

In both of the classes of theories described earlier, local information (from aperture edges) is used to constrain the movement of an entire contour or object within an aperture by applying a global constraint over continuous space. The constraints used to resolve the ambiguity are based on object rigidity. An interesting violation of this assumption of rigidity, recently documented by Nakayama and Silverman (1988a, 1988b), occurs when a curved line translates behind an aperture. Nakayama and Silverman found that they could make a translating sinusoidal line appear to be nonrigid by manipulating the line's curvature. We generalize their approach by considering the question of rotation.

Aperture Problem: Rotating Lines

In general, a rotating line viewed through an aperture is simultaneously translating and changing orientation. We are

interested in the case in which a line is rotating rigidly around a fixed center. Can the movement of such a line be perceived accurately?

With Prior Knowledge of a Fixed Center

Suppose that a line in an aperture is rotating rigidly around a center off of the line and that the image is analyzed by an ideal observer who knows that there is a fixed center of rotation but does not know the center's location. An ideal observer, with the prior knowledge that the line rotates around a fixed center, can determine the location of the fixed center precisely from three (or more) different views of the line. For example, for any pair of views of the line, the center of rotation must lie on the line bisecting the angle of their intersection, as shown in Figure 2. For two different pairs of views, two different angle bisectors can be found. The intersection point of these bisectors indicates the location of the fixed center of rotation. Thus, prior knowledge of the existence of a fixed center of rotation eliminates the otherwise infinite number of possible interpretations and invites a unique solution.

This analysis demonstrates an important distinction between translational and rotational motions interpreted with the same prior knowledge. In both cases, a moving line is known to belong to a rigid object. In the case of viewing a translating line through a single aperture, neither human nor ideal observers can identify accurately the actual motion of

the line because the parallel component of motion is absent. In the case of rotation, however, both the parallel and the perpendicular motion components are available over time, providing an opportunity for accurate perception.

Without Prior Knowledge of a Fixed Center

Suppose that an ideal observer does not know that a line is rotating around a fixed center. When knowledge of a fixed center is absent, an infinite number of interpretations are consistent with the observed motion. For the observer to interpret the line's motion uniquely, a constraint must be imposed on the possible motions. If the observer were to use a constraint in accordance with the smoothness theories described previously, the instantaneous center of rotation would be located on the line, as shown in Figure 2. As in the case of translation, a smoothness constraint applied locally within an aperture would favor interpretations that maintain the rigidity of the visible portion of the contour. A smoothness constraint could be applied within an aperture in the following way: As a result of a correspondence between the visible endpoints of the line, the measured velocities of these endpoints are parallel to the edge of the aperture. At any instant, a smoothness algorithm constrains the interpretation solution such that all the points along the contour have velocities approximately parallel to those of the endpoints. In addition, neighboring points also are interpreted as having velocity magnitudes that

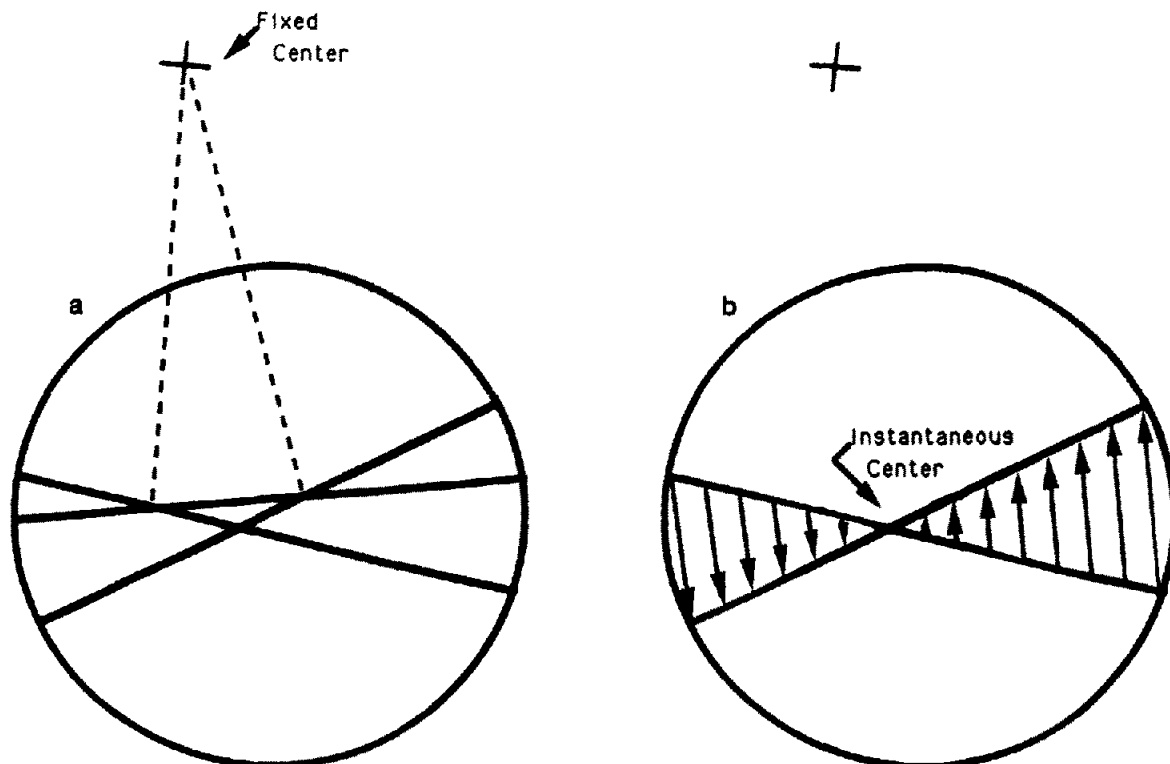


Figure 2. (a) Three consecutive views of a solid line rotating about a fixed center off the line. (An ideal observer could determine the center of rotation from three such views by finding an intersection of each pair of lines.) (b) Location of the center of rotation through the application of the smoothness constraint. (Velocity differences between points on a line are minimized. The center of rotation is that point with a zero velocity vector. This constraint yields a center of rotation located on the line.)

vary smoothly. The point on the contour having a velocity vector of magnitude zero is interpreted as the instantaneous center of rotation. This center is located on the rotating line or on that line's invisible extension. Over time, this center of rotation itself rotates around the fixed center of rotation. The resulting percept is, in fact, a correct description of the actual motion of the object, or more precisely, of the edge of the object.

One of our goals was to determine how people perceive rotational motion. To determine whether human observers can use prior knowledge of object rigidity, we asked subjects to locate the center of rotation for lines rotating about fixed centers under different conditions. In the first experiment, which was designed to measure how accurately humans perceive the location of a fixed center of rotation, subjects saw lines rotating either in apertures that covered the endpoints of the lines or in windows large enough to expose the lines' endpoints. In both cases, subjects were told that the line rotated about a fixed center. The angle of rotation also was varied so that the relationship between angular extent and accuracy of perception could be assessed.

Experiment 1

Method

Subjects. Twenty Stanford University students and researchers participated in this study either as volunteers or for credit toward completion of a class requirement. All subjects were unaware of the hypothesis under investigation.

Apparatus. Stimuli were displayed on a Hitachi RGB 19-in. (48.3-cm) monitor with 1024 × 768 pixel resolution and 60-Hz refresh rate. The monitor was controlled by a Silicon Graphics IRIS 2400 Workstation system. Subjects used a "mouse" device to record their responses.

Stimuli. The two types of stimuli used in the two display conditions of this experiment are shown in Figure 3. In both the aperture and nonaperture conditions, subjects observed a green homogeneous line rotating with an angular velocity of 90°/s in an oscillatory motion about a fixed center. The line width subtended 0.16° of visual angle when viewed from the subjects' viewing distance of approximately 61 cm. In the aperture condition, the rotating line was seen behind a square viewing window, or aperture, with a side length of 6.5° of visual angle. The boundary of the viewing window was outlined with a yellow line. The actual endpoints of the line were hidden throughout

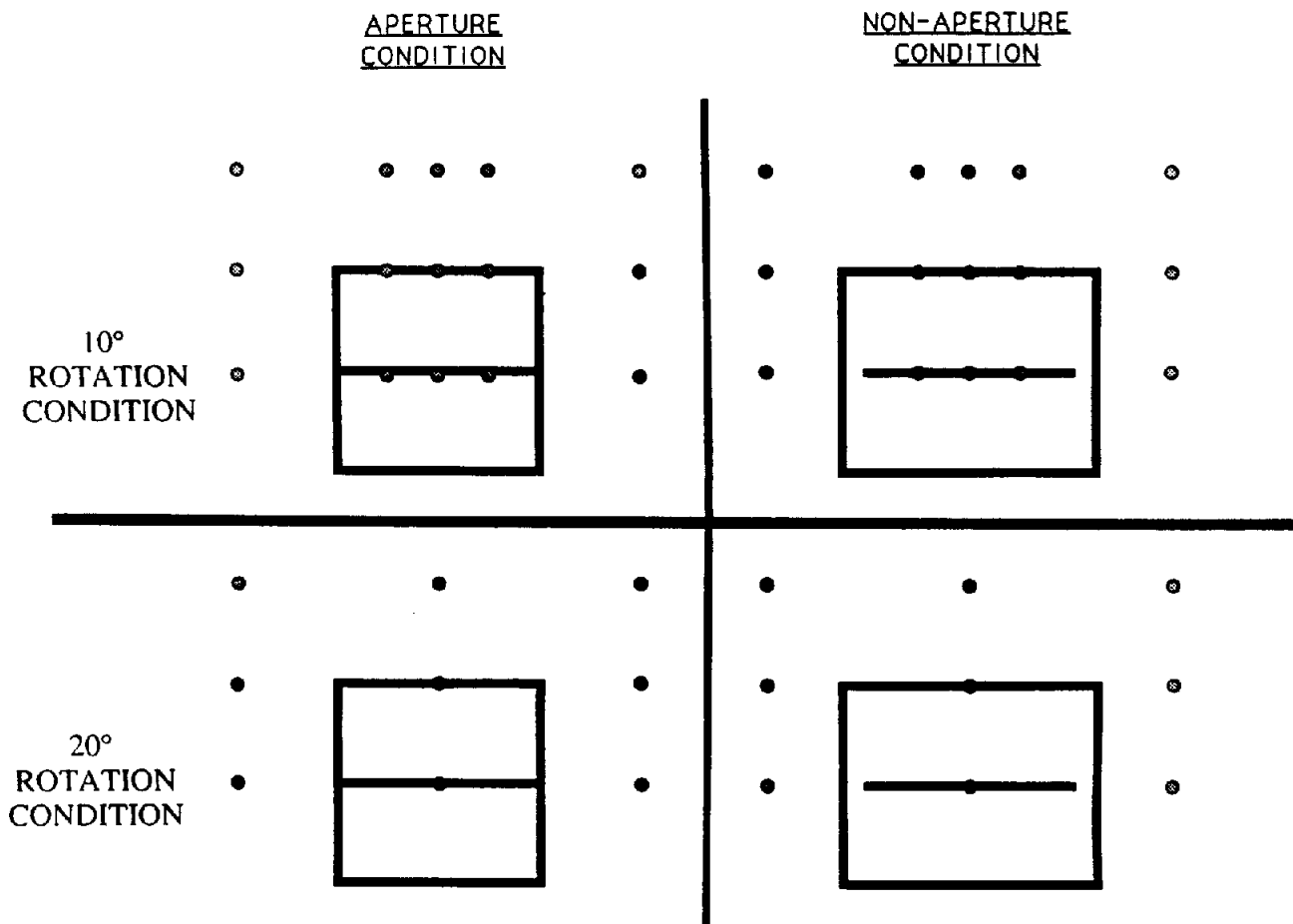


Figure 3. The stimuli for Experiment 1. (Stimuli varied according to their display and rotation conditions. The displays consisted of either aperture or nonaperture conditions. The rotations subtended either 10° of rotation about 1 of 15 possible axes or 20° of rotation about 1 of 9 possible axes. The locations of the centers of rotation are denoted by the shaded circles.)

the rotation, so that the visible endpoints of the line coincided continuously with the edges of the aperture, as though the length of the line extended behind the aperture.

The stimulus in the nonaperture condition was similar to that in the aperture condition, except that the viewing window was longer in the horizontal dimension and did not occlude the endpoints of the rotating line. The height of the viewing window remained 6.5° of visual angle, whereas the width was enlarged to subtend 7.5° . The length of rotating line was 6.5° of visual angle so that its visible length was similar to that in the aperture condition. Again, the endpoints of the line were visible throughout the rotation. In both the aperture and the nonaperture conditions, a red dot (0.16° diameter), the location of which was controlled by a mouse device, was used by subjects to indicate their perceived center of rotation on each trial.

There were also two possible rotation conditions in both the aperture and the nonaperture display conditions. These two rotation conditions differed from each other with regard to the angle of rotation (10° vs. 20° , respectively) and with regard to the number of possible centers of rotation (15 vs. 9, respectively). In both rotation conditions, the rotation had the same uniform angular velocity. In the 10° rotation condition, the line rotated back and forth through 10° around 1 of 15 possible centers of rotation, as shown in Figure 3. The vertical positions of the centers were as follows: (a) 5 of the centers were located along the vertical midpoint of the viewing window, on the stationary location of the horizontally oriented line; (b) 5 centers were located 3.25° above the line, on the yellow outline of the viewing window; and (c) 5 centers were positioned 6.5° above the line, outside of the viewing window. The horizontal positions for each of these vertical locations were the following: (a) the horizontal midpoint of the rotating line; (b) 1.63° to the right of the midpoint; (c) 1.63° to the left of the midpoint; (d) 6.52° to the right of the midpoint; and (e) 6.52° to the left of the midpoint. One of the 15 possible centers of rotation was used in each trial of the 10° rotation condition.

In the 20° rotation condition, the line rotated back and forth through 20° of angle around 1 of 9 possible centers of rotation. These 9 centers of rotation were identical to the 15 centers from the 10° rotation condition, except that the sets of centers located 1.63° to the right and 1.63° to the left of the midpoint were eliminated.

Procedure. Each trial began with a display of the viewing window with the rotating line positioned in the center of the computer screen. After trial onset, subjects were requested to observe the rotating line and to determine the location of the center of rotation. Subjects were told that the location of the center of rotation varied from trial to trial and could be positioned anywhere on the computer screen. The perceived center of rotation was assessed by the method of adjustment. On each trial, subjects were asked to position a red cursor dot, using a mouse device, at the perceived location of the center of rotation. The initial position of the red cursor dot varied randomly between trials. When subjects were satisfied with the position of the cursor, they pressed a button on the mouse. After the button press, the screen cleared and the stimulus of the next trial appeared. Subjects were given as much time as they needed to respond on each trial. The line continued to rotate back and forth around one of the fixed centers until a response was made. Each subject completed 10 trials per center of rotation.

The experiment had a between-subjects design with four possible conditions. Each subject was assigned randomly to either the 10° aperture condition, the 10° nonaperture condition, the 20° aperture condition, or the 20° nonaperture condition. Five subjects were assigned to each of the four conditions, and all subjects saw only stimuli within the assigned condition. All subjects completed 5 practice trials before beginning the experimental trials.

Results

The resulting location estimates of the centers of rotation for each of the four rotation-display conditions, averaged over subjects, are shown in Figure 4. These results indicate that all subjects in the aperture conditions reported seeing the center of rotation for every trial as located either on or very near the rotating line, regardless of the actual center location. Although subjects were inaccurate in the vertical component of their localizations, they were very accurate in their estimates of the horizontal component of the location of the rotation center (along the line). In comparison, subjects in the nonaperture, or control, conditions were much more accurate in their perceptions of the locations (vertical and horizontal) of centers of rotation. However, the control subjects tended to underestimate the distance from the center of rotation to the line. Nonetheless, subjects in the nonaperture condition were clearly able to discriminate more accurately than subjects in the aperture conditions between centers of rotation falling on and off the line.

Although the results of aperture and nonaperture conditions were qualitatively different, the results from the 10° rotation condition were not significantly different from those of the 20° rotation condition. The ability of subjects to localize centers of rotation did not change when the angle of rotation was increased from 10° to 20° . In both the 10° and the 20° aperture conditions, subjects perceived fixed centers of rotation as located on or very close to the rotating line. Moreover, during debriefing subjects were asked whether they ever saw a moving center of rotation, and all subjects reported that they saw only fixed, stationary centers of rotation. Although subjects were fairly accurate in their percepts of the fixed centers of rotation, in both the 10° and the 20° nonaperture conditions they exhibited the same tendency to underestimate the distance between centers of rotation and the line.

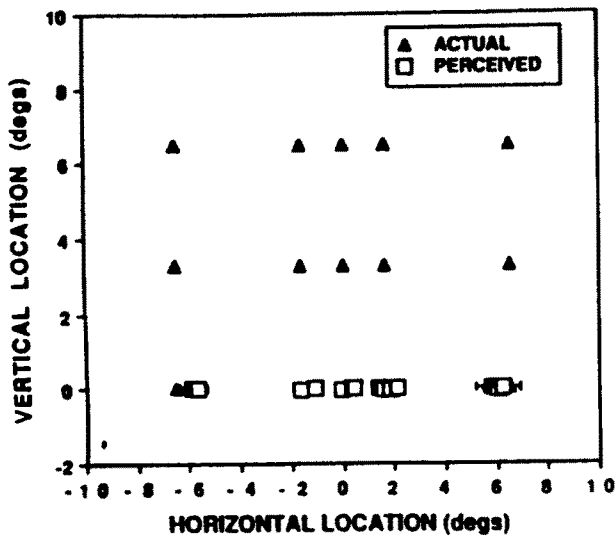
Discussion

Rotation of a line. Unlike an ideal observer with prior knowledge of a fixed center, human observers in this experiment were unable to determine accurately the fixed center of rotation for a rotating line viewed through an aperture. Subjects did not appear to use the global or high-level object rigidity constraint in their interpretations of rotational motion. However, when the same rotating line was viewed through a window that did not hide the endpoints of the line, subjects were able to determine the fixed center of rotation with relatively high accuracy. This pattern of results is consistent with the hypothesis that observers use a constraint, such as local smoothness, to determine the fixed center of rotation for an unmarked line viewed through an aperture. This finding is particularly interesting because an ideal observer using a rigidity or fixed-axis constraint could determine the location of a fixed center of rotation precisely from as few as three views of a rotating line within an aperture.

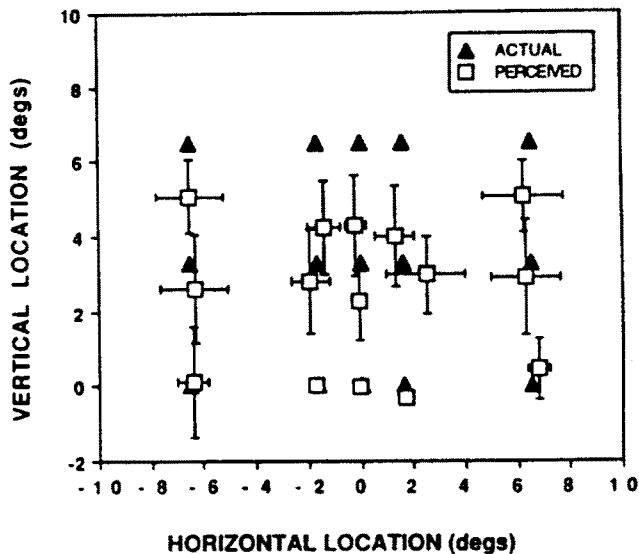
The smoothness constraint can be used to explain the results in the following manner. The explanation rests on the assumption that the motion of visible endpoints of the rotat-

10° ROTATION

APERTURE

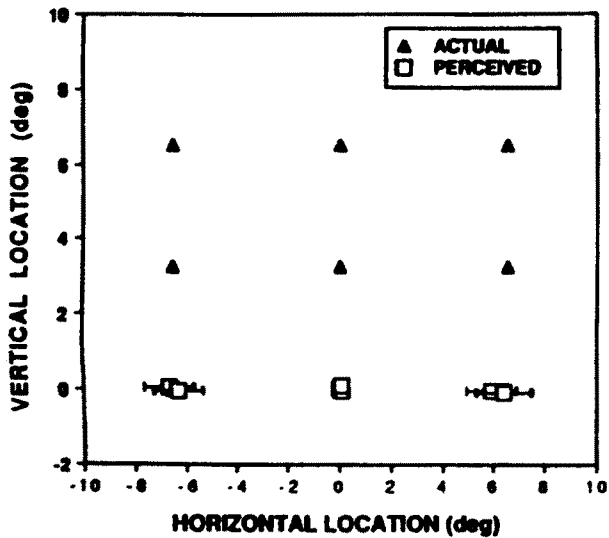


NON-APERTURE



20° ROTATION

APERTURE



NON-APERTURE

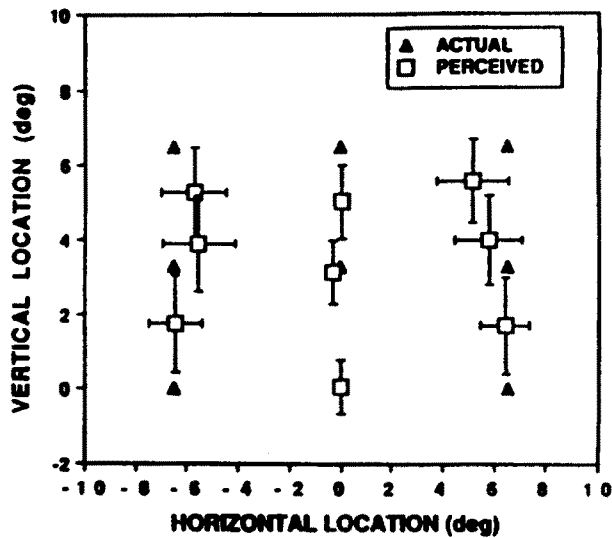


Figure 4. Results for 10° and 20° rotation conditions of Experiment 1. (In the aperture conditions, subjects consistently reported seeing the center of rotation located on the line. In the nonaperture conditions, subjects were significantly more accurate in their perceptions of the distal locations of the centers of rotation. The pattern of results was the same for both the 10° and 20° rotation conditions. Horizontal and vertical bars represent standard deviations.)

ing line is unambiguous, with instantaneous velocities parallel to the aperture edges. After an observer applies the smoothness constraint, all the intervening points must be interpreted as having velocities parallel to those of the endpoints, with magnitudes as similar as possible to those of their neighbors. In other words, application of the smoothness constraint leads the observer to select the interpretation that minimizes two quantities: the differences between the velocities within the velocity field (dV/ds) and the error of the estimated velocity field relative to the unconstrained, measured velocity field [$v_i(s) - u_i(s)$]. In this way, application of the smoothness constraint favors solutions that maintain the rigidity of the visible portion of the rotating line.

In the nonaperture condition, the viewing window never occluded the rotating line segment. Thus, the true intrinsic endpoints of the line were always shown. These endpoints provided all of the information necessary to interpret the line's motion exactly. So, as in the aperture condition, the application of the smoothness constraint to interpret the motion of the completely visible line could have occurred in the following manner: The unique motion of the terminators is propagated along the length of the line; thus, each point is interpreted as having a motion parallel to that of the terminators and as having a magnitude as similar as possible to the magnitude of its neighbors. Because the motion of the endpoints of the line is unique, the motion of the entire line is interpreted accurately.

Although constraints such as maximization of smoothness provide an explanation for people's tendency to see the center of rotation on the line, such constraints do not explain why the subjects did not see the motion of the local center of rotation. In fact our experiment provided a conservative test of observers' ability to perceive this rotational motion. Subjects could have used the red dot cursor as a reference point. A reference point can enable an observer to determine more easily the component of motion parallel to an unmarked line. This parallel component of motion, in conjunction with the perpendicular component of motion (which subjects were able to observe accurately), provides the information necessary to interpret motion correctly. Nevertheless, the reference point did not help observers to accurately perceive the location of the fixed center of rotation for a line in an aperture. Moreover, because no fixation point was provided during the experiment, subjects were free to track the rotating lines with their eyes. Yet, the availability of eye movement information did not eliminate the perceptual bias.

The ability of observers to locate fixed centers of rotation accurately was not significantly influenced by the size of the angle of rotation, at least up to the maximum 20° angle tested. Whether observing 10° or 20° of rotation, subjects in the aperture condition consistently reported that the fixed center of rotation for the unmarked line was located on the line. In the 10° and 20° nonaperture conditions, subjects were able to discriminate axes of rotation that were off the line from those on the line with approximately the same degree of accuracy.

On the basis of these results, however, we cannot exclude the possibility that, at some larger angle of rotation, observers would accurately perceive a fixed center of rotation. For example, in an extreme case, if a line rotated through 360°, observers should be able to locate a center of rotation with

high accuracy in either an aperture or a nonaperture condition. During informal experiments, we determined the angle of rotation above which the visual system could perceive a fixed center of rotation accurately. We found that, for angles of rotation greater than 35°, subjects still saw the center of rotation as located on the line. At these larger angles, however, the center no longer appeared stationary but rather moved back and forth along the length of the visible line segment. This moving center of rotation corresponds roughly to the movement of the line's instantaneous center of rotation. Even with lines rotating through an angle as large as 90°, a person clearly sees the movement of the instantaneous center but does not perceive the fixed center of rotation.

In summary, this experiment demonstrated that human observers have a strong bias to see a center of rotation on or near a moving line and that observers are insensitive to the motion of this apparent center. Subjects were not able to use prior knowledge of a fixed center when that information was presented verbally. Perhaps the accuracy of their judgments would improve if information about the existence of a fixed center of rotation were presented graphically. If this is so, then subjects' accuracy would improve when they observed a rotating rigid object through several apertures. The different information available from spatially separate local measurements and the high-level knowledge about more complex objects should allow observers to use prior knowledge in image interpretation. Because observers often seem to determine object shape from motion (Ullman, 1979), rotational motion of more complex objects seen through more than one aperture might allow for veridical perception. To study quantitatively the interaction between prior knowledge and local constraints, we investigated the perception of rotating rigid two-dimensional objects (e.g., polygons) observed through several spatially separated windows. This investigation is described below.

Rotation of a rigid polygon. The question of how people perceive rotating polygons led us to analyze a recently discovered illusion. The illusion of a nonrigid percept reported by Meyer and Dougherty (1987) seems to indicate that motion information available from multiple apertures does not ensure accurate motion perception. These researchers found that, if an illusory square in a standard subjective contour demonstration is rotated about its center, the rotating square undergoes a nonrigid transformation. A static representation of this stimulus is shown in Figure 5. Although the physical stimulus is always consistent with that of a rigid square, observers report seeing a structurally plastic square that expands and shrinks as it rotates. More specifically, when the corners of the rotating illusory square are visible, the square appears rigid. However, when the corners are not visible, the illusory square appears nonrigid.

Experiment 2

The Meyer and Dougherty (MD) illusion consists of an illusory square partially outlined by straight edges that are visible in the four circular apertures. As the illusory square rotates around its center, these four viewing windows contain either moving corners or homogeneous straight edges. During those periods of rotation when only the straight edges of the

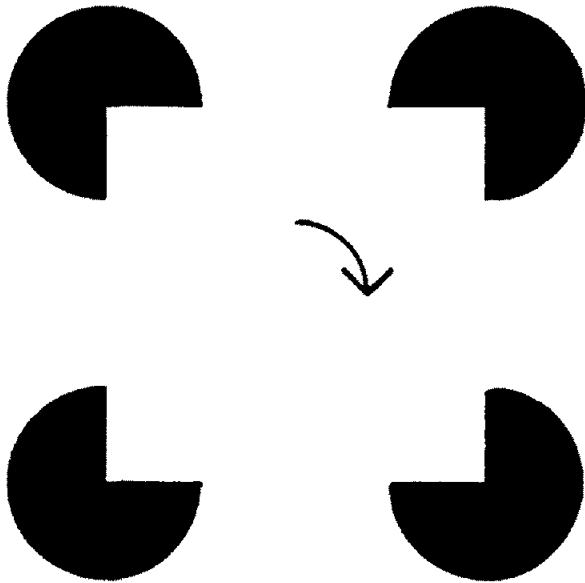


Figure 5. The Meyer and Dougherty illusion. (When the illusory square is rotated about its center, it appears to undergo a phasic size fluctuation. The square appears to shrink and expand as it rotates.)

square are visible, the display is similar to the aperture conditions in our previous experiment. That is, the observer views a plain edge rotating about a fixed center through an aperture. When the corners of the square are visible through these viewing windows, the situation is similar to the nonaperture conditions in this experiment. The presence of the corner allows the subject to extract both the parallel and the orthogonal components of the motion vector and, as a result, to construct an accurate motion interpretation. This analysis is consistent with another characteristic of the MD illusion. When a pattern of large saw teeth is placed on the edge of the illusory square, the square appears much more rigid as it rotates (Meyer & Dougherty, 1987). These distinctive markings may allow the observer to solve continually for both components of motion in the viewing window and thereby to construct a veridical motion interpretation. If the size of the saw teeth is decreased, then the square again appears nonrigid as it rotates, because it becomes more difficult to continuously identify the parallel motion component. Meyer and Dougherty's hypothesis was that the illusion results from the perception of subjective contours. Instead, we postulate that the MD illusion results from the bias of observers toward interpreting the center of rotation of an unmarked line viewed through an aperture as being located on the line, coupled with an inability of observers to overcome this bias by combining motion information across multiple apertures.

To test the ability of human observers to integrate object rigidity information across multiple apertures, we conducted the following experiment. The purpose of this experiment was to test the predictions of our hypothesis regarding the cause of the MD illusion. First, we had to determine whether the MD illusion was restricted to illusory objects or could be extended to partially occluded real objects. According to our hypothesis, any unmarked rotating line segments, whether

they form illusory contours or amodally completed edges, should result in perceived nonrigidity when viewed through multiple apertures. Another issue that this experiment was designed to address was whether motion is necessary for the MD illusion. Our hypothesis suggested that the MD illusion should not persist in the condition of a stationary square with moving apertures. Another possible explanation for the MD illusion might be that the illusion results from changes in brightness. Objects of high luminance are seen as being closer to an observer than are objects of low luminance (Schwartz & Sperling, 1983). As the illusory square of the MD illusion rotates, the amount of exposed area of the viewing windows varies. Because the periodic fluctuation in the amount of exposed area of the viewing windows is correlated directly with the brightness of the overall display, it is possible that brightness variations cause the illusory square to change in apparent depth and that the visual system thus interprets the brightness change as a size change.

To answer these questions, we needed to measure the amount of apparent size fluctuation during rotation. We therefore developed a nulling procedure in which subjects were able to eliminate the perceived size decrease by adding an artificial size increase. The magnitude of the scale increase needed to eliminate the percept of nonrigidity served as a measure of the magnitude of the perceived size decrease during rotation. Using this nulling procedure, we examined a variety of hypotheses underlying the MD illusion in the following experiment.

Method

Subjects. Ten Stanford University students participated in this study for credit toward completion of a class requirement. None of the subjects participated in Experiment 1.

Apparatus. The apparatus for this experiment was the same as that used in Experiment 1.

Stimuli. The stimuli can be classified according to the two conditions of this study. In the rotation condition, a rotating square was viewed through four stationary apertures. In the stationary condition, a stationary square was viewed through four apertures that rotated as a whole about the center of the stationary square. The three possible types of squares are displayed in Figure 6. The first, which we call an illusory square, rotated in front of four square apertures. If the illusory square were to remain stationary in an upright position with one aperture at each corner, each aperture would display 25% of the

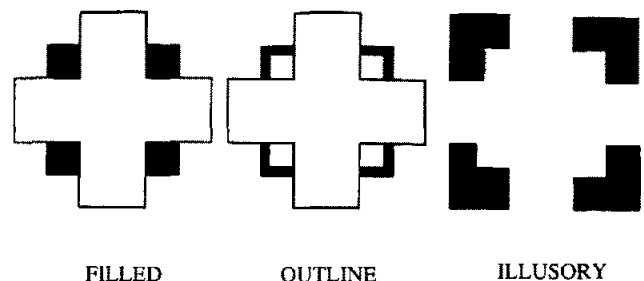


Figure 6. The three types of stimuli for Experiment 2. (The stimuli consisted of a solid square partially occluded by a cross, a similarly occluded outlined square, and an illusory square defined by four square regions.)

length of each side of the illusory square. The illusory squares were defined by green apertures against a dark background. The other two types of squares were both green (real as opposed to illusory) squares, which were partially occluded by a black cross; black was also the color of the background of the display. The second square was filled; the third was an outlined square with a line width of 0.16° of visual angle. Because these "real" squares were partially occluded, they could be seen as whole objects only by amodal completion. If a real square were to remain stationary in an upright position such that every corner was visible behind the stationary cross, then 25% of the area of the square would be visible. All three squares were the same size, with every side subtending 6.5° of visual angle. The squares or apertures, depending upon condition, rotated about their center at a rate of $120^\circ/\text{s}$ (i.e., 3 s per revolution).

During each rotation, the size of the square was physically modified such that the perceptual reduction in size could be approximately compensated for by an appropriate actual enlargement of the square. When the corners of the square were hidden, the instantaneous size of the square at time t was controlled by a scaling factor s :

$$s(\omega t) = 1 + a \left[1 - \frac{(\omega t - 45^\circ)^2}{441^\circ} \right], \quad (2)$$

for $(22^\circ + n \times 90^\circ) \leq \omega t \leq (68^\circ + n \times 90^\circ)$, in which ω is the angular velocity (in deg/s), ωt is the instantaneous orientation of the square, and a is a constant controlled by the subjects. When this factor a was positive, the size of the square increased while the corners were hidden. When the corners of the squares were visible, the scaling factors remained fixed at unity so that the square was actually rigid during this portion of the rotation. Thus, the size of the square changed only when the corners were hidden. We chose this function, which is shown graphically in Figure 7, because it is the simplest polynomial that approximated the perceptual effect: It increased the actual size of the square to a maximum when the corners of the square were at the points farthest from the apertures.

Procedure. Subjects sat with their eyes approximately 61 cm from the computer screen. Each trial began with display of a rotating stimulus positioned in the center of the computer screen. The initial scaling constant a was randomly determined within a range of -0.1 to 0.1 . In this experiment we used the method of adjustment to determine the scaling factor on each trial by moving the mouse device in a vertical direction. Moving up the mouse device increased the scaling constant, whereas moving it down decreased this constant. Each subject was asked to move the mouse device to that point at

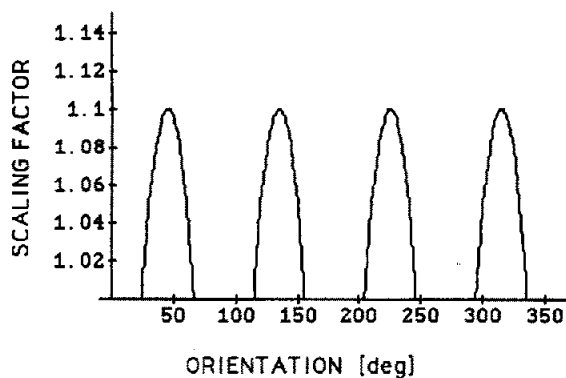


Figure 7. The parabolic function used to control the magnitude of the scaling factor s of the square during rotation. (Subjects varied the amplitude of this function. A scaling factor of 1 corresponds to a constant size of the square, whereas a factor of 2 doubles the length of each side of the square.)

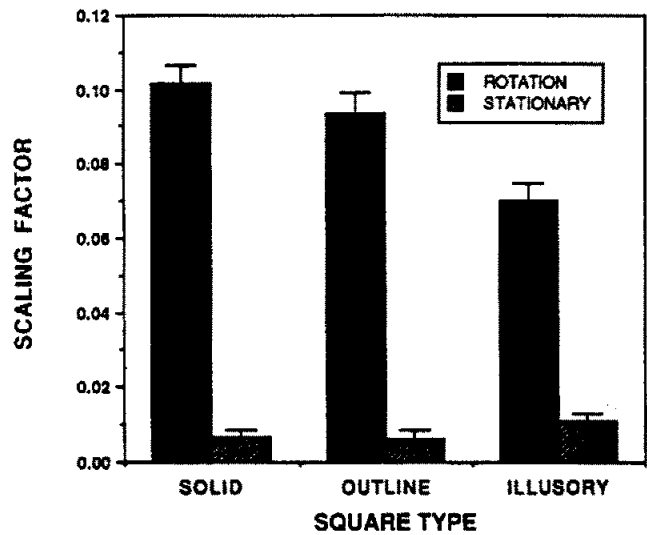


Figure 8. The results from Experiment 2 in terms of scaling constant a . (Subjects added a significantly larger scaling increase for the rotating squares, as compared with the stationary squares. Vertical bars represent standard deviations.)

which the square maintained a constant size throughout the rotation. That is, the subject was asked to find that point at which the phasic size change appeared to be eliminated. Once he or she found this size-change cancellation point, the subject pressed a button on the mouse. After the button press, the stimulus disappeared, and the next stimulus appeared. Each subject completed one block of 60 trials. The order of presentation of the different square types was randomized without replacement for each subject. All subjects completed six practice trials before beginning the experimental trials.

Results

The deviation of the additive scaling constant a from zero that was needed to cancel out the illusory size fluctuation effect was recorded for each trial. This scaling constant served as a measure of the size of the perceived fluctuation. A large, positive scaling constant corresponds to a large perceived size decrease when the corners of the square are hidden. A negative scaling constant would mean a perceived size increase. The scaling constant a for both the rotation and stationary conditions, averaged over all subjects, is shown in Figure 8. The data are presented as the average scaling constant required by subjects to null the perceived size change. A value of 0.1 corresponds to a 10% linear increase of the length of each side of a square. Because area increases as a square of this value, a 0.1 value is equivalent to approximately a 21% change in the perceived area of the rotating square.

All of the scaling constants in the motion condition were positive, which indicates that subjects perceived reductions in square size. In the control (i.e., stationary) condition, the scaling constant was statistically indistinguishable from zero, indicating that this nulling procedure measured changes associated with motion. All subjects reported that they were able to eliminate the size change of the squares with the nulling procedure.

The first major finding is that the loss of rigidity occurred for both real occluded (i.e., amodally completed) and illusory

squares. The nonrigid percept is, therefore, not related to the strength of the perception of subjective edges. To null the perceived size shrinkage, subjects chose positive scaling constants that were significantly greater than zero, $t(388) = 12.50$, $p < .001$. This finding is consistent with our hypothesis that the MD illusion arises from biases in the perception of rotational motion and not from any special characteristics of subjective contours.

The second prediction of our hypothesis, that object motion is necessary for the nonrigidity percept to be obtained, was supported because subjects chose scaling constants for the rotating illusory contour that were significantly larger than those chosen in the stationary illusory square condition, $t(194) = 12.12$, $p < .001$. This same pattern of results was found for both of the partially occluded real squares. That is, subjects required much larger scaling constants to null the perceptual size decrease for both the rotating outlined square relative to its stationary control, $t(194) = 13.90$, $p < .001$, and the rotating filled-in square relative to its stationary control, $t(194) = 18.24$, $p < .001$. Moreover, this finding illustrates that partial occlusion is not responsible for the nonrigidity percept, as both stationary and rotating squares were displayed in the same manner.

The third alternative explanation, that with regard to brightness variation underlying the MD illusion, can be ruled out because the same squares in the rotation and stationary conditions underwent the same brightness fluctuation, yet the size of the nonrigidity effect differed significantly across all pairs of squares. In addition, any difference in the nonrigidity effect that is due to brightness variation should have been smaller for the outline squares than for the solid squares. The data in Figure 8 indicate, however, that the size fluctuation effect for the outline square did not significantly differ from that for the solid square, $t(194) = -1.06$, $p < .292$. These results strongly suggest that brightness is not a significant factor in the illusion.

Discussion

In summary, the nonrigidity effect was strongest when the two-dimensional object had smooth, homogeneous edges and was in rotational motion. These results suggest that the illusion may arise from the ambiguity inherent in observing motion through small apertures, which cannot be eliminated by the application of prior knowledge of object rigidity. The alternative explanations of brightness variation and subjective contour properties were not supported by the results of this experiment. We postulate that human observers use the constraint of local smoothness, or local rigidity, to interpret rotational motion within each aperture and, as a result, perceive centers of rotation as located on each line segment of the square and not in the center of the square itself.

The main result of this experiment was the observation that the apparent size of the square changed in phase with the square's rotational motion. Because the relationship between motion and change in size is critical in perception of three-dimensional space, it seemed plausible that the same mechanism could be responsible for the MD illusion (J. T. Todd, personal communication, 1990). Perhaps the same models that account for reconstruction of spatial structure from mo-

tion would predict our results. In particular, these models use local properties of optic flow in the image (i.e., the curl and divergence) to determine relative structure (e.g., depth) of the three-dimensional world (Koenderink & van Doorn, 1975; Longuet-Higgins & Prazdny, 1980; Regan & Beverley, 1979).

To evaluate the hypothesis that these models of structure from motion, which are based on local analysis of optic flow, could account for the present results, we computed the local derivatives for our stimuli. Both the curl and the divergence were evaluated along the edge of the rotating square. The value of curl characterizes the rotational component of the optic flow and was, at each point, equal to the angular velocity of the rotating square. The divergence indicates expansion and, in the case of the rotating square, can be related to the changes in the visible area of the square. The hypothesis that the perceived change in the size of the object is determined by the divergence is equivalent to the assumption that the visual system uses local information rather than the prior global knowledge. Additional assumptions, however, are required in order to specify how this local information might be integrated over the entire image of an object.

The theory based on local optic flow has one additional shortcoming: It does not account for the fact that the perceived change in size is apparent only during the time intervals when the corners of the square are not visible. To demonstrate the problem, we computed the visible area of the square as a function of time for the rectangular apertures. The resulting area varies as a function of the orientation of the square. It is maximal at the starting position, and it decreases smoothly to the minimum when the square is rotated by 45°. Thus, the visible area is continuously changing, but the percept remains constant during the periods when the corners are visible. Similar failure to account for the data is observed in the condition with stationary square and rotating apertures.

The present experiment, by itself, does not rule out the possibility that the visual system interprets the rotating square as perfectly rigid but translating in depth. In fact, the uniqueness of the rigid interpretation is based on the assumption that the object is rotating in the image plane. The problem with this explanation is that the observers see nonrigid deformation, rather than the perspective transformations that would accompany movement in depth.

Our working hypothesis is that, for smooth, pure rotation motion, observers have difficulty integrating information across different apertures. Rather, they perceive somewhat independent movement in each aperture. We propose that, in the case of smooth, pure rotation, observers see four moving centers of rotation that correspond to the instantaneous centers of rotation for each of the four unmarked sides of the square. This assumption was supported by the findings of Experiment 1 and by subsequent studies in our laboratory. This tendency to see one center of rotation for each side might lead to the perception of nonrigidity in the MD illusion in the following manner: When the perceived center is located on the line as shown in Figure 9, then the predominant component of rotational motion is parallel to the line. This parallel motion is not confirmed by local information and therefore is not seen by the observer as a component of the motion in the aperture. When each center of rotation is perceived to lie along the lines, then the observer perceives a

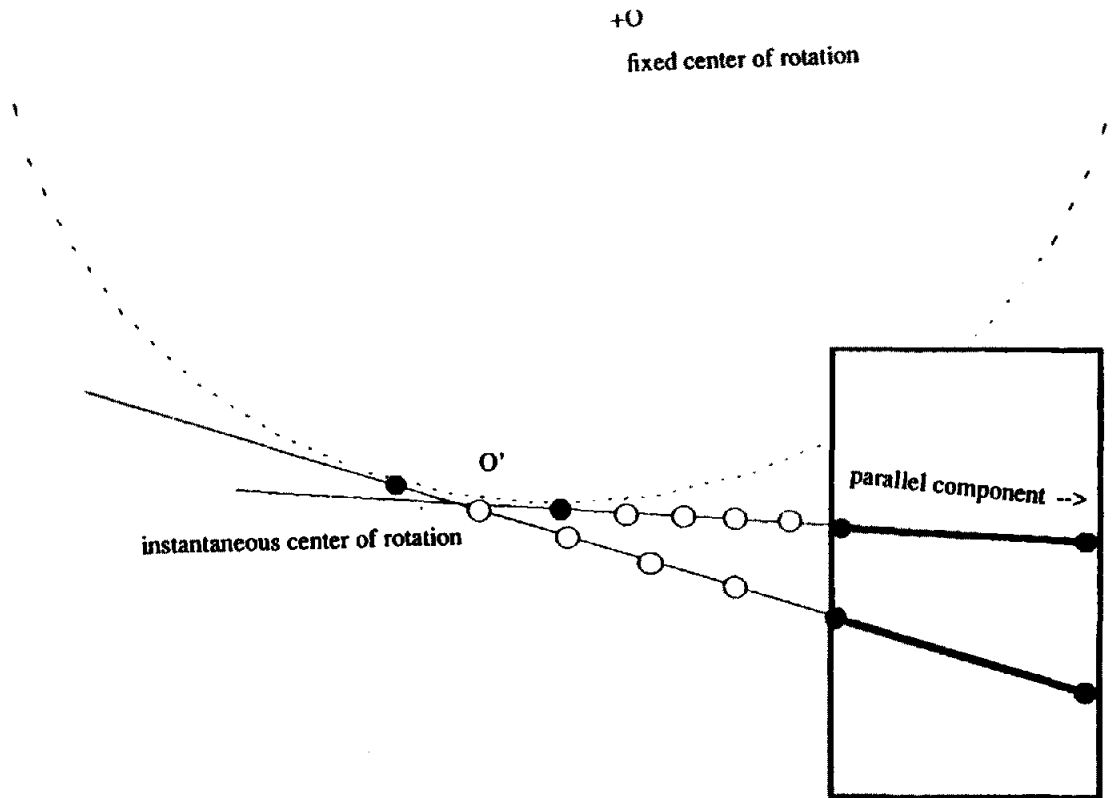


Figure 9. An edge rotating around the fixed center O is perceived as rotating around O' . (In the position shown, the instantaneous motion of O' is mostly parallel to the edge. This parallel component of the motion is not measurable, and the movement of the instantaneous center of rotation is underestimated by the visual system.)

nonrigid shortening of the distance between the line seen in the aperture and the center of rotation.

When the corners of the square are visible, however, the parallel motion information can be interpreted correctly. Because the rotating corners provide both perpendicular and parallel components of motion, the velocity of each point along the length of the line can incorporate the parallel component of motion. As a result, each line segment is perceived to rotate about a fixed center located off of the line (as was shown for the nonaperture conditions in Experiment 1). Because the percepts of motion within each of the apertures are consistent with the same perceived center of rotation, the rotating square appears rigid when its corners are visible.

General Discussion

The general approach of our work was based on the assumption that the visual system applies various constraints to solve the underdetermined problems of image interpretation. Our goal was to determine to what extent high-level, prior knowledge of the rigidity of two-dimensional objects could guide the interpretation of rotating edges and lines. The particular situation of interest arises when the rotating lines are viewed through limited-size apertures that obscure corners. Our findings indicate that human observers cannot easily use

the knowledge of object rigidity to determine the center of rotation. This inability holds for single lines, as well as for multiple lines (edges) defining rigid objects. The case of rotating lines is particularly interesting because an ideal observer with the same knowledge as was held by our subjects could determine the unique solution corresponding to a rigid object.

In the first experiment, we determined how the visual system interprets the rotational motion of a straight line or an edge viewed through an aperture. We discovered a strong bias toward seeing the center of rotation on the line, regardless of prior knowledge. This center was often seen as stationary, even if it was moving in the direction of the line. Observers saw the center as stationary even when they knew that the line was rotating around a stationary center. Thus, we have shown that the prior knowledge of a fixed center of rotation does not override the biases effective at the level of an aperture.

On the basis of these results, in some situations local motion information appears to be more salient than high-level, prior knowledge of objects. We demonstrated in Experiment 2 that these biases for local information prevail even if there is strong visual evidence for a single rigid object. The visual system failed to see a rigid rotating object, although it had sufficient information to resolve the ambiguity resulting from the aperture problem and lacked any conflicting information. Thus,

an object rigidity constraint was ineffective when local measurements did not include sufficient orientation or direction of motion information. Under these conditions, the visual system relies on local computations performed on contours within different apertures, rather than on object rigidity across apertures. For example, "figural perception" (the identification of an object) does not seem to influence whether that object will be interpreted as rigid (Rock, 1981). Instead, the perception of rigidity seems to depend primarily on low-level measurements.

This lack of ability to integrate information across disconnected apertures does not always occur. For example, it is possible to demonstrate, under certain conditions, that the visual system is capable of integrating motion information when the object is translating. Moreover, in a recent study, Lappin, Norman, Loken, and Fukuda (1990) demonstrated that, in some cases, subjects are able to see a moving rigid square when it is viewed through separate apertures. There are several differences between Lappin et al.'s study and our study. The task in Lappin et al.'s study was to judge correlation. It is possible to learn this task even if the object also has elastic motion. Another difference is the type of motion. In the present study, the two-dimensional square was rotating smoothly and relatively slowly. The motion in the study by Lappin et al. was faster and much more localized. Local computation of motion is more difficult if the movement has high temporal frequencies. Further investigation of the differences between these two paradigms may reveal in more detail how the information from different locations is combined to form a unique percept.

Our results suggest that the human visual system must rely on low-level information more than previous theories have predicted. Because of ambiguities arising from the motion of unmarked edges, the visual system tends to disregard this information and to rely on the motion of corners, contour discontinuities, and texture. That is, the visual system may not give much weight to motion information measured from smooth contour segments because these segments only provide information about contour orientation. The visual system may, in fact, estimate which receptive fields are useful for motion interpretations by examining the outputs of several units with different orientations from the same location. Information from units that are indicated to be reliable is used to reconstruct the motion.

References

- Adelson, E., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*, 523-525.
- Braunstein, M., & Andersen, G. (1984). A counterexample to the rigidity assumption in the perception of structure from motion. *Perception*, *13*, 213-217.
- Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, *88*, 171-195.
- Fennema, C. L., & Thompson, W. B. (1979). Velocity determination in scenes containing several moving images. *Computers and Graphic Image Processing*, *9*, 301-315.
- Hildreth, E. (1984). The measurement of visual motion. Cambridge, MA: MIT Press.
- Hildreth, E., & Koch, C. (1987). The analysis of visual motion: From computational theory to neuronal mechanisms. *Annual Review of Neuroscience*, *10*, 477-533.
- Horn, B., & Schunk, B. (1981). Determining optic flow. *Artificial Intelligence*, *17*, 185-203.
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*, 161-180.
- Koenderink, J. J., & van Doorn, A. J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica Acta*, *22*, 773-791.
- Lappin, J. S., Norman, J. F., Loken, K. B., & Fukuda, H. (1990). The visibility of globally coherent motion behind multiple apertures. *Association for Research in Vision and Ophthalmology Abstracts*, *31*, 240.
- Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of moving retinal images. *Proceedings of the Royal Society of London B*, *208*, 385-387.
- Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proceedings of the Royal Society of London B*, *211*, 151-180.
- Meyer, G., & Dougherty, T. (1987, November). *Sawtooth pac people and the realization of illusory edges: Computational, cognitive, and utilitarian implications*. Paper presented at the 28th Annual Meeting of the Psychonomic Society, Seattle, WA.
- Nakayama, K. (1985). Biological image motion processing: A review. *Vision Research*, *25*, 625-660.
- Nakayama, K., & Silverman, G. (1988a). The aperture problem—I. Perception of nonrigidity and motion direction in translating sinusoidal lines. *Vision Research*, *28*, 739-746.
- Nakayama, K., & Silverman, G. (1988b). The aperture problem—II. Spatial integration of velocity information along contours. *Vision Research*, *28*, 747-753.
- Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature*, *317*, 314-319.
- Regan, D., & Beverley, K. I. (1979). Visually-guided locomotion: Psychophysical evidence for a neural mechanism sensitive to flow patterns. *Science*, *205*, 311-313.
- Rock, I. (1981). Anorthoscopic perception. *Scientific American*, *244*, 145-153.
- Schwartz, B., & Sperling, G. (1983). Luminance controls the perceived 3-D structure of dynamic 2-D displays. *Bulletin of the Psychonomic Society*, *21*, 456-458.
- Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, *29*, 619-626.
- Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London B*, *203*, 405-426.
- van Santen, J., & Sperling, G. (1984). A temporal covariance model of human motion perception. *Journal of the Optical Society of America*, *1*, 451-473.
- Wallach, H. (1935). Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung*, *20*, 325-380.
- Wallach, H. (1976). On perceived identity: I. The direction of motion of straight lines. In H. Wallach (Ed.), *On perception* (pp. 201-216). New York: Quadrangle/The New York Times Book Co.
- Wallach, H., & O'Connell, D. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*, 205-217.
- Watson, A., & Ahumada, A. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America*, *2*, 322-341.
- Waxman, A. M., Kamgar-Parsi, B., & Subbarao, M. (1987). Closed-form solutions to image flow equations for 3-D structure and motion. *International Journal of Computer Vision*, *1*, 239-258.
- Yuille, A. L., & Grzywacz, N. M. (1988). A computational theory for the perception of coherent visual motion. *Nature*, *333*, 71-73.

Received July 28, 1989

Revision received October 18, 1990

Accepted July 13, 1990 ■