

The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations

J. FARLEY NORMAN and JAMES T. TODD
Ohio State University, Columbus, Ohio

In two experiments, we evaluated the ability of human observers to make use of second-order temporal relations across three or more views of an apparent motion sequence for the perceptual analysis of three-dimensional form. Ratings of perceived rigidity were obtained in Experiment 1 for objects rotating in depth that were simultaneously subjected to sinusoidal affine stretching transformations along the line of sight or in a direction parallel to the image plane. Such transformations are theoretically interesting because they cannot be detected by analyses that are restricted to first-order temporal relations (i.e., two views), but they can be detected by more conventional analyses of structure from motion in which second-order temporal relations over three or more views are used. The current results show that human observers can perceive stretching transformations of a rotating 3-D object in a direction parallel to the image plane but that they fail to perceive stretching transformations along the line of sight. This result suggests that human observers can make use of some limited second-order temporal information. This finding was confirmed in Experiment 2, in which we investigated the effects of several specific optical consequences of sinusoidal stretching transformations applied in different directions. The results indicate that observers may be sensitive to the sign of acceleration, but that they cannot make use of the precise magnitude of second-order relations necessary to recover euclidean metric structure.

Human observers have the remarkable ability to perceive an object's three-dimensional (3-D) form from its projected pattern of motion within a 2-D visual image. This phenomenon, sometimes referred to as the *kinetic depth effect*, has been widely investigated by perceptual psychologists for over 60 years (Braunstein, 1962; J. J. Gibson & E. J. Gibson, 1957; E. J. Gibson, J. J. Gibson, Smith, & Flock, 1959; Green, 1961; Johansson, 1964; Johansson & Jansson, 1968; Metzger, 1934; Miles, 1931; Wallach & O'Connell, 1953). More recently, this capability has also attracted the attention of researchers in machine vision, who have developed working algorithms for computing an object's 3-D structure from moving 2-D optical patterns.

Most mathematical analyses for computing structure from motion are designed to operate on a small number of identifiable points across a small number of discrete temporal "views." The goal of these analyses is to discover whether there is a rigid (usually) 3-D structure compatible with the positions of those points over time. Ullman (1979) has shown that one can recover 3-D structure from orthographic projections given the motions of

four non-coplanar points moving rigidly over three successive views. Furthermore, it has been mathematically proven (Bennett, Hoffman, Nicola, & Prakash, 1989; Huang & Lee, 1989; Koenderink & van Doorn, 1991; Todd & Bressan, 1990) that a unique 3-D structure cannot be obtained from apparent motion sequences containing only two orthographic views of a moving 3-D object, except in certain specialized cases—as, for example, when a configuration of moving points is confined to a fixed plane (see Hoffman & Flinchbaugh, 1982; Lappin & Love, 1992). On theoretical grounds, it would therefore appear that three distinct orthographic views are both necessary and sufficient for determining 3-D structure from motion.

However, recent psychophysical evidence indicates that these theoretical analyses may have little relevance to the human perception of 3-D shape. Human observers effortlessly perceive 3-D shapes defined by two-view motion sequences, despite the theoretical ambiguity (Braunstein, Hoffman, & Pollick, 1990; Braunstein, Hoffman, Shapiro, Andersen, & Bennett, 1987; Todd, Akerstrom, Reichel, & Hayes, 1988; Todd & Bressan, 1990; Todd & Norman, 1991). Other researchers have produced percepts of 3-D surfaces by using scintillating optical patterns in which no individual point survives for more than two consecutive views (Doshier, Landy, & Sperling, 1989; Todd, 1985). Furthermore, evidence now exists to suggest that the visual mechanisms responsible for perceived 3-D shape may be limited to the information available from two successive orthographic views and cannot take

This research was supported by Grant 89-0016 from the Air Force Office of Scientific Research and BNS-8908426 from the National Science Foundation, the Office of Naval Research, and the Air Force Office of Scientific Research to J. T. We are grateful to Joseph S. Lappin for helpful discussion of issues relating to these experiments. Correspondence should be addressed to J. F. Norman, Department of Psychology, Ohio State University, Columbus, OH 43210.

advantage of the additional information available from a third view (Todd & Bressan, 1990; Todd & Norman, 1991).

Todd and Bressan (1990) and Todd and Norman (1991) have recently found that no significant improvements occurred as the number of views was increased from two to eight for a wide variety of psychophysical tasks. These tasks included discrimination of 3-D lengths, discrimination of angles between line segments oriented in 3-space, discriminations between planar and nonplanar configurations of lines, discrimination between rigid and nonrigid 3-D motion, judgments of amplitude in depth relative to the period of smoothly curved sinusoidal surfaces, and discriminations between differently curved ellipsoids. Other researchers (e.g., Hildreth, Grzywacz, Adelson, & Inada, 1990) have found improvements in performance as a function of the number of views. However, in the displays of Hildreth et al., each apparent motion sequence was presented once, so that the total viewing duration was longer with apparent motion sequences containing many views and shorter with sequences of fewer views. We have found that when this confound is removed and human observers are allowed similar observation times for both long and short apparent motion sequences, performance for tasks with two-view sequences can be as good as that for longer sequences. Observers are then able to maximally sample the limited available information.

If human observers fail to exploit the second-order relations available from three views and actually utilize only the first-order relations, we should take a closer look at what types of 3-D structures are specifiable from three views and what alternative types of 3-D structures, if any, are recoverable from shorter motion sequences. It is important to keep in mind that conventional analyses of structure from motion (see, e.g., Bennett & Hoffman, 1985; Hoffman & Bennett, 1986; Ullman, 1979) are designed specifically to recover an object's structure by computing the 3-D distance between each pair of visible points (see Norman & Todd, 1992). In practice, however, these models are susceptible to noise and may fail for very small measurement errors. Unlike human observers, their performance does not degrade gracefully, and they will not work if their underlying constraints (e.g., assumed rigidity) have been violated.

What type of information is available from two consecutive views of a moving 3-D object? Ullman (1977) showed that the projected trajectories of points undergoing a rigid rotation in depth can always be made parallel by rotating one of the two views in the image plane, but that this is not possible for points undergoing nonrigid motion—except for certain degenerate transformations whose probability of occurrence in natural vision is vanishingly small. Thus, if one can obtain parallel trajectories in the image plane following a rotation of a single view about the line of sight, it is safe to conclude that the object's physical motion in 3-space is rigid. Furthermore, Todd and Bressan (1990) and Koenderink and van Doorn (1991) have shown that a substantial amount of geometric information can

also be recovered. They show that two orthographic views, while insufficient to determine euclidean structure, do allow for the recovery of *affine* structure. Although a continuous family containing an infinite number of different 3-D structures is consistent with optical patterns composed of only two successive orthographic views, this does not mean that there is no useful information present. In particular, Todd and Bressan's analysis shows that, for small angular displacements, two views allow for the recovery of 3-D structure up to an affine stretching transformation along the line of sight (see Norman & Todd, 1992, for a further discussion of affine geometry).

Todd and Bressan's (1990) analysis would therefore predict that any two 3-D objects that could be related by a stretching transformation along the line of sight would be indiscriminable to human observers under two-view apparent motion. We refer to two such objects as being "affine equivalent along the line of sight." Notice that this is more restrictive than general affine equivalence. As an example, two triaxial ellipsoids (i.e., with three semi-axes of different lengths) are affine equivalent in general, since one can be made congruent to the other by a suitable combination of stretching transformations along differing axes. However, two such triaxial ellipsoids are affinely different in our more restricted sense, since they cannot be made congruent by a stretching transformation applied only along the line of sight. The information provided by two successive views can be surprisingly powerful, despite the affine ambiguity. From such first-order temporal relations, it is possible to discriminate between flat and curved surfaces since they are affinely different—a flat surface cannot be made curved through any type of uniform stretching transformation. Also, in general, two similar but differently curved surfaces, such as a paraboloid of revolution (changing curvatures) and a hemispherical shell (constant curvatures), will also be affinely different.

An appropriate way to further evaluate whether the visual perception of 3-D shape is based on euclidean or affine properties is to consider a task analysis. Most, if not all, of the common psychophysical tasks used to investigate perceived 3-D structure could be performed by an observer armed only with knowledge of an object's affine properties (see Norman & Todd, 1992; Todd & Norman, 1991). The main difference between euclidean and affine geometries is the ability in euclidean geometry to compare distance intervals oriented in different directions. In affine geometry, one can compare distance intervals in parallel directions (so one could bisect distance intervals accurately—see Lappin & Fuqua, 1983), but not in different directions. If human observers are able to accurately discriminate between the lengths of lines oriented in different directions, their knowledge of 3-D structure must be based on euclidean properties.

Todd and Bressan (1990) found that observers performed poorly on tasks requiring detection of euclidean properties, such as discriminating between 3-D lengths oriented in different directions or discriminating whether

the angle formed between two 3-D line segments was greater than or less than 90°. Moreover, performance did not improve as the number of views was increased beyond two, suggesting that the judgments were not based on euclidean properties. In contrast, their observers accurately discriminated between planar and nonplanar configurations of line segments, a task requiring only the detection of affine structure. Todd and Norman (1991) subsequently confirmed another prediction of Todd and Bressan's affine structure-from-motion analysis. They found that a pair of ellipsoids differing only by a relative expansion along one axis were discriminable from two-view apparent motion sequences, except when viewed from a degenerate orientation where the objects were affine equivalent along the line of sight, in which case the discrimination became nearly impossible.

The present paper concerns an especially strong test of the hypothesis that the visual mechanisms responsible for perceived 3-D shape are limited to detecting whatever geometric information is obtainable from two successive orthographic views. Specifically, if at every transition from one view to the next there is a continual affine ambiguity such that the visual system cannot distinguish between different objects related by an expansion along the line of sight, it should be possible to create an extended apparent motion sequence of an extremely nonrigid shape deforming as it rotates that would appear to an observer to be completely rigid. That is, if observers do not take advantage of the information provided by many-frame displays, and if their knowledge of three-dimensional shape is based solely on the information present within two successive views, then, at every moment in time, an observer can only determine an object's structure to within an arbitrary stretching factor along the line of sight. If this is correct, we should be able to arbitrarily expand or contract a shape along the line of sight at every frame transition before rotating to the next frame, and the shape should appear unchanged. The global optical pattern over many frames would not be consistent with the rigid rotation of any 3-D shape because of the cumulative effect of the multiple stretching transformations. Nevertheless, Todd and Bressan's (1990) two-view analysis would predict that this global display would be perceived as a rigidly rotating 3-D figure. The present series of experiments was designed to test this hypothesis.

EXPERIMENT 1

Method

Stimulus displays. The stimuli for this experiment were all composed of connected, randomly oriented line segments, similar to the original figures used by Wallach (Wallach & O'Connell, 1953). The wire-frame figures consisted of 12 vertices, which were randomly placed within a cubical volume of 3-space (1,000 cm³; i.e., 10 × 10 × 10 cm). The 3-D figures were rotated in depth either about a Cartesian vertical axis or an axis slanted out of the image plane while undergoing either of two types of nonrigid stretching deformations. In one condition, the rotating 3-D figures were subjected to a nonrigid stretching deformation along an observer's line of sight before rotating to a new orientation, which had a cumulative effect

on the resulting optical flow field, since the objects were not "unstretched" following the rotation. In a different condition, the 3-D figures were subjected to an identical nonrigid deformation, except that this deformation was applied in a direction perpendicular to each observer's line of sight. In particular, for any given frame transition, the 3-D figure was stretched by a scaling factor that varied sinusoidally over time:

$$\text{scaling_factor} = 2.0^{\text{amp}} * \sin(10\theta i), \quad (1)$$

where i represents the frame number across time. The amp parameter controls the amplitude of the stretching deformation. Theta is a constant angle needed for one complete revolution of the figure over the total number of frames. Taking the sine of 10 times that angle indicates that the frequency of the stretching is 10 times higher than the rotation period. The 3-D objects were stretched according to the magnitude of the scaling factor, such that

$$z = z/\text{scaling_factor} \quad (2)$$

or

$$x = x/\text{scaling_factor}, \quad (3)$$

depending on whether the stretching transformation was applied along the line of sight (z -axis) or applied in a horizontal direction in the image plane (x -axis). Following the structural deformation, the surface was rotated in depth. For a rotation about a Cartesian vertical axis, the following transformation was applied:

$$x' = [x * \cos(\theta)] - [z * \sin(\theta)] \quad (4)$$

$$y' = y \quad (5)$$

$$z' = [x * \sin(\theta)] + [z * \cos(\theta)], \quad (6)$$

where $\theta = 3.6^\circ$ (the object made one complete revolution in 100 frames). Similar transformations were used to rotate the figures around an axis slanted with respect to the image plane.

Since the stretching transformations indicated by Equation 2 were always applied only along the z -axis, Todd and Bressan's (1990) affine analysis of structure from motion would predict that the effects of this transformation would not be visible. This two-view affine analysis cannot distinguish between any structures that differ by a stretching transformation along the z -axis.

The apparent motion sequence was composed of 100 individual frames, during which the figure made one complete revolution through 360°. After the figure rotated from the orientation in the first view to the orientation in the last view (the 100th frame), the figure reversed direction and rotated back to the orientation shown in the first view. The figure continuously oscillated in this manner until the observer made an appropriate response.

Apparatus. These apparent motion sequences were displayed on a Silicon Graphics Personal Iris (4D/25 with Turbo graphics) workstation. The displays were viewed monocularly through a viewing hood. The viewing distance was 76.0 cm, such that the 1,280-pixel-wide × 1,024-pixel-high viewing screen subtended 25.22° × 20.29° of visual angle. The 11 line segments composing the random wire-frame figures were 5 pixels wide, each subtending approximately 5.9' of arc. The wire-frame figures were presented as a set of white lines against a dark black background. The rotation of the figure from one view to the next was updated at 30 Hz (i.e., the temporal duration of each view was 33.3 msec).

Procedure. While piloting this experiment, we noticed considerable variation in the appearance of the displays. Depending on the specific display parameters, the depicted objects could appear to rotate rigidly in depth with no detectable deformation; they could appear to rotate and deform simultaneously; or they could appear to undergo severe deformation with no detectable rotation. To measure how each display was perceived within this continuum of possibilities, observers rated the apparent rigidity of the depicted object's motion in 3-D space using a rating scale from 1 to 9. A rating

of 9 indicated *perceived rigid rotation with no detectable deformation*; a rating of 1 indicated *perceived deformation with no detectable rotation*; and a middle rating of 5 indicated *equal amounts of perceived 3-D rotation and deformation*. The observers made their estimates on the numeric keypad on the Silicon Graphics keyboard.

There were a total of 12 experimental conditions obtained by the orthogonal combination of two axes of rotation (Cartesian vertical or an axis slanted 30° out of the image plane), two axes of stretching (along the observer's line of sight or perpendicular to the observer's line of sight), and three magnitudes of the sinusoidal stretching (low, medium, and high; the amp parameter in Equation 1 was 0.2, 0.4, or 0.6, respectively).

To better appreciate the large cumulative effect of these stretching transformations, it is useful to consider how the distances of a pair of points from the axis of rotation change over time for a representative motion sequence as shown in Figure 1. The amp parameter used to generate these trajectories was 0.4, the medium stretching magnitude used in the actual experiment. One can readily see that the structural deformation is quite large. The upper point's radius varies from about 3.6 cm to nearly 8.0 cm.

To quantify the nature of the sinusoidal stretching transformations, we calculated two measures, the average deviation and the maximum deviation relative to the mean radius. These measures are defined as

$$\text{Average deviation} = \frac{\sum_{i=1}^{100} |\text{radius}_i - \text{mean_radius}|}{100} \quad (7)$$

$$\text{Average deviation relative to mean} = \frac{\text{Average deviation}}{\text{mean_radius}} \quad (8)$$

The maximum deviation relative to the mean is a simpler measure. It is the single maximum deviation from the mean radius divided by the mean radius.

We calculated these two measures for amp parameters of 0.2, 0.4, and 0.6. Each number reflects the mean of six individual measurements, the motions of six points over time. Mean average deviations (relative to the mean radius, proportions) were 0.07, 0.15, and 0.24 for amp parameters of 0.2, 0.4, and 0.6, respectively. Similar mean maximum deviations were 0.22, 0.49, and 0.78. For example, the mean maximum deviation in radius for a 0.6 magnitude stretching transformation was 78% of the mean radius—in other words, a severe structural deformation. Note that these numbers reflect the single maximum or average deviation (absolute value) in radius—the peak-to-trough deviations would be approximately double the reported quantities.

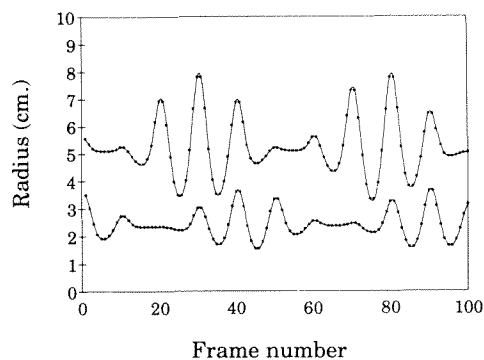


Figure 1. The radii of two points from the axis of rotation, stretched sinusoidally along the line of sight. The depicted deformations are typical of those used in Experiment 1.

Each observer participated in two experimental sessions. Each session consisted of five trials for each of the 12 experimental conditions. Within each session, all 12 conditions were shown in random order before going to the next set of 12 conditions. Ten ratings were therefore obtained after completion of the two experimental sessions.

Observers. Five members of the academic community at Brandeis University served as observers. Among them were the two authors, a professor, and two graduate students. All had participated in prior psychophysical experiments. The two graduate students were naive to the purpose of the experiment.

Results and Discussion

The data combined for all observers are shown in Figure 2. For both rotations about a vertical axis and for an axis slanted out of the image plane, the motions of the objects that were stretched along the observers' line of sight appeared perfectly rigid. This perceived rigidity was maintained for all magnitudes of the nonrigid stretching deformation. However, the figures that were nonrigidly stretched in a direction perpendicular to the line of sight were perceived as nonrigid, with the perceived rigidity decreasing as the magnitude of the stretching deformation was increased. Note, in particular, that for the largest stretching parameters employed in this experiment, the perceived deformations perpendicular to the line of sight were so severe that the rotation component of the displays was barely detectable.

We have confirmed these findings informally over a wide variety of different conditions, including displays depicting random configurations of dots, smoothly curved surfaces, or regular polyhedra (see Norman & Todd, 1991) and with varying rates of stretching and rotation. For all these variations, the appearance of the depicted motion is identical. Objects stretched along the line of sight appear rigid, whereas objects stretched perpendicular to the line of sight appear nonrigid.

In describing the displays, all the observers reported that the 3-D figures stretched along the line of sight appeared to be rotating rigidly with an angular velocity that varied sinusoidally over time. It is important to recognize, however, that this interpretation is not valid. No possible rigid interpretation exists that is compatible with the motions of the vertices of the figure across the 100 views in the motion sequence. Any conventional model of structure from motion that uses the information available from three successive views would be able to track the changing structure of the deforming objects as they rotate.

To demonstrate that these three-view models would be able to detect the nonrigid deformation of the stretch-in-z-axis displays (while human observers do not), we implemented an algorithm developed by Hoffman and Bennett (1986). This algorithm (section 6, "Fixed-Axis Motion: Angular Acceleration") takes as input the projected positions of a set of points in the image plane over three frames. From these positions in the image plane, it recovers the radii of the points relative to the axis of rotation. We input to our implementation of Hoffman and Bennett's model the projected motions of a rigid 3-D figure that ac-

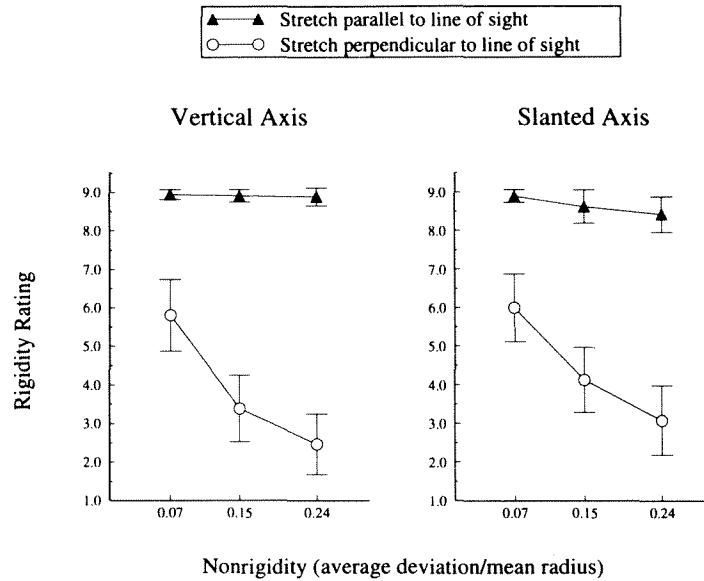


Figure 2. Mean rigidity ratings of the observers in Experiment 1 for each of the two different types of deformations as a function of the actual structural rigidity.

celerated and decelerated during its rotation (the rotation increment θ varied sinusoidally over time). From these projected positions, the algorithm was able to accurately recover the radius of each of the figure's constituent points. First, the positions of several points in Frames 1, 2, and 3 were input to the model. On the basis of these image positions, the radii of the points were calculated. Then, the positions of the same points for Frames 2, 3, and 4 were input, and the radii were again recovered. For every set of three frames, the model calculated the 3-D structure of the points. The recovered radius of each point was stable over time, despite the accelerations and decelerations. Furthermore, the recovered radii were in good quantitative agreement with the radii of the original moving 3-D vertices that generated the 2-D projections.

After establishing the accuracy of our implementation of Hoffman and Bennett's (1986) model, we input the 2-D projected positions of the nonrigidly rotating figures that were stretched along the line of sight. (To satisfy the assumptions of this special case analysis, we used rotations around a vertical axis in the image plane.) The output of the model for the motions of the stretch-in-z displays is shown in Figure 3. As can clearly be seen, the model tracked the changing 3-D structure over time. A comparison with Figure 1, which shows the actual structural deformation of the stretch-in-z displays, indicates that the radii output from Hoffman and Bennett's model vary in the same manner as the variations in the actual 3-D structure used to produce the 2-D optical patterns. However, the magnitude of the variations in radii output from the model are much higher than the actual structural deformations. Using the projected image positions of a nonrigidly rotating figure (with amp parameter of 0.003,

much lower than the 0.2-0.6 used in the experiment), their model produced the varying radii shown in Figure 3. The model detects the nonrigidity but is overly sensitive and exaggerates the variations in structure. The fact that this model cannot track the deforming structure perfectly is not surprising, since it was only designed to be used with rigid motions. The main point to keep in mind is that it can successfully determine the structure of a rigidly rotating object even when the rate of rotation varies over time, and that it can be used, at least qualitatively, to detect the presence of a nonrigid transformation.

At this stage in our investigation, we were convinced that the results provided an unusually strong confirmation of the theoretical model proposed by Todd and Bresnan (1990) and Todd and Norman (1991). This model predicts that the perceptual analysis of structure from mo-

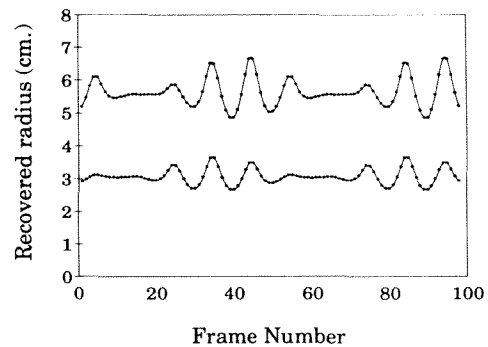


Figure 3. Output of Hoffman and Bennett's (1986) algorithm for the projected image motions of two points that are sinusoidally stretched along the line of sight. Note that the model can successfully detect the structural nonrigidity of the rotating 3-D figures.

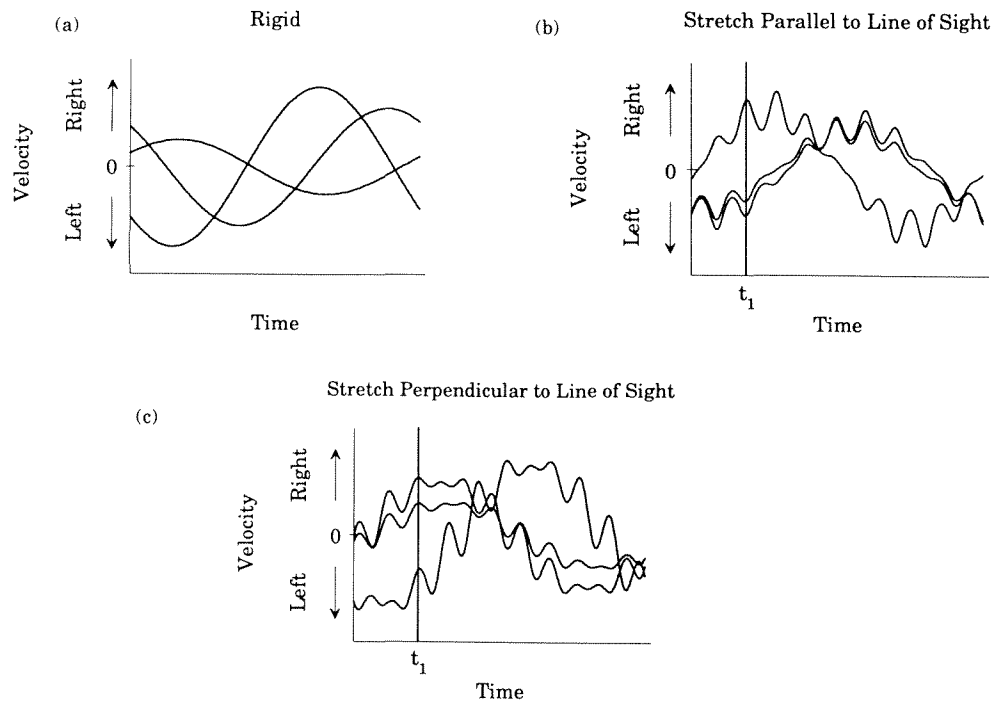


Figure 4. Image velocity of several vertices plotted over time for (a) rigidly rotating objects, (b) objects that are sinusoidally stretched along the line of sight before each rotation, and (c) objects that are sinusoidally stretched in the image plane perpendicular to the line of sight before each rotation.

tion is restricted to first-order temporal relations within two view sequences, and that because of this limitation, it should not be possible to detect stretching transformations along the line of sight. Since a three-view analysis of euclidean structure would be able to detect such transformations, the results of this experiment would appear to indicate that human observers can only make use of the available information within two-view sequences.

On closer examination, however, other aspects of the data are inconsistent with this conclusion. The inconsistency arises from the observers' perceptions of non-rigid motion when the objects were stretched perpendicularly to the line of sight. Any model that is only sensitive to two-view relations would also predict that the stretches perpendicular to the line of sight would also appear rigid. To see why this is true, let us consider how the Ullman (1977) rigidity test would respond to the motions of objects stretched both parallel to and perpendicular to the line of sight. Since both of these stretching transformations produce optical patterns where the points have parallel trajectories between every pair of two views, rigid interpretations do exist for both types of displays. At every transition between views, both types of displays would "pass" Ullman's test for rigidity. The result obtained from this experiment, namely that the stretches parallel to the line of sight appear rigid while similar deformations perpendicular to the line of sight appear nonrigid, is inconsistent with either a three-view analysis or a purely two-view analysis. This outcome suggests that some lim-

ited information exists over time which enables the visual system to detect the nonrigidity in one case, but not in the other (see also the similar findings of Todd, 1982).

Faced with this pattern of results that are neither consistent with conventional analyses of structure from motion nor fully consistent with the affine analysis of structure from motion as proposed by Todd and Bressan (1990), we decided to more closely examine the properties of the 2-D optical patterns over many frames of the motion sequence. In Figure 4, velocity is plotted in the horizontal direction as a function of time for (a) objects whose vertices are rotated rigidly around a vertical axis, (b) objects whose vertices are sinusoidally stretched along the line of sight while rotating around a vertical axis, and (c) objects whose vertices are sinusoidally stretched along a horizontal axis in the image plane, perpendicular to the line of sight while rotating around a vertical axis. Each curve in all three plots represents the velocities of a separate vertex over time. Velocities for only three vertices were plotted to more clearly show the relationships between the different conditions.

Several interesting patterns emerge in these velocity profiles. Rigid rotation in depth about a vertical axis (as shown in Figure 4a) leads to pure sinusoidally varying velocities along the horizontal direction in the image plane. The effects of the high-frequency stretching transformations are clearly evident in panels b and c of Figure 4. The basic low-frequency sinusoidal variations characteristic of rotation in depth have been

amplitude modulated by the high-frequency stretching transformations. There are two salient differences between the velocity profiles of the stretches parallel to the line of sight (shown in Figure 4b) and the velocity profiles for the stretches perpendicular to the line of sight (shown in Figure 4c). The variations in velocities of vertices that are stretched along the line of sight are all in phase—they all speed up or slow down together (see time t_1 for an example). This relation does not exist for stretches perpendicular to the line of sight. In Figure 4c, some vertices are speeding up while others are slowing down.

Another salient difference in the optical patterns caused by stretching in different directions concerns the pattern of amplitude modulation. For stretches parallel to the line of sight, maximal amplitude modulation occurs at the highest velocities, when the rotating vertices cross the line of sight. The opposite is true for stretching transformations in the image plane. In this condition, the maximal amplitude modulation occurs when velocities are slowest, at the extreme end of their 2-D trajectory.

Are either of these two major differences in velocity profiles between stretches parallel to and perpendicular to the line of sight responsible for the perceived nonrigidity of the stretch-in- x display? That is, do the stretches perpendicular to the line of sight appear nonrigid because of differences in either phase relationships between the motions of individual vertices, or patterns of amplitude modulation? This issue is examined in detail in Experiment 2.

EXPERIMENT 2

In this experiment, we examined whether phase relationships among the motions of individual points or type of amplitude modulation for a given point's motion would contribute to perceived rigid motion. Phase and amplitude modulation are concepts that are ill or undefined when only two consecutive views are analyzed. Differences in either phase or pattern of amplitude modulation in the image may be responsible for the large differences in perceived rigidity that exist between the two types of stretching deformations used in Experiment 1.

If the significance of phase and amplitude modulation for the perception of rigidity is to be evaluated properly, it is important to control these variables directly. Toward this end, we created optical displays that were not based on the 2-D projections of a 3-D figure rotating in depth, but in which velocities were directly manipulated as a function of time. This is similar in some respects to the earlier work of Johansson (1964) and Todd (1982), who also directly varied 2-D velocities to create the perception of objects rotating in 3-D space.

Amplitude and phase differences arise in the 2-D projections of a rotating 3-D object as a result of the object's constituent points' having different positions relative to the axis of rotation. Figure 5 shows how these amplitude and phase differences arise. We adopt a conventional coordinate system where x and y represent horizontal and vertical directions in the image plane, respectively, and

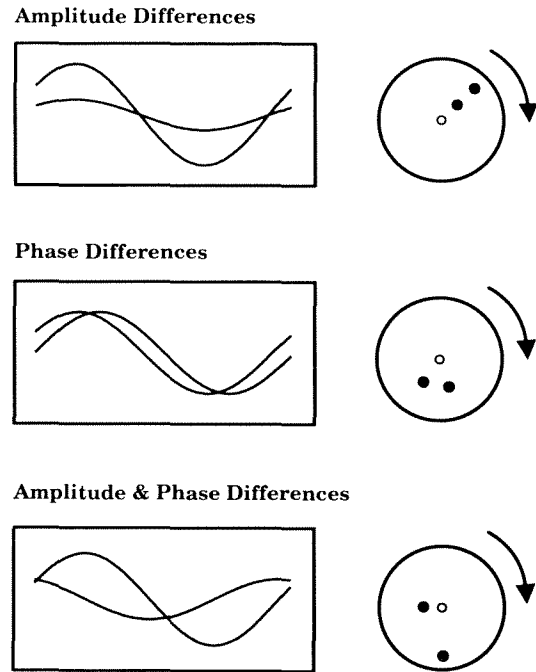


Figure 5. Horizontal image velocities (left half of figure) for each of the rotations around a vertical axis depicted in the right half of the figure. The vertical rotation axis is illustrated by an open circle. Amplitude differences in the image arise as a result of points having differing distances from the rotation axis. Phase differences in the image reflect that the points occupy differing positions in their orbits around the rotation axis. The general situation is shown at the bottom, where differing radii and orbit positions lead to both amplitude and phase differences in the image.

the z direction indicates the perpendicular dimension in depth. The right portion of the figure schematically illustrates two points rotating in depth about the vertical, or y -axis (shown by the open circle). The plots in the left half of the figure show horizontal velocities (i.e., along the x -axis in the image plane) over time for each of the three 3-D situations depicted in Figure 5. Different amplitudes (top panel) arise in the velocity profiles as a result of rotating points that differ in their distances from the axis of rotation. Differing phases (middle panel) indicate that the points are at the same distance from the axis of rotation, but at different positions in their orbits. A normal situation is depicted in the lower panel, where both amplitude and phase differ because of initial differences in orbit position and distances to the axis of rotation. A wider variety of comparisons between 3-D arrangements of points and their resulting velocity profiles following rotation in depth can be found in Metzger (1934).

Method

Stimulus displays. Motion sequences were created for this experiment by generating for each vertex a displacement function with randomly chosen amplitudes and phases, simulating the rotation of an actual 3-D figure. In particular, we created optical patterns based on the following equations. The horizontal position of a point rotating around a vertical axis in depth is a sinusoidal function of time:

$$x = A * \sin(\omega t + \phi), \quad (7)$$

where A controls the amplitude, ω controls the frequency, and ϕ controls the initial phase of the sinusoidal waveform. Differentiating this function with respect to time gives a relation describing how the horizontal velocity varies over time:

$$dx/dt = A\omega * \cos(\omega t + \phi). \quad (8)$$

Generating these 2-D velocity (or displacement) functions directly in this manner made them easy to modify, and, thus, the creation of optical patterns with unusual properties became possible. The effects of phase and amplitude variations that covaried together in the two different stretching conditions employed in Experiment 1 could now be separated.

The base motions of each of the displays' vertices were constructed by using Equations 7 and 8. These base waveforms were then modulated (i.e., multiplied) by three different types of high-frequency "noise" waveforms to create five different types of optical patterns. These three types of noise are shown in Figure 6—in-phase noise variations (top), random counterphase noise variations (middle), and random amplitude noise variations (bottom). The amplitudes of the noise signals are plotted over time. For three of the five types of optical patterns, where the pattern of amplitude modulation was similar to the stretch-in-z displays of Experiment 1, these noise signals (which varied about the magnitude 1.0, unlike normal sinusoids, which vary about zero) multiplied the base low-frequency sinusoidal velocities of each vertex. For two other conditions, where the pattern of amplitude modulation was similar to the stretch-in-x displays of Experiment 1, the noise signals (in phase and random counterphase) multiplied the *positions* of each vertex in the image plane rather than the velocities. Examples of final velocity functions for each of the five different optical patterns after applying these noise modulations are shown in Figure 7. One can readily see that they do not show a close similarity to normal velocity variations in the image plane following the rigid rotation in depth of a 3-D object.

The in-phase condition resembles the stretch-in-z displays used in Experiment 1. All high-frequency velocity variations are in phase—the vertices accelerate or decelerate at the same time. The amplitude modulation (effects of the noise) is greatest at the highest absolute velocities. The counterphase condition preserves the amplitude modulation while disrupting the relative phase relationships of the noise. A random half of the vertices have their phases of noise modulation shifted by 180° (see Figure 6). Notice that some vertices are speeding up while others are slowing down. A comparison between this condition and the in-phase condition isolates and provides a test for the importance of the relative phase of the high-frequency variations.

The next two conditions, counteramplitude modulation and random amplitude, evaluate the importance of the amplitude modulation. The counteramplitude modulation condition disrupts the pattern of amplitude modulation characteristic of the stretch-in-z displays used in Experiment 1. In this case, the effects of the high-frequency noise are greatest when the absolute velocities are slowest. However, it is impossible to create a display with an opposite pattern of amplitude modulation that does not also introduce temporary disruptions of phase (see Figure 7). This occurs because of the large variations in velocity at the slowest speeds. Therefore, another test of amplitude was developed that preserved phase relationships better. The random amplitude condition provides a stronger test of the importance of amplitude of noise variations. In this condition, the amplitude of every cycle of the high-frequency noise for every vertex was chosen at random from some interval (see Figure 6). Modulating normal sinusoidal base motions with this type of noise signal produces a very unusual optical pattern (see Figure 7). The velocity variations from the normal sinusoidal motions that exist across the different vertices (the amplitude of the modulating signal may sometimes be zero) are all in phase. However, the motions of each

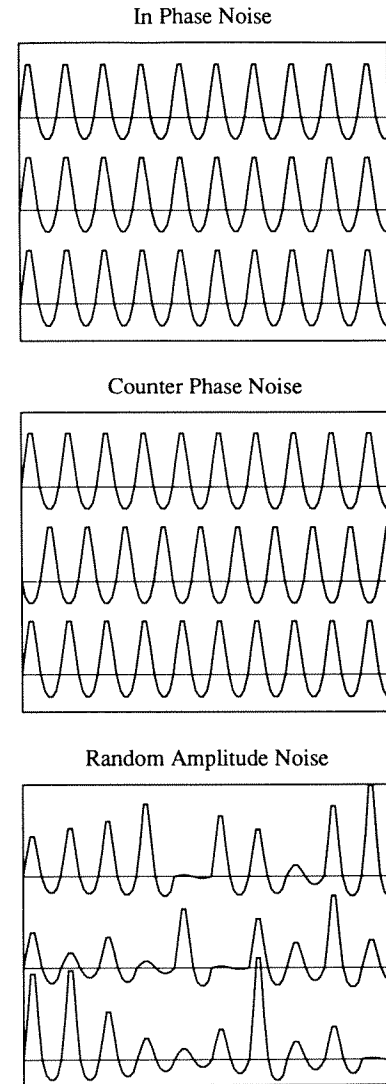


Figure 6. Three types of high-frequency noise signals used to modulate the basic sinusoidal image velocities. Each panel shows noise signals over time for three different vertices in the optical pattern. The top panel shows an example of in-phase noise variations, where the noise increases or decreases in magnitude at the same time for all vertices. The middle panel illustrates counterphase noise signals, where a random half of the vertices have their noise waveforms shifted in phase by 180° relative to the remaining noise signals. Some noise signals are increasing in magnitude, while others are simultaneously decreasing. Random amplitude noise is shown in the lower panel, where the amplitude of the noise signal is chosen at random within some range for every cycle and every vertex in the optical pattern. The magnitudes of the noise signals are thus uncorrelated with the noise waveforms of the other vertices at the same time.

individual vertex have been highly distorted by the random amplitude noise. Sometimes each vertex may be little influenced by the noise, whereas at other times it is influenced a great deal—its motion is unusual and is characterized by sudden accelerations and decelerations that are uncorrelated with the motions of the other vertices. The final condition is counteramplitude modulation with random counterphase. This condition combines two of the previous manipulations and severely disrupts phase and amplitude modulation at the same time.

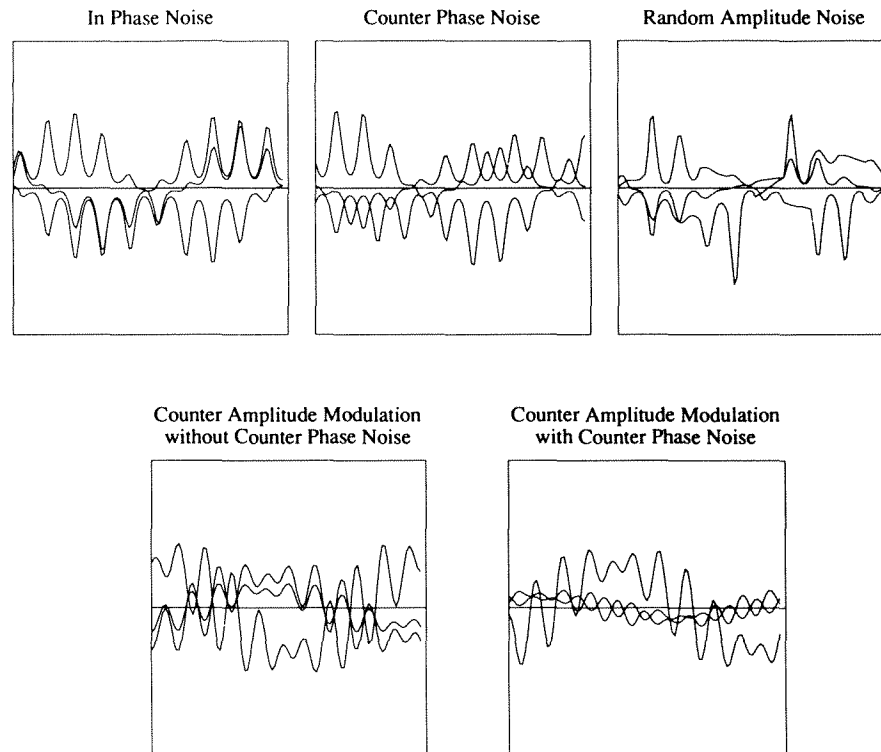


Figure 7. Image velocities of three points as a function of time after the basic low-frequency velocities characteristic of the kinetic depth effect have been modulated by the noise waveforms illustrated in Figure 6. The velocity plots at the top of the figure (amplitude modulation similar to the stretch-in-z displays of Experiment 1) were created by multiplying the low-frequency sinusoidal velocities by the three noise signals depicted in Figure 6. The two bottom velocity plots (amplitude modulation similar to the stretch-in-x displays of Experiment 1) were created by multiplying the positions of the vertices by the top (in-phase) and middle (random counterphase) noise signals of Figure 6, rather than multiplying velocities.

Optical patterns produced by all of these manipulations would be detected as nonrigid by conventional models of structure from motion that utilize three view relations. The results of Experiment 1 demonstrated that human observers fail to integrate much of the information available beyond two views, since the stretch-in-z displays were perceived as rigidly rotating 3-D objects. However, the perceived nonrigidity of the stretch-in-x displays in Experiment 1 suggests that some type of global information about rigid motion can be recovered from many-frame apparent motion sequences. The discussion section of Experiment 1 highlighted two salient differences between the stretch-in-z displays and the stretch-in-the-image-plane displays. These factors covaried together in the displays of Experiment 1. Which of these five noise conditions used in Experiment 2 appear rigid and which appear nonrigid should indicate which of these global 2-D properties is responsible for the perceptual difference.

The stimulus displays were superficially similar to those used in Experiment 1. Eleven connected line segments (12 vertices, each line 4 pixels wide) formed a simulated 3-D object that appeared to rotate around a Cartesian vertical axis. The motions were generated by Equations 7 and 8, where the phase ϕ of each vertex was randomly chosen between 0° and 360° , the amplitude A was chosen at random between 0 and 5.0 cm, and the vertical position in the optical pattern (y -axis in the image plane) was randomly chosen between -5.0 and $+5.0$ cm. (the center of the display screen was $x, y = 0, 0$). The "base" sinusoidal motions generated by Equations 7 and 8 were subsequently modulated (i.e., multiplied) by the

three different noise waveforms shown in Figure 6 to produce the in-phase, random counterphase, and random amplitude conditions. The counteramplitude modulation and counteramplitude modulation with random counterphase conditions were produced by multiplying the positions of the vertices by the in-phase and random counterphase noise signals rather than multiplying the velocities as in the first three conditions. Typical final velocity profiles for each of these conditions are shown in Figure 7.

The apparent motion sequence was composed of 100 individual frames during which the simulated 3-D figure rotated through 360° . Ten cycles of the high-frequency noise modulated this 100-frame motion sequence. The simulated figure oscillated between Views 1 and 100, in such a way that when Frame 100 was reached, the simulated figure reversed its direction of rotation to return to Frame 1. The figure continuously oscillated in this manner until the observer made a response. All other details of the equipment and displays were identical to those used in Experiment 1.

Procedure. The psychophysical task was identical to that used in Experiment 1, except that estimates of perceived rigidity were obtained by using a continuous rating scale from 1 to 9 to indicate the relative perceptual salience of the rotation and deformation components of each display. The observers made their ratings by sliding the computer's mouse—the movement of the mouse was yoked to the movement of a pointer that slid along a visible rating scale presented below the rotating figure. The observers adjusted the position of the mouse until they were satisfied with the rating, then pressed a button on the mouse to initiate a new trial.

There were a total of 20 experimental conditions obtained by the orthogonal combination of five different types of optical patterns and four magnitudes of nonrigidity. The magnitude of each display's nonrigidity was expressed by summing the displacement differences between each vertex's base sinusoidal motion and its motion after modulation by the various noise conditions. These sums for each vertex were themselves summated across vertices to arrive at a single measure expressing how different that optical pattern was from a valid 2-D orthographic projection of an object rotating in depth. Prior to the start of the actual experiment, the relationship between the amplitude of each qualitative type of modulating noise and this measure of nonrigidity was empirically calculated. A mean of these summated displacement differences was calculated for each type of noise by using 50 different randomly determined optical patterns. Amplitudes for each of the noise conditions were used which corresponded to mean summated displacement differences of 75, 100, 125, and 150 cm. These measures reflect the four levels of nonrigidity used in the experiment. A zero would indicate no displacement differences from pure sinusoidal motion of the vertices—that is, rigid rotation in depth. For a given level of nonrigidity, say 100 cm, all displays, despite having qualitatively different motions, would have the same overall amount of nonrigidity (i.e., deviation from normal sinusoidal displacements).

Each observer participated in two experimental sessions. Each session consisted of 10 trials for each of the 20 experimental conditions. The different noise conditions were presented in random order. After completion of both sessions, a total of 20 ratings were obtained for each observer for each experimental condition.

Observers. The observers were five members of the academic community at Brandeis University. The five included the two authors, one professor, one research associate, and one graduate student. The research associate and graduate student were naive to the purpose of the experiment.

Results and Discussion

The results for all observers combined are shown in the left panel of Figure 8. The mean rigidity ratings are plotted as a function of increasing nonrigidity, with each

separate curve representing ratings from a different type of noise modulation. For all observers, there is a clear separation of the curves into two groups: one containing the in-phase and random amplitude modulation conditions (solid symbols), the other containing the counterphase and counteramplitude modulation conditions (open symbols). It is interesting that there was little overlap in the observers' responses across these two groups. This is shown in the right panel of Figure 8, in which the different conditions of each group have been collapsed, with error bars to show the relative variances of the observers' judgments.

Note in these figures that the appearance of rigidity was greatest for the in-phase and random amplitude modulation conditions, which preserve the phase relationships among the high-frequency noise variations. The apparent rigidity of the in-phase displays is not surprising, since they were basically identical to the patterns of optical motion that would be produced by objects rotating rigidly in depth with a sinusoidally varying angular velocity. However, it is important to keep in mind that no such rigid interpretation exists for the random amplitude modulation displays, because of the uncorrelated amplitude variations across the different vertices. The fact that these displays appear rigid provides strong evidence that observers are insensitive to relative amplitudes of the high-frequency modulation functions.

The situation for the remaining three conditions, all of which disrupted phase relationships to varying degrees, was quite different. Displays in these three conditions were perceived to be highly nonrigid, although the rotation component of each display was clearly visible—that is, no mean ratings near 1.0 were obtained. The fact that observers were able to detect the presence of nonrigid distortions in these displays suggests strongly that they were

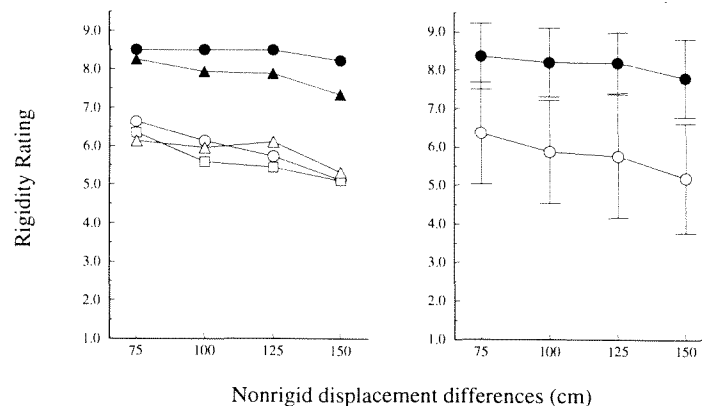


Figure 8. Mean rigidity ratings of the observers in Experiment 2 for five types of optical deformation. The filled circles and filled triangles in the left panel represent the in-phase and random amplitude conditions, respectively, in which the high-frequency noise variations were always in phase. The open squares, and open triangles, in contrast, represent the random counterphase condition and the counteramplitude modulation conditions with and without random counterphase, in which the high-frequency modulations of different vertices were not in phase. The overall means for the in-phase (filled circles) and out-of-phase conditions (open circles) are presented in the right panel with error bars, to indicate the standard deviations of the observers' judgments.

sensitive to relative phase relationships within the moving patterns.

To summarize briefly, the results for all five conditions are consistent with the hypothesis that observers are sensitive to the relative phases but not the relative amplitudes of high-frequency noise modulations of a sinusoidal carrier motion. It is important to keep in mind while considering these findings that the displays were equated in terms of actual 3-D nonrigidity—that is, the average deviation from pure sinusoidal motion characteristic of rotation in depth was identical in all conditions. Moreover, we also performed an additional analysis to show that the observers' responses could not have been determined by the magnitude of 2-D nonrigidity in each display. To test for this possibility, we measured the average variance of 2-D distance across all views within the motion sequence for all pairs of vertices and conditions and correlated these measures with the observers' rigidity ratings. The analysis revealed that perceived rigidity actually *increased* slightly with the magnitude of 2-D nonrigidity, although this trend was not statistically significant (Pearson $r = .389$, 18 df , $p > .05$).

GENERAL DISCUSSION

Many mathematical analyses now exist which show that recovery of complete 3-D structure can occur from the discrete 2-D orthographic projections of a set of moving 3-D points. Recovery of structure is possible if projective correspondences can be established over three views and if the points' 2-D positions can be measured with sufficient accuracy. How many points are required depends on the specific assumptions of each particular model (four points for Ullman's [1979] model, which assumes rigid motion; two points for Hoffman and Bennett's [1986] model, which assumes a fixed or stationary axis of rotation). These analyses also show that two views by themselves are ambiguous and cannot specify a unique 3-D structure. An object's "structure" is typically defined in this context by specifying the positions of each of its constituent points within a polar or Cartesian coordinate system. These theoretical analyses make clear predictions about the necessary and sufficient conditions under which recovery of this kind of 3-D structure can occur. Recovery should be impossible, given only two orthographic views, and it should improve dramatically as the number of views is increased beyond two.

Recently, psychophysicists have attempted to evaluate whether the human perception of 3-D structure from moving patterns behaves as these computational models would predict. The computational models posit that the type of 3-D structure that is perceived is euclidean—therefore, three views are necessary to support that perception. One can test this basic question in either of two ways. First, attempt to determine if the underlying representation is euclidean by using psychophysical tasks that require knowledge of euclidean structural relations. If observers can perform such tasks, presumably their knowledge derives from a representation of euclidean quantities. A sec-

ond way to test whether euclidean relations form the basis for human perception is to not satisfy the necessary conditions. If three views are required, have observers perform tasks with only two views. Then compare that performance with that obtained with longer motion sequences. If the computational models are psychologically valid, performance should be poor with two views and should improve rapidly with additional views.

Both of these techniques to evaluate whether the human perception of 3-D shape depends on detection of euclidean structure have been utilized. For example, Todd and Bressan (1990) tested the ability of observers to compare distance intervals in different directions in 3-D, which requires knowledge of euclidean relations—it cannot be done if observers only have access to affine geometrical properties. They found that observers could not make these kinds of judgments with any reasonable accuracy, and performance did not improve as the numbers of views was increased from two to eight. However, Todd and Bressan (1990) and Todd and Norman (1991) found that other tasks that were possible given knowledge of affine geometrical properties could be performed accurately. In no case was performance substantially improved by adding additional views beyond two. Whatever property was being used to make judgments in these experiments was detectable in two-view motion sequences and therefore was not euclidean in nature. On the basis of these results, Todd and Bressan (1990) and Todd and Norman (1991) claimed that two-view relations contained all of the geometrical information actually perceived by human observers and that optical variables defined over three or more views (such as acceleration, etc.) were not important for the perception of shape.

The results of Experiment 1, where the figures were sinusoidally stretched along the line of sight, would appear to constitute strong evidence confirming the earlier results, suggesting that the perception of 3-D shape was based on affine properties. Second-order temporal information that could have informed observers that these displays were nonrigid deforming 3-D structures was available. Indeed, when we input the motions from these optical patterns to a typical computational model, it had no problem detecting the nonrigid 3-D deformation that appeared perfectly rigid to actual human observers.

However, the results of the stretching transformations perpendicular to the line of sight in the image plane, as well as the results of Experiment 2, show that some optical variables over many views *can* be detected, and that they can influence the perception of rigid motion. In particular, the phase relationships between the velocity changes of a figure's constituent points seem crucial for the perception of rigid motion. The perceived rigidity of the five qualitatively different optical patterns fell into two groups. The two conditions that preserved phase variations were perceived as rigid, whereas the three conditions that disrupted phase were perceived as nonrigid.

One possibility that would explain both the failure of observers to detect the nonrigidity of the stretch-in- z displays used in Experiment 1 and the ability to detect dif-

ferences in the relative phases of the high-frequency variations between different moving points in Experiment 2 is that the human visual system may be able to extract some limited type of second-order temporal information. The simple detection of sign of acceleration would explain the partial sensitivity to second-order relations found in Experiment 2 and would be consistent with the failure to use the magnitude of the accelerations to recover euclidean 3-D relations found in Experiment 1.

Other evidence besides the current results is consistent with the finding that some limited information across many views is recovered. In particular, Todd (1982) showed that some variables, such as frequency, are important for perceived rigidity. Detecting that some points are rotating at a different frequency than others would require more than two views. However, despite the finding that some variables that are defined only over many views influence perceived rigidity of an object's motion in 3-space, there is no evidence to suggest that humans can accurately detect the precise magnitude of second-order temporal relations that is required to calculate euclidean properties related to shape. The magnitudes of the individual motions in the random amplitude condition in Experiment 2 were not in agreement (over three views or more) with the projections of any rotating rigid 3-D object. Nevertheless, these displays were perceived as rigidly rotating 3-D shapes as long as the high-frequency uncorrelated variations in velocity were in phase.

Taken together, all of this evidence (Todd & Bressan, 1990; Todd and Norman, 1991, as well as the current results) indicates that the perceptual knowledge of 3-D structure is probably based on more abstract relationships than those allowed under euclidean geometry—that is, on affine and ordinal relationships in 3-space. When considering stereopsis (also a situation involving two views, simultaneous rather than successive), Julesz (1971) noted that “stereopsis itself gives us the experience of relative depth only. It enables us to rank order the nearness-farness of objects within a region of space around a fixation point” (p. 144). He concluded that “for stereopsis, one must generalize the metric of space from a rigid Euclidean one to a less rigid affine or topological space” (p. 290). We now find that the recovery of 3-D shape from motion may be based on principles similar to those used to recover shape from stereoscopic vision.

REFERENCES

- BENNETT, B. M., & HOFFMAN, D. D. (1985). The computation of structure from fixed-axis motion: Nonrigid structures. *Biological Cybernetics*, *51*, 293-300.
- BENNETT, B. M., HOFFMAN, D. D., NICOLA, J. E., & PRAKASH, C. (1989). Structure from two orthographic views of rigid motion. *Journal of the Optical Society of America A*, *6*, 1052-1069.
- 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., HOFFMAN, D. D., & POLLICK, F. E. (1990). Discriminating rigid from nonrigid motion: Minimum points and views. *Perception & Psychophysics*, *47*, 205-214.
- BRAUNSTEIN, M. L., HOFFMAN, D. D., SHAPIRO, L. R., ANDERSEN, G. J., & BENNETT, B. M. (1987). Minimum points and views for the recovery of three-dimensional structure. *Journal of Experimental Psychology: Human Perception & Performance*, *13*, 335-343.
- DOSHER, B. A., LANDY, M. S., & SPERLING, G. (1989). Kinetic depth effect and optic flow: I. 3D shape from fourier motion. *Vision Research*, *29*, 1789-1813.
- GIBSON, J. J., & GIBSON, E. J. (1957). Continuous perspective transformations and the perception of rigid motion. *Journal of Experimental Psychology*, *54*, 129-138.
- GIBSON, E. J., GIBSON, J. J., SMITH, O. W., & FLOCK, H. (1959). Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology*, *58*, 40-51.
- GREEN, B. F. (1961). Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, *62*, 272-282.
- HILDRETH, E. C., GRZYWACZ, N. M., ADELSON, E. H., & INADA, V. K. (1990). The perceptual buildup of three-dimensional structure from motion. *Perception & Psychophysics*, *48*, 19-36.
- HOFFMAN, D. D., & BENNETT, B. M. (1986). The computation of structure from fixed-axis motion: Rigid structures. *Biological Cybernetics*, *54*, 71-83.
- HOFFMAN, D. D., & FLINCHBAUGH, B. E. (1982). The interpretation of biological motion. *Biological Cybernetics*, *42*, 195-204.
- HUANG, T. S., & LEE, C. H. (1989). Motion and structure from orthographic projections. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *11*, 536-540.
- JOHANSSON, G. (1964). Perception of motion and changing form: A study of visual perception from continuous transformations of a solid angle of light at the eye. *Scandinavian Journal of Psychology*, *5*, 181-208.
- JOHANSSON, G., & JANSSON, G. (1968). Perceived rotary motion from changes in a straight line. *Perception & Psychophysics*, *4*, 165-170.
- JULESZ, B. (1971). *Foundations of cyclopean perception*. Chicago, IL: University of Chicago Press.
- KOENDERINK, J. J., & VAN DOORN, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A*, *8*, 377-385.
- LAPPIN, J. S., & FUQUA, M. A. (1983). Accurate visual measurement of three-dimensional moving patterns. *Science*, *221*, 480-482.
- LAPPIN, J. S., & LOVE, S. R. (1992). Planar motion permits perception of metric structure in stereopsis. *Perception & Psychophysics*, *51*, 86-102.
- METZGER, W. (1934). Beobachtungen über phänomenale Identität. *Psychologische Forschung*, *19*, 1-60.
- MILES, W. R. (1931). Movement interpretations of the silhouette of a revolving fan. *American Journal of Psychology*, *43*, 392-405.
- NORMAN, J. F., & TODD, J. T. (1991). The perception of rigid motion for affine distortions in depth. *Investigative Ophthalmology & Visual Science (Suppl.)*, *32*, 958.
- NORMAN, J. F., & TODD, J. T. (1992). The visual perception of 3-dimensional form. In G. A. Carpenter & S. Grossberg (Eds.), *Neural networks for vision and image processing* (pp. 92-110). Cambridge, MA: MIT Press.
- TODD, J. T. (1982). Visual information about rigid and nonrigid motion: A geometric analysis. *Journal of Experimental Psychology: Human Perception & Performance*, *8*, 238-252.
- 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.
- TODD, J. T., AKERSTROM, R. A., REICHEL, F. D., & HAYES, W. (1988). Apparent rotation in three-dimensional space: Effects of temporal, spatial, and structural factors. *Perception & Psychophysics*, *43*, 179-188.
- TODD, J. T., & BRESSAN, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics*, *48*, 419-430.

- TODD, J. T., & NORMAN, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics*, **50**, 509-523.
- ULLMAN, S. (1977). *The interpretation of visual motion*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.
- ULLMAN, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.
- WALLACH, H., & O'CONNELL, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, **45**, 205-217.

(Manuscript received May 26, 1992;
revision accepted for publication August 18, 1992.)