

The visual perception of rigid motion from constant flow fields

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

Four experiments investigated observers' judgments of rigidity for different types of optical motion. The depicted structural deformations were of two types: (1) those with nonparallel image trajectories that are detectable from the first-order spatiotemporal relations between pairs of views; and (2) those with parallel image trajectories that can only be detected from higher order relations among three or more views. Patterns were composed of smooth flow fields in Experiments 1 and 3, and of wire frame figures in Experiments 2 and 4. For both types of display, the nonrigidity detectable from the first-order spatiotemporal structure of the motion sequence was much more salient than the deformation detectable only from the higher order spatiotemporal structure. These results indicate that observers' judgments of rigidity are based primarily on a two-view analysis, but that some useful information can be obtained under appropriate circumstances from higher order spatiotemporal relations among three or more views.

When a visible object rotates in depth, its projected pattern of optical flow on the retina provides perceptually compelling information about its three-dimensional (3-D) structure. During the past decade, there has been a rapid growth in our theoretical understanding of how optical motion could potentially be analyzed. Much of this research was initiated by the early work of Ullman (1977, 1979), who provided the first computational analysis about the minimum amounts of motion information needed to obtain a unique interpretation of euclidean metric structure. In general, this analysis requires a minimum of three views of four noncoplanar points, though there are a few special cases in which these limits can be reduced (see, e.g., Bennett & Hoffman, 1985; Hoffman & Bennett, 1985, 1986; Hoffman & Flinchbaugh, 1982).

There has also been a considerable amount of psychophysical research during this period to determine how closely the performance of these models compares with that of actual human observers, but these empirical investigations have produced a rather surprising pattern of results. Whereas the computational analysis of euclidean metric structure in most instances requires a minimum of three distinct orthographically projected views, there is a growing amount of evidence to suggest that human observers can obtain compelling kinetic depth effects from two-view apparent motion sequences, and that there is only minimal improvement in performance on most objective response tasks as additional views are added (see, e.g., Braunstein, Hoffman, & Pollick, 1990; Braunstein, Hoffman, Shapiro, Andersen, & Bennett, 1987; Liter, Braunstein, & Hoffman, 1994; Norman & Todd, 1993; Todd,

Akerstrom, Reichel, & Hayes, 1988; Todd & Bressan, 1990; Todd & Norman, 1991).

In light of these psychophysical results, many researchers have looked for new theoretical explanations in order to investigate the potential information that is available from only two frames of an apparent motion sequence (see Bennett, Hoffman, Nicola, & Prakash, 1989; Huang & Lee, 1989; Koenderink & van Doorn, 1991; Todd & Bressan, 1990). One of the earliest and most intuitive of these analyses was developed by Ullman (1977). In his *two-frame* analysis (as opposed to the more well-known "structure from motion theorem"), Ullman noted that every rigid object rotation can be decomposed into two simple components: a rotation about an axis in the picture plane, and a rotation about the line of sight. This second component will hereafter be called *image rotation*. An example of this type of decomposition is shown in Figure 1. Ullman then proved that for any pair of orthographic projections of a rigidly rotating object it is possible to remove the image rotation component to produce a pattern of moving elements with parallel image trajectories (see Figure 1). Although it is mathematically possible for physically nonrigid motions to produce parallel image trajectories as well, the probability of encountering such a motion in natural vision is vanishingly small.

It is important to keep in mind, however, that Ullman's two-view analysis will falsely identify degenerate nonrigid motion with parallel image trajectories as rigid. In order to identify these motions as nonrigid, it would be necessary to employ a multiple frame analysis, like the one proposed by Ullman (1979).

Our goal for the present research was to determine the extent to which *two-frame* and *multiframe* analyses are used by the visual system in order to detect object rigidity. That is, for determining rigidity, does the visual system exploit the information available across three or more views

This research was supported in part by AFOSR Grant F49620-93-1-0116. Correspondence should be addressed to V. J. Perotti, Department of Psychology, Ohio State University, 201 Lazenby Hall, Columbus, OH 43210 (e-mail: perotti.2@osu.edu).

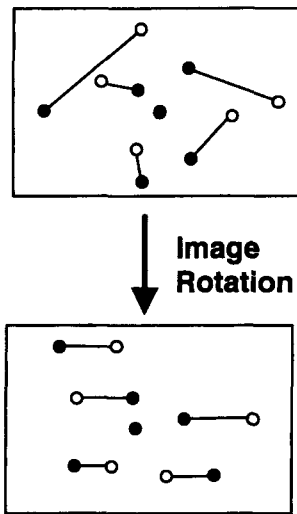


Figure 1. For any pair of images of a rigid object rotating in depth under orthographic projection, it is possible to produce a pattern of parallel trajectories by rotating one image with respect to the other. The connected pairs of open and filled circles represent corresponding points in different views of an apparent motion sequence.

(which is the minimal quantity needed to compute the complete euclidean structure)? In order to investigate this question, it was useful to compare observers' rigidity judgments for three different types of motion: (1) *rigid rotations*, which could be correctly identified as rigid by both two-frame and multiframe analyses; (2) *first-order structural deformations*, which could be correctly identified as nonrigid by both two-frame and multiframe analyses; and (3) *higher order structural deformations*, which could only be identified as nonrigid by a multiframe analysis.

The question of whether two-frame models can account for observers' rigidity judgments as well as multiframe models has previously been addressed by Norman & Todd (1993). They used displays composed of randomly connected line segments, which appeared much like a bent wire hanger, and they analyzed human observers' sensitivity to two distinct types of stretching deformations: those along the line of sight orthogonal to the projection plane, which were perceived as rigid, and those along a horizontal axis parallel to the projection plane, which were perceived as nonrigid. Because neither type of deformation is detectable with the use of a two-frame model, they both fall into the category of higher order structural deformations. Thus, observers' ability to detect the nonrigidity of stretching transformations along a horizontal axis indicates that they must be sensitive to some degree of multiframe temporal information in moving displays.

We designed the research described in the present article in order to reinvestigate the Norman and Todd (1993) result with a different paradigm and different stimuli. In addition to the higher order structural deformations examined by Norman and Todd, displays were included that should have been identifiable as rigid using either a two-

frame or a multiframe analysis, and also displays that could be identified as nonrigid by either analysis. The first type of motion can be generated simply by simulating a rigid object motion, but the different categories of structural deformation require a more careful design.

The projected motion of any smooth surface can be expressed as a flow field, where the signed magnitude of velocity V at each image location is defined as a function of both its 2-D location (x,y) and time t :

$$V = f(x,y,t). \quad (1)$$

There exist certain situations in nature for which the flow field remains constant over time, but these situations are very rare. They occur only when a surface of revolution rotates about its axis of symmetry. In all other cases, if a flow field remains constant over time, it will have no possible rigid interpretations.

A constant flow field can be generated as a set of randomly placed image elements, whose motions are determined by their current 2-D location (x,y) :

$$V = f(x,y), \quad (2)$$

such that the velocity associated with each location remains constant throughout the entire display period. Thus, any dot which after some number of previous movements arrives at position (x,y) will have displacement $f(x,y)$ for the subsequent frame transition. In general, a constant flow field will fall into the category of a higher order structural deformation, provided that one other constraint is satisfied. A two-view motion sequence under orthographic projection will only allow a rigid interpretation if there is a pattern of parallel trajectories following the removal of image rotation. Thus, if observers are insensitive to higher order spatiotemporal relations among three or more views, as has been suggested by Todd and Bressan (1990) and Todd and Norman (1991), then a constant flow field with parallel trajectories ought to be perceived as rigid since the first-order temporal properties (instantaneous velocities) indicate rigid motion. If, on the other hand, observers are sensitive to these higher order spatiotemporal relations, then they should be able to correctly identify the motion as nonrigid. Constant flow fields can also be utilized to depict first order structural deformations. Since both two-frame and multiframe models are sensitive to the first-order temporal relations, a display whose instantaneous velocities have no possible rigid interpretation should be perceived as nonrigid using either type of analysis. An example of such a display is the constant flow field with nonparallel trajectories. Using a two-frame analysis, a constant flow field with nonparallel trajectories should be judged nonrigid for the very reason that the trajectories are not parallel. One final prediction with respect to Ullman's two-frame model is that the presence of image rotation should have no impact on perceived rigidity in any of these cases, since this rotation is removed prior to testing for the occurrence of parallel trajectories.

In Experiment 1, each of the three motion types described above was presented (both with and without image rota-

tion) in an effort to test the predictions of both two-frame and multiframe analyses of structure from motion. A summary of the various display conditions along with the predictions of each model is shown in Table 1.

EXPERIMENT 1

Method

Apparatus. Apparent motion sequences were displayed on a Silicon Graphics 4D/310 VGX workstation. The displays were viewed monocularly at a distance of 74 cm, and head movements were restricted by a chinrest. The display screen subtended a visual angle of $25.9^\circ \times 20.8^\circ$. It had a spatial resolution of $1,280 \times 1,024$ pixels, though the effective resolution was increased approximately tenfold by hardware antialiasing. The frame rate was 30 frames per second. Each stimulus was composed of a specific number of distinct frames that were displayed forward and backward in oscillation for as long as was necessary for observers to rate its apparent rigidity.

Stimulus displays. Every display contained 5,000 points, each of which was assigned an initial 2-D location (x, y) at random in the display area. The coordinate system of the points was assigned so that the center of the screen was the origin $(0, 0)$. The x - and y -coordinates were chosen from the range $(-600.0, 600.0)$ and $(-600.0, 600.0)$ respectively. However, only the pixels within the range $(-300.0, 300.0)$ were displayed at any given moment. When points left this smaller display area because of a displacement, they were moved to the opposite side of the display. The patterns of motion were defined mathematically based on an equation of the form:

$$V = 2 \left[\sin \left(\frac{\left(\frac{ax^2 + by^2 + cxy}{100} + dx + ey \right)}{100} \right) + 2 \right], \quad (3)$$

where V is the horizontal image displacement of a point (in pixels) at position (x, y) , and a, b, c, d , and e are constants chosen at random from the range $[0, 1]$. A stereogram representing a typical 3-D surface of this form is shown in Figure 2. In the experiment, three different types of display were presented: *rigid rotations*, *parallel constant flow fields*, and *nonparallel constant flow fields*.

For the simulation of rigid rotation, the values given from Equation 3 were multiplied by 50 to obtain a z (depth) coordinate for each object point in pixels relative to the axis of rotation. Such an object, which has the form $z = f(x, y)$, is called a Monge surface. At each frame transition, the surface was rotated about a vertical axis and its points were displayed under orthographic projection so that their motions were limited to the horizontal direction.

To create parallel constant flow field displays, Equation 3 defined the displacement for each point at a given (x, y) screen location. For example, assume that on frame 1, point A at location (x, y) has a displacement of 2 pixels in the x direction and a displacement of 0 in the y direction [i.e., $f(x, y) = 2$]. At the next frame transition, the displacement at location (x, y) would remain the same [$f(x, y) = 2$]. However, point A's new location would be $(x+2, y)$ and its new displacement would be given by the function $f(x+2, y)$. The move-

ments of all points in these displays were constrained to the horizontal direction, as was the case for the rigid rotations.

The third type of display (nonparallel constant flow field) was generated in a manner identical to that for the parallel constant flow field case, except that the set of points had independent displacement functions: $dx = f(x, y)$ and $dy = g(x, y)$, where dx and dy are the displacement values (in pixels) in the horizontal and vertical directions, respectively. Each function (f and g) was of the form of Equation 3. For this set of displays, the position of a dot could vary over time both in the horizontal (x) and vertical (y) directions, thus allowing the dots to have nonparallel trajectories. For each display type (rigid rotation, parallel constant flow fields, and non-parallel constant flow fields), an additional condition was included which had an added rotation about the line of sight (image rotation).

The extents of oscillation for rigid rotations were 20° , 40° , or 60° , centered on the frontoparallel view of the depicted Monge surface. For displays with image rotation, the angular velocity about the line of sight was 0.5° per frame transition. The angular velocity about the vertical axis was 1° per frame for the 40° and 60° angular extents, and 0.75° per frame for the 20° rotations. Given the total extent of rotation and the angular velocity, the apparent motion sequences contained 26, 40, or 60 frames in increasing order of sequence length. The same sequence lengths were also used in the constant flow field conditions.

Procedure. Observers rated the perceived rigidity of moving dot patterns by manipulating a handheld mouse that controlled a pointer on a visible scale presented just below the moving display. The motion sequences oscillated continuously until the pointer was set at the desired location and the observer pressed a button on the mouse to initiate a new trial. The rating scale varied from 0 to 1, where 0 designated nonrigid deformation and 1 designated a rigid rotation. The instructions for rating each display were as follows:

If the object appears to be rotating rigidly with no detectable deformation, then respond by adjusting to complete rigid rotation. If the object appears to be deforming with no detectable rigid rotation, then respond by adjusting to the nonrigid end of the scale. If the object appears to have components of both rigid rotation and nonrigid deformation, then respond at a point along the scale which represents the perceived relative weighting of the two components.

The experiment had a total of 18 conditions (3 motion types \times 2 levels of image rotation \times 3 sequence lengths). For each condition, observers made a total of 10 responses.

Observers. Six observers participated in the experiment: 3 were the authors, and 3 others were naive about how the displays were generated.

Results

The combined results for the 6 observers are depicted in Figure 3. An analysis of variance with two orthogonal planned comparisons was used to estimate the effects of the experimental manipulations. Virtually all of the variance in the observers' responses was determined by the comparison of those displays with parallel trajectories (rigid rotation, and parallel constant flow field) versus those displays with nonparallel trajectories (nonparallel constant flow field). The sum of squares for this comparison accounted for 99.96% of the total sum of squares in the experiment. This strongly significant result [$F(1, 5) = 2754.837, p < .01$] indicates that the parallelism of image trajectories is the most important criterion for determining perceived rigidity in these displays.

The main effect of image rotation was insignificant, and so was a comparison of the two cases which had parallel

Table 1
Conditions of Experiment 1

Condition	Prediction	
	Two-Frame	Multiframe
Rigid rotation	rigid	rigid
Parallel constant flow field	rigid	nonrigid
Nonparallel constant flow field	nonrigid	nonrigid

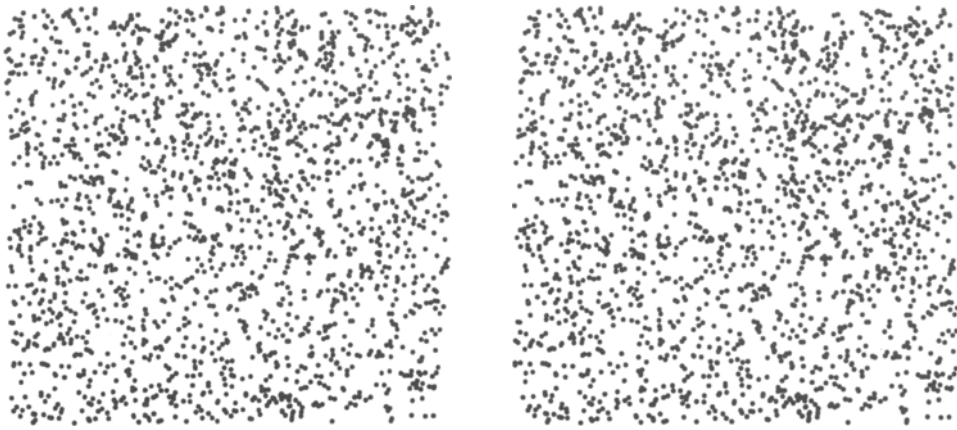


Figure 2. A random dot stereogram depicting an example surface used in Experiment 1.

trajectories before the addition of image rotation (the rigid rotation conditions versus the parallel trajectory constant flow field conditions). The latter result provides strong evidence that human observers preferentially utilize the two-frame velocity information for their rigidity judgments. If the observers in this experiment had been sensitive to multi-frame motion information, they should have been able to detect the structural deformations which were present in all of the constant flow field displays. There are two reasons to conclude that observers in Experiment 1 were unable to detect the nonrigidity of motion in the parallel constant flow field condition. First, it is evident from the second planned comparison that observers did not rate the rigid rotation condition as any more rigid than those displays which had consistent rigidity only in the two-frame motion information present. Second, the mean of the rigidity ratings in the parallel constant flow field conditions was quite high (i.e., .982 on a scale of 0.0 to 1.0). In sum, the displays with paral-

lel trajectories appeared as a group to be almost perfectly rigid.

In order to assess the actual extent of 3-D structural deformation depicted in these constant flow field displays, it is useful to employ a tolerance analysis that has recently been developed by Hogervorst, Kappers, and Koenderink (1996) to test the rigidity of an object's projected motion after it has been transformed to a pattern of parallel trajectories by the removal of image rotation. The basic idea is quite simple and powerful. The horizontal positions of two points are measured over a sequence of successive frames relative to a third point. These are then plotted in phase space, in such a way that their relative positions are represented along orthogonal axes. If the trajectory in this space is anything but an ellipse (or a line) centered on the origin, then the object has no possible rigid interpretation.

Hogervorst et al. (1996) have employed this analysis to account for the findings of Norman and Todd (1993) that rotating objects stretched sinusoidally along the line of sight are perceived as rigid, whereas those stretched perpendicular to the line of sight are perceived as nonrigid. The phase space trajectories for the first type of displays are very close to elliptical, whereas those for the latter are not. If the structure of phase space trajectories provides a perceptually valid metric for scaling the magnitude of nonrigid deformations, as suggested by Hogervorst et al., then that would explain why the displays used by Norman and Todd (1993) appeared so different from one another.

Could a similar explanation be used to account for the results of the present experiment? In applying this analysis to constant flow field displays, it is convenient to choose three points that are vertically colinear in the initial frame of the motion sequence. If the depicted object motion is rigid, the resulting phase space trajectory will form a straight line that passes through the origin. Figure 4 shows two examples of phase space trajectories that were computed from the rigid motion and parallel constant flow field conditions of the present experiment. The left panel of this figure shows a linear trajectory passing through the origin, which is characteristic of rigid rotation in depth.

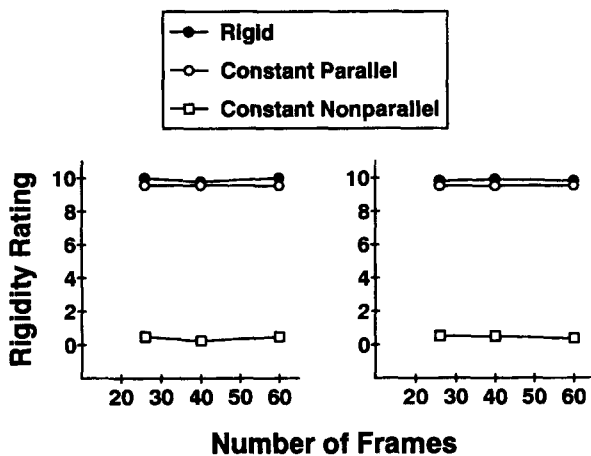


Figure 3. The average rigidity ratings combined over 6 observers for the different conditions of Experiment 1. Ratings for displays with and without image rotation are summarized in the right and left panels, respectively.

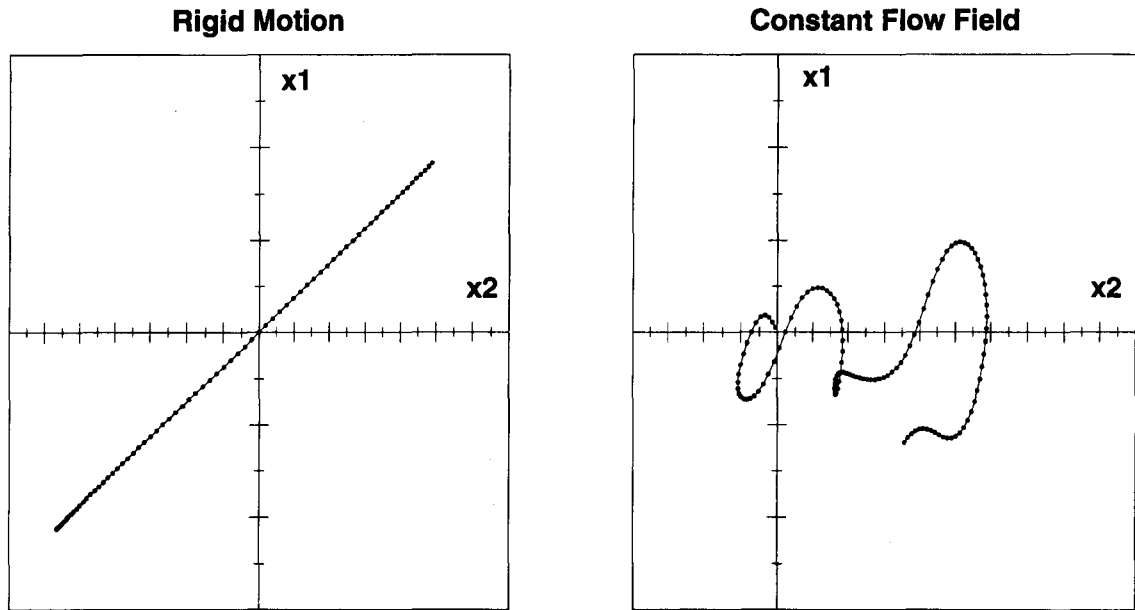


Figure 4. Example phase space trajectories for the rigid motion and parallel constant flow field displays of Experiment 1. Each plot represents the horizontal positions of two points x_1 and x_2 relative to a third point, which are all vertically colinear in the initial frame of the motion sequence. Each major tick along the x_1 and x_2 axes represents 0.5° of visual angle. The pattern of projected motion will only have a mathematically possible rigid interpretation if its phase space trajectories are elliptical (or linear) and centered on the origin.

The right panel, in contrast, shows a typical trajectory from the parallel constant flow field condition. Note in this case that the trajectory is clearly nonlinear and should therefore have been identified as nonrigid if observers had based their ratings on a process similar to the one described by Hogervorst et al.

EXPERIMENT 2

Although the results of Experiment 1 would appear to indicate that observers are unable to exploit higher order spatiotemporal relations among three or more views for analyzing the rigidity of an object's motion, there is other evidence from Norman and Todd (1993) to suggest that these higher order relations can influence observers' perceptions of rigidity in other circumstances. Experiment 2 was designed therefore in an effort to provide some converging operations to address this issue. The stimulus displays employed in this experiment were based on Todd's (1982) analysis of the geometric relations among projected element trajectories in image space—which should not be confused with trajectories in phase space as depicted in Figure 4.

Todd (1982) identified three constraints for a rigid object rotating about a fixed axis: (1) The trajectory of all points on the object must be circular as it moves in depth; (2) the centers of all trajectories must lie on a straight line, which is perpendicular to each plane of rotation; and (3) each point must traverse its trajectory with an identical frequency. There are three corresponding constraints

that are applicable to the optical projections of rigid motions: (1) the projected images of all points will travel in elliptical trajectories; (2) the minor axis of these projected trajectories will lie on a single straight line; and (3) all points must traverse their projected trajectories at a common frequency. By systematically violating these various constraints, it is possible to produce several different classes of nonrigid motion that are represented schematically in Figure 5.

Experiment 2 was a simplification and extension of the Todd (1982) paradigm, designed to isolate the relative importance of various stimulus factors that can influence two-frame and multiframe analyses of structure from motion. All of the displays were generated with a single simulated axis of rotation, perpendicular to the line of sight and in the vertical direction. This simplification forced all projected trajectories to be linear. The goal for this experiment was to evaluate the viability of both two-frame and multiframe models as explanations for human rigidity judgments. As in Experiment 1, observers were presented with three different classes of stimuli: rigid rotations in depth, first-order structural deformations that were detectable with a two-frame analysis, and higher order structural deformations that could only be detected with a multiframe analysis.

Method

Apparatus. The apparatus was identical to that used in Experiment 1.

Stimulus displays. In this experiment the displays were composed of seven vertically separated points, connected by solid lines. The 2-D trajectories of these points were equivalent to the projec-

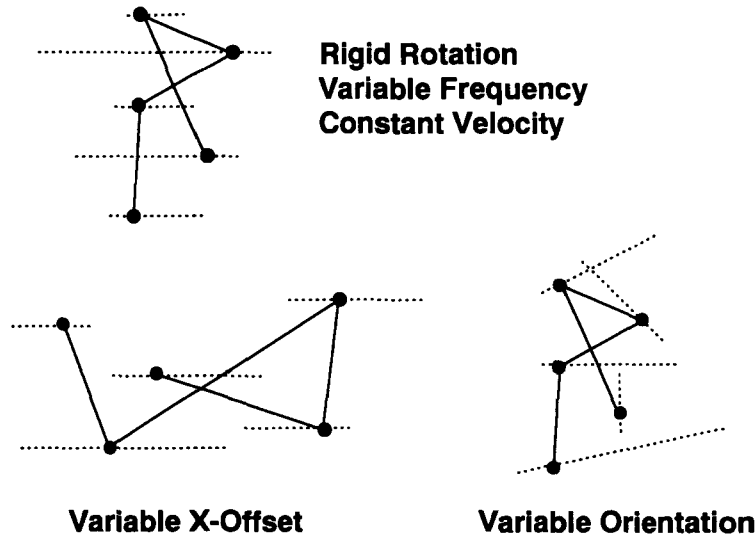


Figure 5. Three patterns of projected element trajectories similar to those used in the various conditions of Experiment 2. The solid lines represent the instantaneous configuration of a moving wire frame figure, and the dotted lines show the projected trajectory of each vertex. The pattern in the upper left is a typical configuration that could occur in the rigid, variable frequency, or constant velocity conditions. The one in the lower left is a typical example from the variable x-offset condition, and the one in the lower right is a typical example from the variable orientation condition. These variable orientation displays were the only ones that could be identified as nonrigid on the basis of a two-frame analysis of structure from motion.

tions of seven distinct points on a 3-D object moving in depth. The equations defining these patterns of motion were simplified versions of those given in Todd (1982):

$$x = A \cos(\omega t + \alpha) \cos(\theta) + B, \tag{4}$$

$$y = -A \cos(\omega t + \alpha) \sin(\theta) + C, \tag{5}$$

where x and y are the pixel locations of a given element at a given moment in time; A is the amplitude of the trajectory in pixels; ω is its frequency; α is the initial phase in the first frame of the motion sequence; θ is the angle of the trajectory relative to the horizontal in radians; B is the horizontal offset of the trajectory from the origin in pixels; and C is the vertical offset in pixels. For all displays, the values of α , A , and C were selected at random on each trial; α ranged from 0 to 2π , A ranged from 100 to 350 pixels, and C ranged from -300 to 700 pixels. For rigid rotations, all trajectories had the same frequency ($\omega = \pi/60$), the same orientation ($\theta = 0$ radians), and the same horizontal offset ($B = 0$).

Three types of nonrigid motion were generated, each of which allowed one of the above factors to vary across the seven points, while the others remained constant. These conditions will be referred to as *varying orientation*, *varying frequency*, and *varying x-intercept*. For varying-orientation displays, θ was randomly assigned one of seven equally spaced values in the range 0 to π . For varying-frequency displays, ω was randomly assigned one of seven equally spaced values in the range $\pi/150$ to $\pi/30$ radians. For varying x-intercept displays, the value of B was randomly assigned one of seven equally spaced values in the range from -400 to 400 pixels. A final nonrigid case was *constant velocity*. This condition was similar to the rigid displays, except that each point moved back and forth along its trajectory at a constant 2-D speed (see Braunstein & Andersen, 1984). The predicted appearances of these different conditions for two-frame and multiframe models are given in Table 2.

The points were connected in a random order by six lines, thus giving the impression of a twisted wire hanger moving in depth. Each

display type was presented with four different sequence lengths composed of 10, 20, 30, or 120 distinct frames. For the rigid rotations, this produced total angular displacements of 30°, 60°, 90°, or 360° over the entire sequence.

Procedure. As in Experiment 1, observers judged the perceived rigidity of moving displays by adjusting a pointer on a visible scale whose endpoints were labeled as *non-rigid deformation* and *rigid rotation*, respectively. The experiment had 20 display conditions: 5 display types (rigid, varying orientation, varying x-intercept, varying frequency, and constant velocity) \times 4 sequence lengths.

Observers. The displays were judged by the same 6 observers who participated in Experiment 1.

Results

Figure 6 presents a summary of the data for all 6 observers. An analysis of variance with four orthogonal planned comparisons was used to estimate the relative significance of each of the experimental factors. A comparison of the varying-orientation condition versus all other conditions was significant [$F(1,5) = 227.75, p < .01$] and accounted for 53% of the total sum of squares in the experiment. This

Table 2
Conditions of Experiment 2

Condition	Prediction	
	Two-Frame	Multiframe
Rigid	rigid	rigid
Varying frequency	rigid	nonrigid
Varying x-intercept	rigid	nonrigid
Constant velocity	rigid	nonrigid
Varying orientation	nonrigid	nonrigid

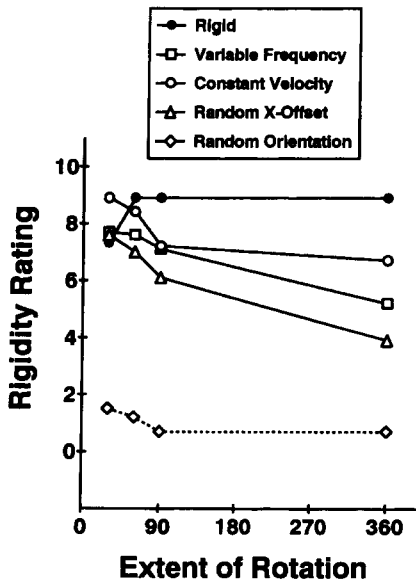


Figure 6. The average rigidity ratings combined over six observers for the different conditions of Experiment 2.

again indicates that the parallelism of image trajectories is a crucial factor for determining perceived rigidity.

A second planned comparison of the *rigid* condition versus the three nonrigid conditions with parallel image trajectories (varying frequency, varying x -intercept, and constant velocity) was also statistically significant [$F(1,5) = 14.215, p < .05$], and accounted for 3% of the total sum of squares. The remaining two comparisons among the different nonrigid conditions were statistically insignificant. The fact that the rigid rotations produced marginally higher ratings than did the nonrigid motions with parallel image trajectories indicates that observers' judgments must have been influenced to some extent by higher order spatiotemporal relations among three or more views in the apparent motion sequences. If their performance had been based solely on first-order relations between pairs of views, then all of the displays with parallel trajectories should have been perceptually equivalent. Since this was not the case, we must conclude that multi-frame temporal characteristics of object motion provide some degree of information for perceived rigidity. However, it is important to keep in mind that the multiframe components of nonrigidity accounted for about an order of magnitude less sum of squares than the two-frame components due to nonparallel trajectories.

Although the multiframe information accounts for much less of the variance in observers' judgments, its presence provides definitive evidence that two-frame analyses are incomplete as explanations for human motion perception. Considered together, these results provide strong confirmation of the findings of Norman and Todd (1993) that observers' rigidity judgments are based primarily on two-frame motion information, but that higher order spatiotemporal relations can have a modest effect under certain conditions.

To assess the extent of physical nonrigidity in these displays we again employed the tolerance analysis of Hoger-vorst et al. (1996). The horizontal positions of two points were measured over a sequence of successive frames relative to a third point, all of which were selected at random from the possible display parameters in the different experimental conditions. The resulting phase space trajectories are shown in Figure 7. Note that for rigid rotation the phase space trajectory is an ellipse centered on the origin. For the varying x -intercept condition, it is also an ellipse though displaced from the origin. For the constant velocity condition, the trajectory forms a six-sided polygon, whose vertices correspond to the moments in time where one of the projected images of a moving element abruptly reverses direction. The most radical deviation from an elliptical phase space trajectory occurs in the varying frequency condition, yet from the results shown in Figure 6, that particular form of nonrigidity was the most difficult for the observers to detect.

EXPERIMENT 3

The previous two experiments showed that displays with parallel trajectories appeared to the observers as compelling 3-D percepts, despite the fact that those displays were not consistent with the extended rotation of any real 3-D object. Experiment 3 was designed to further explore the perceptual similarity between constant flow fields and projected 3-D rotation. In particular, a discrimination paradigm was employed to determine whether the observers could find *any* salient differences between the two qualitatively different types of displays or whether they were perceptually indistinguishable.

Method

Apparatus. The displays were presented on a Silicon Graphics Crimson VGXT workstation. The display area and resolution were identical to those in Experiments 1 and 2. The refresh rate for the motion was 60 frames per second.

Stimulus displays. Experiment 3 required observers to discriminate the rigid and nonrigid motions of smoothly curved random dot surfaces. As in Experiment 1, the rigid displays were simulations of actual 3-D rotations, while the nonrigid displays were constant flow fields created using the same surface function. The rotation of 3-D objects naturally leads to situations in which certain parts of an object occlude other parts. Had we employed the same displays as in Experiment 1, the presence or absence of occlusion contours would have provided an obvious cue for discrimination if observers were provided with response feedback. Thus, in order to prevent this from occurring, the displays were created using a new function to define the patterns of motion. To better understand how the stimulus displays were constructed, it is useful to imagine a set of identical ellipses stacked on top of one another that are all centered on a vertical axis in the image plane. The surface is created by allowing each ellipse to have a different orientation relative to the x - and z -coordinate axes. The orientation (θ , in radians) of each ellipse is a function of its vertical location (y):

$$\theta = \frac{1}{2} \sin(fy^2 + gy + h) + \theta_0, \quad (6)$$

where f , g , h , and θ_0 are randomly selected constants for each unique surface. As long as the minor axis of each ellipse is longer than the

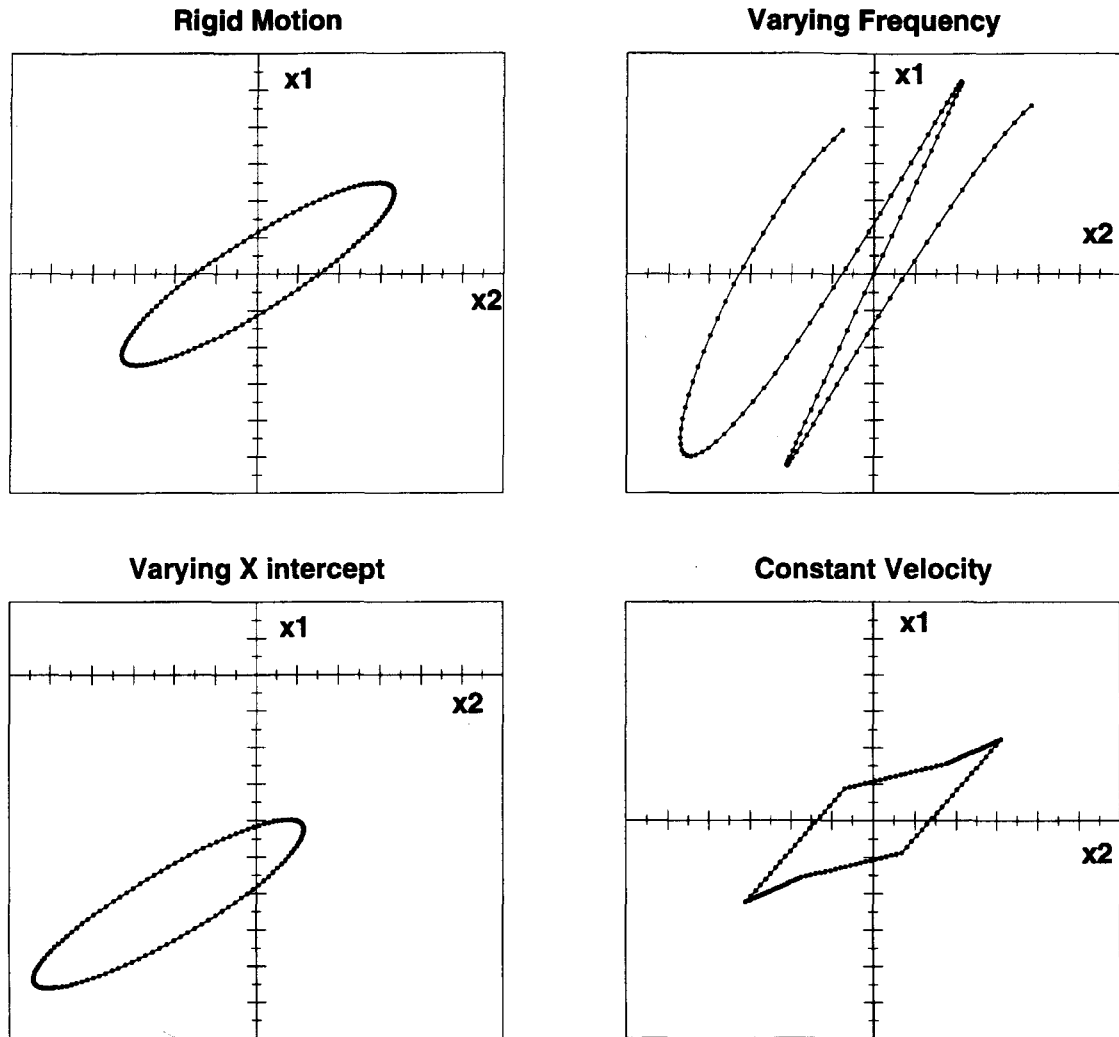


Figure 7. Example phase space trajectories for the different conditions of Experiment 2. Each plot represents the horizontal positions of two points x_1 and x_2 relative to a third point. Each major tick along the x_1 and x_2 axes represents 2.5° of visual angle. The pattern of projected motion will only have a mathematically possible rigid interpretation if its phase space trajectories are elliptical and centered on the origin.

visible display width, this ellipse surface will produce no visible self-occlusion as it rotates about a vertical axis. A stereogram depicting an example ellipse surface is shown in Figure 8.

All displays were based on the ellipse surface described in Equation 6, with the free parameters selected as follows. The origin of the coordinate axes was again at the center of the screen, but for this experiment y values ranged from $(-5.0, 5.0)$ centimeters. The minor and major axes of each ellipse were always set to 10.0 cm and 20.0 cm, respectively. The values for the f , g , and h parameters above had ranges (0 to 0.5), (2 to 6), and (0 to 5), respectively. The initial orientation of all ellipses, θ_0 , had a range of (0 to 2π) radians. For rigid rotations, the surfaces defined by this function were rotated in depth with 0.2° increments at each frame transition. The same function which was used to create the rigid 3-D objects was scaled and used as the velocity for the constant flow fields. There were three different sequence lengths composed of 80, 180, or 280 distinct frames, and the displays could be presented both with and without image rotation. The rate of the image rotation (when present) was 0.2° per frame transition.

The problem of self-occlusion in the rigid rotation condition was not the only one that had to be considered when creating the displays. An additional constraint was the requirement that the nonrigid conditions (constant flow fields) had to have quantifiable amounts of structural nonrigidity. To measure the nonrigidity of the displays, we performed a simulation based on the analysis of phase space trajectories developed by Hogervorst et al. (1996). If two points rotating in depth about a vertical axis are aligned vertically in the initial frame of a motion sequence, their phase space trajectories will trace a line centered at the origin. If the points move within a constant flow field, in contrast, their phase space trajectories will be quite different. A typical example is shown in Figure 9. Note that this trajectory is clearly nonlinear over its entire extent, but that it is approximately linear over limited segments.

In an effort to quantify the deviations from rigidity, we performed a regression analysis for 500 random sets of points for each of three different display lengths (80, 180, or 280 frames). On each trial, three points were selected for comparison and were started in phase. After moving the points for a given number of frames, the phase

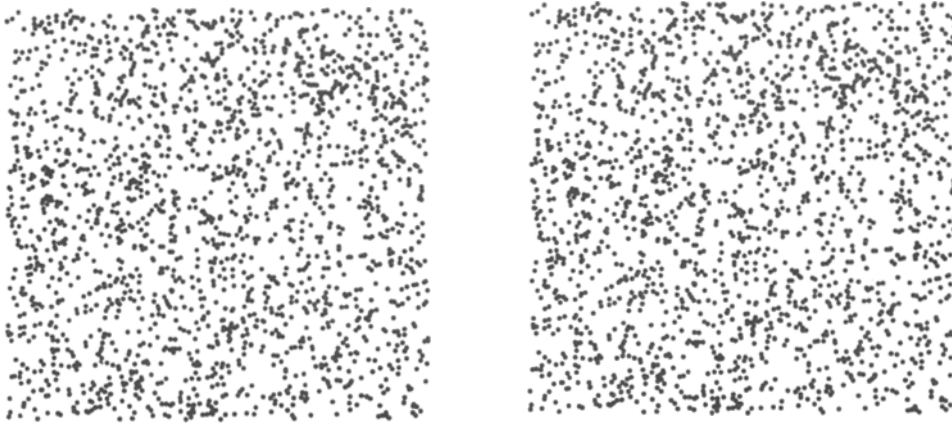


Figure 8. A random dot stereogram depicting a typical surface used in Experiment 3.

space trajectories were tested for linearity using a regression analysis. There are two ways in which any given display might fail to be rigid: one is that the shape of the phase space trajectories is not a line, and the other is that the line is not centered on the origin. Figure 10 depicts the distributions of r^2 values for each of the different sequence lengths used in the experiment. Predictably, the rigidity of the displays decreases as the number of frames increases. Note that for the largest sequence lengths (280 frames) about 45% of the simulated trials had r^2 values less than .8, thus indicating that the simulated displays were indeed nonrigid. It should also be noted that this is a conservative measure, since it does not take into account the distance of the regression line from the origin.

Procedure. On each trial, observers were required to discriminate whether a motion sequence was rigid or nonrigid by pressing an ap-

propriate button on a handheld mouse. Displays oscillated forward and backward through a specified number of frames until the total time displayed was 4.7 sec, after which the screen went blank until the observer made a response. Immediate feedback about the accuracy of the response was then provided in the form of an auditory beep. There were 12 basic conditions: 3 sequence lengths \times 2 levels of image rotation \times 2 motion types, each of which was replicated 200 times.

Observers. The displays were judged by 4 observers, 3 of whom were naive to the purpose of the experiment.

Results

Although the observers were instructed to base their responses on the rigidity of the depicted motion, there were other possible image cues that they could have exploited to improve their performance. It is important to keep in mind that the surfaces used for this experiment were created specifically so that there would be no smooth occlusion boundaries as are typically observed during real 3-D rotation. This was done to ensure that the discrimination was based on perceived 3-D structure and not upon simple 2-D image differences, such as whether self-occlusion regions were present or absent. However, there were other possible confounding factors that *could not* be eliminated.

For example, the global features in a constant flow field (e.g., peaks and valleys) do not change location over time. Suppose that observers had exploited this property in making their discriminations, so that displays with moving features would be judged as rigid, and those with stationary features would be judged as nonrigid. Although this strategy would produce above chance performance for the displays with parallel trajectories, it would not be effective for the displays with image rotation. Thus, we would expect that performance should break down in the latter condition if observers had resorted to a 2-D response strategy.

Figure 11 shows the percentage of rigid responses in each condition, and Figure 12 shows the percentage of correct responses collapsed over display type. For the displays without image rotation there were significant effects of display type and sequence length [$F(1,18) = 31.49, p < 0.01$; and $F(2,18) = 6.859, p < 0.01$, respectively], but neither

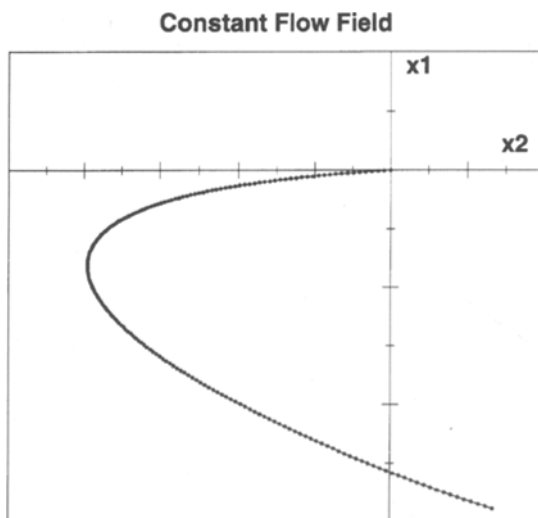


Figure 9. Example phase space trajectories for a typical constant flow field stimulus used in Experiment 3. Each plot represents the horizontal positions of two points x_1 and x_2 relative to a third point, which are all vertically colinear in the initial frame of the motion sequence. Each major tick along the x_1 and x_2 axes represents 0.5° of visual angle. The pattern of projected motion will only have a mathematically possible rigid interpretation if its phase space trajectories are elliptical (or linear) and centered on the origin.

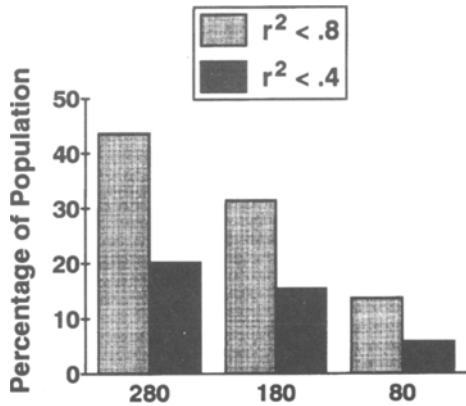


Figure 10. The results of a simulation to evaluate the nonrigidity of the ellipse surface displays of Experiment 3. Under the conditions of this simulation, the image motions of a set of points will only have a rigid interpretation if their trajectory in phase space is a straight line passing through the origin. For each sequence length, 500 phase space trajectories for randomly selected patterns were analyzed using linear regression. The light and dark bars show the percentage of these trajectories that had r^2 values less than .8 and .4, respectively, at each sequence length. Note that this does not take into account the additional requirement for rigid motion that the best-fitting regression line must pass through the origin.

of these factors had a significant effect on performance for the displays with image rotation. For the latter displays, the observers' judgments were no better than chance regardless of sequence length (see Figure 12). Thus, when the availability of secondary image cues was minimized, the nonrigid displays were indiscriminable from the rigid displays. It is also important to note, moreover, that even without image rotation, the observers' performance did not exceed 80% accuracy. Given that the displays included up to 280 distinct frames, and that the observers were provided with immediate feedback after every trial, it seems reasonable to conclude that their ability to make use of higher order spatiotemporal relations among three or more views was extremely limited.

EXPERIMENT 4

Experiment 4 assessed the ability of human observers to discriminate the rigid and nonrigid motions of rotating wire frame figures. The procedure was identical to that used in Experiment 3, though the stimulus displays included four of the different motion conditions from Experiment 2: rigid rotation, varying frequency, linear velocity, and varying x -intercept. The varying orientation condition from that experiment was not included, since the observers' ratings indicated that the displays with first-order structural deformations were clearly discriminable from the other categories. The ranges for all of the parameters in Equations 4 and 5 were equivalent to those used earlier. Three different frame sequence lengths were used (20, 80, or 140). Thus, there were 12 unique conditions, each of which was repeated 75 times for a total of 900 trials. The

4 observers were the same as those who had participated in Experiment 3.

Results

Figure 13 shows the percentage of rigid responses in each condition. Note in this case that there were clear differences among the different display types, such that the non-rigid displays were less likely to be identified as rigid than were the rigid rotations. This finding confirms again that observers were sensitive to at least some aspects of higher order spatiotemporal relations over three or more views, though at least some portion of this performance was prob-

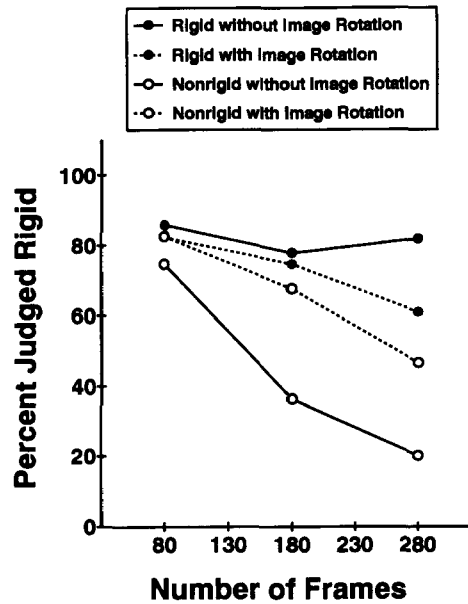


Figure 11. The percentage of trials in each condition that were identified as rigid by the 4 observers of Experiment 3.

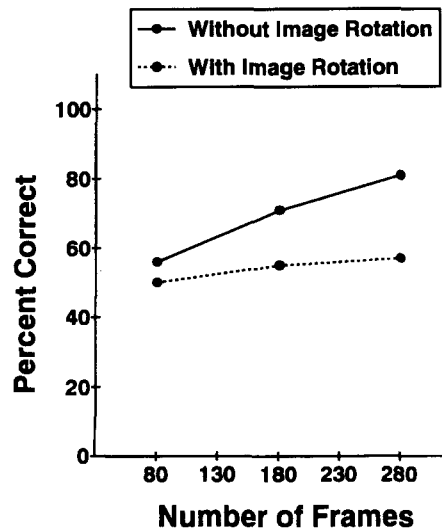


Figure 12. The percentage of correct responses for each sequence length in Experiment 3, collapsed over display type.

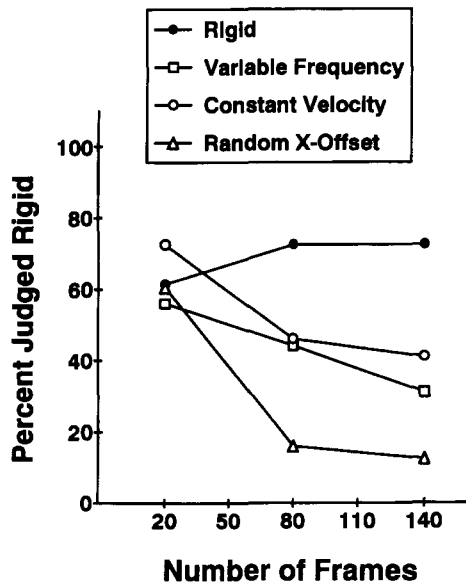


Figure 13. The percentage of trials in each condition that were identified as rigid by the 4 observers of Experiment 4.

ably due to artifactual cues. For example, the constant velocity displays could sometimes be identified by the abrupt transitions as each point changed its direction, and the variable frequency displays could sometimes be identified by the magnitude of their image velocities. In selecting the ranges of parameters for the displays, we made every effort to minimize the salience of these artifactual properties, but there were not enough degrees of freedom available to eliminate them entirely.

Figure 14 shows the percentage of correct responses for each sequence length collapsed over display type. The important thing to note in this figure is that despite the potential presence of artifactual cues and the use of immediate feedback after every trial, the overall level of accuracy did not rise above 75%, even with the longest sequence length of 280 frames. When combined with the results of Experiments 1, 2, and 3, and the earlier work of Todd and Bressan (1990), Todd and Norman (1991), and Norman and Todd (1993), these findings provide strong evidence that perceived rigidity is based primarily on an analysis of first-order spatiotemporal relations that are detectable within two-view motion sequences.

GENERAL DISCUSSION

There are two general types of analyses that have been proposed in the literature for computing an object's 3-D structure from motion: two-frame models that are restricted to the analysis of first-order spatiotemporal relations between pairs of views, and multiframe models that are capable of exploiting the higher order relations among three or more views. One two-frame analysis proposed by Ullman (1977) relies on a decomposition of image displacements into distinct components of optic flow: rotation about the line of sight (image rotation), and rotation

about an orthogonal axis in the picture plane. If the removal of image rotation results in a pattern of parallel trajectories, there is a strong probability that the depicted object is undergoing rigid motion. A similar analysis has also been proposed by Koenderink and van Doorn (1991).

The goal of Experiment 1 was to evaluate the basic predictions of a two-frame analysis of structure from motion: (1) that parallelism of image point trajectories should be a crucial factor for determining an object's perceived rigidity, and (2) that perceived rigidity should be unaffected by image rotation. The results of this experiment were perfectly consistent with the predictions of a two-frame model. Almost without exception, observers based their ratings on whether or not the points in the image were moving along parallel trajectories. Furthermore, the addition of image rotation to the displays had no significant effect on perceived rigidity. The result is also in agreement with earlier psychophysical tests which demonstrated the ability of observers to discriminate rigid from nonrigid motion given only two distinct views (Braunstein et al., 1990; Todd et al., 1988).

The second experiment was designed to provide a converging source of evidence relating to this issue. In a simplification of the Todd (1982) paradigm, geometric relationships between image point trajectories were varied systematically to study the relative salience of two-frame versus multiframe properties. The results confirmed the basic finding from Experiment 1 that parallelism of image trajectories is the single most important factor affecting observer's perceptions of rigidity. The results also demonstrated that observers can obtain useful information from higher order spatiotemporal relations under certain conditions.

Experiments 3 and 4 were designed to investigate the ability of observers to discriminate rigid rotations from nonrigid deformations with parallel image trajectories. One important difficulty with this type of discrimination paradigm is that it is mathematically impossible to eliminate the presence of all possible artifactual cues (i.e., those

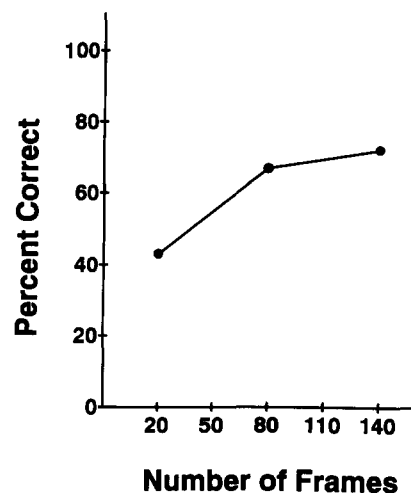


Figure 14. The percentage of correct responses for each sequence length in Experiment 4, collapsed over display type.

that have nothing to do with the analysis of rigidity *per se*). This problem is exacerbated, moreover, by providing immediate feedback after every trial, which increases the probability that observers would discover the utility of these cues to facilitate their performance. What is most striking about the results, however, is that the observers' strategies for performing these discriminations were surprisingly unsuccessful, regardless of whether they were based on an analysis of higher order spatiotemporal relations among three or more views or the detection of artifactual sources of information. Even when the motion sequences contained 280 distinct frames, the observers could only discriminate the rigid and nonrigid displays with an accuracy of about 80%, and their performance was barely above chance when the displays included a component of image rotation.

It is also interesting to consider the observers' subjective impressions while performing these experiments. For Experiments 1 and 2, which included displays with first-order structural deformations, all of the observers reported that they had a high degree of confidence in their ratings—even though they were given no response feedback about the accuracy of their responses. Displays given high ratings appeared clearly rigid, and those given low ratings appeared clearly nonrigid. For Experiments 3 and 4, in contrast, where they were forced to discriminate between rigid motions and higher order structural deformations, all of the observers commented that the task seemed more difficult. With the exclusion of first-order structural deformations, the appearance of nonrigidity in the displays was greatly reduced, and it is likely that the task would have been impossible for naive observers had they not been provided with immediate feedback after every trial.

During the past several years, there have been numerous psychophysical investigations assessing the extent to which human observers can utilize higher order spatiotemporal relations across multiple views of an apparent motion sequence. The basic design of most of these studies has been to compare performance for some type of 3-D judgment using apparent motion sequences composed of varying numbers of discrete frames. The results of this research have been somewhat contradictory, in that some investigators have reported systematic improvements with increasing numbers of views (e.g., Todd, 1982), whereas others have not (e.g., Todd & Bressan, 1990).

To interpret the results of such experiments, many subtleties need to be considered. Suppose, for example, that a researcher presents moving displays at a rate of 100 frames per second and with a 3-D angular displacement of 0.1° per frame. It is clear from previous research that these parameters are adequate to produce compelling kinetic depth effects. However, if two successive frames from such a sequence were presented in continuous alternation, they would not be perceived as 3-D motion because the displacement would be too small and the rate of alternation would be too fast to be detected by the human visual system. Comparing such an inadequate two-frame stimulus with the results obtained over multiple views would be highly misleading about the possible use of higher order spatiotemporal relations in the perception of structure

from motion. The problem here is that a multiple-frame sequence contains many different pairs of views that could potentially be analyzed. Thus, in order to demonstrate that human observers are able to exploit higher order spatiotemporal relations, it is necessary to show that perceptual performance for multiple-view sequences is significantly greater than would be possible for *any* individual pair of those views presented in isolation. This is a difficult criterion to satisfy methodologically, because of the potentially large number of pairs that would have to be considered.

One way to approach this problem, suggested by Todd et al. (1988), Todd and Bressan (1990), and Todd and Norman (1991), is to use only two-view displays whose parameters have been optimized for the best possible perceptual performance. These researchers have noted that several different factors can dramatically affect the perceptual compellingness of two-frame apparent motion sequences. One particularly important factor is the rate of alternation. Two-frame sequences work best when the stimulus onset asynchrony (SOA) between frames is 200 msec or more, whereas longer length sequences appear most compelling with SOAs of 50 msec or less. Another important factor is the structural configuration of a display. Todd et al. (1988) found that the most perceptually coherent two-frame sequences were those that depicted patterns of connected lines or dense configurations of dots on smoothly curved opaque surfaces, and that coherence was significantly decreased for moving patterns composed of random dots in a volume.

It is likely that at least some of the discrepant results concerning the effects of views in perceived structure from motion are due to the choice of these critical stimulus parameters. For example, Lister et al. (1994) have recently reported that observers have difficulty maintaining a coherent perception of structure from motion for two-frame motion sequences of random dots in a volume, and that ratings of subjective depth for such displays increase with the number of views—although they found no effect of views for other types of 3-D judgments. Todd and Norman (1991), in contrast, used two-frame motion sequences depicting dense configurations of dots on smoothly curved opaque surfaces for which observers have no difficulty maintaining a coherent perception of structure from motion. Their results showed that increasing the number of views beyond two had only negligible effects on observers' ratings of subjective depth.

Another important factor that needs to be considered in comparing performance across varying numbers of views is the existence of orientation anisotropies in the visual perception of structure from motion (e.g., see Cornilleau-Péres & Droulez, 1989; Norman & Lappin, 1992; Tittle, Todd, Perotti, & Norman, 1995; Todd & Norman, 1991). When an object rotates in depth, its orientation must necessarily change relative to the observer's line of sight. In comparing results between apparent motion sequences of different lengths, it is important to consider two-view performance across the entire range of orientations that can occur for the multiview displays. For example, Todd and Norman found that different pairs of views produced dramatic variations in performance on a surface discrimination

task, but that observers' responses to the best two-view sequences were indistinguishable from their judgments of multiview displays. Note in this case that it would have been easy to conclude erroneously that there was a significant effect of views if the experimental design had only included a single two-view sequence that was less than optimal (e.g., see Johnston, Cumming, & Landy, 1994).

Still another important factor that can complicate a comparison between two-view and multiview displays in detection tasks is the potential role of probability summation. Suppose, for example, that observers are asked to detect the presence of a nonrigid deformation in an object's depicted motion (see, e.g., Braunstein et al., 1990; Todd & Bressan, 1990). In order to determine the expected outcome from a strictly two-view analysis, it would again be necessary to consider the impact of all possible pairs of views. If the probability of detecting the deformation in a two-frame display is p , then the probability of detecting it from n independent pairs would be $\{1 - (1-p)^n\}$, which increases rapidly with the magnitude of n . Thus, one could only conclude that observers were exploiting higher order spatiotemporal relations if performance increased with the number of views at a rate that exceeded the predicted result based on simple probability summation.

Because of these difficulties in interpreting performance differences between two-view and multiview displays, the research described in the present article was designed to provide a converging methodology for evaluating observer sensitivity to the higher order spatiotemporal relations among three or more frames of an apparent motion sequence. Observers were presented with several different types of displays, many of which involved nonrigid distortions that would be inherently undetectable using only the first-order relations between individual pairs of views. The advantage of this procedure is that the rigid and nonrigid displays could all be presented with arbitrarily long motion sequences using exactly the same timing parameters. Using a two-view analysis of structure from motion applied to any arbitrary pair of views in the sequence, the nonrigid distortions with parallel trajectories should have appeared perfectly rigid. The fact that this was not always the case in Experiments 2 and 4 thus provides clear evidence that observers are able to make use of higher order spatiotemporal relationships among three or more views.

It is also important to recognize, however, that the utility of this higher order information appears to be quite limited. In all of these experiments, the nonrigid distortions that were detectable with a two-frame analysis appeared much more nonrigid than displays whose distortions required an analysis over three or more views. These findings are in general agreement with the similar findings of Norman and Todd (1993) on the perceived rigidity of affine stretching transformations, and with numerous other studies of how the perceived 3-D structure of objects is affected by the number of distinct frames in an apparent motion sequence (LITER et al., 1994; Todd & Bressan, 1990; Todd & Norman, 1991).

Although the results of the present experiments provide clear evidence that higher order spatiotemporal relations have some effect on the visual perception of structure from motion, it cannot be determined from these data how this information is actually used. One plausible hypothesis has recently been suggested by Hogervorst et al. (1996). Their model involves an analysis of the phase space trajectories between pairs of moving elements. For objects rotating rigidly about a fixed axis in the image plane, these trajectories must always form an ellipse centered about the origin (cf. Todd, 1982). If not, then the depicted motion must be nonrigid. This analysis was proposed originally to account for the results of Experiment 1 of Norman and Todd (1993), but it does not appear to generalize to the displays employed in the present experiments. Note in Figures 4 and 7, for example, that the two displays that deviated most from an elliptical trajectory were the constant flow fields of Experiment 1 and the variable frequency displays of Experiment 2, yet those were the two types of distortions that appeared perceptually to be the most rigid. The question of why some displays with parallel image trajectories appear rigid while others do not will have to wait for future research.

REFERENCES

- BENNETT, B., & HOFFMAN, D. (1985). The computation of structure from fixed axis motion: Nonrigid structures. *Biological Cybernetics*, **51**, 293-300.
- BENNETT, B., HOFFMAN, D., NICOLA, J., & 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., & ANDERSEN, G. J. (1984). Shape and depth perception from parallel projections of three-dimensional motion. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 749-760.
- 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.
- CORNILLEAU-PÉRÈS, V., & DROULEZ, J. (1989). Visual perception of curvature: Psychophysics of curvature detection induced by motion parallax. *Perception & Psychophysics*, **46**, 351-364.
- HOFFMAN, D., & BENNETT, B. (1985). Inferring the relative three-dimensional positions of two moving points. *Journal of the Optical Society of America A*, **2**, 242-249.
- HOFFMAN, D., & BENNETT, B. (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.
- HOGERVORST, M. A., KAPPERS, A. M. L., & KOENDERINK, J. J. (1996). Structure from motion: A tolerance analysis. *Perception & Psychophysics*, **58**, 449-459.
- HUANG, T., & LEE, C. (1989). Motion and structure from orthographic projections. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **11**, 536-540.
- JOHNSTON, E. B., CUMMING, B. G., & LANDY, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, **34**, 2259-2275.
- KOENDERINK, J. J., & VAN DOORN, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A*, **8**, 377-385.

- LITER, J. C., BRAUNSTEIN, M. L., & HOFFMAN, D. D. (1994). Inferring structure from motion in two-view and multi-view displays. *Perception*, **22**, 1441-1465.
- NORMAN, J. F., & LAPPIN, J. S. (1992). The detection of surface curvatures defined by optical motion. *Perception & Psychophysics*, **51**, 386-396.
- NORMAN, J. F., & TODD, J. T. (1993). The perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. *Perception & Psychophysics*, **53**, 279-291.
- TITTLE, J. S., TODD, J. T., PEROTTI, V. J., & NORMAN, J. F. (1995). Systematic distortions of perceived three-dimensional structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception & Performance*, **21**, 663-678.
- TODD, J. T. (1982). Visual information about rigid and nonrigid motion: A geometric analysis. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 238-251.
- 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.

(Manuscript received September 20, 1994;
revision accepted for publication September 16, 1995.)