In W. Epstein & S. J. Rogers (Eds.) Handbook of perception and cognition, Volume 5: Perception of space and motion. Orlando, FL: Academic Press. pp. 201-226.

The Visual Perception of 3-Dimensional Structure from Motion

James T. Todd

The Ohio State University

One of the most perplexing phenomena in the study of human vision is the ability of observers to perceive the 3-dimensional layout of the environment from patterns of light that project onto the retina. Indeed, were it not for the facts of our day-to-day experiences, it would be tempting to conclude that the perception of 3-dimensional form is a mathematical impossibility, since the properties of optical stimulation appear to have so little in common with the properties of real objects encountered in nature. Whereas real objects exist in 3-dimensional space and are composed of tangible substances such as earth, metal or flesh, an optical image of an object is confined to a 2-dimensional projection surface and consists of nothing more than flickering patterns of light. Nevertheless, for many animals including humans these seemingly uninterpretable patterns of light are the primary source of sensory information about the layout of objects and surfaces in the surrounding environment.

There are many different aspects of optical stimulation that are known to provide perceptually salient information about an object's 3-dimensional form. Some of these properties -- the so called pictorial depth cues -- are available within individual static images. These include texture gradients, linear perspective and patterns of shading. Others are defined by the systematic transformations among a sequence of multiple images, including the disparity between each eye's view in binocular vision, and the optical deformations that occur when objects are observed in motion.

This chapter is designed to review our current knowledge about how observers are able to perceive an object's 3-dimensional structure from image motion. The chapter is organized into several parts: It will provide 1) a formal analysis of the specific patterns of optical transformations that are produced by different types of rigid body motions; 2) an historical overview of various factors that can influence observers' perceptions of structure from motion; 3) a summary of existing computational models of how image motion could be theoretically analyzed; 4) a review of current psychophysical evidence about the psychological validity of these models; and 5) a discussion of several issues that remain for future research.

Optical Projection

In order to appreciate existing research and theory on the visual perception of structure from motion, it is first necessary to consider how the physical motions of objects in 3-dimensional space influence the patterns of optical stimulation at a point of observation. The top portion of Figure 1 shows the geometry that arises when an observer views a set of points in 3-dimensional space through a planar projection surface (i.e., a window). The position of a point **P**(x, y, z) at any given instant of time (t) can be defined within a Cartesian coordinate system, whose origin is located at the point of observation, and whose z-axis is parallel to the line of sight. If the image plane is located a unit distance from the origin along the z-axis, then the projected position **P'**(x', y') of the point is defined by the following equations:

1) $$x' = x / z$$

2) $$y' = y / z$$

If the point has an instantaneous velocity **V(**dx/dt, dy/dt, dz/dt), then its projected velocity **V'**(dx'/dt, dy'/dt) in the image plane is given by:

3) $$dx'/dt = \frac{dx/dt}{z} - \frac{x(dz/dt)}{z^2}$$

4) $$dy'/dt = \frac{dy/dt}{z} - \frac{y(dz/dt)}{z^2}$$

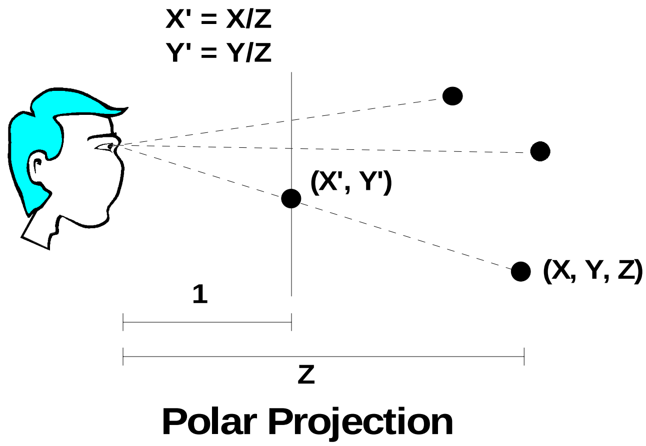
The geometry shown in the upper portion of Figure 1 is often referred to as polar or central projection, because the rays connecting each visible point in 3-dimensional space with its corresponding point in the image plane all converge at the point of observation. This geometry is appropriate to model almost any situation in natural vision. The lower portion of Figure 1 shows another alternative geometry called parallel or orthographic projection that is a reasonable approximation to natural vision whenever an observed object's extension in depth is relatively small in comparison to its distance from the observer. Using parallel projection the image plane positions **P'** and velocities **V'** are defined by the following equations:

5) $$x' = x$$

6) $$y' = y$$

7) $$dx'/dt = dx/dt$$

8) $$dy'/dt = dy/dt$$



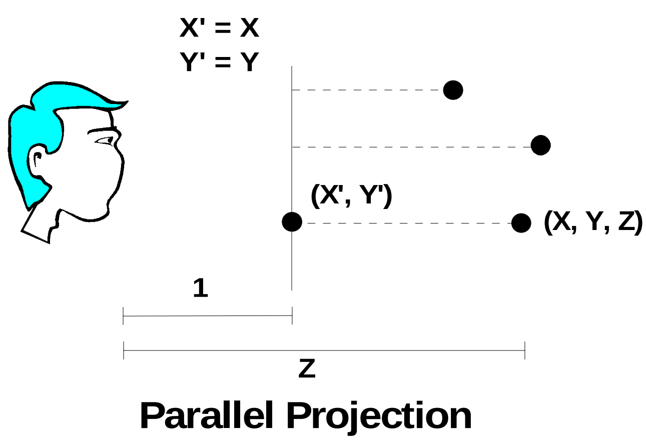


Figure 1 -- Two types of optical projection used in the analysis of structure from motion

Note in Figure 1 that if the depicted configuration of points under polar projection were moved farther and farther from the point of observation, the angular difference between their visual directions would become closer and closer to zero, and the resulting optical pattern would become a closer and closer approximation to a true parallel projection. One way of quantifying the appropriateness of this approximation is to measure the *perspective ratio* of an object's extension in depth relative to its distance from the observer (cf Braunstein, 1962). As a general rule of thumb, any observed object with a perspective ratio smaller than 0.1 can for all practical purposes be considered as a parallel projection.

Figure 2 shows the patterns of optical motion produced by different types of rigid body motions under both parallel and polar projection. Let us first consider the case of rigid translation. Note in Figure 2 that the projected motion of an object translating under parallel projection provides no information about the object's 3-dimensional form. In the special case of translation in depth there is no optical motion at all, and in all other cases the optical motion is identical for every point regardless of its position in depth (see equations 7 and 8). For objects translating under polar projection, however, there is potentially useful information available from the relative image motion of different points. As is evident from Equations 3 and 4, the optical velocities under polar projection are scaled by their positions in depth, such that far away points produce slower projected velicities than do points that are closer to the observer. For components of translation perpendicular to the line of sight this produces a pattern of parallel velocities in the image plane whose magnitudes vary as a function of depth. For components of translation that are parallel to the line of sight, this produces an overall image expansion for translations toward the observer, and an overall image compression for translations away from the observer. Such variations of image velocity as a function of depth are often referred to as *motion parallax.*

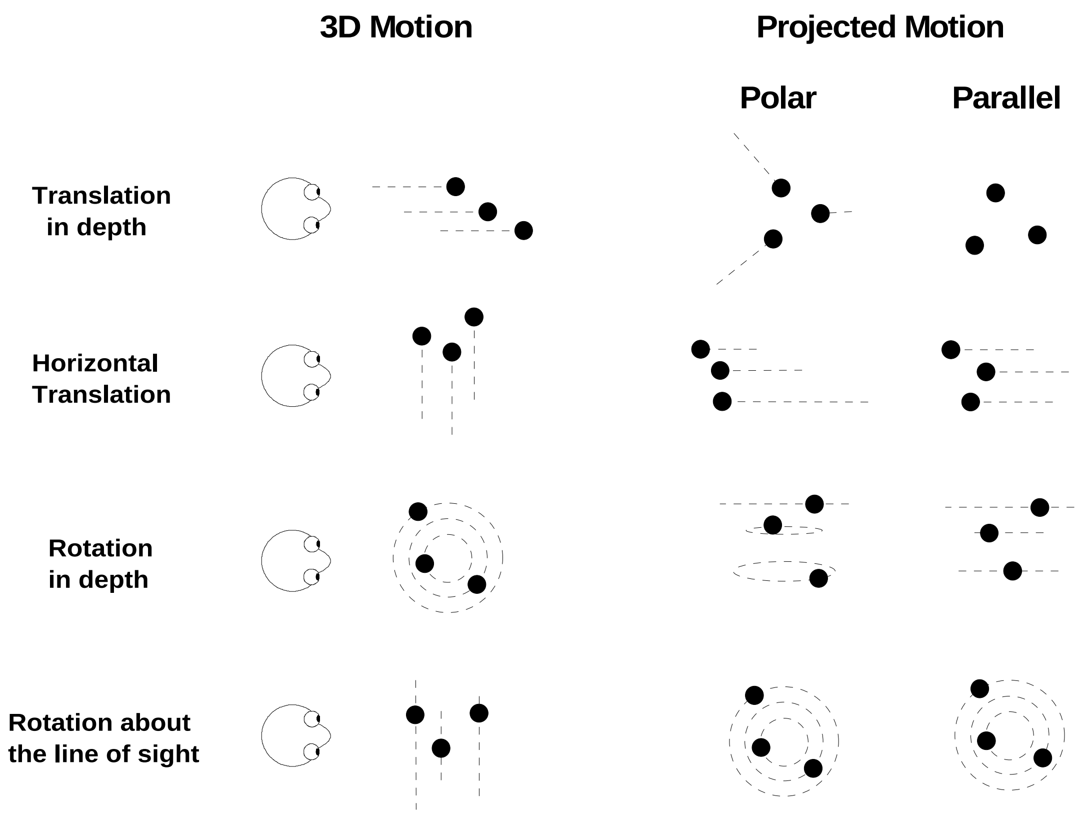


Figure 2 -- The optical projections produced by several different types of rigid body motions

The analysis of rigid body rotations is somewhat more complicated than the analysis of translation. For objects rotating about fixed axes, each point **P'** in the image plane will move in an elliptical trajectory whose minor axis coincides with the optical projection of the axis of rotation (see Todd, 1982). The eccentricities of these elliptical trajectories are determined by the slant of the rotation axis with respect to the image plane. For parallel projections, the trajectories of every point will all have the same eccentricity, whereas for polar projections the eccentricities will vary monotonically along the axis of rotation. For both types of projection, the image plane velocity **V'** is determined by the distance of a point in depth from the axis of rotation. The velocity reaches a maximum in one direction when a point is at its closest position to the observer in depth, and it reaches a maximum in the opposite direction when it is farthest away in depth. A degenerate case of rotary motion can occur when the axis of rotation is coincident with the line of sight. All of the image points in that case move along circular trajectories, and their relative instantaneous velocities are mathematically unrelated to an object's 3-dimensional structure.

Methodological issues

Display generation

The importance of motion for the visual perception of 3-dimensional form was discussed anecdotally in the writings of Mach (1886) and Helmholtz (1910), but a more systematic scientific investigation of this phenomenon did not occur until much later. In order to pursue such an investigation, it was first necessary to develop appropriate technologies for isolating the effects of optical flow from other potential sources of information such as shading, texture or binocular disparity. The earliest experiments on the visual perception of structure from motion used a shadow projection technique to satisfy this requirement (e.g., Metzger, 1934; Wallach & O'Connell, 1953; Gibson & Gibson, 1957; von Fieandt & Gibson, 1959; White & Mueser, 1960; Flock, 1964). Objects were placed on a moving track or turntable between a light source and a translucent display screen. Naive observers on the other side of the screen were then asked to report their subjective impressions while viewing the optical deformations of the objects' projected shadows. The most systematic investigation of this type was performed by Wallach & O'Connell (1953) using shadows cast by a wide variety of objects on a rotating turntable. For many of these displays observers spontaneously reported the perception of solid objects rotating in depth. Wallach & O'Connell named this phenomenon the *kinetic depth effect,* which is sometimes abbreviated in the literature as KDE.

In the early 1960's a new technology was invented that allowed researchers to create images of 3D objects on a computer controlled cathode ray tube (CRT). This technology was exploited almost immediately to study the perception of structure from motion in a pioneering series of experiments by Green (1960), Braunstein (1962, 1966, 1968) and Johansson (1964). When computer graphics first became available, the creation of 3D motion displays was a rather arduous process. Because laboratory computers in the 1960's were not fast enough to simulate and display complicated 3D motions in real time, each frame of the motion sequence was photographed individually so that the sequence could be played back in an experimental setting using a standard motion picture projector. As computer technology has advanced, however, this limitation has long since vanished, and most modern research in this area is now performed using real time displays.

Response tasks

There are several different response tasks that have been employed over the years to assess observers' perceptions of 3-dimensional structure from motion. In many of the earliest experiments on this topic, observers were simply asked to report their subjective experiences while viewing various types of moving displays e.g., see Metzger, 1934; Wallach & O'Connell, 1953; Gibson & Gibson, 1957; von Fieandt & Gibson, 1959; White & Mueser, 1960);. This is a good "quick and dirty" method to reveal qualitative aspects of an observer's perceptions, but it cannot provide precise quantitative information about an object's perceived shape or the nature of its perceived motion. Other common response tasks that are designed to overcome this difficulty include magnitude estimations of specific 3-dimensional properties such as depth or slant, or rating the perceived rigidity or coherence of an object's motion (e.g., see Braunstein, 1962, 1966, 1968; Gibson & Gibson, 1957; Green, 1960). More recent investigations have also employed discrimination procedures to measure how accurately observers can distinguish rigid from nonrigid motion (e.g., see Todd, 1982; Braunstein, Hoffman & Pollick, 1990) or to detect small differences in various aspects of 3-dimensional structure (e.g., see Braunstein, Hoffman, Shapiro, Andersen & Bennett, 1987; Todd & Bressan, 1990; Todd & Norman, 1991; Norman & Lappin, 1992).

One important issue that needs to be considered in selecting a response task is the extent to which it encourages observers to rely on artifactual sources of information (Braunstein & Todd, 1990; Sperling, Landy, Dosher & Perkins, 1989). This is especially true for experiments that employ discrimination procedures with response feedback. Suppose, for example, that observers are shown two objects oscillating back and forth about a vertical axis, and are asked to discriminate which one has the largest extension in depth. If both objects rotate at the same angular velocity, then their relative extensions in depth will covary linearly with the relative range of their projected image velocities. Thus, with the benefit of immediate response feedback, an observer could potentially learn to perform this task accurately, without knowing anything about the relative 3-dimensional structures of the depicted objects. One way of revealing if observers' judgments are based on artifactual sources of information is to ask them to describe their strategies for performing the experimental task. A better technique, however, is to identify these potential artifacts in advance and to systematically control for them. In the experiment described above, for example, the covariation between depth and image velocity can be eliminated by having the objects rotate at different angular velocities.

General factors that can affect perceived 3D structure from motion

Different types of motion

In the four decades that have elapsed since Wallach & O'Connell (1953) first reported their experiments on the kinetic depth effect, numerous other researchers have investigated how observers' perceptions of moving displays are influenced by a wide variety of stimulus variables. One important class of variables that has been studied extensively involves the specific nature of an objects depicted motion. It has been shown, for example, that observers' perceptions can be significantly influenced by the orientation of an object's axis of rotation (Green, 1961; Todd, 1982; Loomis & Eby, 1988, 1989). Ratings of depth and rigidity are highest for axes that are parallel to the image plane, and are reduced significantly for axes that are precessing or slanted in depth. Indeed, when the axis of rotation is at its most extreme possible slant such that it is coincident with the line of sight, a projected pattern of rotary motion generally produces no impression of 3-dimensional structure whatsoever. The one exception to this general rule occurs for image plane rotations of closed concentric contours, which can produce a compelling illusion of a moving 3-dimensional object (Musatti, 1924; Wallach, Weisz & Adams, 1956; Zanforlin, 1988; Proffitt, Rock, Hecht & Shubert, 1992). This phenomenon is often referred to as the *stereokinetic effect*.

There are several other aspects of an object's motion that can also influence observers' perceptions of its 3-dimensional structure. For example, the amount of perceived depth for a rotating object tends to increase with angular velocity (Todd & Norman, 1991) or the angular extent of rotation (Loomis & Eby, 1988, 1989; Liter, Braunstein & Hoffman, 1994). Similarly, the perceived rigidity or coherence of an object can vary significantly with the length of an apparent motion sequence and the timing between each frame (Todd, Akerstrom, Reichel, & Hayes, 1988; Todd & Bressan, 1990; Todd & Norman, 1991). For minimal apparent motion sequences consisting of only two frames presented in alternation, perceived rigidity is greatest when there are relatively long time intervals of about 200 msec between each frame transition. As the length of an apparent motion sequence is increased, however, the optimal time interval can be reduced to as low as 50 msec or less.

Perspective

Another important stimulus factor that can influence the perception of structure from motion is the amount of perspective used to create the experimental displays (Braunstein, 1966; Todd, 1984; Dosher et al, 1989). Objects viewed with high levels of perspective appear to have a greater extension in depth than those viewed under parallel projection. If the perspective is larger than what is appropriate for the actual viewing distance, then objects may sometimes appear to distort nonrigidly as they rotate. For motion parallax displays of objects undergoing translation, perspective is necessary to get any perceived depth at all. Perspective can also have an effect on the perceived direction of rotation (Andersen & Braunstein, 1983; Braunstein, 1977a; Braunstein, Andersen, & Riefer, 1982). For objects viewed under parallel projection the direction of rotation is completely ambiguous unless other information is available, such as the accretion or deletion of texture at occlusion boundaries. Under strong polar perspective, however, the true direction of rotation can be perceptually specified by the image plane velocity differences between near and far surface regions.

Structural Configuration

Still another important class of variables that can influence the perception of structure from motion includes specific aspects of a depicted object's structural configuration (Green, 1961; Braunstein, 1962; Petersik, 1980; Todd et al., 1988; Dosher et al., 1989). For example, it has been reported that patterns of connected lines produce greater perceived coherence than patterns of random dots, and that opaque surfaces appear more coherent than volumes or transparent surfaces. Increasing the number of moving elements can increase perceived coherence up to a point, but coherence can break down for very large numbers of elements on transparent surfaces or in random volumes, probably because of difficulties in matching elements over time.

An object's perceived 3-dimensional structure can also be affected by how it is oriented in space. Research has demonstrated that the perception of surface curvature exhibits a large anisotropy with respect to how an object is oriented relative to the direction of rotation. Observers are most sensitive to curved surfaces when they are curved in a direction parallel to the axis of rotation, (Cornilleau-Peres, & Droulez, 1989; Norman & Lappin, 1992). Similarly, other experiments have shown that object discriminations and magnitude estimates of perceived shape can be dramatically influenced by how an object is oriented with respect to the observer's line of sight (Reichel & Todd, 1990; Todd & Norman, 1991; Tittle, Todd, Perotti & Norman, 1994).

Conflicting sources of information

One final class of variables that is known to have a significant effect on observer's perceptions of structure from motion is the presence of conflicting cues from other potential sources of information. These effects can be observed informally by viewing the projected image of a rotating object while alternately opening and closing one eye. When such a display is observed monocularly it appears to have a greater extension in depth than when it is observed binocularly, because, in the latter case, there is conflicting information from binocular disparity to indicate that the pattern is confined to the image plane.

Most of the existing research on how other sources of information can conflict with motion has been primarily concerned with information about static slant, such as texture gradients, linear perspective or patterns of occlusion (Ames, 1951; Braunstein, 1968, 1971; Braunstein & Payne, 1968; Braunstein & Stern, 1980; Mingolla, & Todd, 1981). A particularly compelling phenomenon arising from conflicts between motion and static slant information can be obtained by direct viewing of a rotating trapezoid, as was first reported by Ames (1951). When a static trapezoid is observed monocularly in the fronto-parallel plane, it is typically perceived as a rectangle slanted in depth. When the object begins to rotate, however, the pattern of image motion provides conflicting information about its true orientation. This conflict can result in some startling perceptual illusions. Under appropriate conditions, a trapezoid in continuous rotation will appear instead to oscillate. It may also appear to undergo severe nonrigid deformations or even to pass through other solid objects.

Theoretical Analyses

Most research on the visual perception of structure from motion that was performed prior to 1980 was basically exploratory in nature. The closest thing resembling a theory that was available to these early researchers to guide their empirical investigations was a hypothesis by Wallach & O'Connell (1953) that the perceptually relevant information for the kinetic depth effect consisted of simultaneous changes in the projected lengths and angles of moving line elements. Unfortunately, however, this hypothesis proved to be of little value for explaining how moving patterns are perceptually analyzed or even to predict their perceptual appearance. For example, it was demonstrated early on that compelling kinetic depth effects can be obtained from length changes alone (Metzger, 1934; White & Musser, 1960; Johansson & Jansson, 1969; Borjesson & vonHofsten, 1972, 1973; Braunstein, 1977b), thus providing strong evidence that simultaneous changes in length and angle do not constitute a necessary condition for perceiving structure from motion. Nor do they constitute a sufficient condition, as can be demonstrated by the fact that a pattern of line segments whose lengths and angles are changed at random in different directions will appear as nothing more than random deformations in the image plane.

The computational analysis of multiple-view motion sequences

The first computational analysis of how it might be possible to determine an object's 3-dimensional structure from its pattern of projected motion was developed by Ullman (1977, 1979, 1983). Ullman's analysis was designed to be used with configurations of discrete points rotating in depth under parallel projection, in which corresponding images of each point could be identified over successive intervals of time. Given these conditions, he was able to prove that a unique rigid interpretation could be obtained (up to a reflection in depth) provided that a display contains at least three distinct views of four noncoplanar points. Subsequent analyses have confirmed that these minimum numbers of points and views are both necessary and sufficient for arbitrary configurations of moving elements, but that these minimal conditions can vary somewhat in certain special case situations (e.g., see Bennett & Hoffman, 1986; Hoffman & Bennett, 1985, 1986). For example, if the motion is constrained to be at a constant angular velocity about a fixed axis of rotation that is parallel to the image plane, then a unique rigid interpretation can be obtained from three views of only two identifiable points (Hoffman & Bennett, 1986).

Ullman (1984) later developed an incremental rigidity scheme for computing structure from motion that was designed to cope with objects whose motions are not perfectly rigid. This analysis attempts to maintain an internal model of an observed object and to modify that model at each instant of time by the minimal amount of change required to account for the observed transformation. One problem with this approach is that it does not always converge on a stable interpretation when applied to rigid configurations, and that even when it does converge it may only do so after an object has rotated in depth for several complete oscillations (Grzywacz & Hildreth, 1987).

Still another class of models have been developed by Webb & Agaarval (1981) and Todd (1982) to compute an object's 3-dimensional structure from the relative trajectories of each element's projected motion, rather than their positions at discrete moments in time (see also Johansson, 1974). These models have identified specific conditions for distinguishing rigid from nonrigid motion, and they are applicable to both parallel and polar projections. They share one of the limitations of incremental rigidity schemes, however, by requiring that an object's motion be observed over a sufficiently long interval of time to specify the structure of each element's projected trajectory.

<u>The computational analysis of two-view motion sequences</u>

In contrast to the analyses of multiple frame apparent motion sequences described above, other theorists have focused their attentions on the potential information that is available within the instantaneous field of image velocities defined by 2-frame motion sequences. The qualitative structure of these instantaneous velocity fields was first described by Gibson (1950, 1967, 1979), who was primarily concerned with how different patterns of optical flow provide information about an observer's movements relative to a fixed rigid environment. A more formal mathematical description was later developed by Gibson, Olum, & Rosenblatt (1958) and Gorden (1965).

There are several possible strategies that have been proposed for computing an object's 3-dimensional structure from its instantaneous field of image velocities under polar projection. For example, if the velocity of the observer is known, then that knowledge can be used to compute the position of any visible point in a rigid environment (e.g., see Lee, 1974; Nakayama & Loomis, 1974). Without knowledge of the observer's velocity, a unique 3-dimensional interpretation can still be obtained from the optical displacements of a small number of identifiable points (e.g., see Longuet-Higgins, 1981; Nagel, 1981), provided that they are viewed with a sufficient amount of perspective. Another possible strategy first proposed by Koenderink & van Doorn (1975, 1977, 1986) and Koenderink (1986) makes use the differential invariants within smooth velocity fields to determine certain aspects of local surface structure such as relative slant or the sign of Gaussian curvature. Similar analyses have also been developed by Longuet-Higgins & Prazdny (1980) and Waxman & Ullman (1985) that use differential invariants of optical flow to determine an object's complete euclidean metric structure.

For object's viewed under parallel projection, an instantaneous field of image velocities is somewhat less informative, but it does contain sufficient information to distinguish rigid from nonrigid motion, and to constrain an object's 3-dimensional structure to a one parameter family of possible rigid interpretations (Ullman, 1977, 1983; Bennett, Hoffman, Nicola, & Prakash, 1989; Huang & Lee, 1989; Todd & Bressan, 1990; Koenderink & van Doorn, 1991). A unique 3-dimensional interpretation can be obtained, however, for the special case of object motions that are confined to a fixed plane (Hoffman & Flinchbaugh, 1982; Lappin, 1990). It is also possible to obtain a unique solution for more general configurations by imposing additional constraints (e.g., see Aloimonos & Brown,1989).

Empirical tests of the Computational Models

As computational models for determining 3-dimensional structure from motion have begun to proliforate in the literature, there has been a complimentary effort to empirically investigate how closely the capabilities and limitations of these models correspond with those of actual human observers. Much of this research has centered on two basic issues: What are the specific aspects of object structure that an analysis of motion should be designed to compute, and what assumptions (i.e., constraints) are required to be satisfied for an analysis to function effectively? Whereas most existing computational models are designed generate precise metrical descriptions of 3-dimensional form within a narrowly constrained context, there is a growing amount of evidence to suggest that human perception is primarily concerned with more qualitive aspects of object structure and that it can function effectively over a surprisingly broad range of viewing conditions.

Perceived 3D structure from nonrigid configurations

One important limitation of existing computational analyses of 3-dimensional structure from motion is that they all involve some form of rigidity hypothesis. Most models require that an object's motion must be globally rigid, though there are others that can tolerate euclidean bendings that preserve distances measured along a surface (e.g., Koenderink & van Doorn, 1986) or piecewise rigid motions composed of locally rigid parts whose relative spatial arrangements can deform arbitrarily (e.g., Hoffman & Flinchbaugh, 1982; Todd, 1982). As a general rule, these models do not degrade gracefully. That is to say, if there is no possible rigid interpretation, then they will be unable to determine anything at all about a depicted objects's 3-dimensional structure.

There is considerable evidence to indicate, however, that the perceptual processes of actual human observers do not behave in this manner. There have been numerous studies reported in the literature that have employed various types of nonrigid displays, including both locally rigid bending transformations (Jansson, 1977; Jansson & Johansson, 1973; Jansson & Runeson, 1977) and stretching transformations that are locally nonrigid (e.g., Todd, 1982, 1984; Braunstein & Andersen, 1984; Cutting, 1987; Braunstein, Hoffman & Pollick, 1990). What is surprising about these displays from the perspective of current theory is that they can produce compelling kinetic depth effects of objects moving in 3-dimensional space -- albeit nonrigidly. Indeed, in those studies that have examined perceived 3-dimensional structure rather than rigidity (i.e., Todd, 1984; Braunstein & Andersen, 1984), the results have indicated that a nonrigid stretching transformation in one direction can have little or no effect on perceived surface curvature in an orthagonal direction.

Effects of number of views

Perhaps the most well known result from theoretical analyses of structure from motion is that a unique euclidean interpretation of an arbitrary configuration of points under parallel projection requires a minimum of three distinct views. This finding defines the minimum amount of information required for an ideal observer who can measure the projected position of each point and perform all subsequent computations on those measures with perfect accuracy. It would not be surprising, however, if additional amounts of information were required for real observers, whose perceptual processes may be less than perfect.

In an effort to compare the performance of this theoretically ideal observer with the processes of human perception, numerous investigators have examined how the number of distinct frames in an apparent motion sequence influences observers' judgements of rigidity (Lappin, Doner & Kottas,1980; Doner, Lappin, & Perfetto, 1984; Petersik, 1987; Todd, Akerstrom, Reichel & Hayes, 1988; Braunstein, Hoffman & Pollick, 1990) and 3-dimensional form (Braunstein, Hoffman, Shapiro, Andersen & Bennett, 1987; Hildreth, Grzywacz, Adelson & Inada, 1990; Todd & Bressan, 1990; Todd & Norman, 1991; Liter, Braunstein & Hoffman, 1994). Although it would be reasonable to expect based on current computational models that the perception of structure from motion should require a minimum of three distinct views, the empirical results have shown clearly that 2-frame motion sequences provide sufficient information to obtain compelling kinetic depth effects. Moreover, if a 2-frame sequence is presented in continuous alternation to eliminate any confounds with stimulus duration, increasing the sequence length with additional views has little or no effect on objective response tasks (Todd & Bressan, 1990; Todd & Norman, 1991; Liter, Braunstein & Hoffman, 1994). Simular results can also be obtained by limiting the lifetimes of individual points for objects viewed in continuous rotation if the overall level of noise is equated in all displays (Todd, 1985; Husain, Treue & Andersen, 1989; Treue, Husain & Andersen, 1991; Dosher, Landy & Sperling, 1990).

In order to make sense of these seemingly impossible results, it is important to recognize that the theoretical limits on the number of distinct views needed to compute 3-dimensional structure from motion are only applicable to the analysis of euclidean distance relations between arbitrary pairs of points. Other computational models have shown that 2-frame motion sequences are theoretically sufficient to perform tasks that do not require a precise determination of euclidean distance relations, including many types of object discrimation and the detection of nonrigid deformations (Ullman, 1977, 1983; Bennett, Hoffman, Nicola, & Prakash, 1989; Huang & Lee, 1989; Todd & Bressan, 1990; Koenderink & van Doorn, 1991).

It is interesting to note in surveying the psychophysics literature on this topic, that observers typically perform with high levels of accuracy on tasks that are theoretically possible with only two distinct views, but that performance can deteriorate dramatically for tasks that require a 3-view analysis of euclidean distance relations (Todd & Bressan, 1990; Todd & Norman, 1991; Norman & Todd, 1993). Such findings provide strong evidence that the human visual system may be incapable of performing the higher order time derivatives needed to analyze euclidean structure, and that it must rely instead on whatever information is available within the first order field of image velocities (see Todd, 1981).

Planar Motion and Perspective

There are some potential exceptions to the preceding conclusions that are useful to consider. Let us assume, as suggested above, that the perceptual analysis of structure from motion can only make use of first order velocity information that is available within 2-frame motion sequences. Such information would be mathematically insufficient to compute the euclidean metric structure of arbitrary configurations, but it would make it possible -- at least in principle -- to obtain a unique rigid interpretation of an object's 3-dimensional form in certain special-case situations. One such special case occurs for object motions that are confined to a fixed plane (Hoffman & Flinchbaugh, 1982; Lappin, 1990). Lappin & Love (1993) and Lappin & Ahlstrom (1994) have recently argued that human observers can indeed discriminate euclidean distance relations in this situation, though this result has been challanged by Pizlo and Salach-Golyska (1994) as arising from artifactual sources of information.

A second special case to consider includes objects viewed under strong polar perspective. Although perspective is known to increase the perceived depth of moving displays, the evidence does not suggest that it produces an accurate perception of euclidean metric structure. Most of the experiments relating to this issue have been specifically concerned with translatory motion perpendicular to the line of sight either with or without concomitant observer head movements (Gibson, Gibson, Smith & Flock, 1959; Flock, 1964; Farber & McKonkie, 1979; Rogers & Graham, 1979, 1982; Braunstein & Andersen, 1981; Rogers & Collett, 1989; Ono, Rivest & Ono, 1986; Braunstein & Tittle, 1988; Ono & Steinbach, 1990; Braunstein, Liter & Tittle, 1993; Caudek & Proffitt, 1993). The most typical pattern of results is that perceived depth is systematically underestimated relative to its true simulated value. There have also been a few studies on the accuracy of perceived 3-dimensional structure for objects rotating in depth under strong polar perspective. The results have shown that observers are quite good at estimating relative distance intervals in a given direction (Lappin & Fuqua, 1983), but that performance deteriorates dramatically for distance intervals that are oriented in different directions (Todd, Norman, Perotti & Tittle, 1993; Tittle, Todd, Perotti, & Norman, 1994).

Problems for Future Research

The scaling problem

Although there is a growing amount of evidence that computational models of 2-frame motion sequences under parallel projection share many of the properties of how human observers perceive structure from motion, there are other aspects of the psychophysical data that these models cannot explain. It is important to keep in mind that a 2-frame display of an arbitrary configuration under parallel projection does not have a unique rigid interpretation. It will either have no possible rigid interpretation at all, or an infinite one parameter family of possible interpretations. This would explain why observers typically exhibit large errors in judgements of euclidean metric structure from motion, and why they are unable to discriminate different structures within the one parameter family even when a motion sequence contains more than two distinct frames (e.g., see Todd & Bressan, 1990; Todd & Norman, 1991; Liter, Braunstein & Hoffman, 1994). The aspect of the data that is hard to explain, however, is the existence of systematic biases in observers' magnitude estimations of perceived depth. If the available information is infinitely ambiguous, then why should an object appear to have any specific depth at all? To the extent that it does, there would have to be some other constraint or heuristic at work to restrict the set of possible perceptual interpretations.

There have been a number of hypotheses proposed in the literature about what this constraint might be. For example, Loomis & Eby (1988, 1989) have proposed that perceived depth from motion is scaled by the amount of shear that is present within the overall pattern of optical flow (cf Koenderink & van Doorn, 1975, 1977, 1986). Others have suggested that perceived depth is scaled by the amount of compression within the overall flow pattern (Braunstein, Liter & Tittle, 1993), the magnitude of relative motion following the removal of image curl (Liter, Braunstein & Hoffman, 1994), or a compactness constraint, in which it assumed that an object's depth will be approximately equal to its width (Caudek & Proffitt, 1993). Unfortunately, however, none of these suggestions is particularly general. Each one was devised to account for the results of a relatively small set of experiments, but none of them is consistent with the entire corpus of data that has been reported in this area.

The perceptual representation of 3-dimensional form

The fact that observers can misperceive an object's extension in depth relative to its width while correctly identifying that it is undergoing rigid rotation leads to an interesting conundrum. Suppose, for example, that an observer underestimates distances in depth by a factor of two (e.g., see Wagner, 1985). If such an observer were to view a rotating ellipsoid whose extension in depth is twice its width at a particular moment in time, it should appear at that moment as a sphere. At a later point in its rotation cycle, however, its width would be twice its depth and it should appear as a flattened ellipsoid. Why wouldn't this change in shape be perceived as a nonrigid deformation? This puzzle was first noted by Helmholtz in considering the systematic distortions of stereoscopic space, but it is also applicable to the visual perception of structure from motion.

One possible resolution of this conundrum, first suggested by Gibson (1979), is that euclidean metric distances in 3-dimensional space are not a primary component of an observer's perceptual experience. This hypothesis has been developed more fully in a recent series of papers by Todd & Reichel (1989), Todd & Bressan (1990), Todd & Norman (1991), Norman & Todd (1992, 1993) and Tittle et. al. (1994). These authors have presented evidence that an observer's knowledge of 3-dimensional form may involve a hierarchy of different perceptual representations. Their findings indicate that observers are quite accurate and reliable at judging an object's topological, ordinal, or affine properties, and that perception of rigid motion occurs when these properties remain invariant over time. Although observers can exhibit a conceptual understanding of euclidean metric structure, this knowledge may be more cognitive than perceptual. The available psychophysical evidence suggests that if observers are required to make judgments about lengths or angles of visible objects in 3-dimensional space, they will resort to using ad hoc heuristics, which typically produce low levels of accuracy and reliability, and which can vary unpredictably among different individuals or for different stimulus configurations.

Analyses of different types of optical deformations

All of the research described thus far has been exclusively concerned with the optical displacements of identifiable features such as reflectance contours or the edges of polyhedra, for which multiple views of any given image point must all correspond to the same physical point in 3-dimensional space -- what is sometimes referred to as the condition of projective correspondence (Todd, 1985). In natural vision, however, the overall pattern of optical stimulation can contain a variety of other structures such as occlusion contours, cast shadows, and smooth variations of surface shading, for which this condition need not be satisfied. This can have important theoretical implications for the analysis of 3-dimensional structure from motion. When objects move or are viewed stereoscopically, these different aspects of optical structure do not always change in the same way, and analyses that are designed to be used with one type of optical deformation will not in general be appropriate for others.

Consider, for example, the occlusion contour that forms the silhouette of a human head. If the head rotates in depth about a vertical axis, the optical contour that bounds its projection will be systematically deformed, but the locus of surface points to which it corresponds will also be continuously changing -- i.e., for a frontal view the occlusion contour will pass through the ears, and for a profile view it will pass through the nose. Analyses that assume projective correspondence will be of little use with this type of optical deformation, even as a local approximation. Indeed, it is often the case that the optical motion of the bounding contour will be in one direction while the projected motion of any identifiable point on that contour is in the opposite direction. (see Todd, 1985).

There are other types of image structure for which motions of the observer and motions of the observed object produce different patterns of optical deformation. When an observer moves in an otherwise rigid environment, visible objects will all maintain a constant relationship with their sources of illumination. Because shadow borders and gradients of lambertian shading in this context remain bound to fixed positions in 3-dimensional space, their resulting patterns of optical deformation will satisfy the condition of projective correspondence, and can therefor be analyzed using conventional techniques for determining structure from motion. When an object moves relative to its light source, however, shadow borders and

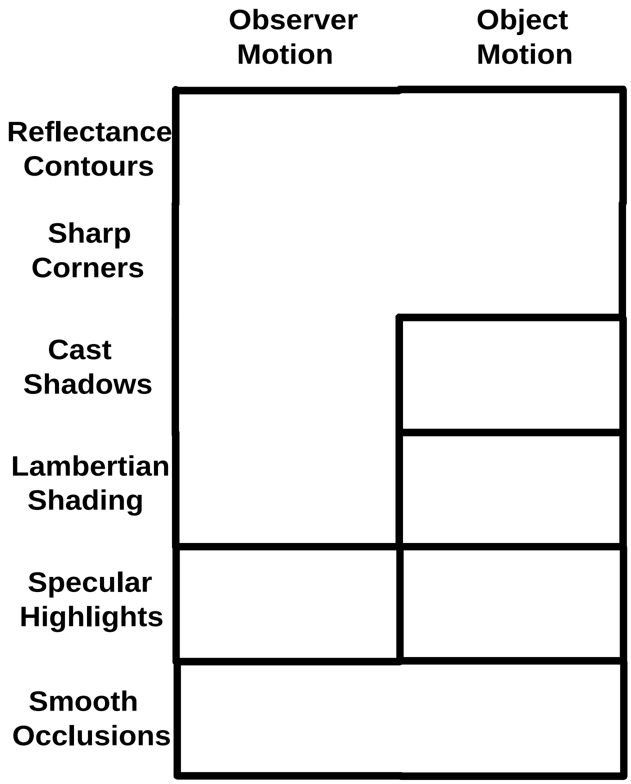


Figure 3 -- Different types of optical deformation

gradients of shading will move over its surface. Because this violates the condition of projective correspondence, existing computational models would be unable to generate a correct rigid interpretation.

Figure 3 provides a summary of all of the different categories of optical deformation described above. The rows of this figure represent different types of optical structure, while the columns are used to distinguish observer motion from object motion. Note that some of the borders between cells in this table have been removed. These open areas define classes of deformation that are formally equivalent. For example, the deformations of reflectance contours, sharp corners, cast shadows and lambertian shading caused by observer motion, and the deformations of reflectance contours and sharp corners caused by object motion are all formally equivalent in that they satisy the condition of projective correspondence. This is the category on which most of the existing research in this area has been focussed.

Although the remaining cells in Figure 3 have attracted much less attention they have not been ignored altogether. There have been several demonstrations reported in the literature that human observers can obtain compelling kinetic depth effects from the optical deformations of smooth occlusion contours (Todd, 1985; Cortese & Andersen, 1991; Pollick, Giblin, Rycroft & Wilson, 1992, Norman & Todd, 1994) and there have also been a few mathematical analyses of how this might be theoretically possible (Koenderink & van Doorn, 1977; Giblin & Weiss, 1987;

Cipolla & Blake, 1990). There is some evidence to suggest that the optical deformations of shadows and shading may provide useful information as well (Todd, 1985; Norman & Todd, 1994), but the generality of this evidence remains to be determined. One important factor that has limited research on these topics is the difficulty of creating controlled laboratory displays of moving shaded images. This difficulty is quickly diminishing, however, with the continuing advance of computer graphics technology, so that this is likely to be a more active area of research within the next several years.

References

Aloimonos, J., & Brown, C.M. (1989). On the kinetic depth effect. Biological Cybernetics, 60, 445-455.

Ames, A. (1951) Visual perception and the rotating trapizoidal window. Psychological Monographs, 65, (7, Whole No. 324).

Andersen, G. J. & Braunstein, M. L., (1983) Dynamic occlusion in the perception of rotation in depth. Perception & Psychophysics, 34, 356-362

Bennett, B. & Hoffman, D. (1986) 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, 6, 1052-1069.

Borjesson, E. & von Hofsten C. (1972) Spatial determinants of depth perception in two dot patterns. Perception & Psychophysics, 11, 263-268.

Borjesson, E. & von Hofsten C. (1973) Visual perception of motion in depth: Application of a vector model to three dot motion patterns. Perception & Psychophysics, 13, 203-208.

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. (1966) Sensitivity of of the observer to transformations of the visual field. Journal of Experimental Psychology, 72, 683-689.

Braunstein, M. L. (1968) Motion and texture as sources of slant information. Journal of Experimental Psychology, 78, 247-253.

Braunstein, M. L. (1971) Perception of rotation in figures with rectangular and trapezoidal features. Journal of Experimental Psychology, 91, 25-29.

Braunstein, M. L. (1977a) Perceived direction of rotation of simulated three-dimensional patterns. Perception & Psychophysics, 21, 553-557.

Braunstein, M. L. (1977b) Minimal conditions for perception of rotary motion. Scandanavian Journal of Psychology, 18, 216-223.

Braunstein, M.L., & Andersen, G.J. (1981) Velocity gradients and relative depth perception. Perception & Psychophysics, 29, 145-155.

Braunstein, M.L., & Andersen, G.J. (1984). Shape and depth perception from parallel projections of three dimensional motion. Journal of Experimental Psychology: Human Perception and Performance, 10, 749-760.

Braunstein, M. L., Andersen, G. J., & Riefer, D. M. (1982) The use of occlusion to resolve ambiguity in parallel projections. Perception & Psychophysics, 31, 261-267.

Braunstein, M.L., Hoffman, D.D., & Pollick, F.E. (1990). Discriminating rigid from nonrigid motion. Perception & Psychophysics, 47, 205-214.

Braunstein, M.L., Hoffman, D.D., Shapiro, L.R., Andrsen, G.J., & Bennett, B.M. (1987). Minimum points and views for the recovery of three-dimensional structure. Journal of Experimental Psychology: Human Perception and Performance, 13, 335-343.

Braunstein, M. L., Liter, J. C., & Hoffman, D. D. (1994) Inferring structure from two-view and multi-view displays. Perception, in press.

Braunstein, M. L. Liter, J. C. & Tittle, J. S. (1993) Recovering three-dimensional shape from perspective translations and orthographic rotations. Journal of Experimental Psychology: Human Perception and Performance, 19, 598-614.

Braunstein, M. L. & Payne, J. W. (1968) Perspective and the rotating trapezoid. Journal of the Optical Society of America, 58, 399-403.

Braunstein, M. L. & Stern, K. R. (1980) Static and dynamic factors in the perception of rotary motion. Perception & Psychophysics, 4, 313-320.

Braunstein, M.L., & Tittle, J. S. (1988) The observer relative velocity field as the basis for effective motion parallax. Journal of Experimental Psychology: Human Perception and Performance, 14, 582-590.

Braunstein, M.L., & Todd, J.T. (1990). On the distinction between artifacts and information. Journal of Experimental Psychology: Human Perception and Performance, 16, 211-216.

Caudek, C. & Proffitt, D. R. (1993) Depth perception in motion paralax and stereokinesis. Journal of Experimental Psychology: Human Perception and Performance, 19, 32-47.

Cipolla, R., and Blake, A. (1990). The dynamic analysis of apparent contours. Proceedings of the Third International Conference of Computer Vision, 616-623.

Cortese, J. M., and Andersen, G. J. (1991). Recovery of 3-D shape from deforming contours. Perception and Psychophysics, 49, 315-327.

Cornilleau-Peres, V., & Droulez, J. (1989). Visual perception of curvature; Psychophysics of curvature detection induced by motion parallax. Perception & Psychophysics, 46, 351-364.

Cutting, J.E. (1987). Rigidity in cinema seen from the front row, side aisle. Journal of Experimental Psychology: Human Perception and Performance, 13, 323-334.

Doner, J., Lappin, J.S., & Perfetto, G. (1984). Detection of three-dimensional structure in moving optical patterns. Journal of Experimental Psychology: Human Perception and Performance, 10, 1-11.

Dosher, B.A., Landy, M.S., & Sperling, G. (1989). Ratings of kinetic depth in multidot displays. Journal of Experimental Psychology: Human Perception and Performance, 15, 816-825.

Dosher, B.A., Landy, M.S., & Sperling, G. (1990). Kinetic depth effect and optic flow -- I. 3D shape from Fourier motion. Vision Research, 29, 1789-1814.

von Fieandt, K. & Gibson, J. J. (1959) The sensitivity of the eye to two kinds of continuous transformation of a shadow pattern. Journal of Experimental Psychology, 57, 344-347.

Farber, J. M. & McKonkie, A. B. (1979) Optical motions as information for unsigned depth. Journal of Experimental Psychology: Human Perception and Performance, 15, 494-500.

Flock, H. (1964) Some sufficient conditions for accurate monocular perceptions of surface slants. Journal of Experimental Psychology, 67, 560-572.

Giblin, P., and Weiss, R. (1987). Reconstruction of surfaces from profiles. Proceedings of the IEEE First International Conference on Computer Vision, 136-144.

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.

Gibson, J. J. (1950) The perception of the visual world. Boston: Haughton Mifflin.

Gibson, J. J. (1967) The senses considered as perceptual systems. Boston: Haughton Mifflin.

Gibson, J. J. (1979) The ecological approach to visual perception. Boston: Haughton Mifflin.

Gibson, J. J. & Gibson, E. J. (1957) Continuous perspective transformations and the perception of rigid motion. Journal of Experimental Psychology, 54, 129-138.

Gibson, J. J., Olum, P. & Rosenblatt, F. (1958) Parallax and perspective during aircraft landings. American Journal of Psychology, 68, 372-385.

Gorden, D. A. (1965) Static and dynamic visual fields in human space perception. Journal of the Optical Society of America, 55, 1296-1303.

Green, B.F., Jr. (1961). Figure coherence in the kinetic depth effect. Journal of Experimental Psychology, 62, 272-282.

Grzywacz, N., & Hildreth, E. (1987). Incremental rigidity scheme for recovering structure from motion: Position-based versus velocity-based formulations. Journal of the Optical Society of America, A4, 503-518.

Helmholtz, H. von (1910) Treatise on Physiological Optics (Translated by J. P. Southall, 1925). Dover, New York.

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. & 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, 1-13.

Hoffman, D.D., & Flinchbaugh, B.E. (1982). The interpretation of biological motion. Biological Cybernetics, 42, 195-204.

Huang, T., & Lee, C. (1989). Motion and structure from orthographic projections. IEEE Transactions on Pattern Analysis and Machine Intelligence, 11, 536-540.

Husain, M., Treue, S., & Andersen, R.A. (1989). Surface interpolation in three-dimensional structure-from-motion perception. Neural Computation, 1, 324-333.

Johansson, G. (1964) Perception of motion and changing form. Scandinavian Journal of Psychology, 5, 181-208.

Johansson, G. (1974) Visual perception of rotary motions as transformations of conic sections. Psychologia, 17, 226-237.

Johansson, G. & Jansson, G. (1968) Perceived rotary motion from changes in a straight line. Perception & Psychophysics, 6, 193-198.

Jansson, G. (1977) Perceived bending and stretching motions from a line of points. Scandinavian Journal of Psychology, 18, 209-215.

Jansson, G. & Borjesson, E. (1969) Perceived direction of rotary motion. Perception & Psychophysics, 6, 19-26.

Jansson, G. & Johansson, G. (1973) Visual perception of bending motion. Perception, 2, 321-326.

Jansson, G. & Runeson, S. (1977) Perceived bending motion from a quadrangle changing form. Perception, 6, 595-600.

Koenderink, J. J. (1986) Optic flow. Vision Research, 26, 161-179.

Koenderink, J.J., & van Doorn, A.J. (1975). Invariant properties of the motion parallax field due to the motion of rigid bodies relative to the observer. Optica Acta, 22, 773-791.

Koenderink, J.J., & van Doorn, A.J. (1977). How an ambulant observer can construct a model of the environment from the geometrical structure of the visual flow. In G. Hauske & F. Butenandt (Eds.), Kybernetik (pp. 224-247). Munich: Oldenberg.

Koenderink, J.J., & van Doorn, A.J. (1986). Depth and shape from differential perspective in the presence of bending deformations. Journal of the Optical Society of America A, 3, 242-249.

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. (1990). Perceiving metric structure of environmental objects from motion, self-motion and stereopsis. In R. Warren and A.H. Wertheim (Eds.), The perception and control of self-motion (pp. 541-576). Hillsdale, NJ: Lawrence Erlbaum.

Lappin, J. S. & Ahlstrom, U. B. (1994) On the scaling of visual space from motion: In response to Pizlo and Salach-Golyska. Perception & Psychophysics, 55, in press.

Lappin, J.S., Doner, J.F., & Kottas, B.L. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. Science, 209, 717-719.

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. (1993). Metric structure of stereoscopic form from congruence under motion. Perception & Psychophysics, 51, 86-102.

Lee, D. N. (1974) Visual information during locomotion. In R. B. MacLeod & H. Pick (Eds.), Perception: Essays in honor of James Gibson. Ithaca, N. Y.: Cornell University Press.

Liter, J. C., Braunstein, M. L., & Hoffman, D. D. (1994) Inferring structure from motion in two-view and multi-view displays. Perception, in press.

Longuet-Higgins, H.C. (1981). A computer algorithm for reconstructing a scene from two projections. Nature, 293, 133-135.

Longuet-Higgins, H.C., & Prazdny, K. (1984). The interpretation of a moving retinal image. Proceedings of the Royal Society of London B, 208, 385-397.

Loomis, J.M., & Eby, D.W. (1988). Perceiving structure from motion: Failure of shape constancy. In Proceedings from the second international conference on computer vision (pp. 383-391). Washington, D.C.: IEEE.

Loomis, J.M., & Eby, D.W. (1989). Relative motion parallax and the perception of structure from motion. In Proceedings from the workshop on visual motion (pp. 204-211). Washington, D.C.: IEEE.

Mach, E. (1962) The analysis of Sensations (English translation 1962; originally published in 1886). Dover, New York.

Metzger, W. (1934) Tiefinericheinungen in optichen bewengungsfelden. Psychologische Forschung, 20, 195-260.

Mingolla, E., & Todd, J. T. (1981). The rotating square illusion. Perception and Psychophysics, 29, 487-492.

Musatti, C. L. (1924). Sui fenomeni stereocinetici. Archivio Italiano de Psicologia, 3, 105-120

Nagel, H.-H. (1981). On the derivation of 3D rigid point configurations from image sequences. Proceedings of the IEEE Conference on Pattern Recognition and Image Processing (pp. 103-108). New York: IEEE Computer Society Press.

Nakayama, K., & Loomis, J. M. ((1974) Optical velocity patterns, velocity sensitive neurons , and space perception: A hypothesis. Perception, 3, 53-80.

Norman, J.F., & Todd, J.T. (1993) The Perceptual analysis of structure from motion for rotating objects undergoing affine stretching transformations. Perception & Psychophysics, 3, 279-291.

Norman, J. F. & Lappin, J. S. (1992) The detection of surfaces defined by optical motion. Perception & Psychophysics, 51, 386-396.

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. Cambridge, MA: MIT press. pp. 93-110.

Norman, J. F. & Todd, J. T. (1993) The Perception of rigid motion in depth from the optical deformations of shadows and occlusion boundaries. Journal of Experimental Psychology: Human Perception and Performance, in press.

Ono, M., Rivest, J. & Ono, H. (1986) depth perception as a function of motion parallax and absolute distance information. Journal of Experimental Psychology: Human Perception and Performance, 12, 331-337.

Ono, H., & Steinbach, M. J. (1990) Monocular stereopsis with and without head movement. Perception & Psychophysics, 48, 179-187.

Petersik, J.T. (1979). Three-dimensional constancy: Coherence of a simulated rotating sphere in noise. Perception & Psychophysics, 25, 328-335.

Petersik, J.T. (1980). The effects of spatial and temporal factors on the perception of stroboscopic rotation simulations. Perception, 9, 271-283.

Petersik, J.T. (1987). Recovery structure form motion: Implications for a performance theory based on the structure-from-motion theorem. Perception & Psychophysics, 42, 355-364.

Pizlo, Z. & Salach-Golyska, M. (1994) Is vision metric: Comment on Lappin and Love (1992). Perception & Psychophysics, 55, in press

Pollick, F. E., Giblin, P. J., Rycroft, J., and Wilson, L. L. (1992). Human recovery of shape from profiles. Behaviormetrika, 19, 65-79.

Proffitt, D. R., Rock, I., Hecht, H. & Shubert, J. (1992). The stereokinetic effect and its relation to the kinetic depth effect. Journal of Experimental Psychology: Human Perception and Performance, 18, 3-21.

Reichel, F.D., & Todd, J.T. (1990). Perceived depth inversion of smoothly curved surfaces due to image orientation. Journal of Experimental Psychology: Human Perception and Performance, 16, 953-664.

Rogers, B. J. & Collett, T. S. (1989) The appearance of surfaces specified by motion parallax and binocular disparity. Quarterly Journal of Experimental Psychology, 41A, 697-717.

Rogers, B., & Graham, M. (1979) Motion Parallax as an independent cue for depth perception. Perception, 8, 125-134.

Rogers, B., & Graham, M. (1982) Similarities between motion parallax and stereopsis in human depth perception. Vision Research, 22, 216-270.

Sperling, G., Landy, M.S., Dosher, B.A., & Perkins, M.E. (1989). Kinetic depth effect and identification of shape. Journal of Experimental Psychology: Human Perception and Performance, 15, 826-840.

Tittle, J. S., Perotti, V. J., Todd, J. T., & Norman, J. S. (1993). The perception of relative surface orientation from binocular disparity and motion. Investigative Ophthalmology & Visual Science, 34, 1132.

Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1993) A heirarchical analysis of alternative representations in the perception of 3D structure from motion and stereopsis. Journal of Experimental Psychology: Human Perception and Performance, submitted.

Todd, J. T. (1981). Visual information about moving objects. Journal of Experimental Psychology: Human Perception and Performance, 7, 795-810.

Todd, J. T. (1982). Visual information about rigid and nonrigid motion: A geometric analysis. Journal of Experimental Psychology: Human Perception and Performance, 8, 238-251.

Todd, J. T. (1984). The perception of three-dimensional structure from rigid and nonrigid motion. Perception and Psychophysics, 36, 97-103.

Todd, J. T. (1985). The perception of structure from motion: Is projective correspondence of moving elements a necessary condition? Journal of Experimental Psychology: Human Perception and Performance, 11, 689-710.

Todd, J.T., Akerstrom, R.A., Reichel, F.D., & Hayes, W. (1988). Apparent rotation in 3-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.

Todd, J. T., Norman, J. F., Perotti, V. J., & Tittle, J. S. (1993) The discrimination of 3D length from motion and stereopsis. Investigative Ophthalmology & Visual Science, 34, 1131.

Todd, J.T., & Reichel, F.D. (1989). Ordinal structure in the visual perception and cognition of smoothly curved surfaces. Psychological Review, 96, 643-657.

Treue, S., Husain, M., & Andersen, R.A. (1991). Human perception of structure from motion. Vision Research, 31, 59-76.

Ullman, S. (1977). The interpretation of visual motion. Ph.D. Thesis, Massachusetts Institute of Technology.

Ullman, S. (1979). The interpretation of visual motion. Cambridge, MA: MIT Press.

Ullman, S. (1983). Recent computational studies in the interpretation of structure from motion. In J. Beck & A. Rosenfeld (Eds.) Human and machine vision (pp. 459-480). New York: Academic Press.

Ullman, S. (1984). Maximizing rigidity: The incremental recovery of 3-D structure from rigid and nonrigid motion. Perception, 13, 255-274.

Wagner, M. (1985) The metric of visual space. Perception & Psychophysics, 38, 483-495.

Wallach, H., & O'Connell, D.N. (1953). The kinetic depth effect. Journal of Experimental Psychology, 45, 205-217.

Wallach, H., Weisz, A., & Adams, P. A. (1956) Circles and derived figures in rotation. American Journal of Psychology, 69, 48-59.

Waxman, A., & Ullman, S. (1985). Surface structure and three-dimensional motion parameters from image flow kinematics. International Journal of Robotics Research, 4, 79-94.

Webb, J. A. & Aggarwal, J. K. (1981) visually interpreting the motions of objects in space. Computer, 8, 40-46.

White, B & Mueser, G. (1960) Accuracy of reconstructing the arrangement of elements generating kinetic depth displays. Journal of Experimental Psychology, 60, 1-11.

Zanforlin, M. (1988). The height of the stereokinetic cone: A quantitative determination of a 3-D effect from 2-D moving patterns without a "rigidity assumption." Psychological Research, 50, 162-172.