CV REFERENCE NO. 16
IEEE Workshop on Motion: Representation and Analysis,
May 1986

INEXACT VISION

William B. Thompson

University of Minnesota
Minneapolis, MN 55455

Joseph K. Kearney

University of Iowa
Iowa City, Iowa 52242

ABSTRACT

Approximate descriptions of object shape, scene layout, and object and observer motion are useful for many tasks. The conventional approach to low level vision has been to recover precise, detailed description of scene properties. In this paper, we argue that in some circumstances image analysis is simplified when approximate, qualitative information is sought. This is especially true for motion analysis. Methods for precisely determining surface shape and object motion have led to complex, numerically unstable estimation procedures. These procedures are often based on unrealistic assumptions about the scene and unrealistic expectations about the capabilities of motion estimation techniques. Better results can often be obtained by approaches that bypass precise analysis and directly determine general properties of structure and motion. We discuss the relationship between quantitative and qualitative representations and outline a range of measurement scales for describing scene properties. Finally, we give three examples of methods for inexact vision.

1. Introduction.

Research in visual motion understanding has been devoted almost exclusively to the determination of the precise geometric properties of scenes. Numerous studies report methods to obtain depth, surface orientation, and relative velocity from the motion of object points on the image. This approach is computationally intensive, prone to error, and in many cases unnecessary. These methods often make strong assumptions about the environment which rarely hold in realistic situations. Not surprisingly, demonstrations of this work with real images are almost nonexistent. Focusing the analysis on more qualitative effects and utilizing less precise representations may lead to solutions which produce useful information in an effective manner without placing restrictive constraints on the scene.

We propose that motion analysis should follow the *principle of least effort*. Informally, this principle says that simple, error tolerant computations should precede more complicated, error sensitive processes. Very difficult tasks should be done only where necessary and should build on the results of the simple solutions. We do not dismiss the importance and usefulness of detailed, quantitative analysis. Rather, we believe that general solutions will require a balance between quantitative and more qualitative information.

Quantitative measures of object shape, scene structure, and trajectories are represented as numerically valued, geometric and temporal attributes. Such representations are attractive because they can easily capture most of what is potentially relevant about scene structure. Qualitative representations use a fixed set of categories to describe geometric properties. In a sense, qualitative descriptions constitute a mapping from the continuous space of possible geometric structures to a discrete space of possible descriptions. Such a mapping may appear to limit the information content of the description due to reduced precision. In fact, the precision associated with quantitative measures is often an illusion due to the limits on accuracy with which these measures can be obtained. Quantitative descriptions can be expressed using a variety of *measurement scales*, many of which capture only partial information about a given property. The existence of these different scales provides a range of possible representational precisions from purely qualitative to exact and complete quantitative descriptions.

The use of qualitative representations at high levels of abstraction has long be accepted as natural means to describe the important characteristics of a scene. In this paper, we argue that less precise representations can be beneficially employed even in the low level analysis of images.

2. Current approaches.

The study of visual motion has been typically broken into 3 stages of processing. In the first stage the two-dimensional motion of points or features on the image is estimated. The second stage of processing, which we are examining in this paper, interprets the two-dimensional motion to recover three-dimensional structure or motion. Later stages may use the three-dimensional information to locate and describe objects, to determine the position of the observer in a known environment, or to understand events such as object or observer trajectories or time to collisions. Figure 1 outlines the approaches currently taken for the analysis of visual motion. Three approaches for estimating motion on the image are given in the first column. The second column presents the most common goals for systems which recover world information from image motion. The last column shows tasks for which this information might be relevant. Solid lines indicate an extensive body of work linking elements from two columns. Dashed lines indicate limited results.

This work was supported by National Science Foundation Grant DCR-8500899 and Air Force Office of Scientific Research contract AFOSR-85-0382.

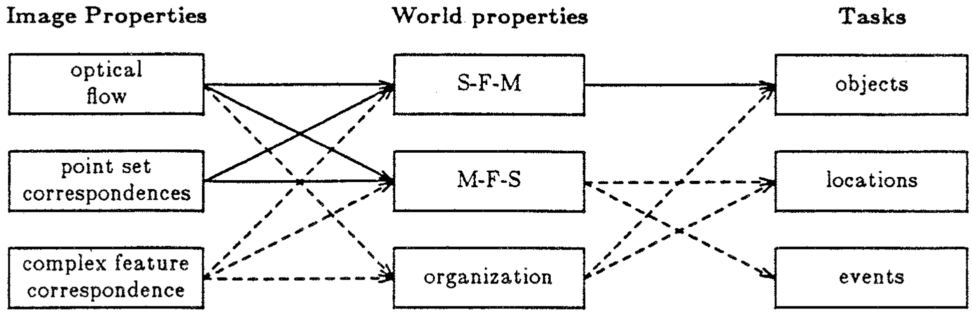

Figure 1.

Image motion may be estimated for all points in the image, for a sparse set of feature points, or for complex structural features. *Optical flow* is the image plane velocity of the projection of visible surface points. Optical flow is in principle defined everywhere over the image plane. *Point set correspondences* are sparse sets of corresponding points in multiple images of the same object. Points in different frames are in correspondence if they arise from the same surface point on the object. *Structural features* are based on image structures more complex than points. (For example, [1] uses the time variation of image contours.) The correspondence between structural features across frames must be established. Visual motion is then characterized by changes in these features.

We are not concerned in this paper with how image cues are computed. We must, however, be aware of general properties of these computations. In particular, perfect estimates of image cues are not possible. We will use the term *noise* to refer to any deviation of an estimate cue from the true value defined by the three-dimensional structure and temporal characteristics of the scene. With few exceptions, the literature on motion perception ignores issues of accuracy and reliability. A small number of papers do point out the difficulties associated with noise (e.g. [2,3]). Even fewer attempt to determine actual measures of accuracy (e.g. [4]).

Several recent papers describing methods for determining surface structure and object trajectories based on optical flow have included empirical analyses of sensitivity to errors in flow estimates [5,6]. Sensitivity is examined by observing the answer obtained in the presence of random, multiplicative distortions of idealized, synthetically derived input. Noise values in the range of 5% to 20% have been considered. Our own experience suggests that the noise associated with estimations of optical flow has a complex dependency on the magnitude of the flow. Noise in correspondence-based methods is principally additive, depending on positional inaccuracies of structures being matched and on the likelihood of a mismatch. Positional accuracy in each frame is not a function of displacement between frames. While the frequency of mismatch may depend on the magnitude of the true flow value, the magnitude of the error induced typically does not. For gradient-based methods, error analysis is more complex. In general, absolute error can be large for both very small and very large inter-frame displacements. While error analyses have not been done for methods estimating optical flow at contours [7] or for determining parameters of motion directly from image changes [8], we expect similar effects.

3. Problems with the current approach.

Most of the structure-from-motion and motion-from-structure techniques are based on systems of non-linear equations. These systems are numerically unstable. Small errors in the input (e.g. image cues) result in large errors in the estimates of geometric properties. Furthermore, many of the proposed techniques require iterative solutions and appear to depend on good starting guesses. Such methods may be good for refining a geometric description obtained by other means, but are not likely to be useful for an initial analysis.

Techniques based on optical flow usually depend on spatial (and less often temporal) derivatives of flow. Similarly, methods based on point correspondences ultimately depend on the differences in two-dimensional disparity between sets of points. Differentiating values characterized by additive noise greatly increases the relative noise in the resulting values. When combined with numerically unstable computations, the end result is not likely to be meaningful.

Problems associated with differential measures are reduced as the magnitude of the difference increases. As a result, it is desirable to measure such values over large areas of the image. The larger the neighborhood used, however, the more likely the presence of a discontinuity in the geometric property being estimated. For example, [3] argue that parameters of observer motion are best determined by an optimization scheme which depends on flow values over the full field of view. The problem with this approach is that as the field of view is expanded, it is more and more likely that moving objects will be visible. Basing a determination of sensor motion on a combination of visual cues generated by the background and by independently moving objects will not give a correct estimation of either motion.

Various smoothing (regularization) techniques are currently popular. One attraction of these techniques is that they allow an analysis based on both local and global image properties. Fitting smooth surfaces to intrinsically discontinuous values seems to unnecessarily complicate the problem, particularly as the immediately following step is often to attempt to recover the discontinuities from the now smoothed and therefore continuous data. The problem is even worse when the discontinuities in geometric properties also lead to incorrect estimation of image image properties. For example, the Horn and Schunck method for estimating optical flow [9] fails whenever there are discontinuities present in optical flow [10]. The global enforcement of smoothness causes this failure to be propagated over the whole image.

4. Desirable properties.

The processes that interpret image motion estimates to obtain a representation of some world property should have the following attributes:

High signal-to-noise ratio.

The process should be applied only where there is sufficient information to support the computation. Differential effects should only be measured across regions that contain a substantial variation in the attribute to be measured. Local computations should only be made at boundaries where the property being measured changes significantly.

Restricted interpretation.

The process should produce only the information required by later stages of processing. Geometric properties often need be determined with only limited precision. Frequently, 2 or 3 values are sufficient: in front/behind, curved/flat, convex/concave, approaching/receding. Reducing precision increases reliability.

Minimal assumptions.

The process should not depend on restrictive assumptions about the nature of the scene (e.g. planar surfaces) or the accuracy of the input (e.g. exact optical flow values). Such assumptions are almost certainly going to be wrong. Unrealistic assumptions are **only** justifiable when it can be shown that useful answers can obtained in realistic situations despite violations in the assumptions. In general, motion analysis methods should assume general motion and general surfaces.

5. Qualitative vs. quantitative representations.

The choice of a representation scheme has a great impact on the generality, reliability, and efficiency of algorithms that operate on images. In other areas of Artificial Intelligence, qualitative descriptions of physical properties and events have generated much interest. [11]. A well defined qualitative representation may capture much or all of the significant aspects of the situation. In this section we examine inexact representations of information obtained from images.

Information about many different attributes of scenes can be obtained from dynamic image sequences. A *representation* assigns values of an attribute for a set of entities in the scene. Entities may range from visible surface points, through structures such as surfaces and volumes, to whole objects. Values may be assigned for all or only a small number of the scene entities. Information about motion and geometry can be represented with varying degrees of accuracy. If the attribute is measured with high precision we say that the representation is quantitative. In passing, we note that the resolution of the representation is highly dependent on the accuracy of the measurements from which the values are obtained. In vision, we are always limited by the spatial resolution and dynamic range of the sensor. Typically, sensors can distinguish between a few hundred brightness levels at a few tens of thousands of image points. Hence, even a quantitative brightness scale may contain only 100 or so discrete values.

If an attribute can take only a small number of values we say that the representation is qualitative. Qualitative representations define a set of equivalence classes on the quantitative scale. We will assume qualitative values correspond to disjoint, continuous intervals on the quantitative scale. Usually, the intervals of adjacent qualitative values abut each other on the quantitative scale.

Typically, boundaries on a qualitative scale are chosen such that they correspond to significant boundaries on the corresponding quantitative scale. For example, curvature could be represented with three values (+1,0,-1) indicating convex, flat, or concave in quantitative representation. Because the precision required of qualitative descriptions is less than for a corresponding quantitative description, qualitative information can often be more reliably and more efficiently determined than the underlying quantitative information. And for many tasks qualitative information is all that's necessary.

It is sometimes the case that quantitative properties are needed to perform some task. When possible, precise determination of scene properties should follow the estimation of more easily obtained qualitative properties. For example, the location of surface boundaries significantly restricts likely surface shape. The image of the boundary constrains the possible locations of the corresponding extremal contour. Even though this specification is not unique, presumptions about compactness of objects and smoothness of surfaces can be used to estimate likely shapes. In some situations, motion information can actually be used to find the surface orientation at the boundary. Because the computational complexity associated with estimating precise quantitative properties is usually large, it is also desirable to make such determinations only when necessary. By combining a focused analysis with the principle of least effort, it may be possible to obtain reliable results.

6. Measurement scales.

An attribute can be represented in many different measurement scales. For example, range to a surface point may be described in terms of absolute distance (e.g. 23 feet) or in terms of an ordinal relationship (e.g. point **a** is closer than point **b**). A clear understanding is needed of the nature of the representational scale corresponding to particular analysis algorithms. In addition, an important open question in low-level vision is the manner in which measurements of the same or similar properties in different scales can be combined.

A representation specifies the value of a scene property to some degree of precision. The representation can be viewed as constraining the value of the property, but except in the case of an exact representation, the property is not limited to a single value. The degree to which the properties of objects are constrained by the representation is determined by the type of the underlying scale. The type of a scale is determined by the allowable transforms on the measurements [12]. The measurements in the transformed representation must have the same properties as the original representation. Although a large number of scale types are possible, only five occur commonly.

Absolute Scales

No transformations are permissible on a absolute scale. Surface orientation and depth are often measured on an absolute scale. If an observer is translating with known velocity through a stationary world, then the range to visible surface points can be determined absolutely.

Ratio Scales

Ratio scales are unique up to a multiplicative scale constant. If an observer is translating with unknown velocity through a stationary world, then the range to visible surface points can be determined to within a scale constant dependent on observer speed. Such a range determination is an example of a ratio scale. If ratio scale information about scene properties is sufficient, less information is required about the circumstances under which imagery is obtained. One potential problem arises when multiple ratio scales are used. The origin of a ratio scale is unaffected by scaling and hence is known exactly. In vision systems, ratio scales for a particular property determined by different methods often have different origins, making comparisons difficult.

Interval Scales

Interval scales preserve information about the difference between the physical entities. Any positive linear transformation can be applied to an interval scale. For example, consider a moving sensor observing a collection of objects close to one another, but distant from the sensor. If it is known that the objects are moving in the same directions, then image velocities provide an interval scale of three dimensional velocity. In particular, the difference in magnitude of the image velocities of two points is proportional to the difference the magnitudes of their three-dimensional velocities. If the velocities of three points on the image have magnitudes $v(x)$, $v(y)$, and $v(z)$ then $v(x)-v(y)=v(y)-v(z)$ implies that the respective 3-dimensional differences are equal.

Ordinal Scales

Ordinal scales contain information only about the ordering of objects. Any order preserving transform can be applied to an ordinal scale. Ordinal depth scales can be derived from interposition information. For example, consider two moving surfaces moving in different directions that share a common edge in the image. The accretion or deletion of surface points at the boundary of the surfaces indicates which surface is closest to the viewer. The appearing or disappearing points must lie on the more distant, occluded surface. Note that this depth information is locally defined. The ordinal relations are not transitive across surfaces. For example, consider the how four sides of a box can be folded over and interlocked to close the top of the box (figure 2). The upper left surface is in front of the lower left surface which is in front of the lower right surface which is in front of the upper right surface which is in front of the upper left surface....

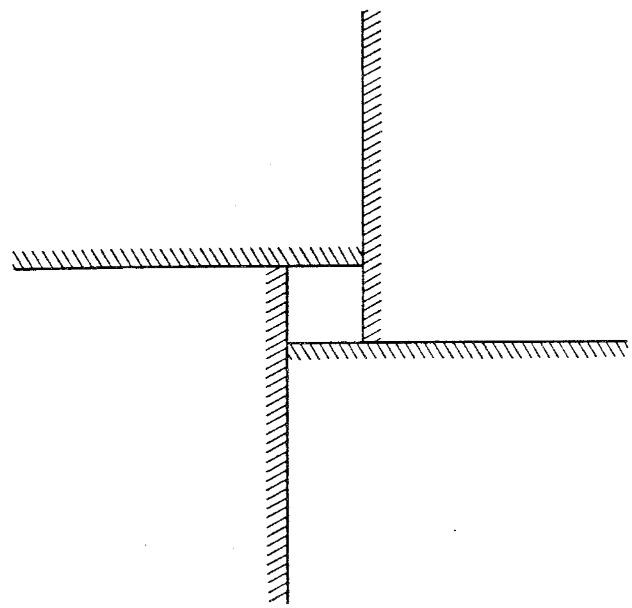

Figure 2.

Nominal Scales

In a nominal scale measurements are simply symbolic labels. Any one-to-one transform can be applied to a nominal scale. The segmentation of an image into regions that correspond to object surfaces results in a nominal scale. Each point is labeled according to the surface to which it belongs. The labels, however, bear no relationship to one another. Any set of unique labels is adequate. Motion information can often provides a strong cue for segmentation. Nearby points on the image with similar motion will usually lie on the same surface. Even if we nothing about the three dimensional shapes and motions of objects, we can often use image motion as a cue for segmentation.

It is important to recognize the scale underlying a measurement system as it defines the meaningfulness of operations on the resultant values. For example, the average of a group of ordinal measurements is meaningless and only the sign of the difference between two values on an ordinal scale has meaning.

Work in motion analysis has been almost exclusively focused at quantitative representations of three-dimensional motion and geometry. The determination of absolution and ratio scales of motion and shape has received the most attention. While this work is important, there is much qualitative information that can also be exploited. Qualitative information can often be obtained more robustly and more efficiently than quantitative information and requires fewer assumptions about the environment. It is usually the case that the greater the scale constrains the properties of objects, the greater the world must be constrained to obtain the measurements. It is likely that successful systems will employ representations of many types.

7. Examples of inexact vision.

In this section, we briefly outline three examples of techniques for analyzing visual motion which satisfy at least some of the desirable properties listed above. The first example deals with the discovery of surface boundaries, and the use of this information to determine aspects of relative surface position and surface shape. Thus, a purely qualitative analysis is followed by the extraction of geometric information. The second example illustrates a method for the detection of objects likely to collide with the observer. One simple solution to this problem depends on assumptions about object and observer motion that do not hold in general. However, these assumptions do hold when a collision is imminent. The third example outlines a new method for determining the direction of observer motion based solely on visual data. While the analysis is quantitative, it is structured in such a way that irrelevant information required by other approaches drops out of the formulation, and thus only relevant information is actually computed.

7.1. Boundary determination.

Most of the work examining the recovery of three dimensional structure for motion information has been directed toward the description of surface shape. The less ambitious goal of determining surface boundaries is easier to achieve. The importance of finding the location in an image of the extremal contours of surfaces is well recognized. An extensive literature in edge detection is motivated in large part by the presumed correspondence between discontinuities in some image property and discontinuities in a related scene property. Edge detection is in fact the most common form of qualitative image analysis -- an edge is either there or it's not. Motion-based segmentation can not only find boundaries that are difficult to locate in a single view, but it can also provide much more information about the structure of the scene. We have previously reported on two different techniques for finding dynamic occlusion boundaries [13,14]. In both cases, the methods not only signal the presence of a boundary, but provide information about the location of the occluding and occluded surfaces at the boundary. Such a determination is important to any process which analyzes the shape of the boundary. While the shape of the edge provides significant information on the structure of the occluding surface, it says little or nothing about the structure of the surface being occluded. This approach can also determine the ordinal depth relationships at a boundary.

Reasonable boundary detection and analysis results are obtained even when the optical flow fields on which the detection is based contain substantial amounts of noise. We feel that this success comes from two features of the techniques. Only simple classification operations are required. First, a yes/no determination is made as to the existence of the boundary. Then, the two sides of the boundary are classified as occluding or occluded. In addition, the qualitative nature of the analysis makes it possible to integrate information over relatively large neighborhoods of the flow field without undue distortions in the final result. The estimation of quantitative point properties of scene structure is complicated by variations in the property over neighborhoods used to perform the calculation. Analysis using limited precision qualitative descriptions is often quite tolerant of these neighborhood variations.

More generally, we argue that visual motion analysis should depend principally on *significant* variations in properties such as optical flow. Human perception of structure-from-motion seems to depend more on spatial *differences* in optical flow than on the flow values themselves. Simultaneous and successive contrast effects and Craik-O'Brien-Cornsweet illusions appear in depth-perception due to motion parallax [15,16]. Human vision reconstructs the shape of surface from motion based on spatial differences in optical flow, not the actual magnitude of flow. (This is in clear contrast to surface interpolation methods such as [17].) By determining surface structure using a computational model such as [18], it might be possible to combine the analysis of boundaries and important surface features, while reducing the sensitivity to noise and to systematic distortions due to effects such as observer motion.

7.2. Collision detection.

A method of detecting collisions is well known to navigators: "If the bearing becomes constant and the distance is decreasing, the two vessels are on collision courses.... [19]" This allows a simple qualitative test for collision based on optical flow. The approach of the object is signaled by a positive divergence in the optical flow field while the "constant bearing" is marked by a zero average flow in the same region of the image [20]. The test is simple and effective because it does not depend in any way on actually determining the relative motion of the object and observer. As a result, it can be implemented in an efficient and accurate manner.

One significant complication arises, however. The result quoted from [19] is actually true only when there is no relative acceleration between the object and observer. Neither object nor observer can be changing speed or direction, and the observer cannot be rotating. Such assumptions are of course not generally valid. It is the case, though, that the assumptions dealing with changing speed and/or direction are less important as collision is imminent and detection becomes more important.

If accelerations are bounded, changes in trajectory have less effect on the likelihood of a collision as an object and observer approach one another. For an object and observer on a collision course, the range of paths the object and observer can take to avoid a collision becomes increasingly limited as the object nears the observer. A nearby object approaching the observer at high velocity is likely to collide with the observer.

7.3. Four-eyed vision.

Considerable attention has been paid in both the perceptual psychology and computer vision communities to the problem of estimating sensor trajectories based on visual motion. Vision can in principle be used to estimate five parameters of instantaneous sensor motion -- three rotational velocities and the direction of translation. (Translational speed cannot be estimated unless there is some independent range cue in the imagery.) In practice the problem is ill-conditioned, making it difficult to determine five distinct parameters.

Often, it is sufficient to know only the direction of translation. The most straightforward way to do this is to *first* solve for the rotational parameters. Once rotation is known, it is easy to convert the optical flow field into a field that would result if only translation were occurring. The direction of translation is indicated by the focus of expansion in this translational flow field. A better approach is to estimate translation directly, without computing additional information not necessarily needed. (Rotation, if required, can be estimated using inertial mechanisms or by examining optical flow in the direction of translation.) [21] shows how the effects of rotation on optical flow can be minimized by applying a local spatial difference operator to the flow field. The use of a very wide field of view can lead to the same effect without the problems associated with local spatial differencing.

Four-eyed vision is a simplification of the more general case of vision using a 180° field of view and a coordinate system based on spherical projection. Four distinct regions at the periphery of the field of view are used in the analysis. Each region can be thought of as the image associated with a separate camera pointing in the appropriate direction. Each camera points directly away from a common center, with pairs of cameras pointing in exactly opposite directions. The imagery from each camera is used to estimate optical flow at the center of each image. The imagery from a single camera is thus used only to determine two parameters specifying the direction and speed of flow. The rotational component of flow for two images making up a pair is the same. The translational components will point in opposite directions, with a magnitude dependent of the range to the surface generating the flow for each camera. Subtracting the two flow from opposing images values cancels the rotational component, leaving only an average translational component. The translational components for the two pair can then be used to solve for the focus of expansion. In situations where four-eyed vision is practical, it provides a simple solution to an otherwise difficult problem. Even when four-eyed vision is not used, it is important to note that the accurate determination of sensor translation in all but noise-free situations requires a wide angle of view [22].

8. Needed research.

The interpretation of image motion can not be studied in isolation. Both *competence-based* and *task-based* analysis is crucial to the development of effective motion analysis techniques. Competence-based analysis involves the determination of what is possible to compute and, equally important, the accuracy of such computations. Task-based analysis focuses on what information actually needs to be computed at each stage of processing. To determine the potential three-dimensional information that can be recovered from image motion, we must understand the characteristics and limitations of the motion estimation techniques. The characteristics of the motion estimates determine in large part the success of later stages of processing. Methods of motion analysis must not rely of estimates that are impossible to obtain. We also need a better understanding of the sensitivity of algorithms utilizing these values to errors in the estimates.

To determine what three-dimensional information is valuable we must study the tasks that will use the recovered geometric information. What information is needed to perform specific tasks such as position determination and collision avoidance? The ability to recover qualitative information and the usefulness of qualitative information must be studied. The implications of determining scene properties in different measurement scales must be better understood, and methods developed for integrating information in different scales.

It is not surprising that so few motion analysis techniques have been demonstrated on real imagery. The heavy focus on the determination of precise geometry and trajectory information leads to methods intolerant of the variability and imprecision in actual images. By reducing the precision required of the analysis, substantial improvements in robustness may be obtained. We are not suggesting that precise estimation methods should be abandoned for more qualitative techniques. We do advocate, however, a more heterogeneous approach in which information is determined to the minimal precision necessary.

BIBLIOGRAPHY

[1] A.M. Waxman and K. Wohn, "Contour evolution, neighborhood deformation and local image flow: Planar surfaces in motion," technical report CS-TR-1394, Computer Vision Laboratory, University of Maryland, April 1984.

[2] J.W. Roach and J.K. Aggarwal, "Determining the movement of objects from a sequence of images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. PAMI-2, pp. 554-562, November 1980.

[3] A.R. Bruss and B.K.P. Horn, "Passive Navigation," *Computer Vision, Graphics, and Image Processing,* v. 21, n. 1, pp. 3-20, 1983.

[4] J.K. Kearney and W.B. Thompson, "Gradient based estimation of disparity," *Proc. IEEE Conf. on Pattern Recognition and Image Processing,* June, 1982.

[5] A.M. Waxman and S. Ullman, "Surface structure and three-dimensional motion from image flow kinematics," *International Journal of Robotics Research,* vol. 4, pp. 72-94, Fall 1985.

[6] K. Wohn and A.M. Waxman, "Contour evolution, neighborhood deformation and local image flow: Curved surfaces in motion," technical report CS-TR-1531, Computer Vision Laboratory, University of Maryland, July 1985.

[7] E.C. Hildreth, *The Measurement of Visual Motion,* Cambridge: MIT Press, 1984.

[8] "Direct passive navigation," to appear in *IEEE Trans. Pattern Analysis and Machine Intelligence.*

[9] B.K.P. Horn and B. Schunck, "Determining optical flow," *Artificial Intelligence,* vol. 17, pp. 185-203, 1981.

[10] B. G. Schunck and B.K.P. Horn, "Constraints on optical flow," *Proc. IEEE Conference on Pattern Recognition and Image Processing,* pp. 205-210, Aug. 1981.

[11] *Qualitative Reasoning About Physical Systems,* D.B. Bobrow, ed., MIT Press, 1985.

[12] C. H. Coombs, R. M. Dawes, and A. Tversky, *Mathematical Psychology,* Prentice-Hall, Inc., Englewood Cliffs, NJ, 1970.

[13] K.M Mutch and W.B. Thompson, "Analysis of accretion and deletion at boundaries in dynamic scenes," *IEEE Trans. Pattern Analysis and Machine Intelligence,* March 1985.

[14] W.B. Thompson, K.M. Mutch, and V.A. Berzins, "Dynamic occlusion analysis in optical flow fields," *IEEE Trans. Pattern Analysis and Machine Intelligence,* July 1985.

[15] M. Graham and B. Rogers, "Simultaneous and successive contrast effects in the perception of depth from motion-parallax and stereoscopic information," *Perception,* vol. 11, pp. 247-262, 1982.

[16] B. Rogers and M. Graham, "Horizontal - vertical anisotropies in the perception of 3-D depth surfaces," presented at meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Fla., 1982.

[17] W.E.L. Grimson, *From Images to Surfaces,* Cambridge: The MIT Press, 1981.

[18] B.K.P. Horn, "Determining lightness from an image," *Computer Graphics and Image Processing,* vol. 3, pp. 277-299, 1974.

[19] N. Bowditch, *American Practical Navigator,* Washington DC: U.S. Navy Hydrographic Office, 1958.

[20] W.B. Thompson, K.M. Mutch, and V.A. Berzins, "Analyzing Object Motion Based on Optical Flow," *Proc. Seventh International Conference on Pattern Recognition,* July, 1984.

[21] J.H. Rieger and D.T. Lawton, "Sensor motion and relative depth from difference fields of optic flows," *Proc. 8th Int. Joint Conf. on Artificial Intelligence,* pp. 1027-1031, August 1983.

[22] G Adiv, "Inherent ambiguities in recovering 3-D motion and structure from à noisy flow field.8(rq *Proc. IEEE Conf. on Computer Vision and Pattern Recognition,* June 1985.