EDGE DETECTION IN OPTICAL FLOW FIELDS

William B. Thompson
Kathleen M. Mutch
Valdis Berzins

Computer Science Department
136 Lind Hall
University of Minnesota
Minneapolis, MN 55455

ABSTRACT

Optical flow is potentially valuable as a source of spatial information. Current techniques provide flow fields which are noisy and sparse, making the recovery of spatial properties difficult at best. This paper describes a technique for locating discontinuities in optical flow, which typically correspond to object boundaries. A simple blurring interpolator is used to smooth out noise and produce denser fields. Discontinuities are found by locating the vector field equivalent of zero crossings in the Laplacian of a scalar field. The technique is illustrated by applying it to realistic vector fields which are both noisy and sparse.

1. INTRODUCTION

The concept of optical flow - vector fields characterizing the two dimensional changes in time varying imagery - has received increased attention recently as a source of spatial information. With few exceptions (e.g. [1]), research has focussed on two non-overlapping tasks. Several techniques have been developed for obtaining vector fields from image sequences [2,3]. At the same time, methods have been derived for computing spatial properties from optical flow fields [4,5,6,7,8]. These methods presume that the flow field is known with high accuracy over a dense sampling of points in the field of view. In a realistic environment, current techniques provide vector fields which are far from ideal, being both noisy and sparse. A need exists to develop techniques for estimating spatial information which will function on realistic vector fields. This paper examines one such technique for locating discontinuities in optical flow. Such discontinuities typically correspond to object boundaries.

Errors may arise from many sources in the determination of vector fields [9]. Sensor noise may be magnified and propagated in certain instances. Some techniques make assumptions, such as linear intensity gradients or locally constant vector fields, which frequently may be violated. In addition, some matching techniques never attempt to match more than a very sparse set of image points [10]. Spatial interpretation techniques requiring accurate input will probably not function well when provided with sparse and error prone vector fields. As Bruss and Horn noted [11], techniques which rely upon local derivatives of a vector field (e.g. [4]) will only magnify any existing error. If the original vector fields contain significant error, such a technique will not function properly. We need to develop techniques which will be relatively insensitive and adaptive to noisy and sparse vector field problems.

In realistic imagery there is an additional reference frame problem. Both camera and object translations and rotations are possible. The result is ambiguous vector fields, where a given vector pattern could represent many different spatial/motion combinations. Some vector field analysis techniques resolve this problem by limiting the type of allowable motion to, for example, object translation only [12], or camera translation only [13]. Frequently these limitations may be violated in important or interesting image sequences. It is desirable for analysis techniques to be insensitive to the problems arising from unconstrained motion.

One way to deal with the noise problem is to use global, rather than local, analysis. Motion parameters can be determined by some form of global optimization dependent on values over the whole field of view (e.g. [11]). These motion parameters can then be used in a more local analysis to determine spatial properties. This technique is useful when the observer is moving relative to a static environment. When some significant portion of the scene is moving or when multiple moving objects are viewed by a stationary observer, the analysis breaks down because no single set of motion parameters is valid over the whole field of view. In such situations, it is desirable for an operator to be localized so that it is within object boundaries and at the same time perform some averaging over an area to reduce noise problems. Since object sizes are unpredictable, it may be useful to apply an operator of variable size, producing an analysis with several degrees of locality.

The technique for edge detection described in this paper attempts to deal with the problems of

This work was supported by the National Science Foundation under Grant MCS-81-05215, and by a Louise Dosdall Fellowship from the University of Minnesota.

noisy, sparse vector fields and arbitrary motion. A simple blurring interpolator is used to smooth out noise and produce denser fields. In general, arbitrary motion produces smoothly varying vector fields within an object boundary, and a discontinuity in the vector field at object edges. Our technique relies upon locating these discontinuities, and so involves no restrictions on object motion. Two examples are provided which demonstrate the performance of the technique on sparse, noisy vector fields obtained from real imagery.

2. METHOD

A vector field can be described by the two component quantities of direction and magnitude. When the motion is limited to camera translation, vector direction will vary slowly across the entire image. At edges, discontinuities will occur in magnitude only, and for most analytical purposes the vector field can be reduced to a scalar field [13,14]. When arbitrary motion is allowed, the discontinuity may occur in either magnitude or direction, or both. The vector field in this case cannot be simplified to a scalar field.

Our analysis of discontinuity detection for two-dimensional vector fields is similar to that of Marr and Hildreth in the scalar case [15]. Discontinuity in a discrete image field means that the variability on either side of an edge is much less than the variability across the edge. If the edge is approximately linear and variability along the edge is suitably constrained, then the search for discontinuities can be decomposed into a separate analysis of the scalar fields corresponding to the x and y components of the optical flow. (These constraints are generalizations of the "linear variability" assumption used in [15].) A discontinuity in either component corresponds to a discontinuity in the original field. Smoothing is performed to reduce the effects of noise and to serve as a bandpass filter on width of the edges. In the case of sparse vector fields, smoothing serves an additional function of interpolation. A gaussian smoothing kernel is used, since it is optimal with respect to the condition of simultaneous locality in the space and frequency domains.

A discontinuity in either of the smoothed scalar fields will result in a peak in the first derivative of the field, and a zero crossing in the second derivative. The Laplacian is a more convenient operator than the second derivative, however, since it is invariant with respect to coordinate rotations, allowing a one pass search for edges with arbitrary orientations. For scalar fields that vary linearly parallel to straight edges, the Laplacian is equal to the second derivative taken in the direction of greatest variation. The Laplacian operator applied to a smoothed function has the additional advantage of being closely approximated by the difference of two gaussian functions [15].

Actual discontinuities are found by recombining the componentwise Laplacians into a vector field and then searching for the vector field analog of a scalar zero crossing. At an edge, there will be a zero crossing in at least one component of this difference field, and a value of zero in the other component. Both components may have a zero crossing. In either case, adjacent vectors will reverse direction when an edge lies between them.

3. IMPLEMENTATION

A sparse vector field was obtained from two adjacent images in a sequence using a point matching technique [10]. The vector field was separated into its component x and y scalar fields. The fields were blurred by two different gaussian kernels with a ratio of standard deviations on the order of 1:1.6. The results were subtracted to form an estimate of the Laplacian of the smoothed fields [15]. The differenced component fields were next recombined, and the resulting vector field was searched for reversals in adjacent vectors. Two thresholds were enforced in the search for vector reversals. Theta, the angle separating adjacent vectors, was allowed to vary in a range around the ideal of 180 degrees. A lower threshold was placed on the combined lengths of adjacent vectors to ensure that the slope of edges was significant.

Figures 1 and 2 show the results when this technique was applied in two different motion situations. No gray scale information was used to assist the vector field technique in locating edges. In figure 1, the vectors on the elephant and tiger are of approximately equal magnitude, but differ in direction. In figure 2, the vectors on the two animals are of different magnitude but lie in approximately the same direction. The edges overlaid on the vector fields indicate that the method works relatively well in both cases.

4. DISCUSSION

The technique described in this paper illustrates one aspect of a more general concept of local reference frames. The larger of the two smoothing kernels acts to form a local average flow. Subtracting a less smoothed field results in information about local deviations away from this average. As we have shown, a directional reversal in this deviation field corresponds to a discontinuity in optical flow. Other surface shape properties such as direction of orientation and sign of curvature also have well defined signatures in the deviation field. As long as only qualitative information is being estimated, a purely local analysis is often sufficient. In such cases, there is no requirement for global motion or camera models.

Another important spatial property concerning the asymmetry of occlusion edges is directly available from the deviation field and computed edge locations. While edges are often described as a boundary between two image regions, occlusion edges are only the boundary of one or the other of the corresponding surfaces. Determining which side of a discontinuity in disparity corresponds to the occluding surface provides information about both scene structure and relative depth.

The key to resolving this edge ambiguity is to note that over time, the edge will have the same disparity as the surface to which it corresponds. Furthermore, only the motion components perpendicular to the edge need be considered, since only they lead to asymmetric changes in the appearance of the boundary. This observation results in a simple computational test. Once an edge has been found, a highly smoothed field may be used to get an estimate of the expected image position of the edge in a subsequent frame. Disparity estimates in this smoothed field will be affected by regions on either side of the edge and hence will in fact be an average of the two actual values. The real edge will be translating over the image with a speed either faster or slower than the estimate. Thus, at the "expected" edge location, the deviation vectors will not in fact exhibit a directional reversal. However, the projection of the deviation vector at this point onto the normal across the edge points toward the occluded surface. Figure 3 presents one possible motion case to which this analysis is applied.

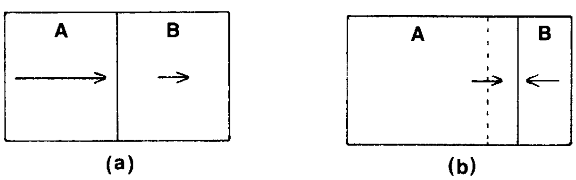

Figure 3. (a) Disparity field obtained from initial image pair, with motion of regions A and B indicated by vectors. (b) Deviation field from a subsequent image pair, when A is the occluding surface. The expected edge is shown as a dotted line, the true edge as a solid line. The deviation vector at the expected edge location points toward the occluded surface.

As with gray scale edges, it is useful to analyze optical flow over a range of resolutions. The deviation fields corresponding to different blurring kernels give information about shape properties at different scales. The use of deviation fields is particularly suited to systems which estimate disparities between image frames using coarse to fine matching [16,14,17]. The results of matching at coarser resolutions can be used to establish local reference frames against which results of finer matches can be compared. Thus, looking across levels of resolution can provide an important source of information about shape.

BIBLIOGRAPHY

[1] H.-H. Nagel, "Representation of moving rigid objects based on visual observations," Computer, v. 14, no. 8, pp. 29-39, 1981.

[2] Special Issue on Computer Analysis of Time Varying Images, Computer, v. 14, no. 8, August, 1981.

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

[4] H.C. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," Proc. R. Soc. Lond., v. B 208, pp. 385-397, 1980.

[5] K. Prazdny, "Egomotion and relative depth map from optical flow," Biol. Cybernetics, v. 36, pp. 87-102, 1980.

[6] S. Ullman, The Interpretation of Visual Motion, Cambridge: MIT Press, 1979.

[7] K. Nakayama and J.M. Loomis, "Optical velocity patterns, velocity sensitive neurons, and space perception," Perception, v. 3, pp. 63-80, 1974.

[8] D.D. Hoffman, "Inferring shape from motion fields," MIT AI Memo No. 592, December, 1980.

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

[10] S.T. Barnard and W.B. Thompson, "Disparity analysis of images," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. PAMI-2, pp. 333-340, July 1980.

[11] A.R. Bruss and B.K.P. Horn, "Passive Navigation," MIT AI Memo No. 645, November, 1981.

[12] R. Jain, "Extraction of motion information from peripheral processes," IEEE Trans. Pattern Analysis and Machine Intelligence, v. PAMI-3, no. 5, pp. 489-503, 1981.

[13] W.F. Clocksin, "Perception of surface slant and edge labels from optical flow: a computational approach," Perception, v. 9, pp. 253-269, 1980.

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

[15] D. Marr and E. Hildreth, "Theory of edge detection," Proc. R. Soc. Lond., v. B 207, pp. 187-217, 1980.

[16] D. Marr and T. Poggio, "Cooperative computation of stereo disparity," Science, vol. 194, pp. 283-287, Oct. 15, 1976.

[17] H.P. Moravec, "Visual mapping by a robot rover", Proc. 6th Int. Joint Conf. on Artificial Intelligence, pp. 598-600, August 1979.

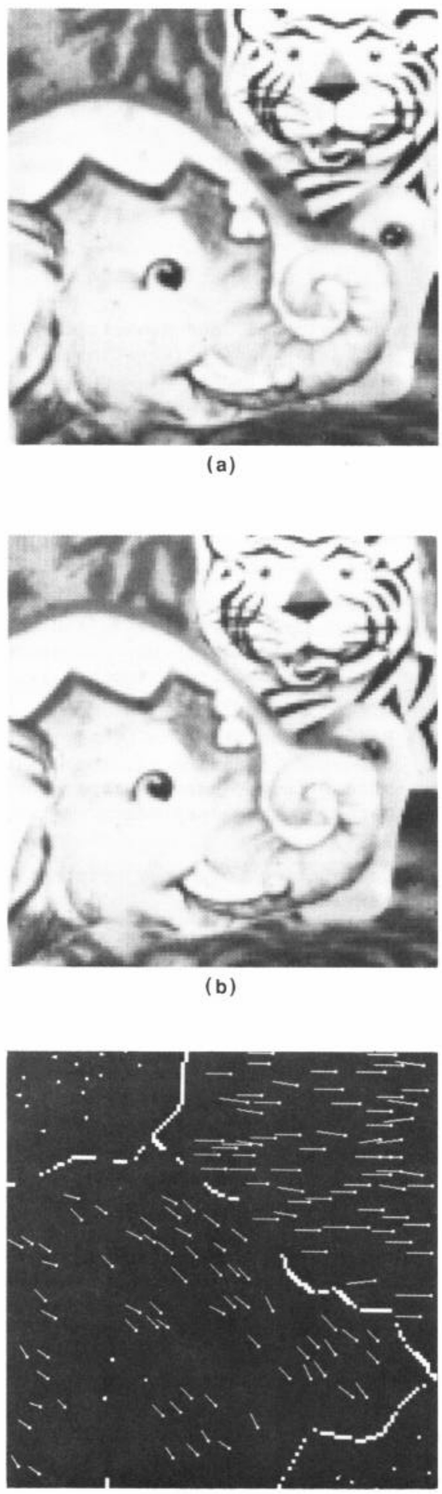

Figure 1. (a) and (b) Original image pair. (c) Vector field obtained with a point matching technique. The edges overlaid on the vector field were obtained with the technique described in the paper.

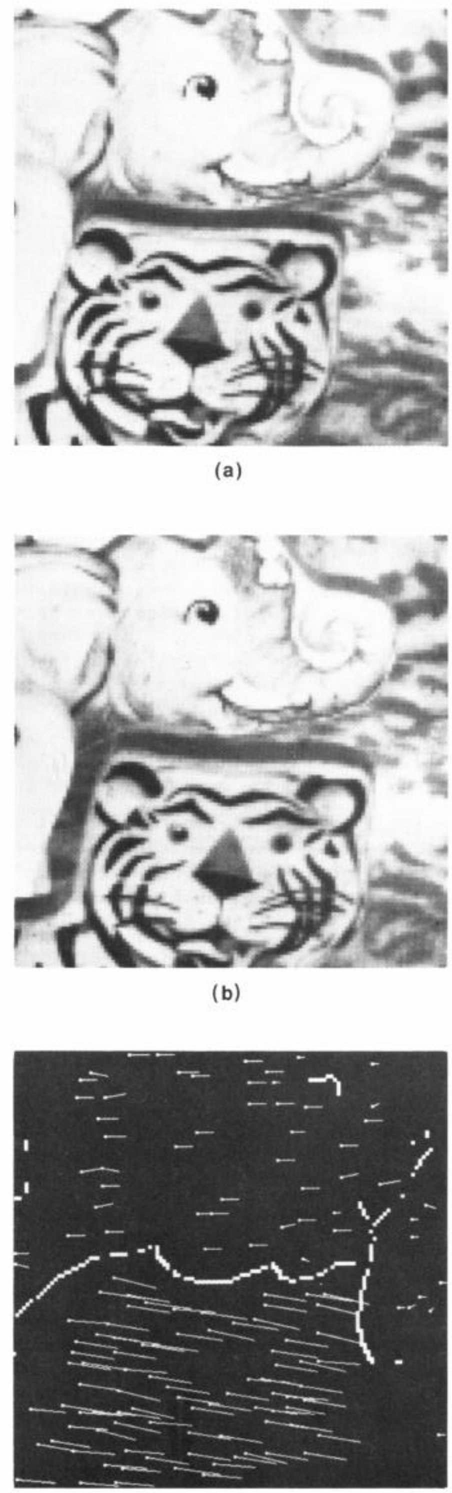

Figure 2. (a) and (b) Original image pair. (c) Vector field obtained with a point matching technique, with edges overlaid.