ABSTRACTS OF THE AAAI WORKSHOP ON PHYSICAL AND BIOLOGICAL APPROACHES TO COMPUTATIONAL VISION 1988

THE $2\frac{3}{4}$-D SKETCH

William B. Thompson *Lincoln G. Craton* *Albert Yonas*

Computer Science Department Institute of Child Development Institute of Child Development

University of Minnesota
Minneapolis, MN 55455

1 Introduction.

It has been known for many years that motion information provides a cue for depth. Two rather distinct types of information are provided. Relative motion of surface points is an indication of the relative depth of the points. (In this article, we will use the term *depth* to indicate the range from the observer to visible surface points.) If the surface points in question are part of the same rigid object, the analysis of relative visual motion leads to the structure-from-motion and motion-from-structure algorithms currently receiving much attention. Motion parallax also generates relative visual motion that provides information about the overall spatial layout of a scene. The second motion cue to depth occurs at dynamic occlusion boundaries. Surfaces on either side of such boundaries are moving visual with respect to one another. Until recently, it was though that the depth cue at dynamic occlusion boundaries was due to the appearance (*accretion*) or disappearance (*deletion*) of surface texture due to the occlud*ed* surface being progressively uncovered or covered by the occlud*ing* surface.

We have shown that there is an alternate source of information for relative depth at dynamic occlusion boundaries. This information comes from the relative motion of the boundary itself with respect to the surfaces on either side. The investigation of this new cue to depth at surface boundaries is an excellent example of the productive interaction between research in computational models of vision and research in perceptual psychophysics. We start by outlining the computational theory of determining depth at boundaries due to motion. Next, we describe experiments designed to determine whether this cue is used in human perception. We finish with a number of open questions raised by this research. In particular, we argue that Marr's $2\frac{1}{2}$-D sketch is inadequate for representing surface boundaries.

2 The Boundary Flow Constraint.

Visual motion can be used to locate surface boundaries [1]. Edges in an image due to motion can arise from far fewer causes than static image cues such as brightness, color, and texture. In particular, a discontinuity in optical flow can occur *only* because there is a corresponding discontinuity in depth and/or two separate objects are moving with respect to one another. Perhaps even more important, motion provides information

This work was supported by AFOSR contract AFOSR-87-0168, NSF Grant DCR-8500899, and NICHD Grant HD-16924.

about the occluding surfaces at a boundary. For an observer undergoing pure translation through an otherwise static scene, over a local region the magnitude of optical flow is inversely proportional to depth. Thus at a boundary, the side of the edge with the larger magnitude of flow is closer, and if there is an overlap of the two surface generating the edge, the closer surface will be occluding the surface generating the smaller magnitude of flow. Observer rotation complicates the analysis, while if objects in the field of view move with respect to each other, there is no direct relationship between magnitude of flow and depth. When general object motion is possible, surfaces corresponding to regions on opposite sides of a boundary may move in arbitrary and unrelated ways.

By considering the flow values on either side of the boundary and the manner in which the boundary itself changes over time, it is usually possible to find which side of the boundary corresponds to the occluding surface, even though the depth to the surfaces on either side cannot be determined. The principal underlying the approach is that the image of the occluding contour moves with the image of the occluding surface. We define *boundary flow* to be the image plane velocity of an occlusion boundary. Boundary flow is distinct from surface flow. The *boundary flow constraint* says that at a point corresponding to the image of an occlusion boundary, the boundary flow is the same as the surface flow of the occluding surface at the boundary [1]. Any image region not satisfying the boundary flow constraint cannot correspond to an occluding surface in the scene. The constraint is particularly useful because surfaces satisfying the constraint can be identified by a classification technique that does not actually require tracking the boundary to determine the actual boundary flow. Our original work described the boundary flow constraint in terms of edges defined by discontinuities in optical flow. In fact, the boundary flow constraint is valid for any edge cue that signals a surface boundary.

3 Boundary Flow in Human Vision.

As far as we are aware, the previous explanations for the perception of depth orderings at boundaries were associated with flow magnitude or with accretion and deletion of surface texture (e.g., [2]). In determining whether or not the boundary flow constraint was actually used by the human vision system, we needed to generate displays in which boundary flow occurred, but there were no accretion or deletion effects. This was done by creating random dot kinematigrams in which two separated random dot patches moved with respect to one another. The patches moved in counter phase — alternately moving towards and away from the center of the display. The patches were sufficiently well separated such that they never overlapped. A vertical "edge" was placed between the two patches. This edge was composed of either a vertical line or a subjective contour generated by line endings. The edge was moved rigidly with one or the other of the surface patches.

Humans ranging from adults to five-month-old infants see the dot patch moving in the same way as an objective contour as being in front of the other dot patch [3,4]. Adults are known to see the effect even more strongly with the subjective contour display. In either case, the impression is of two flat surface, one moving back and forth on top of the other. The moving contour is clearly seen as the edge of this top surface.

4 Implications for Computational Models.

Our results have implications beyond simply adding another depth cue to the list of lower-level vision modules. First of all, it is difficult to easily fit the technique into the computational paradigms of Marr [5] and Barrow and Tennenbaum [6]. Marr and Barrow and Tennenbaum suggest a computational architecture with a bottom-up, linear data flow. Use of the boundary flow constraint requires that the boundary be found, the motion of the boundary determined, and the motion of the surrounding surfaces be determined prior to the determination of relative depth. To complicate the computation further, the boundary itself may be signaled only by visual motion. The linear data flow model imposes a predefined ordering on computational operations. It is not clear what ordering could work for boundary flow analysis and still perform adequately for the many other types of low-level computations that are required.

There is an even more important implication. Marr's $2\frac{1}{2}$-D sketch was proposed, in part, as an alternative to the purely 2-D segmentation-based representations that were then popular. The $2\frac{1}{2}$-D sketch was considered as an advantage as it provided 3-D information about surfaces, while not requiring the global organization of the image into "objects". The $2\frac{1}{2}$-D sketch shares one critical deficiency with segmentation-based representations, however. Both are two-dimensional representational structures. Edges in these representations are separations between two regions differing in some visual property. What is missing is any indication of the asymmetric nature of boundaries: edges corresponding to surface boundaries provide information about the occluding surface, but not the occluded surface. Thus, we need something like a $2\frac{3}{4}$-D sketch in which overlapping surfaces can be described.

One explanation of why the subjective contour displays are more effective than the objective contour displays is that the particular subjective contour that was used is a less ambiguous indicator of a depth discontinuity than is the simple straight line which could have arisen from many different causes. The suggestion is that some image cues suggest the existence of an "unsigned" depth boundary [7]. This cues indicate that one surface is in front of another, without indicating which of the surfaces is actually nearer. Cues such as boundary flow can then be used to determine that sign of the depth change. Computational analysis of this sort requires a representation of boundaries more sophisticated than that provided by current models.

References

[1] 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.

[2] G.A. Kaplan, "Kinetic disruption of optical texture: The perception of depth at an edge," *Perception and Psychophysics*, vol. 6, pp. 193-198, 1969.

[3] A. Yonas, L.G. Craton, and W.B. Thompson, "Relative motion: Kinetic information for the order of depth at an edge," *Perception & Psychophysics*, vol. 41, no. 1, 1987.

[4] L.G. Craton and A. Yonas, "Infants' sensitivity to relative motion information for depth at an edge," submitted to *Child Development*.

[5] D.A. Marr, *Vision*, San Francisco: W.H. Freeman and Company, 1982.

[6] H.G. Barrow and J.M. Tennenbaum, "Recovering intrinsic scene characteristics from images," in *Computer Vision Systems*, A.R. Hanson and E.M. Riseman, eds., New York: Academic Press, 1978.

[7] J.M. Farber and A.B. McConkie, "Optical motions as information for unsigned depth," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 5, no. 3, 1979.