INTERPRETING LINE DRAWINGS AS THREE-DIMENSIONAL SURFACES

Harry G. Barrow and Jay M. Tenenbaum

Artificial Intelligence Center
SRI International, Menlo Park, CA 94025

ABSTRACT

We propose a computational model for interpreting line drawings as three-dimensional surfaces, based on constraints on local surface orientation along extremal and discontinuity boundaries. Specific techniques are described for two key processes: recovering the three-dimensional conformation of a space curve (e.g., a surface boundary) from its two-dimensional projection in an image, and interpolating smooth surfaces from orientation constraints along extremal boundaries.

INTRODUCTION

Our objective is the development of a computer model for interpreting two-dimensional line drawings, such as Figure 1, as three-dimensional surfaces and surface boundaries. Line drawings depict intensity discontinuities at surface boundaries, which, in many cases, are the primary source of surface information available in an image: i.e., in areas of shadow, complex (secondary) illumination, or specular surfaces analytic photometry is inappropriate. Understanding how line drawings convey three-dimensionality is thus of fundamental importance.

Given a perspectively correct line drawing depicting discontinuities of smooth surfaces, we desire as output arrays containing values for orientation and relative range at each point on the implied surfaces. This objective is distinct from that of earlier work on interpretation in terms of object models (e.g. [1]) and more basic. No knowledge of plants is required to understand the three-dimensional structure of Figure 1, as can be demonstrated by viewing fragments out of context (through a mask, for example).

Ambiguity and Constraints

The central problem in perceiving line drawings is one of ambiguity: in theory, each two-dimensional line in the image could correspond to a possible projection of an infinitude of three-dimensional space curves (see Figure 2). Yet people are not aware of this massive ambiguity. When asked to provide a three-dimensional interpretation of an ellipse, the overwhelming response is a tilted circle, not some bizarrely twisting curve (or even a discontinuous one) that has the same image. What assumptions about the scene and the imaging process are invoked to constrain to this unique interpretation?

---

* This research was supported by funds from DARPA, NASA, and NSF.

We observe that although all the lines in Figure 1 look fundamentally alike, two distinct types of scene event are depicted: extremal boundaries (e.g., the sides of the vase), where a surface turns smoothly away from the viewer, and discontinuity boundaries (e.g., the edges of the leaves), where smooth surfaces terminate or intersect. Each type provides different constraints on three-dimensional interpretation. At an extremal boundary, the surface orientation can be inferred exactly; at every point along the boundary, orientation is normal to the line of sight and to the tangent to the curve in the image [2]. A discontinuity boundary, by contrast, does not directly constrain surface orientation. However, its local curvature in the image does provide a statistical constraint on the three-dimensional tangent of the corresponding space curve. The local surface normal is constrained only to be orthogonal to this tangent, and is thus free to swing about it as shown in Figure 3.

The ability to infer 3-D surface structure from extremal and discontinuity boundaries suggests a three-step model for line drawing interpretation, analogous to those involved in our intrinsic image model [2]: line sorting, boundary interpretation, and surface interpolation. Each line is first classified according to the type of surface boundary it represents (i.e., extremal versus discontinuity). Surface contours are interpreted as three-dimensional space curves, providing relative 3-D distances along each curve; local surface normals are assigned along the extremal boundaries. Finally, three-dimensional surfaces consistent with these boundary conditions are constructed by interpolation. (For an alternative model, see Stevens [3].) This paper addresses some important aspects of three-dimensional recovery and interpolation (see [1] and [4] for approaches to line classification).

INTERPRETATION OF DISCONTINUITY BOUNDARIES

To recover the three-dimensional conformation of a surface discontinuity boundary from its image, we invoke two assumptions: surface smoothness and general position. The smoothness assumption implies that the space curve bounding a surface will also be smooth. The assumption that the scene is viewed from a general position implies that a smooth curve in the image results from a smooth curve in space, and not from an accident of viewpoint. In Figure 2, for example, the sharply receding curve projects into a smooth ellipse from only one viewpoint. Thus, such a curve would be a highly improbable three-dimensional interpretation of an ellipse.

The problem now is to determine which smooth space curve is most likely. For the special case of a wire curved in space, we conjectured that, of all projectively-equivalent space curves, humans perceive that curve having the most uniform curvature and the least torsion [2]; i.e., they perceive the space curve that is smoothest and most planar. Consistent findings were reported in recent work by Witkin [5] at MIT on human interpretation of the orientation of planar closed curves.

Measures of Smoothness

The smoothness of a space curve is expressed quantitatively in terms of intrinsic characteristics such as differential curvature (k) and torsion (t), as well as vectors giving intrinsic axes of the curve: tangent (T), principal normal (N), and binormal (B). A simple measure for the smoothness of a space curve is uniformity of curvature. Thus, one might seek the space curve corresponding to a given image curve for which the integral of k' (the spatial derivative of k) was minimum. This alone, however, is insufficient, since the integral of k' could be made arbitrarily small by stretching out the space curve so that it approaches a twisting straight line (see Figure 4). Uniformity of curvature also does not indicate whether a circular arc in the image should correspond to a 3-D circular arc or to part of a helix. A necessary additional constraint in both cases is that the space curve corresponding to a given image curve should be as planar as possible, or more precisely that the integral of its torsion should also be minimized.

Integral 1 expresses both the smoothness and planarity of a space curve in terms of a single, locally computed differential measure d(kB)/ds:

$$d(kB/ds)^2 ds = (k'^2 + k^2 t^2) ds \qquad (1)$$

Intuitively, minimizing this integral corresponds to finding the three-dimensional projection of an image curve that most closely approximates a planar, circular arc, for which k' and t are both everywhere zero.

Recovery Techniques

A computer model of this recovery theory was implemented to test its competence. The program accepts a description of an input curve as a sequence of two-dimensional image coordinates. Each input point, in conjunction with an assumed center of projection, defines a ray in space along which the corresponding space curve point is constrained to lie. The program can adjust the distance associated with each space curve point by sliding it along its ray like a bead on a wire. From the resulting 3-D coordinates, it can compute local estimates for curvature k, intrinsic axes T, N, and B, and the smoothness measure d(kB)/ds. An iterative optimization procedure then adjusts distance for each point to determine the configuration of points that minimize the integral in (1).

The program was tested using input coordinates synthesized from known 3-D space curves so that results could be readily evaluated. Correct 3-D interpretations were produced for simple open and closed curves such as an ellipse, which was interpreted as a tilted circle, and a trapezoid, which was interpreted as a tilted rectangle. However, convergence was slow and somewhat dependent on the initial choice of z-values. For example, the program had difficulty converging to the "tilted-circle" interpretation of an ellipse if started either with all z-values in a plane parallel to the image plane or all randomized to be highly nonplanar.

To overcome these deficiencies, we experimented with an alternative approach that involved more local constraints. A smooth space curve can be locally approximated by arcs of circles, and circular arcs project as elliptic arcs in an image. From the principal axes of the ellipse, it is possible to infer the direction and magnitude of the tilt of the circle that generated it. The relative depth at points along a surface contour can thus be determined, in principle, by locally fitting an ellipse (five points suffice to fit a general conic) and then projecting the local curve fragment back onto the plane of the corresponding circular arc of space curve. Assuming orthographic projection, a simple linear equation results, relating differential depth along the curve to differential changes in its image coordinates: $dz = a.dx + b.dy$.

The ellipse-fitting method yields correct 3-D interpretations for ideal image data but, not surprisingly, breaks down due to large fitting errors when small amounts of quantization noise are added. Several alternative approaches that attempt to overcome these problems by exploiting global properties (e.g. symmetry, parallelism) are presently under investigation.

SURFACE INTERPOLATION

Given constraints on orientation along extremal and discontinuity boundaries, the next problem is to interpolate smooth surfaces consistent with these boundary conditions. The problem of surface interpolation is not peculiar to contour interpretation, but is fundamental to surface reconstruction, since data is generally not available at every point in the image. We have implemented a solution for an important case: the interpolation of approximately uniformly-curved surfaces from initial orientation values and constraints on orientation [6].

The input is assumed to be in the form of sparse arrays, containing local estimates of surface range and orientation, in a viewer-centered coordinate frame, clustered along the curves corresponding to surface boundaries. The desired output is simply filled arrays of range and surface orientation representing the most likely surfaces consistent with the input data. These output arrays are analogous to our intrinsic images [2] or Marr's 2.5D sketch [7].

For any given set of input data, an infinitude of possible surfaces can be found to fit arbitrarily well. Which of these is best (i.e., smoothest) depends upon assumptions about the nature of surfaces in the world and the image formation process. For example, surfaces formed by elastic membranes (e.g., soap films) are constrained to minimum energy configurations characterized by minimum area and zero mean curvature; surfaces formed by bending sheets of inelastic material (e.g., paper or sheet metal) are characterized by zero Gaussian curvature; surfaces formed by many machining operations (e.g., planes, cylinders, and spheres) have constant principal curvatures.

Uniformly Curved Surfaces

We concentrate here on surfaces that are locally spherical or cylindrical (which have uniform curvature according to any of the above criteria). These cases are important because they require reconstructions that are symmetric in three dimensions and independent of viewpoint. Many simple interpolation techniques fail this test, producing surfaces that are too flat or too peaked. An interpolation algorithm that performs correctly on spherical and cylindrical surfaces can be expected to yield reasonable results for arbitrary surfaces.

Our approach exploits an observation that components of the unit normal vary linearly across the images of surfaces of uniform curvature. Consider a three-dimensional spherical surface, as shown in Figure 5. The radius and normal vectors are aligned, and so from similar figures we have: $Nx = x/R$, $Ny = y/R$, $Nz = z/R$. A similar derivation for the right circular cylinder is to be found in [6]. The point to be noted is that for both the cylinder and the sphere, Nx and Ny are linear functions of x and y, and Nz can be derived from Nx and Ny.

An Interpolation Technique

We have implemented an interpolation process that exploits the above observations to derive the orientation and range over a surface from boundary values. It uses parallel local operations at each point in the orientation array to make the two observable components of the normal, Nx and Ny, each vary as linearly as possible in both x and y. This could be performed by a standard numerical relaxation technique that replaces the value at each point by an average over a two-dimensional neighborhood. However, difficulties arise near surface boundaries where orientation is discontinuous. We decompose the two-dimensional averaging process into several one-dimensional ones, by considering a set of line segments passing through the central point, as shown in Figure 6a. Along each line we fit a linear function, and thus estimate a corrected value for the point. The independent estimates produced from the set of line segments are then averaged. Only the line segments that do not extend across a boundary are used: in the interior of a region, symmetric line segments are used (Figure 6a) to interpolate a central value; at boundaries, an asymmetric pattern allows values to be extrapolated (Figure 6b).

The interpolation process was applied to test cases in which surface orientations were defined around a circular outline, corresponding to the extremal boundary of a sphere, or along two parallel lines, corresponding to the extremal boundary of a right circular cylinder. Essentially exact reconstructions were obtained, even when boundary values were extremely sparse or only partially constrained. Results for other smooth surfaces, such as ellipsoids, seemed in reasonable agreement with human perception.

Current work is aimed at extending the approach to partially constrained orientations along surface discontinuities, which will permit interpretation of general solid objects.

REFERENCES

1. K. Turner, "Computer Perception of Curved Objects Using a Television Camera," Ph.D. thesis, Department of Machine Intelligence and Perception, University of Edinburgh, Edinburgh, Scotland (1974).

2. H. G. Barrow and J. M. Tenenbaum, "Recovering Intrinsic Scene Characteristics from Images," in Computer Vision Systems, A. Hanson and E. Riseman, eds., pp. 3-26 (Academic Press, New York, New York, 1978).

3. K. Stevens, "Constraints on the Visual Interpretation of Surface Contours," A.I. Memo 522, M.I.T., Cambridge, Massachusetts (March 1979).

4. I. Chakravarty, "A Generalized Line and Junction Labeling Scheme with Applications to Scene Analysis," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-1, No. 2, (April, 1979).

5. A. Witkin, Department of Psychology, M.I.T., Cambridge, Massachusetts (private communication).

6. H. G. Barrow and J. M. Tenenbaum, "Reconstructing Smooth Surfaces from Partial, Noisy Information," Proc. ARPA Image Understanding Workshop, U.S.C., Los Angeles, California (Fall 1979).

7. D. Marr, "Representing Visual Information," in Computer Vision Systems, A. Hanson and E. Riseman, eds. (Academic Press, New York, New York, 1978).

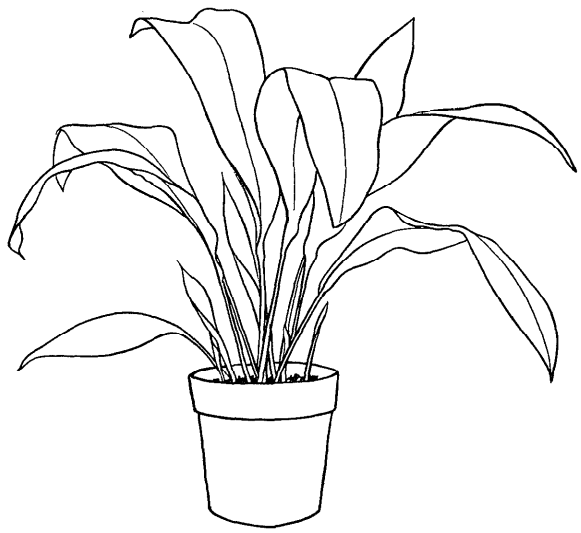


FIGURE 1 LINE DRAWING OF A THREE-DIMENSIONAL SCENE (Surface and boundary structure are distinctly perceived despite the ambiguity inherent in the imaging process.)

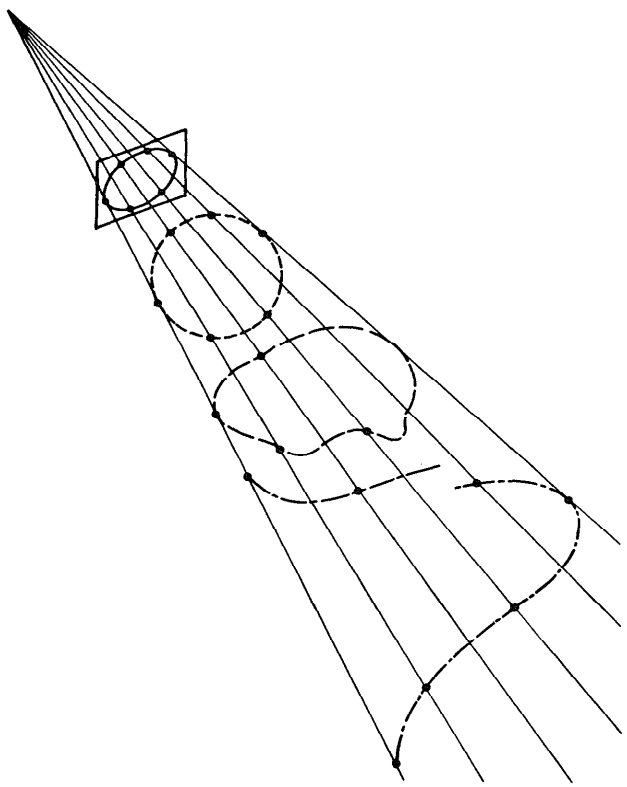


FIGURE 2 THREE-DIMENSIONAL CONFORMATION OF LINES DEPICTED IN A LINE DRAWING IS INHERENTLY AMBIGUOUS (All of the space curves in this figure project into an ellipse in the image plane, but they are not all equally likely interpretations.)

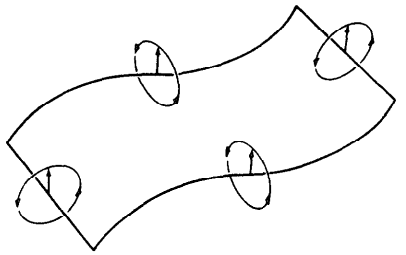


FIGURE 3 AN ABSTRACT THREE-DIMENSIONAL SURFACE CONVEYED BY A LINE DRAWING (Note that surface orientation is constrained to one degree of freedom along discontinuity boundaries.)

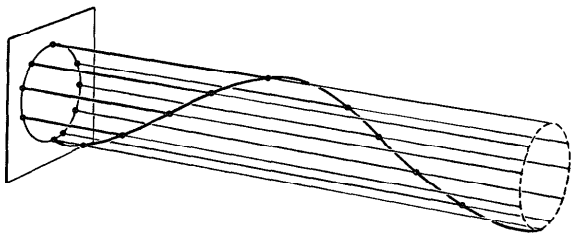


FIGURE 4 AN INTERPRETATION THAT MAXIMIZES UNIFORMITY OF CURVATURE

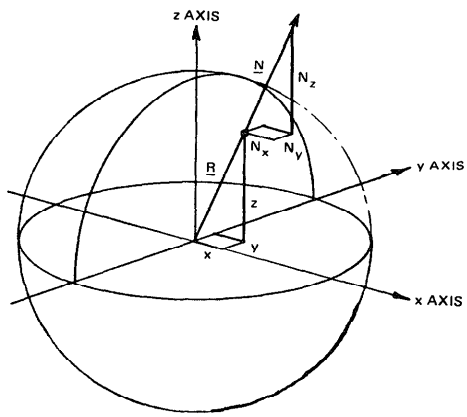
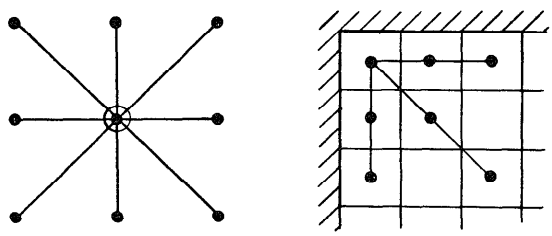


FIGURE 5 LINEAR VARIATION OF N ON A SPHERE



(a) symmetric (b) asymmetric

FIGURE 6 LINEAR INTERPOLATION OPERATORS