11

Interpreting line-drawings as 3 – dimensional surfaces

H. G. Barrow* and J. M. Tenenbaum*
Artificial Intelligence Center
SRI International, USA

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.

1. INTRODUCTION

Surface perception plays a fundamental role in early visual processing, both in humans and machines [1, 2]. Information about surfaces comes from various sources: stereopsis, motion parallax, texture gradient, and shading, to name a few. In all of these, discontinuities of brightness at surface boundaries are crucial because they are comparatively easy to detect, and they provide constraints that allow the image to be interpreted in terms of physically significant events. In many cases such discontinuities provide the primary source of information; for example, in areas of shadow or complex illumination where determination of shape from shading is difficult, or in textureless areas where texture gradient and stereopsis are inapplicable. The value of the discontinuities to human vision is amply demonstrated by our ability readily to interpret line drawings of scenes. Understanding how line drawings convey three-dimensionality is thus of fundamental importance.

Our objective is the development of a computer model for interpreting two-dimensional line drawings, such as Fig. 1, as three-dimensional surfaces and

* Present address: Artificial Intelligence Laboratory, Fairchild Camera and Instrument Corporation, Palo Alto, Ca94304.

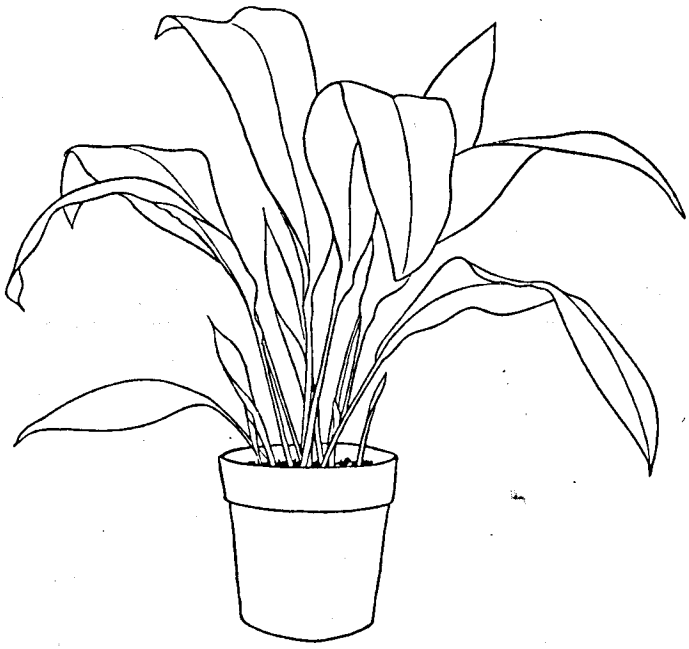

Fig. 1 – Line drawing of a three-dimensional scene. Surface and boundary structure are distinctly perceived despite the ambiguity inherent in the imaging process.

surface boundaries. Specifically, given a perspectively correct line drawing depicting discontinuities of smooth surfaces, we desire arrays containing values for orientation and relative range at each point on the implied surfaces. The interpretation of line drawings as three-dimensional surfaces is distinct from earlier work on interpretation in terms of object models [3-6] and more fundamental. No knowledge of plants is required to understand the three-dimensional structure of Fig. 1, as can be demonstrated by viewing out of context (e.g., through a mask).

A. Ambiguity and Constraints

The central problem in perceiving line drawings is one of ambiguity: in theory, each two-dimensional line in the image corresponds to a possible projection of an infinitude of three-dimensional space curves (see Fig. 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 this unique interpretation?

We observe that although all the lines in Fig. 1, look fundamentally alike, two distinct types of scene event are depicted: extremal boundaries (e.g., the

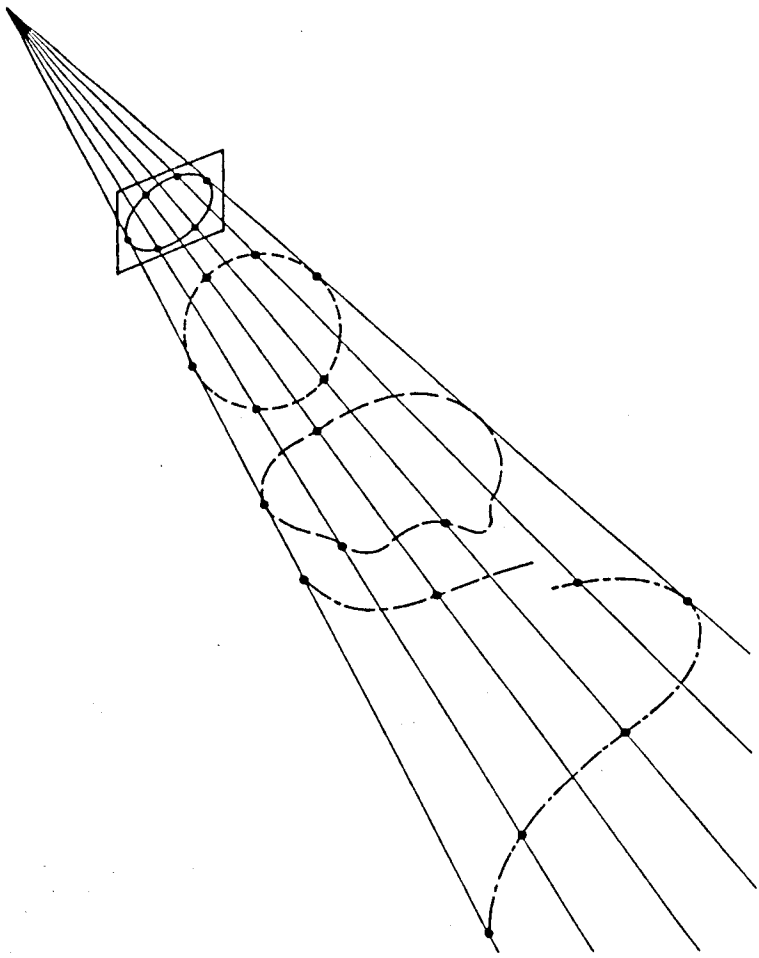

Fig. 2 – Three-dimensional conformation of lines depicted in a line drawing is inherently ambiguous. All of the space curves in this figure project into an ellipse in the image plane, but they are not all likely interpretations.

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 constaints 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 [1].

A discontinuity boundary, by contrast, does not directly constrain surface orientation. However, its local two-dimensional curvature in the image does provide a statistical constraint on the local plane of the corresponding three-

dimensional space curve, and thus relative depth along the curve. Moreover, the surface normal at each point along the boundary is then constrained to be orthogonal to the three-dimensional tangent in the plane of the space curve, leaving only one degree of freedom unknown, i.e., the surface normal is hinged to the tangent, free to swing about it as shown in Fig. 3.

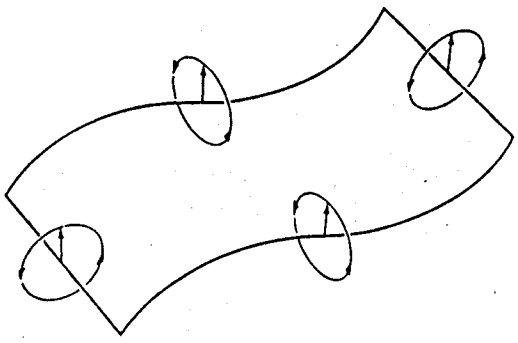

Fig. 3 – An abstract three-dimensional surface conveyed by a line drawing. Note that surface orientation is constrained to one degree of freedom along discontinuity boundaries.

The ability to infer 3D 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 [1]: 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 3D 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 [7].

The following two sections elaborate two key elements of the above model. The first deals with the problem of inferring the three-dimensional conformation of a discontinuity boundary from its image contour. The second presents an approach for interpolating smooth surfaces consistent with orientation constraints along boundaries.

2. 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 Fig. 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 it to determine which smooth curve is most likely. For the special case of a wire curved in space, which can be regarded as a thin, ribbon-like surface, we conjectured that, of all projectively-equivalent space curves, humans perceive that curve having the most uniform curvature and the least torsion [8] i.e., they perceive the space that is smoothest and most planar. The ellipse in Fig. 2 is thus almost universally perceived as a tilted circle. Consistent findings were reported in recent work by Witkin [9] at MIT on human interpretation of the orientation of planar closed curves.

A. 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). k is defined as the reciprocal of the radius of the osculating circle at each point on the curve. N is the vector from the centre of curvature normal to the tangent. B, the vector cross-product of T and N, defines the normal to the plane of curve. Torsion t is the spatial derivative of the binormal and expresses the degree to which the curve twists out of a plane. (For further details, see any standard text on vector differential geometry.)

An obvious 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 Fig. 4). Uniformity of curvature also does not indicate whether a circular arc in the image should correspond to a 3D 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.

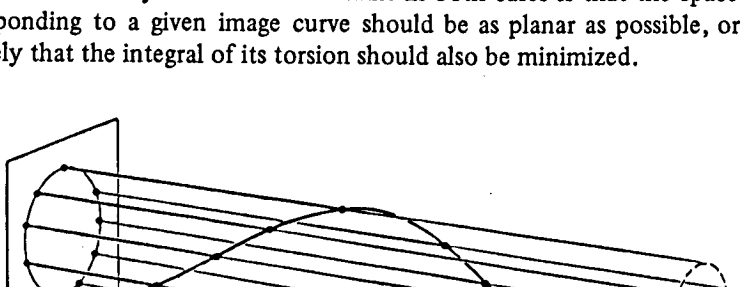

Fig. 4 – An interpretation that maximizes uniformity of curvature.

Integral (1) expresses both the smoothness and planarity of a space curve in terms of a single, locally computed differential measure $d(kB)/ds$. To interpret an image curve, it is thus necessary to find the projectively equivalent space curve that minimizes this integral.

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

Intuitively, minimizing (1) 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.

B. 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 centre of projection, defines a ray in space along which the corresponding space curve point is constrained to lie (Fig. 5). 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 3D 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 was used to determine the configuration of points that minimized the integral in equation (1). The optimization proceeded by independently adjusting each space curve point to minimize $d(kB)/ds$ locally. (Note that local perturbations of z have only local effects on curvature and torsion.)

The program was tested using input coordinates synthesized from known 3D space curves so that results could be readily evaluated. Correct 3D 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 based on ellipse fitting that involved more local constraints. Mathematically, a smooth space curve can be locally approximated by arcs of circles. Circular arcs project as elliptic arcs in an image. We already know that an ellipse in the image corresponds to a circle in three-dimensional space; the plane of the circle is obtained by rotating the plane of the ellipse about its major axis by an angle equal to arc-cos(minor axis/major axis). 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.

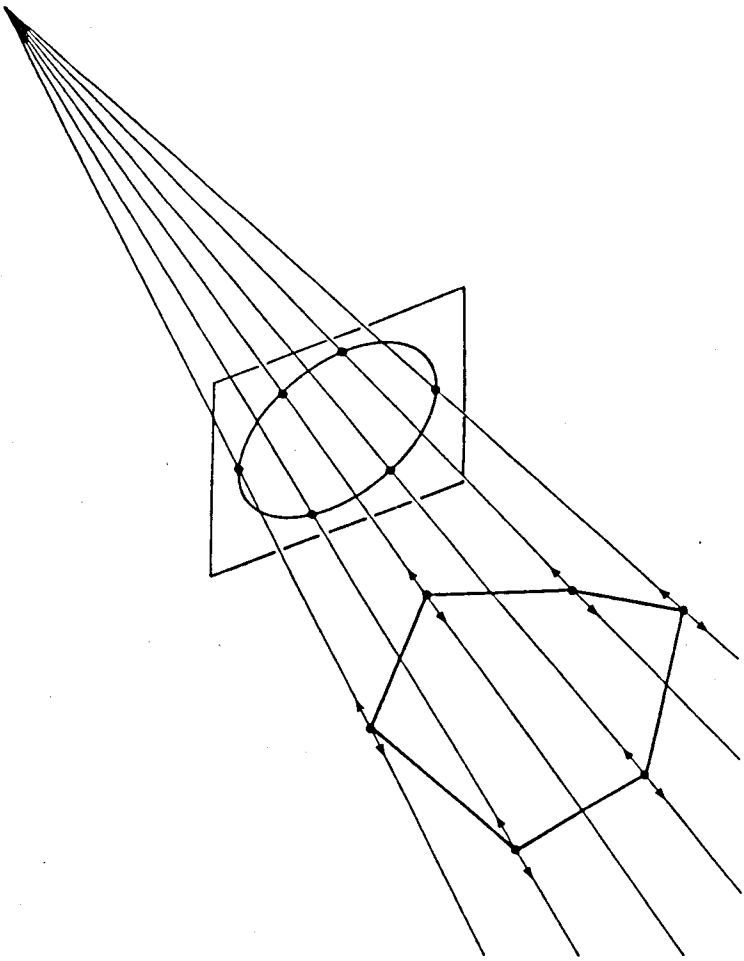

Fig. 5 – An iterative procedure for determining the optimal space curve corresponding to a given line drawing. Projective rays constrain the three-dimensional position associated with each image point to one degree of freedom.

Assuming orthographic projection, a simple linear equation results, relating differential depth along the curve to differential changes in its image coordinates, as shown in equation (2):

$$dz = a \cdot dx + b \cdot dy \tag{2}$$

The ellipse-fitting method yielded correct 3D interpretations for ideal image data but, not surprisingly, broke down owing to large fitting errors when small amounts of quantization noise were added.

Two other possible solutions are currently under consideration: a hierarchical approach in which gross orientation is first determined from large fragments of

an image curve; and a two-dimensional approach, in which refinement of boundary interpretations is integrated with the process of interpolating smooth surfaces over the interior regions. The second alternative is appealing on several grounds. First, it avoids explicit segmentation of the image curve into uniform fragments, a process likely to be both expensive and error-prone. Second, it allows boundary smoothing to propagate across surfaces so that each boundary point is refined by every other, not just those immediately adjacent. Promising preliminary results with integrated boundary refinement and surface interpolation are reported in section 3.

3. 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 are 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 [10].

The input is assumed to be in the form of sparse arrays, containing local estimates of surface range and orientation, in a viewer-centred 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 [1] or Marr's 2.5D sketch [2].

For a 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 [11]; surfaces formed by bending sheets of inelastic material (e.g., paper or sheet metal) are characterized by zero Gaussian curvature [12]; surfaces formed by many machining operations (e.g., planes, cylinders, and spheres) have constant principal curvatures.

A. Uniformly Curved Surfaces

In this paper we focus 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 first the two-dimensional example in Fig. 6. Observe that the unit normal to a semicircular surface cross-section is everywhere aligned with the radius. It therefore follows that triangles *OPQ* and *PST* are similar, and so

$$OP : OQ : QP = PS : PT : TS \ . \tag{3}$$

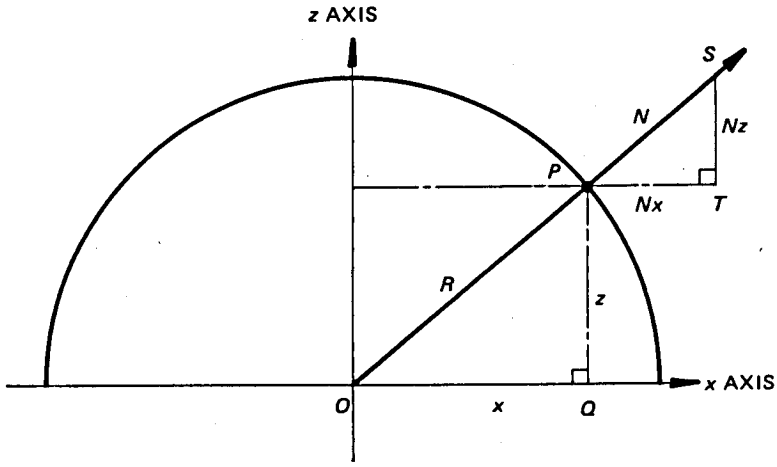

Fig. 6 – Linear variation of N across a semicircle.

But vector *OP* is the radius vector (x,z) and *PS* is the unit normal vector (N_x, N_z). Moreover, the length *OP* is constant (equal to R), and the length *PS* is also constant (equal to unity). Hence,

$$N_x = x/R \quad \text{and} \quad N_z = z/R \ . \tag{4}$$

Now consider a three-dimensional spherical surface, as shown in Fig. 7. Again the radius and normal vectors are aligned, and so from similar figures we have $N_x = x/R \quad N_y = y/R \quad \text{and} \quad N_z = z/R \ .$ (5)

A similar derivation for the right circular cylinder is to be found in [10]. The point to be noted is that for both the cylinder and the sphere, N_x and N_y are linear functions of x and y, and N_z can be derived from N_x and N_y.

B. 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, N_x and N_y each vary as linearly as possible in both x and y. This could be performed by a standard numerical

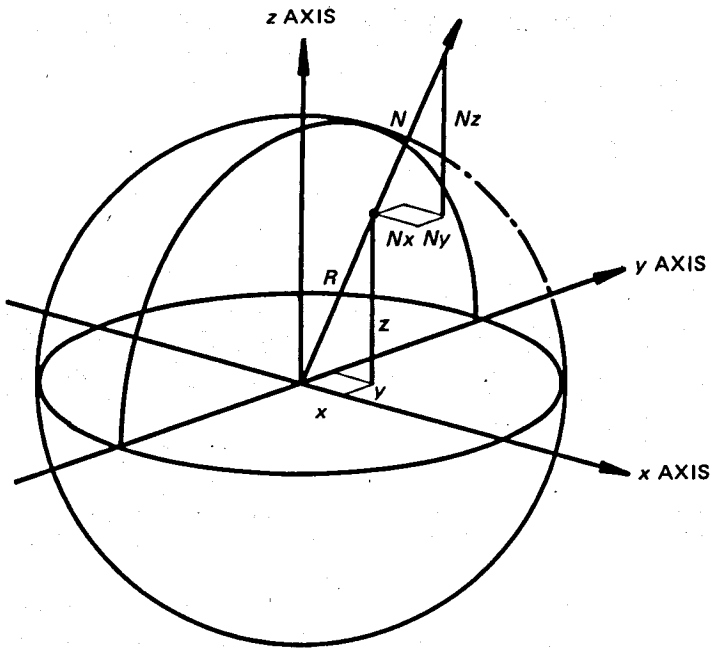

Fig. 7 – Linear variation of N on a sphere.

relaxation technique that replaces the value at each point by an average over a two-dimensional neighbourhood. 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 Fig. 8. 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.

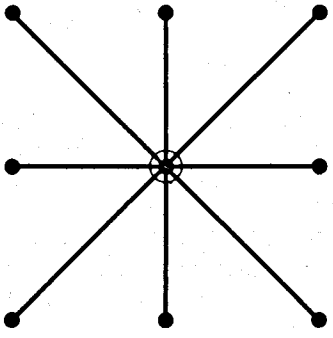

Fig. 8 – Symmetric linear interpolation operators.

Only the line segments that do not extend across a boundary are used: in the interior of a region, symmetric line segments are used (Fig. 8) to interpolate a central value; at boundaries, an asymmetric pattern allows values to be extrapolated (Fig. 9).

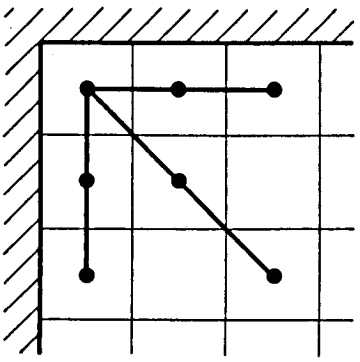

Fig. 9 – Asymmetric linear interpolation operators.

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.

4. SUMMARY

We have made a start toward a computational model for interpreting line drawings as three-dimensional surfaces. In the first section we proposed a three-step model for interpretation, based on constraints on local surface orientation along extremal and discontinuity boundaries. We then described specific computational approaches 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.

Some important open problems remain. Our technique for interpreting a three-dimensional space curve is slow and ineffective on noisy image curves. Also the surface interpolation technique must be extended to handle partially constrained orientations along discontinuity boundaries. Aspects of the problem

that we have not considered here include classification of lines into the type of physical boundary each represents (e.g., extremal or discontinuity boundary) and the initial extraction of line drawings from grey-level imagery. Both of these problems have received much attention in isolation, but we conjecture that satisfactory solutions will not be found until they are considered in the context of a system that understands about surfaces.

Current approaches to surface perception, however, have a fundamental flaw: their dependence on idealized models of lighting and surface reflectance precludes their applicability in real scenes. Research into line-drawing interpretation is therefore significant because of its potential for explaining surface perception without recourse to analytic photometry.

REFERENCES

[1] Barrow, H. G., & Tenenbaum, J. M., (1978). Recovering intrinsic scene characteristics from images, *Computer Vision Systems*, pp. 3-26, (eds. Hanson, A., & Riseman, E.). New York: Academic Press.

[2] Marr, D., (1978). Representing visual information, *Computer Vision Systems*, pp. 61-80), (eds. Hanson, A., & Riseman, E.). New York: Academic Press.

[3] Roberts, L. G., (1965). Machine perception of three-dimensional solids, *Optical and Electro-Optical Information Processing*, (eds. Tippett, J. T., et al.). Cambridge, Mass.: M.I.T. Press.

[4] Falk, G., (1972). Interpretation of imperfect line data as a three-dimensional scene, *Artificial Intelligence*, 4, No. 2, 101-144.

[5] Waltz, D. L., (1972). Generating semantic descriptions from drawings of scenes with shadows, *Technical Report AI-TR-271*, Artificial Intelligence Laboratory, M.I.T., Cambridge, Massachusetts (November 1972).

[6] Turner, K., (1974). Computer perception of curved objects using a television camera, PhD Thesis. Edinburgh: Department of Machine Intelligence, University of Edinburgh.

[7] Stevens, K., (1979). Constraints on the visual interpretation of surface contours, *A.I. Memo 522*. Cambridge, Mass.: Artificial Intelligence Laboratory, M.I.T.

[8] Barrow, H. G., & Tenenbaum, J. M., *op. cit.*, p. 19, para. 4.

[9] Witkin, A., Department of Psychology, M.I.T., Cambridge, Mass. (private communication).

[10] Barrow, H. G., & Tenenbaum, J. M., (1979). Reconstructing smooth surfaces from partial, noisy information, *Proc. ARPA Image Understanding Workshop*. U.S.C., Los Angeles, California:

[11] Almgren, F. J., Jr., & Taylor, J. E., (1976). The geometry of soap films and soap bubbles, *Scientific American*, 82-93.

[12] Huffman, D. A., (1976). Curvature and creases: a primer on paper, *IEEE-TC, C-25, No. 10.*