

3-D SHAPE REPRESENTATION BY CONTOURS

ISAAC WEISS*

Massachusetts Institute of Technology

ABSTRACT

The question of 3-D shape representation is studied on the fundamental and general level. The two aspects of the problem, (i) the reconstruction of a 3-D shape from a given set of contours, and (ii) finding "natural" coordinates on a given surface, are treated by the same theory. We first set a few basic principles that should guide any shape reconstruction mechanism, regardless of its physical implementation. Second, we propose a new mathematical procedure that complies with these principles and offers several advantages over the existing ad hoc treatments. Some general results are derived from this procedure, which conform very well with human visual perception.

I. Introduction

A major component of Image Understanding is associating 3-D shapes with contours. First, given a set of contours, such as may be provided by some image processing device, one wants to infer the shape of the surface that they most likely describe. In the complementary problem, given the surface, one seeks its "natural parametrization", namely contours that will convey its essential characteristics in the most economical, yet reliable way. One known mechanism that performs these tasks well is the human visual system. A few drawn lines can create a surprisingly vivid and convincing impression of a 3-D shape [Barrow & Tanenbaum, 1981].

In this report, we address this problem on the fundamental and general level of first, finding the principles that must guide any process of representing a surface by contours, regardless of its physical implementation, and second, we propose a mathematical mechanism that can perform these tasks, conforming with our principles. Nevertheless, as the eye is the most successful image processing system we have, we shall test the performance of our abstract procedure against its. In the following we shall summarize the requirements that we impose on a surface reconstruction mechanism, and their relation to previous work.

1) A mechanism should build a surface in accordance with the information it has available about it, like boundary or other contours on it, but it should not add extraneous information of its own. It is qualitatively clear that information is closely related to the smoothness of the curve or the surface. A straight line can be described by the coordinates of the end points only, while a more complicated shape will require more information. Thus it is reasonable, and mathematically convenient, to associate minimal information with minimal curvature, and to assume that a surface reconstruction mechanism will look for a surface that will have a minimum overall curvature while fitting the given contours.

2) The reconstruction mechanism should be invariant to rotations and translation, i.e. if the input image is rotated or translated, the output should move by the same amount but not

change its shape. The principle of minimal "energy", assuming that the overall scalar curvature $\int \bar{k}^2$ is minimal, satisfies requirement (1) and (2), but not the following demands, which our proposed measure does.

3) Dimensionlessness, or scale invariance: When the distance between an object and a viewer changes, the object's apparent size changes, but not its shape. Thus, a reconstruction mechanism should yield the same output shape from an input image, regardless of its size. Thus the mathematical method of reconstruction must be dimensionless, or invariant to scale. The minimum energy principle does *not* satisfy this important demand. A simple remedy, examined later, is to multiply the energy by the total length of the contour, but this will not satisfy the rest of our requirements. The principle of maximum compactness, which seeks to maximize the ratio of the area to the square of the perimeter, is dimensionless, but it has only been successfully applied to closed planar curves.

4) Handling of different scales of variations: A variation in the tangent occurring over a small length results in a large curvature. A small bump, or "noise", will have considerable effect on the integral. A sharp corner will totally dominate it, with the integrals value being determined by the exact way the corner is formed, which should be immaterial for determining the shape. Thus, the energy principle is only applicable for very smooth curves, while the compactness criterion is inherently quite insensitive to noise. The energy principle also tends to be insensitive to slowly changing, large scale features, and it will completely ignore straight sections of the curve. Our procedure deals with both small- and large-scale variations quite successfully.

5) We would prefer the same mechanism to handle both aspects of the problem mentioned above, namely finding both the surface and a suitable set of coordinates on it. While previously suggested mechanisms only attempt to find the surface, ours also lead to a "natural" parametrization of it.

This report is an abbreviation of a paper (to appear) which contains the mathematical details.

II. Smooth Contours

As we want to represent the amount of information contained in a curve by its curvature, the natural mathematical quantities to deal with are the derivatives of the tangent vector, such as the curvature vector $\bar{k} = d\bar{t}/ds$. As we demand rotational invariance, we have to use a scalar product of these vectors. This has led to the suggestion of the principle of minimum energy, by which one wants to minimize the integral $\int \bar{k}^2 ds$, with the integral taken over the whole curve, and to its 3-D generalizations. This may be implemented in several ways: first, given the coordinates and tangents of two points on the curve, one can find the curve that will pass through them and will minimize this integral. Second, given a closed boundary such as an ellipse, one can try to interpret it as a projected image of a surface in 3-D. The human eye will usually regard an ellipse as a slanted circle, and we would like an extremum principle to yield

*Now at the Center for Automation Research, University of Maryland, College Park, MD 20742.

the circle as an extremum over the set of all curves compatible with the projected ellipse. Limiting ourselves to planar curves (zero torsion), this means the set of ellipses (produced by different slant angles), with the same major axis as the apparent one.

The energy principle fails in this task as it does not extremize the circle over the set of ellipses. Even worse, as it is not dimensionless, the extremum will depend on the apparent size of the image, contrary to our demands. We now propose a modification of the energy principle, which will work for smooth curves.

Letting Δs be the total length of the curve, we can define the dimensionless variable:

$$l = \frac{s}{\Delta s}$$

where s is the length of the curve lying between a point on the curve and one of its ends. l is a measure of the relative position of a point on the curve, regardless of the curve's length. We define the "action", in analogy to the physical quantity, as the integral:

$$A = \int_0^1 \frac{d\vec{t}}{dl} \cdot \frac{d\vec{t}}{dl} dl$$

This is equivalent to multiplying the "energy" by Δs . As both l and $\vec{t}$ are dimensionless, so is A . (Unlike the physical action). Another advantage of this normalization is related to the way the extremum is found. In general, one can use the Euler-Lagrange equations to obtain a differential equation for $\vec{t}(l)$. But this requires that both the dependent and the independent variables will have fixed values at the end points. Our normalization provides that, as l always runs between 0 and 1, unlike s .

The extremum problem is easily solved in the simple case of a curve with two boundary points. It is convenient to use a polar coordinate system, with the angle ϕ defined as the angle between the tangent $\vec{t}$ and (say) the x axis. We thus have:

$$t_x = \cos \phi, \quad t_y = \sin \phi$$

The action now reads:

$$A = \int_0^1 \left(\frac{d\phi}{dl} \right)^2 dl$$

We shall now show that given two end points with the tangent there, the curve that passes through them is a circular arc. As a special case, extremizing over the set of ellipses with their extreme points fixed, will yield a circle. Our variational variable is now $\phi(l)$, with the boundary conditions $\phi(0) = \phi_1$ and $\phi(1) = \phi_2$, where ϕ_1, ϕ_2 are the inclinations of the curve at the end points. The EL equations yield:

$$\frac{d^2 \phi}{dl^2} = 0$$

with the unique solution:

$$\phi = (\phi_2 - \phi_1)l + \phi_1$$

which is a circular arc.

III. Large- and Small-scale Variations

The simple minimum energy principle, our modification notwithstanding, suffers from a severe shortcoming: It cannot handle sharp turns in the curve. A sharp corner will completely dominate the curve, and the value of the integral will depend on the exact shape of the corner, with point-like edges leading to infinities. Even a small bump will have a considerable influence, and in fact, the smaller the bump, the greater its effect on the integral will be (keeping its shape similar). This makes the energy principle inapplicable for curves with edges, or noise, without elaborate filtering techniques. On the other hand, straight lines are ignored by this principle, regardless of how long they may be.

We propose a new kind of an extremum principle, one of whose advantages is essentially solving this small (and large) scale variation problem. (Another advantage will become apparent when we go over to 3-D). First, we parametrize the curve by a new coordinate α , running along the curve with its limits being 0 and 1. Then we define a new "action" A :

$$A = \int_0^1 \left[\left(\frac{d\vec{t}}{d\alpha} \right)^2 + \left(\frac{dl}{d\alpha} \right)^2 \right] d\alpha$$

Intuitively, one can regard α as a variable which is affected by two forces, generated by the fact that the extremizing drives α to follow l and $\vec{t}$, in analogy to an inertia effect. The second term in A pushes α to be as close as possible to l . In the absence of the first term (i.e. a straight line) α will coincide with l . This makes the first term as close as possible to the notion of curvature. On the other hand, the first term drives α to follow $\vec{t}$, (or ϕ) to some extent and concentrate in regions where $\Delta \vec{t}$ is large (such as a corner), so that $\Delta \vec{t}/\Delta \alpha$ does not get out of control.

As a "physical" analogy, one may think of a spring which can be stretched and bent unevenly along its length. The variable " α " will represent the amount of mass from a point on the spring to one end (as opposed to the spatial length l). The analogy is not perfect, though, as an actual spring usually does not have a dimensionless A . The spring is "ideal" in the sense that it, or part of it, can shrink to an infinitesimal length (like an ideal gas). This is what will happen in a corner. A finite "mass" $\Delta \alpha$ will concentrate in an infinitesimal corner, allowing the finite bending $\Delta \phi$ to spread over it.

In extremizing A , the unknowns are the functions $l(\alpha)$ and $\vec{t}(\alpha)$ (or equivalently $\phi(\alpha)$). Solving for them will give us both $\alpha(l)$, (the distribution of the "mass" along the mass, and $\vec{t}(l)$, which defines the curve. These unknown functions are not always completely independent, which may reduce the number of unknowns, or they may depend on a parameter such as the slant, that will become an unknown to be found by the extremization process. As we shall see, the first term will be rather dominant in sections of the curves having rapid curvature changes over short length, such as bumps, corners, or saw-teeth. The second term will be the major contributor in the large-scale features, having slow variations in curvature, such as the other boundaries of the saw. Thus we have obtained a "natural" way of treating curves with two very different scales of curvature change, without having to use arbitrarily preset filtering widths, as is commonly done.

A simple demonstration of the workings of this procedure can be made for a simple corner. We shall now examine the influence of a corner on our new A , as compared to its contribution to the energy.

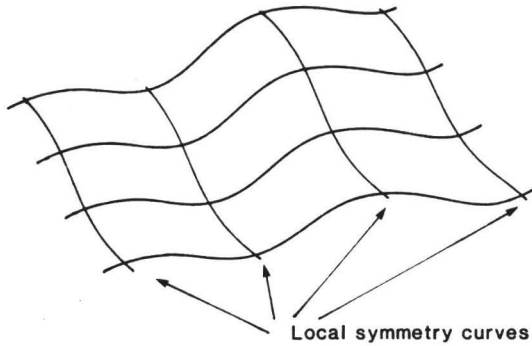


Figure 1. Local symmetry curves

We can build a corner from two "sticks" of length l , making an angle $\Delta\phi$ between themselves. We keep $\Delta\phi$ fixed, and we are only interested now in finding the distribution of the α along this corner. This turns out to be a quite simple extremum problem. The increments of α along the sticks and on their joint is, respectively,

$$\Delta\alpha_{stick} = \frac{l}{2l - \Delta\phi}, \quad \Delta\alpha_{joint} = \frac{\Delta\phi}{2l + \Delta\phi}$$

As we noted before, the second quantity is indeed finite. (In fact, we derived this result by treating the joint as a small circular arc, with length l_{joint} , which tends to a point.) Substituting these quantities in A we obtain in the extremum:

$$A = (1 + \Delta\phi)^2$$

which is independent of the exact shape and size of the corner, only on the total turn $\Delta\phi$.

In comparison, the (modified) energy principle will give for this corner:

$$A_{energy} = \frac{\Delta\phi^2}{\Delta l_{joint}}$$

which tend to infinity as the length of the corner goes to zero.

A bump can be regarded as a series of relatively sharp turns. Thus, given our expression for A at a corner, it is clear that its contribution is not dominating. With filtering, that will reduce the turns $\Delta\phi$ along the bump to small values, the influence of the bump will be quite negligible.

IV. Skew-symmetry

We now collect 4 corners to form a parallelogram, which is a skew-symmetric shape. We shall allow the $\Delta\phi$ -s of the corners to vary, in accordance with with interpreting the parallelogram as a shape with a slant in 3-D. We want to extremize A in respect to these $\Delta\phi$ -s, to find the slant. The problem is complicated slightly by the fact that there is a constraint, namely that the sum of the angle of the 4 corners is equal to 2π . This can be handled by the method of Lagrange multipliers, and the result is that all the corner angles are equal, namely equal to $\pi/2$. Hence, the parallelogram is interpreted by our extremum principle as a slanted rectangle, which is consistent with human perception. This result can be extended to general skew-symmetric shapes.

V. General Surfaces in 3-D

So far we have dealt with planar surfaces that can be described by a one-parameter boundary curve. We now turn to

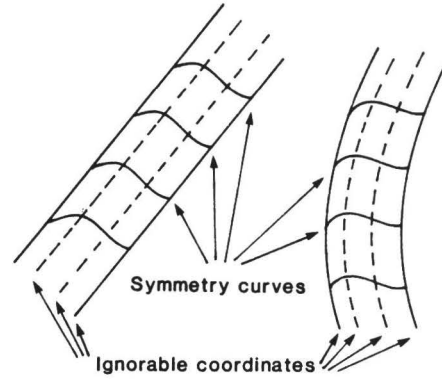


Figure 2. Ribbons with symmetry curves

the general case of a (reasonably) arbitrary surface in 3-D, such a surface can be parametrized by two coordinates α_1, α_2 . Our action A will now be a double integral:

$$A = \int_0^1 \int_0^1 \sum_i \left[\left(\frac{d\vec{r}}{d\alpha_i} \right)^2 + \left(\frac{dl}{d\alpha_i} \right)^2 \right] d\alpha_1 d\alpha_2$$

where the summation is on $i = 1, 2$.

Given a set of known contours, either on the boundary or otherwise, we construct the surface that best fits these contours as the one that extremizes the above integral. Moreover, the α_i will now have a more tangible meaning than in the one-parameter case: they will be the natural coordinates parametrizing the surface. This is another advantage of the new procedure over the energy principle. It is interesting that the variables α_i can serve the dual purpose of both handling bumps (and straight lines) and provide a set of 3-D natural coordinates.

It is easy to prove, for example, that a circular boundary will give rise to a sphere, complete with its longitudinal and latitudinal coordinates. Rather than do that, we shall state a general theorem about curves on surfaces, which will be very useful in finding natural coordinates as well as the surfaces themselves (including the sphere).

We first define a curve of local symmetry, Γ as one which divides the surface in two parts that are symmetric in the vicinity of the curve. Put another way, a reflection, say of the surface lying to right side of Γ will match the left side, near Γ . An example is the center line of any fold, ridge, valley or corrugation on a surface (fig 1), if this line is planar. Another example is the three symmetry lines on an ellipsoid (this is a global symmetry).

Theorem 1: A curve of local symmetry is planar.

proof: Trivial. A reflection of the right side of Γ will not match the left side if Γ is not planar. (fig 1).

Theorem 2: A local network of coordinates, consisting of a local symmetry curve Γ and curves that are orthogonal to it, locally extremizes the action A .

By "locally extremize" we mean that an infinitesimal variation of these curves, in the vicinity of Γ , will leave A unchanged.

We shall discuss consequences of the last theorem. (The proof is omitted.) A very interesting class of surfaces to which our theorem is easily applicable is the one having an "ignorable coordinate", namely, its curvatures will not change as we move along this coordinate. As sub-classes one can mention (a) strips, or pipes, of arbitrary cross section, whether straight, circularly or helically bent, (fig 2), in which the ignorable coordinate is the length of the pipe, and (b) surfaces of revolution, in which the

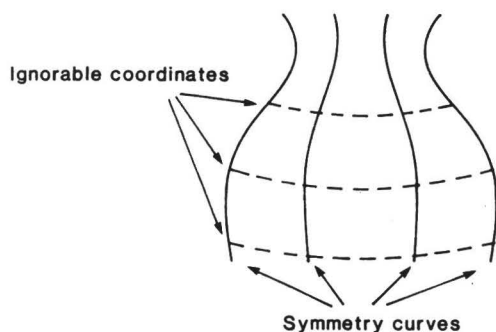


Figure 3. A surface of revolution.

azimuth ϕ is "ignorable". (fig 3). The coordinate *orthogonal* to the the ignorable one can be regarded as a local symmetry curve, as the surface stays the same on its sides. Thus, this curve, and the curves in the ignorable direction that intersect it, form a local network of curves that satisfy the conditions of theorem 1, and thus it extremizes A .

As we can see from the figures, these curves are exactly the ones that our visual intuition would expect as natural, so that we are justified in calling the curves extremizing A the "natural coordinates".

A further consequence of theorem 2 is the ability to make predictions about surfaces when they are not known in advance, without having to actually extremize A using the EL equations. It is clear that generally, the greater the number of extremal local sections a surface has, the more "extremal" the global 2-dimensional integral A will be. (This would be clearer if this "extremum" were a "minimum", which it usually is, but as we have not examined here the second derivatives, we shall stick to the term "extremum".) Thus, a surface that extremizes A will contain as many symmetry curves as are compatible with the initial data. As a consequence, we will tend to obtain surfaces containing planar, spherical or cylindrical parts, exact or approximate, rather than bumpy ones. As a particular example, we can use this general result to conclude, without having to solve the EL equations, that a circular boundary will give rise to a sphere. This is because a sphere is the most symmetric surface, thus having the greatest amount of symmetry curves. More generally, boundary contours such as of fig. 3 will give rise to a surface of revolution, as this surface consists of a collection of local networks (the meridians and parallels) which fits those contours, and similarly for other surfaces.

It is reasonable to assume, that when the conditions of the theorem are satisfied only approximately, e.g. when the curve is only approximately symmetric, a similar parametrization will still take place, with the natural coordinates approximately following the quasi-symmetry curves. Thus our theory is applicable to shapes like generalized cones, as long as the flutings vary slowly on the length-scale of the cone's radius. This also is consistent with human perception.

VI. Relation to Other Work

We have already noted the intrinsic flaws of the energy principle, and its derivatives (such as Barrow and Tenenbaum's), as surface reconstruction mechanisms. The compactness measure of Brady and Yuille [1984] does not suffer from these deficiencies. It has been applied successfully to closed planar curves, but it is hard to see how it can be made a general theory. It should be

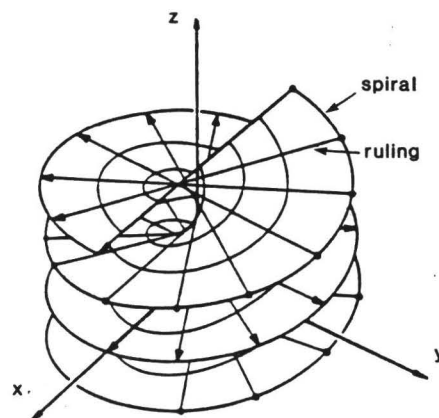


Figure 4. A helicoid

best viewed, perhaps, as a good measure of global symmetry.

Brady & Asada [1984] have studied local symmetries in planar shapes. Strictly speaking, our definition of "local" applies to an infinitesimal vicinity around a curve. In this sense, every straight line on a plane is a local symmetry curve, except at its edges. Brady examined larger vicinities, that extend to the nearby boundaries of the curve, so we shall elevate the type of symmetry he treated to a "regional" symmetry. The fact that Brady's curves are also natural coordinates indicate that the notion of symmetry is of fundamental importance at all spatial scales (global, regional, and local) of shape representation.

In differential geometry terms, our local symmetry curves are planar geodesic ones. Brady *et al* [1984] have shown that planar geodesics are usually "natural" coordinates. However, not all natural coordinates are planar geodesics. For example, the parallels on a surface of revolution, and the spirals on a helicoid (fig. 4), are not planar geodesics. The question why these look natural was left open, as they did not seem to fit any consistent rule. In our theory, these curves are natural coordinates by virtue of their being *orthogonal at each point of their length to planar geodesics*, namely the meridians and the rulings, respectively. Moreover, unlike the previous works, our results come out as a part of a general theory of shape representation derived from sound first principles.

Acknowledgements

The author thanks Michael Brady of MIT's AI Lab for carefully reading earlier drafts of this paper and making very valuable comments and suggestions.

References

- Barrow, H. G. and J. M. Tenenbaum, [1981], "Interpreting line drawings as three-dimensional surfaces", *Artif. Intell.*, 17, 75-117.
- Brady, Michael, J. Ponce, A. Yuille, H. Asada, [1984], "Describing surfaces". To be published.
- Brady, Michael, and H. Asada, [1984], "Smooth local symmetries and their implementation", A. I. Memo No. 757.
- Brady, Michael, and A. Yuille, [1984], "An extremum principle for shape from contour", *IEEE Patt. Anal. & Mach. Int.*, PAMI-6.