MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A. I. Memo No. 711 June, 1983

An Extremum Principle for Shape from Contour

Michael Brady and Alan Yuille

ABSTRACT. An extremum principle is developed that determines three-dimensional surface orientation from a two-dimensional contour. The principle maximizes the ratio of the area to the square of the perimeter, a measure of the compactness or symmetry of the three-dimensional surface. The principle interprets regular figures correctly and it interprets skew symmetries as oriented real symmetries. The maximum likelihood method approximates the principle on irregular figures, but we show that it consistently overestimates the slant of an ellipse.

Acknowledgements. This research was done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the Laboratory's artificial-intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-80-C-0505, and in part by the Systems Development Foundation.

1. Introduction

An important goal of early vision is the computation of a representation of the visible surfaces in an image, in particular the determination of the orientation of those surfaces as defined by their local surface normals [Marr 1982, Brady 1982]. Many processes contribute to achieving this goal, stereopsis and structure from motion being currently the most studied in image understanding. Three other important contributing processes are shape-from-contour, shape-from-texture--gradients, and shape-from-shading. Several psychophysical demonstrations show that shape-from-contour is significantly more powerful than shape-from-texture-gradients [Clark et. al. 1956, Gruber and Clark 1956, Braunstein and Payne, 1969]. Similarly, Barrow and Tenenbaum [1981, Figure 1.3 ff] suggest that shape-from-contour is a more effective clue to shape than shape-from-shading.

In this paper we consider the computation of shape-from-contour. Figure 1 shows a number of shapes that are typically perceived as images of surfaces which are oriented out of the picture plane. The corresponding (sets of) surface normals (up to tilt reversal) are shown to the right. It may be supposed (see for example [Gregory 1973, pages 168ff]) that the slant judgements in Figure 1 are largely determined by familiarity with regular shapes such as circles and squares. Figure 2 strains that hypothesis (though Gregory proposes that we are familiar with the shape of puddles) and suggests that the computation is based on more general knowledge of shapes and surfaces. The method we propose is based on such general knowledge, namely a preference for symmetric, or at least compact, surfaces. Note that the contour does not need to be closed in order to be interpreted as oriented out of the image plane. Finally, Figure 3 shows that, in general, contours are interpreted as curved three-dimensional surfaces.

We develop an extremum principle for determining three-dimensional surface orientation from a two-dimensional contour. Initially, we work out the extremum principle for the cases illustrated in Figures 1 and 2, that is, assuming *a priori* that the contour is closed and that the interpreted surface is planar. Later, we discuss how to extend our approach to open contours and how to interpret contours as curved surfaces as shown in Figure 3.

The extremum principle maximizes a familiar measure of the compactness or symmetry of an oriented surface, namely the ratio of the area to the square of the perimeter. It is shown that this measure is at the heart of the maximum likelihood approach to shape-from-contour developed by Witkin [1981] and Davis, Janos, and Dunn [1982]. The maximum likelihood approach has had some success interpreting irregularly shaped objects. The method is ineffective, however, when the distribution of image tangents is not random, as is the case, for example, when the image is a regular shape, such as an ellipse or

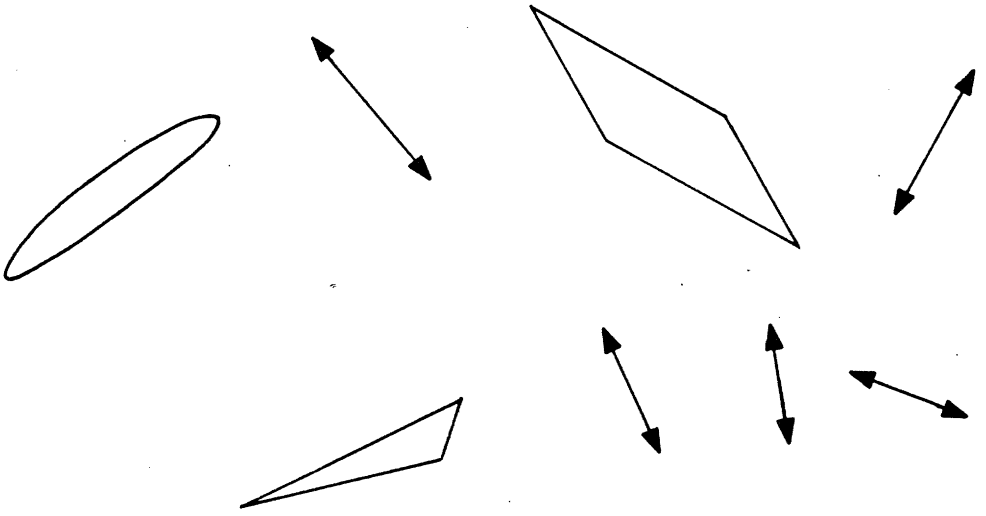

Figure 1
Two-dimensional contours that are often interpreted as planes that are oriented with respect to the image plane. The commonly judged slant is shown next to each shape.

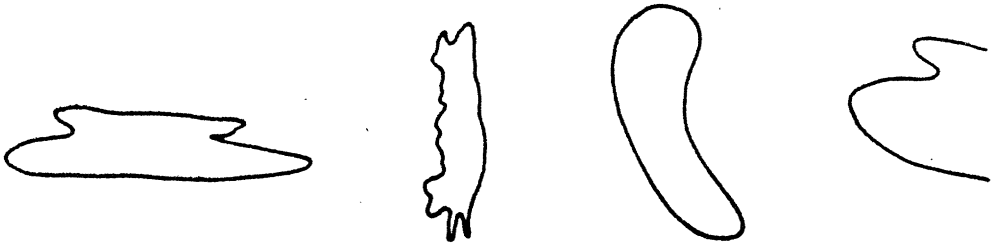

Figure 2
Some unfamiliar shapes that are also interpreted as planes that are oriented with respect to the image plane. The shape on the left is from [Gregory 1973, fig. 10.9], the others from [Witkin 1980, p 29 and 94].

a parallelogram. Our extremum principle interprets regular figures correctly. We show that the maximum likelihood method approximates the extremum principle for irregular figures; but that the maximum likelihood method does not compute the correct slant for an ellipse. Witkin [1981, Figure 5] provides

Figure 3
Some shapes that are interpreted as curved three-dimensional surfaces.

empirical evidence that the maximum likelihood method computes a good approximation to the perceived tilt but *underestimates* the slant. We prove in Appendix A that the maximum likelihood method consistently *overestimates* the slant of an ellipse. A more thorough investigation of the difference between the Extremum Principle and the Maximum Likelihood method is needed.

One class of figures that are readily perceived as lying in a plane other than the image plane are skew symmetries, which are two-dimensional linear (affine) transformations of real symmetries. Kanade [1981, page 424] has suggested a method for determining the three-dimensional orientation of skew-symmetric figures, under the "heuristic assumption" that such figures are interpreted as oriented real symmetries. We prove that our extremum principle necessarily interprets skew symmetries as oriented real symmetries, thus dispensing with the need for any heuristic assumption to that effect. Kanade shows that there is a one-parameter family of possible orientations of a skew-symmetric figure, forming a hyperbola in gradient space. He suggests that the minimum slant member of the one-parameter family is perceived. In the special case of a real symmetry, Kanade's suggestion implies that symmetric shapes are perceived as lying in the image plane, that is having zero slant. It is clear from the ellipse in Figure 1 that this is not correct. Our method interprets real symmetries correctly.

First, we review the maximum likelihood method. In Section 3, we discuss several previous extremum principles and justify our choice of the compactness

measure. In Section 4, we derive the mathematics necessary to extremize the compactness measure, and relate the extremum principle to the maximum likelihood method. In Section 5, we investigate Kanade's work on skew symmetry. One approach to extending the extremum principle to interpret curved surfaces, such as that shown in Figure 3, is sketched in Section 6. In the final section, we relate this work to the psychophysical literature on slant estimation and image understanding work on shape from texture.

2. The Sampling Approach

Witkin [1981] has treated the determination of shape-from-contour as a problem of signal detection. Recently, Davis, Janos, and Dunn [1982] have corrected some of Witkin's mathematics and proposed two efficient algorithms to compute the orientation of a planar surface from an image contour. Witkin's approach uses a *geometric model* of (orthographic) projection and a *statistical model* of (a) the distribution of surfaces in space (statistics of the universe) and (b) of the distribution of tangents to the image contour. We shall adopt the geometric model, but dispense with the statistical model in favor of an extremum principle.

First, the geometric model. Assume that the image plane is horizontal with coordinates (x, y) (see Figure 4) . To obtain a plane with slant σ and tilt τ, we rotate (x, y) by τ in the image plane and then rotate the image plane by σ about the new y axis. Assume that the coordinate frame in the plane (σ, τ) is chosen so that it projects into (x, y). (In Section 4, we describe this transformation more precisely, see also Davis, Janos, and Dunn [1982, p3].)

Now suppose that a curve is drawn in the plane (σ, τ) and denote by β the angle that the tangent makes at a typical point on the curve. Let α be the tangent angle in the image plane at the point corresponding to β. Then α and β are related by:

$$\tan(\alpha - \tau) = \frac{\tan \beta}{\cos \sigma}. \tag{2.1}$$

We now turn to the statistical model, which consists of two assumptions called *isotropy* and *independence*. Isotropy reasonably supposes that all surface orientations are equally likely to occur in nature and that tangents to surface curves are equally likely in all directions. More succinctly, it is assumed that the quantities (β, σ, τ) are randomly distributed and their joint probability density function ("density") $D(\beta, \sigma, \tau)$ is given by [Davis, Janos, and Dunn, 1982]:

$$D(\beta, \sigma, \tau) = \frac{1}{\pi^2} \sin \sigma \tag{2.2}$$

We assume that the ranges of the angles are: $0 \leq \sigma \leq \frac{\pi}{2}, 0 \leq \tau \leq \pi, 0 \leq \beta \leq \pi$. Similarly, the density of (σ, τ) is given by :

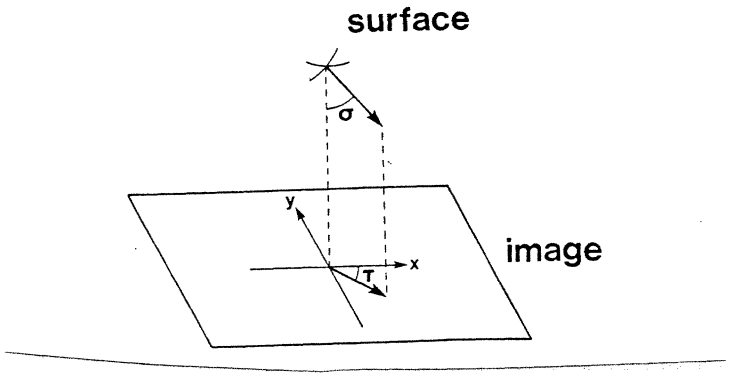

Figure 4
The geometry of orthographic projection. See text for details of notation.

$$D(\sigma,\tau) = \frac{1}{\pi}\sin\sigma \tag{2.3}$$

The independence assumption requires that the image tangents $\{\alpha_i{:}1 \leq i \leq n\}$ are statistically independent. That is, it is assumed that the tangent directions at different points on the image curve are independent. This is only true if the contour is highly irregularly shaped, or if the number of samples is small. In any case, the assumption of independence is an inherent weakness of the sampling approach (see for example [Witkin 1981, p. 36]).

Witkin [1981, p. 25 - 26] shows that the joint density function $D(\alpha,\sigma,\tau)$ is given by:

$$D(\alpha,\sigma,\tau) = \frac{1}{\pi^2}\frac{\sin\sigma\cos\sigma}{\cos^2(\alpha-\tau)+\cos^2\sigma\sin^2(\alpha-\tau)}. \tag{2.4}$$

For the conditional density $D(\alpha|\sigma,\tau)$ we find:

$$D(\alpha|\sigma,\tau) = \frac{1}{\pi}\frac{\cos\sigma}{\cos^2(\alpha-\tau)+\cos^2\sigma\sin^2(\alpha-\tau)}. \tag{2.5}$$

Denote the sample $(\alpha_1,\alpha_2,\ldots\alpha_n)$ by A (the sample is independent by assumption). It has conditional density

$$D(A|\sigma,\tau) = \prod_{i=1}^{n} D(\alpha_i|\sigma,\tau) \tag{2.6}$$

By Bayes' formula we obtain

$$D(\sigma,\tau|A) = \frac{D(A|\sigma,\tau)D(\sigma,\tau)}{\int\int D(A|\sigma,\tau)D(\sigma,\tau)d\sigma\, d\tau} \tag{2.7}$$

Observe that the numerator is independent of σ and τ. The sampling approach takes a random sample A and defines the most likely orientation of the plane (σ,τ) to be that which extremizes $D(\sigma,\tau|A)$. Witkin [1981] quantizes σ and τ, and describes an algorithm to find the maximizing (σ_j,τ_k). Davis, Janos, and Dunn [1982] develop a more efficient algorithm that first estimates σ and τ and then uses those estimates in a Newton iterative process. They provide evidence that their method is more accurate than Witkin's. Curiously, however, they state [Davis, Janos, and Dunn 1982, p 24] that "the iterative algorithm was not used [in the experiments they report] because the intitial estimates (whose computation is trivial) are very accurate and the iterative scheme often failed to converge to the solution".

3. Extremum Principles.

Brady and Horn [1983] survey the use of extremum principles in image understanding. The choice of performance index or measure to be extremized, and the class of functions over which the extremization takes place, are justified by appealing to a model of the geometry or photometry of image forming and constraints such as smoothness. For example, the use of extremum principles in surface reconstruction is based upon surface consistency theorems [Grimson 1981, 1982, 1983, Yuille 1983] and a thin plate model of visual surfaces [Brady and Horn 1983, Terzopoulos 1983].

There are several plausible measures of a curve that might be extremized in order to compute shape-from-contour. First, $\oint \kappa^2 ds$, where κ is the curvature of the contour, has been investigated as a curve of least energy for interpolating across gaps in plane curves [Horn 1981]. Here we seek a measure of a curve that is extremized when the plane containing the curve is slanted and tilted appropriately. Contrary to what appears to be a popular belief, given an ellipse in the image plane, $\oint \kappa^2 ds$ is *not* extremized in the plane that transforms the ellipse into a circle. Appendix B contains a proof of this assertion. Since ellipses are normally perceived as slanted circles, we reject the square curvature as a suitable measure.

Another possible measure is proposed by Barrow and Tenenbaum [1981, p89]. Assuming planarity (the torsion τ is zero), it reduces to

$$\oint \left(\frac{d\kappa}{ds}\right)^2 ds.$$

One objection to this measure is that it involves high-order derivatives of the curve. This means it is overly dependent on small scale behaviour. Consider, for example, a curve which is circular except for a small kink. The circular part of the curve will contribute a tiny proportion to the integral even when the plane containing the curve is rotated. The kink, on the other hand, will contribute an arbitrarily large proportion and so will dominate the integral no matter how small it is compared with the rest of the curve. This is clearly undesirable. For example, it suggests that the measure will be highly sensitive to noise in the position and orientation of the points forming the contour.

A second objection to the measure proposed by Barrow and Tenenbaum is that it is minimized by, and hence has an intrinsic preference for, straight lines, for which $d\kappa/ds$ zero. This means that the measure has a bias towards planes that correspond to the (non-general) side-on viewing position. These planes are perpendicular to the image plane and have slant $\pi/2$.

We base our choice of measure on the following observations.

1. Contours that are the projection of curves in planes with large slant are most effective for eliciting a three-dimensional interpretation.
2. A curve is foreshorted by projection by the cosine of the slant angle in the tilt direction, and not at all in the orthogonal direction.

We conclude that three-dimensional interpretations are most readily elicited for shapes that are highly elongated in one direction. Another way to express this idea is that the image contour has large aspect ratio or is radially asymmetric. The measure we suggest will pick out the plane orientation for which the curve is most compact and most radially symmetric. Specifically our measure is

$$M = \frac{(Area)}{(Perimeter)^2}. \tag{3.1}$$

This is a scale invariant number characterizing the curve. For all possible curves it is maximized by the most compact one, a circle. By compact, we mean most radially symmetric. This gives the measure an upper bound of $1/4\pi$. Its lower bound is clearly zero and it is achieved for a straight line. It follows that our measure has a built-in prejudice against side-on views for which the slant is $\pi/2$.

In general, given a contour, our extremum principle will choose the orientation in which the deprojected contour maximizes M. For example an ellipse is interpreted as a slanted circle. The tilt angle is given by the minor axis of the ellipse. It is also straightforward to show that a parallelogram corresponds to a rotated square. Appendix C discusses the interpretation of several simple shapes.

In particular, an ellipse is interpreted as a slanted circle, a parallelogram as a slanted square, and a triangle as a slanted equilateral triangle. In Section 5 we extend the parallelogram result to the more general case of skewed symmetry.

We note that the quantity M is commonly used in pattern recognition and industrial vision systems [Agin 1980, Pavlidis 1977, Ballard and Brown 1982] as a feature that measures the compactness of an object. Furthermore, we can show that the measure M defined in Eq. (3.1) is at the heart of the geometric model in the maximum likelihood approach.

From Section 2, we see that the maximum likelihood approach maximizes the product of a number of terms of form

$$f(\alpha) = \frac{\cos\alpha}{\cos^2(\alpha - \tau) + \cos^2\sigma\sin^2(\alpha - \tau)}. \tag{3.2}$$

Differentiating Eq. 2.1 with respect to the arc length s_I along the image curve and s_R along the rotated curve respectively we obtain

$$\frac{\kappa_I\, ds_I}{\kappa_R\, ds_R} = \frac{1}{f(\alpha)} \tag{3.3}$$

where κ_I and κ_R are the curvature at corresponding points of the image contour and its deprojection in the rotated plane respectively. In fact, $\kappa_I = d\alpha/ds_I$ and $\kappa_R = d\beta/ds_R$. There is no σ or τ dependence in the numerator of equation (3.3). We can write each term $\kappa\, ds$ as $(ds\, ds)/(\rho\, ds)$ where ρ is the radius of curvature. Now observe that $(\rho\, ds)/(ds\, ds)$ is just a local computation of area divided by perimeter squared! Hence maximizing each $f(\alpha)$ in the maximum likelihood approach is equivalent to locally maximizing area over perimeter squared. In section (4) we will examine this connection more rigorously.

Finally we note that the area, as well as the perimeter, can be obtained by an integral round the contour. If $\mathbf{n}$ is the normal to the curve then it is a straightforward application of Stokes' Theorem to show that

$$(Area)\mathbf{n} = \frac{1}{2}\oint \mathbf{r} \times d\mathbf{r}, \tag{3.5}$$

where $(Area)$ is a scalar quantity, and $\mathbf{r}$ is a vector coordinate system in the plane of the figure. This formula simplifies the calculations and means the perimeter and the area can be computed simultaneously.

4. Extremizing the Measure

We now write down the measure for a curve with arbitrary orientation and then extremize with respect to the orientation. Let the unit normals to the image plane and the rotated plane be $\mathbf{k}$ and $\mathbf{n}$ respectively. The slant σ of the rotated plane is given by the scalar product

$$\cos\sigma = \mathbf{k}\cdot\mathbf{n} \tag{4.1}.$$

Let Γ_R and Γ_I be the contour in the rotated and image planes. A vector $\mathbf{r}$ in the image plane satisfies $\mathbf{r}\cdot\mathbf{k} = 0$, and is the projection of a vector $\mathbf{v}$ in the rotated plane that satisfies $\mathbf{v}\cdot\mathbf{n} = 0$. The projection relationship between $\mathbf{v}$ and its image $\mathbf{r}$ is defined by:

$$\mathbf{r} = \mathbf{k}\times(\mathbf{v}\times\mathbf{k}) = \mathbf{v} - (\mathbf{v}\cdot\mathbf{k})\mathbf{k} \tag{4.2}$$

$$\mathbf{v} = \frac{\mathbf{n}\times(\mathbf{r}\times\mathbf{k})}{(\mathbf{n}\cdot\mathbf{k})} = \frac{(\mathbf{n}\cdot\mathbf{k})\mathbf{r} - (\mathbf{n}\cdot\mathbf{r})\mathbf{k}}{(\mathbf{n}\cdot\mathbf{k})} \tag{4.3},$$

where $\times$ denotes vector product. Now Γ_R and Γ_I have (vector) areas $\mathbf{A}_R$ and $\mathbf{A}_I$ given by

$$\mathbf{A}_R = \frac{1}{2}\oint_{\Gamma_R}\mathbf{v}\times d\mathbf{v} \tag{4.4}$$

$$\mathbf{A}_I = \frac{1}{2}\oint_{\Gamma_I}\mathbf{r}\times d\mathbf{r} \tag{4.5}$$

Observe that the area vectors have the same direction as the normal to the plane containing the area. In particular, $\mathbf{A}_R$ is in direction $\mathbf{n}$ and $\mathbf{A}_I$ is in direction $\mathbf{k}$. Substituting Eq. 4.3 into Eq. 4.4 and using Eq. 4.1, we find

$$\begin{aligned}\mathbf{A}_R &= \mathbf{A}_I + \frac{\mathbf{k}\times(\mathbf{n}\times\mathbf{A}_I)}{(\mathbf{n}\cdot\mathbf{k})} \\ &= \frac{(\mathbf{k}\cdot\mathbf{A}_I)}{\cos\sigma}\mathbf{n}\end{aligned} \tag{4.6}$$

It follows that

$$\|\mathbf{A}_R\| = \frac{\|\mathbf{A}_I\|}{\cos\sigma} \tag{4.7}$$

(We recall that the range of σ guarantees that $\cos\sigma$ is positive.

The perimeter lengths P_R and P_I are given by

$$P_R = \oint_{\Gamma_R}\|d\mathbf{v}\| \tag{4.8}$$

$$P_I = \oint_{\Gamma_I}\|d\mathbf{r}\| \tag{4.9}$$

Substituting Eq. 4.3 into Eq. 4.8 gives

$$P_R = \oint_{\Gamma_R}\left\{(d\mathbf{r})^2 + \frac{(\mathbf{n}\cdot d\mathbf{r})^2}{(\mathbf{n}\cdot\mathbf{k})^2}\right\}^{\frac{1}{2}} \tag{4.10}$$

In general there is no simple relationship between the perimeters analogous to Eq. 4.6 between the areas. Nevertheless, by Eq. 4.7,

$$\frac{\|\Lambda_R\|}{P_R^2} = \frac{\|\Lambda_I\|}{P_R^2 \cos\sigma} \tag{4.11}$$

and so our extremum principle is equivalent to extremizing $\cos^{\frac{1}{2}}\sigma P_R$, which we write as

$$I = \oint_{\Gamma_R} \left\{ (\mathbf{n}\cdot\mathbf{k})d\mathbf{r}^2 + \frac{(\mathbf{n}\cdot d\mathbf{r})^2}{(\mathbf{n}\cdot\mathbf{k})} \right\}^{\frac{1}{2}} \tag{4.12}$$

We extremize this with respect to the orientation $\mathbf{n}$ of the rotated plane, maintaining the constraint that $\mathbf{n}$ is a unit vector by a Lagrange multiplier $\Lambda/4$. This gives

$$\begin{aligned} \mathbf{k}\oint & \left((\mathbf{n}\cdot\mathbf{k})d\mathbf{r}^2 + \frac{(\mathbf{n}\cdot d\mathbf{r})^2}{(\mathbf{n}\cdot\mathbf{k})} \right)^{-\frac{1}{2}} \frac{(\mathbf{n}\cdot\mathbf{k})^2 d\mathbf{r}^2 - (\mathbf{n}\cdot d\mathbf{r})^2}{(\mathbf{n}\cdot\mathbf{k})^2} \\ & + 2\oint \left((\mathbf{n}\cdot\mathbf{k})d\mathbf{r}^2 + \frac{(\mathbf{n}\cdot d\mathbf{r})^2}{(\mathbf{n}\cdot\mathbf{k})} \right)^{-\frac{1}{2}} \frac{(\mathbf{n}\cdot\mathbf{k})(\mathbf{n}\cdot d\mathbf{r})d\mathbf{r}}{(\mathbf{n}\cdot\mathbf{k})^2} \\ & + \Lambda\mathbf{n} = \mathbf{0} \end{aligned} \tag{4.13}$$

Taking scalar products of Eq (4.13) with $\mathbf{k}$ and $\mathbf{n}$, respectively, gives

$$\begin{aligned} 0 &= (\mathbf{n}\cdot\mathbf{k})\Lambda \\ &\quad + \oint \left((\mathbf{n}\cdot\mathbf{k})d\mathbf{r}^2 + \frac{(\mathbf{n}\cdot d\mathbf{r})^2}{(\mathbf{n}\cdot\mathbf{k})} \right)^{-\frac{1}{2}} \frac{(\mathbf{n}\cdot\mathbf{k})^2 d\mathbf{r}^2 - (\mathbf{n}\cdot d\mathbf{r})^2}{(\mathbf{n}\cdot\mathbf{k})^2} \end{aligned} \tag{4.14}$$

$$\begin{aligned} 0 &= \Lambda \\ &\quad + (\mathbf{n}\cdot\mathbf{k})\oint \left((\mathbf{n}\cdot\mathbf{k})d\mathbf{r}^2 + \frac{(\mathbf{n}\cdot d\mathbf{r})^2}{(\mathbf{n}\cdot\mathbf{k})} \right)^{-\frac{1}{2}} \frac{(\mathbf{n}\cdot\mathbf{k})^2 d\mathbf{r}^2 + (\mathbf{n}\cdot d\mathbf{r})^2}{(\mathbf{n}\cdot\mathbf{k})^2}. \end{aligned} \tag{4.15}$$

We use Eq. 4.14 to remove the integral coefficient of $\mathbf{k}$ in Eq. 4.13, allowing us to express Eq. 4.13 as a sum of the second integral and $\mathbf{k}\times(\mathbf{k}\times\mathbf{n})$ times Λ. We now use Eq. 4.15 to eliminate Λ from this new form of Eq. 4.13. We recall that the unit tangent $\mathbf{t}$ is defined by

$$\mathbf{t} = \frac{d\mathbf{r}}{ds},$$

where s is arc length along the contour. It follows that

$$d\mathbf{r} = \mathbf{t}ds.$$

Recalling that $(\mathbf{n}\cdot\mathbf{k}) = \cos\sigma$, we find

$$2\oint\left\{\cos^2\sigma + (\mathbf{n}\cdot\mathbf{t})^2\right\}^{-\frac{1}{2}}(\mathbf{t}\cdot\mathbf{n})d\mathbf{r} = -\mathbf{k}\times(\mathbf{k}\times\mathbf{n})\oint\left\{\cos^2\sigma + (\mathbf{n}\cdot\mathbf{t})^2\right\}^{\frac{1}{2}}ds \tag{4.16}$$

where $\mathbf{t} = d\mathbf{r}/||d\mathbf{r}||$ is the unit tangent to the image contour.

Let the unit vectors in the x and y directions in the image plane be $\mathbf{i}$ and $\mathbf{j}$ and the normal to the image plane $\mathbf{k}$. By definition, $\cos\sigma = \mathbf{k}\cdot\mathbf{k}$. The tilt τ is defined by

$$\cos\tau = \frac{\mathbf{i}\cdot\mathbf{n}}{\sin\sigma} \tag{4.17}$$

$$\sin\tau = \frac{\mathbf{j}\cdot\mathbf{n}}{\sin\sigma}. \tag{4.18}$$

The tangent vector $\mathbf{t}$ and the normal $\mathbf{n}$ can be written:

$$\mathbf{t} = \cos\alpha\mathbf{i} + \sin\alpha\mathbf{j} \tag{4.19}$$

$$\mathbf{n} = \sin\sigma\cos\tau\mathbf{i} + \sin\sigma\sin\tau\mathbf{j} + \cos\sigma\mathbf{k} \tag{4.20}$$

where α is the tangent angle in the image. We now form the scalar products of Eq. 4.16 with $\mathbf{i}$ and $\mathbf{j}$ to obtain (with appropriate cancelling)

$$\begin{aligned} 2\oint(\cos^2\sigma + \sin^2\sigma\cos^2(\alpha-\tau))^{-\frac{1}{2}}\cos(\alpha-\tau)\cos\alpha ds \\ = \cos\tau\oint(\cos^2\sigma + \sin^2\sigma\cos^2(\alpha-\tau))^{\frac{1}{2}}ds \end{aligned} \tag{4.21}$$

and

$$\begin{aligned} 2\oint(\cos^2\sigma + \sin^2\sigma\cos^2(\alpha-\tau))^{-\frac{1}{2}}\cos(\alpha-\tau)\sin\alpha ds \\ = \sin\tau\oint(\cos^2\sigma + \sin^2\sigma\cos^2(\alpha-\tau))^{\frac{1}{2}}ds \end{aligned} \tag{4.22}$$

we can rewrite these equations, after multiplying by factors of $\pm\cos\tau$ and $\pm\sin\tau$, as:

$$\begin{aligned} 2\oint(\cos^2\sigma + \sin^2\sigma\cos^2(\alpha-\tau))^{-\frac{1}{2}}\cos^2(\alpha-\tau)ds \\ = \oint(\cos^2\sigma + \sin^2\sigma\cos^2(\alpha-\tau))^{\frac{1}{2}}ds \end{aligned} \tag{4.23}$$

$$\begin{aligned} 2\oint(\cos^2\sigma + \sin^2\sigma\cos^2(\alpha-\tau))^{-\frac{1}{2}}\cos(\alpha-\tau)\sin(\alpha-\tau)ds \\ = 0 \end{aligned} \tag{4.24}$$

We use Eqs. 4.23 and 4.24 in Appendix D to determine the extremizing orientation for a skew symmetry. We can rewrite these equations in the form:

$$\frac{\partial}{\partial\sigma}\left(\frac{1}{\cos^{\frac{1}{2}}\sigma}\oint(\cos^2\sigma+\sin^2\sigma\cos^2(\alpha-\tau))^{\frac{1}{2}}ds\right)=0 \tag{4.25}$$

$$\frac{\partial}{\partial\tau}\left(\frac{1}{\cos^{\frac{1}{2}}\sigma}\oint(\cos^2\sigma+\sin^2\sigma\cos^2(\alpha-\tau))^{\frac{1}{2}}ds\right)=0 \tag{4.26}$$

to emphasize that they correspond to extremizing with respect to σ and τ.

We can implement the extremum method directly from Eqns. (4.25) and (4.26) by quantizing the tangent angle α_i, slant σ_j, and tilt τ_k, replacing the integral by a sum. This gives good results, even though we use fixed point integers in our edge finder. A multi-level search speeds the algorithm by a factor of ten. For large slant the ratio of the greatest to least value of the expression is large, and the result is numerically well-conditioned. For smaller slants (less than about 45 degrees in the case of an ellipse) the ratio is small and the result poorly conditioned, so that round-off errors can be significant.

To conclude this Section, we show that these equations are similar, though not identical, to those obtained by the maximum likelihood method in the limit as the number of sampled tangents tends to infinity. To see this we recall from Eq. 2.7 that this method involves extremizing $D(A|\sigma,\tau)$ with respect to σ and τ. Since the denominator is independent of σ and τ, this amounts to extremizing $D(A|\sigma,\tau)D(\sigma,\tau)$. This is the same as extremizing $\log D(A|\sigma,\tau)D(\sigma,\tau)$. Using 2.3, 2.5 and 2.6 we obtain:

$$E=n\log\cos\sigma+\log\sin\sigma-\sum_{i=1}^{n}\log(\cos^2(\alpha_i-\tau)+\cos^2\sigma\sin^2(\alpha_i-\tau)). \tag{4.27}$$

where we have ignored factors of π which will vanish on differentiation. Dividing E by n and taking the limit as n tends to infinity gives:

$$F=\log\cos\sigma\oint d\mathbf{r}-\oint\log(\cos^2(\alpha-\tau)+\cos^2\sigma\sin^2(\alpha-\tau))d\mathbf{r}. \tag{4.28}$$

Using the identity:

$$\cos^2(\alpha-\tau)+\cos^2\sigma\sin^2(\alpha-\tau)=\cos^2\sigma+\sin^2\sigma\cos^2(\alpha-\tau) \tag{4.29}$$

gives

$$F=\log\cos\sigma\oint d\mathbf{r}-\oint\log(\cos^2\sigma+\sin^2\sigma\cos^2(\alpha-\tau))d\mathbf{r}. \tag{4.30}$$

This formula is similar to Eqs 4.25 and 4.26. Thus we expect the Extremum Method to give similar results to the Sampling Method when the contour is sufficiently irregular. However, we can show formally that the Sampling Method and the Extremum Method *are not* equivalent. In Appendix A we show that the Sampling Method overestimates the slant of an ellipse. The precise discrepancy between the methods, and its practical consequence for computing shape from contour is currently under investigation.

5. Skew Symmetry

We now consider a more general class of shapes for which the maximum likelihood approach is not effective. Kanade [1981, sec. 6.2] has introduced *skew symmetries*, which are two-dimensional linear (affine) transformations of real symmetries. There is a bijective correspondence between skew symmetries and images of symmetric shapes that lie in planes oriented to the image plane. Kanade proposes the heuristic assumption that a skew symmetry is interpreted as an oriented real symmetry, and he considers the problem of computing the slant and tilt of the oriented plane.

Denote the angles between the x-axis of the image and the images of the symmetry axis and an axis orthogonal to it (the skewed transverse axis) by α and β respectively. The orthogonality of the symmetry and transverse axes enable one constraint on the orientation of the plane to be derived. Kanade uses gradient space (p, q) [Horn, 1977, Brady, 1982] to represent surface orientations. He shows [Kanade 1981, p. 425] that the heuristic assumption is equivalent to requiring the gradient (p, q) of the oriented plane to lie on the hyperbola

$$p_1^2 \cos^2 \frac{(\alpha - \beta)}{2} - q_1^2 \sin^2 \frac{(\alpha - \beta)}{2} = -\cos(\alpha - \beta) \tag{5.1}$$

where

$$\begin{aligned} p_1 &= p \cos\left(\frac{\alpha + \beta}{2}\right) + q \sin\left(\frac{\alpha + \beta}{2}\right), \\ q_1 &= -p \sin\left(\frac{\alpha + \beta}{2}\right) + q \cos\left(\frac{\alpha + \beta}{2}\right). \end{aligned} \tag{5.2}$$

Kanade [1981, p. 426] further proposes that the vertices of the hyperbola, which correspond to the least slanted orientation, are chosen within this one-parameter family. This proposal is in accordance with a heuristic observation of Stevens [1980]. In the special case that the skew symmetry is a real symmetry, that is in the case that $\alpha - \beta = \pm\pi/2$, the hyperbola reduces to a pair of orthogonal lines [Kanade 1981, page 426] passing through the origin. In such cases the slant is zero. In other words, Kanade's proposal predicts that real symmetries are inevitably interpreted as lying in the image plane, and hence

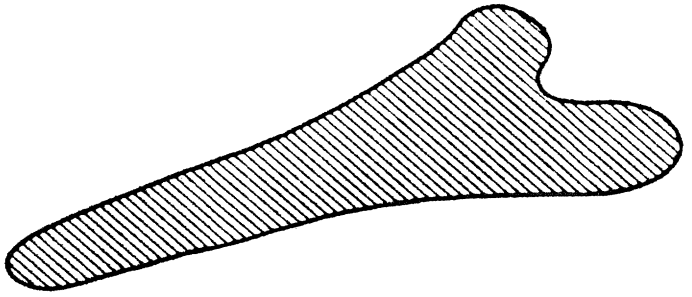

Figure 5
A partition of a skew-symmetric object into regions for the piecewise analysis presented in this section.

having zero slant. Inspection of Figure 1 shows that this is not the case. A (symmetric) ellipse is typically perceived as a slanted circle, particularly if the major and minor axes do not line up with the horizontal and vertical.

Although Kanade's minimum slant proposal does not seem to be correct, there is evidence (for example [Stevens 1980]) for Kanade's assumption that skew symmetries are interpreted as real symmetries. We now show that the assumption can in fact be *deduced* from our Extremum Principle. As a corollary, in Appendix D we determine the slant and tilt of any given skew-symmetric figure; only in special cases does it correspond to the minimum slant member of Kanade's one-parameter family.

Instead of using Eqs. 4.25 and 4.26, we prove the result from first principles since this enables us to use symmetry directly. First, we partition the image into regions as shown in Figure 5. A typical region is shown in Figure 6a and we skew it by an angle α to get Figure 6b. We show that the ratio of the area to the perimeter squared is greatest for Figure 6a with $\alpha = 0$. We show this is also true for the sum of the regions and take the limit as the width of the regions goes to zero to complete the proof.

From Figure 6a we calculate the area $A_i = ||\Lambda||$ and the the portion P_i of the perimeter of the shape to be

$$A_i = 2al + (b - a)l \tag{5.3}$$

and

$$P_i = 2\sqrt{l^2 + (b-a)^2}. \tag{5.4}$$

From Figure 6b we find

$$\begin{aligned} A_i' &= 2al\cos\alpha + (b-a)l\cos\alpha \\ &= A_i\cos\alpha \end{aligned} \tag{5.5}$$

and

$$P_i' = \sqrt{l^2\cos^2\alpha + (b-a+l\sin\alpha)^2} + \sqrt{l^2\cos\alpha^2 + (b-a-l\sin\alpha)^2}. \tag{5.6}$$

We need to show that

$$\frac{\cos\alpha}{P_i'^2} < \frac{1}{P_i^2}. \tag{5.7}$$

This is equivalent to showing

$$\begin{aligned} &(2\cos\alpha - 1)(l^2 + (b-a)^2) \\ &< \left\{l^2 + (b-a)^2 + 2(b-a)l\sin\alpha\right\}^{\frac{1}{2}}\left\{l^2 + (b-a)^2 - 2(b-a)l\sin\alpha\right\}^{\frac{1}{2}}. \end{aligned} \tag{5.8}$$

This condition clearly holds for $\pi/2 \geq \alpha \geq \pi/3$. Assume therefore that $0 \leq \alpha < \pi/3$. Squaring both sides of (5.8), we see that the condition is equivalent to

$$l^4 + (b-a)^4 + l^2(b-a)^2(2 + \frac{\sin^2\alpha}{\cos^2\alpha - \cos\alpha}) > 0. \tag{5.9}$$

On completing the square we see this always holds provided

$$4\cos\alpha - 1 - 3\cos^2\alpha > 0 \tag{5.10}$$

This is so provided $\frac{1}{3} < \cos\alpha < 1$, that is for $0 < \alpha < \frac{\pi}{3}$. It follows that the ratio of the area to the perimeter squared is maximized for each region when $\alpha = 0$.

We now partition the shape into n regions as shown in Figure 5. Let the i^{th} block have area A_i' and perimeter P_i' when the skew angle is α, and denote the area by A_i and the perimeter by P_i when α is zero. It follows from the above results that

$$A_i' = A_i\cos\alpha \tag{5.11}$$

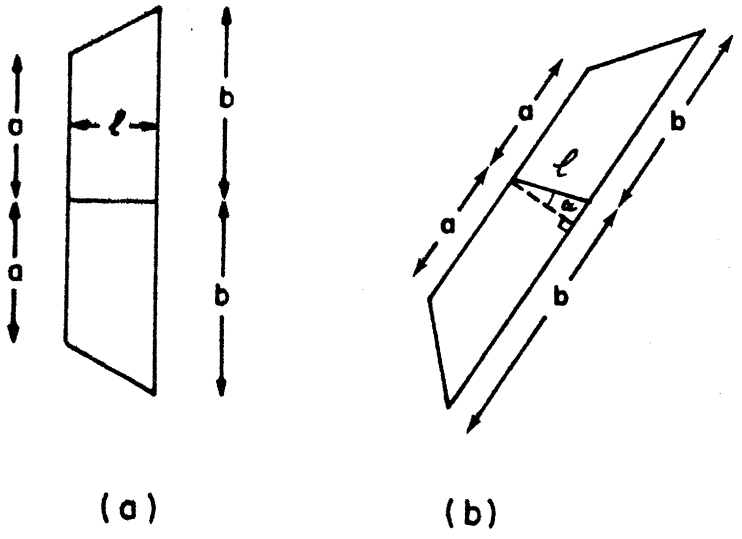

Figure 6
a. A basic region before skewing it. b. The result of skewing the region shown in (a) through angle α.

and

$$\frac{A_i'}{P_i'^2} < \frac{A_i}{P_i^2}. \tag{5.12}$$

We conclude from Eqs. (5.11) and (5.12) that

$$P_i'^2 > P_i^2 \cos\alpha. \tag{5.13}$$

Hence we obtain

$$\sum_{i=1}^{n} P_i'^2 > \cos\alpha \sum_{i=1}^{n} P_i^2 \tag{5.14}$$

$$\sum_{i=1}^{n} A_i' = \cos\alpha \sum_{i=1}^{n} A_i. \tag{5.15}$$

It follows that

$$\frac{\sum_{i=1}^{n} A_i'}{\sum_{i=1}^{n} P_i'^2} < \frac{\sum_{i=1}^{n} A_i}{\sum_{i=1}^{n} P_i^2}, \tag{5.16}$$

and taking the limit as n tends to infinity shows that our extremum principle interprets a skew symmetry as a oriented real symmetry.

6. Strategy and Sensitivity

In Section 3, we suggested that radially asymmetric, elongated, or non-compact, contours most readily elicit three-dimensional interpretations. We then proposed a compactness measure A/P^2, and suggested that perceived surface orientation corresponds to the slant and tilt that maximize the measure. We briefly discussed the role of extremum principles in computer vision (Brady and Horn [1983] present a fuller discussion), and criticized Barrow and Tenenbaum's measure on the grounds that it is highly sensitive to noise. In Sections 4 and 5 we analyzed the extremum principle and showed that it corresponded closely to human perception of oriented planar contours, and that it interprets skew symmetries as oriented real symmetries.

In this Section we return to the discussion in Section 3, concentrating on two additional questions concerning our measure. First, we ask whether there are specific reasons why a visual system should adopt the strategy of extremizing our measure. We suggest that the extremum principle not only determines the orientation of the viewed curve, but provides an estimate of the stability of the interpretation. We relate this to the idea of general viewpoint. Second, we show that the compactness measure is relatively insensitive to noise and to the scale at which the image is sampled. To this end we consider two cases: (i) adding a sinusoid to the contour, (ii) assuming that the image can be modelled in terms of fractals [Mandelbrot, 1982].

Viewing position and stability

Consider viewing a given planar curve from the hemisphere of all possible directions. Consider further the way the image changes when the viewpoint is shifted slightly. A smooth curve will hardly change with a slight change of viewpoint from most viewing directions. We call these viewing directions *stable viewpoints.* Stable viewpoints can be grouped into regions whose boundaries correspond to viewpoints where the contour changes rapidly for a slight change of viewing position. These stable regions will be quite large for images of smooth planar curves and smooth curved surfaces. We suggest that image contours are interpretated as curves that are viewed from stable viewpoints. This is the essence of the *general viewpoint* constraint in computer vision. We are able to estimate how stable a given viewpoint is.

One way to find stable viewing positions is to define a *similarity measure* for viewpoints and then find the extrema of this measure. Sutherland [personal communication] has proposed A/P^2 as a similarity measure, based on empirical studies of animal perception. If this is correct, extremizing A/P^2 corresponds to finding the stable viewing positions of a contour, and, as a result, the stable

interpretations of a contour, defined to be those which observers at general positions are most likely to see. This suggests that assuming general viewing position corresponds to assuming interpretations around which our compactness measure is extremized and where it varies slowly.

We illustrate this idea by considering an ellipse. For an ellipse the stable viewing position corresponds to interpreting the ellipse as an oriented circle. Even if the circle is slanted by as much as thirty degrees, the value of A/P^2 hardly changes, as our algorithm demonstrates, and so the circle is a stable interpretation. If the algorithm is applied to an ellipse with large eccentricity, corresponding to large slant, A/P^2 changes rapidly for slight changes to the viewing position. So if we look at an ellipse from all possible viewing angles it is most likely to appear as a circle and so interpreting an ellipse as a slanted circle is a good strategy.

These results are preliminary, and they will be rigorously developed in a further paper.

Sensitivity of the compactness measure to noise

The second question concerns the sensitivity of the measure to noise and to the scale at which an image is sampled. Since our measure involves fewer derivatives of the contour than measures such as that proposed by Barrow and Tenenbaum, it should be relatively insensitive to noise. Although our present algorithm and the sampling method are essentially equally sensitive, it is possible to develop a less sensitive algorithm to implement our measure. The sampling method inevitably involves calculating tangents, however, and hence is inherently sensitive to noise and scale. For example if the image contour is continuous, but not differentiable, the sampling method cannot (strictly) be applied. Similarly if the orientation of the tangent to the contour has a large high frequency component, the sampling approach will be very sensitive to noise.

Consider the sensitivity to noise of the extremum principle. Clearly, the perimeter P depends on the scale at which the image is viewed, while the area A is much less dependent. It follows that the ratio A/P^2 varies with scale. If a sinusoid is added to a smooth contour, for example, the area will remain approximately the same but the perimeter will change significantly. This does not imply, however, that *the extrema* of A/P^2 vary with scale. For example, it is easy to show that if a sinusoid is added to the contour, the *extremum* of the area over the perimeter squared is effectively unchanged.

Similarly, suppose that the image can be approximated by a fractal [Mandelbrot, 1982]. If we measure the image at a small fractal scale constant l, the perimeter will be given by

$$P(l) = Fl^{(1-d)}. \tag{6.1}$$

where F is independent of l, and $d \geq 1$ is the fractal dimension. Provided l is small, the value of A is essentially independent of l, since the fractal tends to a limit contour, with finite area, as l goes to zero. Now we extremize A/P^2 for all contours which can be projected into the image contour. This is equivalent to extremizing $\log(A/P^2)$ which, using (6.1), we write as

$$\log(\frac{A}{P^2}) = \log A - 2\log F - 2(1-d)\log l. \tag{6.2}$$

We now extremize this over σ and τ (Eqs. 4.24 and 4.25), and note that the term involving l is independent of σ and τ and disappear. It follows that the extrema of A/P^2 are independent of the scale of viewing for an image that can be approximated by a fractal.

7. Interpreting image contours as curved surfaces

Figure 3 shows a number of contours that are interpreted as curved surfaces. In this section we discuss one method for extending our extremum principle to this general case. The key observation, as it was for Witkin [1981], is that our method can be applied locally. To do this, we assume that the surface is locally planar. At the surface boundary, corresponding to the deprojection of the image contour, the binormal coincides with the surface normal. The idea is to compute a local estimate of the surface normal by the extremum principle described in the previous sections and then to use an algorithm, such as that developed by Terzopoulos [1983], to interpolate the surface orientation in the interior of the surface. The method is closely related to that proposed by Brady and Grimson [1981] for perceiving subjective surfaces.

The main question concerns how to apply the extremum principle locally. We are currently investigating the following approach. Consider the circle of curvature to the space curve that is the deprojection of the contour. One way to define the circle of curvature is as the best fitting local circle through three points (or more if one assumes noisy data and makes a least squares estimate) on the space curve near the point in question. The circle of curvature projects into an ellipse. We compute the best fitting ellipse at each point on the image contour, and compute from it a local estimate of the surface orientation by finding the slant and tilt of the corresponding circle, interpreted as the circle of curvature. It is easy to show that so long as the surface foreshortening is differentiable, this is a good estimator of the circle of curvature and hence of the local surface normal.

It requires 5 parameters to define an arbitrary ellipse. Computing the best fitting ellipse in the general case is a complex nonlinear problem best approached using a numerical descent method, though several algorithms have

been published recently for computing best fitting ellipses [Bookstein 1979, Agin 1981, Nakagawa and Rosenfeld 1979, Sampson 1982]. If we assume that the normal to the space curve is not significantly foreshortened, we can compute the ellipse center and major axis from the curvature of the contour at the point in question. The slope of the contour at that point also defines the orientation of the ellipse, leaving a much simpler one-parameter problem.

We note that perceptually the strongest local cues to surface orientation correspond to points of high curvature. This is consistent with our method of locally estimating surface orientation by fitting local ellipses. Recall that our compactness measure was inspired by the observation, for example on ellipses, that large slant is an effective cue to surface orientation and in the case of an ellipse this produces points of high curvature. In fact, we can show that numerical conditioning of the estimator of slant increases monotonically with the slant. Conversely, for straight line portions of a contour, the curvature is zero, and the surface is locally planar. Hence surface orientation does not change along the length of the straight portion.

One issue that remains to be studied is the interface to the surface reconstruction algorithm. Consider the image of a triangle shown in Figure 1. There are, in general, three different perceived orientations of the triangle corresponding to propagating the interpretation of each of the (high curvature) corners. Adjacent corners give inconsistent information, and so it seems necessary to use a labelling approach such as that proposed by Zucker, Hummel, and Rosenfeld [1977].

7. Related work

Gibson [1950] has argued that surface orientation is directly determined by certain "higher order variables" in the proximal stimulus array. What exactly constitutes a "higher order variable" has been the subject of extensive debate. Nevertheless, Gibson and his followers have proposed several such, especially for optical flow and texture gradients. Flock [1964, Flock et. al. 1967] have adopted a Gibsonian perspective on the judgement of slant in texture gradients. He has introduced a "higher order variable" *optical slant* (actually "optical theta" in the original) as a possible basis for monocular slant perception [Flock 1964, Eq. 5] or, at least, as a discriminant for planarity. Flock's work implicitly assumes planarity [Flock 1964, p. 381], and it assumes that the tilt direction has been computed previously. In fact, there are many assumptions in Flock's paper (Freeman [1965, p. 502] counts 13). Moreover, several studies [Clark et. al. 1956, Gruber and Clark 1956, Freeman 1965, Braunstein and Payne 1969] have shown that the determination of surface orientation from texture gradients, for example using optical slant or as in Ikeuchi [1983], is less effective than its determination from bounding contour.

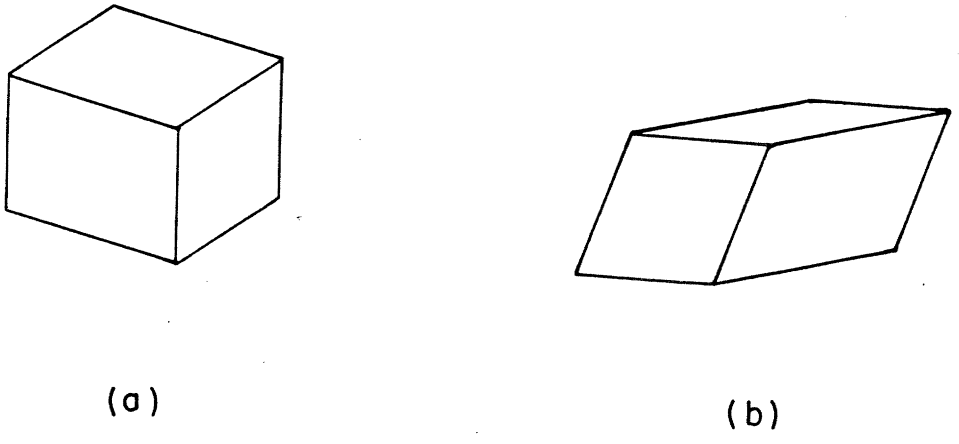

Figure 7
a. A box-like figure typical of those studied by Attneave and Frost [1969]. b. A sheared box.

Attneave [1972, see also Attneave and Frost 1969] describes an approach to determining surface orientation that has similarities to that developed here. First, he argues for a Praegnanz theory of perception, which stresses economy principles in perception. This has meant different things to different researchers. The Gestalt psychologists explicitly noted the link between Praegnanz and minimization principles (for example the soap bubble) in mathematical physics. Attneave [1972, p. 285] suggests that such minima may be computed by "hill--climbing techniques". The extremum principles discussed in Section 3 of this paper and surveyed in [Brady and Horn 1983, Grimson 1981, 1982, 1983, and Terzopoulos 1983] can be considered to be more sophisticated formalizations of similar intuitions.

In fact, Attneave adopts Hochberg's formulation of the Praegnanz theory as the tendency to keep differences to a minimum. In particular he considers three-dimensional interpretations that equalize one or more of angles, lengths of edges, and surface slopes in figures such as Figure 7a. The more of these that are in fact equalized in a particular three-dimensional interpretation of an image, the more likely that interpretation is to be chosen. In fact, the extremum principle developed in this paper will interpret Figure 7a correctly. It will also determine the shear in Figure 7b.

Second, Attneave [1972, Fig. 4] considers the judged orientation of rhombus figures (see Appendix D below for analysis of this case by our extremum

principle.) Surface orientation can be judged reliably for such shapes, more reliably in fact than for the box figures shown in Figure 7, so long as the symmetries of the rhombus do not align with the horizontal and vertical. It is well known that the horizontal and vertical are important for shape description. (see for example Brady [1983] for discussion). Note that surface slant is consistently underestimated.

In order to express depth constraints, Barnard [1982] and Ikeuchi [1983] have proposed alternative projections to orthographic projection used in the development in Sections 3 and 4. The methods we have developed extend straightforwardly to these alternative projections.

Barnard [1982] finds surface orientation from vanishing points of families of parallel edges using central or perspective projection. He considers the projection of an angle, and, after the fashion of Hochberg and Attneave, heuristically assumes equi-angularity to arrive at a three-dimensional interpretation of a triangle. As noted in Appendix C, our method dispenses with the need for such a heuristic assumption. Similar remarks apply to Barnard's study of curvature.

Ikeuchi [1983] proposes a method for determining surface orientation for a surface that is covered with uniformly repeating texture elements. He assumes that the surface element corresponding to the texture element is planar. Most critically it is assumed that the shape of the texture element is known *a priori*. For a particular slant and tilt, the orientation of the projection of the known figure changes. Ikeuchi proposes a measure that is superficially similar to the symmetry measure in Section 4 to determine the slant and tilt of a texture element.

Olson [1974] has studied a version of Ames' trapezoidal window illusion. Surface slant is judged more accurately when the stimulus is moving consistent with the static three-dimensional interpretation than in the purely static case. Similarly, Wallach, Weisz, and Adams [1956] have shown that when an ellipse is rotated in the image plane about an axis passing through its center and normal to the image it is perceived as a spinning oriented circle, like a settling spinning penny. The instantaneous surface orientation of the oriented circle can be computed by our method. Hildreth's forthcoming PhD thesis on the perception of motion discusses the rotating ellipse and proposed a theoretical explanation of human perception of it.

Appendix A: The maximum likelihood method applied to an ellipse

Equation (4.30) gives the value F of the logarithm of the Maximum Likelihood estimator for surface orientation:

$$F = \log\cos\sigma \oint dr - \oint \log(\cos^2\sigma + \sin^2\sigma\cos^2(\alpha - \tau))dr. \qquad (A.1)$$

In this Appendix we investigate the slant computed by the Maximum Likelihood method for the ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1. \qquad (A.2)$$

It is convenient to use the standard parameterization,

$$\begin{aligned} x &= a\cos\theta \\ y &= b\sin\theta, \qquad 0 \leq \theta < 2\pi \\ dr &= (a^2\sin^2\theta + b^2\cos^2\theta)^{\frac{1}{2}}d\theta \end{aligned} \qquad (A.3)$$

Pending a more thorough analysis, we *assume* that the tilt $\tau = \pi/2$ is computed correctly by the Maximum Likelihood method, and restrict attention to slant. Witkin [1981, Figure 5] and our own computational experiments suggests that this is reasonable; in any case, if it is not so, the inequivalence of the Maximum Likelihood and Extremum Principle methods would be assured. Without loss of generality, we assume that $a > b$, and denote a/b by $\lambda > 1$.

Since $\tau = \pi/2$,

$$\begin{aligned} \cos^2\alpha &= \frac{1}{1 + \tan^2\alpha} \\ &= \frac{a^2\sin^2\theta}{a^2\sin^2\theta + b^2\cos^2\theta}, \end{aligned} \qquad (A.4)$$

since $\tan\alpha = dy/dx$. Also,

$$\sin(\alpha - \tau) = \sin(\alpha - \frac{\pi}{2}) = -\cos\alpha, \qquad (A.5)$$

and so Eq. (A.1) reduces to

$$F = \log\cos\sigma \oint dr - \oint \log(1 - \cos^2\alpha\sin^2\sigma)dr. \qquad (A.6)$$

We are interested in extrema of F, and so we consider $dF/d\sigma$, which we write

$$\cot\sigma\frac{dF}{d\sigma} = \oint \frac{a^2\sin^2\theta\cos^2\sigma - b^2\cos^2\theta}{a^2\sin^2\theta\cos^2\sigma + b^2\cos^2\theta}dr, \qquad 1 \leq \lambda \qquad (A.7)$$

Suppose that the Maximum Likelihood method is extremized at $\sigma = \sigma'$. The slant that is necessary to interpret the ellipse as a circle is given by $\cos\sigma = 1/\lambda$. The Maximum Likelihood method overestimates slant if and only if $\sigma' > \cos^{-1}(1/\lambda)$, which is if and only if $\cos\sigma' < 1/\lambda$. This is true if $dF/d\sigma \neq 0$ for all σ such that $1 \leq \lambda\cos\sigma < \lambda$. Denote $\lambda\cos\sigma$ by μ, so that $1 \leq \mu < \lambda$. Substituting in Eq. (A.7) and changing the limits of integration, we find

$$\frac{dF}{d\sigma} = 2b\int_0^{\pi} \frac{\mu^2\sin^2\theta - \cos^2\theta}{\mu^2\sin^2\theta + \cos^2\theta}(\lambda^2\sin^2\theta + \cos^2\theta)^{\frac{1}{2}}d\theta \qquad (A.8)$$

We now split the range of integration into four equal intervals of size $\pi/4$. With suitable changes of variables to bring the intervals of integration to $[0, \frac{\pi}{4}]$, we find

$$\begin{aligned}\frac{dF}{d\sigma} = 4\int_0^{\frac{\pi}{4}} & \frac{\mu^2\cos^2\theta - \sin^2\theta}{\mu^2\cos^2\theta + \sin^2\theta}(\lambda^2\cos^2\theta + \sin^2\theta)^{\frac{1}{2}} \\ & - \frac{\mu^2\sin^2\theta - \cos^2\theta}{\mu^2\sin^2\theta + \cos^2\theta}(\lambda^2\sin^2\theta + \cos^2\theta)^{\frac{1}{2}}d\theta\end{aligned} \qquad (A.9)$$

Over the range of integration, $\cos\theta > \sin\theta$. Algebraic manipulation of Eq. (A.9) shows it to be greater than zero, so that it is certainly not zero. Hence the Maximum Likelihood method overestimates the slant.

Appendix B: Why square curvature is inappropriate

In section 3 we rejected the measure

$$F = \int \kappa^2 ds \tag{B.1}$$

because, despite popular belief to the contrary, it does not interpret an ellipse as a circle. In this appendix we substantiate that claim.

Consider the ellipse defined by Eqs. (A.2) and (A.3). The curvature κ of the ellipse is given by

$$\kappa^2 = \left| \frac{d^2\mathbf{r}}{ds^2} \right|^2 \tag{B.2}$$

$$= \frac{a^2 b^2}{(a^2 \sin^2\theta + b^2 \cos^2\theta)^3} \tag{B.3}$$

where s is the arc length. Substituting Equation (B.3) into (B.1) we have

$$F(a,b) = \int_0^{2\pi} \frac{a^2 b^2}{(a^2 \sin^2\theta + b^2 \cos^2\theta)^{\frac{5}{2}}} d\theta \tag{B.4}$$

Suppose, without loss of generality, that the ellipse is in the image plane and that $a > b$. Suppose further that the square curvature performance index (B.1) is correct and interprets the ellipse as a circle lying in a plane that is slanted with angle σ to the y-axis, where

$$b = a \cos\sigma \tag{B.5}$$

Equivalently, if we set $b' = b/\cos\sigma$, the measure will choose σ such that $b' = a$. Hence if we write $b = \lambda a$, and consider $F(a,b)$ to be a function of λ, the measure should be extremized by $\lambda = 1$. We will now show that this is *not* the case.

$$F(\lambda) = \frac{1}{a} \int_0^{2\pi} \frac{\lambda^2}{(\sin^2\theta + \lambda^2 \cos^2\theta)^{\frac{5}{2}}} d\theta \tag{B.6}$$

Differentiating this with respect to λ gives

$$\frac{\partial F}{\partial \lambda} = \frac{1}{a} \int_0^{2\pi} \frac{2\lambda \sin^2\theta - 3\lambda^3 \cos^2\theta}{(\sin^2\theta + \lambda^2 \cos^2\theta)^{\frac{7}{2}}} d\theta \tag{B.7}$$

Evaluating this expression at $\lambda = 1$ gives

$$\begin{aligned} \left.\frac{\partial F}{\partial \lambda}\right|_{\lambda=1} &= \frac{1}{a}\int_0^{2\pi} (2\sin^2\theta - 3\cos^2\theta)d\theta \\ &= -\frac{\pi}{a} \end{aligned} \tag{B.8}$$

Since the partial of F with respect to λ does not vanish when λ is equal to one, the circle does not extremize the ellipse for the square curvature measure.

Appendix C: The interpretation of some simple shapes

In this appendix we show that our method correctly interprets a number of simple shapes, namely an ellipse, a parallelogram, and a triangle.

Ellipse

Suppose the ellipse is given by equations (A.2) and (A.3). It is easy to show that the area A and perimeter P are given by

$$\begin{aligned} A &= \pi ab \\ P &= \int_0^{2\pi} (a^2 \sin^2\theta + b^2 \cos^2\theta)^{\frac{1}{2}} d\theta \end{aligned} \qquad (C.1)$$

Maximizing A/P^2 is equivalent to minimizing $P/\sqrt{A}$. We set $b = \lambda a$, (λ is the eccentricity of the ellipse,) and define $P/\sqrt{A} = f(\lambda)$. We find

$$f(\lambda) = \frac{1}{\sqrt{\pi}} \int_0^{2\pi} (\frac{\sin^2\theta}{\lambda} + \lambda \cos^2\theta)^{\frac{1}{2}} d\theta \qquad (C.2)$$

By the same argument presented in appendix B, the ellipse will be interpreted as a circle provided $\lambda = 1$ is a minimum of $f(\lambda)$. Changing the variable of the integral to $\phi = \theta + \frac{\pi}{2}$ we find that

$$f(\lambda) = f\left(\frac{1}{\lambda}\right) \qquad (C.3)$$

which implies that $\lambda = 1$ extremizes $f(\lambda)$. It is also clear that extrema occur in pairs of the form $\lambda, 1/\lambda$. Furthermore, both are stationary points or one is a maximum and the other a minimum. Observe that $f(\lambda)$ tends to infinity as λ tends either to zero or infinity. It follows that a sufficient condition for $\lambda = 1$ to be a global minimum is that all pairs of extrema be stationary points. Suppose that this is *not* the case, and let λ_0 be the smallest extremum that is not a stationary point. Since $f(\lambda)$ tends to infinity as λ tends to zero, λ_0 must be a minimum and $1/\lambda_0$ a maximum. But $1/\lambda_0$ is the largest non-stationary extremum, and so, by the same argument as above, it must be a minimum. This contradiction establishes the result.

Parallelogram and triangle

In section 5 we showed that a skew symmetry is always interpreted as an oriented symmetry by our method. In particular, a parallelogram is interpreted as a rectangle. By the same argument, a rectangle is a skewed symmetry of a square. Hence our method interprets a parallelogram as a square.

Similar reasoning shows that a triangle is interpreted as a skewed isosceles triangle, which is interpreted as a skewed equilateral triangle. The axes of the skewed symmetry join a vertex to the midpoint of the opposite side. Hence our method interprets a triangle as an oriented equilateral triangle.

Appendix D: The slant and tilt of a skewed symmetry

In this appendix we calculate the slant σ and the tilt τ that correspond to the oriented real symmetry that is the interpretation of a skewed symmetry whose skew angle is δ.

As a simple, though instructive, example, consider a rhombus of side a and included angle γ (Figure 8). To find the extremizing tilt, we substitute the data from Figure 8 into Eq. 4.24, and find

$$\frac{\sin\tau\cos\tau}{\left\{1-\sin^2\sigma\sin^2\tau\right\}^{\frac{1}{2}}}+\frac{\sin(\tau-\gamma)\cos(\tau-\gamma)}{\left\{1-\sin^2\sigma\sin^2(\tau-\gamma)\right\}^{\frac{1}{2}}}=0 \qquad (D.1)$$

We can rewrite this in the form

$$\left\{\cos^2(\tau-\gamma)-\cos^2\tau\right\}\ \left\{\cos^2(\tau-\gamma)+\cos^2\tau + (1-\cos^2\sigma)(1-\cos^2(\tau-\gamma))(1-\cos^2\tau)-1\right\} = 0 \qquad (D.2)$$

We assume first that the first factor is zero. It follows that

$$\cos(\tau-\gamma)=\pm\cos\tau \qquad (D.3)$$

and so

$$\gamma-2\tau=n\pi. \qquad (D.4)$$

Since $0\leq\tau\leq\pi$ and $0<\gamma<\pi$, there are two possible solutions, namely

$$\tau=\frac{\gamma}{2} \qquad (D.5)$$

$$\tau=\frac{\gamma+\pi}{2} \qquad (D.6)$$

Observe that the tilt direction is one of the axes of symmetry of the rhombus shown in Figure 8.

Having solved for the tilt, we now solve for the extremizing slant σ using Eq. 4.23. Recalling from Eq. D.4 that $\sin^2(\tau-\gamma)=\sin^2\tau$, we find upon substitution into Eq. 4.23

$$\frac{1}{2}\left\{1-\sin^2\sigma\sin^2\tau\right\}^{\frac{1}{2}}=\cos^2\tau\left\{1-\sin^2\sigma\sin^2\tau\right\}^{-\frac{1}{2}} \qquad (D.7)$$

which we solve to get

$$\cos^2\sigma = \frac{\cos^2\tau}{\sin^2\tau}. \qquad (D.8)$$

The requirement that $|\cos\sigma| < 1$ picks out either Equation (D.5) or (D.6), so we get a unique solution. In this case the skew angle is given by

$$\delta = \frac{\pi}{2} - \gamma. \qquad (D.9)$$

To summarize, if $\delta < 0$ we get

$$\begin{aligned} \tau &= \frac{\pi}{4} - \frac{\delta}{2} \\ \cos\sigma &= \frac{1+\sin\delta}{\cos\delta}, \end{aligned} \qquad (D.10)$$

and if $\delta > 0$,

$$\begin{aligned} \tau &= \frac{3\pi}{4} - \frac{\delta}{2} \\ \cos\sigma &= \frac{\cos\delta}{1+\sin\delta}, \end{aligned} \qquad (D.11)$$

These formulae were derived for an equal-sided parallelogram but they will clearly apply to the more general case and a rotation through the angles given by Equations (D.10) and (D.11) will unskew any symmetry. It should be noted that τ is taken to be zero on the axis of symmetry, as in Figure 8.

To conclude this Appendix, we consider the case that the second factor in Eq. (D.2) is zero and the first factor non-zero. We introduce the angle ψ by analogy with Eq. (2.1) defining orthographic projection:

$$\tan^2\psi = \cos^2\sigma\tan^2(\gamma-\tau) \qquad (D.12)$$

Then

$$\begin{aligned} \sin^2(\gamma-\tau) &= \frac{\tan^2\psi}{\cos^2\sigma+\tan^2\psi} \\ \cos^2(\gamma-\tau) &= \frac{\cos^2\sigma}{\cos^2\sigma+\tan^2\psi} \end{aligned} \qquad (D.13)$$

where, without loss of generality, we suppose $0 < \psi < \pi/2$. Now by assumption, the second factor in Eq. (D.2) is zero:

$$\cos^2\tau + \cos^2(\gamma-\tau) + \sin^2\sigma\sin^2\tau\sin^2(\gamma-\tau) - 1 = 0 \qquad (D.14)$$

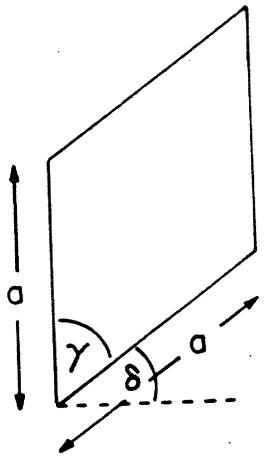

Figure 8
A typical skew symmetric figure, namely a rhombus of side a and included angle γ, with skew angle $\delta = \frac{\pi}{2} - \gamma$. The tilt direction is aligned with one of the axes of symmetry of the rhombus, determined by the angle δ.

Using Eq. (D.13), we find

$$\tan^2\tau\tan^2\psi = 1,$$

from which we deduce

$$\begin{aligned} \sin\tau &= \cos\psi \\ \cos\tau &= \mu\sin\psi, \qquad \mu = \pm 1, \end{aligned} \tag{D.15}$$

since the ranges of the variables can be assumed to be $0 < \sigma < \pi/2; 0 < \tau < \pi; 0 < \psi < \pi/2; 0 < \gamma < \pi$. From Eq. (D.13) we find that

$$\sin(\gamma-\tau)\cos(\gamma-\tau) = \frac{\nu\tan\psi\cos\sigma}{\cos^2\sigma + \tan^2\psi}, \qquad \nu = \pm 1, \tag{D.16}$$

and, since $\sin\tau\cos\tau$ has the same sign as $\sin(\gamma-\tau)\cos(\gamma-\tau)$, $\mu = \nu$. From Eq. (D.1) we deduce

$$\begin{aligned} \frac{\sin\tau\cos\tau}{1-\sin^2\sigma\sin^2\tau^{\frac{1}{2}}} &= \frac{-\sin(\gamma-\tau)\cos(\gamma-\tau)}{1-\sin^2\sigma\sin^2(\gamma-\tau)^{\frac{1}{2}}} \\ &= \frac{\nu\sin\psi\cos\psi}{1-\sin^2\sigma\cos^2\psi^{\frac{1}{2}}} \end{aligned} \tag{D.17}$$

So far in this Appendix we have only used one of the constraints derived in Section 4, namely Eq. (4.24) (from which we derived Eq. (D.1)). We now use the second constraint Eq. (4.23), which we can write in the form

$$\begin{aligned}\frac{2\cos^2\tau}{(1-\sin^2\sigma\sin^2\tau)^{\frac{1}{2}}} &- (1-\sin^2\sigma\, 2\tau)^{\frac{1}{2}} \\ &+ \frac{2\cos^2(\gamma-\tau)}{(1-\sin^2\sigma\sin^2(\gamma-\tau))^{\frac{1}{2}}} \\ &- (1-\sin^2\sigma\sin^2(\gamma-\tau))^{\frac{1}{2}} = 0\end{aligned} \tag{D.18}$$

After some algebraic manipulation, we deduce

$$\cos\sigma = -\tan^2\psi,$$

from which it follows that $\cos\sigma$ is negative, which is impossible in the range under consideration.

Acknowledgements

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's Artificial Intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643, the Office of Naval Research under contract number N0014-80-C-0505, and the System Development Foundation. The authors thank Ruzena Bajcsy, Chris Brown, John Canny, Eric Grimson, Ellen Hildreth, Tommy Poggio, Demetri Terzopoulos, and Andy Witkin for their comments.

References

Agin, G. J. "Computer vision systems for industrial inspection and assembly," *Computer* **13** (1980), 11-20.

Agin, G. J. "Fitting ellipses and general second-order curves," The Robotics Institute, Carnegie Mellon University, CMU-RI-TR-81-5, 1981.

Attneave, Fred. "Representation of physical space," *Coding processes in human memory*, Melton, A. W., and Martin, F., eds., John Wiley, New York, 1972, 283 - 306.

Attneave, Fred, and Frost, R. "The determination of perceived tridimensional orientation by minimum criteria," *Perception and Psych.* **6** (1969), 391 –396.

Ballard, D. H., and Brown, C. M. *Computer vision*, Prentice-Hall, New Jersey, 1982.

Barnard, Steven. "Interpreting perspective images," SRI International TR 271, 1983.

Barrow H. G. and Tenenbaum J. M.. "Interpreting line drawings as three dimensional surfaces," *Artificial Intelligence* **17** (1981), 75-117.

Bookstein, Fred L. "Fitting conic sections to scattered data," *Computer Graphics and Image Processing* **9** (1979), 56-71.

Brady, Michael. "Computational approaches to image understanding," *Computing Surveys* **14** (1982), 3-71.

Brady, Michael. "Criteria for representations of shape," *Human and Machine Vision*, eds. J. Beck and A. Rosenfeld, Academic Press, 1983, .

Brady, Michael and Grimson, W. E. L. "The perception of subjective surfaces," MIT, AI Memo 666, 1981.

Brady, Michael, and Horn, B. K. P. "Rotationally symmetric operators for surface interpolation," *Computer Graphics and Image Processing* **22 (1)** (1983), 70 - 95.

Braunstein, Myron L., and Payne, John W. "Perspective and form ratio as

determinants of relative slant judgements," *J. Exp. Psychology* **3** (1969), 584-590.

Clark, W. C., Smith, A. H., and Rabe, Ausma. "The interaction of surface texture, outline gradient, and ground in the perception of slant," *Canad. J. Psych.* **10** (1956), 1 - 8.

Davis, L. S., Janos, Ludwig, and Dunn Stanley. "Efficient recovery of shape from texture," Computer Vision Laboratory, University of Maryland, TR-1133, 1982.

Flock, Howard R. "A possible optical basis for monocular slant perception," *Psych. Review* **71** (1964), 380 – 391.

Flock, Howard R., Graves David, Tenney James, and Stephenson Bruce. "Slant judgements of single rectangles at a slant," *Psychon. Sci.* 7 (1965), 57 – 58.

Freeman Jr., Robert B. "Ecological optics and visual slant," *Psych. Review* **72** (1965), 501 – 504.

Freeman Jr., Robert B. "Effect of size on visual slant," *J. Exptl. Psych.* **71** (1966), 96 – 103.

Gibson, J. J. *Perception of the visible world*, Houghton Miflin, Boston, 1950.

Gregory, Richard L. *Eye and Brain, 2nd ed.*, McGraw-Hill, New York, 1973.

Grimson, W. E. L. *From images to surfaces: a computational study of the human early visual system*, MIT Press, Cambridge, 1981.

Grimson, W. E. L. "A computational theory of visual surface interpolation," *Phil. Trans. Roy. Soc. Lond B* **298** (1982), 395-427.

Grimson, W. E. L. "An implementation of a computational theory of visual surface interpolation," *Comp. Vision Graphics, and Image Processing* **22 (1)** (1983), 39 - 70.

Gruber, H. E., and Clark, W. C.. "Perception of slanted surfaces," *Percep. and Motor Skills* **6** (1956), 97-106.

Horn, B. K. P. "Understanding image intensities," *Artificial Intelligence* **8** (1977), 201-231.

Horn, B. K. P. "The curve of least energy," MIT, AIM-610, 1981.

Ikeuchi, Katsushi. "Shape from regular texture," *Artificial Intelligence* **(to appear)** (1983), .

Kanade, T. "Recovery of three-dimensional shape of an object from a single view," *Artificial Intelligence* **17** (1981), 409-460.

Marr D.. *Vision*, Freeman, San Francisco, 1982.

Nakagawa, Yasuo, and Rosenfeld, Azriel. "A note on polygonal and elliptical approximation of mechanical parts," *Pattern Recognition* **11** (1979), 133 - 142.

Olson, R. K. "Slant judgements from static and rotating trapezoids correspond to rules of perspective geometry," *Perception and Psych.* **15** (1974), 509 – 516.

Pavlidis, T. *Structural Pattern Recognition*, Springer, New York, 1977.

Sampson, Paul, D. "Fitting conic sections to very scattered data: an iterative refinement of the Bookstein algorithm," *Computer Graphics and Image Processing* **18** (1982), 97 - 108.

Stevens, K. A. Surface perception from local analysis of texture (also, AI-TR--512), MIT, 1980.

Terzopoulos, D. "Multi-level reconstruction of visual surfaces," *Computer Graphics and Image Processing* (1983).

Wallach, Hans, Weisz, Alexander, and Adams, Pauline Austin. "Circles and derived figures in rotation," *Am. J. Psychology* **69** (1956), 48 – 59.

Witkin, Andrew P. Shape from contour, PhD Thesis, MIT, Artificial Intelligence Laboratory, also: AIM-TR-589, 1980.

Witkin, Andrew P. "Recovering surface shape and orientation from texture," *Artificial Intelligence* **17** (1981), 17-47.

Yuille, Alan. "Zero crossings on lines of curvature," MIT Artificial Intelligence Laboratory, MIT-AIM 718, 1983.

Zucker, S. W., Hummel, R. A., and Rosenfeld A. "An application of relaxation labelling to line and curve enhancement," *IEEE Trans. Comput.* **C-26** (1977), 394-403, 922-929.