Psychological Review
2002, Vol. 109, No. 1, 91–115

0033-295X/02/$5.00 DOI: 10.1037//0033-295X.109.1.91

A Contour Propagation Approach to Surface Filling-In and Volume Formation

Peter Ulric Tse
Max Planck Institute for Biological Cybernetics

A new approach to surface and volume formation is introduced in response to the question, "Why do some silhouettes look 3 dimensional (3D) and others look 2D?" The central idea is that form information can propagate away from a "propagable segment" (PS) of occluding contour that could have projected onto the image from the visible portion of a cross-section of a surface. A key property of a PS is that it exhibits abrupt curvature changes where it meets the rest of the occluding contour. An algorithm is described for filling in curved surfaces from a PS: When copies of a PS are propagated into the interior, they act as cross-sectional surface contours that also exhibit abrupt curvature changes with the rest of the occluding contour. The result is a nonmetric coding of 3D-shape in terms of local ordinal surface curvature and orientation relationships that is scale, translation, and rotation invariant.

Recovering three-dimensional (3D) form from the inherently ambiguous 2D retinal image is perhaps the most fundamental problem faced by the visual system. To solve this problem, multiple shape-from-*x* systems have evolved to recover 3D shape from the various cues to form found in the image. Thus, under normal viewing conditions, the problem of recovering 3D shape from the 2D image can be solved using multiple mutually constraining cues. In some cases, however, just a single cue can be used to generate a distinct percept of 3D form. Silhouettes are an example of a single-cue stimulus because the only information available to generate 3D form from a silhouette lies in the shape of its contour. For this reason, silhouettes are perhaps better than any other class of stimuli for isolating and modeling properties of the shape-from-contour system (e.g., Kimia, Tannenbaum, & Zucker, 1995; Richards & Hoffman, 1985; Zhu & Yuille, 1996). If we can understand how the visual system generates curved surfaces from the contours of silhouettes, we may be in a better position to understand how shape is coded for stimuli in which multiple cues are present.

Interestingly, some silhouettes, like those in Figure 1a, look 3D, whereas others, like those in Figure 1b, look flat. Silhouettes that look 3D can look like volumes (i.e., closed surfaces plus the space that they enclose; Albert & Tse, 2000; Tse, 1998, 1999a, 1999b, 1999c; Tse & Albert, 1998) or open curved surfaces that do not enclose space. How does the visual system generate a representation of a curved surface for the interior of a silhouette in which there is no explicit form information available? Although there are no local cues to surface orientation in the interior of a silhouette, most observers perceive Point A on the large "bump" in Figure 1a to have less surface *slant* than Point B. (Note: terms in italics are defined in Appendix B.) What contour relations does the visual system use to generate variations in perceived surface orientation for regions far from the contour? The visual system must solve this form recovery problem by knowing which aspects of the shape of the 2D image contour are most informative about 3D shape in the world and how local shape information constrains the construction of surfaces.

To most observers, the ellipse in Figure 1b tends to look like a flat disk on a ground plane, whereas the similar silhouette (constructed by joining two half-ellipses of different aspect ratio along their common major axis) shown at the bottom of Figure 1a tends to look like a bump lying on a ground plane (Kristjansson & Tse, 2001; Tse & Albert, 1998). Even though the aspect ratios of the two ellipses from which the bump in Figure 1a was constructed differ very little, it looks volumetric, whereas the ellipse in Figure 1b looks flat. It is remarkable that changing the *occluding contour* in this subtle way leads to such a drastic shift in perceived form. Any algorithm that generates 3D form from silhouettes must account for this difference between a bump and an ellipse.

My goal is to understand those aspects of contours that are most informative about shape and to describe an algorithm that the shape-from-contour system may be using to recover 3D shape. I raise the possibility that the visual system may seek out "propagable" segments of occluding contour that could project from segments of *rim* lying on a *planar cut* (Figure 2a) or cross-section of a volume. Once an inference of such a cross-section has been made, this cross-sectional information can be generalized over ambiguous portions of the image to generate the percept of a volume. Although several authors have discussed surface filling-in from the boundary (e.g., Grossberg & Mingolla, 1985; Paradiso & Nakayama, 1991), this has generally been described, at least implicitly, as occurring over flat surfaces. The algorithm described here amounts to a mechanism for filling in curved surfaces from

Peter Ulric Tse, Max Planck Institute for Biological Cybernetics, Tübingen, Germany.

This work was first presented at the Association for Research in Vision and Opthalmology 1998 Annual Meeting and was funded by McDonnell-Pew Grant 98-49CNS and by the Max Planck Society. I thank Patrick Cavanagh, Ken Nakayama, and Nikos Logothetis for their kind support while I was thinking through these ideas and writing them up. Jack Beusmans, Robert Bjork, David Jacobs, James Todd, Johan Wagemans, and Qasim Zaidi made helpful comments and criticisms.

Correspondence concerning this article should be addressed to Peter Ulric Tse, who is now at the Department of Psychological and Brain Sciences, Dartmouth College, Moore Hall, Hanover, New Hampshire 03755. E-mail: peter.tse@dartmouth.edu

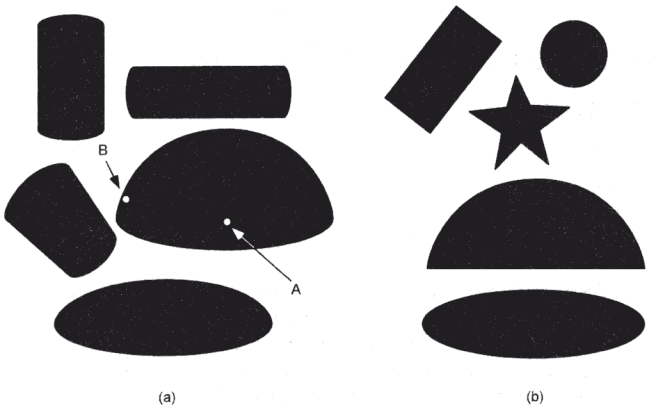


Figure 1. Why do the silhouettes in Figure 1a look 3D and those in 1b look 2D? Note that the surface orientation of the point marked A appears less slanted than that at B.

the boundary. Geometric properties of propagable segments of occluding contour are described. I show that locations that undergo *abrupt curvature changes* along a contour are particularly revealing about shape. Potentially propagable segments of contour lie between two points of abrupt curvature change along a silhouette's contour. Once identified, they can be propagated by multiple affine transformations of the propagable contour segment into the interior of a silhouette. Both a propagable segment and its copies can be interpreted as the visible portions of cross-sections through a volume and can, therefore, be used to recover a volume from a silhouette. Before presenting details of this algorithm, it is useful to offer some theoretical background on the fundamental problem to be solved.

Background to the Contour Propagation Algorithm

I constructed the silhouette shown in Figure 3a to have at least two 3D form interpretations. These are depicted in Figure 3b and c with curves corresponding to a stack of parallel, equidistant planar cuts. Why are the two volumetric interpretations in Figure 3 (b and c) perceived, whereas the infinity of other possible 2D or 3D shapes that could have projected to the same image in Figure 3a are not? The answer that is proposed here is that there are two different segments of occluding contour (shown in Figure 3a) that can be propagated into the interior, and these lead to mutually exclusive 3D form interpretations. They are mutually exclusive in the sense that the 3D form depicted in Figure 3b cannot be superimposed, in a 3D sense, on top of the form depicted in Figure 3c. The problem of how a propagable segment deforms as it propagates over a surface is addressed later. My initial goal is to define an algorithm that can specify which segments of occluding contour are propagable and which are not.

The geometric relationship between *surface contours* and the surfaces from which they project must be constrained if the visual system is to recover 3D surface layout from the 2D image. In other words, the visual system must make assumptions about the relationship between the curvature of contours in the image and the curvature of surface curves in the world if it is to recover 3D form. Otherwise, the problem of shape recovery from contours would be underconstrained and the image open to too many incorrect form interpretations.

Stevens (1981, 1986) and Knill (1992) argued that the visual system is biased to attribute most, if not all, of the curvature of a surface contour to the underlying surface curvature (see also Todd & Reichel, 1990). This means that the visual system assumes that the contour in the image projects from a curve on the surface that has no more curvature than is afforded it by the underlying surface. This point follows from the assumption of nonaccidentality (Barrow & Tenenbaum, 1981; Binford, 1981; Freeman, 1994; Nakayama & Shimojo, 1992). The visual system makes the implicit assumption that the image is not due to one of the accidental object arrangements or one of the few accidental viewpoints from which an object's surface layout is not derivable from its projected contours and other image cues. In other words, it assumes that the information available in the image is sufficient to recover 3D surface layout. A corollary of the assumption of nonaccidentality

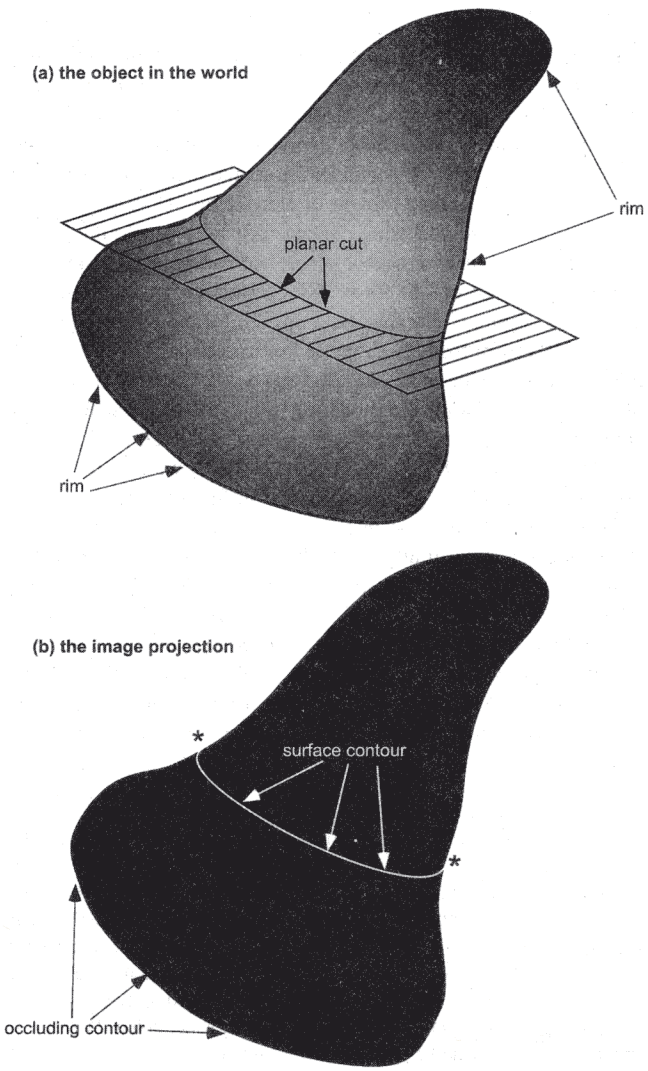


Figure 2. In Figure 2a an object is sliced by a plane. In 2b the surface contour that projects from the planar cut in 2a is shown. The occluding contour in 2b projects from the rim in 2a.

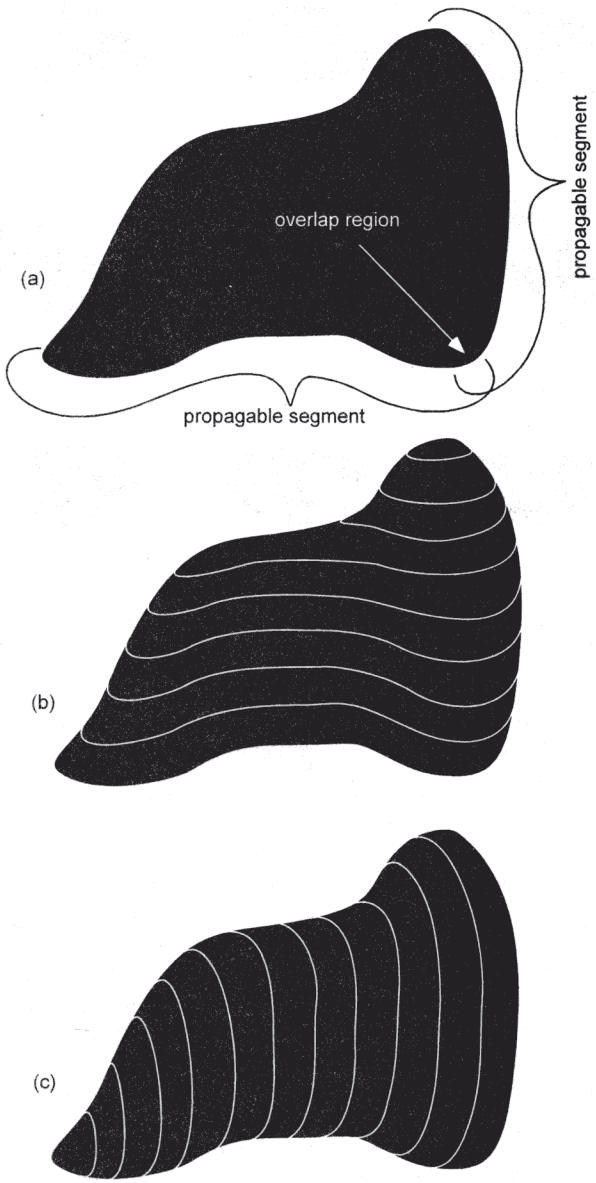


Figure 3. When Figure 3a is seen as a 3D shape, two different shapes are commonly seen. These are shown in 3b and 3c. The different shapes arise because they are generated from the two different propagable segments indicated in 3a.

is that the visual system does not infer undulations of the 3D surface in the absence of evidence for such (Richards, Koenderink, & Hoffman, 1987; Tse, 1998). Conversely, the visual system assumes that the curvature of surface contours in the image is due to the curvature of the underlying projecting surface rather than to the curvature of the curve itself that lies on that curved surface. Of course, there are surface curves that have more curvature than is afforded them by the underlying surface, such as spirals or those that intersect themselves (e.g., see Figure 4a). However, in the absence of image evidence that surface contours have more curvature than is consistent with other cues to surface curvature, it is assumed that no excess curvature is present.

For certain segments of surface contour (described next), the assumption of no excess curvature is equivalent to an assumption that the surface curves that project to those surface contours arise from an intersection of a plane with the curved surface (Tse, 1998). This "planar cut" assumption is weaker than the assumption that a surface contour projects from a line of *principal curvature*, as suggested by Stevens (1981), and is also weaker than the assumption that a surface contour projects from a *geodesic*, as suggested by Knill (1992; Knill, Kersten, & Mamassian, 1996). This is because a surface curve corresponding to a locus of intersection between any plane and a curved surface will have a curvature entirely resulting from the surface curvature of the curved surface, because a plane is not curved at all. In most cases, such curves will conform to neither lines of principal curvature nor geodesic curves. If we consider a single point on a surface, there will be only two lines centered on that point that are lines of principal curvature. It is unnecessarily stringent to require surface contours to project from such lines or sums of lines because a surface curve can pass through a surface point in a nonprincipal direction and yet still have no more curvature than is afforded it by the underlying surface. Similarly, if we consider two points on a surface, there will generally exist only one geodesic path (of shortest distance along the surface) connecting them. Again, it is unnecessarily stringent to require surface contours to project from geodesics because there are other paths that connect those two points, and these paths have no more curvature than is afforded them by the underlying surface. In particular, the set of paths defined by planar cuts that pass through those two points will have no excess curvature. However, planar cuts that slice through a surface at an oblique angle will tend to project to surface contours with relatively higher curvature than those that slice through the surface at a transverse or right angle. Because oblique cuts may introduce more uncertainty into the interpretation of the curvature of a surface contour, there might be a preference for short over long paths defined by planar cuts, in which case the planar cut assumption would approximate a fuzzy version of a geodesic assumption. This would be equivalent to a *prior* on interpreting surface contours as having arisen from planar cuts that pass normally through a surface rather than obliquely.

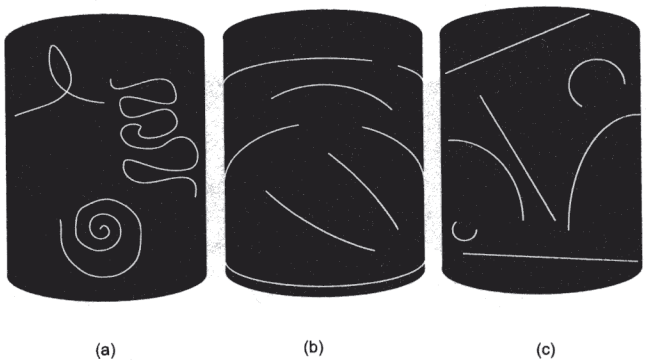


Figure 4. Because the surface contours in Figure 4a have abrupt curvature changes or intersect themselves, they are not assumed to project from a planar cut of the surface. The surface contours in 4b can project from planar cuts of a cylindrical surface. Those in 4c lack abrupt changes in curvature but are inconsistent with occluding contour cues to 3D shape and, therefore, cannot lie on a cylinder's surface.

The visual system may similarly assume that certain segments of occluding contour project from rim segments that lie on a planar cut of the volume that projected the image. Because a given segment of rim on an object would count as a surface contour if viewed from a different viewpoint, this is not very different from what has just been claimed. The important question is what kinds of segments of occluding contour or surface contour are assumed to project from a planar cut. Later it is shown that a segment of occluding contour that lacks abrupt changes in curvature will be taken by the visual system to project from a segment of rim that lies in a plane. In Figure 5a, for example, the segments labeled A, B, C, and D can each be assumed to lie in a plane because of their lack of curvature discontinuities. Indeed, these segments were created by segmenting the boundary at points of curvature discontinuity. In Figure 5b, possible planes have been depicted to make explicit this key idea. The planes are depicted as they might look under perspective projection to enhance the 3D effect, although I do not address the issue of perspective deformations until later.

In contrast to occluding contours, a segment of surface contour that lacks abrupt changes in curvature may or may not be taken by the visual system to project from a mark or boundary on the surface that lies on a planar cut of that surface, depending on its consistency with other cues to 3D form. Thus, the surface contours shown in Figure 4a will be excluded from potentially lying on planar cuts because they possess abrupt changes in curvature or because they intersect themselves. However, the surface contours shown in Figure 4b may be interpreted as planar cuts because they not only have no abrupt changes in curvature, but they are also consistent with other cues to 3D form found in the occluding contour. The surface contours shown in Figure 4c, however, will not be perceived as lying on a planar cut because, although they lack curvature discontinuities, they are not consistent with occluding contour cues to the overall 3D form of the object.[1] These surface contours might, therefore, be interpreted as curves drawn on the surface with excess curvature, as in Figure 4a, or separate objects attached to the surface of the object.

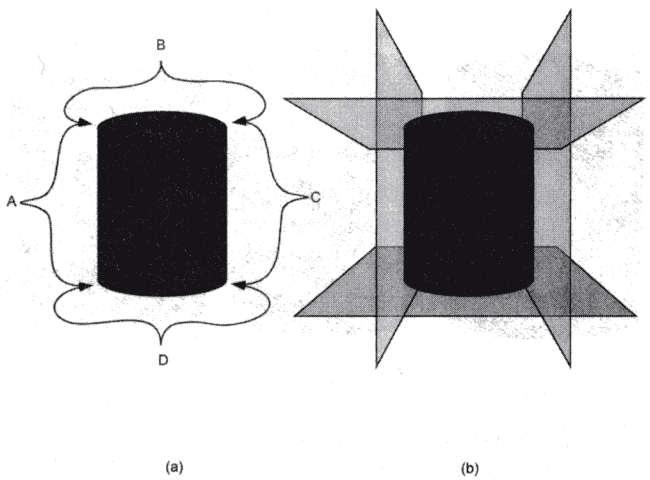


Figure 5. (a) The occluding contour segments A, B, C, and D each lack abrupt curvature changes and, therefore, each project from a rim segment that lies on a common plane. (b) Possible planes.

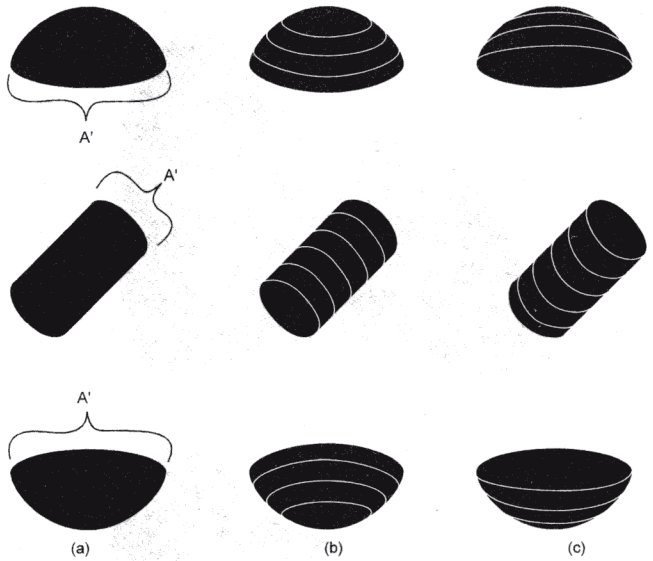


Figure 6. The propagable segments A′ in Figure 6a can generate two different solutions for the silhouettes, as shown in Figure 6b and c.

Geometric Properties of Propagable Contour Segments and the Contour Propagation Algorithm

Occluding contour segments that project from a portion of rim lying on a planar cut have geometric properties that other segments of occluding contour do not. Such segments carry information about the shape of a volume's cross-section that can be used to infer the shape of visible surfaces even in regions that lack local cues to surface curvature. Propagable segments of occluding contour can be used to generate a set of surface contours that can tell us about surface relationships for surface regions that project to the interior of a silhouette. These surface contours are ones that would be visible if planar cuts were made through the object's surface by planes parallel to the plane that contains the rim segment that projects to the propagable segment of occluding contour in the image. The basic idea is shown in Figure 6. In Figure 6a the portion of occluding contour marked "A′" propagates along the length of the silhouette. (The prime indicates that the symbol refers to an entity in the image rather than an entity in the world. Symbols without primes refer to entities in the world such as the rim or a curve lying on a volume's surface. Symbols with primes refer to entities in the image, such as occluding or surface contours.) The two possible propagations given the cross-sectional information available are shown in Figure 6b and c. In the "bump" cases, the cross-sections change only their size, not their shape. In Figure 3b and c, the size and shape of the inferred cross-section change over the course of propagation. In this section, I investigate

[1] Surface contours inconsistent with planar cuts may serve to flatten the appearance of a silhouette if they are interpreted as lying on the surface. Indeed, most observers tend to see Figure 4a as most 3D (i.e., cylindrical) in appearance, and Figure 4c as least. On the other hand, surface contours with abrupt curvature changes can be taken to lie on planar cuts through a volume if occluding contour cues suggest the existence of a corner in the interior of the silhouette.

the geometric properties of propagable segments of occluding contour. I address the problem of size and shape deformation of propagated segments later.

I make several assumptions to simplify the initial geometric analysis and later relax them. I investigate the properties of silhouettes that correspond to volumes that have no surface tangent discontinuities, such as corners, and have no holes (i.e., genus zero) and assume orthographic projection. I also assume that the occluding contour of a silhouette corresponds to a single unbroken loop of rim. (Of course, this rim loop need not lie on a plane. Later I consider more natural volumes that have rims composed of discontinuous curves and loops). Therefore, I do not consider silhouettes projected from volumes whose visible surfaces surround regions of self-occluded surface. Last, I assume a nonaccidental viewpoint.

The outline of a silhouette is a closed loop of contour in the image. If this loop is cut at two different points R1′ and R2′, two image contour segments A′ and B′ will result, as shown in Figure 7. Although there are countless ways of cutting such an image loop into two contour segments A′ and B′, only a small number of contour segments A′ can be projections from a portion of rim lying on a planar cut. Occluding contour segments that project from a planar portion of rim have several geometric properties that other occluding contour segments generically do not have. These are described later. An intuitive explanation is given in the discussion after each claim. Where indicated, a more detailed analysis can be found in Appendix A.

> Property 1: The projection of a planar cut corresponding to the visible portion of a cross-section through a volume will generically lack tangent discontinuities with the occluding contour projected from the rim of a volume with everywhere differentiable surfaces. (For proof, see Item 1 in Appendix A.)

The geometric relationships discussed here can be better visualized by comparing the planar cut of a volume in the world with its image projection. Figure 2A depicts a volume in the world sliced by a plane, and Figure 2B depicts the image projection of this situation. Although there is a surface tangent discontinuity where the plane meets the surface of the volume in the world, the corresponding surface contour in the image merges with the bounding contour of the silhouette without linear tangent discontinuities at the two points indicated by asterisks in Figure 2B.

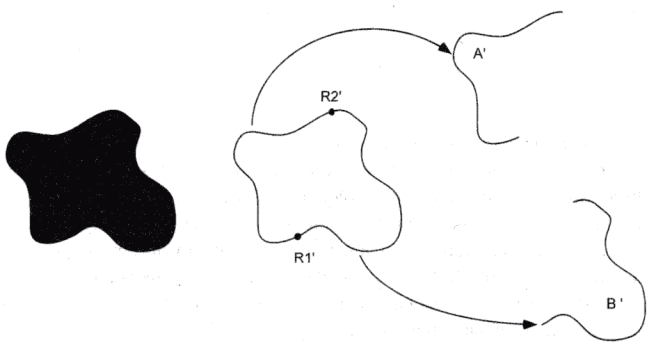


Figure 7. The silhouette on the left can be segmented into two halves: A′ and B′ at R1′ and R2′, respectively.

> Property 2: The image projection of a planar cut will generically exhibit abrupt curvature changes with the occluding contour projected from the rim. (For proof, see Item 2 in Appendix A.)

A useful way to think about points of abrupt curvature change is as follows. The radius of the circle that has the same curvature as the curvature at some point P on an image curve is given by the inverse of that curvature (Courant & John, 1989). There will be an abrupt curvature change at P when these "circles of curvature" change their radii discontinuously as they pass through P. For example, consider an ellipse. The circles of curvature will change smoothly into one another as we move around the ellipse. In contrast, the bump silhouette will have abrupt radii changes at its two "corners." Curves that contain inflection points will normally have such breaks because the circle of curvature will jump from one side of the curve to the other as it passes through the inflection point. However, because the visual system is more sensitive to the presence of local maxima of positive or negative contour curvature than it is to inflection points (Attneave, 1954), we can disregard such changes in sign and focus on abrupt changes in the absolute value of the radius of curvature. An abrupt curvature change can be defined as a point at which the second derivative along the contour is not well defined. This is a *second-order discontinuity.*

> Property 3: A segment of occluding contour that lacks curvature discontinuities projects from a segment of rim lying on a plane. (For proof, see Item 3 in Appendix A.)

Consider the case of a linear occluding contour segment. Assuming a generic view, a line projects from a linear segment of rim, which must lie in a plane. Similarly, consider a segment of occluding contour that corresponds to a segment of an ellipse. Assuming a nonaccidental view, this will project from an elliptic (and therefore planar) segment of rim because only an ellipse will generically project to an ellipse under orthographic projection. The orientation of the plane on which the rim segment lies is, however, ambiguous. Note that a segment of occluding contour that contains curvature discontinuities between its two endpoints can project from a segment of rim that lies on a plane, but this will not generically be true.

A corollary of Property 3 is the following: Large curvature discontinuities in the occluding contour are more informative about non-coplanarity of adjacent rim segments than are small curvature discontinuities. If two adjacent rim segments lie on different planes, but these planes are very close to being coplanar, then the abruptness of the corresponding curvature change in the image will be small, even if mathematically abrupt. In general, the abruptness of a given occluding contour curvature change will increase as the angle increases between the planes on which corresponding rim segments lie and will decrease as the angle decreases (for proof, see Item 4 in Appendix A). Because the visual system has limited resolution, it may not be able to detect very small curvature discontinuities. The visual system is most likely sensitive to large curvature discontinuities because these are likely to correspond to pairs of rim segments that lie on planes that deviate greatly from coplanarity. When planar cuts are almost coplanar, they describe almost the same cross-section. However, when planar cuts approach orthogonality, they provide two different cross-sections, and two cross-sections are more informative than one. Large curvature discontinuities, therefore, provide more

information about potential cross-sections than small ones. The visual system is thus likely to focus on large curvature discontinuities or their associated maxima of positive curvature rather than slight curvature discontinuities.[2]

In searching for planar cuts of the rim specified by the occluding contour, the visual system may search for local maxima of positive curvature because abrupt curvature changes arising from a planar cut will tend to lie on a local maximum of positive contour curvature in the image or in the immediate neighborhood of such a maximum. For example, for a cylinder (Figure 8, left panel) the curvature discontinuities R1′ and R2′ generated by a planar cut lie right on local maxima of positive contour curvature. For the silhouette of a cone, however (Figure 8, right panel), curvature discontinuity points are somewhat displaced from the local maxima of positive contour curvature. Rather than search for every point of abrupt curvature change along a silhouette's occluding contour, the visual system may just search for local maxima of positive curvature along the bounding contour of a silhouette and then consider only points of curvature discontinuity in their neighborhood. If the resolution of the visual system is low, it may just use segments of occluding contour between such maxima because mathematical discontinuities are typically eliminated as the image is blurred, but curvature maxima will generally be robust to blurring. This is consistent both with Attneave's (1954) insight that local maxima of positive curvature along a contour are particularly revealing about shape and with recent neurophysiological work demonstrating preferential macaque V4 cell tuning to maxima of curvature at specific locations along the closed contour of a silhouette (Pasupathy & Connor, 2001; see also Pasupathy & Connor, 1999). Of course, it is an empirical question how abrupt a curvature discontinuity or how sharp a local maximum of positive curvature must be in order for the visual system to take it as potentially informative about the possibility of a planar cut. It is also an empirical question at what resolution or resolutions the visual system searches for such informative points. The bounding contour of a silhouette of a volume will only have a finite number of such local maxima of positive curvature at a given resolution. The number of segments of occluding contour between pairs of these maxima will, therefore, also be finite. One or more of these segments can then be considered for propagability.

I now describe a contour propagation algorithm that exploits this contour information. Once a segment of the occluding contour with the above properties has been specified, it can be propagated into the interior of a silhouette. The basic idea is simple. Step the specified segment A′ into the interior of the silhouette and expand or contract it until it just touches the sides of the silhouette. Call these new surface contours "prop(A′)." A detailed description of the algorithm is given in Item 5 of Appendix A. Other constraints on contour propagation are described later, but this is the main idea. Note that the algorithm describes the propagation of occluding contour information across the image, not the propagation of planar cuts across the volume itself. However, the prop(A′) contours are interesting only because they help us recover the shape of a volume or curved surface by telling us about cross-sections where there are no local cues to cross-sectional shape. Of course, propagation in this manner will not give a unique shape solution to a silhouette in a metric sense because the precise depth, slant, and *tilt* of planar cuts will remain uncertain (for a discussion of this point, see Items 6 and 7 in Appendix A). However, the need to be consistent with propagable or partially propagable segments of occluding contour imposes strong constraints on possible shapes consistent with an image.

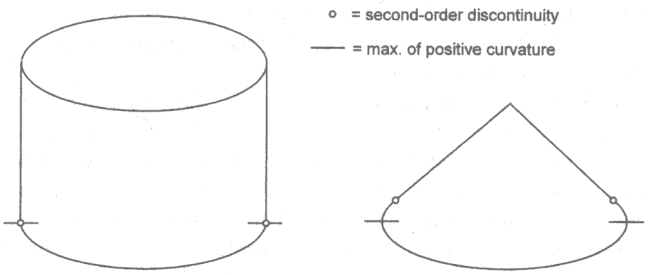


Figure 8. Local maxima (max.) of positive contour curvature and points of second-order discontinuity overlap for the base of a cylinder but not for the base of a cone.

Propagating A′ amounts to generating surface contours in the interior of the silhouette that correspond to the image projection of a series of planar cuts of the volume that are parallel to the plane on which the rim segment A lies. If A′ is propagable into the interior of the silhouette, then the surface orientation relationships pertaining at A will also apply in the surface neighborhood of A that projects to the interior of the silhouette. A more general contour propagation algorithm will have to take into account how the propagating surface contour deforms under perspective projection and in interaction with the rest of the occluding contour B′. These issues are discussed later. First, however, let us consider some other properties of A′ that determine whether it is a propagable segment of occluding contour or not.

> Property 4: If prop(A′) extends outside of A′ or B′, then A′ is not propagable.

If prop(A′) extends beyond A′ or B′, then there will be sections of prop(A′) that do not lie within the silhouette. This is impossible because surface contours cannot lie outside of occluding contours, just as a cross-section of a volume cannot lie outside that volume. An example of what is disallowed is shown in Figure 9a and b.

> Property 5: A propagable segment of contour A′ will have an overall radius of curvature that is larger than that for the nonpropagating segment B′.

[2] It might behoove the visual system to segment occluding contours into linear and quadratic segments because these will always lack curvature discontinuities, and will always project from linear and quadratic segments of rim on a volume, assuming orthographic projection. Sensitivity to linear and quadratic curvature could underlie sensitivity to abrupt changes of curvature at multiple scales. Because the "bump" silhouette is composed of two half ellipses of different aspect ratio merged along their common major axis, and arcs of an ellipse in the image generically project from rim segments that lie on a common plane, we can assign the upper half of that silhouette to one plane and the lower half to another, without specifying the precise slants of those planes.

If we fit[3] one arc of an ellipse to A′ and another to B′, the A′ arc will appear less curved in the image than the B′ arc. Intuitively, this must be the case because, when A′ is taken to be the projection of the near edge of a volume's cross-section, the volume itself is presumed to occlude the far side of that cross-section. The image projection of the rest of the volume B′ must, therefore, bulge away from A′. In Figure 10a, B′ must arch above the dotted far side of A′ to obscure it. When the segment is taken to be the projection of the far edge of the cross-section as in Figure 10b, B′ must still bulge so that A′ can propagate. If B′ did not have a smaller overall radius of curvature than A′, it would not bulge out enough for A′ to propagate. That is, there would be no volume over which the cross-section corresponding to A′ could propagate.

A further constraint on propagated contour segments is the following: An A′ in the image that is fit with an ellipse whose aspect ratio approaches 1 projects from a segment of rim that lies approximately in the frontoparallel plane. This follows from the simple observation that an arbitrary ellipse will only project to a circle in the image given an accidental viewpoint. In general, an ellipse that is not a circle will project to an ellipse that is not a circle. Thus, assuming a generic view, if A′ is best fit by a circle, it projects from a segment of rim that would also be best fit with a circle that lies in the frontoparallel plane.[4] An example of this is shown in Figure 11a. The silhouette is typically interpreted in two ways (Figure 11b and c). In both cases, the two A′ project from rim segments A that appear to lie approximately in the frontoparallel plane (Figure 11d and e). Projections of planar cuts would correspond to a stack of surface contours that imply an approximately cylindrical surface. Moreover, because successive inferred planar cuts correspond to cross-sections that must occlude one another, under one interpretation the cylinder seems to point into the page and under the other interpretation it appears to jut out of the page.

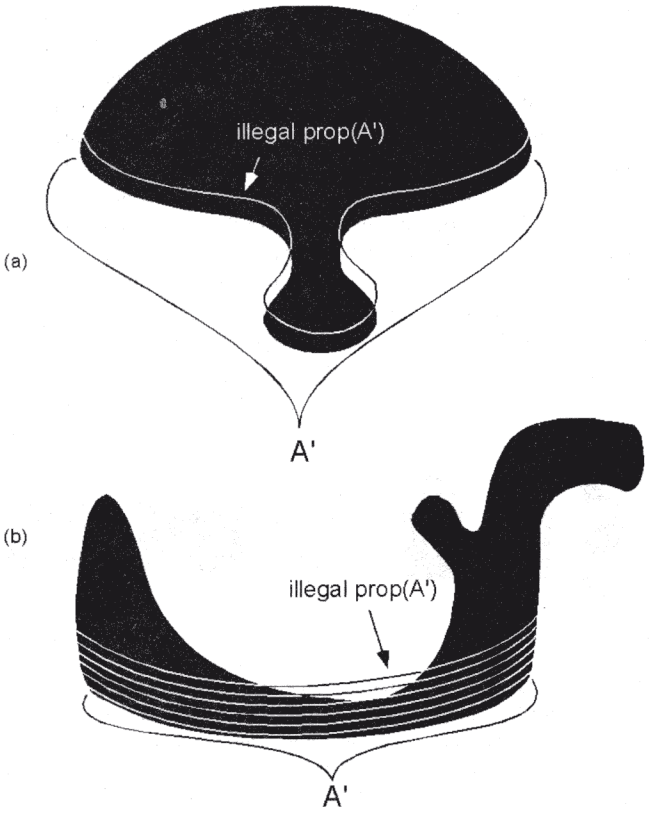


Figure 9. (a) Prop(A′) cannot cross A′. (b) Prop(A′) cannot cross B′. The portion of the occluding contour that is not A′ is B′. prop = propagation function.

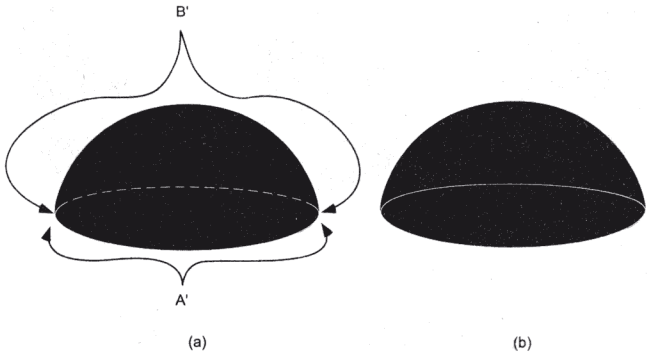


Figure 10. B′ must bulge above the ellipse fit to A′ in either the viewed-from-above case (a) or the viewed-from-below case (b) so that A′ has a region over which it can propagate.

The five properties described previously set propagable segments of occluding contour apart from nonpropagable segments. When prop(A′) satisfies these conditions, then A′ is a propagable segment of occluding contour. An example of the algorithm for locating propagable contour segments in the image is shown in Figure 12. In Figure 12a, prop(A′) is first-order continuous and second-order discontinuous with B′ and can, therefore, count as the image projection of a volume's planar cut. In Figure 12b, A′ is not propagable because prop(A′) is first-order discontinuous with B′. In Figure 12c, prop(A′) is first-order continuous and second-order discontinuous with B′, but A′ has a smaller overall radius of curvature than B′ and is, therefore, not propagable. Of course, it is possible that an entirely flat object will project to an image that has propagable segments of occluding contour. Indeed, all the silhouettes considered here are such objects. The fact that we tend to see these silhouettes as 3D demonstrates just how powerful propagable segments are as 2D cues to 3D form. In the absence of some other image cue indicating that a propagable segment is not a valid cue to a volume's cross-section, the visual system appears to assume that it is.

[3] An ellipse can be filled using some procedure in the image such as minimization of least squares, taking into account that usually only about half of a planar cut is visible because of self-occlusion.

[4] There are other examples of image contour configurations that can be assumed to lie in a frontoparallel plane assuming a nonaccidental view. A right angle in the image will generically project from a right angle in the frontoparallel plane because a right angle in the world will not project to a right angle from any other plane under perspective projection (Richards, Jepson, & Foldman, 1996). Similarly, horizontal (vertical) contours in the image will generically project from horizontal (vertical) lines in the frontoparallel plane. If a non-frontoparallel line in the world projects to a horizontal (vertical) line in the image, then a slight vertical (horizontal) movement of the head will make the line nonhorizontal (nonvertical) under perspective projection.

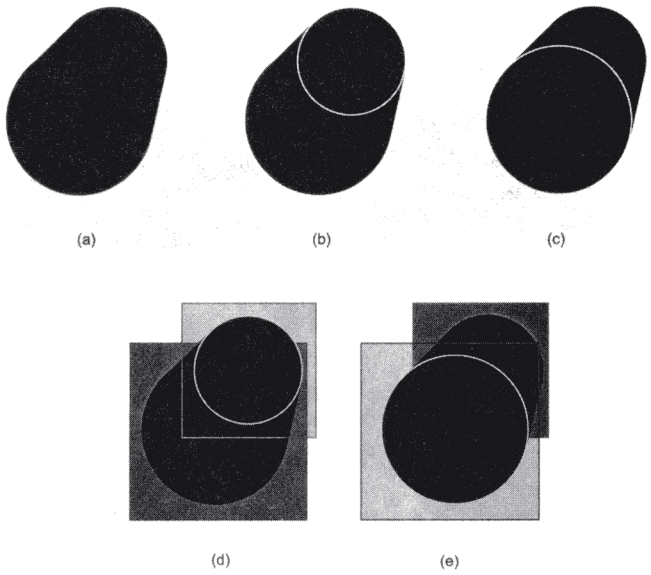


Figure 11. If the silhouette in Figure 11a is given a 3D interpretation, it tends to be given either the interpretation in 11b or the interpretation in 11c. The front and back circles are assumed to project from segments of rim in the frontoparallel plane as shown in 11d and 11e.

Multiple Propagable Contour Segments

In some cases, such as shown in Figure 3A, alternative propagable segments will overlap. In general, the volume interpretation generated by propagating one of these propagable segments will be different than that generated by propagating the other segment. Because a given volume in the world can only have one 3D form, the visual system generates a volume interpretation consistent with only one of these segments. Because different segments of partially overlapping occluding contour that are propagable will generally imply different surface layouts, a given point on the contour can belong to only one propagable segment at a time.

The global form inferred from a silhouette must always be consistent with all local propagations. For the "coral" depicted in Figure 13a, this is not a problem because the local propagations imply mutually consistent cross-sections and, therefore, one or more globally consistent volume solutions. When two inferred cross-sections differ in shape, intermediate cross-sections can be interpolated using a morphing algorithm. Problems interpreting a silhouette arise when local propagations are not consistent with one another, when they imply mutually inconsistent cross-sections, or when one segment implies a volume and the other a flat or curved nonvolumetric surface. For example, in Figure 13b, the silhouette is consistent with a cylindrical cross-section along its base but is not consistent with this same cross-section at the top, where there is just a straight line. Because the visual system assumes a nonaccidental view, this line projects from a linear segment of rim. Surfaces in the neighborhood of such a segment of rim must have zero curvature and, therefore, either be flat or cylindrical. If the surface curvature is cylindrical, then the visible portions of horizontal planar cuts will be lines and vertical planar cuts will be curved. However, this is orthogonal to the cylindrical curvature implied by the base, where horizontal planar cuts would appear curved and vertical planar cuts would be lines. Because interpreting the upper portion of the silhouette as projecting from a surface with cylindrical curvature leads to an inconsistency, the inferred surface *collapses* into a flat one in the neighborhood of the top-most linear segment of rim. The overall shape may then be curved on the bottom and wedgelike at the top, like the head of an axe. This collapse would also explain why the silhouettes in Figure 1b look flat, whereas those in Figure 1a look volumetric. Another example is shown in Figure 13c. In this case, the triangular top of the silhouette is consistent both with a triangular flat projecting surface and a conical projecting surface. Because the bottom of the silhouette implies a cylindrical cross-section and the top is not inconsistent with that cross-section, propagation of that cross-section all the way to the top is possible. The most likely perceived shape will, therefore, be a cylinder with a cone on top. The silhouette in Figure 13d is locally consistent with a cylindrical surface solution in the neighborhood of any of its four ends. However, the surface at the center of this cross cannot be both cylindrical along one axis and cylindrical along the perpendicular axis. Therefore, the surface curvature in the center is ambiguous and collapses into a flat surface to resolve the inconsistency. In Figure 13e, the base is consistent with a volumetric solution but the top is not. This is not a problem because the cross-section can be a curved line for all transversal planar cuts, and the perceived form will be a curved rectangular sheet rather than a volume. Note that this silhouette is still subject to at least two interpretations: one where the middle hump appears closer and the other where it appears farther away than other parts of the surface. In Figure 13f,

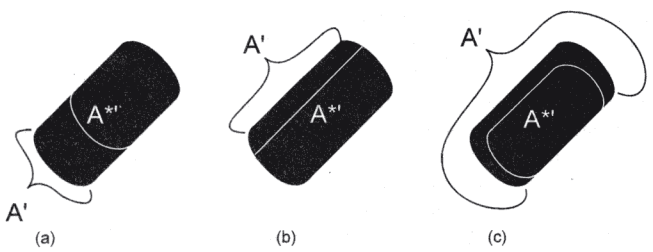


Figure 12. The A*′ (i.e., copy of the propagable contour segment A′) in Figure 12a is legal, but those in 12b and 12c are illegal for reasons explained in the text.

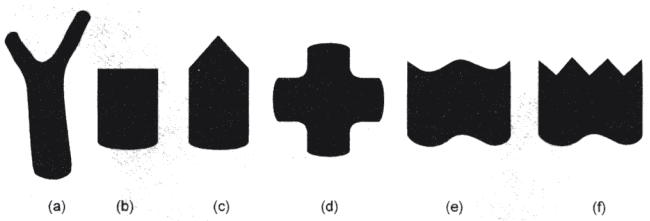


Figure 13. (a) This silhouette has three propagable segments. A global solution must be consistent with all local propagations. In Figure 13b the cross-section at the bottom is consistent with an ellipse, whereas at the top it must be a line. The perceived shape may, therefore, be an axehead. In Figure 13c the top triangle is ambiguous but consistent with the cross-section dictated by the shape of the base. The center of Figure 13d may appear flat because the four propagable segments imply mutually inconsistent surface curvatures. The silhouette in Figure 13e is only consistent with a curved sheet and (f) is ambiguous.

however, the interpretation is unclear because the base is consistent with a volumetric solution, but the top is probably not. Thus, to have a distinct open or closed curved surface solution, all propagations must imply consistent and globally morphable local cross-sections.

The Problem of Prop(A′) Deformation

By propagating cross-sectional information, the algorithm makes the implicit assumption that the cross-section in ambiguous distal portions of a volume bears some relationship to the cross-sectional information carried by the contour segment A′. The volume can be given by the integration of all similarly shaped cross-sections over the height or length of the volume, as depicted in Figure 7. In the case shown in Figure 7b, the shape of the inferred cross-section remains constant, although it changes its size over the course of propagation away from A′ and different amounts of that planar cut become visible as it propagates away from A′. In the cases shown in Figure 3b and c, however, the shape of the image projection of the planar cut changes as it propagates away from A′.

How does the visual system determine how the shape of prop(A′) will change with distance from A′? Clearly, the relationship between the shape of A′ and that of the rest of the occluding contour is essential. In image terms, the prop(A′) can be assumed to change shape in a way that is consistent with B′. One problem is that there may be multiple consistent solutions. There is, however, information in the shapes of A′ and B′ about surface curvature and, by implication, cross-sectional shape that can constrain solutions.

One important constraint was discovered by Koenderink and van Doorn (1976; Koenderink, 1984), who showed that the sign of occluding contour curvature at a given point corresponds to the sign of the surface curvature at the corresponding point on the rim. This insight forms the basis of the "codon" theory of shape recovery from silhouettes (Richards & Hoffman, 1985; Richards et al., 1987). Although this constrains possible solutions, there are still infinitely many possible volumes that could project to any given silhouette. Because the two possible interpretations of Figure 3a that are shown in Figure 3b and c have the same occluding contours, the sign of surface curvature must be the same at each corresponding point of rim on the two volumes. As Figure 3B and C demonstrates, however, volumes with different forms can satisfy the necessary sign equivalence between surface curvature at the rim and occluding contour curvature.

In general, a smooth closed loop, such as the outline of the silhouettes considered here, must have an even number of inflection points where the sign of curvature changes from positive to negative (Beusmans, Hoffman, & Bennett, 1987). In order to close, any change in curvature sign will have to be offset by a sign change in the opposite sense.[5] An inflection point on the contour projects from the point on the rim where a curve lying on the surface composed of points of zero surface curvature intersects the rim. A small number of possible classes of volume solutions can be generated by linking pairs of inflection points with curves of zero curvature (Beusmans et al., 1987). These curves of zero curvature cannot cross (Koenderink & van Doorn, 1980, 1981) for volumes with surfaces that are differentiable everywhere. As Figure 14 shows, the curves of zero surface curvature separating the regions

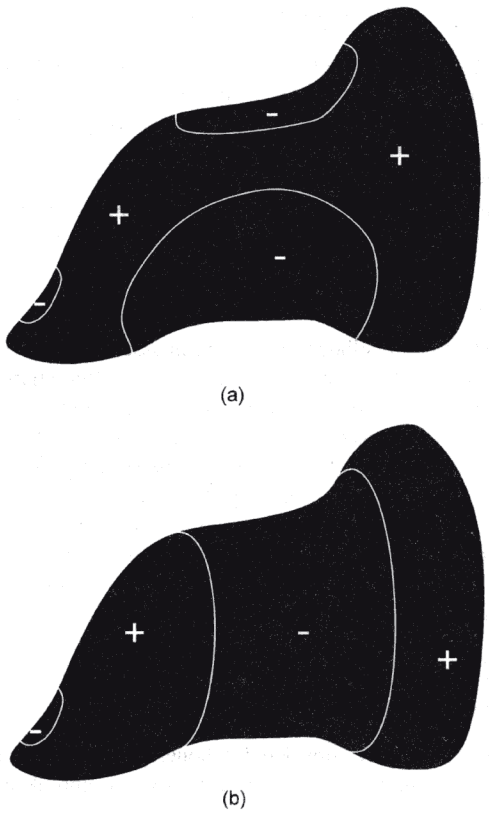


Figure 14. Regions of negative surface curvature may remain separate (a) or they may link up in (b). Figure 14a corresponds to the solution shown in Figure 3b and 14b corresponds to the volume solution shown in 3c.

of positive and negative surface curvature are different for the two form solutions shown in Figure 3. These solutions ignore the shape of the back surface of the volume because it is not known to what degree the surface curvatures of the backs of volumes are represented by the visual system (Tse, 1999a; van Lier, 1999). Given a propagable occluding contour segment A′, the visual system may be able to estimate the shape of prop(A′) on the basis of surface curvature information implicit in the shape of B′. A more detailed consideration of the information carried by the shape of B′ and how B′ may induce changes in the shape of prop(A′) can be found in Items 8 and 9 of Appendix A.

In summary, once a propagable segment of occluding contour has been identified, it will typically change shape as it propagates because the shape of corresponding cross-sections will generally vary over a volume. One way a propagable segment's shape can deform is through interpolation or morphing with other known cross-sections. Other cross-sections can be inferred from the shape of B′. In particular, the relationship discovered by Koenderink and van Doorn (1976; Koenderink, 1984) between contour and surface curvature sign imposes important constraints on possible cross-sections.

[5] Intuitively, this can be understood as follows. Imagine an ant walking along the contour with the inside of the silhouette to its left. To end up at its starting position, any time the ant turns to the right, it will have to turn to the left to undo this movement away from closure.

Generalizing the Algorithm

The algorithm for contour propagation as it has been described so far is limited by the initial simplifying assumptions and is not ideally suited to deal with silhouettes generated by blackening images of real objects. This section discusses ways of extending the prop(A′) algorithm so it can apply to silhouettes generated under situations of perspective projection, self-occlusion, surface nonsmoothness, and curved medial axes.

Generalizing to Cornered Objects

Up to this point, silhouettes have been assumed to be the image projections of volumes with surfaces that are differentiable everywhere. However, silhouettes can also project from volumes that have corners. It is easy to extend the algorithm to deal with cases, such as the silhouette of a cube, where the surface has corners by allowing the endpoints of the prop(A′) to have first-order discontinuities with the nonpropagating part of the bounding contour B′ for A′ that are first-order discontinuous with B′. Such endpoints would obviously still be points of curvature discontinuity between B′ and any given prop(A′).

Generalizing to Contour Tangent Discontinuities That Imply Self-Occlusion

So far I have only considered silhouettes that look volumetric and whose occluding contour projects from an unbroken loop of rim. However, many silhouettes look like volumes with inferred rims that are not unbroken loops, such as the "hat" shown in Figure 15. The existence of contour tangent discontinuities in such silhouettes may be taken as potential T-junctions that could serve as cues to surface self-occlusion and thus 3D form relationships. Many authors have suggested that surface completion behind an occluder occurs because of good continuation (Wertheimer, 1923) among image contours that terminate at an image tangent discontinuity such as a T-junction (e.g., Kellman & Shipley, 1991; Takeichi, Nakazawa, Murakami, & Shimojo, 1995; Tse, 1999a; Wouterlood & Boselie, 1992). The initiating conditions for potential contour and surface interpolation are local tangent discontinuities.[6] Initiated contour interpolations can link up at a distance with other interpolated contours if they meet the condition of good contour continuation. Because the contours terminating at L-junctions in the "hat" silhouette are continuous, two surfaces at different depths, one occluding the other, may be interpolated. The hat in Figure 15a has two interpretations that are consistent with such an interpolation. These are shown in Figure 15b and c. Cross-sectional information also appears to play a role in these types of silhouettes because the cross-section of the lower part of the inferred volume appears to have a cross-section whose aspect ratio is the same as that of the upper elliptical portion of the silhouette.

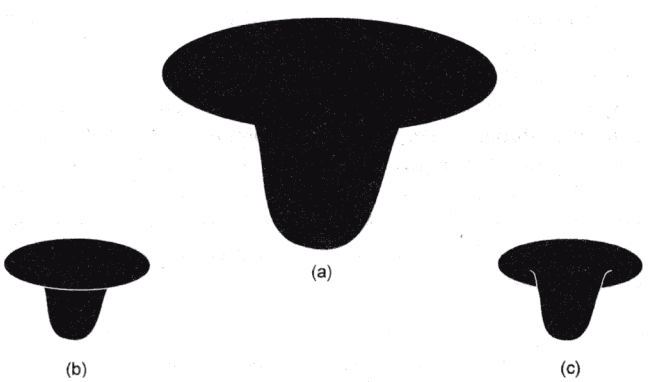


Figure 15. Silhouettes that look 3D can project from objects whose rim is discontinuous. (b and c) Two possible interpretations of Figure 15a.

Generalizing to Perspective Projection

The visual system may assume that the overall aspect ratio of a cross-section will be preserved in distal portions of a volume in the absence of image cues that it is not preserved. That is, if a propagable segment of image contour is fit with an ellipse, then the visual system may assume that the ellipses fit to distal planar cut projections have the same aspect ratio, have parallel axes, and project from planar cuts that lie on parallel planes through the volume. Consider the silhouette projected from a cylinder. The interpolated ellipses generated from the propagable segment of occluding contour at the base all have the same aspect ratio and orientation. There are two propagations consistent with this ellipse and two concomitant possible percepts (Figure 2). This type of propagation would be consistent with an assumption of orthographic projection but is inconsistent with the perspective projection that, in fact, underlies image formation on the retina.

Propagation of planar cut information consistent with a given perspective may require at least two propagable segments of the silhouette's occluding contour that do not overlap. For example, the silhouette shown in Figure 16a could be the image projection of a cylinder. Note that the two propagable contour segments A1′ and A2′ have different aspect ratios, as indicated by the pair of ellipses in Figure 16a. This is not consistent with the orthographic projection of a cylinder. The set of propagated surface contours in the image—the set of all prop(A′)—would have different aspect ratios even though the stack of planes defining the planar cuts of the volume may be parallel in the world and the set of cross-sections across the volume would be identical in the world as well. Depending on which lip of the inferred cross-section is presumed to be in front, the two propagations shown in Figure 16b and c are possible. To interpolate how the set of prop(A′) will be deformed by perspective, the aspect ratio of a given prop(A′) between A1′ and A2′ could be given by a linear weighted average of the aspect ratios of A1′ and A2′.

The visual system may make other assumptions to help interpret the distortions introduced by a perspective projection. One may be that an A′ projected onto the image from a planar cut through a convex volume whose plane passes over the eyes will generally bulge upward, whereas the opposite will be true for planar cut planes that pass below the eyes (see Figure A7). The visual system can perhaps gauge the level of its viewpoint with respect to an object by considering the degree of upward and downward bulging for two propagable contour segments A1′ and A2′.

[6] The process of contour and surface interpolation is not a simple bottom-up process, because the "mergeability" of volumes determines which continuous contours can be linked (Tse 1999a).

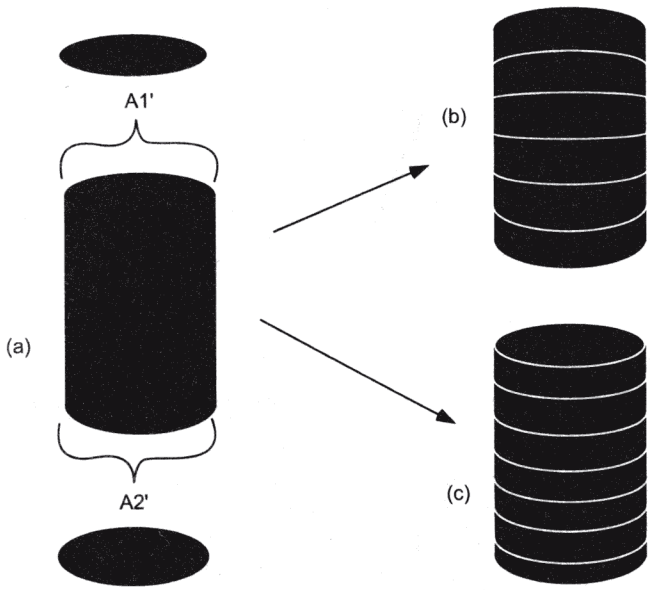


Figure 16. A1′ and A2′ have different aspect ratios. This difference is taken as a perspective cue. (b and c) Two solutions consistent with the silhouette in Figure 16a.

Note that perspective cues can be interpreted as information about object size, object orientation in the world, or viewer position with respect to the object. The visual system must overcome this ambiguity using world knowledge and image cues beyond the shape of the occluding contour because the occluding contour is inherently ambiguous. The gauging of viewing position from perspective cues implicit in the occluding contour may interact with certain priors that the visual system may have on assumed viewing position. For example, there may be a preference for right angles at inferred corners (Richards et al., 1996), placement on ground planes (Albert & Tse, 2000; Tse, 2000), or canonical shapes such as cuboids, cones, and cylinders given ambiguous image data. Such preferences would presumably reflect the statistics of the world (e.g., Knill & Richards, 1996) learned through experience or internalized over the course of evolution.

Generalizing to Curved Medial Axes

There has been an implicit assumption so far that cross-sectional information is propagated in such a way that successive cross-sections are due to parallel planes. This need not be the case because cross-sections need not be parallel in the world in order to convey information about 3D form. For example, in the "candy cane" shown in Figure 17a, the cross-sections are not parallel, as shown in the two solutions drawn in Figure 17b and c. The manner in which successive prop(A′) surface contours are generated could take the curvature of a silhouette's medial axis into account. For the case of the candy cane, the medial axis generated by the most common algorithm for generating medial axes in the image is shown in Figure 17d. This algorithm links the centers of all circles that have grown within the interior of the silhouette until they make contact with as much of the bounding contour of the silhouette as possible (Blum, 1973; Kovacs, Feher, & Julesz, 1998; Kovacs & Julesz, 1994; compare also Siddiqi, Kimia, Tannenbaum, & Zucker, 2001). Given the medial axis shown in Figure 17d and a propagable segment A′ approximated by an ellipse whose major axis makes an angle X with the tangent to the medial axis, the prop(A′) can be propagated such that their major axes continue to maintain the same angle X with the tangent to the medial axis. This modification of the algorithm would allow for solutions such as those shown in Figure 17b and c.

Although a planar cut approach to volume recovery from silhouettes is not inconsistent with a medial axis account, there are instances in which a medial axis account gives a wrong shape description. Because the center-of-circles algorithm described previously operates over the image and not over the volume that projects to the image, it will tend to give a useful medial axis description only for cases in which a long object in the world casts a long silhouette in the image. In cases in which a bloblike object casts a long image, as in the bump shown at the bottom of Figure 1a, or a long object casts a bloblike image, the medial axis approach will tend to give wrong answers. For example, the medial axis shown in Figure 18a implies a long Brazil nut-shaped object. Because observers tend to see a round bump rather than an elongated 3D shape when seeing the silhouette alone, the medial axis approach by itself is an inadequate basis for a visual shape code. Except where long objects cast long silhouettes, there will not generically be a relationship between the medial axis of the object in a 3D sense and the medial axis in a 2D sense. The axis of rotational symmetry of a bump in the world would cast the line shown in Figure 18b instead, and this line is perpendicular to the medial axis shown in Figure 18a. Note that this line is the path traced out by the centers of the prop(A′).

There may be other ways that the occluding contour can interact with a medial axis. Implicit in the algorithm is the idea that rim

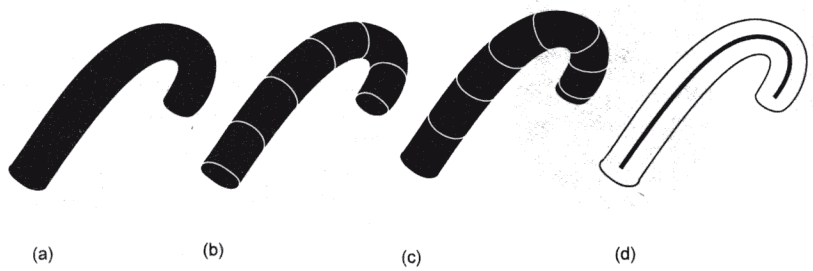


Figure 17. (a) The candy cane silhouette has at least two solutions (b and c). The propagated segment may interact with the medial axis (d).

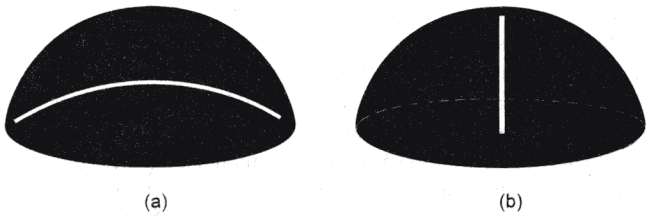


Figure 18. (a) Most medial axis algorithms would result in this axis. This is a wrong solution because it does not correspond to any axis of the perceived 3D shape. The axis of symmetry of a bump would project to the line shown in Figure 18b.

segments rotated around the medial axis in a 3D sense sweep out a volume. Given only one rim segment, such an operation would typically lead to corresponding occluding contour segments that are the mirror image of one another across the medial axis in a 2D sense. An example of this is depicted in Figure 19b for the silhouette shown in Figure 19a. Symmetrical silhouettes may look 3D in the absence of a propagable segment of occluding contour if (a) they could be the image projection of a volume that could have been generated through revolution in this way and (b) the volume thus generated could not project to a silhouette with a propagable contour. The existence of symmetry itself may, therefore, imply a circular cross-section in the absence of a segment of occluding contour that specifies a noncircular cross-section.[7]

Shape From Surface Contours

For static images, occluding contours probably offer the strongest constraints on 3D form because other cues, such as shading or texture, appear to be captured or dominated by contour cues (cf. Christou, Koenderink, & van Doorn, 1996; Koenderink, van Doorn, & Christou, 1996; Koenderink, van Doorn, Christou, & Lappin, 1996). As line drawings and silhouettes demonstrate, occluding contours can be a sufficient cue to 3D shape, although they are not a necessary cue.[8] Indeed, the bounding contour carries more information than shading cues when recognizing objects across rotations in depth (Hayward, 1998; Hayward & Tarr, 1997; Hayward, Tarr, & Corderoy, 1999). Inferring cross-sectional information from the shape of the occluding contour may raise 3D form possibilities that are further constrained by other image cues,

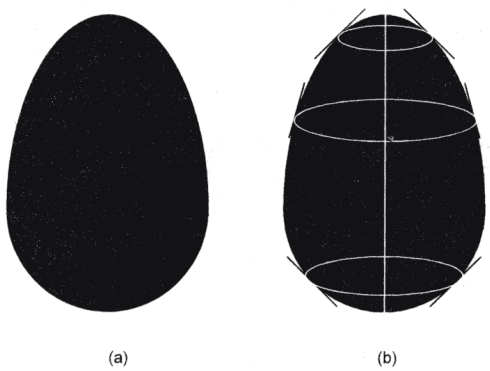


Figure 19. (a and b) Some silhouettes that lack propagable segments may appear 3D because of symmetry.

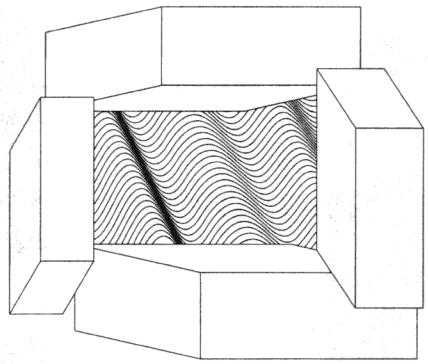


Figure 20. Surface contours are sufficient to generate a 3D percept.

and other image cues may help determine which segments of occluding contour provide cross-sectional information.

The shape-from-contour system can recover shape from surface contours as well as it can recover shape from occluding contours. There are examples, such as shown in Figure 20, in which there are no occluding contours at all, and still we see an undulating surface. Claims (e.g., Li & Zaidi, 2000) that 3D shape can only be recovered when a texture has most of its energy along directions of principal curvature are disproven by this type of example (Todd & Oomes, 2001). In this case, the surface contours are assumed to correspond to planar cuts of the underlying surface, and undulations of the former are taken to correspond to undulations in the latter. Note that points along a single surface contour that are lower than other points along that surface contour correspond to points on the curved surface that are lower in the world, assuming orthographic projection. Which surface contour is seen to be furthest in front is ambiguous. It could be either the top-most or bottom-most contour in Figure 20.

In general, the manner in which surface contours intersect the occluding contour yields important 3D shape information. Because surface contours are presumed to correspond to a planar cut in the absence of evidence to the contrary, surface contours that meet the occluding contour without first-order discontinuities imply that the rim lies on a differentiable portion of surface. However, surface contours that intersect the occluding contour with first-order discontinuities, such as shown in Figure 21, imply that the rim lies on a corner or that the surface has an edge at the corresponding points in the world.[9]

Surface contours can be interpreted as corners and can serve as constraints on how propagable segments of occluding contour propagate. Thus, the surface contours in Figure 22a appear to

[7] Note that even a circular silhouette can look like a sphere when placed in the appropriate 3D context, such as on the image of a ground plane. In this case, there is no medial axis and no propagable segment of occluding contour.

[8] Just as there is no necessary and sufficient image cue for amodal completion (Tse, 1999a), there appears to be no necessary and sufficient image cue for 3D form. While occluding and surface contours are strong cues to the recovery of 3D shape, they are by no means necessary cues because shape can be seen in the absence of either.

[9] They could also imply partial submersion of the surface in an occluding medium.

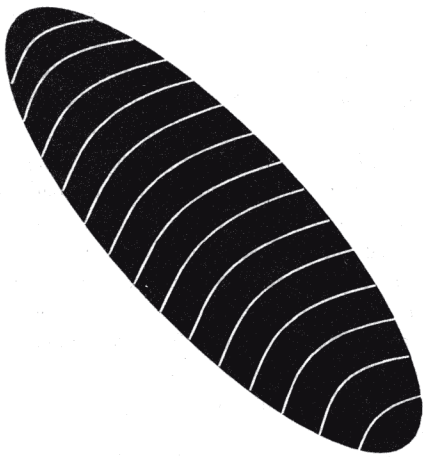


Figure 21. Because the surface contours have tangent discontinuities with the occluding contour, the occluding contour is seen as projecting from a rim that lies on a sharp corner or edge.

comprise a corner on the surface of a chevron. A surface contour will be interpreted as a corner when it intersects with a sharp tangent discontinuity along the occluding contour, as here, because a surface tangent discontinuity at the rim will project to a contour discontinuity assuming a nonaccidental view. Conversely, the absence of such a correlation (i.e., a Y or arrow junction) in the image indicates that the surface contour does not correspond to a surface tangent discontinuity or corner in the world. In the case shown in Figure 22B, the entire occluding contour serves as the projection of a cross-section, and this information can be propagated inward in interaction with the presumed corners.

General Discussion

Comparison With Other Shape Descriptions

There are several shape formation models in the literature. The primary ones are (a) codon theory (Richards & Hoffman, 1985; Richards et al., 1987), (b) a parts-based or structural approach such as the geon theory of Biederman (1987; Biederman & Gerhardstein, 1995; Hummel & Biederman, 1992), (c) a medial axis approach (e.g., Blum, 1973; Kovacs & Julesz, 1994), and (d) an approach premised on the complete recovery of visible surface orientations and depths (Marr, 1982). In this article I have outlined another approach to shape perception, surface filling-in, and volume formation premised on the ideas of contour propagation and the recovery of cross-sectional information. This contour propagation approach is offered as a solution to the specific problem of how 3D form might be seen in silhouettes. It is not meant to replace existing theories and, indeed, could supplement versions of any of the four major theories. It is important to keep in mind that the representation of 3D shape used by the visual system may not be monolithic. Shape codes could involve aspects of more than one of these theories, depending on the particular problem to be solved. For the particular problem of recovering 3D shape from 2D contours, a contour propagation approach may result in a more useful and less ambiguous shape description than those of other theories. It is, therefore, instructive to compare the contour propagation algorithm with each of the four major shape theories.

The codon theory (Richards & Hoffman, 1985; Richards et al., 1987), discussed before in relation to Figure 14, imposes useful constraints on which 3D shapes can be inferred from a silhouette. The bump, for example, inferred from the silhouette shown in Figure 1A, is one solution available in a large family of solutions consistent with a codon description. However, there is nothing in the codon theory itself that would explain why observers see a bump here rather than a Brazil nut. Thus, the codon approach offers useful constraints on possible shapes but does not predict which specific shape of all possible shapes will be perceived. The contour propagation approach, in contrast, predicts more precisely which specific shapes will be perceived from a given silhouette and can predict which silhouettes will look 3D and which will look flat. The codon theory makes mistakes because it is built on the flawed assumption that all silhouettes arise from 3D objects. Thus, an elliptical silhouette is claimed to look like an ellipsoid when to most observers it, in fact, looks like a flat hole or disk lying on a ground plane. Contour propagation predicts correctly that an elliptical silhouette, and indeed all the silhouettes in Figure 1B, will look flat.

The geon theory of Biederman (1987) is built on the idea that objects can be represented in view-invariant terms as an assemblage of primitive parts called geons. This theory is not able to infer a bump from the silhouette shown in Figure 1A because there are no parts in a bump, and a bump is not one of the volumetric primitives that Biederman posited. It would be futile to create a new shape primitive for every possible 3D object that lacks obvious parts because there are countless partless objects. Imagine walking into a cave full of arbitrarily shaped rock formations, none of which are segmentable into parts. We can clearly see the shape of these stone formations without being able to specify their parts. A confusion may arise because geons have perhaps been regarded as a solution to two independent problems: (a) shape formation, on the one hand, and (b) object recognition, on the other. Although 3D parts may be a useful way to index a shape in memory to recognize something, 3D shapes need to be constructed before they can be

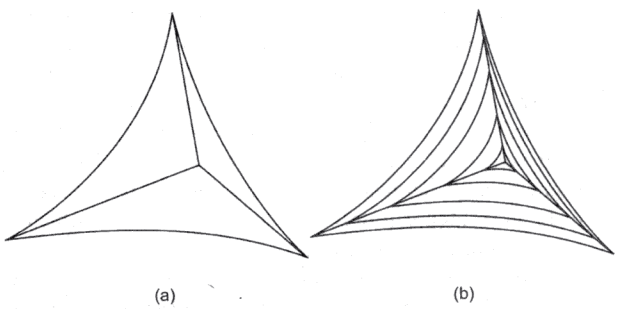


Figure 22. Surface contours can specify corners of the surface, especially when they intersect points of tangent discontinuity at the occluding contour. The whole occluding contour acts as a propagable segment in this example.

segmented into 3D parts.[10] It is, therefore, unlikely that shapes are themselves constructed from a small alphabet of primitive simple shapes. However, once surfaces and volumes have been constructed, it is reasonable that these are segmented at regions of local minima of surface curvature (compare Hoffman & Richards, 1984; Hoffman & Singh, 1997) and that these parts may serve as an index for the matches to memory that underlie recognition. A solution to the primary problem of shape formation should not be limited to combinations of primitive volumes because many shapes lack a distinct volume entirely, such as the surface of the ocean or swirls of smoke. In contrast to a geon-based approach to shape formation, the contour propagation approach is solely an account of (a) shape formation that makes no claims about how (b) object recognition takes place. According to the contour propagation approach, arbitrary curved surfaces and volumes can be generated from surface contours and occluding contours directly in the absence of any visual shape primitives.

Relevant here is a large body of empirical evidence that shows that object recognition is view dependent (e.g., Bülthoff & Edelman, 1992; Edelman, 1997; Tarr, 1995; Tarr & Bülthoff, 1995; Tarr & Pinker, 1989; Tarr, Williams, Hayward, & Gauthier, 1998; Wallis & Bülthoff, 1999) in contrast to the predictions of geon theory. These authors generally argue that objects are represented and stored as a series of views. However, what comprises a view is not clear. At one extreme a view might just be a 2D image. This extreme would have difficulty accounting for the various constancies (i.e., indifferences to image transformation) expressed by the visual system. For example, an object defined by contours alone, motion alone, or texture alone will tend to look like it has the same shape across these cues. Moreover, an object viewed from various distances and under various lighting conditions will generally appear to have the same shape, although particular images will be very different from each other. A more moderate stance is that a view is a collection of features, in which "feature" refers to any diagnostic combination of light and shade, color, form, and so on (Wallis & Bülthoff, 1999). Such features, even if they do not explicitly represent 3D shape or depth information, may implicitly capture 3D information, because viewpoint invariant recognition could emerge if all views of an object are matched to the same node in a distributed neural network (e.g., Bülthoff, Edelman, & Tarr, 1995; Poggio & Edelman, 1990). If network models can be built that match correctly, it may become difficult to experimentally distinguish whether the visual system constructs explicit representations of 3D shape or whether it only acts as though it did. At the other extreme, a view might include an explicit representation of 3D shape. Tarr and Kriegman (2000), for example, suggested that a view is a span of viewpoints over which the qualitative shape description, in terms of occluding contour relationships, does not change. This converges to a certain extent with the revised version of geon theory (Biederman & Gerhardstein, 1995; Hummel & Biederman, 1992), according to which recognition will be view invariant only over a set of views for which a given collection of geons is visible.

The contour propagation algorithm is consistent with these more recent stances of both the viewpoint-dependent (Tarr & Kriegman, 2000) and viewpoint-invariant (Biederman & Gerhardstein, 1995) schools. The nonmetric shape description that emerges from contour propagation in terms of relative slant, tilt, and depth is viewpoint dependent, because surface orientation and position at a given point on a surface are coded not only in ordinal relation to the orientations and depths of neighboring points on that surface (i.e., more or less slanted, closer or farther) but also in relation to the position of the observer. Indeed, by their very nature, slant, tilt, and depth are meaningful descriptions only with respect to a given viewing position. However, 3D structures (e.g., holes, protrusions, parts, corners, valleys, indentations) and the particular spatial relationships that hold among them (e.g., hole below pinnacle above bulge) that can be discerned from a given viewpoint are *intrinsic* to the object and can underlie a viewpoint-invariant representation of shape because these same structures will be visible from many other viewpoints. Even if a geon description *per se* is not used by the visual system for recognition, it is likely that some other structural description is. In general, the visual system attempts to recover the intrinsic properties of objects (e.g., surface reflectance, material substance, 3D shape) because these are more or less constant, whereas extrinsic properties (e.g., lighting, shading, shadows, distance, orientation) are constantly changing. Both intrinsic and extrinsic information can be derived from the image, and probably both types are stored and used for various tasks, including recognition.

Medial axis theory is inadequate because current algorithms calculate axes in the image, not in the world. Yet it is not clear how to generate medial axes in a 3D sense. Take, for example, a bump. If we generalize the center-of-the circle algorithm to a center-of-the-sphere algorithm (such that a sphere would grow from within the volume until it maximally "kissed" the inside surfaces of the bump), the "medial axis" would be a slightly curved disklike shape embedded in the bump (compare Mohr & Bajcsy, 1983; Nackman, 1982). This approach is not useful because it requires that we have the 3D shape description of the bump already, and this is just what we are trying to recover from the image. If we limit ourselves to determining axes in the image, then the medial axis approach can give wrong solutions, as shown in Figure 18a. In contrast to a medial axis approach to shape formation, the contour propagation approach places emphasis on existing contours that convey local cross-sectional and surface curvature information that can be generalized across the image. Nonetheless, the contour propagation approach is not incompatible with a medial axis approach. Note that inferred cross-sectional information may interact with a medial axis (Figure 17).

Metric surface recovery theories (e.g., Marr, 1982; Marr & Nishihara, 1978/1992; see also Gibson, 1950; Gibson & Robinson, 1935) maintain that perceived shape depends on recovering precise values of depth and surface orientation in viewer-centered coordinates for every point of a visible surface.[11] This approach to shape recovery is incorrect because the shape code underlying visual perception is not metric (i.e., euclidean). More recently, an exten-

[10] A local minimum of image contour curvature does not necessarily correspond to a local minimum of surface curvature and is, therefore, not a reliable means of discerning object parts. A helix or spring, for example, would project to a silhouette with many L-junctions arising from self-occlusion. However, the 3D object has no parts to speak of.

[11] About the time that Marr hit on the idea, Gibson (1979) came to repudiate his earlier (Gibson, 1950) notion that form perception is based on a determination of local depths and orientations because he found that subjects are in fact poor at judging even the slant of planar surfaces.

sive literature has emerged showing that there is substantial variance and inaccuracy when observers try to specify depth and orientation values for positions on a surface, even when given varied or multiple sources of visual information (e.g., shading: Erens, Kappers, & Koenderink, 1993a, 1993b; Koenderink, van Doorn, & Kappers, 1992; Todd & Mingolla, 1983; visual contours: Koenderink, van Doorn, Kappers, & Todd, 1997; texture: Reichel, Todd, & Yilmaz, 1995; Todd & Akerstrom, 1987; motion: Todd & Bressan, 1990; Todd & Norman, 1995; binocular disparity: McKee, Welch, Taylor, & Bowne, 1990; Koenderink, Kappers, Todd, & Norman, 1996). Most shape-from-*x* cues, including motion parallax, perspective, texture gradients, surface contours (e.g., Knill, 1992; Todd & Reichel, 1990), occluding contours (e.g., Koenderink, 1984), highlights, shading, or shadows, can only provide information about the *sign* of surface curvature (Beusmans et al., 1987). These shape-from-*x* cues can perhaps provide information about relative surface orientation or curvature but not about absolute surface orientation or curvature. Two shape-from-*x* cues, disparity (Carman & Welch, 1992; Gregory, 1970; Julesz, 1971; Marr & Poggio, 1977) and motion (Ullman, 1979), can, in principle, provide information about absolute surface curvature at any visible point of an object provided that certain reasonable assumptions, such as object rigidity, are adopted. However, the visual system does not seem to exploit this image information fully because perceived shape is not coded metrically (e.g., Koenderink et al., 1992; Koenderink, van Doorn, & Kappers, 1994, 1995; Koenderink, van Doorn, Christou & Lappin, 1996; Reichel et al., 1995; Todd & Norman, 1995) at least insofar as depth and surface orientation are not precisely represented. It appears that the visual system may be satisfied with a fairly inaccurate representation of 3D form rather than the precise one hoped for in Marr's (1982) program.

After researchers rejected Marr's (1982) program for the metric recovery of surface orientation and distance, it was not clear what type of shape description could or should replace it. Just because a metric description is not attainable does not mean that precise shape information cannot be recovered from the image. However, none of the major contour-based approaches to form perception besides Marr's offer a program for the recovery of precise shape information from the image. Certainly, geons (Biederman, 1987), codons (Richards et al., 1987), or medial axes (e.g., Kovacs et al., 1998) are not capable of uniquely or metrically specifying the curved internal structure of a surface. However, these other approaches cannot even give a precise ordinal or relational description of the curved internal structure of a surface. There seems to be a gap between the precise but metric description of shape that Marr sought and the imprecise, nonmetric shape descriptions that have been offered in its place. One goal of the current work, then, has been to try to help bridge this gap by developing an account of how precise shape relations that specify the ordinal or affine structure of a surface can be recovered from image contours. In contrast to a metric approach, the contour propagation approach to shape formation is inherently relational and ordinal. For example, we can tell which of two neighboring points has more slant without being able to specify the precise metric value of slant at either point. The relationships it describes are, moreover, rotation, scale, and translation invariant. A more detailed description of how contour propagation implies that shape is coded in terms of relative angle, slant, tilt, and curvature can be found in Items 6 to 9 of Appendix A.

In sum, of all these approaches, only the contour propagation approach provides a shape description that accounts for why the silhouette in Figure 1a looks like a bump rather than some other shape. Indeed, none of the other theories of how shape is coded can explain why we see the surface shapes that we do in the interior of a silhouette. The planar cut approach developed here argues that we see shape in the interior of a silhouette because of contour propagation. Contour propagation in interaction with other image cues, in turn, specifies how planar cuts are maintained and deformed over a volume. This also explains why some silhouettes look 3D and others look flat. Contour propagation can interact with constraints provided by codons or medial axes but offers new constraints on 3D form perception and the nature of the shape code used by the shape-from-contour system. In particular, the process of contour propagation can be thought of as the filling in of a curved surface from the boundary.

Future Directions and Unanswered Questions

The contour propagation approach can be taken in many directions in future work. Of immediate interest is the application of the algorithm described here in the context of computer vision. Problems may arise in its practical application that were not foreseen on theoretical grounds. I have ignored the hard problems of local edge extraction and global contour formation. In practice, these are extremely difficult problems that probably require top-down processing (e.g., Lee, Mumford, Romero, & Lamme, 1998; Leopold & Logothetis, 1999). Also, the problem of how arbitrary curves might be fit with lines and arcs of an ellipse was not addressed in detail. In practice, very slight deformations of a contour may require a radically different fit depending on the nature of the fitting rule. For example, if the bump silhouette is deformed slightly along its base as in Figure 23a, an entirely new curvature discontinuity is introduced. A similar problem might be introduced by the irregular base of a tree trunk such as the one depicted in Figure 23b. One way to deal with this problem is to fit quadratic segments at multiple scales. Each of the small irregular curves at the base of the tree trunk could be fit with a small arc of a quadratic curve. However, the best fitting large arc would correspond to the

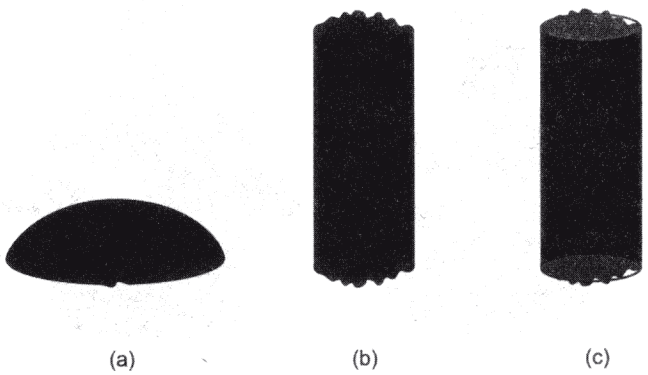


Figure 23. Perturbing a propagable segment creates abrupt changes in contour curvature that must be discounted. (a) The silhouette still looks like a bump. (b) The "tree stump" may still appear to have an approximately circular cross-section. This requires smoothing (c).

one shown in Figure 23c. This would amount to a smoothing process.

Another potential problem that might arise in the practical application of this algorithm may occur in dealing with parts of an image. Consider the silhouette shown in Figure 24a. The best approximation to this might be the four separate subsilhouettes shown in Figure 24b. The propagation algorithm could be applied after an image-based segmentation into parts (e.g., Hoffman & Richards, 1984; Hoffman & Singh, 1997) at points of deep concavity along the contour so that local propagable segments of occluding contour could propagate within each part without interference from other more distal propagations.

Another direction for future work will involve investigating how contour propagation interacts with other cues to 3D form. There are, no doubt, many interesting ways in which the shape-from-contour system can constrain and be constrained by the other shape-from-*x* systems. I have limited this discussion to shape cues available in static contours. This approach is not ecological. In the real world, an animal would probably move its head if it had any doubt about the layout of an object's surfaces. An interesting avenue of research involves examining the interactions between the shape-from-contour and shape-from-motion systems. The manner in which occluding contours change as an object rotates relative to an observer is an important cue to 3D form. For example, when the bounding contour of a silhouette deforms such that a tangent discontinuity suddenly appears, this can be taken as a cue for the sudden appearance of a self-occluding or previously occluded portion of surface. When the occluding contour becomes less curved, the underlying surface must also become less curved. How the visual system integrates these changes over time to recover surface layout is an important problem (see Norman, Dawson, & Raines, 2000; Raines & Norman, 1999). Even though information recovered from any particular view must be in a viewer-centered coordinate frame, the integration of relational information over a succession of views probably aids in the construction of a representation of shape in a nonmetric but object-centered frame (cf. Edelman & Weinshall, 1991; Foldiak, 1991; Perrett & Oram, 1993; Wallis & Bülthoff, 1999; Wallis & Rolls, 1997).

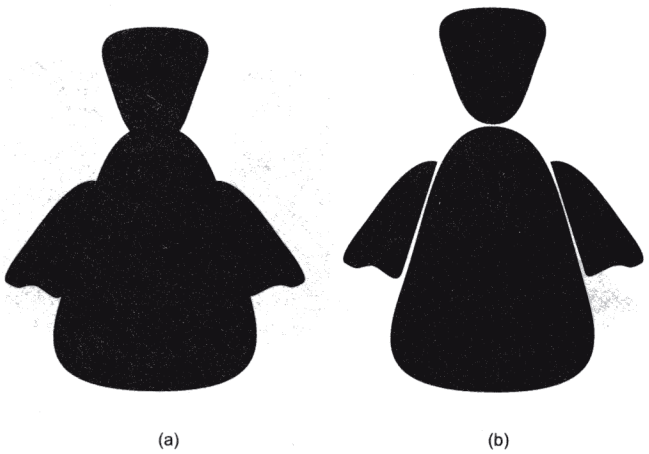


Figure 24. Segmentation (b) based on deep concavities within the image (a) may precede filling-in from propagable segments so that local propagations do not interfere.

Psychophysical experiments are needed to test the idea of contour propagation. Experiments should be able to ascertain whether a process of propagation actually takes place over time or whether occluding contours simply impose constraints on shape solutions that emerge as complete wholes under multiple constraints simultaneously. If contour propagation is a process that takes place over time, individuals may be able to judge ordinal relationships (e.g., closer/farther) more quickly when the judgment is made for two points in the interior of a silhouette that lie near a propagable segment of occluding contour than for points that lie far from a propagable segment.

One prediction of the contour propagation account of shape formation is that the visual system is highly sensitive to abrupt changes in curvature along a contour. Indeed, the visual system may be sensitive to abrupt changes of curvature without being sensitive to tangent discontinuities *per se*. Tangent discontinuities such as those found at T- or L-junctions are just a special case of an abrupt curvature change. Certainly, in the periphery, the visual system is not likely to make a distinction between tangent and curvature discontinuities because the spatial resolution in the periphery is low. A tangent discontinuity will always be a curvature discontinuity, but a curvature discontinuity need not be a tangent discontinuity. Because curvature discontinuity is the more general concept, it is possible that the visual system does not have special tangent discontinuity detectors such as receptive fields tuned to T-, X-, or L-junctions. Experimentally, curvature discontinuities are found to be cues for rapid shape analysis because curvature discontinuities pop out among curves that lack curvature discontinuities (Kristjansson & Tse, 2001). It is an empirical question how abrupt a curvature discontinuity must be in order to pop out. If the visual system is tuned to curvature, it is likely that it is tuned to multiple curvatures at multiple scales (Zucker, Dobbins, & Iverson, 1989, 1992). The same applies to curvature discontinuities.

The visual system is probably also sensitive to the higher order statistics of contours such that an irregular bounding contour is attributed an abrupt curvature change only where the overall curvature changes abruptly. For example, in Figure 25 the standard silhouette, Silhouette a, has been randomly "fractalized" along its length to an ever larger degree as one moves from Silhouettes b to f. Some of these modified silhouettes can still look volumetric, even though the underlying surface may now appear lumpy or crystalline. This demonstrates that the notions of abrupt curvature change and planar cut need to be developed further to encompass the role played by these higher order contour statistics. One way to accomplish this would be to process contours at several spatial scales and average the results. At some point, as the size of the fractals increases, the random deviations from the contour of Silhouette a are no longer discounted as noise masking important abrupt curvature changes. Instead, they become appendages of the silhouette in their own right.

Finally, a planar cut assumption may be an instance of a more general assumption made by the visual system. Namely, deviations from linear relations in the image are due to deformations caused by an underlying surface (or other spatiotemporal structure in the world). The visual system would then be sensitive to deformations not only of lines projected onto a curved surface (the planar cut assumption) but also of linear relationships of other kinds. It would calculate the inverse of these deformations to recover the world structure that caused the deformation, much as it must discount

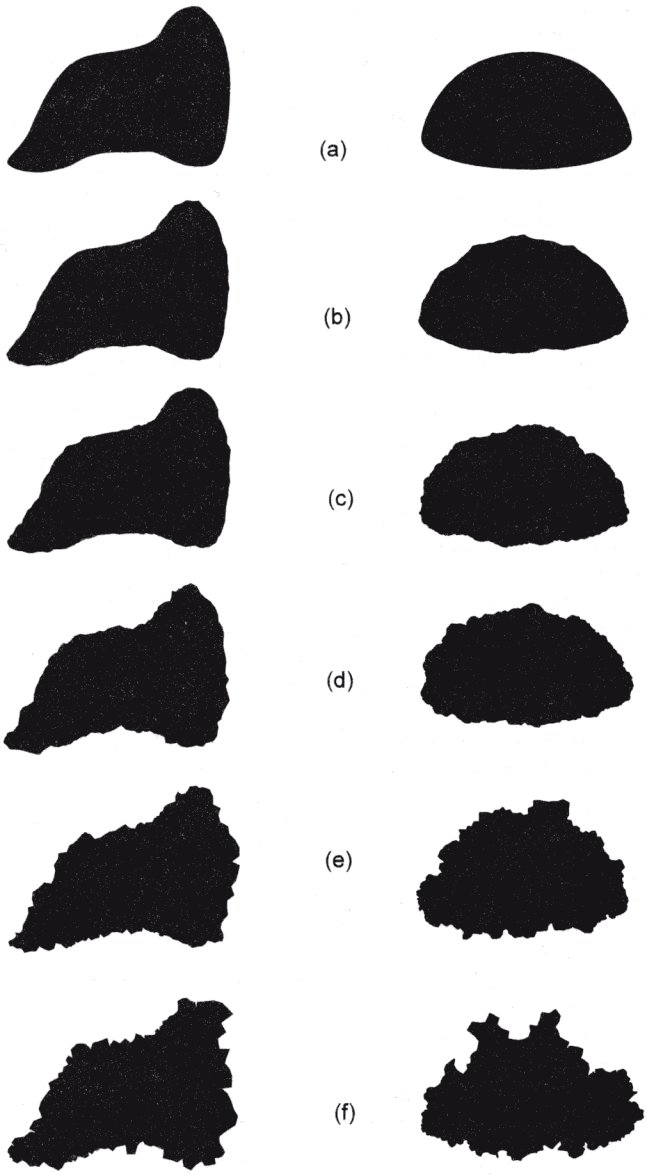


Figure 25. (a) The silhouettes have been subjected to greater and greater random deviations from the original contour for Silhouettes b–f.

image deformations introduced by perspective projection in order to recover 3D layout and shape correctly. For example, the visual system can recover surface shape from the deformations created when a uniform texture is projected or painted onto a curved surface (Todd & Oomes, 2001). A typical deformation of this class is foreshortening. How much and what types of texture nonuniformity can the visual system tolerate before it misinterprets texture "deformations" to be deformations caused by the underlying surface? Todd and Reichel (1990; Figure 2) created stimuli that can be thought of as the intersection of a "wrinkled" plane (e.g., a piece of paper that has been crumpled and then flattened again) with a smooth curved surface. These stimuli look similar to Figure 20 but are crossed by jagged rather than by smooth planar cuts. Such deviations from linearity do not necessarily hamper the recovery of underlying surface shape, suggesting that the visual system is able to discount noise in extracting linear relations and their surface-induced deformations. In the end, it is an empirical question what kinds of linear relationships and deformations the visual system can use to solve the inverse recovery problem.

Conclusion

I have outlined a new account of shape coding, surface filling in, and volume formation based on the idea of propagating occluding contour information to distal areas that may lack local cues to 3D form. The central idea is that the visual system seeks out and uses planar cut information to generate a 3D form percept. In particular, a propagable segment of occluding contour propagates and thereby fills in a curved surface within the image boundary of an object in interaction with that boundary and other image cues. A propagable segment provides information about the shape of a cross-section that can be used to infer the shapes of cross-sections elsewhere. The result is a nonmetric coding of 3D shape in terms of local ordinal surface curvature, depth, and orientation relationships that is scale, translation, and rotation invariant.

References

Albert, M., & Tse, P. (2000). The role of surface attachment in perceived volumetric shape. *Perception, 29,* 409–420.

Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review, 61,* 183–193.

Barrow, H. G., & Tenenbaum, J. M. (1981). Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence, 17,* 75–116.

Belhumeur, P. N., Kriegman, D. J., & Yuille, A. L. (1999). The Bas-relief ambiguity. *International Journal of Computer Vision, 35,* 33–44.

Beusmans, J. M. H., Hoffman, D. D., & Bennett, B. M. (1987). Description of solid shape and its inference from occluding contours. *Journal of the Optical Society of America A, 4,* 1155–1167.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94(2),* 115–147.

Biederman, I., & Gerhardstein, P. C. (1995). Viewpoint-dependent mechanisms in visual object recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1506–1521.

Binford, T. O. (1981). Inferring surfaces from images. *Artificial Intelligence, 17,* 205–244.

Blum, H. (1973). Biological shape and visual science: Part 1. *Journal of Theoretical Biology, 38,* 205–287.

Bülthoff, H. H., & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proceedings of the National Academy of Sciences USA, 92,* 60–64.

Bülthoff, H. H., Edelman, S. Y., & Tarr, M. J. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex, 5,* 247–260.

Carman, G. J., & Welch, L. (1992). Three-dimensional illusory contours and surfaces. *Nature, 360,* 585–587.

Christou, C., Koenderink, J. J., & van Doorn, A. J. (1996). Surface gradients, contours and the perception of surface attitude in images of complex scenes. *Perception, 25,* 701–713.

Courant, R., & John, F. (1989). *Introduction to calculus and analysis I.* New York: Springer Verlag.

Edelman, S. (1997). Computational theories of object recognition. *Trends in Cognitive Sciences, 1,* 296–304.

Edelman, S., & Weinshall, D. (1991). A self-organizing multiple-view representation of 3D objects. *Biological Cybernetics, 64,* 209–219.

Erens, R. G., Kappers, A. M., & Koenderink, J. J. (1993a). Estimating local

shape from shading in the presence of global shading. *Perception & Psychophysics, 54,* 334–342.

Erens, R. G., Kappers, A. M., & Koenderink, J. J. (1993b). Perception of local shape from shading. *Perception & Psychophysics, 54,* 145–156.

Faugeras, O. (1995). Stratification of 3-D vision: Projective, affine, and metric representations. *Journal of the Optical Society of America A, 12,* 465–484.

Fermuller, C., & Aloimonos, Y. (1996). Ordinal representation of visual space. In *Proceedings of the ARPA Image Understanding Workshop,* 897–904.

Foldiak, P. (1991). Learning invariance from transformation sequences. *Neural Computation, 3,* 194–200.

Freeman, W. T. (1994). The generic viewpoint assumption in a framework for visual perception. *Nature, 368,* 542–545.

Gibson, J. J. (1950). The perception of visible surfaces. *American Journal of Psychology, 63,* 367–384.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Hillsdale, NJ: Erlbaum.

Gibson, J. J., & Robinson, D. (1935). Orientation in visual perception: The recognition of familiar plane forms in differing orientations. *Psychology Monographs, 46*(6, Whole No. 210).

Gregory, R. L. (1970). *The intelligent eye.* London: Weidenfeld & Nicholson.

Grimson, W. E. L. (1982). A computational theory of visual surface interpolation. *Philosophical Transactions of the Royal Society of London, B, B298,* 395–427.

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review, 92,* 171–211.

Hayward, W. G. (1998). Effects of outline shape in recognition. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 427–440.

Hayward, W. G., & Tarr, M. J. (1997). Testing conditions for viewpoint invariance in object recognition. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 1511–1521.

Hayward, W. G., Tarr, M. J., & Corderoy, A. K. (1999). Recognizing silhouettes and shaded images across depth rotation. *Perception, 28,* 1197–1215.

Hilbert, D., & Cohn-Vossen, S. (1952). *Geometry and the imagination* (P. Nemeny, trans.) New York: Chelsea.

Hoffman, D. D., & Richards, W. (1984). Parts of recognition. *Cognition, 18,* 65–96.

Hoffman, D. D., & Singh, M. (1997). Salience of visual parts. *Cognition, 63,* 29–78.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99,* 480–517.

Julesz, B. (1971). *Foundations of cyclopean perception.* Chicago: University of Chicago Press.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23,* 141–221.

Kimia, B. B., Tannenbaum, A. R., & Zucker, S. W. (1995). Shapes, shocks, and deformations: I. The components of shape and the reaction-diffusion space. *International Journal of Computer Vision, 15,* 189–224.

Knill, D. C. (1992). Perception of surface contours and surface shape: From computation to psychophysics. *Journal of the Optical Society of America A, 9,* 1449–1464.

Knill, D. C., Kersten, D., & Mamassian, P. (1996). Implications of a Bayesian formulation of visual information for processing for psychophysics. In D. Knill & W. Richards (Eds.), *Perception as Bayesian inference* (pp. 239–286). Cambridge, United Kingdom: Cambridge University Press.

Knill, D. C., & Richards, W. (1996). *Perception as Bayesian inference.* Cambridge, United Kingdom: Cambridge University Press.

Koenderink, J. (1984). What does the occluding contour tell us about solid shape? *Perception, 13,* 321–330.

Koenderink, J. J., Kappers, A. M. L., Todd, J. T., & Norman, J. F. (1996). Surface range and attitude probing in stereoscopically presented dynamic scenes. *Journal of Experimental Psychology: Human Perception & Performance, 22,* 869–878.

Koenderink, J. J., & van Doorn, A. J. (1976). The singularities of the visual mapping. *Biological Cybernetics, 24,* 51–59.

Koenderink, J. J., & van Doorn, A. J. (1980). Photometric invariants related to solid shape. *Optica Acta 27,* 981–996.

Koenderink, J. J., & van Doorn, A. J. (1981). A description of the structure of visual images in terms of an ordered hierarchy of light and dark blobs. In *Proceedings of the Second International Visual Psychophysics and Medical Imaging Conference.* New York: Institute of Electrical and Electronics Engineers.

Koenderink, J. J., & van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America, A, 8,* 377–385.

Koenderink, J. J., van Doorn, A. J., & Christou, C. (1996). Shape constancy in pictorial relief. In J. Ponce, A. Zisserman, & M. Herbert (Eds.), *Object recognition in computer vision, II* (pp. 151–164). New York: Springer-Verlag.

Koenderink, J. J., van Doorn, A. J., Christou, C., & Lappin, J. (1996). Shape constancy in pictorial relief. *Perception 25,* 155–164.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. L. M. (1992). Surface perception in pictures. *Perception & Psychophysics, 52,* 487–496.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. L. M. (1994). On so-called paradoxical monocular stereoscopy. *Perception, 23,* 583–594.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. L. M. (1995). Depth relief. *Perception, 24,* 115–126.

Koenderink, J. J., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (1997). The visual contour in depth. *Perception & Psychophysics, 59,* 828–838.

Kovacs, I., Feher, A., & Julesz, B. (1998). Medial-point description of shape: A representation for action coding and its psychophysical correlates. *Vision Research, 38,* 2323–2333.

Kovacs, I., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature, 370,* 644–646.

Kristjansson, A., & Tse, P. (2001). Curvature discontinuities are cues for rapid shape analysis. *Perception & Psychophysics, 63,* 390–403.

Lee, T. S., Mumford, D., Romero, R., & Lamme, V. A. F. (1998). The role of the primary visual cortex in higher level vision. *Vision Research, 38,* 2429–2454.

Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: Changing views of perception. *Trends in Cognitive Sciences, 3,* 254–264.

Li, A., & Zaidi, Q. (2000). Perception of three-dimensional shape from texture is based on patterns of oriented energy. *Vision Research, 40,* 217–242.

Lowe, D. G. (1987). Three-dimensional object recognition from single two-dimensional images. *Artificial Intelligence, 31,* 355–395.

Marr, D. (1982). *Vision.* New York: Freeman.

Marr, D., & Nishihara, H. K. (1992). Visual information processing: Artificial intelligence and the sensorium of sight. In S. M. Kosslyn and R. A. Andersen (Eds.), *Frontiers in cognitive neuroscience* (pp. 165–186). Cambridge, MA: MIT Press. (Original work published 1978)

Marr, D., & Poggio, T. A. (1977). Cooperative computation of stereo disparity. *Science, 194,* 283–287.

McKee, S. P., Welch, L., Taylor, D. G., & Bowne, S. F. (1990). Finding the common bond: Stereoacuity and the other hyperacuities. *Vision Research, 30,* 879–891.

Mohr, R., & Bajcsy, R. (1983). Packing volumes by spheres. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 5,* 111–116.

Nackman, L. R. (1982). Curvature relations in three-dimensional symmetric axes. *Computer Graphics Image Processing, 20,* 43–57.

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science, 257,* 1357–1363.

Norman, J. F., Dawson, T. E., & Raines, S. R. (2000). The perception and

recognition of natural object shape from deforming and static shadows. *Perception, 29,* 135–148.

Paradiso, M. A., & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research, 31,* 1221–1236.

Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology, 82,* 2490–2502.

Pasupathy, A., & Connor, C. E. (2001). *Shape representation in area V4: Position-specific tuning for boundary conformation.* Manuscript submitted for publication.

Perrett, D., & Oram, M. W. (1993). Neurophysiology of shape processing. *Image and Vision Computing, 11,* 217–333.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize 3D objects. *Nature, 343,* 263–266.

Raines, S. R., & Norman, J. F. (1999). The perception of ordinal depth relationships from deforming boundary contours. *Investigative Opthalmology and Visual Science, 40,* 4.

Reichel, F. D., Todd, J. T., & Yilmaz, E. (1995). Visual discrimination of local surface depth and orientation. *Perception & Psychophysics, 57,* 1233–1240.

Richards, W. A., & Hoffman, D. D. (1985). Codon constraints on closed 2-D shapes. *Computer Vision Graphics and Image Processing, 32,* 265–281.

Richards, W. A., Jepson, A., & Feldman, J. (1996). Priors, preference, and categorical percepts. In D. Knill & W. Richards (Eds.), *Perception as Bayesian inference* (pp. 93–122). Cambridge, United Kingdom: Cambridge University Press.

Richards, W. A., Koenderink, J. J., & Hoffman, D. D. (1987). Inferring three-dimensional shapes from two-dimensional silhouettes. *Journal of the Optical Society of America, A, 4,* 1168–1175.

Rosenholtz, R., & Koenderink, J. (1996). Affine structure and photometry. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 790–795.

Shapiro, L., Zisserman, A., & Brady, M. (1995). 3D motion recovery via affine epipolar geometry. *International Journal of Computer Vision, 16,* 147–182.

Siddiqi, K., Kimia, B. B., Tannenbaum, A. R., & Zucker, S. W. (2001). On the psychophysics of the shape triangle. *Vision Research, 41,* 1153–1178.

Stevens, K. A. (1981). The visual interpretation of surface contours. In J. M. Brady (Ed.), *Computer vision* (pp. 47–74). Amsterdam: North-Holland.

Stevens, K. A. (1986). Inferring shape from contours across surfaces. In A. P. Pentland (Ed.), *From pixels to predicates* (pp. 93–110). Norwood, NJ: Ablex.

Takeichi, H., Nakazawa, H., Murakami, I., & Shimojo, S. (1995). The theory of the curvature-constraint line for amodal completion. *Perception, 24,* 373–389.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study of the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin and Review, 2,* 55–82.

Tarr, M. J., & Bülthoff, H. H. (1995). Is human object recognition better described by geon-structural-descriptions or by multiple-views? *Journal of Experimental Psychology: Human Perception and Performance, 21,* 1494–1505.

Tarr, M. J., & Kriegman, D. J. (2000). *Toward understanding human object recognition: Aspect graphs and view-based representations.* Manuscript submitted for publication.

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21,* 233–282.

Tarr, M. J., Williams, P., Hayward, W. G., & Gauthier, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience, 1,* 275–277.

Todd, J. T., & Akerstrom, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 242–255.

Todd, J. T., & Bressan, P. (1990). The perception of 3-dimensional affine structure from minimal apparent motion sequences. *Perception & Psychophysics, 48,* 419–430.

Todd, J. T., & Mingolla, E. (1983). Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 583–595.

Todd, J. T., & Norman, J. F. (1995). The visual discrimination of relative surface orientation. *Perception, 24,* 855–866.

Todd, J. T., & Oomes, A. H. J. (2001). *Three degenerate conditions for the perception of surface shape from texture.* Manuscript submitted for publication.

Todd, J. T., & Reichel, F. D. (1990). Visual perceptions of smoothly curved surfaces from double-projected contour patterns. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 665–674.

Tse, P. U. (1998). Illusory volumes from conformation. *Perception, 27,* 977–994.

Tse, P. U. (1999a). Complete mergeability and amodal completion. *Acta Psychologica, 102,* 165–201.

Tse, P. U. (1999b). A planar cut approach to volume recovery from silhouettes. *Investigative Ophthalmology & Visual Science, 40,* 4.

Tse, P. U. (1999c). Volume completion. *Cognitive Psychology, 39,* 37–68.

Tse, P. U. (2000). The sawtooth illusion. *Perception, 29,* 874–876.

Tse, P. U., & Albert, M. (1998). Amodal completion in absence of image tangent discontinuities. *Perception, 27,* 455–464.

Ullman, S. (1979). *The interpretation of visual motion.* Cambridge, MA: MIT Press.

Ullman, S., & Basri, R. (1991). Recognition by a linear combination of models. *IEEE Transactions Pattern Analysis Machine Intelligence 13,* 992–1006.

van Lier, R. (1999). Investigating global effects in visual occlusion: From a partly occluded square to the back of a tree-trunk. *Acta Psychologica, 102,* 203–220.

Wallis, G., & Bülthoff, H. (1999). Learning to recognize objects. *Trends in Cognitive Sciences, 3,* 22–31.

Wallis, G., & Rolls, E. T. (1997). A model of invariant object recognition in the visual system. *Progress in Neurobiology, 51,* 167–194.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der gestalt [Laws of organization in perceptual forms]. *Psychologische Forschung, 4,* 301–350.

Wouterlood, D., & Boselie, F. (1992). A good-continuation model of some occlusion phenomena. *Psychological Research/Psychologische Forschung, 54,* 267–277.

Zhu, S. C., & Yuille, A. L. (1996). FORMS: A flexible object recognition and modeling system. *International Journal of Computer Vision, 20,* 187–212.

Zucker, S. W., Dobbins, A., & Iverson, L. (1989). Two stages of curve detection suggest two styles of visual computation. *Neural Computation, 1,* 68–81.

Zucker, S. W., Dobbins, A., & Iverson, L. (1992). Connectionism and the computational neurobiology of curve detection. In S. Davis (Ed.), *Connectionism: Theory and practice. Vancouver studies in cognitive science, Vol. 3* (pp. 277–296). New York: Oxford University Press.

(Appendixes follow)

Appendix A
Detailed Explanation of Geometric Properties

1. *The projection of a planar cut corresponding to the visible portion of a cross-section through a volume will generically lack tangent discontinuities with the occluding contour projected from the rim of a volume with surfaces that are smooth everywhere.* A more general version of this claim (Tse & Albert, 1998) is the following: For a point at which the locus of intersection of two interpenetrating volumes meets the rim of either volume, the tangent lines of the projected contours will be identical at the image projection of that point. Because the line of sight of the observer grazes the surface at all points on the rim, the eye of the observer lies in the tangent plane to the surface at rim points. Therefore, in the projection to the retinal image, the tangent plane at a rim point collapses to a line L′, because the tangent plane is viewed edge on. Consider a differentiable curve D on the surface passing through a rim point R. The tangent line to D at R must project either to L′ or to a single image point on L′. In the latter case, the tangent line coincides with the observer's line of sight, implying that the observer has an accidental view. It follows that the tangents to two such curves D1 and D2 on the surface that intersect at R will both project from R onto L′ in the image, assuming a generic view. Because the locus of intersection of two smooth surfaces is a smooth curve contained in both surfaces, any junction of that locus with the differentiable curve specified by the rim itself will not generically project tangent or *first-order discontinuities* onto the retinal image. It follows as a special case that the surface contour, which is the image projection of a locus of intersection generated by a plane that intersects a volume, will lack tangent discontinuities where it meets the occluding contour projected from the rim of the volume. Similarly, a segment of occluding contour A′ that projects from a segment of rim that lies on a planar cut (of a volume that is everywhere differentiable) is first-order differentiable where it meets the occluding contour B′ that projects onto the image from the rest of the rim. That is, the first derivative is uniquely defined at points R1′ and R2′ where A′ meets B′. Indeed, the same reasoning implies that a surface contour that projects from any surface marking, reflectance boundary, or shadow will not generically have a tangent discontinuity with the occluding contour that projects from the rim.

2. *The image projection of a planar cut will generically exhibit abrupt curvature changes with the occluding contour projected from the rim.* Consider D1′ and D2′, the image projections of two curves D1 and D2 lying on the surface in the neighborhood of R, which is their point of intersection on the rim, that projects to R′. Note that the rate of curvature change per unit of contour distance for D1′ is independent of that for D2′ because D1 and D2 are themselves arbitrary and independent curves. Even if there are no sudden changes in the rate of curvature change per unit length of contour along either D1′ or D2′, there will generically be an abrupt curvature change at the image point R′ because of this independence. That is, the rate of curvature change will suddenly shift from D1′ to D2′ as one passes through R′ even though there is no contour tangent discontinuity at R′, as proven under Item 1. If there is no curvature discontinuity at R′, then the observer has an accidental view, or the planar cut includes the rim on both sides of R, which can be thought of as an accidental arrangement rather than an accidental view.

As a special case, the image projection of a planar cut will generically have a curvature discontinuity where it meets the image projection of the rim because the shape of a volume's cross-section can vary independently of the shape of the rim. Thus, with regard to the shape of a silhouette's outline, a segment of occluding contour A′ that projects onto the image from a segment of rim A that lies on a planar cut of the volume will generically meet the image contour segment B′ at R1′ and R2′ with curvature discontinuities.

In particular, let D1 coincide with the rim in the neighborhood of R and let D2 coincide with the locus of intersection lying on the surface resulting from a transverse planar cut. The suddenness with which the rate of curvature change shifts as one moves from D1′ to D2′ is enhanced by the fact that typically, as one approaches R′ from D2′, the image becomes increasingly compressed the closer one is to R′. For D1 that lies in the frontoparallel plane, this is because the surface normal for points along D2 becomes closer and closer to lying in the frontoparallel plane as one approaches R. That is, the surface normals tend toward perpendicularity with the line of sight as one approaches R. This follows because points of D2 within some finite neighborhood of R will lie closer to the observer the further one is from R for volumes whose surfaces are everywhere differentiable. Thus, the image projections of two points that are a fixed unit of distance apart on D2 will appear closer to one another on D2′ as the pair approaches R′. This will effectively accelerate the rate of curvature change for D2′ but not for D1′. As D1 deviates from lying in the frontoparallel plane, this effect of foreshortening will diminish. Although the rim is in no way bound to lie in the frontoparallel plane, a segment of rim that does not lie in the frontoparallel plane will project to a shorter contour in the image than a segment of rim of equal length that does lie in the frontoparallel plane. Thus, more occluding contour will tend to project from rim segments that lie on less slanted planes than lie on more highly slanted planes. In the absence of other image information implying otherwise, there may, therefore, be a bias toward placing D1 in or near the frontoparallel plane.

3. *A segment of occluding contour that lacks curvature discontinuities projects from a segment of rim lying on a plane.* Let us say a segment of occluding contour that lacks a curvature discontinuity projects from connected segments of rim that lie on two or more nonidentical planes. Assume that there are a finite number N of such planes and that the angle between any two of them is noninfinitesimal. These N rim segments will define a curve lying on the 3D surface that projects to the image. A slight change in viewpoint should give rise to a curvature discontinuity in the image for the contour that projects from this curve on the surface precisely because its segments lie on different planes, as proven before. Because we assume a nonaccidental viewpoint, it follows that a lack of curvature discontinuities in an image contour segment arises from a coplanar rim segment.

Note, however, that if N is allowed to go to infinity, then cases could arise in which a nonplanar segment of rim projects to a segment of occluding contour that lacks abrupt changes in curvature. This is because the planes corresponding to adjacent planar rim segments could differ by only an infinitesimal angle and no curvature discontinuities would arise in the image. However, because the resolution of the visual system is limited, it will not be able to detect curvature discontinuities that are smaller than some threshold. It will, therefore, not be able to distinguish nearly coplanar from coplanar rim segments, if the deviation from coplanarity is small. A corollary of the nonaccidental view assumption is that the visual system assumes that all information necessary to infer x is present in the image. In other words, if x is not detected in the image, the existence of x is not assumed. It follows from the assumption of nonaccidentality that segments of occluding contour that have subthreshold curvature discontinuities will be taken to lack curvature discontinuities. They will, therefore, be taken to project from a coplanar segment of rim.

4. *Let D coincide with the segment of rim on one side of a rim point R and let A be a segment of rim lying on a planar cut that intersects D at R. In general, the abruptness of the curvature change as one moves from D′ to A′ in the image decreases as the planes on which D and A lie approach one another.* If we only consider the portion of D in the immediate neighborhood of R, it can be estimated by a line segment. Assume that D lies on the least slanted plane that contains this line. Imagine that the planes that contain D and A are perpendicular to one another. As the plane that contains A approaches the plane that contains D, A will approach D. If the angle between these two planes approaches 180 degrees, A′ will increasingly appear to align with D′ in the neighborhood of R′. Therefore, the abruptness of the curvature change between D′ and A′ decreases. An example of this is shown in Figure A1a. As the abruptness of the curvature

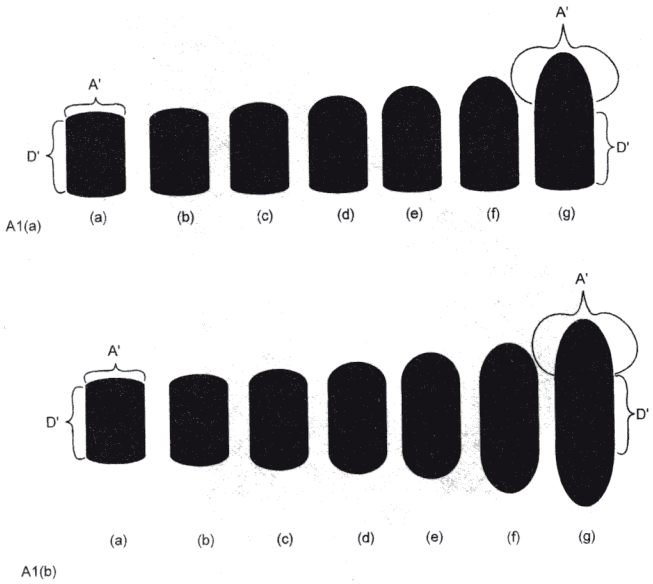


Figure A1. (a and b) The segment of rim that projects to A′ has increasingly less slant as it progresses from a to d. From e to g, it is placed in the frontoparallel plane. As the planes on which adjacent rim segments A and D converge, the abruptness of the curvature discontinuity as one goes from A′ and D′ in the image increases.

discontinuity between A′ and D′ decreases as we go from a to g, the plane on which A lies appears to approach the plane on which D lies. Beyond a certain point, the curvature discontinuity is no longer abrupt enough to indicate a transversal planar cut, and the silhouette begins to look either flat or, if still volumetric, rounded at its end such that A and D are approximately coplanar. Note, however, that the decrease in the abruptness of curvature change between D′ and A′ could also be due to a change in viewing angle rather than a change in shape. The parts of Figure A1a appear to change shape under constant viewing angle because the propagable segment of occluding contour at the base of the silhouette remains constant. This may imply that the slant of the ground plane with respect to the viewer remains constant. However, in Figure A1b, that also changes, so that, as we move from (a) to (d), we seem to be viewing a cylinder from different heights or a cylinder that is rotating toward us. Note that the visual system may be expressing a bias here not only to interpret ambiguous silhouettes in terms of familiar shapes, such as cylinders, but also to interpret shapes as resting at right angles to the inferred ground plane (cf. Richards et al., 1996; Tse, 2000). Beyond a certain point, however, the degree of abruptness in curvature change between D′ and A′ decreases to such an extent that the silhouette is no longer consistent with a cylinder interpretation. This happens beyond (d) when A′ is best fit with an ellipse that is circular. As described in the discussion of Property 5, a circular A′ would tend to imply an A that lies in the frontoparallel plane. However, if this were the case, then the D would have to be foreshortened and D′ would be relatively short as in Figure A1. Because this is not the case here, the favored interpretation is a "pill." These pill-shaped silhouettes look either flat or like volumetric pills whose A and D lie in approximately the same plane.

5. *A propagable segment of image contour can be propagated away from its image location in such a way that it will continue to lack tangent discontinuities with the occluding contour segment B′ and will continue to have curvature discontinuities with B′.* What follows is the core of the propagation or filling-in algorithm. Let A be a segment of rim that lies on a plane whose image projection A′ has curvature but not tangent discontinuities with B′, where B′ is the image projection of the rest of the rim apart from A. A′ meets B′ at two points, R1′ and R2′. In the image, construct a copy of A′ and call it A*′. At first A*′ exactly overlaps A′. Now displace or propagate A*′ while keeping A′ in place, as shown in Figure A2a. Without rotating A*′ displace its endpoint E1*′ (which was at R1′) a small distance Δs along B′ away from R1′. "Without rotation" here means the following: Take the two endpoints E1′ and E2′ on A′ and connect them with a line L′. After displacement of A*′, the corresponding endpoints E1*′ and E2*′ of A*′ can be connected by a line L*′ that is parallel to L′. After displacement, A*′ now meets B′ at at least one point R1*′ a distance Δs from R1′. The endpoint of A*′ that was not displaced along B′ will lie in the interior of the silhouette, on B′, or outside of the silhouette. If this endpoint E2*′ does not already lie on B′, expand or contract A*′ uniformly without rotation until this endpoint lies on B′. This endpoint will lie on B′ at a point R2*′ a distance Δk from R2′, as shown in Figure A1b. The distance Δs need not be the same as Δk because the orientations of B′ in the neighborhood of R1′ or R2′ may differ. In Figure A2a A*′ does not need rescaling to touch B′ at both of its endpoints. However, in Figure A2c, A*′ extends outside of the silhouette and needs to be rescaled in order to yield the prop(A′) shown in Figure A2d.

If we had displaced A*′ from A′ along B′ by moving E2*′ to R2*′ instead of displacing E1*′ to R1*′, and expanded or contracted A*′ until the endpoint E1*′ met B′, E1*′ would still meet B′ at R1*′ because it must move along L*′ as it contracts or expands, and L*′ intersects B′ at only R1*′ and R2*′. This algorithm therefore has a unique solution.

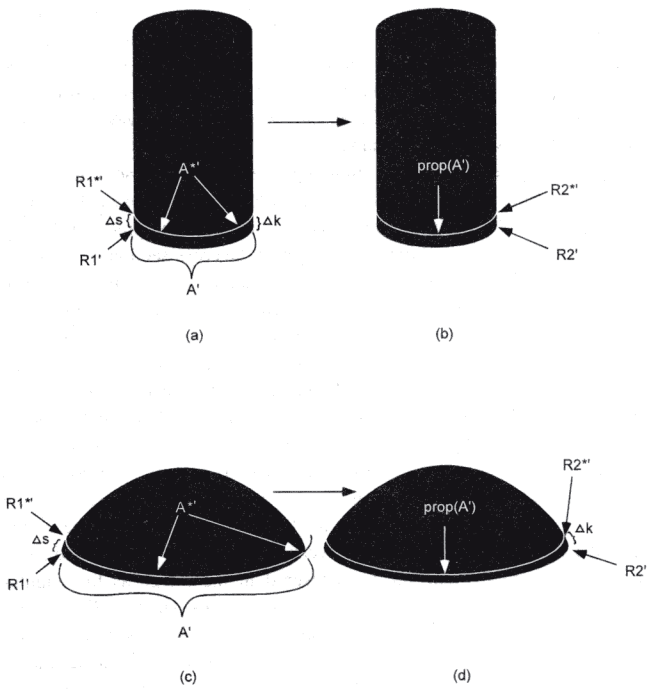


Figure A2. The A*′ in Figure A2a exactly coincides with prop(A′) in A2b, but the A*′ in A2c must be scaled to generate the prop(A′) shown in A2d. prop = propagation function.

(*Appendixes continue*)

Define a function "prop" such that prop(A′) equals the expanded or contracted version of A*′. Propagable segments of image contour A′ have the following property: If A′ shared no tangent discontinuities with B′ at R1′ and R2′, prop(A′) will not share tangent discontinuities with B′ at R1*′ and R2*′. Similarly, if A′ shared curvature discontinuities with B′ at R1′ and R2′, prop(A′) will share curvature discontinuities with B′ at R1*′ and R2*′. If prop(A′) has tangent discontinuities or lacks curvature discontinuities with B′, A′ is not a propagable segment of contour, at least for the special class of silhouettes under discussion now. This follows from Items 1 and 2 because prop(A′), like A′, is supposed to be the image projection of a planar cut of a volume with surfaces that are differentiable everywhere. It should, therefore, also have all the defining properties of propagability described in these claims. One can apply the prop function iteratively such that the next displacement of A′ after prop(A′) would be given by prop(prop(A′)) and so forth. If Δs is allowed to go the limit of ds, then a smooth curved surface can be inferred from the integration of all prop(A′). If at some point propagation is no longer possible, then A′ will only be partially propagable into the interior of the silhouette.

6. *A propagable segment cannot provide either metric slant or tilt information about the plane from which it projects onto the image.* A circle lying flat on a plane will project to an ellipse in the image whose major axis lies along a line whose angle relative to vertical in the image gives the tilt of the plane. This is because the ellipse results from foreshortening of the circle in a direction perpendicular to the tilt of the plane. Fit A′ (or prop(A′)) with an arc of an ellipse. If we assume that A′ projects from a circular cross-section, the major axis of the fitted ellipse will yield precise tilt information about the plane on which the cross-section lies. For arbitrary cross-sections, however, the tilt of the plane on which the cross-section lies is only ambiguously specified by A′. If the cross-section is elongated rather than circular, the angle of the major axis of the ellipse that is fit to A′ in the image will not generically yield precise tilt information. The visual system is, therefore, confronted with an ambiguity in interpreting the causes of an image. It can either attribute the shape of A′ to the shape of A or to compression as a result of foreshortening (or some combination of both causes). Given only A′, there is no way of knowing which factor has caused the shape of A′.

Similarly, given A′, slant is ambiguous because an ellipse in the image can project from multiple ellipses in the world placed on planes of arbitrary slant. Because the shape of the cross-section cannot be precisely known given only A′, the slant of the plane on which the cross-section lies can also not be precisely specified. Thus, a given A′ cannot provide metric slant or tilt information about the plane from which it projects onto the image.

7. *A′ and prop(A′) can offer relative slant, tilt, and surface curvature information about the surface of a volume under certain assumptions.* Consider a planar cut A* of a volume that does not lie on the volume's rim, as shown in Figure A3. A* lies on a plane P_{A^*}. The image projection of A* is A*′, a curve lying inside the corresponding silhouette, as shown in Figure A4. (A*′ need not be an instance of prop(A′)). P_{A^*} intersects two rim points R1* and R2*. The tilt of the plane P_{R1^*} that is tangent to the volume's surface at R1* containing the line of sight is known because the line of sight lies in both P_{R1^*} and the vertical plane. P_{R1^*} projects to the line $L_{R1^{*\prime}}$ in the image. Tilt can be measured in the image as the angular difference between vertical and $L_{R1^{*\prime}}$. Moreover, the slant of P_{R1^*} is zero degrees (i.e., perpendicular to the frontoparallel plane), assuming orthographic projection. Similarly, P_{R2^*} projects to a line $L_{R2^{*\prime}}$ in the image. Thus, the slant and tilt of surface tangents to points on the rim are precisely specifiable from image information. Note that although the slant of P_{R1^*} and P_{R2^*} is 0 degrees in both cases their tilts may be different.

Can the precisely known slant and tilt of P_{R1^*} and P_{R2^*} be used to estimate slant and tilt information for points on A* between R1* and R2*? The answer is no for the general case because neither the precise tilt nor slant of P_{A^*} can be determined from the image. For the general case in which R1* and R2* lie at different depths, the precise tilt cannot be known for P_{A^*} because the line L′ can be the same in the image for different slants

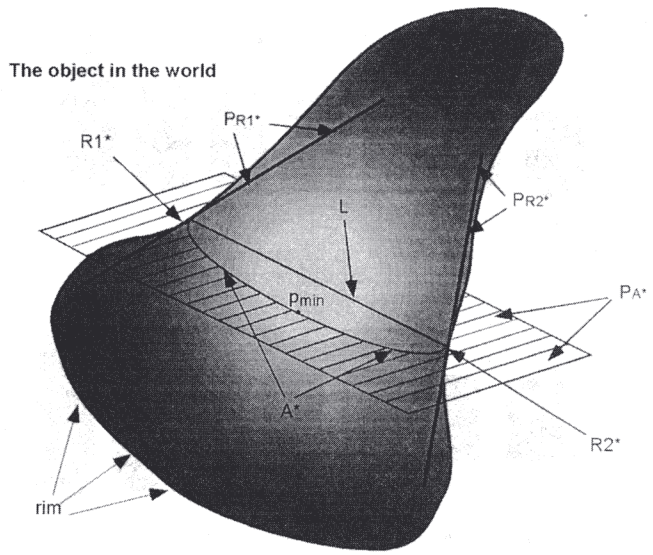


Figure A3. Plane P_{A^*} slices through the object in the world as shown. The line of sight lies in planes P_{R1^*} and P_{R2^*}. Labeled quantities are discussed in the text.

of P_{A^*}. (P_{A^*} must pass through the line L connecting R1* and R2*, and this line projects to L′.) For example, L′ might be horizontal because R1* and R2* lie at the same depth and height, or L′ might be horizontal because R1* and R2* lie at different depths and different heights.

The situation is more tractable for the special case in which R1* and R2* are assumed or known to lie at the same depth. Consider the angle between P_{A^*} and P_{R1^*}. At R1* this angle is ø1*. The angle ø1* can be measured in the image as the angle ø1*′ between L′ and $L_{R1^{*\prime}}$, where L′ is the line connecting R1*′ and R2*′. At R2* the corresponding angle is ø2*. The angle ø between the tangent plane to the volume's surface (at points along A*) and P_{A^*} can be assumed to vary smoothly between ø1* and ø2* in the absence of image cues to the contrary. Thus, ø can be generated at each point on A* by some weighted average of ø1* and ø2*. The angle for some point p on A* that lies x% of the way between R1* and R2* along A* might be given, for example, by a formula such as $øp^* = ((100 - x)ø1^* + (x)ø2^*)/100$. This would amount to translating the tilt of P_{R1^*} or P_{R2^*} into slant. Relative slant and tilt information could be estimated in various ways. As the tangent to A*′ approaches the extreme given by $P_{R1^{*\prime}}$, for example, the corresponding surface normal to A* would approach zero slant and the tilt of P_{R1^*}. Also, because L is presumed to lie in the frontoparallel plane for this special case, the point on A* closest to the observer would be the one farthest from L. This would project to the point on A*′ that is farthest away from L′. These points are shown as points $p_{\min}$ and $p_{\min}{}'$ in Figures A3 and A4. They correspond to points at which the surface is locally least slanted. In the image, the tangent to $p_{\min}{}'$ will have the slope of L′. There could be several points along A*′ with this tangent, each counting as a local point of least slant.

The problem of recovering relative surface curvature information along A* is comparably much simpler than recovering relative slant and tilt in the general case in which R1* and R2* do not lie at a common depth. Relative surface curvature information along A* would be revealed by the relative sharpness of contour curvature along A*′. Within a small neighborhood, the contour curvature introduced by foreshortening could be ignored, and points on A*′ with high contour curvature would correspond to points on A* with high surface curvature.

If it is true that shape is coded relationally, then two objects that have the same shape relationships or ratios over relevant dimensions should appear to have the same shape even when their metric or absolute values of size,

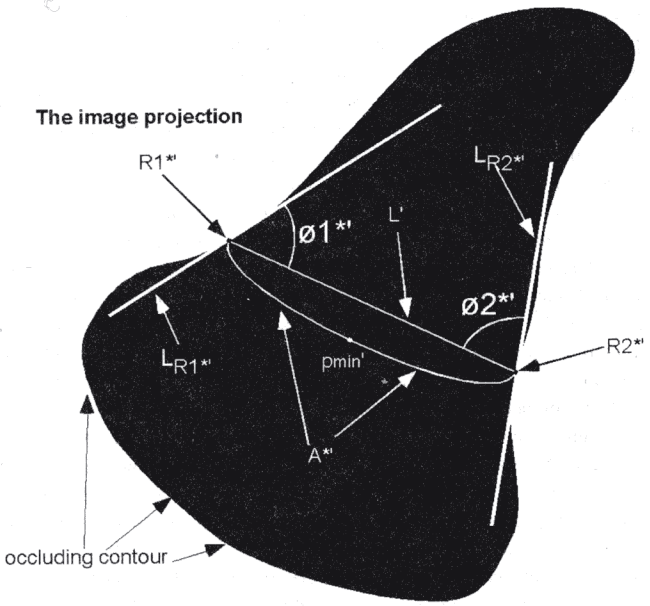


Figure A4. This figure corresponds to the image projection of Figure A3. Labeled quantities are discussed in the text.

slant, and so forth differ radically. A common illusion demonstrating that this is so is a cameo or relief such as Lincoln's face on a penny. Even though a cameo has completely different metric values of orientation and curvature along its curved surfaces than a real face has in profile, the important slant, tilt, size, and angle relationships are preserved, and it is this fact that makes cameos seem so realistic. Belhumeur, Kriegman, and Yuille (1999) proved that no visual system can recover metric information about a surface from its shading or shadows in principle. This is because any linear scaling of, for example, a relief's surface along the line of sight (plus a plane) will result in a surface that is capable of casting the same image as the original unscaled relief, assuming orthographic projection and a movable light source. Thus, any visual system, animal or machine, can only recover the shape of a surface from shadows or shading up to a scale factor in the direction of the line of sight. This result implies that any coding of shape-from-shading or shape-from-shadows must be nonmetric. Not surprisingly, then, psychophysical evidence indicates that the human visual system only recovers relevant shape relationships rather than precise surface depths and orientations (Koenderink, van Doorn, & Christou, 1996). Although the nature of such a shape code must be nonmetric, it is not clear what the nature of the nonmetric representation used by the human visual system might be. The three main contenders are (a) an affine representation (i.e., those in which affine transformations are equivalent; Koenderink & van Doorn, 1991; Rosenholtz & Koenderink, 1996; Shapiro, Zisserman, & Brady, 1995; Ullman & Basri, 1991); (b) a projective representation (Faugeras, 1995); and (c) an ordinal representation (Fermuller & Aloimonos, 1996).

A relational rather than metric coding of shape would also underlie the fact that objects do not typically seem to change their shape with distance, rotation, or scaling, even though the retinal image will change size under such operations. An example of the type of relationship that could underlie translation, rotation, and size constancy is the angle subtended by pairs of contour tangents to the occluding contour in the image. The angle between two contour tangents will be the angle between the two projecting surface tangent planes at corresponding rim points, because the line of sight lies in the tangent plane to the surface for rim points. For example, the angle subtended between the two tangents to the silhouette shown in Figure A5 would remain constant regardless of distance, rotation, or size assuming orthographic projection and assuming that the corresponding points on the rim of the object remain the same. (Even assuming a perspective projection, the rim points would tend to move short distances over a volume's surface relative to changes in object distance or size. Thus, even under perspective projection, relative angles would in most cases change only very little with scaling or distance of the object that casts the silhouette.) Figure A5b might be the image projection of the same object that projects to A5a if we were to move it further away or uniformly shrink it. Note that the angle between the two tangents remains constant despite the change in size from Figure A5a to A5b. Relations among occluding contours, therefore, reveal invariant information about surface tangent relationships regardless of size or distance. Once surface tangent information is available for points at the rim, the visual system may link known surface tangents in the smoothest possible way.

8. *Let C′ be a contour segment between but not including two inflection points I1′ and I2′ that has positive contour curvature everywhere. There will be a planar cut G that projects to a curve G′, which links two points on C′ such that G′ has no inflection points.* Because C′ has positive contour curvature everywhere, it has no inflection points. C′ is the projection of a segment of rim C. If a region of negative surface curvature N crossed C, there would be inflection points along C′ because the sign of contour curvature corresponds to the sign of surface curvature at the rim (Koenderink, 1984). Because C′ has positive curvature everywhere, there is no region N that meets C. Assuming a nonaccidental view, there must be a noninfinitesimal distance Δx separating any region N and C. Any planar cut G that passes through two points on C that does not pass through N will only pass through regions of visible surface that have positive surface curvature. (Because we are assuming a nonaccidental view, we also assume that there are no islands of concavity or regions of negative curvature that do not meet the rim and reveal themselves by inflection points along the occluding contour.) One such planar cut is given by the cross-section that is defined by the following three points: two points on C (C1 in the neighborhood of I1, and C2 in the neighborhood of I2 that project to C1′ and C2′ in the neighborhood of I1′ and I2′, respectively) and one visible nonrim point P on the surface between C and N that projects to P′, as shown in Figure A6. G′ will generically have inflection points when G does. Because all the curvature of G defined by this plane is due to the curvature of the underlying surface, and the underlying surface has a constant positive sign of curvature, G will not have inflection points. Therefore, G′ will not have inflection points.

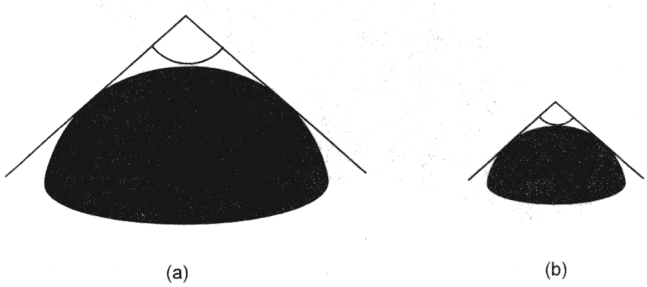


Figure A5. The angle between pairs of tangents to the occluding contour is the same as the angle between corresponding tangent planes at the rim.

(*Appendixes continue*)

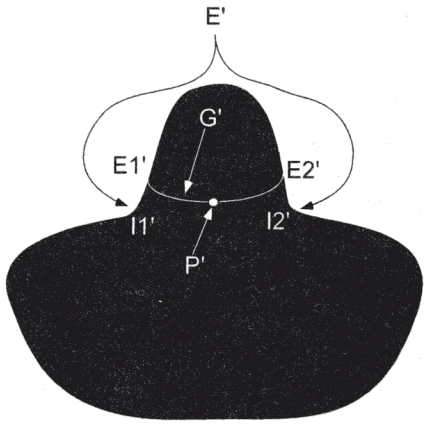


Figure A6. G′ lacks inflection points because the corresponding planar cut G only passes through regions of positive surface curvature.

A planar cut projecting onto a surface contour in the interior of a silhouette that satisfies the prior conditions will not have inflection points. If the plane defining the planar cut passes above the line of sight, the surface contour in the image will appear to bulge upward from points C1′ and C2′. If the plane passes below the line of sight, it will appear to bulge downward from C1′ and C2′, as shown in Figure A7. Of course, if the plane included the line of sight, the surface contour that projects from the planar cut would be a line segment in the image, but that would be an accidental view. This is why the top-most surface contour in Figure 3b has no inflection points even though the prop(A′) in the neighborhood of A′ do have inflection points. As the prop(A′) propagate, they interact with the rest of the occluding contour B′ in such a way that prop(A′) deforms to be consistent with the surface curvature implied by B′.

9. *Let C′ be a contour segment between but not including two inflection points I1′ and I2′ that contains two and only two inflection points I3′ and I4′. Moreover, let the curvature between I3′ and I4′ be negative. There will be a planar cut G that projects to a surface contour G′, which links two points on C′ such that G′ has at least two inflection points.* Because the curvature between I3′ and I4′ is negative, the corresponding surface curvature at the rim must be negative. Because the curvature between I1′ and I3′ (or between I2′ and I4′) is positive, the corresponding surface curvature at the rim must also be positive. A region of negative surface curvature N must cross the rim segment that lies between the points I3 and I4. Assuming a nonaccidental view, N will continue away from the rim some noninfinitesimal distance Δx. Consider the cross-section defined by the following three points: two points on C (C1 in the neighborhood of I1 and C2 in the neighborhood of I2 that project to C1′ and C2′ in the neighborhood of I1′ and I2′, respectively) and one visible nonrim point P on the surface within N, as shown in Figure A8. The planar cut G given by this cross-section will pass through a region of positive surface curvature, then through N, and then again through a region of positive surface curvature. Any point in N will have one direction of positive principal curvature and one direction of negative principal curvature. G can be chosen so that it will have inflection points if some of its tangents lie in the direction of negative principal curvature. G′ will, therefore, have at least two inflection points corresponding to the points at which G crosses from N into the regions of positive surface curvature outside of N.

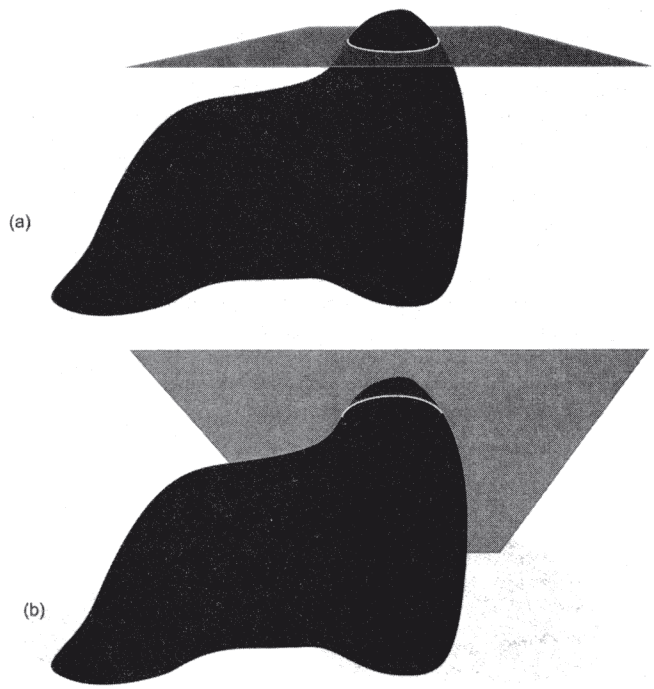


Figure A7. (a and b) If the intersecting plane passes below (above) the eyes of the viewer, the image projection of the planar cut will bulge downward (upward).

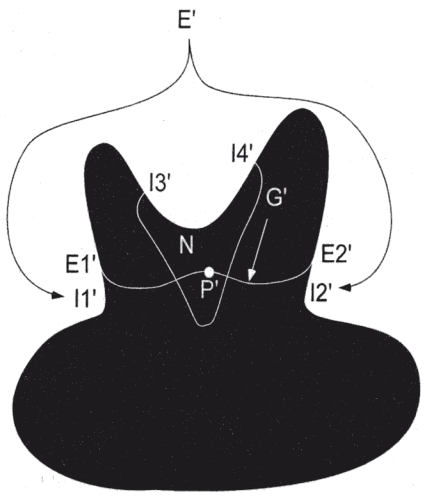


Figure A8. G′ has at least two inflection points because the planar cut G has inflection points where it crosses from regions of positive surface curvature to negative.

Appendix B

Glossary of Terms

Abrupt curvature changes: Curvature changes along an occluding contour revealed by discontinuous or nearly discontinuous tangents to the contour.

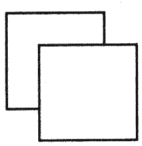
Contour curvature: The first derivative along an image contour at a point P on the contour is the tangent to the contour at P. The second derivative is the curvature of the contour at P and is given by the inverse of the radius of a circle that has the same curvature at P as the contour, when the contour and circle are traversed in the same sense. Let us say that the interior of a silhouette lies on our left side as we move along its bounding contour. Segments of contour where the next point that we traverse lies to the left (right) of the instantaneous tangent of the point where we are now have positive (negative) contour curvature. Inflection points along a smooth contour have zero curvature and separate segments with positive and negative contour curvature.

Curvature discontinuity: See "second-order discontinuity."

First-order discontinuity: A first-order discontinuity along a contour occurs at a point at which the first derivative or tangent is not well defined. For example, there is no unique tangent at a corner along a contour.

Geodesic: The shortest path along a surface that connects two points on that surface.

Nonaccidentality: A nonaccidental or "generic" view is one for which a slight shift of viewpoint in an arbitrary direction will not lead to a radical change in image structure. Assuming such a viewpoint, a straight (curved) line in the image projects from a straight (curved) line segment of rim in the world (cf. Grimson, 1982; Lowe, 1987). Instead of talking about accidental views, it would be better to talk of accidental images because accidental alignments can result from an accidental view or an accidental arrangement of objects in the world. An accidental arrangement is one for which a slight rearrangement of the objects would result in an image with radically different spatial or topological relationships. For example, if the image below results from a large square partially occluding another square, slightly moving the objects in the world will result in more or less the same image. However, if this image is due to a large square abutting an L shape at the same depth plane, then a slight rearrangement of the objects in the world would result in an image that cannot be mistaken for an image of two squares.



Occluding contour: Occluding contours are the self-occlusion contours projected onto the image from the "rim" of an object. The rim lying on the surface of a volume, relative to a particular viewpoint, is the set of points at which the observer's line of sight grazes the surface. For a volume with a smooth surface (i.e., differentiable as many times as necessary), the rim is composed of a set of smooth curves and loops that divide the visible and self-occluded surfaces of the volume. Except for special cases like spheres, the rim almost never lies in a common plane in the world. Occluding contours, however, lie entirely in the image plane.

Planar cut: The intersection of a plane and a volume is that volume's cross-section. A planar cut is the loop or set of closed nonintersecting loops lying on the surface of a volume defined by the outline of the cross-section. In this article, I only consider the visible portions of planar cuts and their image projections.

Principal curvature: See "surface curvature."

Prior: A prior assumption about the structure of the world that biases how inherently ambiguous image data are interpreted.

Rim: See "occluding contour."

Second-order discontinuity: A second-order discontinuity is a point along a contour where the second derivative is not defined. In this article, the focus is on abrupt curvature changes where the curvature changes discretely on either side of a contour point.

Slant: The slant of a plane is the angle (between 0° and 90°) that its surface normal makes with the line of sight. Without changing its slant, a plane can be rotated between 0 and 360 degrees around the line of sight. This will vary its tilt.

Surface contour: A surface contour projects from a visible curve, boundary, scratch, or other discontinuity on a surface. A surface contour exists in the image, whereas the curve on the surface of an object that projects to a surface contour exists in the world.

Surface curvature: For any point on a smooth nonplanar surface, there is a direction where the surface curves the most and one where it curves the least. These directions of principal curvature are always perpendicular, and the product of these two curvatures is the Gaussian or surface curvature of the surface at that point (Hilbert & Cohn-Vossen, 1952). A point on a surface will either have positive, negative, or zero curvature.

Tilt: See "slant."

Received July 30, 1999
Revision received May 1, 2001
Accepted June 12, 2001 ■