International Journal of Computer Vision, 12:2/3, 113–136 (1994)

Recovering Shape by Purposive Viewpoint Adjustment

KIRIAKOS N. KUTULAKOS AND CHARLES R. DYER
Computer Sciences Department, University of Wisconsin–Madison, 1210 W. Dayton St., Madison, WI 53706
KYROS@CS.WISC.EDU
DYER@CS.WISC.EDU

Received June 29, 1992; revised June 1, 1993.

Abstract
We present an approach for recovering surface shape from the occluding contour using an active (i.e., moving) observer. It is based on a relation between the geometries of a surface in a scene and its occluding contour: If the viewing direction of the observer is along a principal direction for a surface point whose projection is on the contour, surface shape (i.e., curvature) at the surface point can be recovered from the contour. Unlike previous approaches for recovering shape from the occluding contour, we use an observer that *purposefully* changes viewpoint in order to achieve a well-defined geometric relationship with respect to a 3-D shape prior to its recognition. We show that there is a simple and efficient viewing strategy that allows the observer to align the viewing direction with one of the two principal directions for a point on the surface. This strategy depends on only curvature measurements on the occluding contour and therefore demonstrates that recovering quantitative shape information from the contour does not require knowledge of the velocities or accelerations of the observer. Experimental results demonstrate that our method can be easily implemented and can provide reliable shape information from the occluding contour.

1 Introduction

There has been considerable interest in recovering information about the structure of a scene from sequences of images, assuming an observer in motion—e.g., work on optical flow (Horn 1986) and shape-from-motion (Aloimonos et al. 1987). One common feature of these approaches is the use of known viewer motion in order to recover quantitative properties of the scene such as surface curvature (Cipolla & Blake 1992). Recently, however, there has been considerable interest in employing simple observer behaviors that either make the recovery of scene properties easier and more efficient (Ballard 1989, 1991; Ballard & Brown 1992; Krotkov & Bajcsy 1993; Wixson & Ballard 1993; Rimey & Brown 1993), or combine simple behaviors in order to perform complex tasks such as tracking, navigation, and obstacle avoidance (Aloimonos 1990; Coombs & Brown 1993; Papanikolopoulos et al. 1993; Grosso & Ballard 1993; Nelson & Aloimonos 1989; Brooks 1986, 1989). These approaches rely on maintaining specific geometric relationships between the observer and the environment. However, without knowledge of viewer motions they only recover qualitative information about the viewed scene—e.g., relative-depth ordering with respect to the fixated scene point (Ballard & Ozcandarli 1988).

This article presents a new approach which combines the above two paradigms in order to recover surface curvature from the occluding contour. We show that shape for a selected point on the viewed surface can be recovered from the occluding contour of two views. The only requirements are that (1) the surface point is projected on the occluding contour in both views, and (2) the viewing direction of the observer for one of the views has a specific relationship with the surface geometry at the selected point. The main idea of our approach is to use an active (i.e., moving) observer that purposefully changes viewpoint in order to achieve such a well-defined geometric relationship with respect to a 3-D shape prior to its recognition. We show that this relationship is characterized by specific image-computable quantities and enables an analysis similar to the one by Krotkov (1987). In addition, our approach does not require any knowledge of observer velocities or object models, and assumes the use of a world-centered coordinate frame (Ballard 1989).

It is well known that the occluding contour is a valuable source of information about surface shape (Barrow & Tenenbaum 1981; Giblin & Weiss 1987; Koenderink 1984; Marr & Nishihara 1978; Ulupinar & Nevatia 1988). The occluding contour is the projection of the one-dimensional set of points separating the visible from the hidden parts of the surface. It is also defined as the projection of the visible rim, the one-dimensional set of visible surface points at which the line of sight of the observer is tangent. There have been several approaches to deriving information about surface geometry from the occluding contour. These approaches are based on three important properties of the contour's geometry:

1. The geometry of the occluding contour is surface-dependent.
2. The occluding contour is the projection of a limited set of surface points.
3. The geometry of the occluding contour is viewpoint-dependent.

The dependency of the occluding contour's geometry on the surface has been investigated by several researchers. It has been shown that the geometry of the occluding contour severely constrains the underlying surface geometry (Barrow & Tenenbaum 1981; Giblin & Weiss 1987; Koenderink 1984; Richards et al. 1988). The major problem, though, is that the occluding contour is created through a projection process and, in general, the information it provides is ambiguous, that is, several different surface rims can project to the same occluding contour. This has been the major motivation for using the occluding contour to derive qualitative rather than quantitative information about the surface shape (Koenderink 1984; Richards et al. 1988; Leyton 1988; Malik 1987). On the other hand, approaches that derive quantitative shape descriptions from the occluding contour assume that additional information is available in order to adequately constrain the shape-recovery process (Cipolla & Blake 1992; Giblin & Weiss 1987). Such approaches require other means for deriving the additional shape information. For example, Cipolla and Blake (Cipolla & Blake 1992) used a moving observer to change viewpoint and measured the relative accelerations of image features near the occluding contour in order to measure surface curvature. Even though their approach was not very sensitive to errors in viewer motion, they assumed the existence of image features near the occluding contour and knowledge of the observer's translational velocity.

The second property of the occluding contour suggests that it can provide only a limited amount of information about the complete shape of the surface. Indeed, only a one-dimensional set of surface points projects to the occluding contour. If the only information available is the occluding contour of the surface from a particular viewpoint, we can at best derive only a qualitative description of the entire surface. For example, Barrow and Tenenbaum (1981) attempted to constrain the recovery problem using additional smoothness assumptions about the surface, but these assumptions do not hold in general.

The dependency of the occluding contour on viewpoint has been used to resolve both of the above ambiguities. A slight change in viewpoint will affect the geometry (i.e., curvature) and possibly the topology of the rim, and hence the occluding contour. Moreover, the set of rim points changes and therefore new constraining information about the surface shape becomes available. It has been shown that if we know how the geometry of the occluding contour changes with viewpoint, we can derive a parameterization of the surface and determine its shape (Giblin & Weiss 1987). Furthermore, algorithms for tracking such contours over a sequence of images are becoming increasingly more sophisticated (Blake et al. 1993). The issue here is how to accurately measure such changes in the contour's geometry with small viewpoint changes. For example, we must be able to measure the velocity and acceleration of surface points entering and leaving the rim (Cipolla & Blake 1992; Vaillant & Faugeras 1992), a problem that requires first- and second-order differentiation operations, and hence is sensitive to noise.

The basic assumption used by all of the above approaches was that the viewpoint is *arbitrary*. This means that the viewpoint is not related in any way to the geometry of the rim or the occluding contour. (For example, there are surfaces for which their rim is planar when viewed from a particular set of directions (Marr & Nishihara 1978).) This is a reasonable approach, however, only if the observer cannot control viewpoint. When the observer has the ability to control the viewing direction, the choice of viewpoint(s) does not have to be arbitrary. Our approach uses an active observer to obtain a view based on the observed object's geometry in order to recover exact shape information from the occluding contour.

1.1 Active Shape Recovery

Our goal is to actively derive a quantitative shape description for surface points in the vicinity of the rim. We accomplish this goal by using properties of the occluding contour. The basic step of our approach involves selecting a point on the rim and recovering the surface shape (i.e., principal curvatures and principal directions) at that point. In addition, we present a strategy for applying this shape-recovery step to neighboring surface points. The surface description is therefore incrementally extended by successively including new points on the rim and recovering the surface geometry for those points.

The main step of our approach is based on a relation between the geometries of a surface in a scene and its occluding contour: If the viewing direction of the observer is along a principal direction for a selected surface point whose projection is on the contour, the corresponding principal curvature at the point can be recovered. Hence, even though, in general, surface curvature computation from the occluding contour of a single view is an underconstrained problem, for any given point there do exist viewing directions that make this recovery problem well defined. If the observer can move to one of those special viewing directions, the ambiguities caused by the projection process can be resolved. We show that the observer can in fact deterministically find these special viewing directions by simply maximizing or minimizing a geometric quantity of the occluding contour (curvature at a point) while changing viewing direction in a constrained way. Furthermore, we show that we can recover the shape of the surface at the selected point (i.e., both principal curvatures) from the occluding contour of one additional view for which the selected point is projected onto the contour. Thus an active observer selects a point on the surface rim and purposefully moves to one of the special viewpoints in order to make shape recovery a well-defined problem.

The significance of our method lies in the use of purposive observer motion to achieve and maintain purely geometrical relations between a surface and its occluding contour in order to recover surface shape. Hence, there is no need to perform any velocity or acceleration measurements in the vicinity of the rim, a process requiring point-to-point correspondences in the images and precise knowledge of viewer motion. Furthermore, since there is a well-defined procedure to reach the desired viewpoint, the observer does not need to perform a complicated search in order to find it (Bajcsy 1988; Hager & Mintz 1987).

Even though our approach is limited to the recovery of surface shape in the vicinity of a single point on the rim, we show that there is an important special case for surfaces of revolution, for which we can derive shape information for the complete set of rim points. In this case the observer actively "aligns" itself with the viewed surface in order to find a viewpoint giving complete surface information (i.e., one perpendicular to the surface's axis of rotation).

We also present an extension to the above approach that recovers the shape of points in the vicinity of the rim. After the shape of a selected rim point is recovered, the observer changes viewpoint in order to bring a new surface point onto the rim and to recover its shape. Since our basic shape-recovery step involves aligning the observer's viewing direction with one of the principal directions at the new point, it is important for this visual alignment process to require only small viewpoint adjustments. We show that if (1) the new point selected is in the normal plane of the previously selected point, and (2) the new point is sufficiently close to the previously selected point, these adjustments will in fact be small and their extent will depend entirely on the intrinsic properties of the surface. This is a major difference from approaches using "passive" motion, where the points selected for reconstruction cannot be controlled.

The rest of this article is organized as follows. The next section reviews basic terminology. Section 3 discusses the relation between the geometries of the occluding contour and the surface, and presents the major result enabling us to actively recover surface geometry from the occluding contour. Section 4 uses this result to describe the main shape-recovery step of our approach. Our results are then extended in the following two sections. Section 5 discusses shape recovery for the case of surfaces of revolution, and section 6 describes the viewing strategy used to select a new point for shape recovery. Finally, section 7 presents experimental results on synthetic and real images to demonstrate the applicability of our theoretical results.

2 Viewing Geometry

Let S be a smooth, oriented surface in $\Re^3$, viewed under orthographic projection along a viewing direction ξ. Viewing directions can be thought of as points on the unit sphere S^2, with s and t being the slant and tilt of the camera respectively (figure 1).

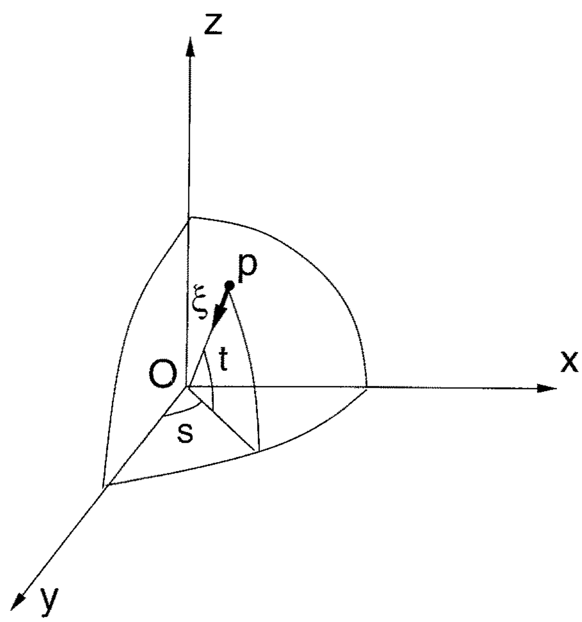


Fig. 1. Viewing direction ξ is represented as the point $p = (\cos t \cos s, \cos t \sin s, \sin t)$ with $s \in [0, 2\pi)$, $t \in [0, \pi)$.

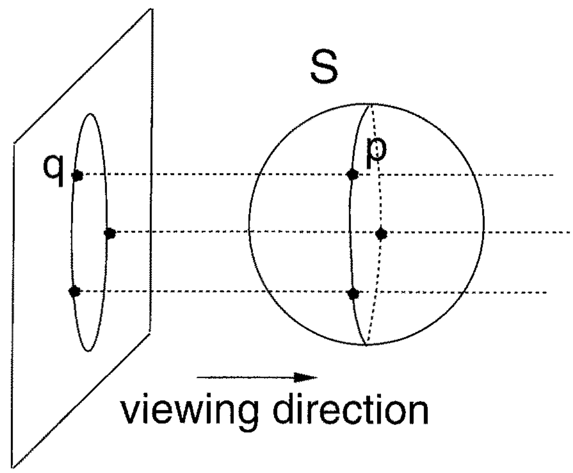


Fig. 2. Point p on the rim of the sphere S is projected to the occluding contour point q on the image plane.

Let x be a parameterization of S and $p = x(u, v)$ be a point on S. The partial derivatives $x_u(p)$, $x_v(p)$ of x with respect to u and v define $T_p(S)$, the plane tangent to S at p. The *rim* of S is the set of those points p for which $T_p(S)$ contains a line parallel to ξ. The *occluding contour* of S is the projection of the visible rim on the image plane (figure 2). The shape of the occluding contour depends on S and the viewing direction. Our goal is to use this contour information to recover a description for the parameterization x at points of S in the vicinity of the corresponding rim points.

Local surface shape (i.e., curvature) is completely expressed by the first and second fundamental forms of S with respect to x (doCarmo 1976; Koenderink 1990). Specifically, let $N(p) : S \to S^2$ be the Gauss map of S, assigning a unit normal vector $N(p)$ in the direction of the vector product $x_u \wedge x_v$ at every point $p \in S$. The normal section of S along a direction ξ in $T_p(S)$ is the plane curve produced by intersecting S with the plane of ξ and $N(p)$. The second fundamental form, $\mathrm{II}(p)$, gives an expression for the curvature of this curve at p. $\mathrm{II}(p)$ has a single maximum and minimum, k_{n_1} and k_{n_2}, along two orthogonal directions, e_1 and e_2, respectively. These directions are called the principal directions at p. We can use the two quantities k_{n_1}, k_{n_2}, called the principal curvatures of S at p, to compute the curvature of the normal sections along any other direction using Euler's formula:

$$k_n(\phi) = k_{n_1} \cos^2 \phi + k_{n_2} \sin^2 \phi \qquad (1)$$

where ϕ is the angle between the new direction and e_1. Hence, we can recover the local shape of S completely from the principal curvatures of S. A qualitative description of the local shape of S can be given by looking at the sign of their product $K = k_{n_1} k_{n_2}$, the Gaussian curvature of S at p (figure 3).

Our goal is to recover the principal directions and principal curvatures at selected points on S. We focus on the general case where p is not an umbilic point, that is, a point where any pair of orthogonal directions on $T_p(S)$ is a pair of principal directions; recovering the local shape of the surface at umbilic points is then straightforward. In the vicinity of nonumbilic points there exists a special parameterization $x(u, v)$ of S such that the tangents to the curves $x(u, v_0)$ and $x(u_0, v)$ (u_0, v_0 constant) are along the principal directions. These curves are called *lines of curvature* and their properties are intrinsically related to the underlying surface. Therefore they serve as a natural basis for describing a surface (Brady et al. 1985; Stevens 1981). In the rest of this article, x will refer to such a parameterization.

The geometry of a point on the occluding contour and the information we can derive from it depend on whether the point is a projection of an elliptic, hyperbolic, or parabolic point. This qualitative classification is therefore especially important in order to evaluate the results of our approach.

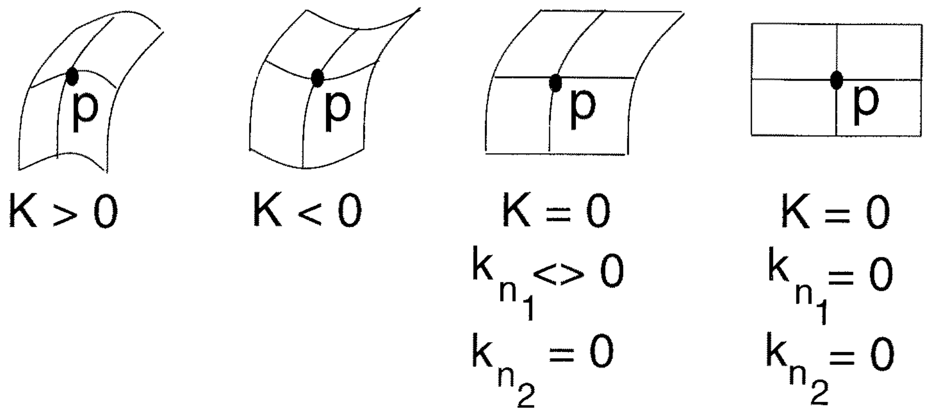


Fig. 3. Classification of the surface point p based on K: (a) p is elliptic, (b) p is hyperbolic, (c) p is parabolic, and (d) p is planar.

3 Local Surface Geometry from Occluding Contour

The problem of recovering surface geometry from the occluding contour has been mainly studied under the assumption that the viewing direction is *arbitrary*. This means that the viewing direction is not related in any way to the geometry of the rim or the occluding contour. For example, there are surfaces for which their rim is planar when viewed from a particular set of directions (Marr & Nishihara 1978). The assumption that the viewing direction is arbitrary immediately excludes such viewpoints from consideration since the rim is not always planar (Koenderink 1990). This is a reasonable assumption, however, only if the observer cannot control the direction of gaze. Unfortunately we can only derive a limited amount of information from the occluding contour when this assumption is in effect.

Let p be a point on the visible rim of S when viewed from direction ξ and let q be its projection on the image plane. There are three main results describing what can be recovered from the shape of the occluding contour under orthographic projection and from an arbitrary viewing direction:

1. We can recover the surface normal and the tangent plane at p from ξ and the tangent to the occluding contour at q. This is because, by definition, $T_p(S)$ contains both ξ and the tangent to the occluding contour (Barrow & Tenenbaum 1981).
2. Let k_o be the curvature of the occluding contour at q. Then k_o and the Gaussian curvature K of S at p have the same sign (Koenderink 1984; Brady et al. 1985).
3. If k_n is the normal curvature of S at p along ξ, then $K = k_n k_o$ (Giblin & Weiss 1987; Koenderink 1984; Brady et al. 1985).[1]

Similar results hold for perspective projection where the plane of projection is not positioned at infinity, and for the case where $k_o = 0$ (Brady et al. 1985). Because K is defined as the product of two curvatures on the surface (i.e., k_{n_1}, k_{n_2}), these results suggest that if we know k_o then we only need to measure one curvature on the surface instead of two. In fact k_n and k_o determine the second fundamental form at p. This was the main idea behind the surface reconstruction approach of Cipolla and Blake (1992).

The above results are important but they also imply that if we have no additional information about the shape of the viewed surface, the information provided by the occluding contour is primarily qualitative. However, when the observer can actively control the viewing direction, we can exploit the existence of directions that allow the derivation of complete information about the surface. We show this by presenting three simple corollaries to a result of Blaschke (Koenderink 1990). Blaschke's result is analogous to Euler's formula and relates the curvature of the occluding contour with the principal curvatures of S at the rim:

Theorem 1 (Blaschke): *Let ϕ be the angle between ξ and the principal direction e_1 at p. If $K \neq 0$,*

$$k_o^{-1}(\phi) = k_{n_1}^{-1} \sin^2 \phi + k_{n_2}^{-1} \cos^2 \phi \quad (2)$$

Corollary 1: *If ξ is along e_1, then $k_o = k_{n_2}$.*[2]

Corollary 2: *Let ξ, ξ' be two distinct viewing directions in $T_p(S)$ from which p is visible, and let k_o, k_o' be the curvatures of the occluding contour at the corresponding projections of p. If* (1) $K \neq 0$ *at p,* (2) $\xi = e_1$, *and* (3) *the angle between ξ and ξ' is known, then we can compute k_{n_1}, e_2, and K at p.*

Corollary 3: *Let p be a point on the visible rim of S with $K \neq 0$. Let $\phi \in [-\pi, \pi)$ be the angle between* $\xi \in T_p(S)$ and e_1. (1) *If p is elliptic and nonumbilic, the function $k_o(\xi)$ takes its minimum and maximum values only when ξ coincides with one of the principal directions.* (2) *If p is hyperbolic, $k_o(\xi)$ is well-defined only when $|\phi| < \arctan\sqrt{(k_{n1}/-k_{n2})}$ for $|\phi| < \pi/2$, or $\pi - |\phi| < \arctan\sqrt{(k_{n_1}/-k_{n_2})}$ for $|\phi| \geq \pi/2$. For these directions, $k_o(\xi)$ takes its maximum value when ξ coincides with e_1 and it has not minimum value.* (3) *If p is umbilic, $k_o(\xi)$ is constant.*

Proofs. If ξ is along e_1, $\phi = 0$ in equation (2). Corollary 1 immediately follows. For corollary 2, note that k_{n_1} is derived using corollary 1, and that ϕ is known. Since $K \neq 0$, equation (2) is well defined and we can use it with k_{n_1} as the unknown. The other principal direction is also computable since e_2 must lie on $T_p(S)$ and be perpendicular to ξ. In fact, ξ is perpendicular to the rim even though this does not hold in general (Koenderink 1990).

Finally, the derivative of $k_o(\phi)$ is

$$k_o'(\phi) = \left(\frac{1}{k_{n_2}} - \frac{1}{k_{n_1}}\right) \frac{\sin 2\phi}{(k_{n_1}^{-1}\sin^2\phi + k_{n_2}^{-1}\cos^2\phi)^2} \quad (3)$$

In the case of an elliptic, nonumbilic point, $k_{n_1} \neq k_{n_2}$ and therefore k_o' becomes 0 for $\phi = 0$ or $\phi = \pi/2$, that is, when ξ is along a principal direction. If p is hyperbolic, the expression in the denominator tends to 0 as ϕ approaches $\arctan\sqrt{(k_{n_1}/-k_{n_2})}$ which is the angle between e_1 and the asymptote of the surface at p. In the interval $[\arctan\sqrt{(k_{n_1}/-k_{n_2})}, -\arctan\sqrt{(k_{n_1}/-k_{n_2})} + \pi]$ p becomes occluded and therefore $k_o(\phi)$ is undefined. In the interval where $k_o(\phi)$ is defined, equation (3) shows that $k_o(\phi)$ has a maximum only for $\phi = 0$ or $\phi = \pi$. Finally, when p is umbilic, $k_{n_1} = k_{n_2}$ by definition, and $k_o'(\phi)$ is identically zero. □

Corollary 1 suggests that the principal directions at p form a special set of directions providing explicit information about surface geometry in the vicinity of p. Now assume that we are viewing a point p from a particular viewing direction and can measure the curvature of the occluding contour at p's projection. If somehow we can adjust our viewing direction to coincide with a principal direction at p and know what this adjustment is, corollary 2 shows that we can derive the second fundamental form of S at p. This solves the shape-recovery problem for p. The most important result is given by corollary 3. It shows that the problem of finding the principal directions at a point can be treated as a simple maximization (or minimization) problem. We describe the implications of this result in the next section and show how it can be used by an active observer to find the principal directions at p.

4 Recovering the Local Geometry of a Surface Point

The basic step of our surface reconstruction approach is to select a point on the occluding contour and recover the local surface geometry for its corresponding rim point. We do not address the point-selection problem directly. The reason for this is that we cannot decide a priori which point on the occluding contour will prove the most useful. This will depend on the context in which the approach is used. However, there are specific types of points for which our reconstruction method may not work. Therefore, our task will be to select a point on the rim for which we can ensure that our approach is effective. Below, we first outline the main ideas, and in section 6 we present more details.

4.1 The Active Reconstruction Approach

Suppose we have selected a point p on the rim of surface S. For simplicity we will assume that p is at the origin. We first consider the case where p is nonumbilic. Corollary 3 says that if p is a hyperbolic point or a nonumbilic elliptic point, there are only two viewing directions in $T_p(S)$ for which k_o obtains a local maximum value and two directions for which k_o obtains a local minimum. Our goal is to find one of these directions since they correspond to e_1 and e_2. We discuss the problem of finding e_2; e_1 is treated similarly.

Viewing directions in the plane $T_p(S)$ can be thought of as points on a unit circle, C, defined by the intersection of the sphere S^2 centered at p with $T_p(S)$ (see figure 4). As the observer changes viewing direc-

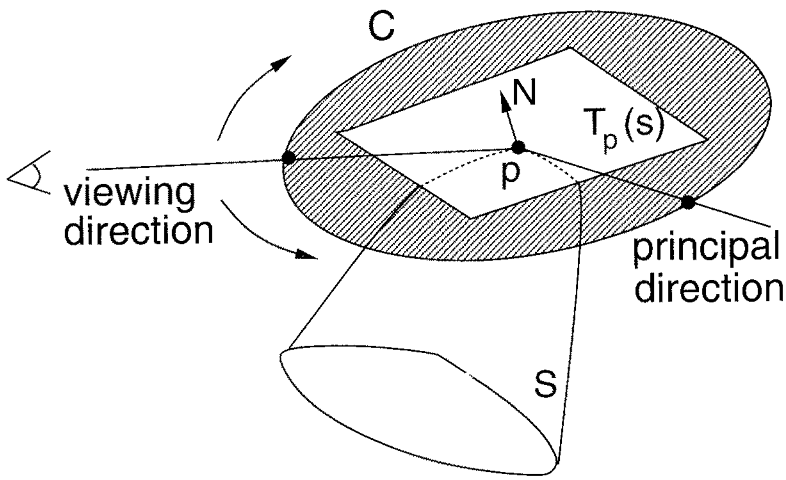


Fig. 4. Viewing directions on $T_p(S)$. Viewing directions correspond to points on the unit circle C lying on $T_p(S)$ and centered at p.

tion on $T_p(S)$, the corresponding point moves on C. Our goal is to smoothly move the point on C until the viewpoint with maximum k_o is found. To do this we must answer two questions: (1) Which direction should the observer move on the unit circle, and (2) how can the observer detect when the viewing direction is equal to e_2?

We have only two possibilities for moving on the unit circle, either clockwise or counter-clockwise. Obviously, we prefer the minimal-motion solution in which the desired extremum is attained with the smallest possible change in viewing direction. In particular, if we move in the direction of increasing k_o, the first extremum we reach is a maximum. It easily follows from the local geometry of elliptic and hyperbolic points that this strategy will in fact produce the smallest viewing direction change (figure 5). On the other hand, parabolic points do not have this property.

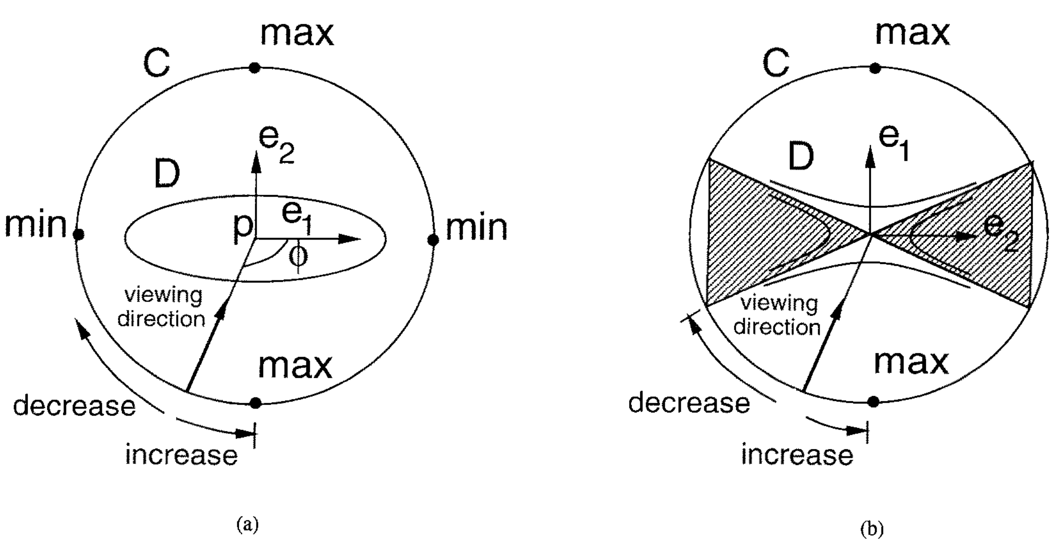


Fig. 5. Finding the principal directions. Top views of the tangent plane are shown. The axes represent the principal directions, and the origin corresponds to the point of contact, p, with the surface. The viewing direction makes an angle ϕ with the first principal direction. C represents the set of viewing directions on the tangent plane, and D is Dupin's indicatrix for p. (a) p is an elliptic point. Clockwise change in viewing direction decreases k_o. The viewing direction can change by at most $\pi/2$ before a local minimum or a local maximum is reached. (b) p is hyperbolic. The only achievable extremum is a local maximum, obtained in this case by a counter-clockwise rotation. Shaded areas, delimited by the asymptotes of the point, represent the directions where p is occluded. The maximum viewing direction change before an extremum is found, in this case, decreases to the angle between e_1 and the asymptotic directions. (Note that the axis labels have been reversed.)

The second question, detecting when the viewing direction is equal to e_2, is partly answered by corollary 3. It says that we can detect this event by detecting a local maximum of k_o. However, in order to detect this local maximum, p must be visible; k_o cannot be measured otherwise. The visibility of p is affected by the local surface geometry at p as well as by the global geometry of S. Ignoring for a moment the case where p is occluded by some distant point on S, we arrive at the following two conclusions: (1) If p is elliptic, we can align the viewing direction with either e_1 or e_2. Furthermore, the maximum possible direction change before the alignment takes place is $\pi/2$ (figure 5a). (2) If p is hyperbolic, we can align the viewing direction only with e_1. The maximum possible direction change in this case is determined by the point's asymptotes (figure 5b).

The problem of recovering the local surface geometry at an umbilic point on the visible rim is even simpler. Corollary 1 suggests that if we know p is umbilic, we can recover the local surface shape at p by simply measuring the curvature of the occluding contour at p's projection. It therefore suffices to find a way of detecting that p is umbilic. Corollary 3 shows that this can be done by determining whether k_o remains constant as the viewing direction moves on the circle C.

These results suggest a simple algorithm to align the observer's viewing direction with e_2 and recover the local surface shape at p:

Step 1. Perform a small change of viewing direction on $T_p(S)$ and measure the difference between the previous and current value of k_o. If it increases, continue to change the viewing direction in the same way so that e_2 will be reached first. If it decreases, move the viewing direction in the opposite way. If it remains constant, stop moving; k_{n_1} and k_{n_2} are both equal to k_o (i.e., p is umbilic).

Step 2. Continue moving in the same direction until k_o reaches a maximum. This viewpoint corresponds to e_2 and therefore the observer can stop moving and use the current value of k_o for k_{n_1}.

Step 3. Measure the total change of viewing direction between the initial and final directions. Corollary 2 says that this angle along with k_{n_1} and the initial value of k_o can be used to determine k_{n_2}.

The above algorithm assumes that the observer can measure relative changes in viewing direction. For elliptic points this requirement can be relaxed at the expense of additional motion: The observer can recover the principal curvatures at the selected point by moving to the viewpoints corresponding to the maximum *and* the minimum value of k_o. Hence, in this case shape recovery can be achieved without relying on any quantitative measurements involving the observer's motion; the observer must simply be able to control viewpoint around the selected point by moving clockwise and counterclockwise on its tangent plane.

4.2 Selecting Surface Points for Reconstruction

Any observer motion minimizing or maximizing k_o must take into account the effects of global surface geometry: Irrespective of its local structure, p may become occluded by distant points on S. The following proposition shows that (a) there are at least some points on the visible rim of S that cannot be occluded by S if the observer changes direction as described above, and (b) these points are easily detected on the occluding contour (figure 6).

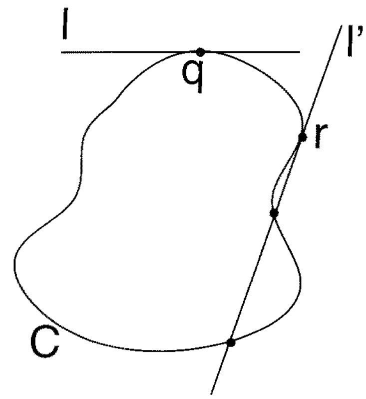


Fig. 6. Determining the complete visibility of rim points. Since the tangent at q does not intersect the occluding contour elsewhere, q corresponds to a rim point visible from every direction in its tangent plane.

Proposition 1: (1) Let p be a visible, elliptic point on the rim of a smooth surface S when viewed from direction ξ under orthographic projection. Let q be the projection of p in the image plane and let l be the tangent to the occluding contour at q. Then, p is visible from every direction on $T_p(S)$ iff l does not intersect the occluding contour and is not tangent to it at any point other than q.

(2) Let C be the occluding contour of S when viewed from direction ξ. Then there is at least one point on S projected in C that is visible from every direction in $T_p(S)$.

Proof: (1) (*Only If*) Consider the intersection of S with $T_p(S)$. If the intersection contains only the point p, then the intersection of any line $m \in T_p(S)$ with S will either be empty or equal to p. Recall that while changing viewing direction, $T_p(S)$ is viewed edge-on and its projection is the line l. Since l is the projection of all lines in $T_p(S)$ (except those lines parallel to ξ), it follows that l will only intersect the occluding contour at q.

(*If*) Assume there is a viewing direction in $T_p(S)$ from which p is not visible. Let ξ' be the first such direction while moving clockwise (figure 7). The viewing direction ξ' must contact S at p and at at least one more point, say s. Now consider the intersection of S with $T_p(S)$. The intersection will consist of a set of closed curves and isolated points. Since p is elliptic there must exist a small disk in $T_p(S)$ centered at p that does not contain any other points of S. Therefore p and s must be in different components. We distinguish

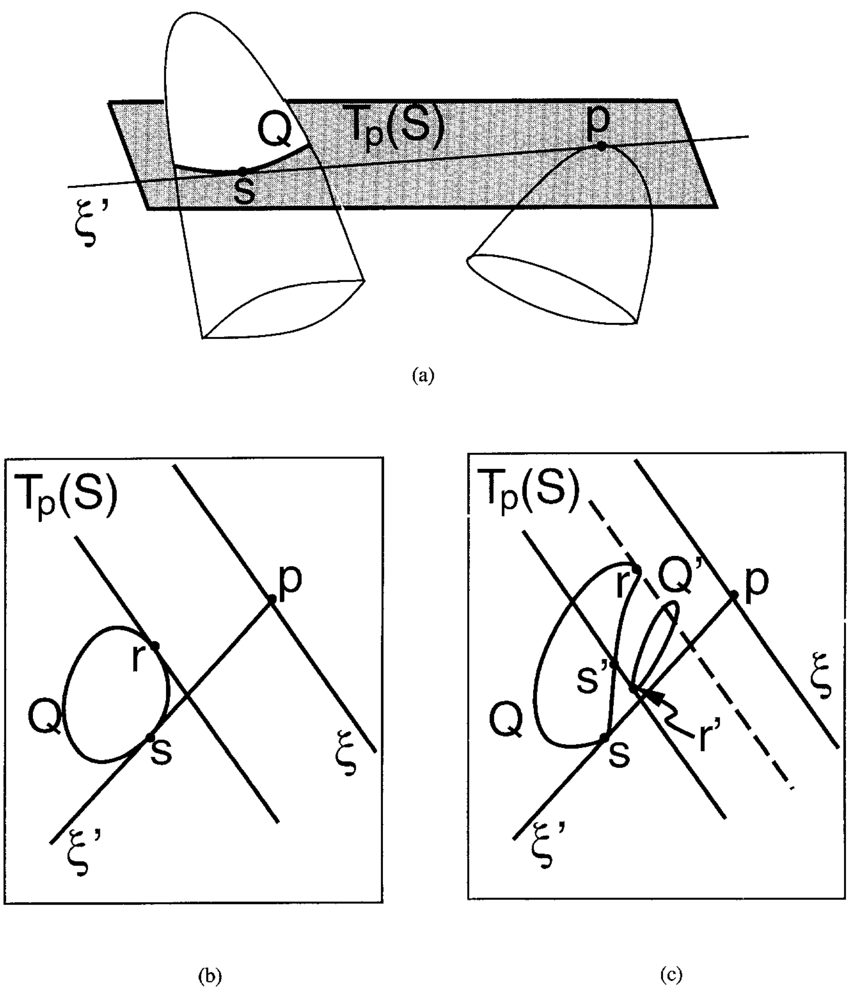


Fig. 7. The effects of global occlusion. (a) A "side" view of $T_p(S)$. Viewing direction ξ' is the first direction in which p becomes occluded. s is the point occluding p from that direction. Q is the component of $S \cap T_p(S)$ containing s. (b) "Top" view of the tangent plane of S at p, shown in (a). ξ is the original viewing direction. (c) "Top" view of the tangent plane $T_p(S)$ for a surface S in which $S \cap T_p(S)$ contains two components, Q and Q'.

two cases, namely whether s is an isolated point or a point on a curve. If s is an isolated point then $T_p(S)$ must be tangent to S at s. But then s is also part of the rim when viewed along the original direction ξ. In addition, s must be on the visible rim because it is the first point that occludes p when changing viewing direction from ξ to ξ'. This implies that the projection of an imaginary line joining s and p will contact the occluding contour at two points, the projections of s and p.

If s is not an isolated point, let $Q \subseteq S \cap T_p(S)$ be the closed curve containing s. Now consider the family of lines parallel to ξ. A line of the family will contact Q, say at point r. Without loss of generality assume that r is the first such contact point when Q is traced in a counter-clockwise fashion starting from s. This point, by definition, must be on the rim of S when the viewing direction is ξ. If it is also on the visible rim (figure 7b), the projection of the imaginary line joining r and p must intersect the occluding contour at at least two points (i.e., at the projections of r and p).

Now suppose r is not on the visible rim. To treat this case, note that by definition, s must be visible when the viewing direction is ξ. Let s' be the first occluded point on Q when Q is traced in a counter-clockwise fashion starting from s (figure 7c). The point r' occluding s' must necessarily belong to the visible rim. Therefore, the projection of the imaginary line connecting r' and p intersects the occluding contour at at least two points.

(2) Consider any point q on C that is also contained on the convex hull of C. Since C cannot be a straight line, q is, by definition, the only point in common between C and the tangent at q.

But then q also satisfies the conditions of (1) above. □

Figure 8 shows the results of applying proposition 1 to the occluding contour of a candlestick. The proposition implies that the only points ensuring the correctness of the algorithm are elliptic. However, this is a *necessary* requirement for the absence of occlusion but not a *sufficient* one. This means that there are cases where the geometry of hyperbolic points can be recovered with our approach. In fact, shape recovery for hyperbolic points requires less observer motion on average since the extent of the visibility of these points is limited by their asymptotic directions.

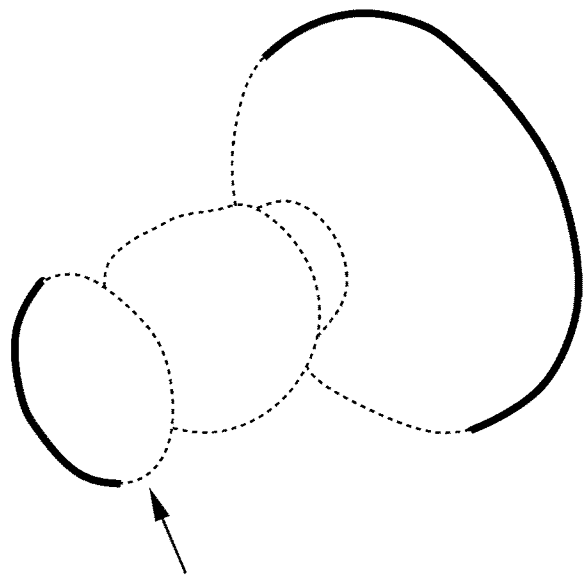


Fig. 8. Selecting points for surface recovery. Solid lines on the occluding contour of a candlestick show the points that cannot become occluded while changing viewing direction in their tangent planes. The arrow indicates the point having the greatest (absolute) curvature of all acceptable points.

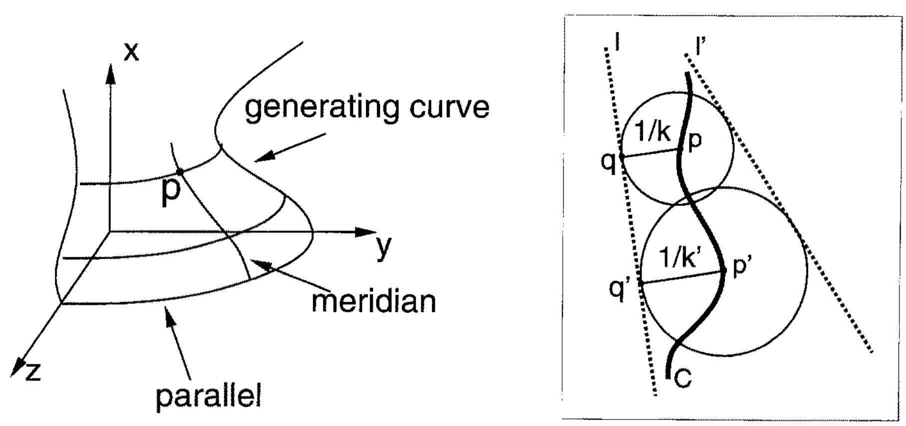
5 Surfaces of Revolution

In the last section we presented an algorithm for recovering the shape of a single point on the surface rim. However, there are surfaces for which the local shape of a single rim point reveals global properties of the surface. Surfaces of revolution present an ideal example of such surfaces. The properties and appearance of surfaces of revolution and their generalizations have been studied extensively (Marr & Nishihara 1978; Ulupinar & Nevatia 1988; doCarmo 1976; Brady et al. 1985; Stevens 1981; Horaud & Brady 1987; Ponce & Chelberg 1987; Richetin et al. 1991). Here, we focus on the specific relation between their global structure and the local shape of points on the rim.

Surfaces of revolution are formed by rotating a planar curve around a straight axis that does not meet the curve. Therefore, we can completely describe a surface of revolution by the axis and the generating curve. Approaches for recovering the axis of a surface of revolution have been mainly geared toward detecting symmetries in their occluding contour or outline (Marr & Nishihara 1978), or utilizing their viewpoint-invariant properties (Horaud & Brady 1987; Ponce & Chelberg 1987; Ponce et al. 1989). The problem with detecting symmetries in the occluding contour is that the existence of such symmetries depends on viewpoint. On the other hand, the identification, detection, and utilization of viewpoint-invariant properties is a nontrivial task. For example, in (Ponce et al. 1989) the axis was recovered using a Hough transform-based technique. However, such a technique largely depends on the number of rim points actually detected. In addition, the axis is severely foreshortened for near-top views of a surface of revolution (i.e., when the viewing direction is almost parallel to the axis of rotation), limiting the applicability of methods relying on a single, arbitrary view to precisely recover the axis. Our active approach neatly lends itself to these problems in order to make them easier to handle. The idea is that if the viewer can align the viewing direction with a principal direction of a rim point, then shape and symmetry analysis of the occluding contour becomes especially simple. This is because one of the principal directions corresponds to a "side" view of the surface (i.e., a view for which the viewing direction is perpendicular to the axis of rotation). If the generating curve of the surface can be written in the form $y = f(x)$, then the recovery of the axis of rotation allows us to recover the generating curve directly from a side view. Even further, the surface rim from such a view is guaranteed to be completely visible. In the rest of this section we focus on surfaces of revolution whose generating curve has this property.

Consider a point p on the rim of a surface of revolution when viewed from an arbitrary direction (figure 9a). The two principal directions at p correspond to the tangents to the parallel and the meridian passing through p. Since the parallel is a planar curve, if the visual ray is tangent to the parallel at p it is contained in the plane of the parallel. Hence, it is perpendicular to the axis of rotation and the view corresponds to a side view of the surface. The occluding contour from such a side view is symmetric. Therefore, the axis of rotation (as well as the generating curve) can be recovered by simply using existing symmetry-seeking approaches (e.g., Horaud and Brady (1987)) which are well defined for such a viewpoint. However, the direction and position of the axis can also be constrained by recovering the principal curvatures corresponding to the parallels for two points on the rim of a side view (figure 9b). This approach is similar to the one used by Richetin et al. (1991) where the geometry of the occluding contour at two parabolic points was used to hypothesize the pose for surfaces that are straight homogeneous generalized cylinders.

The observer must choose between moving toward the principal direction of minimum curvature or moving



(a) (b)

Fig. 9. (a) A surface of revolution. The x-axis is the axis of rotation. (b) Constraining the axis of rotation from the principal curvatures of two rim points from a side view. Curve C is a segment of the occluding contour corresponding to a side view. Since the view is a side view, C belongs to a generating curve at the surface. Let the principal curvature corresponding to the parallels at p and p' be k and k', respectively. The distance of p and p' from the projection of the axis of revolution, l, is $1/k$ and $1/k'$, respectively. Let q and q' be the projections of the points of intersection of the axis with the planes of the parallels at p and p', respectively. The axis must be perpendicular to the lines along pq and $p'q'$. Therefore the axis must be tangent to the circles of radius $1/k$ and $1/k'$, centered at p and p' respectively. There are at most two such lines that do not intersect the generating curve C, namely l and l'. The direction of the axis is the normal of the plane defined by the viewing direction and the line through pq.

toward the one of maximum curvature. Although the curvature extremum corresponding to a side view for a selected point is not known a priori, this choice is easy if the visible rim contains hyperbolic points. Recall that the only principal direction from which these points are visible is the direction of maximum curvature and that if the generating curve of the surface can be written in the form $y = f(x)$, then these points must be visible from a side view of the surface. Hence, the observer can select one hyperbolic point and align the viewing direction with the principal direction of maximum curvature (corollary 3).

It is also easy to show a more general property of surfaces of revolution with this type of generating curve: If the viewing direction smoothly changes on the tangent plane of a selected rim point, this point will not become occluded if the viewing direction is approaching the direction of a side view. This fact can be used to decide how to change viewing directions on the tangent plane in order to approach a side view of the surface when no hyperbolic surface points are visible.

Our discussion above deals with a specific type of surface of revolution. However, it can be generalized to an arbitrary surface of revolution and to the case of straight homogeneous generalized cylinders where the axis is perpendicular to the cross-section. Consider the case where the observer selects a rim point that belongs to a parallel that is also a geodesic. If such parallels exist on the surface, our approach can be used to obtain both the top and the side views of the surface as well as its axis. Consider the case where the generating curve of the surface of revolution cannot be written in the form $y = f(x)$ (e.g., a torus). In this case, we can still recover the axis of rotation from the side view using occluding contour symmetries, and find points on the generating curve for which the tangent to the generating curve is parallel to the axis. These points belong to parallels that are geodesics and their principal directions correspond to the side and top views of the surface. Therefore, the observer can align the viewing direction with the top view for any type of surface of revolution.

Points on the rim that belong to geodesic parallels are also important because they can be used to recover the axis of surfaces of revolution and straight homogeneous generalized cylinders. The surface normals at these points lies on the plane of the parallels. The viewing direction corresponding to a side view also belongs to this plane. Therefore, we can recover the plane of the parallels. Since the axis of rotation is normal to this plane we can also recover the direction of the axis of the surface. It is in fact possible to detect such a point on the rim if it exists (without having already determined the axis).

The next section extends our basic shape-recovery step by (1) selecting a new point on the surface in the vicinity of the previously selected point, and (2) applying the shape-recovery step presented in section 4 to the new point. We also briefly discuss how this two-step approach can be used to select rim points that belong to geodesic parallels.

6 Extending Surface Recovery to Neighboring Points

Our main objective is to recover the complete shape description for a single rim point. In this section we consider an extension to this approach—selecting a new point and applying the shape-recovery process to that point. We must consider two important issues in order to demonstrate the effectiveness of such an extension:

1. The extent of the viewing direction adjustments needed to align the viewing direction with one of the principal directions at the newly selected point.
2. The extent of the viewing direction adjustments *required* by our basic shape-recovery algorithm in order to produce reliable shape information for the newly selected point. This is because if the viewing direction adjustment is close to zero, then numerical problems are introduced in the calculations of the principal curvatures from corollary 2.

We will discuss the issue of selecting new points for shape recovery based on these two issues. The process has as a primary goal the removal of the first point from the rim and its replacement by a new point at which the first step will again be applied.

Let p be the previously selected point. After applying the shape-recovery step, the viewing direction ξ of the observer is aligned with one of the principal directions at p, say e_2. We have seen that if we change directions in $T_p(S)$, p will not leave the rim. Therefore, we must change viewing directions in some other plane containing e_2. The important issues here are (a) which plane should be selected for changing the viewing direction, and (b) how much should the viewing direction change in that plane? The motivation for our approach is to ensure that the shape-recovery step for the new point will need only small viewing direction adjustments. In other words, we require that the new viewing direction does not form a large angle with one of the principal directions at the new point.

Suppose we have selected a particular plane P passing through p and containing e_2, and that we continu-

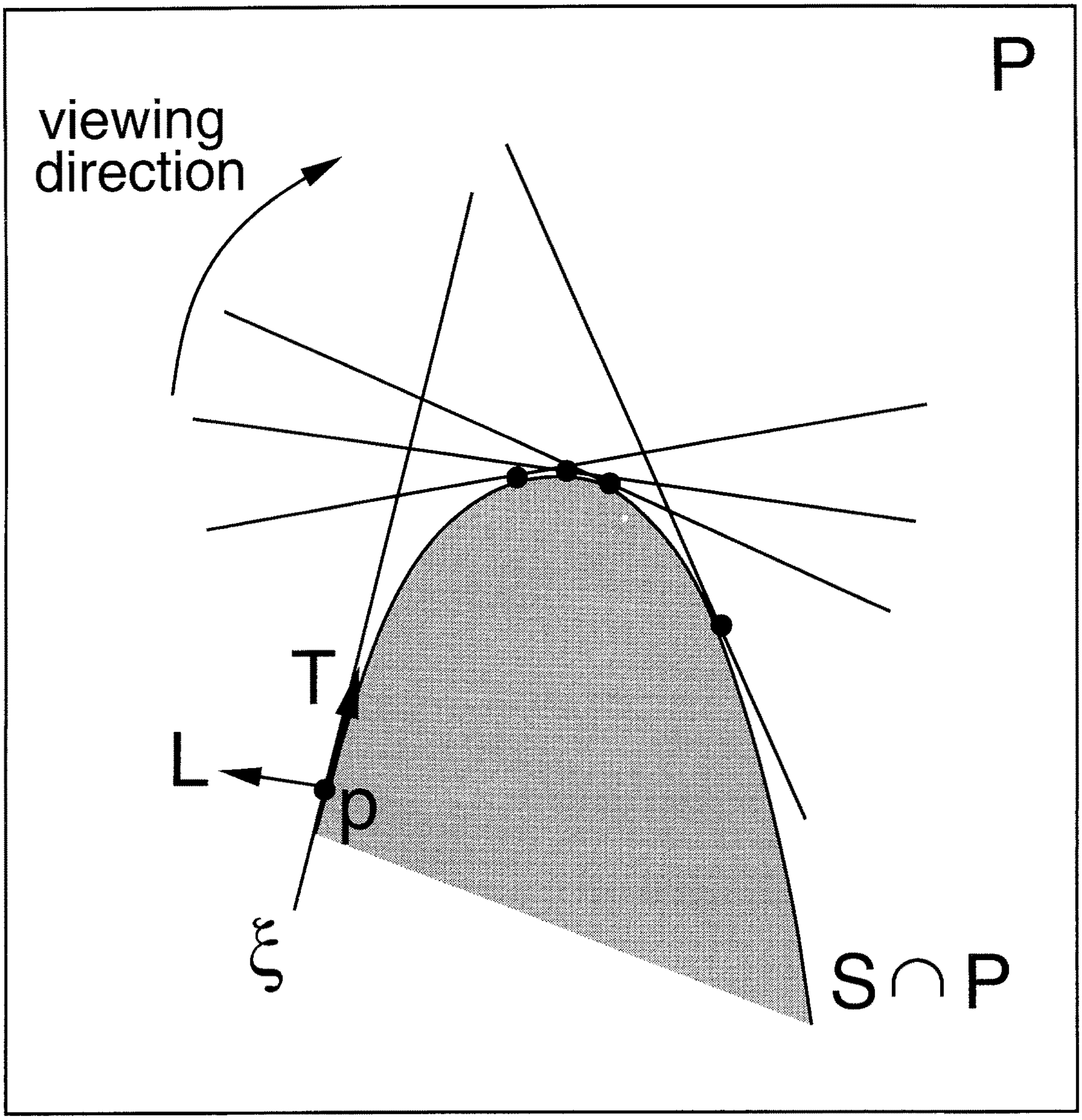


Fig. 10. Removing p from the rim. The figure shows the intersection of a selected plane P with the surface. The viewing direction ξ changes in P and the visual rays graze the surface along the curve $\beta(s) = S \cap P$.

ously change viewing direction in that plane. As the viewing direction changes, the visual ray contained in P will graze the surface along a curve $\beta(s)$ also contained in that plane (figure 10). Now suppose that we stop at a new viewing direction ξ'. The visual ray will now be tangent to $\beta(s)$ at some new surface point. The shape-recovery step will now be applied to this point, attempting to align the viewing direction with e_2 at the point. We must therefore examine how the angle of $\beta'(s)$ with e_2 varies with s. The basic idea is to examine the properties of $\beta(s)$ in light of the following efficiency and reliability requirements:

- The efficiency requirement is that $\beta(s)$ should always form an angle with e_2 that is as close to 0 as possible. This means that we require $\beta(s)$ to approximate a line of curvature.
- The reliability requirement is that $\beta(s)$ should form an angle of at least ϕ^*, for some predetermined constant ϕ^* that depends on the reliability of the

shape-recovery step. This means that we require $\beta(s)$ to form an angle of at least ϕ^* with the lines of curvature corresponding to e_2.

The compromise between these two requirements is to require $\beta(s)$ to form an angle of exactly ϕ^* with the corresponding lines of curvature. This means that $\beta(s)$ is a *loxodrome* for the surface, that is, a line on S that forms a constant angle with the lines of curvature. Therefore we should trace S along such a curve while changing viewing directions.

We show in the appendix that if the selected plane P is the normal plane (i.e., the plane defined by the viewing direction and the surface normal at p) and if the change of viewing direction on this plane is small, then the viewing direction adjustments during the shape-recovery step will in fact be smooth and depend entirely on intrinsic properties of the surface. Specifically, we show that these adjustments are (to a first approximation) proportional to the geodesic curvature of the lines of curvature at p and inversely proportional to the normal curvature of the lines of curvature at p. This is an important result because it allows us to predict the performance of our active viewing strategy based on knowledge of the intrinsic properties of the surface.

As an example, consider the case of surfaces of revolution. Suppose that the viewing direction of the observer is aligned with the principal direction corresponding to the parallels. Now suppose that the observer changes viewing direction on the normal plane at p and eventually selects a new point p' for shape recovery, as outlined above. If p belongs to a geodesic parallel, no viewing direction adjustments will be necessary during the shape-recovery step at p', that is, the viewing direction is also tangent to the parallel through p'. On the other hand, if the geodesic curvature of the parallel through p is nonzero, some viewing direction adjustments will be necessary. In fact it can be shown that if the observer repeats this process and selects points $p', p'', p''', \ldots$, these points will asymptotically approach a geodesic parallel if such a parallel exists. To illustrate this, let us assume for simplicity that the axis of the surface of revolution is vertical and the point initially selected is p. Then, the new point selected will be on a parallel below the parallel through p if the surface normal is pointing upward. Therefore if there is a geodesic parallel below the parallel through p, the points selected will approach that geodesic parallel.

7 Experimental Results

In this section we demonstrate the applicability of our active shape-recovery approach. We have implemented a prototype system that (1) automatically selects points on the rim of an object, (2) tracks these points while changing viewing direction on their tangent plane, and (3) computes the curvature of the occluding contour at the selected points in order to detect the viewpoints where it obtains an extremum value. We have applied our algorithms to simulated scenes and have also performed some preliminary experiments with a real scene. Figures 11 and 12 show the objects used in our experiments.

7.1 Simulated Scenes

The simulations were performed by synthetically generating images given a polyhedral object model and an observer constrained to move only in a circle of viewing directions around the object.[3] The occluding

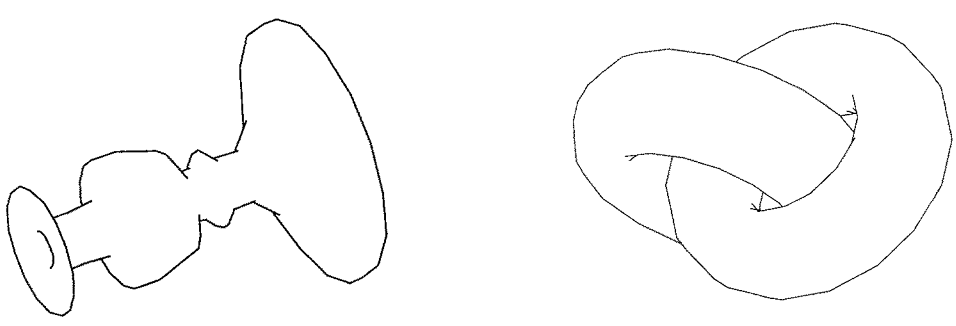


Fig. 11. Models of a candlestick and two tori used for the simulations.

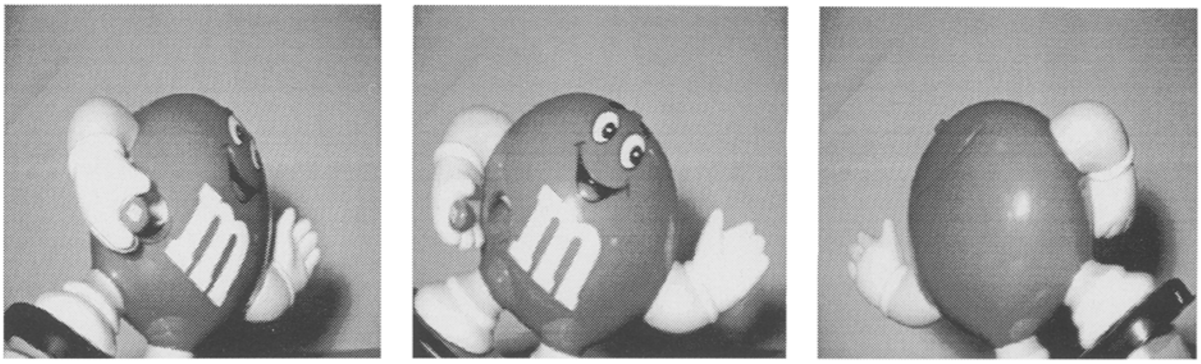


Fig. 12. A sequence of 120 frames used in our experiments. Frames 1, 30, and 120 are shown.

contour of the model for the current viewing direction was displayed and updated as the viewing direction changed smoothly. In our examples, the viewing direction was changed on a plane Π defined by a horizontal line in the image and the viewing direction (i.e., the projection of this plane to the image is a horizontal line).

Recall that the process of aligning the viewing direction with a principal direction at a point requires that the viewing direction changes on the point's tangent plane (figure 4). Hence, the one degree of freedom in rotation allowed us to detect the principal directions only for points tangent to Π. Our system automatically identified these points by finding the points where the occluding contour was tangent to a horizontal line. These points were automatically detected, labeled, and subsequently tracked while the viewing direction changed smoothly (figure 13).

Point tracking was performed using a simple algorithm. Note that since the viewing direction changes on the plane tangent to the selected points, they will always remain on the rim and the occluding contour will always be tangent to a horizontal line at these points. We use this observation to track points in subsequent frames by searching for points on the occluding contour that have horizontal tangents in the vicinity of the previously selected point. Occlusion is detected when this simple tracking step fails.[4] Figure 13 shows some of the tracked points for the two models. The points were initially selected and labeled for the viewing direction $\xi = 0$. Note that after a rotation of 3.93 radians the only unoccluded points are the points 0 and 6, exactly as predicted by proposition 1 (i.e., the tangents to the occluding contour at these points do not intersect the contour).

Curvature computations were performed by first approximating the occluding contour in the neighborhood of the selected points using cubic B-splines (Conte & de Boor 1972). The curvature was measured at the points where the tangent to the splines was horizontal. Even though splines have the effect of smoothing high curvature parts of a curve, we found that even with the actual rim curvatures being underestimated the curvature maxima were very distinct. In the case of polyhedral models, smooth viewpoint changes can result in an arbitrary number of model vertexes entering and exiting the polyhedral rim. Hence, the shape of the rim changes in a very discontinuous fashion, a problem not encountered with smooth surfaces where topological changes of the rim are not as frequent. This fact resulted in discontinuities in the curvature estimates, which ideally should vary continuously with viewpoint. However, figure 14 shows that the major peaks and valleys of the curvature estimates are clearly visible even in the presence of the discontinuities caused by the polyhedral approximation.

Figure 14 also shows how the absolute value of the curvature of the occluding contour at the selected points varies with viewpoint. Note that the candlestick and the torus are surfaces of revolution. Therefore, a "side" view corresponds to the viewing direction that is a principal direction for all points on their rim (i.e., the direction is tangent to the surface parallels). This is illustrated by the fact that the curvature maxima and minima occur at approximately the same viewing directions for the selected points. View 2 of the candlestick and the tori shows the occluding contour from the viewpoint of maximum curvature for point 0. The views in fact correspond to side views of the surfaces as expected. Also, note that in the case of the candlestick,

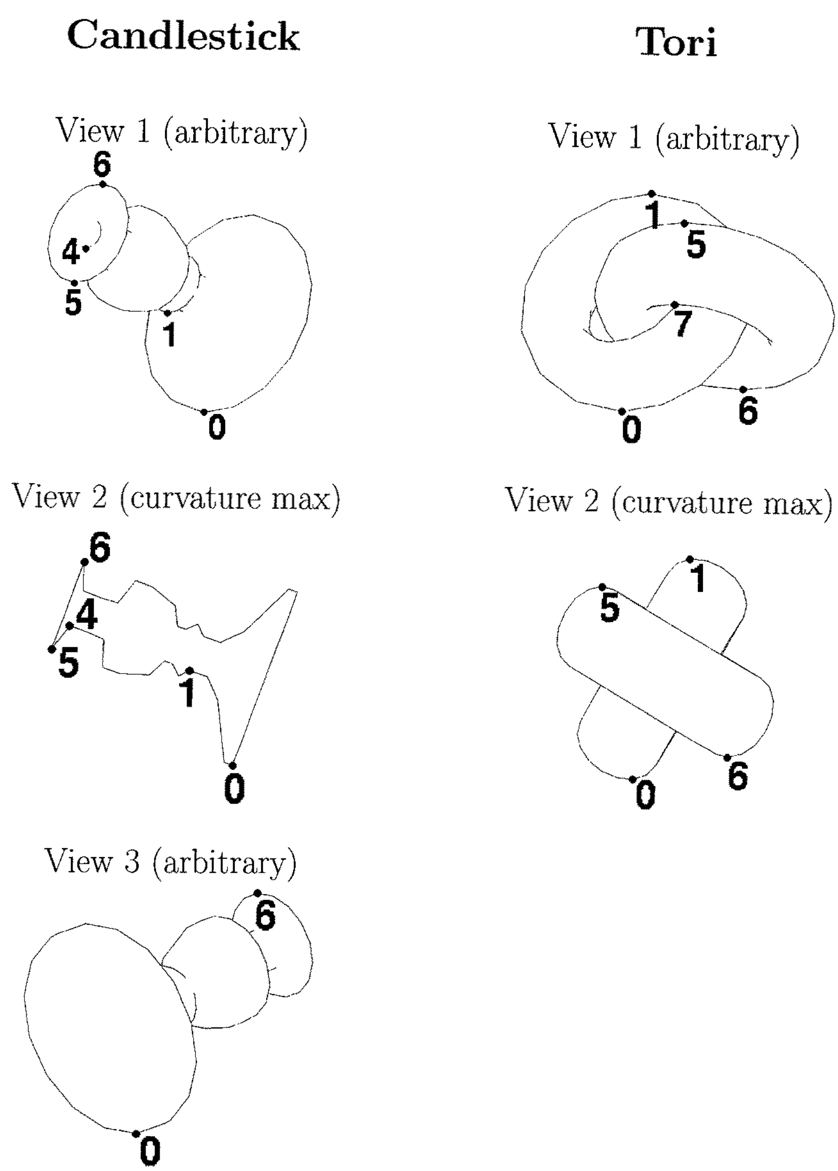


Fig. 13. Snapshots of the occluding contours of the two models as the viewing direction changes. The numbered points are the points being tracked.

the curvature maxima are much larger than the minima (over an order of magnitude), whereas in the case of the tori the extremal values are not very different. This is because the difference in the values of the principal curvatures at the selected points on the candlestick is much larger than for the two tori.

7.2 A Real Scene

In order to perform preliminary experiments with a real scene we extended the simple tracking and curvature estimation algorithms used in our simulations, and applied them to the sequence shown in figure 12. The sequence was produced by manually rotating an object after placing it on a horizontal turntable. The amount of the object's rotation between frames was assumed unknown. As in our simulations, this object motion allows the observer's viewing direction to be aligned only with the principal directions of points on the object whose tangent planes are horizontal.

The occluding contour of the object was tracked across frames using a simplified implementation of a B-spline snake (Cipolla & Blake 1992). The snake

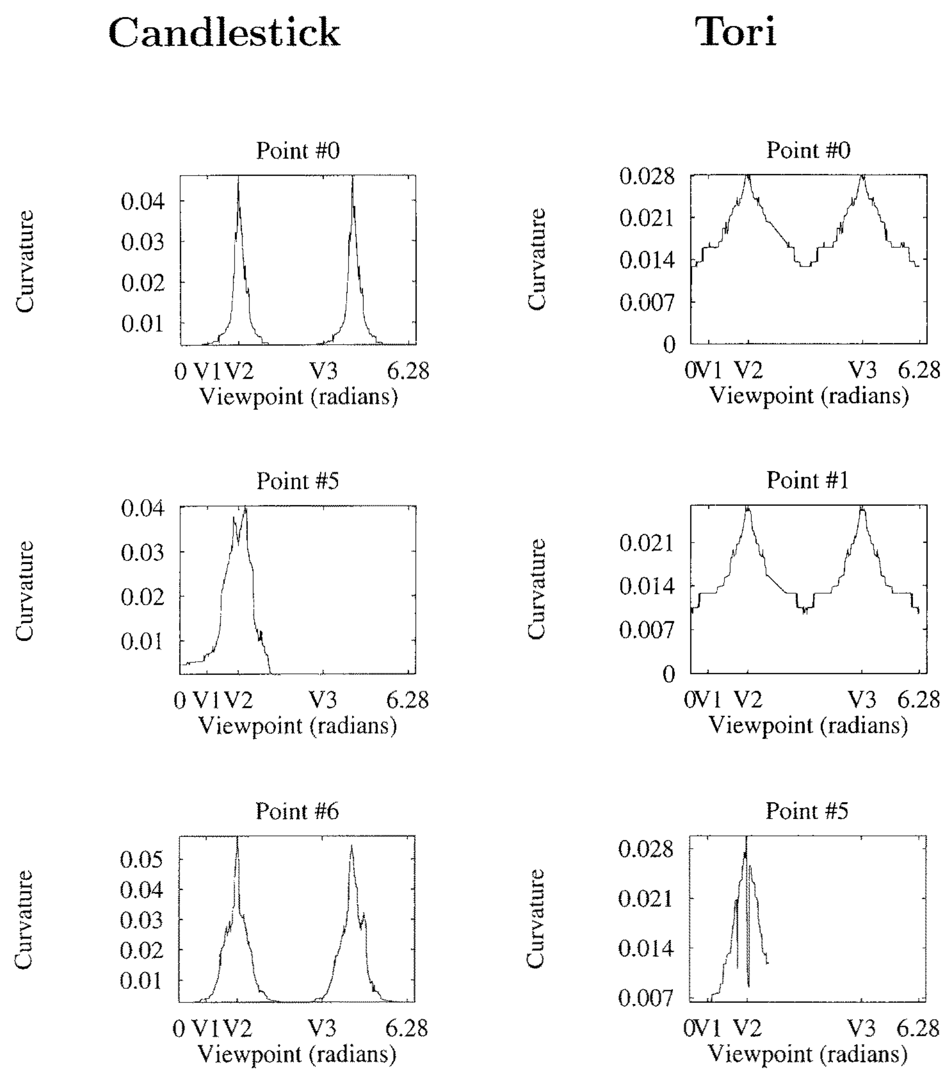


Fig. 14. Variation of the absolute curvature with respect to viewpoint at the selected points on the occluding contour. The models were rotated a total of 2π radians. The curves for point 5 on the candlestick and on the tori end at the viewpoint where their occlusion is detected. *Left column*: Viewpoints $V1$, $V2$, $V3$ correspond to views 1, 2, and 3, respectively of the candlestick in figure 13. *Right column*: Viewpoints $V1$ and $V2$ correspond to views 1 and 2, respectively of the tori in figure 13.

was interactively initialized near the object's contour. Point tracking was again performed by tracking the point on the snake whose tangent is horizontal (figure 15). Figure 16a shows the variation of the curvature of the snake at the tracked point for one run of the tracking process. Figure 17 shows the views of the object corresponding to the minimum and maximum measured curvature.

Curvature measurements were noisy mainly because of the snake's tracking behavior, which depended on the initial positioning of the snake and did not always lead to accurate approximations of the object's contour. Clearly, the curvature estimation process can be improved (especially near viewpoints corresponding to curvature maxima) by paying closer attention to the snake's tracking behavior. Our purpose here is simply to illustrate that the theoretically predicted curvature variation at the tracked point can be observed in practice. Since the only computations apart from snake tracking involve measuring the curvature of the snake at a single point, our active reconstruction approach is amenable to a real-time implementation; snake trackers operating at video rates are already becoming available (Blake et al. 1993).

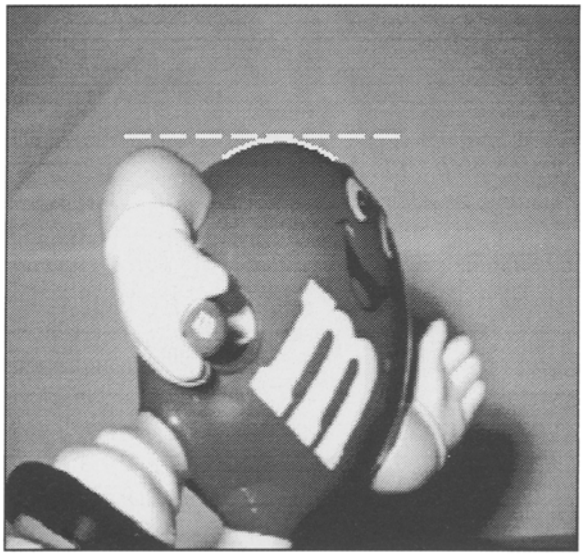


Fig. 15. The point being tracked is the snake point whose tangent is horizontal.

Figure 16a shows that it is necessary to incorporate measurements from multiple adjacent views in order to detect the principal directions of the tracked point. This implies that the observer must move past the viewpoint corresponding to the point's principal direction before the curvature extremum can be realiably detected. Figures 16b and 18 give the results from another run of the tracking and curvature measurement process. In this run, the contour's curvature was averaged over nine frames. Although the snake approximated the object's contour in a different manner, the principal directions of the tracked point are again easily distinguishable.

We are currently investigating an extension to our approach for improving the principal curvature measurements at the selected point by incorporating information from multiple frames during the observer's motion, and for more accurately localizing the point's principal directions. More specifically, when the relative changes in viewing direction between frames can be accurately measured, the principal curvatures and principal directions at the selected point can be predicted using three viewing directions on the tangent plane of the tracked point that satisfy the reliability requirement of section 6 (i.e., they are not too close to each other) (Kriegman 1993). This observation leads to a prediction-verification scheme for improving the accuracy of our active approach, whereby predictions during the process of aligning with a principal direction are evaluated against the outputs of the contour curvature estimator and the extremum detector. We expect this process to be useful primarily when the observer's viewing direction is close to the principal direction of maximum curvature where contour curvature measurements tend to be more reliable.

8 Concluding Remarks

We have demonstrated that an active observer can follow a very simple viewing strategy to recover exact shape information at selected rim points. Furthermore, this strategy is based purely on the computation of a simple property of the occluding contour (curvature at a point). Our experimental results show that this strategy is readily implementable and because of its simplicity and its low computational requirements, is very suitable for real-time implementation.

The use of an active observer is the most crucial aspect of our approach. The observer's ability to *purposefully* change viewpoint makes it possible to reach the special viewpoint where the shape of the occluding contour provides complete and exact surface-shape information. Moreover, our approach demonstrates that

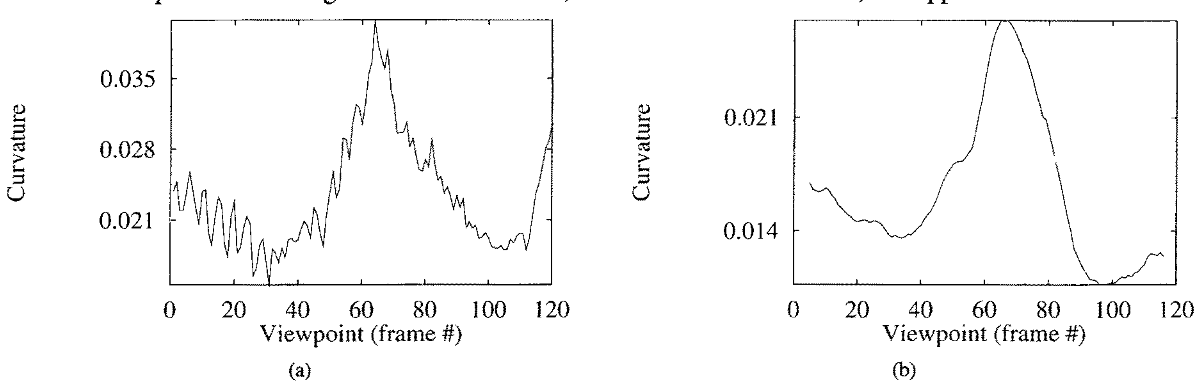


Fig. 16. (a) Curvature variation with viewpoint. (b) Curvature estimates averaged over nine frames.

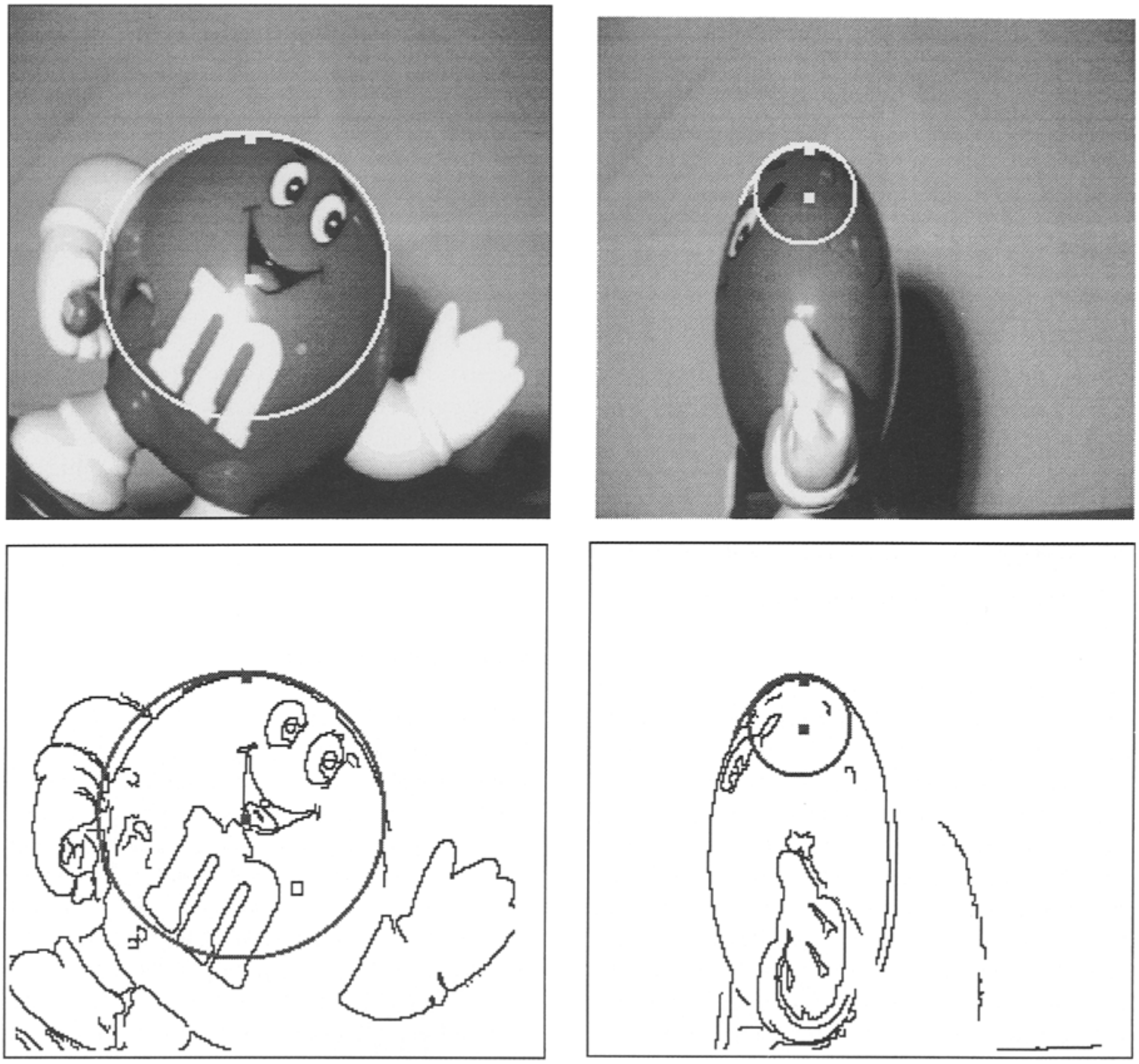


Fig. 17. Viewpoints corresponding to the global minima and maxima of the curvature measurements. Also shown is the computed osculating circle at the tracked point, i.e., the circle that is tangent to the tracked point and has radius equal to $1/k_o$.

recovering quantitative shape information from the occluding contour by an active observer does not necessarily require knowledge of the velocities or accelerations of the observer, but only knowledge of the observer's viewing direction. The reason is that observer motion is not used to merely change the shape of the occluding contour (as in existing approaches), but it is used to change it in a well-defined way, factoring out the need for differential measurements involving observer motion. This is a major step toward qualitative, active vision, allowing the use of a world-centered coordinate frame and requiring knowledge of only relative viewing-direction changes.

Current limitations of the approach are (1) the use of orthographic projection, (2) the requirement that viewing directions change on arbitrary planes, and (3) its applicability to only elliptic or hyperbolic surface points. We believe, however, that our active approach of moving toward viewpoints that are closely related to the geometry of the viewed surfaces is a very important and general one. Consider, for example, the problem of obtaining a "face-on" view of a planar curve (or a texture element). This problem has been studied extensively in the past and several approaches exist that *hypothesize* face-on views, based on information from a single viewpoint—e.g., (Brady & Yuille 1984; Kanade

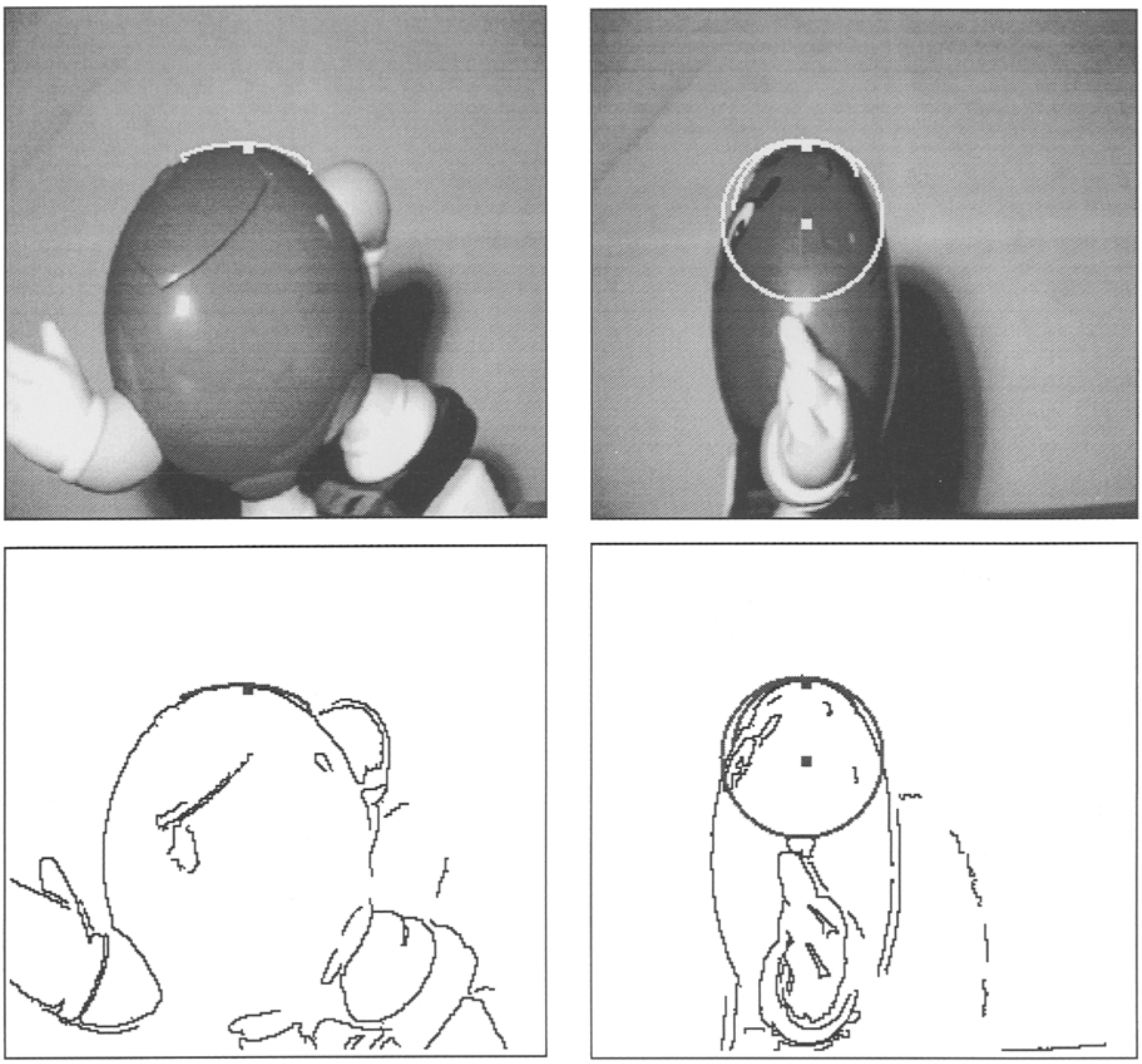


Fig. 18. Viewpoints corresponding to the global curvature minima and maxima for a different run of the tracking process. The extrema were found after averaging the curvature measurements over nine consecutive frames. The osculating circle for the view corresponding to the detected curvature minimum is not shown because it is not fully contained in the image. Note that both curvature measurements have been underestimated, and the global curvature minimum now corresponds to the back side of the object.

1981). We are currently investigating an approach similar to the one presented in this article that enables the observer to change viewpoint in order to obtain a face-on view of a planar curve. We are also trying to extend our results to perspective projection and are investigating possible uses of this approach in the context of active surface exploration (Kutulakos & Dyer 1993).

Acknowledgments

The support of the National Science Foundation under Grant Nos. IRI-9002582 and IRI-9220782 is gratefully acknowledged. We also wish to thank the reviewers for their helpful comments and suggestions.

Appendix: The Extent of Viewing Direction Adjustments

Let us assume we have recovered the principal curvatures at p and the viewing direction ξ is along the principal direction e_2 at p. Now assume that the observer changes viewing direction on the plane of ξ and the surface normal at p in order to introduce new points to the rim. We show that the viewing direction adjustment that will be needed during the shape recovery step is proportional (at a first approximation) to k_{g_2} and inversely proportional to k_{n_2}, the geodesic and normal curvatures of the line of curvature corresponding to e_2. This is an important result because it allows us to predict the performance of this active viewing strategy based on intrinsic properties of the viewed surface. It follows that the performance of our strategy smoothly degrades as the surface becomes more complicated (i.e., k'_{g_2} and k'_{n_2} become large). We first present some concepts from differential geometry for the study of curves on surfaces.

The Local Geometry of Surface Curves

Let $\alpha(s) : I \to \Re^3$ be a curve parameterized by arc length (i.e., $|\alpha'(s)| = 1$). Consider the unit tangent and unit normal vector, $t(s)$ and $n(s)$ respectively, at point $\alpha(s)$. We can describe the curve with two quantities, its curvature $\kappa(s)$ and torsion $\tau(s)$, where $\alpha''(s) = \kappa(s)n(s)$ and $[t(s) \wedge n(s)]' = \tau(s)n(s)$. The vectors $t(s)$, $n(s)$, $t(s) \wedge n(s)$ describe an orthogonal coordinate frame, the *Frenet frame* centered at $\alpha(s)$. This coordinate frame can be used to locally describe the curve based on the values of κ and τ at $\alpha(s)$.

Now let S be a smooth, oriented surface, and let $\bar{\alpha}(s)$ be a smooth curve on S. We can locally describe $\bar{\alpha}(s)$ using a coordinate frame similar to the Frenet frame called the *Darboux frame* (figure 19). Consider a point p on $\bar{\alpha}(s)$. The Darboux frame is defined by $N(p)$, the normal to the surface, $T(p)$, the tangent to $\bar{\alpha}(s)$, and $V(p) = N(p) \wedge T(p)$. Note that the $T - V$ plane is the plane tangent to S. The vector $\bar{\alpha}''(s)$ defining the curvature of $\bar{\alpha}(s)$ can be analyzed in terms of two components, a tangential component (i.e., on $T_p(S)$) in the direction of V, and a normal component in the direction of N. Therefore we can define the curvature of $\bar{\alpha}(s)$ in terms of the curvatures of its projections k_g, k_n on the tangent plane of S and on the $T - N$ plane, respectively. k_g is called the *geodesic curvature of* $\bar{\alpha}(s)$ and k_n is the curvature of the normal section of S in the direction of T. Intuitively, the geodesic curvature measures how far off the $T - N$ plane the curve actually lies. We show that the geodesic curvature of the lines of curvature is closely related to the strategy employed by the observer to select new points for shape recovery. Intuitively, the geodesic curvature of the lines of curvature measures how the arc length of a curve in the e_1 direction changes as one moves along the e_2 direction.

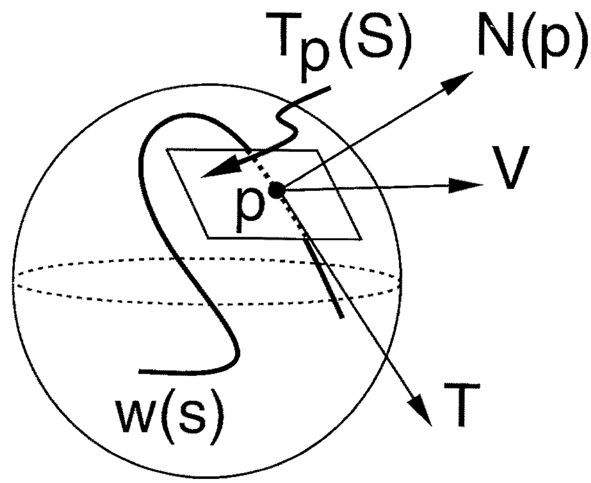


Fig. 19. The Darboux frame. $w(s)$ lies on a sphere S, and p is a point of $w(s)$. $N(p)$ is the surface normal, $T = w'(s)$, and $V = N(p) \wedge T \in T_p(S)$. The Darboux frame is the orthonormal coordinate frame of T, V, N.

The curve $\bar{\alpha}(s)$ can be locally described by the vectors T, N, V and their derivatives. These derivatives can also be expressed in terms of the three frame vectors:

$$\frac{dT}{ds} = k_g V + k_n N \tag{4}$$

$$\frac{dV}{ds} = -k_g T - \tau_g N \tag{5}$$

$$\frac{dN}{ds} = -k_n T + \tau_g V \tag{6}$$

where τ_g is called the *geodesic torsion of* $\bar{\alpha}$.

The Dependence of the Viewing Direction Adjustments on k_{g_2}

Intuitively, the dependence on k_{g_2} is not unexpected: Recall that k_{g_2} measures how far off the plane of ξ and $N(p)$ the line of curvature actually lies. On the other hand, the curve $\beta(s)$ traced by the visual ray that originally passed through p lies on that plane (figure 20). Therefore, one should expect a connection between the angle of $\beta'(s)$ and e_2 and k_{g_2}. The following result shows that there is a very simple relation between them:

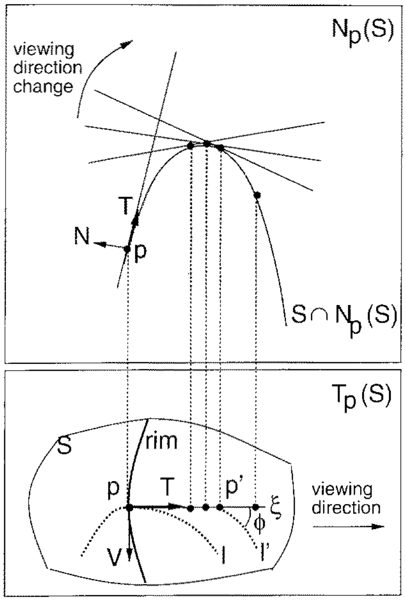


Fig. 20. Changing directions on the *T-N* plane. Top: $N_p(S)$ is the *T-N* plane. The visual ray initially grazes the surface at p in the direction T. N is the surface normal at p. As the viewing directions change on this plane, the visual ray traces the curve $\beta(s) = S \cap N_p(S)$. p' is the new point selected for shape recovery. Bottom: A view of the tangent plane at p. The plane $N_p(S)$ and the traced curve are viewed edge-on. The change in viewing direction stops when the visual ray grazes p'. l and l' are the lines of minimum curvature passing through p and p', respectively. The shape-recovery step will require a rotation by an angle ϕ on the tangent plane at p' in order to align the viewing direction ξ with e_2 at p'.

Proposition 2. (1) Let $\beta(s)$ be the intersection of S with the plane defined by ξ and $N(p)$ ($\beta(0) = p$, $\beta'(0) = \xi$). If ξ is along the principal direction e_2 then

$$k_{g_2} = \frac{d\phi}{ds} \tag{7}$$

where k_{g_2} is the geodesic curvature of the line of curvature along e_2 at point p, and $\phi(s)$ is the angle between $\beta'(s)$ and the second principal direction at $\beta(s)$.

(2) Let $\xi' = \beta'(s)$ be the new viewing direction on the plane of ξ and $N(p)$. If θ is the angle between ξ and ξ', then for values of θ close to 0 we have

$$\phi(\theta) \approx \frac{k_{g_2}}{k_{n_2}} \sin\theta \tag{8}$$

Proof: (1) The Darboux trihedron for $\beta(0)$ is composed of the vectors $T(0) = \beta'(0)$, $N(\beta(0)) = N(p)$, and $V(0) = N(\beta(0)) \wedge T(0)$, where $N(\cdot)$ is the Gauss map for the surface. We use a second-order Taylor series expansion and equations (4)–(6) to find $T(s) = \beta'(s)$ with respect to $T(0) = \xi$:

$$T(s) - T(0) \approx s\frac{dT}{ds} + \frac{s^2}{2}\frac{d^2T}{dT^2} \tag{9}$$

$$\approx s(k_gV + k_nN) + \frac{s^2}{2}(k_gV + k_nN)' \tag{10}$$

$$\approx -\frac{s^2}{2}(k_g^2 + k_n^2)\,T + \left[sk_g + \frac{s^2}{2}(k_g' + k_n\tau_g) \right] V + \left[sk_n + \frac{s^2}{2}(k_n' - k_g\tau_g) \right] N \tag{11}$$

where all coefficients of s are evaluated at $\beta(0)$. Now note that $\beta(s)$ is always on the $T - N$ plane and therefore $T(s) \cdot V(0) = T(0) \cdot V(0) = 0$, or $[T(s) - T(0)] \cdot V(0) = 0$ for all s. Constraining the *V*-component of equation (11) to be identically equal to zero we get

$$k_g = 0 \tag{12}$$

The geodesic curvature k_g can be expressed in terms of the geodesic curvatures of the lines of curvature using Liouville's formula:

$$k_g = k_{g_1}\cos\psi + k_{g_2}\sin\psi + \frac{d\psi}{ds} \tag{13}$$

where ψ is the angle between $\beta'(0)$ and e_1. But $\beta'(0)$ is equal to e_2, and therefore $\psi = \pi/2$. Noting that $\phi = \pi/2 - \psi$ and combining equations (12) and (13), we get the desired result. □

(2) ξ' will be tangent to $\beta(s)$ for some s. Therefore, $\xi' = T(s)$. We use equation (11) to get a first-order approximation of s for values close to 0:

$$s \approx \frac{[T(s) - T(0)] \cdot N}{k_{n_2}} \tag{14}$$

Note that $[T(s) - T(0)] \cdot N$ equals $\sin\theta$, where θ is defined as above. Now using equation (7) and a first-order approximation for $\phi(\theta)$ we get the desired result. □

Finally, we can draw three conclusions from equation (8):

1. If $k_{g_2} = 0$, no viewing direction adjustments will be done during the shape recovery phase if the viewing direction changes in the plane of ξ and $N(p)$ in the second step. The curve $\beta(s)$ traces a part of the line of curvature associated with e_2. This can happen only if that line of curvature is also a geodesic.
2. $\phi(\theta)$ can grow arbitrarily large with decreasing values of k_{n_2}. If k_{n_2} is close to 0, the surface is locally flat in the e_2 direction. Therefore, in such a case the approximation is not valid. However, this problem is inherent to the use of the occluding contour for shape recovery in the case of almost flat surfaces. The reason is that if the surfaces are locally flat, surface points will enter and leave the rim at arbitrarily large rates. This problem will also exist for methods that measure image velocities in the vicinity of the rim—e.g., (Cipolla & Blake 1992)—since they require that the image points or features are not widely separated on the surface.
3. Equation (8) can also be used as a means to approximate k_{g_2}: After a small rotation by θ in the plane of $N(p)$ and ξ, the shape-recovery step will produce a value for $\phi(\theta)$. Hence, we can use the equation to approximate k_{g_2}. This means that we will be able to completely describe the line of curvature corresponding to e_2 in the vicinity of the previously selected point.

Notes

1. The quantity k_n is also referred to as the *radial curvature* of S at p.
2. This is also mentioned in (Koenderink 1990).
3. Our use of polyhedral models was only for convenience in generating the occluding contour. The implementation of our algorithms did not exploit the polyhedral property of the models.
4. Occlusion occurs either when the tracked point becomes occluded by a distant point on the object's surface, or when the observer's viewing direction becomes aligned with an asymptote of the tracked point. In the former case, the tracked point will project to a T-junction, while in the latter case a cusp will be formed. The ability to distinguish between these two cases can be used to verify the "sidedness" of the occluding contour, that is, on which side of the contour the surface lies. For example, if the selected point is erroneously assumed elliptic, the contour will cusp at the point's projection when the shape-recovery algorithm of section 4 is applied to that point.

References

Aloimonos, Y. 1990. Purposive and qualitative active vision, *Proc. 10th Intern. Conf. Patt. Recog.*, Atlantic City, NJ, pp. 346-360.

Aloimonos, Y., Weiss, I., and Bandyopadhyay, A. 1987. Active vision, *Proc. 1st Intern. Conf. Comput. Vis.*, London, pp. 35–54.

Bajcsy, R. 1988. Active perception, *Proc. IEEE* 76(8):996–1005.

Ballard, D.H. 1989. Behavioural constraints on animate vision, *Image Vis. Comput.* 7(1):3–9.

Ballard, D.H. 1989. Reference frames for animate vision, *Proc. 11th Intern. Joint Conf. Artific. Intell.*, Detroit, pp. 1635–1641.

Ballard, D.H. 1991. Animate vision, *Artificial Intelligence* 48:57–86.

Ballard, D.H., and Brown, C.M. 1992. Principles of animate vision, *Comput. Vis. Graph., Image Process.: Image Understanding* 56(1):3–21.

Ballard, D.H., and Ozcandarli, A. 1988. Eye fixation and early vision: Kinetic depth, *Proc. 2nd Intern. Conf. Comput. Vis.*, Tarpon Springs, FL, pp. 524–531.

Barrow, H.G., and Tenenbaum, J.M. 1981. Interpreting line drawings as three-dimensional images, *Artificial Intelligence* 17:75–116.

Blake, A., Curwen, R., and Zisserman, A. 1993. A framework for spatio-temporal control in the tracking of visual contours, *Intern. J. Comput. Vis.*, 11(2):127–145.

Brady, M., and Yuille, A. 1984. An extremum principle for shape from contour, *IEEE Trans. Patt. Anal. Mach. Intell.* 6(3):288–301.

Brady, M., Ponce, J., Yuille, A., and Asada, H. 1985. Describing surfaces, *Comput. Graph. Image Process.* 32:1–28.

Brooks, R.A. 1986. A robust layered control system for a mobile robot, *IEEE J. Robotics Automat.* 2(1):14–23.

Brooks, R.A. 1989. A robot that walks: Emergent behaviors from a carefully evolved network, *Proc. IEEE Robotics Automat. Conf.*, Scottsdale, AZ, pp. 692–696.

Cipolla, R., and Blake, A. 1992. Surface shape from the deformation of apparent contours, *Intern. J. Comput. Vis.* 9(2):83–112.

Conte, S.D., and de Boor, C. 1972. *Elementary Numerical Analysis*. McGraw-Hill: New York.

Coombs, D. and Brown, C. 1993. Real-time binocular smooth pursuit, *Intern. J. Comput. Vis.*, 11(2):147–164.

doCarmo, M.P. 1976. *Differential Geometry of Curves and Surfaces*. Prentice-Hall: Englewood Cliffs, NJ.

Giblin, P., and Weiss, R. 1987. Reconstruction of surfaces from profiles, *Proc. 1st Intern. Conf. Comput. Vis.*, London, pp. 136–144.

Grosso, E., and Ballard, D.H. 1993. Head-centered orientation strategies in animate vision, *Proc. 4th Intern. Conf. Comput. Vis.*, Berlin, pp. 395–402.

Hager, G., and Mintz, M. 1987. Searching for information, *Proc. Workshop on Spatial Reasoning and Multi-Sensor Fusion*, Charles, IL, pp. 313–322.

Horaud, R., and Brady, M. 1987. On the geometric interpretation of image contours, *Proc. 1st Intern. Conf. Comput. Vis.*, London, pp. 374–382.

Horn, B.K.P. 1986. *Robot Vision*. MIT Press: Cambridge, MA.

Kanade, T., 1981. Recovery of the three-dimensional shape of an object from a single view, *Artificial Intelligence* 17:409–460.

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

Koenderink, J.J. 1990. *Solid Shape*. MIT Press: Cambridge, MA.

Kriegman, D.J. 1993. Personal communication.

Krotkov, E. 1987. Focusing, *Intern. J. Comput. Vis.* 1(3):223–237.

Krotkov, E., and Bajcsy, R. 1993. Visual behaviors for reliable ranging: Cooperating focus, stereo, and vergence, *Intern. J. Comput. Vis.*, 11(2):187-203.

Kutulakos, K.M., and Dyer, C.R. 1993. Global surface reconstruction by purposive control of observer motion. *Proc. Conf. Comput. Vis. Patt. Recog.*, Seattle, WA, to appear.

Leyton, M. 1988. A process grammar for shape, *Artificial Intelligence* 34:213–247.

Malik, J. 1987. Interpreting line drawings of curved objects, *Intern. J. Comput. Vis.* 1(1):73–103.

Marr, D., and Nishihara, H.K. 1978. Visual information processing: Artificial intelligence and the sensorium of sight, *Technology Review* 81:2–23.

Nelson, R.C., and Aloimonos, Y. 1989. Obstacle avoidance using flow field divergence, *IEEE Trans. Patt. Anal. Mach. Intell.* 11(10):1102–1106.

Papanikolopoulos, N.P., Khosla, P.K. and Kanade, T. 1993. Visual tracking of a moving target by a camera mounted on a robot: A combination of control and vision, *IEEE Trans. Robotics Automat.* 9(1):14–35.

Ponce, J. and Chelberg, D. 1987. Finding the limbs and cusps of generalized cylinders, *Intern. J. Comput. Vis.* 1(3):195–210.

Ponce, J., Chelberg, D., and Mann, W.B. 1989. Invariant properties of straight homogeneous generalized cylinders and their contours, *IEEE Trans. Patt. Anal. Mach. Intell.* 11(9):951–966.

Richards, W., Koenderink, J.J., and Hoffman, D.D. 1988. Inferring 3D shapes from 2D silhouettes. In *Natural Computation*, W. Richards, ed., pp. 125–136, MIT Press: Cambridge, MA.

Richetin, M., Dhome, M., Lapreste, J.T., and Rives, G. 1991. Inverse perspective transform using zero-curvature contour points: Application to the localization of some generalized cylinders from a single view, *IEEE Trans. Patt. Anal. Mach. Intell.* 13(2):185–192.

Rimey, R.D., and Brown, C.M. 1993. Control of selective perception using Bayes nets and decision theory, *Intern. J. Comput. Vis.*, this issue.

Stevens, K.A. 1981. The visual interpretation of surface contours, *Artificial Intelligence* 17:47–73.

Ulupinar, F., and Nevatia, R. 1988. Using symmetries for analysis of shape from contour, *Proc. 2nd Intern. Conf. Comput. Vis.*, Tarpon Springs, FL, pp. 414–426.

Vaillant, R., and Faugeras, O.D. 1992. Using extremal boundaries for 3-D object modeling, *IEEE Trans. Patt. Anal. Mach. Intell.* 14(2):157–173.

Wixson, L.E., and Ballard, D.H. 1993. Exploiting world structure to efficiently search for objects, *Intern. J. Comput. Vis.*, this issue.