Building Global Surface Models
By Purposive and Qualitative Viewpoint Adjustment

Kiriakos N. Kutulakos
kyros@cs.wisc.edu
Computer Sciences Department
University of Wisconsin
Madison, Wisconsin 53706 USA

W. Brent Seales
seales@ms.uky.edu
Computer Sciences Department
University of Kentucky
Lexington, Kentucky 40536 USA

Charles R. Dyer
dyer@cs.wisc.edu
Computer Sciences Department
University of Wisconsin
Madison, Wisconsin 53706 USA

ABSTRACT

We present an approach for recovering a global surface model of an object from the deformation of the occluding contour using an active (i.e., mobile) observer able to control its motion. In particular, we consider two problems: (1) How can the observer's viewpoint be controlled in order to generate a dense sequence of images that allows incremental reconstruction of an unknown surface, and (2) how can we construct a global surface model from the generated image sequence? We achieve the first goal by *purposefully* and *qualitatively* controlling the observer's instantaneous direction of motion in order to control the motion of the visible rim over the surface. We achieve the second goal by using a stationary calibrated trinocular camera rig and a mechanism for controlling the relative position and orientation of the viewed surface with respect to the trinocular rig. Unlike previous shape-from-motion approaches which derive quantitative shape information from an *arbitrarily* generated sequence of images, we develop a collection of simple and efficient viewing strategies that allow the observer to achieve the global reconstruction goal by maintaining specific geometric relationships with the viewed surface. These relationships depend only on tangent computations on the occluding contour. To demonstrate the feasibility and effectiveness of our approach we apply the developed algorithms to synthetic and real scenes.

1 Introduction

There has been considerable interest in approaches that recover information about the structure of a scene from sequences of images, assuming an observer in motion (e.g., work on optical flow and shape-from-motion[6]). One common feature of these approaches is that they use a dense sequence of images produced by an observer undergoing arbitrary, but known, motion. Unfortunately, arbitrary motion implies that the parts of the scene reconstructed are not under the observer's control.

In this paper we combine the shape-from-motion paradigm with the purposive and qualitative vision paradigm[2] in order to recover a *global* surface description of an object by purposefully controlling the viewpoint of an active observer and by using the occluding contour. In particular, we answer two questions: (1) How can the observer's viewpoint be controlled in order to generate a dense sequence of images that allows incremental reconstruction of the surface, and (2) how can we build a global surface model from the generated sequence of images?

Contrary to previous shape-from-motion approaches which attempt to derive quantitative local surface shape information using an *arbitrarily* generated sequence of images, the central idea of our approach comes from studying the global surface reconstruction problem at two levels of abstraction: At the higher, *qualitative* level we consider the task of "intelligently" controlling the observer's viewpoint in order to generate a dense image sequence permitting global reconstruction. At the lower, *quantitative* level we consider the task of building a global model of the viewed surface from the generated sequence. We achieve our higher-level goal by employing simple observer behaviors[3] for controlling viewpoint that allow the observer to maintain specific geometric relationships with the viewed scene[9,11]; we achieve our lower-level goal by using a calibrated trinocular camera rig and a mechanism for controlling the relative position and orientation of the viewed surface with respect to the trinocular rig.

Our viewpoint control behaviors are developed in the context of three increasingly more general reconstruction tasks: The *local reconstruction task*, where the observer must control viewpoint to reconstruct a patch around a selected surface point; the *incremental reconstruction task*, where the observer must control viewpoint to iteratively "grow" a reconstructed region on the surface; and the *global reconstruction task*, where an entire reconstructible region on the surface must be reconstructed. All these tasks require the use of multiple viewpoint control behaviors that must be appropriately integrated.

Our work builds directly on work on the occluding contour by Giblin and Weiss,[8] Cipolla and Blake,[6] Vaillant and Faugeras,[15] and Koenderink.[10] The occluding contour is the projection of the visible rim, the set of visible points at which the line of sight is tangent to the surface. Under continuous observer motion the visible rim slides over the surface, affecting the geometry of the occluding contour and revealing shape information for the parts of the surface over which the visible rim slides. This evolution of the visible rim depends on the local and global shape of the surface as well as the motion of the observer.

Previous shape-recovery approaches using the occluding contour studied how its deformation over a dense sequence of frames can be used to recover local shape information for points projecting to the occluding contour.[6,15] Until now, however, this work has not been extended to reconstruct complete surface models and assumes arbitrary observer motion: Since the contour's deformation can be used only for reconstructing the parts of the surface over which the continuously-evolving visible rim slides, the shape recovery capabilities of the observer become severely constrained if the observer's motion cannot be controlled.[†]

In this paper we show that by combining a collection of qualitative viewpoint control behaviors with a method for recovering quantitative shape information from the occluding contour a description of the viewed surface can be incrementally constructed. This allows us to provably perform the local reconstruction and the incremental reconstruction tasks when the the surface is unknown, is non-convex, and even self-occludes.[‡] Our approach therefore uses a very simple principle: Since any attempt at reconstruction from the occluding contour using arbitrary observer motion will miss parts of the surface, we simply let the surface itself determine how to view it. It is quite fortunate that the occluding contour provides all the information necessary to achieve the global reconstruction goal: Recent results demonstrate that the occluding contour can be efficiently tracked,[4] and that shape information can be recovered from the deformation of the occluding contour even with a monocular observer.[6,7] Furthermore, our experimental results show that the occluding contour can be reliably detected in edge images, can be used to recover quantitative shape information and, moreover, can guide the purposive viewpoint adjustment process. Consequently, instead of using mechanisms that require sophisticated sensors (e.g., range sensors or laser scanners) to reconstruct the scene from a single viewpoint or a small number of viewpoints,[1,13,16] we control viewpoint in a simple and efficient manner that allows us to use the occluding contour for reconstruction.

The significance of our approach lies in the use of an active observer that purposefully controls viewpoint in order to achieve and maintain a well-defined relationship with respect to a 3D shape prior to its recognition. This relationship is characterized by simple image-computable quantities derived from the occluding contour (tangent at a contour point). We show how the motion

[†]This is a major reason why previous work on surface reconstruction from the occluding contour has not considered the problem of recovering a description of an entire surface. Experimental work on the subject has either been applied to uncluttered scenes containing geometrically-trivial surfaces such as ovoids, or has been applied to more complicated surfaces such as surfaces of revolution under the assumption that the motion of the observer was not arbitrary (e.g., the observer moves roughly perpendicular to the axis of revolution[6]).

[‡]For a general treatment of the global reconstruction task see recent work by Kutulakos and Dyer.[12]

of the visible rim over the surface can be characterized qualitatively and how the observer can control this motion by controlling viewpoint in a simple and efficient manner. This makes it possible to force the visible rim to slide over previously-unreconstructed parts of the surface. Furthermore, we show how the generated sequence of images can be used to reliably construct a complete model of the surface and discuss the limitations of current reconstruction approaches that assume arbitrary observer motion. Our purposive approach is therefore a major step toward efficient and reliable global model reconstruction in complex scenes whose structure cannot be predicted beforehand.

The rest of this paper is organized as follows. The next section reviews basic terminology. Section 3 presents the basic step of our approach, developing a collection of simple and qualitative viewing strategies that allow the shape of the surface to be recovered in a whole patch of points around a selected visible rim point. The strategies developed follow principles similar to those presented by Kutulakos and Dyer.[11] These basic strategies are used in Section 4 to define an incremental reconstruction strategy that guides the observer's motion in order to expand the set of surface points reconstructed. To demonstrate the applicability and effectiveness of our approach, we apply the developed strategies to synthetic scenes. Finally, Section 5 presents our approach to the model construction problem and Section 6 presents results from applying this approach to real scenes, briefly discussing in the context of these results the limitations of approaches using arbitrary observer motion for model construction.

2 Local surface shape from occluding contour

Let S be an opaque, smooth and oriented surface bounding a volume V in $\Re^3$. We assume that S is viewed under spherical projection. The space of viewpoints in this case is identical to the set of observer positions, i.e., the set $\Re^3 - V$. If $\mathbf{x}(s,t)$ is a local parameterization of a patch Π of S, the partial derivatives $\mathbf{x}_s(p), \mathbf{x}_t(p)$ of $\mathbf{x}$ with respect to s and t define $T_p(S)$, the plane tangent to S at $p \in \Pi$. The local shape of the surface at p is determined by the first and second fundamental forms $I(p), II(p)$ at p which can be computed from $\mathbf{x}$.[5]

The *rim* of S is the set of surface points for which $T_p(S)$ contains the line segment connecting p and the observer's viewpoint, c. The visible rim consists of the rim points that are visible. The *occluding contour* of S is the projection of the visible rim on the

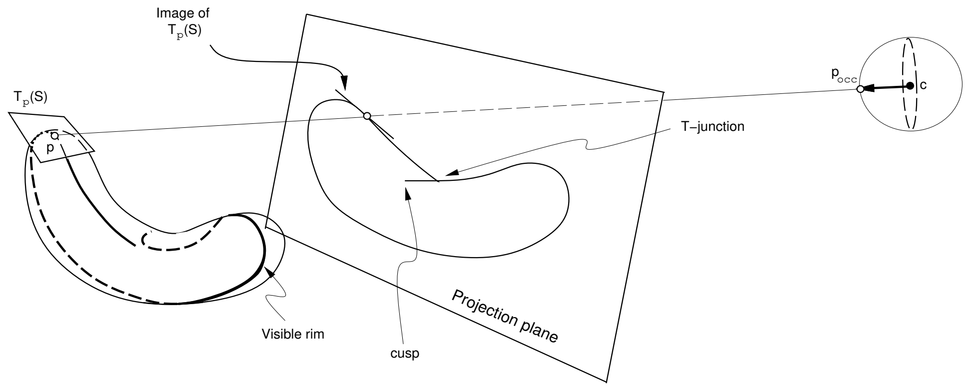


Figure 1: Viewing geometry. The projection, p_{occ}, of a point p on the spherical image can be thought of either as a point on the unit sphere or as a unit vector in the direction of the ray passing through p and the observer's viewpoint, c. The visible rim and the occluding contour corresponding to the projection of a bean-shaped surface are shown.[8] The occluding contour consists of a single curve whose endpoints are a T-junction and a cusp. For simplicity, we show the visible rim projected to a planar image perpendicular to p_{occ}.

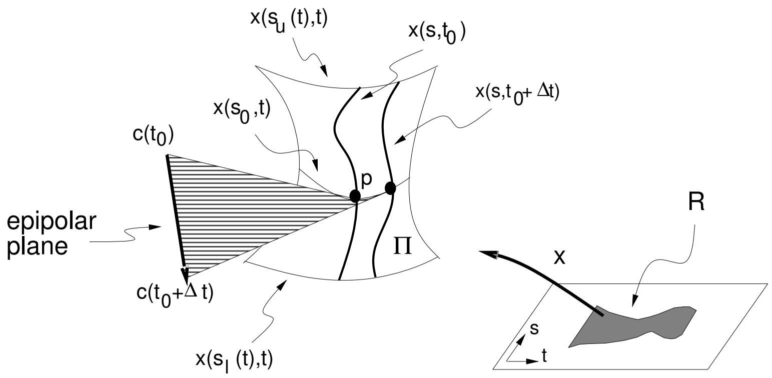


Figure 2: Epipolar parameterization for a surface patch Π around point p. The parameterization $\mathbf{x}$ defines a mapping from R to Π. Curves $\mathbf{x}(s, t_0)$ and $\mathbf{x}(s, t_0 + \Delta t)$ are curves on the visible rim of the surface corresponding to observer positions $c(t_0)$ and $c(t_0 + \Delta t)$, respectively. The tangent to the curve $\mathbf{x}(s_0, t)$ for $t = t_0$ is along the line through $c(t_0)$ and p. The curve's normal is in the epipolar plane defined by the direction of motion, $\mathbf{v}(t)$, and the line connecting $c(t_0)$ and p.

image (Figure 1). Generically, the occluding contour is a collection of open and closed smooth curves for almost all positions of the observer. The endpoints of open occluding contour curves are either cusps or T-junctions.[10] The shape and topology of this collection of curves depends on S and the observer's viewpoint. Under continuous observer motion and when the topology of the visible rim does not change, the visible rim "slides" over the surface, tracing patches. These surface patches can be represented using the *epipolar parameterization* which describes the motion of the visible rim over the surface (see Figures 2 and 8(a)). This parameterization has been used by several researchers[6,8] to derive the fundamental forms of the surface for all points lying on these patches from the deformation of the visible rim's projection, i.e., the occluding contour.

3 Local reconstruction by purposive observer motion

In our approach, both global and incremental reconstruction is performed by integrating a collection of qualitative viewpoint control behaviors, each of which performs an instance of the following local reconstruction task: Suppose the observer is at position c, and let p be a visible rim point on the surface that is identified by its projection, p_{occ}, on the occluding contour. The task of the observer is to continuously control viewpoint, starting from point c, in order to reconstruct the surface in *some* neighborhood Π of p using the epipolar parameterization.[†] Given a collection of basic behaviors that provably perform this task, incremental surface reconstruction is achieved by appropriately combining them (Section 4).

Our approach is based on the following three observations:

- If p is not the endpoint of a visible rim curve and the topology of the visible rim does not change in the neighborhood of p when the observer's viewpoint is infinitesimally perturbed, the motion of the observer can be controlled so that there exists a neighborhood Π of p such that the surface in Π can be described by the epipolar parameterization.
- If p is the endpoint of a visible rim curve, the epipolar parameterization cannot describe the surface in the neighborhood of

[†]The formulation of the local reconstruction task is similar to the one reported by Kutulakos and Dyer,[11] where a visible rim point p was selected and the surface shape at p was recovered by controlling the observer's viewpoint. The difference here is that the shape of the surface in a whole patch around p is recovered, instead of just the shape at p.

p. Although p belongs to the visible rim when viewed from a two-dimensional set of viewpoints on the tangent plane of the surface at p, the position of p within the visible rim curve it belongs to is *not* fixed when the viewpoint of the observer varies in this set: There are positions on the tangent plane of p such that p belongs to the visible rim and is not the endpoint of a visible rim curve when viewed from those positions.

- The point p and the observer's viewpoint may be such that the occluding contour's topology changes in the neighborhood of p under an infinitesimal viewpoint perturbation. For all points p except those lying on a collection of special sets (which are curves for generic surfaces), the observer can move to other viewpoints on p's tangent plane at which the contour's topology does not change in the neighborhood of p when these viewpoints are infinitesimally perturbed. The special sets for which this cannot be achieved bound the reconstructible surface regions.[12]

Based on these observations, for any given viewpoint c we distinguish four types of visible rim points: *Ordinary points*, which are not endpoints of a visible rim curve; *cusp points*, which are visible rim endpoints projecting to a cusp on the occluding contour; *T-junction points*, which are visible rim endpoints projecting to a T-junction on the occluding contour; and *degenerate points*, where the occluding contour's topology changes in the vicinity of their projection if the viewpoint c is infinitesimally perturbed. The basic step of our approach is to control the motion of the observer in order to deal with the case where p is an ordinary visible rim point; this strategy is described in Section 3.1. The other three cases are treated by (1) controlling the position of the observer in order to reach a viewpoint where p is an ordinary visible rim point, and (2) applying the basic step in order to recover the shape of the surface in a neighborhood of that point. Due to lack of space, in Section 3.2 we only present how the first of these three cases is handled. The other two cases can also be handled by following a similar approach.[12]

3.1 Local Reconstruction Around Ordinary Points

This section presents the crucial link between the ability to control the motion of the observer and the problem of guiding the surface reconstruction process. We assume that the observer is initially positioned at point c and has selected an ordinary visible rim point p. The task of the observer is to move in a way that forces the visible rim to slide over all points in a neighborhood of p. This corresponds to the tasks of controlling motion in order to (1) force the visible rim to slide over all points in this neighborhood that are occluded from viewpoint c, and (2) force the visible rim to slide over all points in this neighborhood that are visible from viewpoint c.

To see how these two tasks can be performed, suppose the observer moves in a continuous fashion by tracing a smooth curve $c(t), c = c(0)$, and let $\mathbf{v} = c'(t)$ be the instantaneous direction of motion. Given a segment $\beta(t)$ of the visible rim at viewpoint $c(t)$, the epipolar parameterization allows us to define the segment $\beta(t + \delta t)$ of the visible rim at $c(t + \delta t)$ that corresponds to $\beta(t)$. The following result provides a qualitative characterization of the motion of the visible rim over the surface (Figure 3):

THEOREM 3.1 (VISIBILITY TRANSITION DYNAMICS). *Suppose* $\mathbf{n}(p)$ *is the surface normal at point p. If* $\beta^+(t)$ *and* $\beta^-(t)$ *are two visible rim segments satisfying* $\mathbf{n}(p) \cdot \mathbf{v}(t) > 0$ *and* $\mathbf{n}(p) \cdot \mathbf{v}(t) < 0$, *respectively, then*

1. *all points of* $\beta^+(t + \delta t)$ *are occluded from position* $c(t)$,
2. *all points of* $\beta^-(t + \delta t)$ *are visible from position* $c(t)$,
3. *all visible rim points satisfying* $\mathbf{n}(p) \cdot \mathbf{v}(t) = 0$ *will remain on the rim at* $c(t + \delta t)$.

When the observer continuously changes viewpoint along a smooth curve $c(t)$ ($t_{start} \leq t \leq t_{end}$), and the topology of the visible rim does not change, the segment $\beta(t)$ will trace a patch Π on the surface around p that can be described using the epipolar parameterization. The boundary of this patch consists of the segments $\beta(t_{start}), \beta(t_{end})$ contained in the visible rim at viewpoints $c(t_{start}), c(t_{end})$, respectively, and the traces of the endpoints of $\beta(t)$. The endpoints of $\beta(t)$ will either be points satisfying $\mathbf{n}(p) \cdot \mathbf{v}(t) = 0$, or will be the endpoints of a visible rim curve. The following simple, qualitative strategy for controlling the motion of the observer can now be used to perform the local reconstruction task for an ordinary point p (Figure 4):

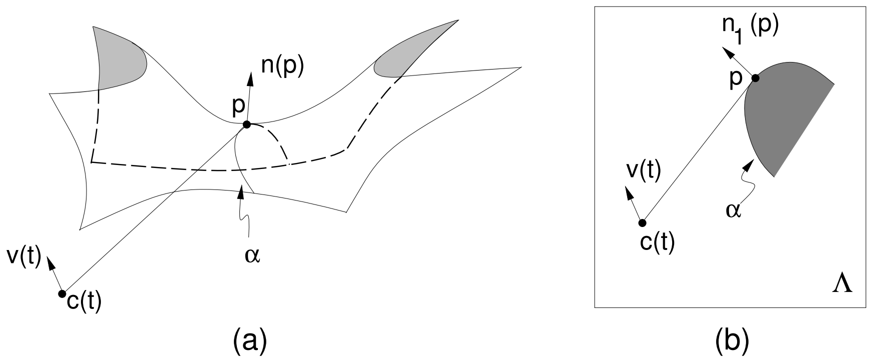


Figure 3: Inducing the visibility of points in a neighborhood of a hyperbolic point p. (a) The curve α is the curve of intersection of S with the epipolar plane Λ. If $\mathbf{v}(t) \cdot \mathbf{n}(p) > 0$, the visible rim point contained in Λ will move toward the previously-occluded portion of α under an infinitesimal viewpoint change along $\mathbf{v}(t)$. (b) A face-on view of the plane Λ. The outward normal $n_1(p)$ of α at p is the projection of $\mathbf{n}(p)$ on the plane Λ. The geometry of the intersection of S with the epipolar planes corresponding to visible rim points *close* to p is also similar to the one shown.

Ordinary Patch Reconstruction Strategy

Step 1: Select a point p on the visible rim that is not the endpoint of a visible rim curve. This selection is done indirectly by selecting p's projection, p_{occ}, on the occluding contour. Point p_{occ} must not be the endpoint of a smooth occluding contour curve.

Step 2: Compute the surface normal at p. The normal is given by $\mathbf{T} \wedge p_{occ}$, where $\mathbf{T}$ is the tangent to the occluding contour at p_{occ}.[10]

Step 3: (Reconstructing the occluded points near p.) Select a direction $\mathbf{v}$ that satisfies the inequality $\mathbf{n}(p) \cdot \mathbf{v} > 0$. Perform a small viewpoint adjustment along $\mathbf{v}$ while monitoring the deformation of the occluding contour curve that initially contains p_{occ}.

Step 4: (Reconstructing the visible points near p.) Move back to the initial viewpoint and reapply Step 3 by selecting a vector $\mathbf{v}$ that satisfies the inequality $\mathbf{n}(p) \cdot \mathbf{v} < 0$.

3.2 Local Reconstruction Around Cusp Points

In this section we consider the local reconstruction task around cusp points by studying the geometrical changes on the occluding contour as the observer moves on the selected point's tangent plane. The main idea is similar to the one used by Kutulakos and Dyer[11]: If the observer moves along specific directions on the tangent plane of the selected point p, p will remain on the visible rim but will cease to be the endpoint of the visible rim curve containing it. Hence, after the observer performs such a viewpoint adjustment, the reconstruction problem around p is transformed to the reconstruction problem around an ordinary visible rim point.

In particular, let p_{occ} be the projection of p and let $C(s)$ be the occluding contour curve containing p_{occ}, with $p_{occ} = C(s_0)$. The following result characterizes the special directions for adjusting the observer's viewpoint that force p to become an ordinary point (Figure 5):

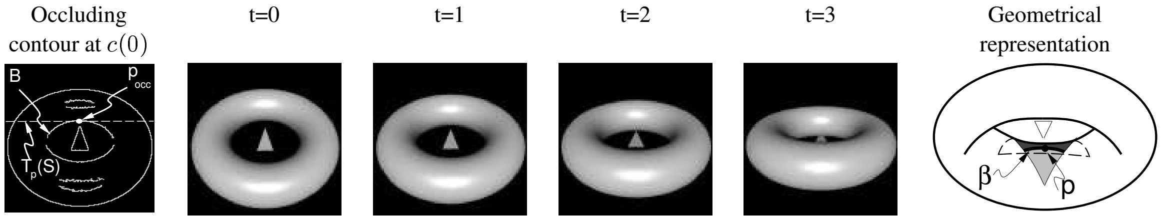


Figure 4: Reconstructing a patch around an ordinary visible rim point on a torus. The leftmost image shows the edges detected in the initial view, $t = 0$. The small triangle in the middle of the torus points toward the direction of the line connecting the initial viewpoint of the observer and the center of the torus. The point selected is point p, shown in the rightmost figure, in which the torus is viewed from below. The point is selected by selecting its projection, p_{occ}, on the occluding contour from the initial viewpoint, $c(0)$. The figure shows the views of the surface as the observer executes Step 3 of the Ordinary Patch Reconstruction Strategy. The tangent to the occluding contour at p_{occ} is horizontal and, hence, the projection of the surface normal at p in the image is a vertical line. The observer adjusts viewpoint by moving vertically downward. B is the projection of the visible rim segment β that contains p and moves toward the initially-occluded portion of the torus in the neighborhood of p according to Theorem 3.1. Here the observer moves until the visible rim curve $\beta(t)$ corresponding to β disappears, i.e., until β becomes occluded. The endpoints of $\beta(t)$ are T-junction points. Since $\beta(t)$ disappears in this case, after the execution of Step 3 the patch reconstructed on the surface is bounded by the curves traced by the endpoints of $\beta(t)$ and by $\beta(0)$ (i.e., a triangle-like patch). The patch is shown as the lightly-shaded area on the rightmost figure. Step 4 completes the reconstruction process around p by reconstructing a patch on the other side of $\beta(0)$ (shown as the darkly-shaded area on the rightmost figure).

THEOREM 3.2. *Let* $\mathbf{T}$ *be the tangent to the occluding contour at point* p_{occ} *defined as the limit* $\lim_{s \to s_0} C'(s)$*. If the observer performs an infinitesimal viewpoint change along a direction* $\mathbf{v}$ *in* $T_p(S)$ *such that* $\mathbf{v} \cdot \mathbf{T} > 0$*,* p *will be an ordinary visible rim point at the observer's new viewpoint.*

The proof of Theorem 3.2 extends the analysis of Kutulakos and Dyer[11] and is omitted due to lack of space. Theorem 3.2 gives a qualitative strategy for adjusting the viewpoint of the observer in order to recover the shape in a neighborhood of p. We only need to answer two additional questions to completely specify such a strategy: how to move and when to stop. The only computation necessary for selecting the direction of motion is the tangent to the occluding contour at p_{occ}. This decision is particularly simple if the distance between the observer and p is kept constant. In this case, the motion decision corresponds to deciding whether to move clockwise or counterclockwise on the unit circle centered at p. Although p will become an ordinary point after an infinitesimal viewpoint adjustment, such an adjustment can leave p arbitrarily close to the visible rim's endpoint. A simple stopping condition is to continue moving on $T_p(S)$ until the distance between p's projection and the two endpoints of the occluding contour curve containing it is maximized.[†] This analysis results in the following simple viewing strategy (Figure 6):

Cusp Patch Reconstruction Strategy

Step 1: Compute the tangent T to the occluding contour at the selected occluding contour point p_{occ}.

Step 2: Compute the normal $\mathbf{n}(p)$ of the surface at the corresponding visible rim point p.

Step 3: Determine whether a clockwise or counterclockwise motion $\mathbf{v}$ on $T_p(S)$ satisfies the inequality $\mathbf{v} \cdot \mathbf{T} > 0$.

[†] This strategy gracefully degrades when p approximates parabolic points on the surface, where the angle between the point's asymptotes can tend to zero, or when the line connecting p and c approaches a line with one third-order and one second-order contact with the surface.[10,12]

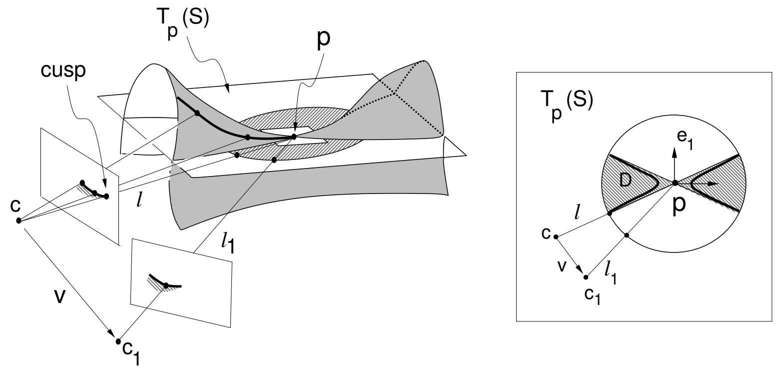


Figure 5: (a) Moving on the tangent plane of a cusp point p. The bold curve on the surface is the visible rim curve containing p. c is the initial position of the observer. $\mathbf{v}$ is a motion direction on $T_p(S)$ satisfying the conditions of Theorem 3.2: p is an ordinary visible rim point when viewed from viewpoint c_1. (b) A top view of the tangent plane at p is shown. e_1 is the first principal direction of the surface at p. D is the Dupin indicatrix at p. The line segment l connecting c and p is along an asymptote of D. p is occluded when l lies in the shaded areas.

Step 4: Consider the occluding contour curve containing the projection of p as consisting of two segments, C_l, C_r, one to the left and one to the right of p's projection. Move while fixating at p until the length of the shortest of the two segments is maximized.

Step 5: Apply the Ordinary Patch Reconstruction Strategy to reconstruct a surface patch around p.

4 Incremental surface reconstruction

An incremental reconstruction strategy must guide the motion of the observer so that new patches on the surface are reconstructed. In order to achieve this using only the occluding contour we need to answer two questions: (1) How can the observer force points on the boundary of the already-reconstructed patches to lie on the visible rim, and (2) how can the observer control its motion so that new patches around those points can be recovered? We briefly explain below how these questions can be answered using the strategies developed in the paper.

The first question can be answered by considering the fact that the boundaries of the already-reconstructed patches were points on the visible rim from previous viewpoints. From the discussion at the end of Section 3.1 we know that the points on this boundary are either the endpoints of a visible rim curve, or are ordinary visible rim points from some previous position of the observer. Hence, it suffices for the observer to move back to the position where a given boundary point belongs to the visible rim. This can be achieved by recording, along with each occluding contour image, the viewpoint corresponding to that image during the execution of the Ordinary Patch Reconstruction Strategy. Since there is a correspondence between the points on the reconstructed patch boundaries and the images they project to, this information is sufficient to guide the observer to the position where a particular boundary point was on the visible rim. Furthermore, since any patch boundary point can be forced to become a visible rim point, and the strategies developed in Section 3 can be used to reconstruct a surface patch around any visible rim point, the second question is easily answered by the strategies already developed in this paper. These considerations lead to the following strategy (Figure 7):

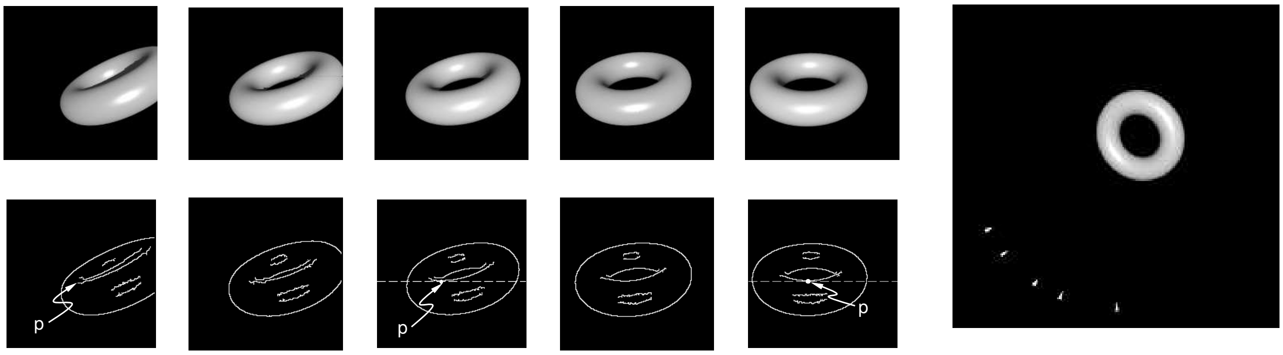


Figure 6: Forcing a cusp point on the visible rim of a torus to become an ordinary point. The upper-left view corresponds to the initial position of the observer. The lower-left view was derived by applying an edge detector to the image above it. The tangent to the selected cusp point p is horizontal. The observer moves on the tangent plane of the selected point, which in this case is a horizontal plane perpendicular to the plane of the page. The sequence of images shows views of the surface from consecutive positions along the observer's path. The rightmost image shows a top view of the tangent plane at p (i.e., "looking down" on the torus), and the five viewpoints where the observer is moving, from left to right. Point p remains on the rim throughout the motion and can be tracked by tracking the occluding contour point whose tangent is horizontal (middle view). The sequence clearly shows that p becomes an ordinary visible rim point after the viewpoint adjustments.

Incremental Reconstruction Strategy

Step 1: If there exists a portion of the surface that has not been reconstructed, select a point p on its boundary and let c be the viewpoint for which p projected to the occluding contour.

Step 2: Move to c.

Step 3: Use the patch reconstruction strategy appropriate for performing the local reconstruction task around p, and continue with Step 1.

5 Building a global surface model

In the previous sections we described a set of qualitative viewing strategies enabling the generation of a sequence of images that permits incremental reconstruction of the viewed surface. To demonstrate the applicability of our viewing strategies we briefly describe in this section the *quantitative* aspects of our approach, namely the construction of a global surface model from the generated sequence of images.

The strategies described in the previous sections can be combined directly with the reconstruction approach of Cipolla and Blake[6] which employs a mobile monocular observer to reconstruct the surface from the deformation of the occluding contour. However, our viewing strategies are not restricted to this specific method for recovering shape from the occluding contour. In particular, we extend the work of Vaillant and Faugeras[15] and employ a stationary calibrated trinocular camera rig. The rig allows the surface to be viewed from three viewpoints that are close to each other. The images from the stereo rig together with

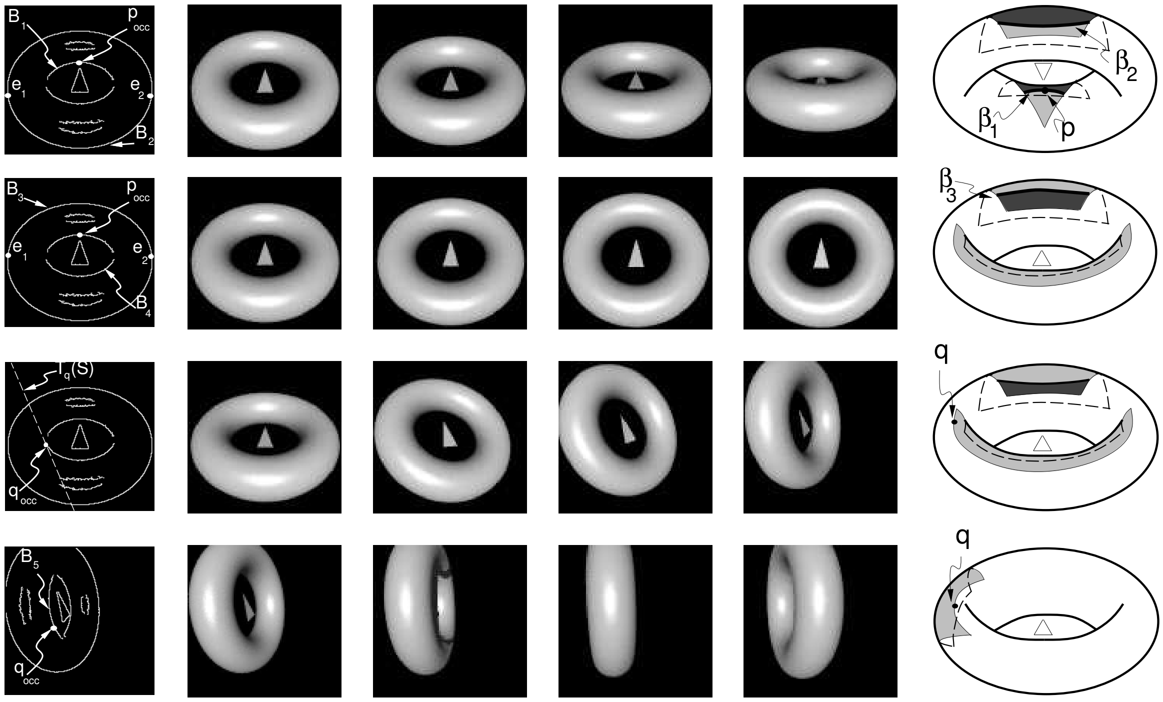


Figure 7: Incremental reconstruction of the surface of a torus. The point selected initially is point p, which is an ordinary visible rim point. The small triangle in the middle of the torus points toward the direction of the line connecting the initial viewpoint of the observer and the center of the torus. *Rows 1 & 2:* These rows of images show consecutive views of the torus as the observer executes the Ordinary Patch Reconstruction Strategy. In the first row the observer moves vertically downward starting from the initial position, while in the second row the observer moves vertically upward after moving back to the initial position. There are four visible rim segments sliding on the torus as the observer moves. Two of the segments, whose projections are B_1 and B_2, slide over initially-occluded parts of the surface during the observer's downward motion. The other two segments, whose projections are B_3, B_4, slide over initially-visible parts of the surface. Points e_1, e_2 are the endpoints of both B_2 and B_3. The figures to the right of these views show the patches traced by these segments. Lightly-shaded patches correspond to the patches reconstructed during the observer's downward motion. Darkly-shaded patches correspond to the patches reconstructed during the observer's upward motion. *Row 3:* The observer selects a point q on the boundary of the reconstructed strip. The point is a cusp point corresponding to the viewpoint, c', of the surface in the first row, second image from the left. The observer executes Steps 1-4 of the Cusp Patch Reconstruction Strategy to force q to become an ordinary point. *Row 4:* The observer executes the Ordinary Patch Reconstruction Strategy to reconstruct a patch for the visible portion of the neighborhood of q, as seen from c'. The figure on the right shows the patch traced by the visible rim segment projecting to B_5.

the calibration information (the relative geometry of the cameras to each other and to the scene) are sufficient to discriminate between occluding contour edges and surface discontinuities, and enable the recovery of the position, surface normal, and surface curvature for a dense set of points that lie on the visible rim corresponding to the three camera viewpoints.[15]

Purposive viewpoint control is achieved by appropriately rotating the viewed surface. Such a rotation corresponds to an instantaneous motion vector $\mathbf{v}$. A correspondence between the geometries of the patches reconstructed from the occluding contour using (1) a monocular observer in motion and (2) a stationary trinocular rig and an object in motion can be achieved by ensuring that the epipolar parameterizations in both formulations are roughly equivalent. In particular, we ensure that the epipolar planes defined by the trinocular rig are approximated by the epipolar plane defined by the instantaneous motion vector: If $\mathbf{v}$ is the motion vector corresponding to the desired viewpoint adjustment, we can find a rotation transformation R such that if R is applied to the viewed surface before the viewpoint adjustment, the desired adjustment will now be along a vector $\mathbf{v}'$ that defines an epipolar plane approximating those of the trinocular rig.

Details of our reconstruction approach and its robustness are described elsewhere.[14] Here we focus on the two primary problems related to the global model construction process, namely (1) the problem of expressing each of the reconstructed points in a common coordinate system and (2) the problem of fitting a surface model to these points, normals, and curvatures that have been reconstructed. We briefly describe our approach to these two problems below.

5.1 Fusing multiple reconstructed patches

Since viewpoint is controlled using the qualitative strategies developed in Sections 3 and 4, we assume here that the motion of the viewed surface is known[†] and that the trinocular camera rig is calibrated. Using the known motion assumption we can express the coordinates of each reconstructed patch in terms of the initially reconstructed surface patch. More specifically, the points in patch Π_i are expressed in the coordinates of patch Π_{i+1} by applying the motion transformation $\Pi_{i+1} = \mathbf{M}_i \Pi_i$. A set of transformations $\mathbf{M}$ for all patches Π_i is what is needed. Since these transformations are known, all reconstructed patches can be expressed in a common coordinate system.

In practice, we use a calibrated horizontal rotating table for making a scan of an object, by rotating the table about a vertical axis. This process corresponds to the process of reconstructing a patch on the surface of the viewed object, as described in Section 3.1. The motion of the table can be described in the coordinates of the table as a simple rotation about the table's vertical axis. The table is calibrated so that the exact rotation (in degrees) is controlled with an accuracy to the one-hundredth of a degree.

5.2 Surface fitting

In our experiments we have generated complete polyhedral surface meshes from the reconstructed data points by taking slices along planes perpendicular to the table's axis of rotation. These slices are connected together in order to form a complete surface mesh across the set of data points. Figures 8(a)-(c) show a reconstructed sphere that has been sliced and meshed in this way. We have made no effort to smooth data at each slice taken in this mesh. Rather, we connect the data in each planar slice, and then connect the slices by triangulation. Figure 8(a) shows all of the reconstructed points after they have been placed in a common coordinate system. Figure 8(b) shows the slices formed by intersecting planes with this data set. Figure 8(c) is the final surface mesh calculated from (b) by connecting slices together and triangulating.

There are many possible approaches to the problem of fitting a surface description to reconstructed data points. Our approach using a reconstructed polyhedral mesh is useful since, at each vertex in the mesh, we have a surface normal that has been either computed from the reconstruction process directly or approximated from the adjacent facets of the mesh. Thus we can manipulate this global mesh with standard graphics tools such as a raytracer or polygon shader.

A shaded view gives a good visual impression of the quality of the reconstructed surface when compared to the original image. Further, smooth surface patches can easily be fit to the underlying polyhedral mesh. For example, Figure 8(d) shows an original image from the input sequence and Figure 8(e) shows the reconstructed surface that has been rendered from the same 3D viewpoint with the same surface characteristics.

[†] However, Seales and Faugeras[14] describe an alternative approach that discards the known motion assumption and recovers the motion of the surface from the motion of fixed edges on that surface.

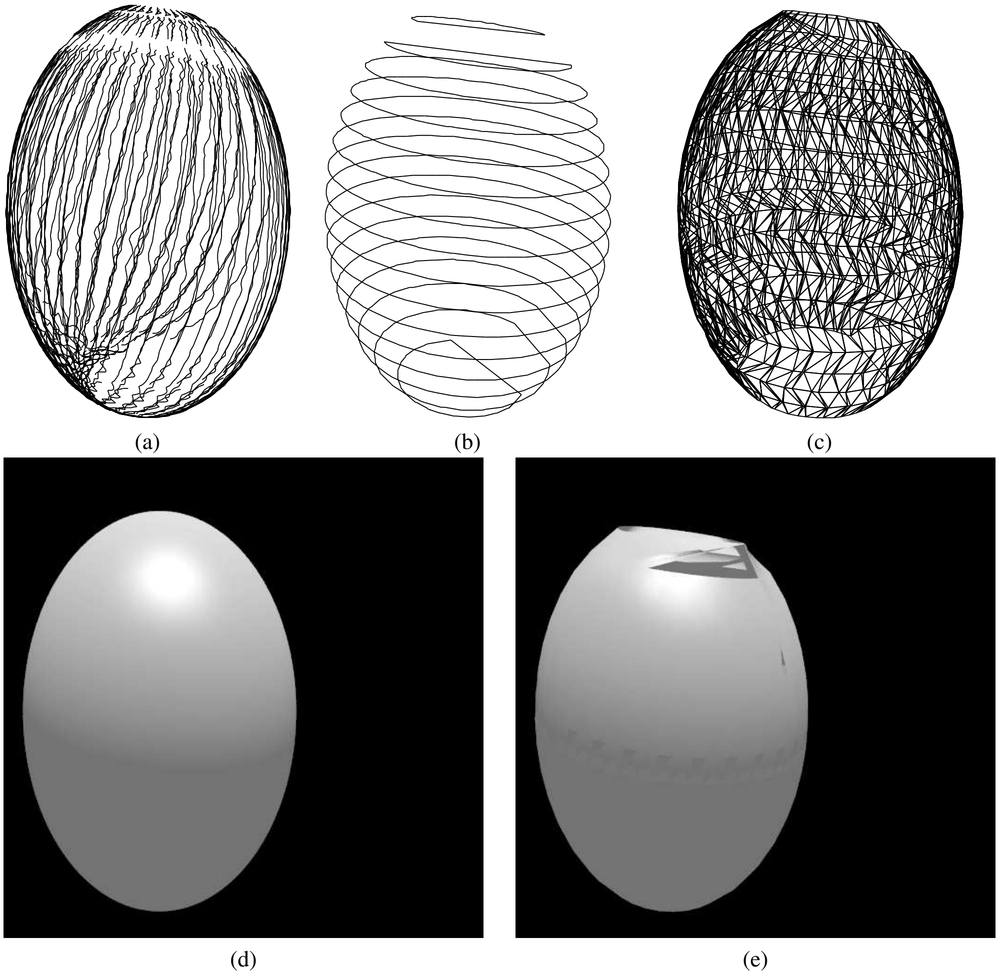


Figure 8: (a) Reconstructed points from an image sequence of a rotating sphere, put in the same coordinate frame. The curves formed by the points correspond to the s-parameter curves of the epipolar parameterization. (b) Slices of data points computed from the reconstructed sphere data. (c) The mesh of triangles computed from the slices in (b). (d) One view of the sequence of the sphere used as input. (e) A rendered view of the reconstructed surface mesh, from the same viewpoint and with surface properties identical to those of the sphere in (d).

6 Experimental results

The experimental results were acquired by positioning a trinocular camera rig, mounted on the INRIA[†] mobile robot, in front of a rotating, calibrated table. First a calibration grid was placed on the table in front of the camera and several images were taken in order to calibrate the camera rig and compute the table's axis of rotation in the coordinate frame of the camera rig. Then, an object was placed on the table and a set of 15 image triplets was taken, each triplet separated by a 12 degree rotation of the

[†] Institut National de Recherche en Informatique et Automatique.

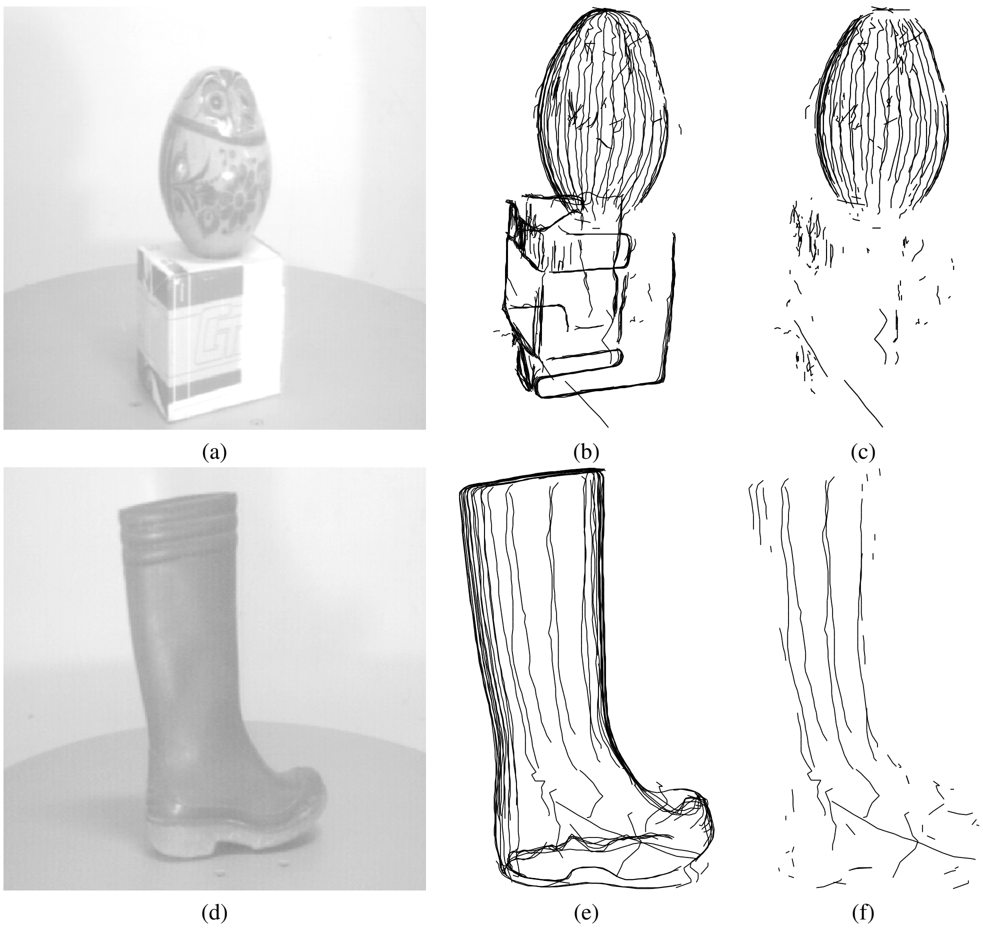
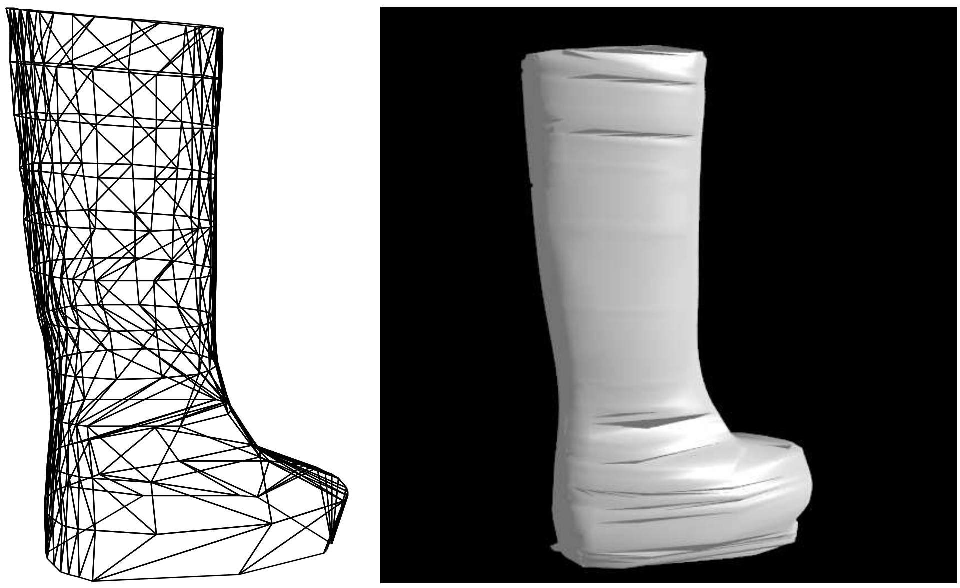


Figure 9: Results of the reconstruction process. The upper row shows results from a sequence of views of a ceramic owl placed on a box. The lower row shows results for a sequence of views of a rubber boot. (a),(d) An image from the sequence. (b),(e) Recovered 3D points for the entire sequence. (c),(f) Edges labeled as occluding contour edges.

table. This complete 180 degree rotation usually allows a cylindrical patch on the object surface to be reconstructed. The rotation angle was selected arbitrarily and we expect that a finer sampling will give a better reconstruction. In our simulations we have found that, for the synthetic sphere example, the mean error in the position of reconstructed points is less than 1 millimeter, and the mean error in the direction of the reconstructed surface normal is about 3 degrees. The interested reader is referred to recent work by Seales and Faugeras.[14]

The reconstruction results for two sequences of 15 views are shown in Figure 9. The ceramic owl was placed on a box in order to produce edge images containing both fixed and occluding contour edges. On the other hand, the dark rubber child's boot was chosen since it does not suffer from the problems of specularities and surface markings.

Our results for the owl sequence show that the occluding contour edges can be reliably distinguished from the fixed edges using our approach. The 3D points projecting to the edges in the owl and the boot sequences are also reconstructed very



(a) (b)

Figure 10: (a) The computed surface mesh from the boot sequence. (b) A rendered view of the boot surface mesh.

accurately. Although Figure 9(f) shows that inaccuracies exist in the classification of occluding contour and fixed edges in the boot sequence, this is because the ratio of the baseline of the cameras to the distance of the cameras from the boot was small; that is, the camera was too far away to detect the occluding contour where the boot's curvature was higher.†

Figure 10 shows rendered views of the complete surface mesh that was recovered from the sequence of the boot. Because we have recovered surface normals at most of the points (those that are on the occluding contour), we can render the mesh using interpolated shading techniques such as Phong shading. The mesh shown in Figure 10 is completely closed and can be rotated and viewed from any direction. By comparing the rendered model with the original images from the sequence (e.g., Figure 9(b)) one can see that the overall accuracy of the reconstruction is very good.

The results of our reconstruction experiments emphasize the importance of controlling viewpoint based on the occluding contour. In particular, our results reinforce the practical validity of the developed viewpoint control strategies since the occluding contour can be reliably detected in edge images, and illustrate the need for employing such strategies: By simply rotating the boot about a vertical axis after placing it on the rotating table as shown in Figure 9 we cannot determine whether or not the boot is in fact "filled", i.e., it contains no cavity on top for inserting the foot. A simple application of the Incremental Reconstruction Strategy of Section 4 for extending the reconstruction process would involve rotating the table again after lying the boot on its side on the rotating table.

† The inaccuracies are caused by the fact that we use a thresholding approach to perform the classification, where edges corresponding to surface points with high curvature are considered to be fixed edges. As expected, Figure 9(f) shows that the edges classified as occluding contour edges were the ones for which the boot's curvature is smaller.

7 Concluding remarks

We have demonstrated that an active observer can follow a set of simple and qualitative viewing strategies to incrementally reconstruct the surface of an unknown object. Furthermore, these strategies are based purely on the computation of a simple property of the occluding contour (tangent at a point) and can be combined with a robust surface reconstruction algorithm in order to build a global model of the viewed object. Our experimental results show that the developed strategies are readily implementable, they rely on visual information that can be robustly computed, and because of their simplicity and low computational requirements they are suitable for real-time implementation.

The use of an active observer is the most crucial aspect of our approach. The ability to *purposefully* change viewpoint enables the observer to generate an image sequence that allows incremental reconstruction of the surface from the occluding contour. The reason is that observer motion is not used merely to force the motion of the visible rim over the surface (as in existing approaches), but it is used to *control* the visible rim's motion in a well-defined and qualitative manner, forcing it to slide over previously-unreconstructed parts of the surface. This is a major step toward qualitative, active vision, establishing an important link between the ability to purposively and qualitatively control viewpoint and the global three-dimensional model construction task.

8 Acknowledgements

The support of the National Science Foundation under Grant Nos. IRI-9022608 and IRI-9220782 and of the Robot Vision group at INRIA-Sophia Antipolis, France, under ESPRIT and INSIGHT grants is gratefully acknowledged.

9 REFERENCES

[1] A. Lynn Abbott and Narendra Ahuja. Active surface reconstruction by integrating focus, vergence, stereo, and camera calibration. In *Proc. 3rd Int. Conf. on Computer Vision*, pages 489–492, 1990.

[2] Yiannis Aloimonos. Purposive and qualitative active vision. In *Proc. Int. Conf. on Pattern Recognition*, pages 346–360, 1990.

[3] Dana H. Ballard and Christopher M. Brown. Principles of animate vision. *Computer Vision, Graphics, and Image Processing: Image Understanding*, 56(1):3–21, 1992. Special Issue on Purposive, Qualitative, Active Vision.

[4] Andrew Blake, Rupert Curwen, and Andrew Zisserman. A framework for spatio-temporal control in the tracking of visual contours. *Int. J. Computer Vision*, 1993. First Special Issue on Active Vision. To appear.

[5] Manfredo P. Do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall Inc., Englewood Cliffs, NJ, 1976.

[6] Roberto Cipolla and Andrew Blake. Surface shape from the deformation of apparent contours. *Int. J. Computer Vision*, 9(2):83–112, 1992.

[7] R. Curwen, A. Blake, and A. Zisserman. Real-time visual tracking for surveillance and path planning. In *Proc. 2nd European Conf. on Computer Vision*, pages 879–883, 1992.

[8] Peter Giblin and Richard Weiss. Reconstruction of surfaces from profiles. In *Proc. 1st Int. Conf. on Computer Vision*, pages 136–144, 1987.

[9] Seth Hutchinson. Exploiting visual constraints in robot motion planning. In *Proc. IEEE Robotics Automat. Conf.*, pages 1722–1727, 1991.

[10] Jan J. Koenderink. An internal representation for solid shape based on the topological properties of the apparent contour. In Whitman Richards and Shimon Ullman, editors, *Image Understanding 1985-86*, pages 257–285. Ablex Publishing Co., Norwood, NJ, 1987.

[11] Kiriakos N. Kutulakos and Charles R. Dyer. Recovering shape by purposive viewpoint adjustment. In *Proc. Computer Vision and Pattern Recognition*, pages 16–22, 1992.

[12] Kiriakos N. Kutulakos and Charles R. Dyer. Global surface reconstruction by purposive control of observer motion. Technical Report 1141, Computer Sciences Department, University of Wisconsin— Madison, March 1993. Available via ftp from ftp.cs.wisc.edu.

[13] Jasna Maver and Ruzena Bajcsy. Occlusions as a guide for planning the next view. *IEEE Trans. Pattern Anal. Machine Intell.*, 15(5):417–433, 1993.

[14] W. B. Seales and O. D. Faugeras. Global surface reconstruction from the extremal boundary. Technical report, University of Kentucky, Lexington, Kentucky, 1992.

[15] Régis Vaillant and Olivier D. Faugeras. Using extremal boundaries for 3-d object modeling. *IEEE Trans. Pattern Anal. Machine Intell.*, 14(2):157–173, 1992.

[16] P. Whaite and F. P. Ferrie. From uncertainty to visual exploration. *IEEE Trans. Pattern Anal. Machine Intell.*, 13(10):1038–1049, 1991.