To appear in: Proc. CVPR'94, Seattle, WA

Global Surface Reconstruction by Purposive Control of Observer Motion

Kiriakos N. Kutulakos
kyros@cs.wisc.edu

Charles R. Dyer
dyer@cs.wisc.edu

Computer Sciences Department
University of Wisconsin
Madison, Wisconsin 53706

Abstract

What real-time, qualitative viewpoint-control behaviors are important for performing global *visual exploration tasks such as searching for specific surface markings, building a global model of an arbitrary object, or recognizing an object? In this paper we consider the task of* purposefully *controlling the motion of an active, monocular observer in order to recover a global description of a smooth, arbitrarily-shaped object using the occluding contour. By studying the epipolar parameterization, we develop two basic behaviors that allow reconstruction of a patch around any point in a reconstructible surface region. These behaviors rely only on information extracted directly from images (e.g., tangents to the occluding contour), and are simple enough to be executed in real time. We then show how global surface reconstruction can be provably achieved by (1) integrating these behaviors to iteratively "grow" the reconstructed regions, and (2) obeying four simple rules.*

1 Introduction

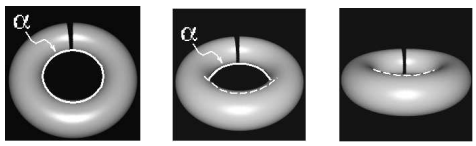
In this paper we use an active (i.e., mobile) observer that *purposefully* controls its viewpoint to derive a global, three-dimensional description of an object. Our approach solves the following problem: How should the viewpoint of the observer be controlled in order to generate a dense sequence of images maximizing the area of the region reconstructed on the object's surface? We call this the *global surface reconstruction task*. We consider this task for smooth surfaces of arbitrary shape; the object is unknown, can be non-convex, and can self-occlude. We use a shape-from-motion module for extracting surface shape information [1]; we assume that the object is stationary and the observer is able to freely move on a sphere around it.

Our goal is to control the observer's viewpoint so that global surface reconstruction is guaranteed. When trying to perform tasks that depend on an object's appearance, provable correctness is critical: The appearance of objects can drastically change depending on the observer's viewpoint, making *ad hoc* viewpoint control mechanisms unpredictable and incomplete. Furthermore, like many others (e.g., [2, 3]), we are interested in mechanisms that tightly couple sensing and action and allow real-time execution. This is because our purpose is to exploit the ability to quickly control viewpoint to achieve simple and efficient solutions for a given task, rather than rely on the availability of large amounts of computational power [3]. To emphasize this point we refer to such mechanisms as *viewpoint control behaviors*.

Very little work has been published on the use of viewpoint control behaviors for reconstruction, exploration or recognition tasks. However, the few recent approaches taking advantage of viewpoint-control behaviors demonstrate their importance and generality. We showed in an earlier paper [4] that the shape recovery problem for smooth surfaces becomes considerably simplified if the observer uses a simple viewpoint-control behavior to move to a special viewpoint, which for the case of surfaces of revolution corresponds to their side view. The work of Wilkes and Tsotsos [5] illustrates how the ability to purposefully and quickly change viewpoint can simplify the task of object recognition in a simple world of Origami objects. Grosso and Ballard [6] are currently designing a head-eye system capable of implementing such viewpoint controls. Recent work by Blake *et al.* [7] showed that shorter paths can be achieved in robotic navigation tasks if the shape of the obstacles is taken into account during navigation.

Apart from the above approaches, viewpoint control for performing various tasks has been treated as a complex and computationally-intensive optimization problem (i.e., "where to look next"), where the best next viewpoint is searched for within the space of all possible viewpoints [8]. In tasks involving global surface reconstruction, the viewpoint-control mechanisms assumed that a three-dimensional representation of the visible surfaces can be recovered independently from each viewpoint [8–10], ruling out their applicability in more qualitative visual exploration tasks. Another disadvantage of such mechanisms is that they do not take into account how the global geometry of the surface (e.g., self-occlusions) affects the correctness of the global reconstruction algorithms. This makes the reconstructed regions unpredictable. Moreover, in the context of animate [3] and purposive [2] vision, their major drawback is their inability to exploit real-time viewpoint control to simplify the reconstruction process (e.g., by utilizing shape-from-motion approaches, which have been shown to considerably simplify shape recovery computations [1]), as well as to simplify the viewpoint control process itself.

The support of the National Science Foundation under Grant No. IRI-9220782 is greatfully acknowledged.



(a) (b) (c)

Figure 1: Forcing the visible rim curve α, shown as a white circle in (a), to slide over the dark curve drawn on the torus.

In this paper we use the occluding contour to control viewpoint and to derive shape information. The occluding contour is the projection of the visible rim, the one-dimensional set of visible surface points at which the line of sight is tangent. The occluding contour is a rich source of shape information [11, 12]. Moreover, contrary to approaches that use sophisticated sensing mechanisms to reconstruct the scene from a single viewpoint or a small number of viewpoints [9, 10], recent results demonstrate that the occluding contour can be reliably detected in edge images [13], and that shape (e.g., curvature) can be efficiently and accurately recovered from the occluding contour [1, 7].

An important property of the visible rim is that it depends on the viewpoint of the observer and the shape of the surface: Under continuous observer motion the visible rim "slides" over the surface and may change its connectivity, affecting the geometry and topology of the occluding contour and revealing shape information for the parts of the surface over which the visible rim slides. We exploit this property by formulating global surface reconstruction as the qualitative task of controlling the observer's viewpoint so that the visible rim slides over the maximal connected and reconstructible surface regions intersecting the visible rim at the initial viewpoint. This formulation is particularly important because it separates the issue of controlling the observer's viewpoint from the issue of reconstructing the surface, i.e., processing the images produced during the observer's motion. Consequently, the mechanisms developed in this paper can be used for other visual exploration tasks (e.g., inspection) where the central problem is how to control viewpoint to make new points on an object's surface visible without necessarily reconstructing its surface.

Our approach is developed in the context of three increasingly more general reconstruction tasks: *The local surface reconstruction task*, where the observer must control viewpoint to reconstruct a patch around a selected visible rim point; *the incremental surface reconstruction task*, where the observer must control viewpoint to iteratively "grow" the reconstructed patches on the surface; and the *global surface reconstruction task*, where the entire reconstructible regions intersecting the visible rim at the initial viewpoint are reconstructed. In [14] we showed how the observer can control viewpoint to perform the local and incremental surface reconstruction tasks under the assumption that topological changes on the visible rim do not occur in the vicinity of the reconstructed patch. In this paper we consider the global surface reconstruction task, showing how to control viewpoint to ensure that the reconstructed patch covers an entire reconstructible region on the surface. This involves paying specific attention to the global shape of the surface and to the topological changes of the visible rim that may occur during the observer's motion.

Intuitively, the main difficulty in solving the global surface reconstruction task is that although the observer has some control over the motion of the visible rim over the surface, this control is not complete; the motion of the visible rim also depends on the shape of the surface itself. In addition, the visible rim's topology can change unexpectedly, further complicating the reconstruction process. To illustrate the difficulties involved in globally reconstructing an object, consider the torus shown in Figure 1(a). In order to reconstruct the torus, the observer must force the visible rim to slide over all points along the dark curve drawn on its surface. One way to proceed is to move downward, forcing the smooth curve α, which is part of the visible rim, to slide over the segment of the dark curve that is initially occluded. As the observer moves downward, however, α shrinks to a point and disappears (Figure 1(c)), changing the topology of the visible rim and making any further downward motion ineffective. The observer must now move differently to continue the reconstruction process. Similar difficulties occur due to geometrical changes in the visible rim (i.e., even when no topological changes of the visible rim occur).

This simple example illustrates that a number of basic steps are necessary to achieve global surface reconstruction. It is therefore necessary to ask how many times such actions need to be executed, whether the whole surface is always reconstructed, and whether the reconstruction process is guaranteed to terminate. These issues are precisely the reasons why provably-correct viewpoint control behaviors are necessary: Since the answers to these questions are not evident even for geometrically-simple surfaces such as the torus, such behaviors are necessary if one hopes to use them for reconstructing the surface of real objects.

2 Local Shape from Occluding Contour

Let S be a smooth, generic surface bounding a finite volume in $\Re^3$. We assume S is viewed under spherical projection and that the observer moves on a *motion sphere* surrounding the surface. The *visible rim* of S at a viewpoint c is the set of visible surface points p for which the tangent plane, $T_p(S)$, contains the line segment connecting p and c. The *occluding contour* is the projection of the visible rim on the image. For almost all positions of the observer, the visible rim is a collection of closed and open smooth curves whose endpoints project to cusps or T-junctions on the occluding contour [12].

The shape and topology of the visible rim and the occluding contour depends on S and the observer's viewpoint. A suitable surface parameterization relating the shapes of S, the visible rim, and the occluding contour is the *epipolar parameterization* [1, 11]. Intuitively, the epipolar param-

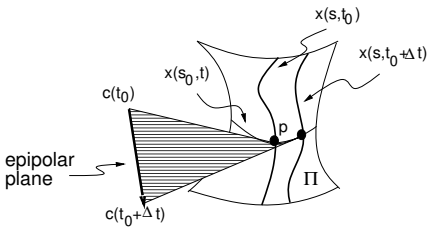


Figure 2: The epipolar parameterization. Curves $\mathbf{x}(\mathbf{s}, \mathbf{t_0})$ and $\mathbf{x}(\mathbf{s}, \mathbf{t_0} + \Delta\mathbf{t})$ are curves on the visible rim of the surface corresponding to observer positions $c(t_0)$ and $c(t_0 + \Delta t)$, respectively. The tangent to the curve $\mathbf{x}(\mathbf{s_0}, \mathbf{t})$ 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, $\mathbf{v}(\mathbf{t})$, and the line $c(t_0)p$.

eterization captures the idea that under continuous motion of the observer (and when the topology of the occluding contour does not change), the set of points comprising the visible rim consists of smooth curves that "slide" over the surface. This allows the non-concave parts of the surface to be considered as a collection of patches, each of which is a family of visible rim curves (Figure 2). The epipolar parameterization was used in [1, 11] to recover the fundamental forms of all points in Π from the deformation of the occluding contour during the observer's motion.

The crucial point in the definition of the epipolar parameterization is that the epipolar parameterization imposes four strong constraints on the observer's ability to recover a patch Π around a visible rim point p:

Epipolar Reconstructibility Constraints

C0: p must be visible from some viewpoints on its tangent plane.

C1: p must not be the endpoint of a visible rim curve. This is because it was assumed that p is not on Π's boundary.

C2: If $\mathbf{v}(\mathbf{t})$ is the observer's velocity, $T_p(S)$ must not contain $\mathbf{v}(\mathbf{t})$. This is because in that case point p remains on the visible rim.

C3: The topology of the visible rim curve containing p must not change in the neighborhood of p under an infinitesimal motion of the observer. Only a finite collection of curves on the surface cannot satisfy this constraint. These curves bound the surface points not satisfying constraint C0 (Section 4).

The Epipolar Reconstructibility Constraints show that the epipolar parameterization cannot be used to describe the surface in the neighborhood of every visible rim point. They also show that the surface patch Π depends on how the visible rim curve $\mathbf{x}(\mathbf{s}, \mathbf{t_0} + \Delta\mathbf{t})$ slides over the surface when Δt varies continuously. Consequently, the dynamics of the visible rim curves determine the patches reconstructed. These dynamics depend on the local and global shape of the surface as well as the observer's motion.

The Epipolar Reconstructibility Constraints characterize the reconstructible regions on the surface, i.e., they tell us what is the most we can expect from any viewpoint-control behavior that uses the occluding contour for reconstruction:

Reconstructible surface regions: The reconstructible regions are the maximal connected sets of points for which all Epipolar Reconstructibility Constraints can be simultaneously satisfied.

We will see that the reconstructible regions are bounded by points that can satisfy constraint C0 but not constraint C3. Constraint C3 applies only to surface points belonging to visual event curves, briefly discussed below.

2.1 Visual Event Curves

The topology of the occluding contour of a smooth surface is stable for almost all viewpoints. Results from singularity theory show that the space of viewpoints can be partitioned into a collection of maximal connected cells within which the occluding contour's topology remains constant [12, 15]. *Visual events* occur when the observer's viewpoint belongs to the boundaries of these cells. An infinitesimal perturbation of such a viewpoint results in changes in the occluding contour's topology. A catalog of the visual events can be found in [15]. For generic surfaces, the visual events are associated with a collection of surface curves, called the *visual event curves*.[1] The visual event curves relevant to our analysis are (1) the parabolic surface curves, (2) the curves $\tau(s)$, associated with triple-point events, such that the line segment $\tau(s)c$ touches the surface at three distinct points for some viewpoint c, (3) the curves $\gamma(s)$, associated with tangent-crossing events, such that the line segment $\gamma(s)c$ touches the surface at two distinct points with identical tangent planes for some viewpoint c, and (4) the curves $\sigma(s)$, associated with cusp-crossing events, such that the line segment $\sigma(s)c$ touches the surface at two distinct points and is an asymptote at $\sigma(s)$ for some viewpoint c. A subset of these curves bounds the reconstructible regions. We characterize this subset precisely in Section 4.

3 Local Surface Reconstruction

In this section we consider the local surface reconstruction task: Suppose the observer is at position c, and let p be a visible rim point on the object's surface that is identified by its projection, p_{occ}, on the occluding contour. The observer's task is to continuously control viewpoint, starting from point c, in order to recover the local shape of the surface for all points in *some* neighborhood Π of p. We use the following two observations:

- If p is the endpoint of a visible rim curve, the epipolar parameterization cannot describe the surface in the neighborhood of p. However, there are other viewpoints on p's tangent plane at which p is not the endpoint of a visible rim curve, i.e., at which Epipolar Reconstructibility Constraint C1 is satisfied.

- The point p and the observer's viewpoint may be such that the occluding contour's topology changes in the

[1] See [15] for details on the definition of these curves.

neighborhood of p under an infinitesimal viewpoint perturbation. For all points p except those lying on a subset of the visual event curves, the observer can satisfy Epipolar Reconstructibility Constraint C3 by moving to other viewpoints on p's tangent plane at which the contour's topology does not change in the neighborhood of p if these viewpoints are infinitesimally perturbed.

Based on these observations, for any given viewpoint we distinguish four types of visible rim points: *Ordinary points*, which satisfy constraint C3 and are not endpoints of a visible rim curve; *cusp points* and *T-junction* points, which satisfy constraint C3 and are visible rim endpoints projecting to a cusp and a T-junction on the occluding contour, respectively; and *degenerate points*, which are visible rim points not satisfying constraint C3. These four types of visible rim points are exhaustive and give rise to four instances of the local surface reconstruction task.

To perform the local surface reconstruction task we use a basic behavior that controls the observer's motion to deal with the case where p is ordinary. The other three cases are treated by (1) reaching a viewpoint where p is ordinary, and (2) using the basic behavior in order to recover the shape of the surface in a neighborhood of that point. Below we outline the behavior handling the case of ordinary points for completeness (see [14] for details), and focus on the case of degenerate points which is central for obtaining the global results of Sections 4 and 6. The interested reader is refered to [14, 16] for a detailed treatment of the remaining cases.

3.1 Reconstruction Around Ordinary Points

Ordinary Patch Reconstruction Behavior

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 an occluding contour curve.

Step 2: Compute the surface normal at p. The normal is given by $\mathbf{T} \wedge \mathbf{p_{occ}}$, where $\mathbf{T}$ is the tangent to the occluding contour at p_{occ} [1].

Step 3: (Reconstructing the occluded points near p.) Select a direction $\mathbf{v_1}$ for moving on the motion sphere that satisfies the inequality $\mathbf{n(p)} \cdot \mathbf{v_1} > \mathbf{0}$. Change viewpoints along $\mathbf{v_1}$ while continuously 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 another direction of motion $\mathbf{v_2}$ that satisfies the inequality $\mathbf{n(p)} \cdot \mathbf{v_2} < \mathbf{0}$.

Two components of the above behavior are purposely left unspecified: The choice of directions $\mathbf{v_1}$ and $\mathbf{v_2}$ in Steps 3 and 4, respectively, and the condition for terminating the observer's motion in these two steps. In Section 6, where we consider the global surface reconstruction task, we show that when executing the Ordinary Patch Reconstruction Behavior the observer must obey a number of rules that "ground" these steps.

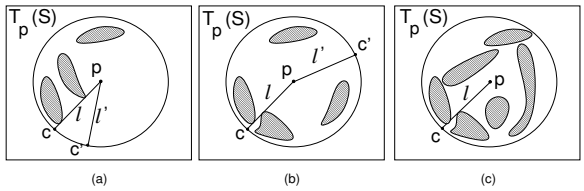


Figure 3: Surface geometry around a degenerate point p. A top view of the tangent plane of p is shown. Shaded regions correspond to the intersections of $T_p(S)$ with the object. In this example, p belongs to a visual event curve associated with a triple-point event: the line through p and the observer's viewpoint, c, touches the surface at three points. (a) A small viewpoint change on $T_p(S)$ makes p ordinary. (b) The geometry of the intersection $T_p(S) \cap S$ forces p to become occluded when small viewpoint changes are performed. However, there are viewpoints on $T_p(S)$ at which p is ordinary. (c) The geometry of $T_p(S) \cap S$ forces p to be occluded at all viewpoints except c.

3.2 Reconstruction Around Degenerate Points

Suppose that the observer is moving along a curve $c(t)$, and that the topology of the visible rim changes in the vicinity of a degenerate visible rim point p at viewpoint $c = c(t_0)$. This can happen only if the line connecting c and p has a high order contact with the surface or if it contacts the surface at multiple points [12], i.e., if and only if p belongs to a visual event curve. If the observer makes an infinitesimal viewpoint change to a new viewpoint, c', on $T_p(S)$, the line connecting c' and p will either have lower order contact with the surface or will touch the surface at fewer points. Hence, if p is visible from the new viewpoint, p will become an ordinary visible rim point (Figure 3(a)). Unfortunately, p may no longer be visible (Figure 3(b)). In this case, in order to make p ordinary the observer must move to distant viewpoints on $T_p(S)$ from which p is visible. We therefore need to specify how the observer should move and when to stop.

The first question can be answered by moving either clockwise or counterclockwise on a circle in $T_p(S)$ around p. The direction of motion on this circle is not important. The observer should stop when p becomes ordinary. It is easy to see that if there is an open arc of viewpoints on the observer's motion circle from which p is visible, any viewpoint on that arc guarantees that p is an ordinary visible rim point at that viewpoint. To completely specify the observer's motion it remains to give a way of detecting when p becomes visible again. A quantitative approach to this problem is to first determine the three-dimensional coordinates of p, and then continuously check, during the observer's motion, if any visible rim point with tangent plane coincident to $T_p(S)$ matches those coordinates. To avoid dependence on such coordinate information we simply observe that the occlud-

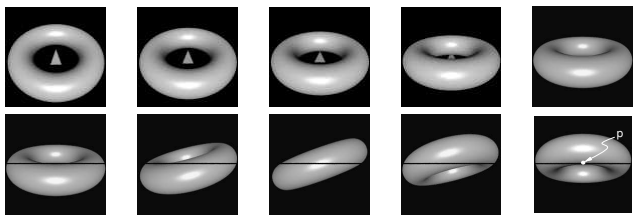


Figure 4: Forcing a degenerate point on the torus to become ordinary. *Top row:* The observer moves downward until the visible rim segment pointed by the triangle shrinks to a degenerate point p. The tangent plane at p is horizontal. The visual event corresponding to the disappearance of that segment is a tangent-crossing event. *Bottom row:* Moving on $T_p(S)$ in order to make p ordinary. The black horizontal line is the projection of $T_p(S)$ in the image. The observer performs a clockwise viewpoint change on $T_p(S)$ until an ordinary visible rim point with tangent plane identical to $T_p(S)$ is detected.

ing contour must be tangent to $T_p(S)$ at p's projection. This leads to the following qualitative viewpoint control behavior for reconstructing the surface around p (Figure 4):

Degenerate Patch Reconstruction Behavior

Step 1: Let $p(t_0-\delta t)$ be the visible rim point at position $c(t_0-\delta t)$ that is matched to p by the epipolar parameterization. Compute the tangent plane at p as the limit $\lim_{\delta t \to 0} T_{p(t_0-\delta t)}(S)$.

Step 2: Perform a small counterclockwise motion on $T_p(S)$. If p remains visible, set $q = p$ and continue with Step 5. Otherwise, return to $c(t_0)$.

Step 3: Perform a small clockwise motion on $T_p(S)$. If p remains visible, set $q = p$ and continue with Step 5. Otherwise, return to $c(t_0)$.

Step 4: Move clockwise on $T_p(S)$ while continuously monitoring the occluding contour, until either $c(t_0)$ is reached or there is an ordinary visible rim point q whose tangent plane coincides with $T_p(S)$.

Step 5: If $c(t_0)$ is reached, stop. Otherwise, apply the Ordinary Patch Reconstruction Behavior to reconstruct a patch around q, and continue with Step 4.

Using the above behavior, local reconstruction is not achieved for p if and only if p is occluded from all but a finite set of viewpoints on its tangent plane. Such points never become ordinary during the observer's motion on $T_p(S)$ (Figure 3(c)). This is not a limitation of the Degenerate Patch Reconstruction Behavior; Epipolar Reconstructibility Constraint C3 cannot be satisfied, and there are simply no observer motions that force the visible rim to slide over a neighborhood of such a point to allow reconstruction.

4 The Reconstructible Surface Regions

The analysis of the local surface reconstruction task gives us a way to characterize the reconstructible regions on the

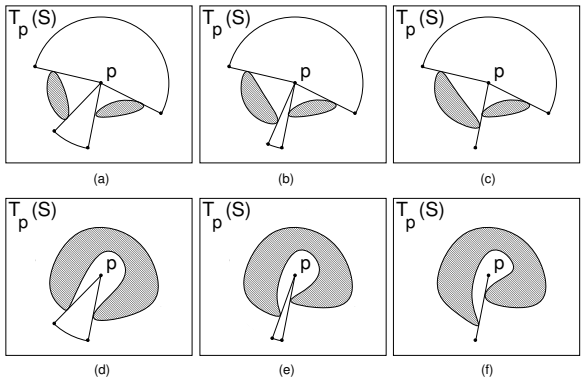


Figure 5: The visibility arcs of a point p. A top view of the tangent plane of p is shown. Shaded areas correspond to the intersections of $T_p(S)$ with the object. (a)-(c) Approaching a visual event curve τ associated with a triple-point event. Point p has two visibility arcs. As p approaches τ, one of the visibility arcs of p degenerates to a point. In this case, the point p in (c) belongs to τ, but is not contained in the boundary of a reconstructible surface region; the neighborhood around p can be reconstructed by moving to a viewpoint in the remaining visibility arc of p. (d)-(f) Approaching a visual event curve τ associated with a triple-point event. Point p now has one visibility arc. As p approaches τ, the visibility arc of p degenerates to a point. In this case, p asymptotically approaches the boundary of a reconstructible surface region.

surface by characterizing their boundaries. In particular, the patch reconstruction behaviors in Section 3 allow us to reconstruct a surface patch around all surface points except for (1) points that are never visible from viewpoints on their tangent plane, and (2) points on visual event curves that are visible only from a finite number of viewpoints on their tangent plane. For these points, there is no path the observer can follow that forces the visible rim to slide over their neighborhood. This leads directly to the following characterization of the reconstructible regions on the surface [16]:

> **Reconstructible surface regions:** The reconstructible surface regions are the maximal connected sets of points that are visible from a one-dimensional set of viewpoints on their tangent plane.

Theorem 1 (Reconstructible region boundaries) *A point p is on the boundary of a reconstructible surface region if and only if (1) it belongs to either a parabolic curve bounding a surface concavity, a curve $\tau(t)$ associated with a triple-point event, a curve $\gamma(t)$ associated with a tangent-crossing event, or a curve $\sigma(t)$ associated with a cusp-crossing event, and (2) is visible from only a finite number of viewpoints on its tangent plane.*

An intuitive description of Theorem 1 can be given as follows. To each point p in a reconstructible surface region we can associate a collection of *visibility arcs*. These arcs are simply the open and connected one-dimensional sets of viewpoints on p's tangent plane from which p is visible (Figure 5). When p asymptotically approaches one of the

above visual event curves, the length of at least one of its visibility arcs decreases, diminishing to zero (Figure 5(a)-(c)); this can only happen for the visual event curves listed above. Now, if the length of *all* visibility arcs of p diminishes to zero, the visual event curve approached by p belongs to the boundary of the reconstructible regions on the surface (Figure 5(d)-(f)). The visual event curves listed in Theorem 1 are therefore *potential boundaries* of a reconstructible surface region. They bound such regions only if they contain points with no visibility arcs.

5 Incremental Surface Reconstruction

The goal of the global surface reconstruction task is to reconstruct the reconstructible surface regions that intersect the visible rim at the initial viewpoint. To achieve this, we incrementally "grow" the patches initially reconstructed on the surface. The following composite behavior is based on the observation that the boundaries of the already-reconstructed patches were endpoints of the visible rim at previous viewpoints. See [14] for more details.

Incremental Reconstruction Behavior

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 at which p projected to the occluding contour.

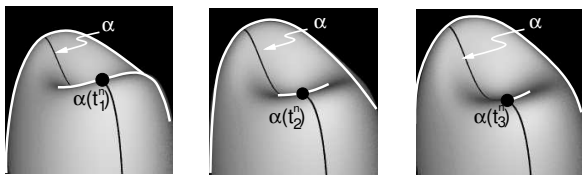
Step 2: Move to c, use the viewpoint control behavior appropriate for performing the local surface reconstruction task around p, and continue with Step 1.

The specific algorithm for selecting the points p on the boundary of the already-reconstructed surface regions is not important for guaranteeing their successive expansion. However, in order to perform the global surface reconstruction task the observer must obey an additional rule when doing this selection; we discuss this rule in Section 6.

6 Global Surface Reconstruction

What kinds of behaviors can perform global surface reconstruction? In the Introduction we motivated the need for behaviors for which (1) the reconstruction process does not terminate (i.e., at least one of the component behaviors is executed an infinite number of times) if and only if there is no finite-length path the observer can follow that allows global surface reconstruction (*finite termination*), and (2) if the reconstruction process terminates, the reconstructed points must be the union of the reconstructible regions intersecting the visible rim at the initial viewpoint, and if it does not terminate the reconstructed points must asymptotically approach that set (*completeness*).

In this section we show that global surface reconstruction can be achieved by (1) using the Incremental Reconstruction Behavior, while (2) obeying a number of simple rules that constrain how that behavior is executed. Motivated by our characterization of the reconstructible regions and the examples of Figures 1 and 6, we develop these rules by considering the following three increasingly more general global reconstruction tasks:



(a) (b) (c)

Figure 6: Difficulties involved in globally reconstructing a dimple-shaped surface. In the n-th iteration of the Incremental Reconstruction Behavior the observer is moving upward in order to reconstruct points in the neighborhood of point $\alpha(t_1^n)$ on α, which lies on the visible rim. The visible rim eventually slides to the right, making the observer's upward motion ineffective for reconstructing the surface in the vicinity of $\alpha(t_3^n)$.

- **Semi-global curve reconstruction task:** Suppose a curve is drawn on the surface so that it intersects the visible rim at the initial viewpoint (Figure 6(a)). The task of the observer is to reconstruct the segments of this curve that are connected, reconstructible, intersect the visible rim at the initial viewpoint, and terminate on a visual event curve that *potentially* bounds a reconstructible surface region, as in Figures 5(c),(f).
- **Global curve reconstruction task:** Suppose a curve is drawn on the surface so that it intersects the visible rim at the initial viewpoint. The task of the observer is to reconstruct the segments of this curve that are connected, reconstructible, intersect the visible rim at the initial viewpoint, and terminate on the *boundary* of a reconstructible region.
- **Global surface reconstruction task:** Global surface reconstruction is a generalization of the global curve reconstruction task in the following sense. It is equivalent to reconstructing for every surface curve that intersects the visible rim at the initial viewpoint, a connected, reconstructible segment terminating on the boundary of a reconstructible region.

By obeying the rules we develop in this section, the observer can "ground" the steps in the Incremental Reconstruction Behavior and the Patch Reconstruction Behaviors that we left unspecified in their earlier presentation. In the following, due to space limitations, we only consider the semi-global curve reconstruction task and the global surface reconstruction task. We keep our analysis at a fairly intuitive level, working through specific examples to motivate the rules used. Further insight into the theorems, an analysis of the global curve reconstruction task, and formal proofs of correctness can be found in [16].

6.1 Semi-Global Curve Reconstruction Task

Recall that during the execution of the Incremental Reconstruction Behavior the observer selects a point p on the

boundary of the already-reconstructed regions and appropriately controls motion to reconstruct a new patch around p. To achieve semi-global curve reconstruction, the length of the curve segment reconstructed at each iteration must diminish if and only if it asymptotically approaches a visual event curve *potentially* bounding a reconstructible region. The following theorem gives the three rules the observer must obey for the semi-global curve reconstruction task.

Theorem 2 (Semi-global curve reconstruction rules) *Let α be a finite-length curve drawn on the surface. If α intersects the visual event curves at most a finite number of times, and the following three rules are obeyed by the observer, the observer will provably perform the semi-global curve reconstruction task:*

Rule 1: *When choosing the point p on which to apply the Patch Reconstruction Behaviors, always select a point of intersection of α with the visible rim.*

Rule 2: *Always execute the Ordinary Patch Reconstruction Behavior after first moving to a viewpoint c corresponding to the middle of a visibility arc of p.*

Rule 3: *When executing the Ordinary Patch Reconstruction Behavior to reconstruct a patch around p starting from a viewpoint c, move around the surface on the normal plane at p and stop only if c is reached again, or if the endpoint of the segment of α being reconstructed coincides with a cusp, T-junction or degenerate visible rim point.*

The explanation of Rules 1 and 3 is simple and is omitted. The key to the solution of the semi-global curve reconstruction task lies in Rule 2, which constrains the long-range effect of the observer's motion on the set of reconstructed curve points. More specifically, suppose the viewpoints corresponding to Figures 6(a)-(c) are $c(t_1^n)$, $c(t_2^n)$ and $c(t_3^n)$, respectively, and the line through $c(t)$ and $\alpha(t)$ is $l(t)$. To achieve semi-global curve reconstruction, the length of the segment between $\alpha(t_1^n)$ and $\alpha(t_3^n)$ must diminish if and only if $\alpha(t_1^n)$ asymptotically approaches a visual event curve that potentially bounds a reconstructible surface region. Now consider Figure 7. Since $\alpha(t_3^n)$ is a cusp point, the line $l(t_3^n)$ is along an asymptote at $\alpha(t_3^n)$. Therefore, if $\psi(t)$ is the angle between $l(t)$ and the corresponding asymptote at $\alpha(t)$, we can conclude that a necessary and sufficient condition for the curve point $\alpha(t_3^n)$ to become a cusp visible rim point is that $\psi(t)$ goes to zero as t approaches t_3^n.

Clearly, if $\psi(t_1^n)$ is large, the length of the segment between $\alpha(t_1^n)$ and $\alpha(t_3^n)$ will also be large. It is therefore necessary to require $\psi(t_1^n)$ to be large. But how large can we make $\psi(t_1^n)$? If $\psi(t_1^n)$ is too large, the line $l(t_1^n)$ may approach the other asymptote at $\alpha(t_1^n)$; the best we can do is to ensure that $c(t_1^n)$ is in the *middle* of the visibility arc, which in this case is bounded by the two asymptotes at $\alpha(t_1^n)$. At that viewpoint, $\psi(t_1^n)$ will form equal angles with both asymptotes at $\alpha(t_1)$. Obeying Rule 2 is quite easy: The observer determines the extent of the visibility arc containing $c(t_1^n)$, and then moves to the middle of that arc. To measure the extent of the visibility arc, the observer can simply move on $T_{\alpha(t_1^n)}(S)$ first in a clockwise and then in a counterclockwise direction, terminating its motions when a cusp or T-junction is formed at the projection of $\alpha(t_1^n)$.

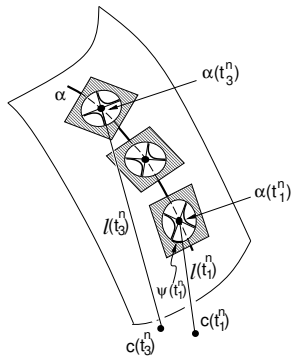


Figure 7: Geometry of the reconstruction of a segment of curve α in Figure 6. Viewpoints $c(t_1^n), c(t_3^n)$ correspond to Figures 6(a) and (c), respectively. The tangent plane and Dupin's indicatrix of points $\alpha(t_1^n)$ and $\alpha(t_3^n)$ is also shown.

By following Rules 1-3, the observer is able to achieve semi-global curve reconstruction: The distance between $\alpha(t_1^n)$ and $\alpha(t_3^n)$ diminishes if and only if the visibility arc at $\alpha(t_1^n)$ degenerates to a point. This occurs only when $\alpha(t_1^n)$ approaches one of the visual event curves in Theorem 1.

6.2 Global Surface Reconstruction Task

In the global surface reconstruction task the observer must reconstruct not only points lying on a single curve drawn on the surface that intersects the visible rim at the initial viewpoint, but must also reconstruct points lying on *every* such curve that can be drawn on the surface. The following theorem shows how this task can be performed using the Incremental Reconstruction Behavior and the Ordinary and Degenerate Patch Reconstruction Behaviors (Figure 8):

Theorem 3 (Global surface reconstruction rules) *If the following four rules are obeyed, the observer will provably perform the global surface reconstruction task:*

Rule 1: *Always choose the point p on which to apply the Patch Reconstruction Behaviors so that the boundary of the already-reconstructed surface region expands uniformly.*

Rule 2: *Always execute the Ordinary Patch Reconstruction Behavior after first moving to a viewpoint c corresponding to the middle of a visibility arc of p.*

Rule 3: *When executing the Ordinary Patch Reconstruction Behavior starting from an initial viewpoint c, move on a great circle around the surface and stop only if c is reached again or the visible rim segment initially containing p (and all visible rim segments splitting from or merging with it) disappears.*

Rule 4: *To reconstruct a patch around p, apply the Degenerate Patch Reconstruction Behavior independently of whether p is ordinary, cusp, T-junction, or degenerate.*

Rule 2 is identical to the one used in Theorem 2 treating the semi-global curve reconstruction task. Rule 3 is a generalization of the corresponding rule in Theorem 2 reconstruction task in the following sense. When performing

Incremental Reconstruction Behavior	
Rules that must be obeyed	
Rule 1:	Select points uniformly on the boundary of the reconstructed region
Rule 4:	Always execute the Degenerate Patch Reconstruction Behavior independently of the type of p
Degenerate Patch Reconstruction Behavior	
Rule that must be obeyed	
Rule 2:	Always move to the middle of a visibility arc
Ordinary Patch Reconstruction Behavior	
Rule that must be obeyed	
Rule 3:	Move on a great circle around the surface and stop when either the initial position is reached again, or the visible rim segment initially containing p (and all segments splitting from it or merging with it) disappear

Figure 8: Behaviors used to perform the global surface reconstruction task. Also shown are the rules that must be obeyed when each of the behaviors are executed.

the semi-global curve reconstruction task, as in the example in Figure 6(a)-(c), the observer was required to stop only after an endpoint of the visible rim "slid over" the curve drawn on the surface. In the global surface reconstruction task, the same rule must hold for *every* curve initially intersecting the visible rim that we can draw on the surface. This would require the observer to move upward until the visible rim segment in Figure 6(a) containing $\alpha(t_1^n)$ disappears (or, equivalently, the two segments β_1, β_2 in Figure 6(c) disappear), or until the initial viewpoint is reached.

The key to solving the global surface reconstruction task is Rule 4. Rule 4 requires the observer to reconstruct several patches around the selected point p, by moving to the middle of *all* the visibility arcs of p and then executing the Ordinary Patch Reconstruction Behavior starting at each one of those viewpoints. By obeying this rule, the area of the reconstructed patch diminishes only if *all* visibility arcs at p diminish. Since this occurs only when p approaches a reconstructible region boundary, global surface reconstruction is guaranteed. Enforcing Rule 4 implies that the reconstruction process is simplified: The Ordinary Patch Reconstruction Behavior and the Degenerate Patch Reconstruction Behavior are sufficient to perform global surface reconstruction (Figure 8).

Rule 1 is a generalization of the corresponding rule in Theorem 2. It simply takes care of the fact that if the reconstruction process does not terminate and the reconstructed region is expanded in only one direction, some pieces of its boundary will never be expanded.

By obeying Rules 1-4, the observer guarantees global surface reconstruction. The Incremental Reconstruction Behavior terminates after a finite number of steps precisely when the whole surface is reconstructible. Otherwise, the set of points reconstructed converges to the reconstructible surface regions intersecting the visible rim at the initial viewpoint. Furthermore, in this case, *no* behavior can achieve global surface reconstruction in a finite number of steps: When the visible rim touches the boundary of a reconstructible region, it touches it at exactly one point, making it impossible to reconstruct the surface in every neighborhood of such a curve in a finite number of steps.

7 Concluding Remarks

We have demonstrated that an active monocular observer can use a combination of simple viewpoint control behaviors to recover a global description of a smooth, arbitrarily-shaped surface from the occluding contour. The regions that are reconstructed on the surface can be accurately characterized, and depend only on qualitative, global shape properties of the surface. Current limitations of the approach are (1) the use of an observer constrained to move on a sphere, and (2) the inability to reconstruct the entire surface of objects with concavities, which is inherent to the use of the visible rim as the primary source of shape information.

References

[1] R. Cipolla and A. Blake, "Surface shape from the deformation of apparent contours," *IJCV*, 9(2): 83–112, 1992.

[2] Y. Aloimonos, "Active vision revisited," in *Active Perception*, pp. 1–18, Lawrence Erlbaum Associates, 1993.

[3] D. H. Ballard and C. M. Brown, "Principles of animate vision," in *Active Perception*, pp. 245–282, Lawrence Erlbaum Associates, 1993.

[4] K. N. Kutulakos and C. R. Dyer, "Recovering shape by purposive viewpoint adjustment," *IJCV*, 12(2): 113–136, 1994.

[5] D. Wilkes and J. K. Tsotsos, "Active object recognition," in *Proc. CVPR*, pp. 136–141, 1992.

[6] E. Grosso and D. H. Ballard, "Head-centered orientation strategies in animate vision," in *Proc. 4th ICCV*, 1993.

[7] A. Blake, A. Zisserman, and R. Cipolla, "Visual exploration of free-space," in *Active Vision*, pp. 175–188, MIT Press, 1992.

[8] C. I. Connoly, "The determination of next best views," in *Proc. IEEE Robotics Automat. Conf.*, pp. 432–435, 1985.

[9] P. Whaite and F. P. Ferrie, "From uncertainty to visual exploration," *IEEE T-PAMI*, 13(10): 1038–1049, 1991.

[10] J. Maver and R. Bajcsy, "Occlusions as a guide for planning the next view," *IEEE T-PAMI*, 15(5): 417–433, 1993.

[11] P. Giblin and R. Weiss, "Reconstruction of surfaces from profiles," in *Proc. 1st ICCV*, pp. 136–144, 1987.

[12] J. J. Koenderink, *Solid Shape*. MIT Press, 1990.

[13] R. Vaillant and O. D. Faugeras, "Using extremal boundaries for 3-d object modeling," *IEEE T-PAMI*, 14(2): 157–173, 1992.

[14] K. N. Kutulakos, W. B. Seales, and C. R. Dyer, "Building global object models by purposive viewpoint control," in *Proc. 2nd CAD-Based Vision Workshop*, pp. 169–182, 1994.

[15] S. Petitjean, J. Ponce, and D. J. Kriegman, "Computing exact aspect graphs of curved objects: Algebraic surfaces," *IJCV*, 9(3): 231–255, 1992.

[16] K. N. Kutulakos and C. R. Dyer, "Global surface reconstruction by purposive control of observer motion," Tech. Rep. 1141, CS Department, University of Wisconsin— Madison, 1993. Available via ftp from ftp.cs.wisc.edu.