Object Exploration By Purposive, Dynamic Viewpoint Adjustment

Kiriakos N. Kutulakos
Charles R. Dyer

Vladimir J. Lumelsky

Computer Sciences Department
University of Wisconsin
Madison, Wisconsin 53706

Mechanical Engineering Department
University of Wisconsin
Madison, Wisconsin 53706

Technical Report #1124
November 1992

Abstract

We present a viewing strategy for exploring the surface of an unknown object (i.e., making all of its points visible) by purposefully controlling the motion of an active observer. It is based on a simple relation between (1) the instantaneous direction of motion of the observer, (2) the visibility of points projecting to the occluding contour, and (3) the surface normal at those points: If the dot product of the surface normal at such points and the observer's velocity is positive, the visibility of the points is guaranteed under an infinitesimal viewpoint change. We show that this leads to an object exploration strategy in which the observer *purposefully* controls its motion based on the occluding contour in order to impose structure on the set of surface points explored, make its representation simple and qualitative, and provably solve the exploration problem for smooth generic surfaces of arbitrary shape. Unlike previous approaches where exploration is cast as a discrete process (i.e., asking where to look next?) and where the successful exploration of arbitrary objects is not guaranteed, our approach demonstrates that dynamic viewpoint control through directed observer motion leads to a qualitative exploration strategy that is provably-correct, depends only on the dynamic appearance of the occluding contour, and does not require the recovery of detailed three-dimensional shape descriptions from every position of the observer.

Index terms: Purposive and qualitative vision, autonomous exploration, moving point of observation, surface visibility, provably-correct algorithms, generic smooth object surfaces, global surface geometry, occluding contour, visual events

---

The support of the National Science Foundation under grants IRI-9022608 and IRI-9196106 is gratefully acknowledged.

1 Introduction

There has been considerable interest in active vision approaches that control the viewpoint of the observer (i.e., position) by following the principle of least commitment [1]. These approaches attempt to recover a three-dimensional description of objects in the scene from the current position before deciding where to move next [2]. Consequently, they require considerable processing on each image, and sophisticated sensor configurations are required for deriving detailed information about the geometry of the object from each position [3,4]. On the other hand, although previous approaches following the purposive vision paradigm [5] have been used to control viewing parameters such as focus, vergence and fixation for tasks like obstacle avoidance and tracking, the ability to direct the observer's motion for exploring an unknown object has not been exploited.

In this paper we apply the purposive vision paradigm to the problem of controlling a moving observer in order to visually explore the surface of an object prior to recognition. We consider *object exploration* to be the problem of controlling viewpoint in order to induce the visibility of all points on the object surface that are occluded when the exploration process is initiated. Our goal is to design a strategy that operates in real time and has provably-correct properties, i.e., the motion of the observer is based on the geometry of the object.

The main idea of our approach is to use an active (i.e., mobile) observer that purposefully directs its motion in order to achieve and maintain a specific geometric relationship with the object [6, 7]. We show that the ability to control the viewpoint of the observer in a continuous fashion through directed observer motion considerably simplifies the mathematical formulation of the object exploration problem. Furthermore, by exploiting the continuous interaction between motion and visual processing, the observer can explore the object's surface by extracting sparse, efficiently-computable information (i.e., the 2D occluding contour) from each image.

The motivation for the problem we consider comes from our efforts to identify the fundamental abilities that an observer must possess in order to visually gather information about an unknown object by controlling viewpoint. In particular, a key phenomenon motivating any such object exploration process is that of occlusion. Our formulation is motivated by the need to develop an exploratory *behavior* [8–11] for inducing the visibility of previously occluded parts of the object. This type of behavior is particularly important for a number of tasks, such as generating three-dimensional object-centered descriptions that cover the entire surface of a non-convex object, searching for surface markings during object recognition, avoiding occlusion during object grasping [12, 13], and object model acquisition by obtaining a series of aspects [14] that completely describes the object. In addition, although we assume in our analysis that the observer is moving and the object is stationary, our analysis also applies to the case where the observer is stationary and manipulates the object being explored (e.g., by grasping it or putting it on a turntable).

1.1 Related Work

In order to completely explore the surface of an object, an observer must be able to perform two tasks: (1) to change viewpoint in order to induce the visibility of points on the object surface that were occluded from all previous viewpoints, and (2) to determine when the exploration goal has been achieved. In order for the observer to perform these two tasks it must represent in some way the set of surface points becoming visible during the course of the exploration process or, alternatively, the set of surface points that remain occluded. Previous approaches that controlled viewpoint in order to induce the visibility of occluded parts of an object considered the process to be discrete: Viewpoint changes were determined by a sequence of observer positions [4, 15–18] generated by going through a three-step cycle: (1) Recover a detailed description of the visible or the occluded portions of the object from the current position of the observer, (2) merge this description with those produced from all previous positions, and (3) move to a new position.

An important characteristic of these strategies is that they try to select a sequence of viewpoints where the appearance of the object is as different as possible from those at all previous viewpoints. For example, Connoly [15] attempted to find a viewpoint that maximizes the area of object surface points that will be visible for the first time. Unfortunately, predicting changes in object appearance between two arbitrary viewpoints requires detailed information about the geometry of the object and makes viewpoint selection a computationally-expensive search problem [4,15] not amenable to real-time implementation. In addition, the observer does not have any information about the geometry of the object within the volume of points that are occluded from all previous viewpoints and, hence, the observer cannot accurately characterize the set of surface points becoming visible for the first time from a particular viewpoint. To date, there has been no attempt to characterize the class of objects for which the developed strategies successfully accomplish the exploration task, i.e., that all points on the object surface become visible.

We feel that theoretical work on the object exploration problem is necessary for two reasons. First, the visual complexity of an object can be significant even when it is bounded by a simple non-convex surface (e.g., a sphere with a dent [19]). This makes it very hard to predict the effectiveness or the performance of heuristic approaches to the exploration problem where, for example, substantial amounts of occlusion may be present. This is also the case for approaches using general optimization criteria for controlling the viewpoint of the observer [4] since their properties are highly dependent on the optimization criterion used, and their correctness and performance are hard to predict for objects bounded by surfaces with complicated geometry. Second, a theoretical analysis will provide the basis for designing practical, real-time strategies for performing the exploration task and for characterizing their effectiveness (e.g., by considering which assumptions used in the theoretical analysis can be violated and studying their effects based on object geometry).

Finally, most previous approaches consider the task of exploring an unknown object equivalent to the task of building a three-dimensional object description. One of the central aspects of our formulation of the object exploration problem lies in the distinction we make between the task of exploring an object and the task of building a three-dimensional description of the object. This is particularly desirable if, for example, an exploratory behavior is used for finding specific markings on the object's surface. On the other hand, if the exploration process is used for the purpose of building a three-dimensional object description, dynamic control of the observer's viewpoint implies that successive views of the object will be very similar, making the shape recovery problem well-posed and the fusion of multiple views easier [20]. Furthermore, by employing a moving observer our approach is well-suited for building dynamic, viewer-centered object representations where the dynamic, viewpoint-dependent appearance of objects is modelled explicitly [21].

1.2 Dynamic, Purposive Object Exploration

In our approach, viewpoint control and visual processing are continuous, interdependent processes that occur simultaneously. The approach is based on two important observations:

- Dynamic viewpoint control corresponds to motion decisions that induce infinitesimal position changes.
- If the motion of the observer is directly related to the geometry of the visible object surfaces, the object appearance under an infinitesimal position change will change in a predictable manner.

These observations suggest that dynamic control of the observer's viewpoint can force the evolution of the appearance of the object to be locally predictable in time. Hence, this evolution can be controlled by directing the observer's motion. In [6] we showed that by appropriately controlling the evolution of the object appearance based on sparse and quickly-computable visual information (curvature measurements on the occluding contour), the mathematical formulation of the problem of recovering surface shape information becomes considerably simplified and leads to a qualitative, active vision strategy for solving it. In this paper we apply the same basic principle in order to design a strategy for addressing the object exploration problem using shape information derived from the occluding contour.

The central concept in our approach is the *exploration frontier*. This is the one-dimensional set of points on the surface that bounds the points seen so far during the exploration process. We address the exploration problem by studying the dynamic evolution of the exploration frontier under directed observer motion. The exploration frontier provides all the information necessary for a moving observer to direct its motion so that previously-occluded points on the surface become visible for the first time, or to determine whether the exploration process is complete. These two properties of the exploration frontier allow the provably-complete exploration of an arbitrary surface when this boundary representation is used to guide the motion of the observer.

Unlike previous strategies that move the observer to a sequence of arbitrarily-distanced positions between which prediction of the evolution of this frontier is impossible, by dynamically controlling the viewpoint of the observer we can predict the dynamics of the exploration frontier and can plan the instantaneous motion of the observer accordingly. Hence, instead of choosing viewpoints that give maximally-different views of the object, our approach is to do the opposite: We exploit the similarity between the appearances of the object from viewpoints along the observer's path that are infinitesimally close to each other in order to accurately and efficiently predict and control the dynamics of the exploration frontier. Dynamic viewpoint control is therefore a natural way to address information-gathering tasks that must be performed in real-time and are applied to unknown objects whose global structure cannot be predicted beforehand.

The significance of our approach lies in (1) considering the object exploration task as a continuous process where viewpoint control and visual processing are tightly coupled, and (2) using a moving observer whose motion is purposefully controlled so that the representations of the object are simple, qualitative, and do not require reconstructing the shape of the visible portions of the object surface from each viewpoint. Our approach demonstrates that purposive, dynamic control of the observer's viewpoint reduces the amount of visual information that must be processed, simplifies the object exploration process, and leads to a provably-correct, efficient exploration strategy.

The rest of the paper is organized into three parts. The first part, presented in Section 2, develops the basic tools necessary for studying the exploration problem. It defines the exploration frontier and presents the crucial link between (1) the ability to dynamically control the observer's viewpoint through directed motion, (2) the problem of inducing the visibility of points on a smooth surface, and (3) the occluding contour. The main result of Section 2 is Theorem 2.1, which can be thought as a formal analog of Gibson's *law of reversible occlusion* [22]. The second part of the paper, presented in Section 3, builds on this basic result to formally develop an exploration strategy for the case where the object is bounded by a single, smooth, generic surface. The third part of the paper, Section 4, focuses on a specific aspect of the exploration process. In particular, we present a simple strategy that allows a moving observer to explore a cylindrical strip of points on the surface by directing the motion of the observer based on the dynamic shape of the occluding contour.

2 Dynamic Evolution of the Exploration Frontier

Any solution to the exploration problem must address the issues of determining if the object is completely explored and, if it is not, ensuring that positive progress is always made. We attack these two issues by studying how the motion of the observer affects the evolution of the exploration frontier.

Suppose that the observer, starting from an initial position i, traces a continuous path in an attempt to explore an object bounded by a smooth surface. This path can be parameterized by a continuous curve $c(t)$ of finite length, where $c(0) = i$. At any position $c(t)$ along this path the exploration goal is only partially fulfilled; points on the surface will either be *explored* if they were visible from a previous position $c(s)$ on the path, $s \leq t$, or will otherwise be *unexplored*. The *exploration frontier* consists of a collection of closed curves bounding the set of explored points. We study its evolution by considering the geometrical and topological changes taking place in this collection of closed curves when the observer is at a position $c(t)$ and moves with an instantaneous velocity $v(t)$.

The basic idea behind our approach is that the evolution of the exploration frontier can be studied by looking at how the motion of the observer affects the geometry and topology of the occluding contour. The *occluding contour* is defined as the projection of the visible rim into the image (Figure 1). The *rim* is the set of points on the surface for which the line connecting them to the observer's position is tangent to the surface at those points. The *visible rim* consists of the rim points that are visible.

The connection between the evolution of the exploration frontier and the dynamics of the occluding contour is a very important one because the geometry and topology of the occluding contour has been widely studied (e.g., [24]). More importantly, this connection allows us to use the occluding contour and the visible rim as our basic forms of visual input: The occluding contour provides all the information necessary for recovering surface shape by a moving observer [20, 23, 25], and the dynamics of its appearance can be used for substantially constraining algorithmic search when exploration is combined with recognition tasks [26]. From a practical standpoint, the recovery of sparse, qualitative visual information such as the occluding contour imposes a relatively light computational burden on the observer, making the occluding contour suitable for problems requiring its real-time detection and tracking [20, 27, 25].

To see how the evolution of the exploration frontier and the dynamics of the occluding contour are related, note that changes in the exploration frontier occur when unexplored points on the surface become visible. The evolution of the exploration frontier therefore depends entirely on the set of surface points becoming visible due to the motion of the observer. A point p on the surface is visible if the open line segment connecting it to the position of the observer does not intersect the surface; otherwise p is occluded. Changes in the visibility state of p occur only when the line segment connecting it to the position of the observer is tangent to the surface at a point p' (possibly coincident with p) and does not pierce the surface elsewhere. This locus of points is usually called the *occlusion boundary* (Figure 2). It is a collection of curves on the

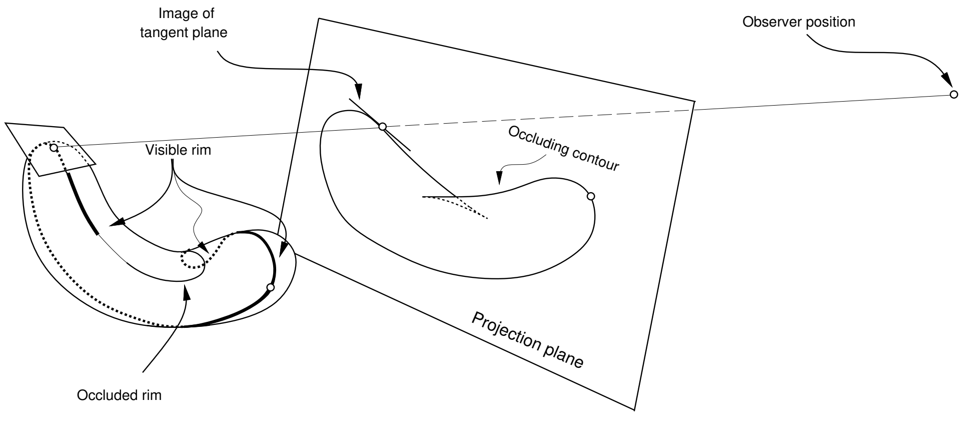

Figure 1: The visible rim and the occluding contour corresponding to the projection of a bean-shaped surface on a planar image (adapted from [23]). In this example part of the rim is occluded by the surface.

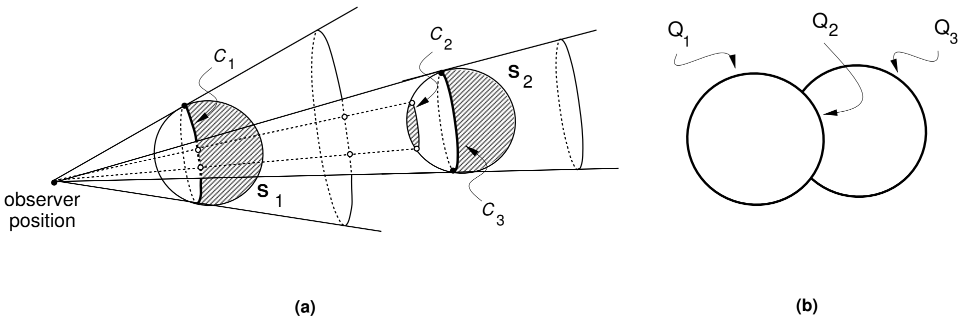

Figure 2: The relation between the occlusion boundary and the visible rim. (a) Sphere S_1 partially occludes sphere S_2 (a side view is shown). Curves C_1 and C_3 correspond to the rim. The occlusion boundary also contains C_2, which is the intersection of the sphere S_2 with the conical surface defined by the position of the observer and C_1. (b) The appearance of the two spheres from the position of the observer. Q_1 and Q_3 are the projections of C_1 and the visible points of C_3 respectively. The projection, Q_2, of C_2 overlaps with that of Q_1.

surface that contains all points becoming visible or occluded due to the instantaneous direction of motion of the observer. Furthermore, the occlusion boundary is a superset of the visible rim and its projection on the image is identical to the occluding contour.[1]

We study the connection between the evolution of the exploration frontier and the dynamics of the occluding contour in two steps. In the first step we consider how the set of points on the occlusion boundary becoming visible is affected by (1) the instantaneous direction of motion of the observer, and (2) the geometry of the surface. This step forms the basis of our approach and is presented in Section 2.1. In the second step, presented in Section 3, we consider how the motion of the observer can be controlled in order to control the evolution of the exploration frontier, determine which of the points on the occlusion boundary are unexplored, and hence, guide the exploration process.

[1]In order to emphasize the transitional visibility state of the points on the occlusion boundary, we will consider these points as being neither visible nor occluded. Under this stricter definition for point visibility, the occlusion boundary may contain unexplored points.

2.1 Inducing Visibility by Purposive Observer Motion

This section presents the crucial link between the ability to control the observer's viewpoint in a continuous fashion, the problem of inducing the visibility of points on a smooth surface, and the occluding contour. In the following we assume the spherical projection model for image formation. In the spherical projection model the visible points on the surface are projected to a sphere of fixed radius that is centered at the observer's position. Unlike imaging models where the points on the surface are projected to a plane (e.g., Figure 1), the spherical projection model permits an arbitrary field of view for the observer which simplifies the analysis of the exploration problem.[2]

We consider the observer to be positioned at a point $c(t)$ and to have an instantaneous velocity $v(t)$. Points on the surface becoming visible due to the instantaneous velocity of the observer must be contained on the occlusion boundary. The projection of this boundary is identical to the occluding contour. The line connecting a point p on the occlusion boundary to the camera lens center must be tangent to the surface at at least one point p_{rim} on the visible rim. We can determine whether the points p and p_{rim} become visible by looking at the surface normal at p_{rim}: (Figure 3(a)):

Theorem 2.1 (Visibility transition dynamics) *Let $c(t)$ be the position of the observer and let p_{rim} and p (possibly identical to p_{rim}) be occlusion boundary points on a smooth surface such that the line segment connecting $c(t)$ and p is tangent to the surface only at p_{rim}. Suppose that c is differentiable at t, with $c'(t) = v(t)$. If $N(p_{rim})$ is the outward surface normal at p_{rim} at time t,*

1. *If $N(p_{rim}) \cdot v(t) > 0$, p and p_{rim} will become visible*
2. *If $N(p_{rim}) \cdot v(t) < 0$, p and p_{rim} will become occluded*

[2] The assumption of an arbitrary field of view, however, is not inherently necessary to our approach. When the field of view is restricted, the only additional mechanism required is a fixation and tracking mechanism used to track a segment of the occluding contour (e.g., see [20]).

3. *If* $N(p_{rim}) \cdot v(t) = 0$, *(a)* p_{rim} *will either remain on the occlusion boundary or become occluded, and (b) the visibility of* p *is completely determined by the local shape of the surface at* p_{rim} *and by* $v(t)$.

Proof. First suppose that the observer does not move on the tangent plane of the surface at p_{rim}, i.e., $N(p_{rim}) \cdot v(t) \neq 0$. Since p_{rim} and p are occlusion boundary points, their visibility is determined by the sign of $N(p_{rim}) \cdot [p_{rim} - c(t)]$ [24]. In particular, for any visible surface point x we have $N(x) \cdot [x - c(t)] < 0$. Changes in the visibility state of p and p_{rim} occur due to (1) changes in the sign of this dot product, or (2) the existence of a surface point y, not in the neighborhood of p_{rim}, which occludes either p or p_{rim}. Since the line segment connecting p and $c(t)$ is tangent to the surface only at p_{rim}, the visibility state of p and p_{rim} under an infinitesimal position change depends only on changes in the sign of the above dot product.

Since p_{rim} is a visible rim point, the dot product is zero at time t. Therefore, the visibility of p_{rim} (and p) under an infinitesimal position change depends on the sign of the derivative $\{N(p_{rim}) \cdot [p_{rim} - c(t)]\}'$. We have:

$$\begin{aligned}
\frac{d}{dt}\{N(p_{rim}) \cdot [p_{rim} - c(t)]\} &= N(p_{rim}) \cdot \frac{d}{dt}[p_{rim} - c(t)] + [\frac{d}{dt}N(p_{rim})] \cdot [p_{rim} - c(t)] \\
&= -N(p_{rim}) \cdot c'(t) \\
&= -N(p_{rim}) \cdot v(t)
\end{aligned}$$

If the observer moves on the tangent plane of p_{rim}, p and p_{rim} may become occluded by points in the neighborhood of p_{rim}. In particular, p_{rim} will remain on the rim; it will be visible (i.e. it will remain on the occlusion boundary) unless the line connecting $c(t)$ and p_{rim} is along an asymptotic direction of the surface at p_{rim} [6]. Following the same analysis as in [6] it can be shown that if p_{rim} is hyperbolic, p will be occluded by a point in the neighborhood of p_{rim}. If p_{rim} is elliptic, p will remain visible but leave the occlusion boundary. If p_{rim} is parabolic, the visibility of p also depends on the direction of motion on the tangent plane at p_{rim}. $\square$

A simple, qualitative strategy for controlling the motion and, hence, the viewpoint of the observer now emerges:

Dynamic Viewpoint Control Algorithm

Step 1: Select a point p on the occlusion boundary that satisfies the conditions of Theorem 2.1 and whose visibility is desired. This selection is done indirectly by selecting p's projection, p_{occ}, on the occluding contour.[3]

Step 2: Compute the surface normal for the rim point p_{rim} projecting to p_{occ} using the tangent of the occluding contour at p_{occ}.

Step 3: Determine the instantaneous direction of motion, $v(t)$, of the observer so that $N(p_{rim}) \cdot v(t) > 0$.

If $N(p_{rim}) \cdot v(t) > 0$, all points projecting to a *segment* of the occluding contour that contains p_{occ} will become visible. In general, points on the occlusion boundary projecting to other segments of the occluding contour may also become visible. Under continuous motion of the observer and when the topology of the occlusion boundary does not change, these segments will trace strips on the surface (Figure 3(a)). In the following, we refer to the segments of the occlusion boundary that satisfy the above inequality as the *strip-generating segments* at time t. The strips produced by these segments can be considered as families of curves on the surface, indexed by the time parameter t [20], forming a two-dimensional subset of the surface. We study the problem of controlling the dynamics of these strips in the following section.

[3] For example, it can be shown that if $c(t)$ is a stable position, the theorem holds for any point that projects to an occluding contour point, p_{occ}, at which the occluding contour does not self-intersect.

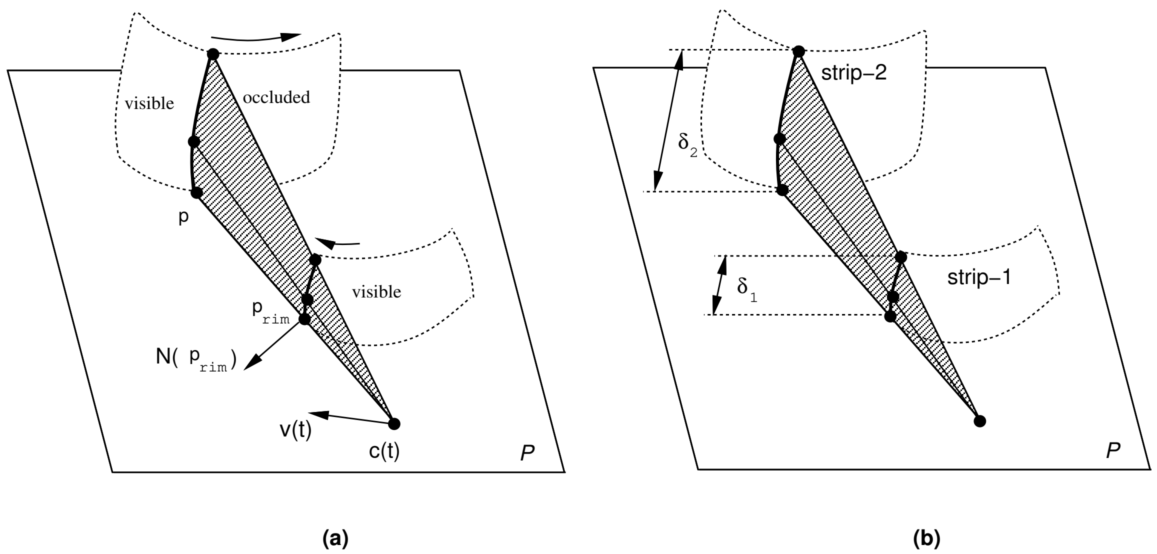

Figure 3: The dynamics of the occlusion boundary. The occlusion boundary is shown in bold. (a) The plane P is defined by the instantaneous velocity $v(t)$ of the observer and the normal at the rim point p_{rim}. The points p and p_{rim} become visible due to the motion of the observer. The occlusion boundary traces two strips on the surface; the arrows indicate the motion of one of the endpoints of the occlusion boundary segments tracing these strips. Only the portions of the strips *above* the motion plane are shown, i.e., the strip points with positive distance from the motion plane. (b) If the observer is restricted to move on the plane P, the distance δ_a corresponding to each of these strips can be computed from the distances δ_1, δ_2. Note that $\delta_2 > \delta_1$ and hence the distance δ_1 (which is the distance of a point on the visible rim from P) can be used to estimate the distance δ_2. In general, δ_1 will not be the maximum distance from P of the points on the occlusion boundary segment C.

3 Exploring Objects Bounded by Generic Surfaces

In this section we consider the problem of exploring an object bounded by a closed, smooth, generic surface. Section 3.1 presents our approach to structuring the evolution of the exploration frontier by purposefully controlling the motion of the observer. Section 3.2 presents the resulting formulation of the exploration problem. Section 3.3 considers the issues involved in the design of a provably-correct exploration strategy.

3.1 Structuring the Exploration Frontier by Purposive Observer Motion

The exploration frontier is the boundary of the set of explored points on the surface. We have seen that this set is the union of the strips of the surface points becoming visible. When the observer moves along an arbitrary, continuous path, these strips may intersect each other, self-intersect, split, merge, and new strips may be created. The dynamics of these strips are dictated by the instantaneous velocity of the observer, the changing topology of the occlusion boundary, and the local shape of the surface in its neighborhood. The goal of our approach is to find ways in which the dynamics of these strips can be controlled by controlling the observer's motion in order to control and represent the boundary of their union, i.e., the exploration frontier.

In our analysis we assume that the object is bounded by a smooth and generic surface [19]. Generic surfaces are surfaces whose topological and geometrical characteristics are stable under infinitesimal perturbations. They correspond to a very general class of surfaces since an infinitesimal perturbation of a non-generic smooth surface will make it generic. The absence of complicated degeneracies on such surfaces (e.g., planar surface patches) makes them ideal for an initial investigation of the object exploration problem, allowing us to focus on the general issues.

In order to impose structure on the shape of the strips and the exploration frontier we base our exploration strategy on the following:

- The observer controls its motion in order to impose structure on a *single* strip at a time. This simplifies the problem of controlling the observer's motion during the exploration process by focusing the attention of the observer on a single subtask.

- The observer controls its motion with respect to a selected *motion plane* so that the strip of points becoming visible is topologically equivalent to a cylinder with one of its bounding curves strictly on each side of this plane.

The main idea is to develop an exploration strategy whereby the observer controls its motion with respect to a family of parallel planes in order to completely cover the surface with partially overlapping cylindrical strips. The strategy has two phases. The first phase is a *cylindrical strip exploration phase* in which the observer explores a cylindrical strip on the surface by moving primarily on the selected motion plane. In the second phase, the *inter-plane motion phase*, the observer first decides whether to move above or below the current motion plane and then moves in the chosen direction to a different plane in the family in order to explore a new strip on the surface.

A major advantage of this strategy is that the simple topological structure of the strips makes their description particularly simple: They can be described by two distances, the minimum distances δ_a and δ_b of their bounding curves above and below the motion plane from the motion plane itself. Furthermore, the shape of the strip of points becoming visible depends on the dynamic appearance of the occlusion boundary and, consequently, the occluding contour. This fact makes it possible to control the motion of the observer based on the appearance of the occluding contour in order to force the shape of the strips to be cylindrical, with $\delta_a, \delta_b \neq 0$ (see Section 4.1).

From a practical standpoint, the advantage of using cylindrical strips as defined above is that their description can be derived from the dynamic appearance of the occluding contour. Since the strips are defined as families of segments of the occlusion boundary, the boundaries of the

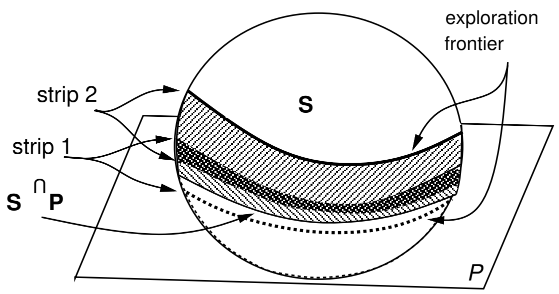

Figure 4: Exploring an object bounded by a convex surface S in strips. P corresponds to the current motion plane. The next motion plane is chosen so that it is above the current motion plane. The boundary of the current strip will be fully contained within the next strip if the next motion plane has distance less than δ_a from the current motion plane. The upper boundary of the current strip and the lower boundary of the next strip are not needed to describe the exploration frontier. The arrows indicate the bounding curves of the two strips and the exploration frontier after the cylindrical strip exploration phase phase is completed for the next motion plane.

strips correspond to the curves traced by the endpoints of their defining segments. Therefore, the computation of δ_a and δ_b can be performed by computing, at every position of the observer, the distance of these endpoints from the motion plane (Figure 3(b)) using an approach such as the one by Cipolla and Blake [20].

3.2 Formulating the Object Exploration Problem

The cylindrical shape of the strips and the structured motion of the observer can be used to simplify the structure of the exploration frontier considerably. This is particularly evident if, for example, we restrict ourselves to the exploration of objects bounded by convex surfaces (Figure 4). In the general case where the surface of the object to be explored is non-convex, our approach (1) extends the cylindrical strip exploration phase to handle changes in the geometry and topology of the occlusion boundary as detected by changes in the geometry and topology of the occluding contour, and (2) describes the surface in a way that is compatible with the notion of exploring the surface by a collection of overlapping cylindrical strips. We use results from the differential topology of surfaces that allow us to represent the surface as a graph and to reduce the exploration process to an abstract graph exploration problem [28].

The graph representation of the surface we employ is a direct generalization of a tree-like representation developed by Koenderink for the qualitative description of gray-scale images [29]. To see how such a representation can be derived for a given surface S, fix a plane P_0 in the family of parallel motion planes selected by the observer. Members of this family can be indexed by d, the distance of a plane in the family from P_0. For generic surfaces, $S \cap P_d$ is a collection of closed curves and isolated points.[4] As d varies continuously, the connectivity of these curves changes at discrete values of d; curves merge, separate, disappear, and new curves are created (Figure 5(a)). The changes in the topology of $S \cap P_d$ occur at values of d corresponding to planes that are tangent to the surface. The intersection of the surface with any plane P_d that is not tangent to the surface consists of a collection of simple closed curves. For each curve, $\alpha_d(s)$, in the collection there is an open interval (d_1, d_2) of maximum length such that the function $F(d, s) = \alpha_d(s)$ corresponding to the family $\mathcal{A} \equiv \alpha_{d \in (d_1, d_2)}(s)$ is a homotopy, i.e., for any pair of numbers d, d' with $d_1 < d < d' < d_2$, the function F corresponds to a continuous deformation of $\alpha_d(s)$ into $\alpha_{d'}(s)$. Furthermore, the family $\mathcal{A}$ corresponds to a cylindrical piece of the surface with its boundaries being the curves $\alpha_d(s)$ and $\alpha_{d'}(s)$.

We can think of the surface as being composed of a collection of such families $\mathcal{A}_1, \ldots, \mathcal{A}_m$ of simple closed curves, where each family corresponds to a function F_j that is a homotopy. Considered as a set of points on the surface, the family $\mathcal{A}_j$ is an open set; the boundary of this set is contained in the intersections $S \cap P_{d_1}$ and $S \cap P_{d_2}$ and is either an isolated point of tangency of the surface with P_{d_1} or P_{d_2}, or it is a closed curve containing such a point. This leads to a natural, qualitative representation of the surface as an undirected graph G, which we call the *surface graph*, that is defined as follows (Figure 5(b)):

[4] Non-generic surfaces may have a whole region of contact with the plane P_d.

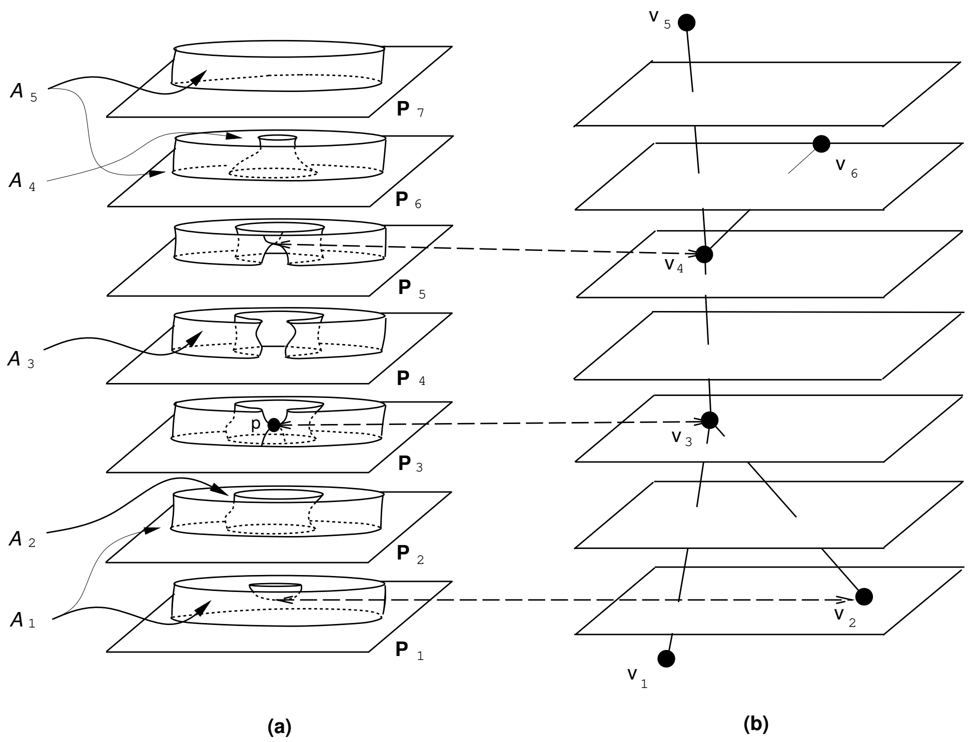

Figure 5: (a) Topological changes in $S \cap P_d$ for a generic cylindrical-shaped surface S with a concavity. The surface can be partitioned into a number of cylindrical-shaped slices of various widths each corresponding to the families $\mathcal{A}_1, \ldots, \mathcal{A}_5$. (b) The surface graph of S. Nodes in the graph are associated with points of tangency of S with a plane from the family $P_{d \in \Re}$. For example, v_3 corresponds to the (hyperbolic) point p contained in the boundaries of $\mathcal{A}_1, \mathcal{A}_2$, and $\mathcal{A}_3$. v_2 corresponds to the (elliptic and concave) point where a plane from the family $P_{d \in \Re}$ between P_1 and P_2 is tangent to the surface and bounds $\mathcal{A}_2$ from below. The observer has no information about the structure of the surface graph (i.e., its vertices and their connectivity) when the exploration process is initiated.

- Each point of tangency $p_d \in P_d$ is represented by a vertex v_{p_d} in G.

- Two vertices $v_{p_{d_1}}, v_{p_{d_2}}$ in G with $d_1 \neq d_2$ are connected by an edge whenever a family $\mathcal{A}_j$ whose corresponding set of points on the surface contains p_{d_1} and p_{d_2} in its boundaries. Since the boundaries of such a set lie on exactly two distinct parallel planes, each family $\mathcal{A}_j$ contributes exactly one edge to the graph.

The surface graph is particularly simple in the case of generic surfaces. Specifically, the surface in the neighborhood of the points of tangency with the planes $P_{d \in \Re}$ can be parameterized by a height function H, measuring the distance of the points in the neighborhood from P_0. Tangency points correspond to *critical values* of the height function, i.e., to images of points on P_0 where the gradient $\vec{\nabla} H$ vanishes [29–31]. A basic result from Morse theory is that the surface in the neighborhood of the critical values of H can be approximated by a height function that is a quadratic polynomial. With probability 1, the surface is locally convex or concave in the neighborhood of points of tangency that are isolated in $S \cap P_d$, and is saddle-shaped for points of tangency that belong to curves in $S \cap P_d$.[5] Consequently, vertices in G have either one or three incident edges depending on whether or not they correspond to a point of tangency that is isolated in $S \cap P_d$ (Figure 5(b)).

Our exploration strategy is to induce visibility for the points on the surface corresponding to the families $\mathcal{A}_j$ one at a time. The set of points corresponding to each $\mathcal{A}_j$ becomes explored by ensuring that the strips of points becoming visible during the exploration process form overlapping cylindrical strips that cover this set. This task is analogous to the task of exploring an object bounded by a convex surface. Furthermore, since each $\mathcal{A}_j$ corresponds to an edge in the surface graph, there is an analogy between exploring all points corresponding to $\mathcal{A}_j$ and traversing an edge in the surface graph. Hence, the object exploration problem reduces to the problem of traversing all edges of the (unknown) surface graph G (Figure 5(b)).

There are several advantages in reducing the object exploration problem to the exploration of the surface graph. First, once the reduction is made, correctness, convergence, and efficiency issues can be examined from a graph-theoretical standpoint. The importance of this ability is amplified by the close relation between the surface graph and the *qualitative* geometry of the surface it represents. For example, the number of vertices in the graph cannot exceed the

[5]The set of orientations for the family $P_{d \in \Re}$ for which there is a plane in the family tangent to a parabolic surface point is a zero-measure subset on the sphere of possible plane orientations.

number of points on the surface having the same normal, a number that depends on the folds of the Gauss map [19]. Currently, several algorithms exist that solve variations of the abstract graph exploration problem (e.g., [28]).

Second, the approach allows the use of a qualitative representation of the exploration frontier containing only the information necessary for solving the exploration problem: The exploration frontier is represented by the set of vertices in the graph adjacent to at least one traversed and one untraversed edge. The tangency points corresponding to these vertices are detected by the observer and serve as surface shape features guiding the exploration process. Third, at the end of the exploration process the observer will have constructed the complete surface graph, and hence, will have computed a qualitative description of the surface that enables subsequent processing to compute qualitative properties of the surface such as its genus.

Although the object exploration problem becomes simple once it is reduced to the exploration of the surface graph, the graph reduction process itself is far from trivial and requires the use of several visually-guided motion control strategies that ensure the correctness of the reduction. Section 3.3 presents the main issues involved in the reduction process.

3.3 Reducing Object Exploration to Exploration of the Surface Graph

In order to model the exploration of the object as a graph exploration process, we need to establish a precise correspondence between the steps of the two processes. Specifically, the graph exploration problem can be formulated as follows: An automaton is initially positioned at a vertex v_0 of a finite graph and its goal is to traverse every edge of the graph. The graph is unknown to the automaton in the sense that it does not have any initial information about the number or connectivity of the vertices in the graph. Furthermore, when the automaton is positioned at a vertex v, it can determine the edge e in the graph last used to reach v and the edges incident to v that were previously used to leave v. The automaton can leave v by traversing a previously untraversed edge incident to v and then marking it traversed, or it can backtrack by traversing the edge e in the opposite direction.

Existing algorithms for exploring an arbitrary, unknown graph rely on depth-first search (e.g., Tarry's algorithm [28]). They rely on three important assumptions: (1) each edge in the graph needs to be marked traversed only once, (2) when the automaton is positioned at a vertex v, it can determine whether an edge incident to that vertex was previously marked traversed, and (3) the automaton has the ability to backtrack to the previous vertex on its path. In order to reduce the object exploration problem to the above graph exploration problem we must ensure that the observer's viewing strategy satisfies these three assumptions.

The first assumption requires that when the observer chooses to extend the set of explored points by exploring points corresponding to the family $\mathcal{A}_j$ (i.e., the automaton chooses to traverse the unmarked edge e_j), the observer must be able to control its motion so that *all* points corresponding to $\mathcal{A}_j$ become explored. Even though in the graph exploration problem the traversal of an edge is an atomic operation, in the object exploration problem the set of points corresponding to $\mathcal{A}_j$ must be explored by generating a finite-length path for the observer. Furthermore, the observer must be able to determine when this exploration process is completed, i.e., that a new vertex is reached in the graph. These issues are addressed during the cylindrical strip exploration phase, and are considered in Section 4.

The second assumption requires that when the observer determines that all points corresponding to a family $\mathcal{A}_j$ have been explored, it can also decide whether all points in the families adjacent to A_j in the surface graph have also been explored. In addition, if one or more of these families contains unexplored points, the observer must be able to initiate a sequence of cylindrical strip exploration phases in order to induce the visibility of all points corresponding to these families.

The third assumption requires the observer to trace a path that leads to the position corresponding to the vertex last visited in the surface graph. It is possible to show that this can be achieved without maintaining a representation of the continuous path that led the observer to the current position. Rather, the observer can reach that position by following a polygonal path containing one point from each motion plane selected during a cylindrical strip exploration phase.

4 Cylindrical Strip Exploration Phase

In the cylindrical strip exploration phase, the observer is positioned on a plane P_d belonging to an *a priori* specified family of parallel planes $P_{d\in\Re}$. The goal of this phase is to explore a cylindrical strip of points contained in a prespecified family $\mathcal{A}$ such that the distance of its boundary above or below the motion plane is greater than zero. The family is represented by the projection of a visible point $p_{\mathcal{A}}$ contained in $\mathcal{A} \cap P_d$ (Figure 6). Without loss of generality, we assume that the observer is required to explore a cylindrical strip of points so that the distance of its boundary above P_d (i.e., in the direction of the plane's normal) is greater than zero. The information determined at the end of the cylindrical strip exploration phase consists of (1) a distance δ_a, indicating that all points in $\mathcal{A}$ above P_d at a distance less than or equal to δ_a have been made visible, (2) whether or not the tangent plane bounding $\mathcal{A}$ above P_d has distance less than δ_a from P_d (i.e., that the edge in the surface graph corresponding to $\mathcal{A}$ has been traversed), and (3) a point $p'_{\mathcal{A}} \in \mathcal{A}$ on $P_{d+\delta_a}$, whenever the tangent plane bounding $\mathcal{A}$ has distance greater than δ_a (which will serve as the point $p_{\mathcal{A}}$ for the next cylindrical strip exploration phase to be performed on $P_{d+\delta_a}$). In the following, we assume that P_d is not tangent to the surface,[6] and use P_d and $\mathcal{A}$ to refer to the motion plane and the selected family of points on the surface, respectively.

Although this phase is straightforward in the case where the viewed surface is convex, the issues involved for arbitrary surfaces are nontrivial. The resulting exploration process is event-based: The observer performs specific actions depending on the occurrence of specific changes in the geometry and topology of the occluding contour. In particular, the observer must perform the following:

Enforce the cylindrical strip property. In order to ensure that a cylindrical strip of points corresponding to a family $\mathcal{A}$ is explored, the observer must be able to determine which strip-generating segments (Section 2.1) trace strips that actually contain points in $\mathcal{A}$ and

[6] An infinitesimal perturbation of d is sufficient to ensure this condition.

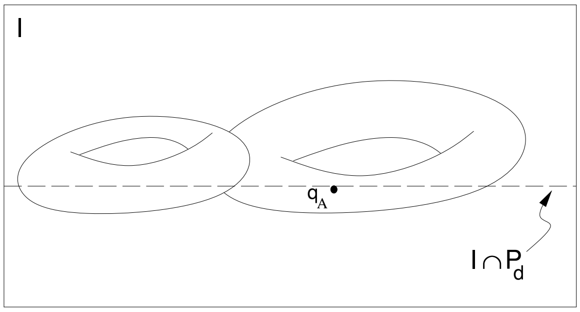

Figure 6: Representing the family $\mathcal{A}$ of surface points. The figure shows the occluding contour of two torus-like surfaces. The dotted line represents the intersection of the motion plane P_d with the image I which is drawn as a plane here, for simplicity. Point $q_{\mathcal{A}}$ is the projection of $p_{\mathcal{A}}$. Since there are two curves in the intersection of the surfaces with P_d, $q_{\mathcal{A}}$ is used to determine the curve corresponding to the selected family $\mathcal{A}$.

ensure that their union is connected and topologically equivalent to a cylinder (Figure 7(a)).

Determine the strip width. From any single view the observer can obtain surface shape information only for points that lie on the visible rim. Furthermore, the distance of the boundaries of $\mathcal{A}$ from P_d is not known *a priori*. This makes it impossible to determine exactly the boundary of the cylindrical strip corresponding to the family $\mathcal{A}$ and, hence, this boundary must be estimated while also ensuring that all points in $\mathcal{A}$ become explored (Figure 7(b)).

Ensure finite execution. The result of a cylindrical strip exploration phase is a distance δ that directs the observer to perform the next cylindrical strip exploration phase on the plane $P_{d+\delta}$. Clearly, the sequence $\sum \delta_i$ of distances resulting from successive phases should never converge, since this would imply that the exploration process would involve performing an infinite number of cylindrical exploration phases.[7] Although Section 4.1 shows that the cylindrical property of the strip can be guaranteed by moving strictly on the motion

[7] Note that this requirement is much stronger than simply requiring δ_i to be non-zero.

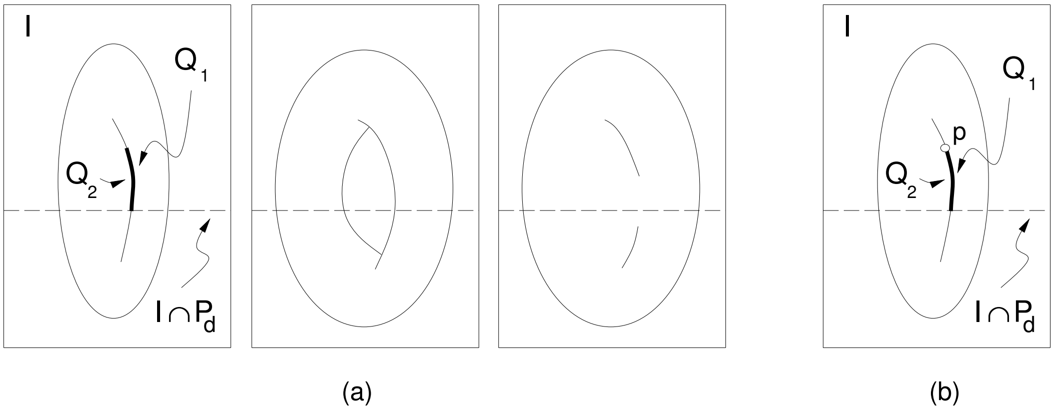

Figure 7: Difficulties encountered during the cylindrical strip exploration phase. (a) The leftmost view shows the projections of the strip-generating segments (shown in bold) above P_d. There are two such segments, Q_1 and Q_2, with Q_1 lying on the visible rim and Q_2 being the intersection of the surface with the conical surface defined by Q_1 and the observer's position (see also Figure 2). Although two strips are traced (i.e., by Q_1 and Q_2), the observer cannot determine from this view whether the surface points belong to a single family $\mathcal{A}$ or not. Hence, the observer cannot determine whether an infinitesimal motion will make visible one or two sets of points in $\mathcal{A}$.Two possible evolutions of the occluding contour are shown in the middle and right views, corresponding to a torus and a bean-shaped surface. (b) Although the observer may be able to determine the distance of Q_1's endpoint, p, from the motion plane, the observer cannot determine whether or not p belongs to family $\mathcal{A}$. Hence, the boundary of the set of visible points in $\mathcal{A}$ above P_d cannot be determined exactly: In this case, the observer assumes that p does belong to $\mathcal{A}$.

plane P_d, it can be shown that the observer must also move above P_d to guarantee this convergence property.

In the following section we focus on the basic component of the cylindrical strip exploration phase, namely a strategy that controls the motion of the observer so that the strip of points corresponding to $\mathcal{A}$ that is explored is guaranteed to be cylindrical.

4.1 Controlling Motion to Enforce the Cylindrical Strip Property

Enforcing the cylindrical property of the set of points explored is the most basic task in the cylindrical strip exploration phase. Our approach is based on the fact that the boundary of the strip on the motion plane must be a closed, connected curve. The observer purposefully controls its motion so that all points on the curve α in $S \cap P_d$ corresponding to the family $\mathcal{A}$ are made visible. Since the surface is assumed generic and the motion plane is not tangent to the surface, this curve is simple, closed and smooth. In order to perform this task, we must be able to answer three questions: (1) which of the visible points on $S \cap P_d$ are contained in α at time t, (2) how to move in order to make more points on α visible, and (3) how to determine when all points on α have been explored?

In this part of our analysis we model the observer as a point. Modelling the observer as a point assumes that the observer has considerable freedom in moving arbitrarily close to the object to be explored (e.g., the observer can go through arbitrarily small holes on the object). In general, this condition is necessary to guarantee the complete exploration of an arbitrary object. Even though this assumption is not realizable in practice, it allows us to focus on the relation between continuous observer motion and the evolution of the exploration frontier without having to worry about the finite dimensions of the observer. In practice, the assumption can be relaxed by employing heuristic strategies that answer the above questions under specific assumptions about the object (see Section 4.1.1).

The first question above is answered by taking advantage of the connectivity of α. From the initial position, the observer can determine a connected set K of visible points belonging to α by looking at the projection $q_{\mathcal{A}}$ of $p_{\mathcal{A}}$ (Figure 8(b)). If some points in α are occluded from the current position, K will be bounded by one or two occlusion boundary points. Similarly, the occlusion boundary segments containing these points bound a connected set of visible points in $\mathcal{A}$. During the cylindrical strip exploration phase the observer can determine which points belong to α by using the fact that α is connected: The observer controls its motion so that it ensures that the connectivity of the points in K is propagated to the points becoming visible on P_d.

The second question is answered by using the algorithm presented in Section 2.1 to circumnavigate α and arrive back at the initial position. There are two ways to move on P_d to achieve this: clockwise or counterclockwise. This corresponds to selecting a *guide point*, g, which is either the left or the right endpoint of K, and controlling the instantaneous direction of motion on P_d so that g becomes visible (Figure 8(a)). Recall that g may be an occlusion boundary point that does not belong to the visible rim (Section 2.1). Here we consider in detail the case where g is a visible rim point; the case where g is an occlusion boundary point not on the visible rim can be treated in a similar fashion.

The third question of determining whether the exploration process is complete corresponds to determining when the circumnavigation of α is complete without having to explicitly represent the set of points on α. A solution is to detect when $c(t)$ crosses the line connecting $c(0)$ and g, and to move toward point $c(0)$ until the point is reached or a collision with the surface becomes imminent. Below we answer the first two questions by focusing on the geometry of the circumnavigation problem and on the necessary observer motion controls.

Suppose that $c(0)$ and $g(0) = g$ satisfy the conditions of Theorem 2.1, and the observer performs an infinitesimal motion in the direction $v(t)$ on P_d in order to make g visible (using the Dynamic Viewpoint Control Algorithm in Section 2.1). The idea of the circumnavigation process comes from the fact that if a new point in the vicinity of $g(0)$ appears on the visible rim at time δt and the conditions of the theorem still hold, the observer can apply the same algorithm to the new point $g(\delta t)$. As the observer continuously applies the algorithm to determine the instantaneous direction of motion for making $g(t)$ visible, both $c(t)$ and $g(t)$ trace curves on P_d.

Theorem 2.1 and the Dynamic Viewpoint Control Algorithm are sufficient to define an algorithm for controlling the observer's motion for any portion of the observer's path for which the conditions of the theorem are true and for which the mapping $f : c(t) \rightarrow g(t)$ is continuous. The motion of the observer during the circumnavigation can therefore be studied and described in terms of a *circumnavigation event*, corresponding to observer positions where the Dynamic Viewpoint Control Algorithm cannot be applied. Each time a circumnavigation event occurs, the observer must take a specific action. Circumnavigation events can be described in two equivalent manners, either by considering the relation between the geometry of $S \cap P_d$ and the point $c(t)$, or by considering the dynamic appearance of the occluding contour during the motion of the observer. The following theorem characterizes these events using the first approach. It also describes the changes in the occluding contour associated with these events when $c(t)$ is a stable viewing position (Figure 8):

Theorem 4.1 (Circumnavigation events) *Circumnavigation events occur iff:*

- *(Event 1) $g(t)$ is an inflection point on α, or*
- *(Event 2) the line segment $L(t)$ connecting $c(t)$ and $g(t)$ is also tangent to a curve in $S \cap P_d$ at a point other than $g(t)$.*

Furthermore, if $c(t)$ is a stable viewing position, the changes in the occluding contour's geometry correspond to:

- *a cusp crossing P_d, or*
- *a T-junction crossing P_d,*

respectively.

Proof. By definition, circumnavigation events occur iff (1) f is not continuous, or (2) the assumptions of Theorem 2.1 do not hold.

Discontinuities of f occur iff the line $L(t)$ connecting $c(t)$ and $g(t)$ is along a direction θ for which the *pedal curve* [23] $w(\theta)$ of the smooth curve α has a cusp at θ. In [23] Giblin and Weiss showed that cusp points on $w(\theta)$ occur iff θ is the direction of an inflectional tangent, i.e., when $L(t)$ is tangent at an inflection point of α, i.e., Event 1 (Figure 8(a) at time t_1).

To see the changes in the occluding contour that correspond to this event, let $g(t)$ be a visible rim point for which $L(t)$ is tangent to α at an inflection point. From Meunsier's Theorem [30], the normal curvature of S along $L(t)$ is 0, i.e., $L(t)$ is along an asymptotic direction of S at $g(t)$.

Hence, if $c(t)$ is a stable position, $g(t)$ is a cusp point on the visible rim [24]; the circumnavigation event in this case corresponds to a cusp point crossing the motion plane[8] (Figure 8(c)).

Now, without loss of generality, suppose that the observer is circumnavigating α in the clockwise direction. The conditions of Theorem 2.1 hold iff $L(t)$ is tangent to $S \cap P_d$ only at $g(t)$. A circumnavigation event therefore occurs when $L(t)$ is tangent to more than one point in $S \cap P_d$. When $c(t)$ is stable, $L(t)$ can be tangent to at most two points and, therefore, $L(t)$ is a bitangent ray. If k is the second point of contact, it must be that α in a neighborhood N of k is convex and that N lies to the left of $L(t)$ (Figure 8(a) at time t_3).

Since $c(t)$ is stable, this circumnavigation event corresponds to a T-junction in the occluding contour with $L(t)$ being the associated bitangent ray [19]. This T-junction must have been either above or below P_d before the event (Figure 8(d)). $\square$

4.1.1 A Circumnavigation Algorithm

The algorithm for circumnavigating α uses the Dynamic Viewpoint Control Algorithm in Section 2.1 for inducing the visibility of surface points when the algorithm is well-defined (i.e., when the position of the observer does not correspond to a circumnavigation event), and specifies the actions taken by the observer in order to handle the two types of circumnavigation events.

The first type of event can be handled easily. When the line $L(t)$ connecting $g(t)$ with $c(t)$ touches an inflection of α, the curve in the neighborhood of $g(t)$ is concave on one side of $L(t)$ and convex on the other. The observer can therefore deduce that the set of points in the concave part of α is connected to the set of points traced by $g(t)$. The circumnavigation process continues by selecting a new point h on the occlusion boundary that bounds the set of visible points of α on the side of the concavity (Figure 8(a,c) at time t_1).

The occurrence of the second type of event poses special problems for the circumnavigation

[8]If the position of the observer is not stable, this circumnavigation event corresponds to a change in the topology of the occluding contour [19]. In particular, since all points in the vicinity of $g(t)$ on α become visible with an infinitesimal change of position, this event can only correspond to a split of the contour at point $g(t)$.

process because the point $g(t)$ can potentially become occluded by an infinitesimal change in the observer's position. In order to avoid the problems caused by this type of event, we control the distance of the observer to $g(t)$ so that such an event never occurs. The idea is to ensure that the observer controls its distance from $g(t)$ so that it passes "between" occlusion boundary points that can cause this event to occur. By also taking taking care of the case where the initial position of the observer corresponds to such an event, the occurrence of this event can be completely ignored.

We now have a simple algorithm for performing the circumnavigation of α. Assume that D and ϵ are fixed positive constants with $\epsilon < 1$.

Step 1: Determine a connected set K of visible points on α contained in $\mathcal{A}$ from the projection of $p_{\mathcal{A}}$.

Step 2: Select a direction for circumnavigating α. (Suppose the clockwise direction is chosen.) Let $g_{occ}(0)$ be the projection of the guide point $g(0)$ (i.e., the left endpoint g of K).

Step 3: (Handling a circumnavigation event of the second type at the initial position.) If there is a visible rim point between $c(0)$ and $g(0)$, move along $L(0)$ toward $g(0)$ until that point is behind the observer.

Step 4: Let $k(t)$ be the visible rim point to the left of $g(t)$ on P_d such that (1) the surface in the vicinity of $k(t)$ is to the left of the line connecting $c(t)$ and $k(t)$, and (2) $k(t)$ is the closest such point, if such a point exists (Figure 8(a) at time t_2).

Step 5: (Continuous control of the direction of motion.) Apply the Dynamic Viewpoint Control Algorithm in Section 2.1 for choosing an instantaneous direction of motion that induces the visibility of the guide point $g(t)$, until either $c(0)$ is reached, a circumnavigation event of the first type occurs, or the following inequalities are not satisfied:

$$d_2(c(t), g(t)) < \epsilon d_2(c(t), k(t)) \tag{1}$$

$$d_2(c(t), g(t)) < D \tag{2}$$

Step 6: (Handling circumnavigation events of the first type.) If $g(t)$ is a point of inflection, define h to be the occlusion boundary point bounding the concavity to the left of $L(t)$. Use h as the new guide point and continue at Step 5.

Step 7: (Avoiding circumnavigation events of the second type through distance adjustments.) If Eq. (1) is not satisfied, move closer to $g(t)$ along $L(t)$ to a position where it is satisfied.

Step 8: (Ensuring the observer does not get arbitrarily far away from α.) If Eq. (2) is not satisfied, move along $L(t)$ to a point where it is satisfied.

This algorithm guarantees that all points on α become visible, but does not necessarily ensure that the strip of points explored on S has non-zero width. However, it is easy to show that this property can be enforced by ensuring that $v(t)$ is never chosen to be along $L(t)$ in Step 5, and by applying the Dynamic Viewpoint Control Algorithm to points in the vicinity of $g(t)$ when a circumnavigation event of the first type occurs.

Finally, note that satisfying Eq. (1) is the only part of the algorithm that may not be implementable if the observer is not a point. The inability to satisfy this inequality implies that the guide point $g(t)$ may inevitably become occluded during the circumnavigation process by the occurrence of a circumnavigation event of the second type. However, if such an event occurs at time t_{ev}, a simple heuristic solution in this case is to circumnavigate the surface by moving on P_d until $c(t)$ crosses the line $L(t_{ev})$. If $g(t)$ is visible, the circumnavigation of α can then be resumed, thereby avoiding the issue of satisfying Eq. (1). This simple heuristic strategy is sufficient to allow the observer to explore cylindrical strips on surfaces such as a torus (Figure 8(e)).

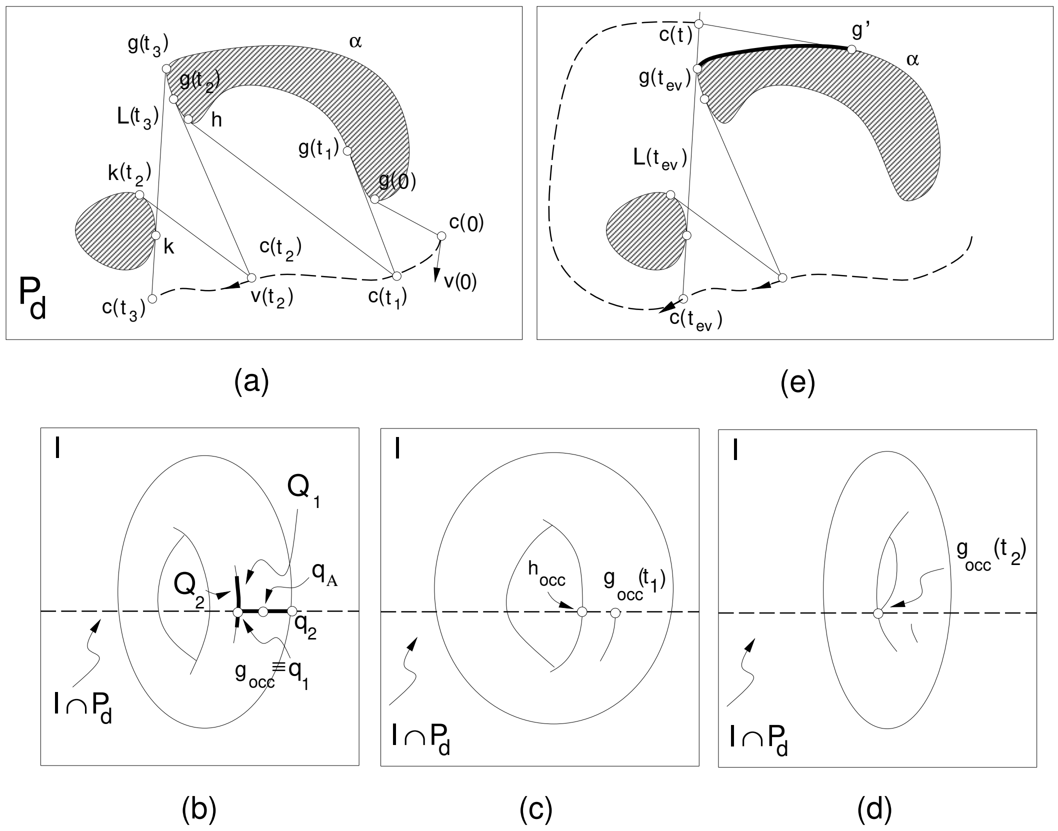

Figure 8: Circumnavigating curve α. The viewed surface a torus with a dent. (a) Circumnavigation events during a circumnavigation path on P_d. Times $0, t_1$ and t_3 correspond to the views of the surface shown in (b),(c) and (d), respectively. The dashed line is a portion of the observer's path. Times $t = 0$ and $t = t_2$ represent typical positions along the observer's path (i.e., where the mapping f is continuous and the conditions of Theorem 2.1 are satisfied). (b) View of the surface when the circumnavigation process is initiated on P_d. There are two strip-generating segments, Q_1 and Q_2 (behind Q_1) whose projections are shown in bold. The dashed line corresponds to the intersection of P_d with the image I (shown here as a plane). The bold horizontal line corresponds to the projection of the points in K. Points q_1, q_2 are the projections of K's endpoints. The guide point g is selected indirectly by selecting its projection g_{occ} on the occluding contour. Here g corresponds to a clockwise circumnavigation of α. The observer assumes that only the segment Q_1, containing g, traces points in $\mathcal{A}$. (c) A circumnavigation event corresponding to a cusp crossing. The point projecting to $g_{occ}(t_1)$ is an inflection point of α. Step 6 of the circumnavigation algorithm selects point h (shown in (a) at time t_1) as the new guide point by selecting its projection h_{occ} on the occluding contour. (d) A circumnavigation event corresponding to a T-junction crossing. The line $L(t_3)$ connecting $c(t_3)$ and $g(t_3)$ is a bitangent ray. (e) Circumnavigation for observers with non-zero dimensions. At time t_{ev} a circumnavigation event of the second type occurs. The observer circumnavigates the occluding curve on P_d until the line $L(t_{ev})$ is crossed. Since $g(t_{ev})$ is visible, $g(t_{ev})$ defines a connected set K' of visible points on α that are connected to $p_{\mathcal{A}}$ (shown in bold). Hence the circumnavigation process can continue by using the endpoint g' of K' as the new guide point.

5 Concluding Remarks

This paper has shown an important link between (1) the problem of inducing the visibility of occluded points on a smooth object surface, and (2) the ability to purposefully control the observer's viewpoint in a continuous fashion by directed observer motion. We have demonstrated that this connection leads to a provably-correct strategy for exploring objects bounded by arbitrary smooth surfaces. Furthermore, this strategy uses quickly-computable quantities derived from the occluding contour (distance to a point) and requires that only a qualitative representation of the surface be maintained during the exploration process.

The use of an active observer that *purposefully* and *dynamically* controls viewpoint is the most crucial aspect of our approach. This ability allows the observer to impose structure on the set of points becoming visible, making it possible to direct the exploration process using quickly-computable visual information in an image and a qualitative object surface representation. Moreover, our approach demonstrates that object exploration does not require the recovery of a detailed three-dimensional description of the object being explored from every viewpoint. The reason for this is that continuous, directed observer motion allows the evolution of the object appearance to be locally predictable in time, thus eliminating the need for detailed shape descriptions to guide the process. This is a major step towards qualitative, active vision, allowing provably-correct, exploratory behaviors to control viewpoint based on qualitative shape information.

Although the approach presented here for solving the object exploration problem is theoretical, our results have important practical consequences. For example, the simple viewing strategy presented for exploring cylindrical strips of an object surface can be easily implemented as a front-end motion control module for other procedures that recover surface curvature information from the deformation of the occluding contour [20, 27]. We are currently in the process of implementing this exploration strategy for exploring and recovering the shape of cylindrical strips on synthetic object surfaces. At a theoretical level, we are also investigating strip exploration strategies that guarantee the convergence of the exploration process.

References

[1] D. Marr, *Vision.* Freeman, 1982.

[2] P. Whaite and F. P. Ferrie, "From uncertainty to visual exploration," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 10, pp. 1038–1049, 1991.

[3] S. A. Hutchinson and A. Kak, "Planning sensing strategies in a robot work cell with multi-sensor capabilities," *IEEE Trans. Robotics Automat.*, vol. 5, no. 6, pp. 765–783, 1989.

[4] J. Maver and R. Bajcsy, "Occlusions and the next view planning," in *Proc. IEEE Robotics Automat. Conf.*, pp. 1806–1811, 1992.

[5] Y. Aloimonos, "Purposive and qualitative active vision," in *Proc. Int. Conf. on Pattern Recognition*, pp. 346–360, 1990.

[6] K. N. Kutulakos and C. R. Dyer, "Recovering shape by purposive viewpoint adjustment," in *Proc. Computer Vision and Pattern Recognition*, pp. 16–22, 1992.

[7] S. Hutchinson, "Exploiting visual constraints in robot motion planning," in *Proc. IEEE Robotics Automat. Conf.*, pp. 1722–1727, 1991.

[8] R. Bajcsy and M. Campos, "Active and exploratory perception," *CVGIP: Image Understanding*, vol. 56, no. 1, pp. 31–40, 1992.

[9] D. H. Ballard and C. M. Brown, "Principles of animate vision," *CVGIP: Image Understanding*, vol. 56, no. 1, pp. 3–21, 1992. Special Issue on Purposive, Qualitative, Active Vision.

[10] J. S. Bay, "A fully autonomous active sensor-based exploration concept for shape-sensing robots," *IEEE Trans. Syst. Man Cybern.*, vol. 21, no. 4, pp. 850–860, 1991.

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

[12] C. Laugier, A. Ijel, and J. Troccaz, "Combining vision based information and partial geometric models in automatic grasping," in *Proc. IEEE Robotics Automat. Conf.*, pp. 676–682, 1990.

[13] J. Y. Zheng *et al.*, "Active camera controlling for manipulation," in *Proc. Computer Vision and Pattern Recognition*, pp. 413–418, 1991.

[14] M. Seibert and A. M. Waxman, "Adaptive 3-d object recognition from multiple views," *IEEE Trans. Pattern Anal. and Machine Intell.*, vol. 14, no. 2, pp. 107–124, 1992.

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

[16] G. Hager and M. Mintz, "Searching for information," in *Proc. Workshop on Spatial Reasoning and Multi-Sensor Fusion*, pp. 313–322, 1987.

[17] S. Lee and H. Hahn, "An optimal sensing strategy for recognition and localization of 3-d natural quadric objects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, no. 10, pp. 1018–1037, 1991.

[18] M. J. Swain and M. Stricker, eds., *Promising Directions in Active Vision*, University of Chicago, November 1991. Written by the attendees of the NSF Active Vision Workshop.

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

[20] R. Cipolla and A. Blake, "The dynamic analysis of apparent contours," in *Proc. 3rd Int. Conf. on Computer Vision*, pp. 616–623, 1990.

[21] W. B. Seales and C. R. Dyer, "Viewpoint from occluding contour," *CVGIP: Image Understanding*, vol. 55, no. 2, pp. 198–211, 1992.

[22] J. J. Gibson, *The Ecological Approach to Visual Perception.* Houghton Mifflin, 1979.

[23] P. Giblin and R. Weiss, "Reconstruction of surfaces from profiles," in *Proc. 1st Int. Conf. on Computer Vision*, pp. 136–144, 1987.

[24] J. J. Koenderink, "An internal representation for solid shape based on the topological properties of the apparent contour," in *Image Understanding 1985-86* (W. Richards and S. Ullman, eds.), pp. 257–285, Norwood, NJ: Ablex Publishing Co., 1987.

[25] R. Vaillant and O. D. Faugeras, "Using extremal boundaries for 3-d object modeling," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 14, no. 2, pp. 157–173, 1992.

[26] R. Basri and S. Ullman, "The alignment of objects with smooth surfaces," in *Proc. 2nd Int. Conf. on Computer Vision*, pp. 482–488, 1988.

[27] A. Blake, M. Brady, R. Cipolla, Z. Xie, and A. Zisserman, "Visual navigation around curved obstacles," in *Proc. IEEE Robotics Automat. Conf.*, pp. 2490–2495, 1991.

[28] V. J. Lumelsky, "A comparative study on the path length performance of maze-searching and robot motion planning algorithms," *IEEE Trans. Robotics Automat.*, vol. 7, no. 1, pp. 57–66, 1991.

[29] J. J. Koenderink and A. J. van Doorn, "A description of the structure of visual images in terms of an ordered hierarchy of light and dark blobs," in *Proc. Second Int. Visual Psychophysics and Medical Imaging Conf.*, pp. 173–176, 1981.

[30] M. P. D. Carmo, *Differential Geometry of Curves and Surfaces.* Englewood Cliffs, NJ: Prentice-Hall Inc., 1976.

[31] L. R. Nackman, "Two-dimensional critical point configuration graphs," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 6, no. 4, pp. 442–450, 1984.