3D Surface Reconstruction Using Occluding Contours

EDMOND BOYER AND MARIE-ODILE BERGER

CRIN–CNRS/INRIA *Lorraine, BP 239, 54506 Vandœuvre les Nancy Cedex, France*

boyere@loria.fr

berger@loria.fr

Abstract. This paper addresses the problem of 3D surface reconstruction using image sequences. It has been shown that shape recovery from three or more occluding contours of the surface is possible given a known camera motion. Several algorithms, which have been recently proposed, allow such a reconstruction under the assumption of a linear camera motion. A new approach is presented which deals with the reconstruction problem directly from a discrete point of view. First, a theoretical study of the epipolar correspondence between occluding contours is achieved. A correct depth formulation is then derived from a local approximation of the surface up to order two. This allows the local shape to be estimated, given three consecutive contours, without any constraints on the camera motion. Experimental results are presented for both synthetic and real data.

1. Introduction

Recovering three-dimensional shapes from visual data is an important intermediate goal of many vision systems used in robotics, surveillance, guidance and modelling. In such applications, surface reconstruction is often used as an intermediate step in the more important task of 3D representation and recognition.

For non-polyhedral objects, rich and robust information on the shape are provided by the occluding contours. In fact, if some strong *a priori* knowledge on the object is available such as parametric descriptions, then a single view allows shape recovery (Ponce et al., 1989; Zerroug and Nevatia, 1993), object pose estimation and object recognition (Kriegman and Ponce, 1990; Glachet et al., 1992; Forsyth et al., 1992). Otherwise, for any smooth object, a sequence of occluding contours must be considered to recover the shape.

The problem of reconstructing surfaces from occluding contours was first considered by Koenderink (1984), who attempted to infer 3D curvature properties from occluding contours. Later, Giblin and Weiss (1987) proved that surface reconstruction can be

performed for planar motion under the assumption of orthographic projection. The problem of reconstructing under perspective projection for general known motions has been tackled by Cipolla and Blake (1990), and Vaillant and Faugeras (1992). Both approaches lead to depth computation under the assumption of continuous observations of the occluding contours and are based on differential analysis. Unfortunately, the use of differential computations with discrete observations requires numerical approximations and produces numerical instabilities. With the aim of further constraining the problem, the occluding contour is supposed to be locally part of a circle. In addition, the plane containing the circle is supposed to be known: in (Cipolla and Blake, 1990), the epipolar plane is used whereas Vaillant and Faugeras (1992) used the radial plane. While keeping the circle model, Szeliski and Weiss (1993) proposed to improve the reconstruction process by computing the epipolar curves on the whole surface together with an estimate of uncertainty. The use of estimation theory leads to a more robust shape recovery using a linear smoother but the basic reconstruction method remains unchanged.

The approaches described above only allow a 3D depth map of points or curves belonging to the surface to be obtained. In the work of Zhao and Mohr (1994), the authors attempt to recover the global 3D surface description from the occluding contours in a single stage by use of B-spline patches. This approach introduces a direct regularisation of the reconstructed surface, but it requires a complete *a priori* parametrisation of the surface which is usually not available. Therefore it is preferable to first recover a more accurate 3D depth map before computing a 3D surface description.

The task considered in this paper is to deal with the reconstruction problem directly from a discrete point of view. We first address the problem of the mapping between successive occluding contours and we give new results in the case of the epipolar correspondence. Then, a correct depth formulation is given when the occluding contours are observed at discrete times. The main interest of such a formulation is to avoid reconstruction constraints—unlike (Cipolla and Blake, 1990; Vaillant and Faugeras, 1992; Szeliski and Weiss, 1993; Joshi et al., 1995) no reconstruction plane is locally imposed—so that non-linear motion can be easily taken into account. Moreover, it gives a general solution to the reconstruction problem which is always defined except when the camera motion is in the viewing direction. Our approach is based on a local approximation of the surface up to the second order, allowing a linear estimation of depth to be derived, given any set of three occluding contours.

This paper is organised as follows: the epipolar correspondence induced between two discrete observations is studied in Section 2. In Section 3, the depth formulation is established, enabling the discrete observations of occluding contours to be handled. In Section 4, our approach is compared to previous methods. The algorithm performances are studied in Section 5 and significant reconstruction results on real data are shown in Section 6, before concluding.

2. Viewing Geometry

Let $\mathcal{S}$ be a smooth curved surface: $\mathcal{S}$ is at least C^2 and not locally plane. Let also P be a point on $\mathcal{S}$. We assume that the imaging system is based on the pinhole model (i.e., perspective projection). Therefore, the vector position r of P can be written

$$r = C + \lambda T, \qquad (1)$$

where C is the camera centre position, T the unit viewing direction and λ the depth of the point P along the
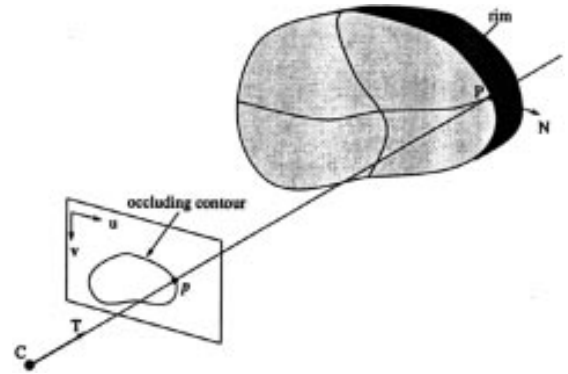

Figure 1.   Rim and occluding contour under perspective projection.

viewing direction. For a given camera position there is a locus of points on the surface $\mathcal{S}$ where the normal N is perpendicular to the viewing direction. This set of points is called the *rim* (also known as the *limb*) and its projection onto the image plane is called the *occluding contour* (also known as the *extremal boundary* or the *profile*), as shown in Fig. 1.

As the camera moves around $\mathcal{S}$ the occluding contours sweep a surface in the space of parameters which is called the *spatio-temporal surface* (Faugeras, 1993; Giblin and Weiss, 1995). Since the camera centre position is a function of time t, this surface can be parameterised by (s, t), where the parameter s describes the position on the occluding contours and the parameter t corresponds to time. However, such a parametrisation is not uniquely defined (Cipolla and Blake, 1990): curves of constant t are the occluding contours but curves of constant s have no physical interpretation. Until now, the most generally accepted parametrisation of the spatio-temporal surface is the *epipolar parametrisation* which yields a mapping between successive occluding contours called the *epipolar correspondence*. The advantage of this parametrisation is that it leads to a local parametrisation of the surface $\mathcal{S}$ which can always be used except when the occluding contour is singular or when the camera motion is in the plane tangent to the surface (Giblin and Weiss, 1995). Furthermore, the epipolar correspondence between points on successive occluding contours constrains the reconstruction problem which becomes a linear estimation, as shall be seen later.

2.1. Epipolar Correspondence

The epipolar parametrisation is derived from the epipolar geometry in stereo-vision. It leads to a planar
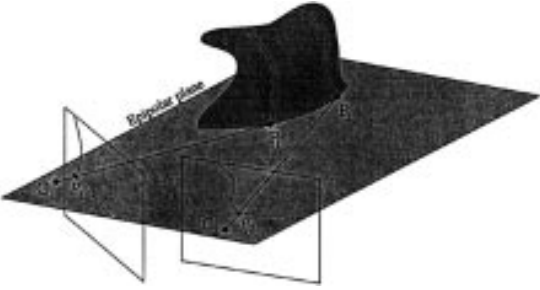
Figure 2.   Two epipolar correspondents p_1 and p_2.

correspondence between two successive occluding contours as shown in Fig. 2. Hence, two points p_1 and p_2 on two successive occluding contours are epipolar correspondents if they both belong to the epipolar plane defined by the two camera centres and one of the points. Such correspondences between points on the spatio-temporal surface generally induce correspondences between points on the surface $\mathcal{S}$. For a continuous motion of the camera, curves of constant s on the spatio-temporal surface induce curves on the surface $\mathcal{S}$. These curves are called *epipolar curves*. In this section we focus on the local behaviour of the epipolar correspondence between two occluding contours of $\mathcal{S}$.

Consider two occluding contours $\mathcal{O}_1$ and $\mathcal{O}_2$ of $\mathcal{S}$ for two successive camera centre positions C_1 and C_2. We denote by $T_1(s_1)$ the viewing direction for points p_1 on $\mathcal{O}_1$ where the parameter s_1 describes the position on $\mathcal{O}_1$, and equivalently by $T_2(s_2)$ the viewing direction for points p_2 on $\mathcal{O}_2$ where the parameter s_2 describes the position on $\mathcal{O}_2$. The epipolar correspondence leads to the following matching constraint between two image points p_1 and p_2 on $\mathcal{O}_1$ and $\mathcal{O}_2$:

$$F(s_1, s_2)_{\mid (p_1, p_2)} = 0, \qquad (2)$$

with:

$$F(s_1, s_2) = (T_1(s_1) \wedge T_2(s_2)) \cdot (C_1 - C_2).$$

In the general case, epipolar correspondents p_1 and p_2 are image projections of two different points P_1 and P_2 on $\mathcal{S}$. This *non-stationarity* property of occluding contours can be used to discriminate such contours from others (Vaillant and Faugeras, 1992). However, there is a circumstance where the property is not verified. This occurs when the camera motion $(C_1 - C_2)$ is in the tangent plane of $\mathcal{S}$ at P_1 or P_2. In this situation, points p_1 and p_2 are image projections of the same fixed point P ($P = P_1 = P_2$), and rims of $\mathcal{S}$ for

camera positions C_1 and C_2 intersect at P. We call such a point P a *multiple point* of the sequence considered.

Definition.   A multiple point of an image sequence of $\mathcal{S}$ is a point P where two or more consecutive rims of the sequence intersect.

Consequently, if P is a multiple point of an image sequence of $\mathcal{S}$ then the epipolar plane at P is tangent to $\mathcal{S}$. In addition, if the n camera centre positions of the image sequence are aligned, then P is of *multiplicity n*.

Remark.   In the literature, the locus of rim points where the epipolar plane is tangent to the surface is called *the frontier* (Giblin and Weiss, 1995; Cipolla et al., 1995). For a linear camera motion, the frontier, if it exists, is restricted to isolated points. Thus, multiple points represent the frontier for linear camera motions going through successive camera centre positions of the sequence. Since these rim points are fixed features on the surface, they can be used to derive constraints on the camera motion (Rieger, 1986; Porill and Pollard, 1991; Cipolla et al., 1995; Joshi et al., 1995).

Equation (2) relates parameters s_1 and s_2 of two epipolar correspondents. Now consider the epipolar map:

$$f \colon \mathcal{O}_1 \to \mathcal{O}_2$$

which associates with every point $p_1 \in \mathcal{O}_1$ its epipolar correspondent $p_2 \in \mathcal{O}_2$ such that $F(s_1, s_2)_{\mid p_1, p_2} = 0$. Assuming that the occluding contours are smooth[1], we have:

Proposition 1.   *The epipolar map f is a local diffeomorphism at p_1 if the point P_1 on $\mathcal{S}$, with image projection p_1, is not a multiple point of the sequence considered.*

Proof: This proposition is a direct consequence of the implicit function theorem (Bruce and Giblin, 1984) applied to the function F. We denote by s_{p_1} and s_{p_2} the parameters of two epipolar correspondents p_1 and p_2 on $\mathcal{O}_1$ and $\mathcal{O}_2$, therefore $F(s_{p_1}, s_{p_2}) = 0$. If the occluding contours are smooth, then $T_1(s_1)$, $T_2(s_2)$ are smooth with respect to s_1 and s_2 and the function $F(s_1, s_2)$ is smooth. By the implicit function theorem, there exists a continuous map $f \colon s_1 \mapsto s_2$ which is smooth in a neighbourhood of s_{p_1} provided $\frac{\partial F}{\partial s_2}(s_{p_1}, s_{p_2}) \neq 0$. Likewise s_1 is a smooth function of s_2 in a neighbourhood of s_{p_2} provided $\frac{\partial F}{\partial s_1}(s_{p_1}, s_{p_2}) \neq 0$. Rewriting (2)
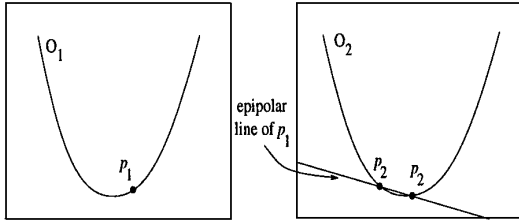
Figure 3. The epipolar correspondence is ambiguous in the proximity of a multiple point: p_1 has two close correspondents p_2.

and taking the derivatives with respect to s_1 and s_2, we obtain:

$$\begin{cases} \frac{\partial F}{\partial s_2}(s_{p_1}, s_{p_2}) = ((C_1 - C_2) \wedge T_1(s_{p_1})) \cdot \frac{\partial T_2}{\partial s_2}(s_{p_2}), \\ \frac{\partial F}{\partial s_1}(s_{p_1}, s_{p_2}) = ((C_2 - C_1) \wedge T_2(s_{p_2})) \cdot \frac{\partial T_1}{\partial s_1}(s_{p_1}). \end{cases}$$

The epipolar plane at p_2 is spanned by $T_2(s_{p_2})$ and $(C_2 - C_1)$, or equivalently by $T_1(s_{p_1})$ and $(C_1 - C_2)$. Thus $\frac{\partial F}{\partial s_2}(s_{p_1}, s_{p_2}) = 0$ if $\frac{\partial T_2}{\partial s_2}(s_{p_2})$ belongs to the epipolar plane at p_2. Since $T_2(s_{p_2})$ and $\frac{\partial T_2}{\partial s_2}(s_{p_2})$ both belong to the tangent plane of $\mathcal{S}$ at P_2, this occurs at points p_1 and p_2 where the epipolar plane is tangent to the surface. In this case, P_2 is a multiple point and $P_1 = P_2$. Likewise $\frac{\partial F}{\partial s_1}(s_{p_1}, s_{p_2}) = 0$ if the epipolar plane is tangent to the surface and thus $P_2 = P_1$. The map $f: s_1 \mapsto s_2$ is therefore a local diffeomorphism at s_{p_1} if the point p_1 is not the projection of a multiple point.   □

Proposition 1 says that given two epipolar correspondents p_1 and p_2, there exist neighbourhoods of p_1 and p_2 where the epipolar map is smooth and has a smooth inverse, provided that p_1 and p_2 are not projections of a multiple point. In fact, in the neighbourhood of the projection of a multiple point, the epipolar map yields two correspondents, as shown in Fig. 3.

The epipolar correspondence is thus ambiguous in the proximity of a multiple point. The criterion one can use to resolve this ambiguity is that the images of P_1, P_2 and the multiple point must appear in the same order along occluding contours (this assumes that occluding contours have the same orientation). Note that this criterion is similar to the ordering constraint in stereo-vision.

2.2. Recovering Depth along the View Line

Our goal is to recover rims from their projections. First, we suppose that under a continuous motion of the camera a complete description of the spatio-temporal surface is available. Then, (1) can be derived according to time t and by taking the scalar product with the normal N to the surface, we obtain:

$$\lambda = \frac{-\frac{\partial C}{\partial t} \cdot N}{\frac{\partial T}{\partial t} \cdot N}. \tag{3}$$

This is the depth formula for points on rims as defined in (Cipolla and Blake, 1990). Assuming that we have a parametrisation of $\mathcal{S}$, we can therefore compute $\frac{\partial T}{\partial t}$ from the spatio-temporal surface and then recover the depth for points on the spatio-temporal surface where $\frac{\partial T}{\partial t} \cdot N \neq 0$. Unfortunately, only discrete information (occluding contours at discrete times $t_i, i \in [1, \ldots, m]$) are available and (3) can not be applied directly. In fact, depth can be computed only by approximation. In this section, we discuss such an approximation.

Now consider two successive occluding contours at times t_1 and t_2. Let P_1 and P_2 be two points on the rims of $\mathcal{S}$ at t_1 and t_2. Using the notations of the previous section, we have:

$$\begin{cases} r_1 = C_1 + \lambda_1 T_1 \text{ at } P_1, \\ r_2 = C_2 + \lambda_2 T_2 \text{ at } P_2. \end{cases}$$

Thus, denoting by Δx the difference $\Delta x = x_1 - x_2$, where x is any of r, C, λ and T, gives:

$$\Delta r = \Delta C + \lambda_1 \Delta T + \Delta \lambda T_2,$$

and by taking the scalar product with the normal N_2 to the surface at P_2, we obtain the following depth formula for point P_1:

$$\lambda_1 = \frac{-\Delta C \cdot N_2}{\Delta T \cdot N_2} + \frac{\Delta r \cdot N_2}{\Delta T \cdot N_2}. \tag{4}$$

If the image projections p_1 and p_2 of P_1 and P_2 are matched according to the epipolar correspondence, then the first term of the above expression corresponds to a triangulation formula. Indeed, $\frac{-\Delta C \cdot N_2}{\Delta T \cdot N_2}$ is the distance from the camera centre C_1 to the viewlines intersection (see Fig. 4). This value is therefore the depth of a virtual point with image projection p_1 and p_2. Hence, it can be computed from measurements in two images by using a stereo formula. On the other hand, the second term of (4) depends on the distance Δr between surface points P_1 and P_2. This value can not be computed, *a priori*, from measurements in two images. A first approach would consist in omitting this term in (4).
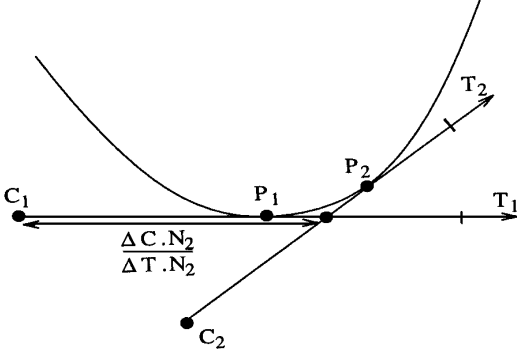
Figure 4. Intersection of tangents in the epipolar plane.

We could then recover the depth of a surface point with only two images. But this approach leads to a stereo reconstruction and implicitly assumes that rims are not view dependent which is of course false. For a smooth surface which is not locally plane, $\Delta r \cdot N_2 \neq 0$ except at a multiple point. Therefore, the second term of (4) should not be omitted when computing depth.

The approach we have developed is based on a local surface model: a second order approximation. Such a model allows $\Delta r \cdot N_2$ to be expressed as a function of local properties of $\mathcal{S}$. Hence, by using more than two images (three for a second order approximation), we can fit locally the surface model to the image measurements and estimate both depth and local properties. This idea is developed in the following section.

3. Reconstruction

We have seen in the previous section that more than two images are needed to estimate the local properties of the surface $\mathcal{S}$. We show here that by using three images, a local approximation up to order 2 of $\mathcal{S}$ can be computed. Moreover, if points between occluding contours of $\mathcal{S}$ are matched according to the epipolar correspondence, this approximation leads to a linear estimation of the position and the curvatures of $\mathcal{S}$ at a point P. Unlike previous methods (Cipolla and Blake, 1990; Vaillant and Faugeras, 1992; Szeliski and Weiss, 1993; Joshi et al., 1995) which use an *a priori* plane in order to estimate the epipolar curve, no assumption is made on the camera motion or the local surface shape. Instead, a local parametrisation of the surface $\mathcal{S}$ is used which in turn leads to a local surface approximation. We first present the local parametrisation (x, y) that is used and the induced local approximation of the surface. We show then that such a parametrisation

allows linear equations to be derived for both depth and normal curvature in the viewing direction.

3.1. The Osculating Quadric

Since $\mathcal{S}$ is a smooth surface, a neighbourhood of a point P on $\mathcal{S}$ can be represented in the form $z = h(x, y)$, where P is the origin of the coordinate frame and the z axis is directed by the normal N of $\mathcal{S}$ at P. Thus the xy plane is the tangent plane to $\mathcal{S}$ at P. Moreover, h is a differentiable function and by taking Taylor's expansion at P, we have:

$$h(x, y) = z$$
$$= \frac{1}{2}(h_{xx}x^2 + 2h_{xy}xy + h_{yy}y^2) + R(x, y),$$

where $R(x, y)$ satisfies:

$$\lim_{(x,y) \to (0,0)} \frac{R(x, y)}{x^2 + y^2} = 0.$$

The quadratic form $\frac{1}{2}(h_{xx}x^2 + 2h_{xy}xy + h_{yy}y^2)$ is known as the Hessian of h at $(0, 0)$ and corresponds to the second fundamental form of $\mathcal{S}$ at P. The quadratic surface $\mathcal{Q}$ defined by

$$\mathcal{Q} = \left\{(x, y, z), z = \frac{1}{2}(h_{xx}x^2 + 2h_{xy}xy + h_{yy}y^2)\right\}$$

approximates $\mathcal{S}$ up to order 2. This surface is called *the osculating quadric* of $\mathcal{S}$ at P and is uniquely defined.

We first assume that the x and y axes are oriented by T and $r_s/|r_s|$ respectively, where T is the viewing direction and r_s the tangent to the rim of $\mathcal{S}$ at P. Since these directions are conjugate (Koenderink, 1984), the first and second fundamental forms[2] of $\mathcal{S}$ at P, in the parametrisation (x, y), are:

$$I_P = \begin{bmatrix} 1 & \cos\theta \\ \cos\theta & 1 \end{bmatrix} \quad II_P = \begin{bmatrix} k_t & 0 \\ 0 & k_s \end{bmatrix},$$

where θ is the angle between T and r_s, k_t is the normal curvature along the viewing direction T, and k_s is the normal curvature of the rim at P. Note that k_t is the normal curvature of the epipolar curve at P. Therefore, the osculating quadric of $\mathcal{S}$ at P is defined, in the parametrisation (x, y), by:

$$\mathcal{Q} = \left\{(x, y, z), z = \frac{1}{2}(k_t x^2 + k_s y^2)\right\}. \quad (5)$$

Figure 5. Epipolar curves.

We denote by $\mathcal{E}_{-1}$ and $\mathcal{E}_1$ the epipolar planes at a point $P \in \mathcal{S}$ corresponding to three successive positions of the camera centre C_{-1}, C and C_1, i.e., planes (C, C_{-1}, P) and (C, C_1, P) (see Fig. 5). The point P is therefore a point belonging to the rim observed from C. For a general motion $\mathcal{E}_{-1}$ and $\mathcal{E}_1$ are different (they are identical for a linear motion). The intersection of one epipolar plane with the surface $\mathcal{S}$ is a curve. By abuse of language, we call these curves epipolar curves[3] and we have then the following property:

Proposition 2. *In any epipolar plane $\mathcal{E}$ at P, the epipolar curve is, up to the order 2, the graph of the following function:*

$$z_{\mathcal{E}} = g(x) = \frac{1}{2}\frac{k_t}{\cos \beta_{\mathcal{E}}} x^2, \qquad (6)$$

where the x axis is directed by $T|_P$, the $z_{\mathcal{E}}$ axis is such that $(x, z_{\mathcal{E}})$ form an orthonormal basis of $\mathcal{E}$ and $\beta_{\mathcal{E}}$ is the angle between the normal N to the surface at P and the projection $N_{\mathcal{E}}$ of N in the epipolar plane: $\cos \beta_{\mathcal{E}} = N \cdot N_{\mathcal{E}}$.

Proof: Near the point P, we have the following description of $\mathcal{S}$:

$$z = \frac{1}{2}(k_t \, x^2 + k_s \, y^2) + R(x, y), \qquad (7)$$

with $\lim_{(x,y)\to(0,0)} R(x, y)/(x^2 + y^2) = 0$. Let $\mathcal{E}$ be an epipolar plane at P and let $\beta_{\mathcal{E}}$ be the angle between the normal N to $\mathcal{S}$ at P and its projection $N_{\mathcal{E}}$ in $\mathcal{E}$. The equation of this plane can be written as:

$$z \sin \beta_{\mathcal{E}} = y \cos \beta_{\mathcal{E}} \sin \theta,$$

where θ is the angle between the x axis and the y axis (or equivalently between T and r_s at P). Now if the $z_{\mathcal{E}}$

axis is defined such that $(x, z_{\mathcal{E}})$ form an orthonormal basis of $\mathcal{E}$, we have:

$$z = z_{\mathcal{E}} \cos \beta_{\mathcal{E}}.$$

The epipolar curve is the intersection of the plane $\mathcal{E}$ with the surface $\mathcal{S}$. Therefore by substituting in (7) and neglecting third order and higher terms we obtain for points on these curves:

$$z_{\mathcal{E}} \cos \beta_{\mathcal{E}} = \frac{1}{2}\left(k_t x^2 + \frac{\sin^2 \beta_{\mathcal{E}}}{\sin^2 \theta} k_s z_{\mathcal{E}}^2\right). \qquad (8)$$

In general, the term $\frac{\sin^2 \beta_{\mathcal{E}}}{\sin^2 \theta} k_s$ in (8) can be considered as bounded since:

1. The case $\sin \theta = 0$ happens only when P is observed along an asymptotic direction which yields a cusp on the occluding contour.
2. The curvature k_s, which is linked to the curvature of the occluding contour, is finite in our context (the observed object is smooth).

Thus, solving (8) for $(x, z_{\mathcal{E}})$ close to $(0, 0)$ yields:

$$\begin{cases} z_{\mathcal{E}} = \frac{1}{2}\frac{k_t}{\cos \beta_{\mathcal{E}}} x^2 + R_{\mathcal{E}}(x), & \cos \beta_{\mathcal{E}} \neq 0, \\ z_{\mathcal{E}} = x = 0, & \cos \beta_{\mathcal{E}} = 0, \end{cases}$$

with: $\lim_{x \to 0} R_{\mathcal{E}}(x)/x^2 = 0$. This shows that up to second order, the epipolar curve is represented by:

$$z_{\mathcal{E}} = \frac{1}{2}\frac{k_t}{\cos \beta_{\mathcal{E}}} x^2, \qquad \cos \beta_{\mathcal{E}} \neq 0.$$

The case $\cos \beta_{\mathcal{E}} = 0$ occurs when P is a multiple point, the epipolar curve is then restricted to a single point P.   □

Remark. The approximation given in Proposition 2 verifies Meusnier's theorem (do Carmo, 1976) which says that the curvature at P of the epipolar curve is $k = \frac{k_t}{\cos \beta_{\mathcal{E}}}$. However, it should be noticed that a local approximation of the epipolar curve based on a circle verifies also Meusnier's theorem. But such an approximation implicitly implies that the surface is locally spherical which is less general than the osculating quadric model.

Note that in the above proposition, the x axis is the one previously defined in the parametrisation (x, y) and is thus independent of the epipolar plane.

Since P is the origin of the x axis, Proposition 2 says that the epipolar curve depends on the position of P in the epipolar plane (i.e., its depth) and on the normal curvature k_t of $\mathcal{S}$ along the viewing direction. Our purpose is to recover the position of a rim point P using three successive occluding contours of $\mathcal{S}$, therefore this can be done by estimating epipolar curves. In the general case there are two different epipolar planes $\mathcal{E}_{-1}$ and $\mathcal{E}_1$ for a point P and three successive camera positions, thus there are also two epipolar curves. Since we can match, in the corresponding images, epipolar correspondents, we know two tangents to each epipolar curve (see Fig. 5). The following section shows how to compute epipolar curves given these tangents.

3.2.   Estimation of the Epipolar Curves

Our goal is to recover the depth λ and the curvature k_t at P. It has been shown that these values are related to the position and the curvature of the epipolar curves. Therefore, the problem is to estimate two of these curves given three tangents. The previous section showed that the epipolar curves are, up to order 2, parabolas of the epipolar planes. Moreover, although the epipolar curves are not in the same plane, by using Proposition 2, linear estimations of both λ and k_t can be derived.

Now consider the epipolar plane $\mathcal{E}_1$. We denote by $\beta_{\mathcal{E}_1}$ the angle between the normal N at P and its projection $N_{\mathcal{E}_1}$ in $\mathcal{E}_1$:

$$\cos \beta_{\mathcal{E}_1} = N \cdot N_{\mathcal{E}_1}.$$

By Proposition 2, the epipolar curve with tangents T and T_1 is defined by (up to order two):

$$z_{\mathcal{E}_1} = \frac{1}{2}k_{\mathcal{E}_1}x^2, \qquad k_{\mathcal{E}_1} = \frac{k_t}{\cos \beta_{\mathcal{E}_1}},$$

where $(x, z_{\mathcal{E}_1})$ form an orthogonal basis of $\mathcal{E}_1$. Let x_{P_1} be the abscissa of P_1 as shown in Fig. 6.

Since P_1 belongs to the epipolar curve, the tangent to $\mathcal{S}$ at P_1 in the direction T_1 goes through the point $(\frac{x_{P_1}}{2}, 0)$ and it follows that:

$$\begin{cases} T_1 \cdot N_{\mathcal{E}_1} = \dfrac{k_{\mathcal{E}_1}\,x_{P_1}^2/2}{\sqrt{\left(k_{\mathcal{E}_1}\,x_{P_1}^2/2\right)^2 + \left(x_{P_1}/2\right)^2}}, \\ \mathrm{sign}(x_{P_1}) = -\mathrm{sign}(T_1 \cdot N_{\mathcal{E}_1}), \end{cases}$$

thus:

$$x_{P_1} = \frac{1}{k_{\mathcal{E}_1}}\frac{-T_1 \cdot N_{\mathcal{E}_1}}{\sqrt{1 - \left(T_1 \cdot N_{\mathcal{E}_1}\right)^2}}.$$
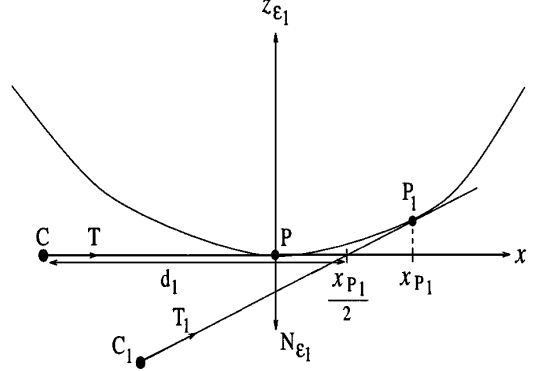

Figure 6.   The epipolar plane $\mathcal{E}_1$.

Since $k_t = k_{\mathcal{E}_1} \cos \beta_{\mathcal{E}_1}$, the above expression can be rewritten as:

$$x_{P_1} = \frac{\cos \beta_{\mathcal{E}_1}}{k_t}\frac{-T_1 \cdot N_{\mathcal{E}_1}}{\sqrt{1 - \left(T_1 \cdot N_{\mathcal{E}_1}\right)^2}}. \qquad (9)$$

T_1, $N_{\mathcal{E}_1}$ and $\cos \beta_{\mathcal{E}_1}$ can be computed from image measurements. Thus, we can determine the distance $\frac{x_{P_1}}{2}$ between the intersection of the tangents and P given the curvature k_t. In addition, it was shown (Section 2.2) that the distance between the camera centre C and the intersection of the tangents can also be computed from image measurements. Hence, if this distance is denoted by d_1, we have the following relation:

$$\lambda = d_1 - \frac{1}{2}x_{P_1}, \qquad (10)$$

which allows the depth at P to be computed given the normal curvature along the viewing direction.

Likewise, since T is also tangent to the second epipolar curve in the epipolar plane $\mathcal{E}_{-1}$, the depth at P can also be computed in the plane $\mathcal{E}_{-1}$. Hence, by (9) and (10):

$$\begin{cases} x_{P_{-1}} = \dfrac{\cos \beta_{\mathcal{E}_{-1}}}{k_t}\dfrac{T_{-1} \cdot N_{\mathcal{E}_{-1}}}{\sqrt{1 - \left(T_{-1} \cdot N_{\mathcal{E}_{-1}}\right)^2}}, \\ \lambda = d_{-1} - \frac{1}{2}x_{P_{-1}}. \end{cases}$$

Consequently, we obtain the following system in two unknowns k_t and λ:

$$\begin{cases} \lambda = d_{-1} + \dfrac{\cos \beta_{\mathcal{E}_{-1}}}{2\,k_t}\dfrac{T_{-1} \cdot N_{\mathcal{E}_{-1}}}{\sqrt{1 - \left(T_{-1} \cdot N_{\mathcal{E}_{-1}}\right)^2}}, \\ \lambda = d_1 + \dfrac{\cos \beta_{\mathcal{E}_1}}{2\,k_t}\dfrac{T_1 \cdot N_{\mathcal{E}_1}}{\sqrt{1 - \left(T_1 \cdot N_{\mathcal{E}_1}\right)^2}}. \end{cases} \qquad (11)$$

Note a crucial property of the above system, i.e., its linearity in $(\frac{1}{k_t}, \lambda)$.

Remark. The connection with the depth formula (4) as written in 2.2 becomes clear if we write (see Fig. 6):

$$r_P - r_{P_1} = -\frac{1}{2} x_{P_1} T$$
$$- \sqrt{\left(k_{\mathcal{E}_1}\, x_{P_1}^2/2\right)^2 + \left(x_{P_1}/2\right)^2} T_1,$$

thus:

$$\Delta r \cdot N_1 = -\frac{1}{2} x_{P_1} T \cdot N_1,$$

and substituting in (4) with $P_1 = P$ and $P_2 = P_1$ gives:

$$\lambda = \frac{-\Delta C \cdot N_1}{\Delta T \cdot N_1} - \frac{1}{2} x_{P_1}.$$

This shows that the second term of (4) is a function of the curvature k_t of $\mathcal{S}$ at P.

The special case $\cos \beta_{\mathcal{E}_{-1}} = 0$ or equivalently $\cos \beta_{\mathcal{E}_1} = 0$ occurs when the normal N at P is orthogonal to the epipolar plane. This implies that P is a multiple point and thus $\Delta T \cdot N_{\pm 1} = 0$. However, if the camera motion is not along the line of sight at P ($\Delta T \neq 0$), the depth at such points can still be computed since in this case two or more different image projections of P are available. This points out that d_{-1} and d_1 should be computed using a robust formula instead of $d_{\pm 1} = \frac{-\Delta C \cdot N_{\pm 1}}{\Delta T \cdot N_{\pm 1}}$ which is not defined at a multiple point (i.e., $\Delta T \cdot N_{\pm 1} = 0$). See the appendix for details on how we compute $d_{\pm 1}$.

Finally, by solving (11) and assuming that camera motions are not along any line of sight, we obtain the following solutions:

$$\begin{cases} \lambda = \frac{d_1 a_{-1} - d_{-1} a_1}{a_{-1} - a_1}, \\ k_t = \frac{a_1 - a_{-1}}{d_{-1} - d_1}, \end{cases} \quad \text{if } a_{-1} \neq a_1, \qquad (12)$$
$$\lambda = d_{-1} = d_1, \quad \text{if } (a_{-1}, a_1) = (0, 0),$$

where:

$$\begin{cases} a_{-1} = \cos \beta_{\mathcal{E}_{-1}} \dfrac{T_{-1} \cdot N_{\mathcal{E}_{-1}}}{\sqrt{1 - \left(T_{-1} \cdot N_{\mathcal{E}_{-1}}\right)^2}}, \\ a_1 = \cos \beta_{\mathcal{E}_1} \dfrac{T_1 \cdot N_{\mathcal{E}_1}}{\sqrt{1 - \left(T_1 \cdot N_{\mathcal{E}_1}\right)^2}}, \end{cases}$$

and:

$$\begin{cases} d_{-1} = \dfrac{-(C - C_{-1}) \cdot ((T \wedge T_{-1}) \wedge T_{-1})}{(T - T_{-1}) \cdot ((T \wedge T_{-1}) \wedge T_{-1})}, \\ d_1 = \dfrac{-(C - C_1) \cdot ((T \wedge T_1) \wedge T_1)}{(T - T_1) \cdot ((T \wedge T_1) \wedge T_1)}. \end{cases}$$

A geometrical interpretation of d_1 and d_{-1} is that they represent the distances from the camera centre position C to the viewing line intersections (see 2.2). To give an interpretation of the terms a_1 and a_{-1}, consider the projection of P_1 and P_{-1} onto the viewing line at P. Intuitively, a_1 and a_{-1} can be seen as the positions of these projected points with respect to P. Hence, if P is a double point, then either a_1 or a_{-1} is zero and if P is a triple point, then both a_1 and a_{-1} are zero.

Remark. The above solutions are not defined if $a_{-1} = a_1$ with $a_1 \neq 0$. This corresponds to situations where the projections of P_{-1} and P_1 onto the viewing line at P are the same. Thus, the contributions of the viewing directions T_{-1} and T_1 in (11) are equal and the system has an infinity of solutions. However, unique solutions for depth and the normal curvature k_t at such points P can still be found. The idea is to first compute the depth at one of the epipolar correspondents P_1 or P_{-1}. This can be done by applying (11) at p_1 (or equivalently p_{-1}). To this aim, two epipolar correspondents to p_1 (or equivalently p_{-1}) must be found. We already have p and we can also use p_{-1} (or equivalently p_1) since p_{-1} and p_1 are epipolar correspondents in that particular case. Once the depth at P_1 has been computed, we can determine the depth at P by the fact that they both belong to the same parabola. Then, the equation given by applying (11) at P gives the normal curvature k_t.

We conclude that the position of a rim point and the normal curvature k_t in the viewing direction can be estimated at any regular point and double point, except if the camera motion is in the viewing direction. For points where more than two rims intersect (i.e., $(a_{-1}, a_1) = (0, 0)$) and with the same exception, depth can be estimated but not the normal curvature. And if the camera motion is along one of the viewing directions T_{-1} or T_1, neither curvatures nor depth at P can be computed.

4. Comparison with Osculating Circle Methods

Most previous works on surface reconstruction from occluding contours (Vaillant and Faugeras, 1992; Cipolla and Blake, 1990; Szeliski and Weiss, 1993; Seales and Faugeras, 1995; Joshi et al., 1995) make use of osculating circle methods. In such methods, a rim point P is reconstructed by estimating the osculating circle at P and its epipolar correspondents on the previous and the next rim. To this aim, the viewing lines

Figure 7. Rims under non-linear camera motion.

Figure 8. A possible situation for epipolar curves in the case of a non-linear camera motion. The point to be reconstructed is the point P while the position estimated by an osculating circle method is $\tilde{P}$.

are projected onto a plane and a circle tangent to these three directions is computed. In (Cipolla and Blake, 1990; Szeliski and Weiss, 1993; Joshi et al., 1995) one of the epipolar planes is used and in (Vaillant and Faugeras, 1992; Seales and Faugeras, 1995) the radial plane is used. These approaches implicitly suppose that the camera motion is linear (Cipolla and Blake, 1990; Szeliski and Weiss, 1993; Joshi et al., 1995) or that the observed surface is locally cylindrical (Vaillant and Faugeras, 1992; Seales and Faugeras, 1995), and assume that the surface remains on the same side of the tangents in the projection plane.

Our approach avoids such constraints, no assumption has to be made on the camera motion, the local surface shape or the side of the tangents on which the surface lies. In addition, it appears that for non-linear camera motions, part of the rim can not be reconstructed by use of an osculating circle method. See for example Fig. 7 where the three successive camera positions C_{-1}, C, and C_1 are not aligned. In this situation, there is part of the rim to be reconstructed where the point P and its epipolar correspondents are not organised as expected for the estimation of the osculating circle. This corresponds to points P where a_1 and a_{-1} have the same sign or, in other words, where the projection of P_{-1} and P_1 onto the viewing line at P are on the same side of P. In such cases, the epipolar curves defined at P can not be considered as part of the same curve and methods based on the osculating circle lead to false solutions since they approximate both epipolar curves with a single planar curve: a circle (see for example Fig. 8).

This shows that epipolar curves have to be estimated as two different curves and that a method based on the osculating circle allows only a partial reconstruction of the rim in the case of non-linear camera motions.

Our work gives a more general solution to the reconstruction problem. Except for the special cases where
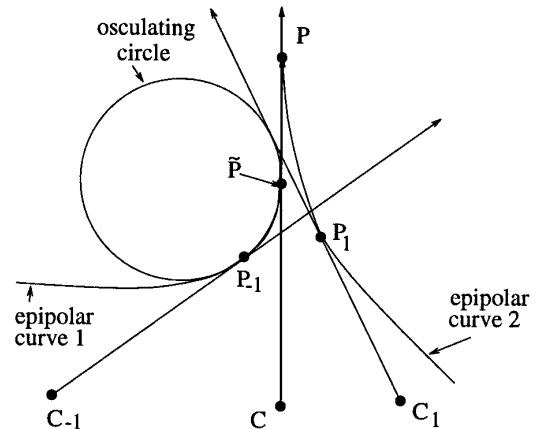
the camera motion is in the viewing direction, depth can be computed at any rim point and for any camera motion. Moreover, it gives a unique solution to the reconstruction problem without knowing on which side of the tangent lines the object lies.

5. Algorithm Performances

In order to determine the accuracy of the reconstruction, we have tested our algorithm on a synthetic sphere. We have chosen synthetic data because it allows computation of exact errors introduced by the reconstruction algorithm. We present here results for plane camera motions.

We suppose that a sphere of radius 200 mm is observed with a camera positioned 1300 mm away from the sphere centre. The intrinsic camera parameters correspond to those of a sony ICX camera. A uniform noise is added to the image measurements with a standard deviation of one pixel in both image plane directions. Since a sphere is observed, the normal curvature in the viewing direction k_t is equal to the inverse of the sphere radius and the depth is constant for points on one rim. Figure 9 shows the rims observed for plane camera rotations around the sphere centre and Fig. 10 gives reconstruction errors for two different rotation angles between successive camera positions.

The depth mean error for a 5 degrees rotation between two camera positions is 1.4 mm and 0.69 mm for a 10 degrees rotation. Note that a depth error of 1 mm yields a distance of 0.0025 mm between the
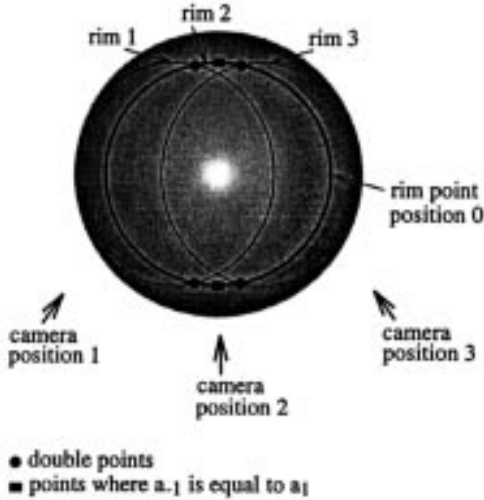
Figure 9.   Rims for plane camera rotations around the sphere centre. The rim to reconstruct is rim 2.

reconstructed point and the sphere. This shows that a high precision can be expected for rim point positions. On the other hand, the second graph shows that the normal curvature computed is far more sensitive to noise than depth. This is not surprising since depth depends on first order differential property whereas curvatures are second order differential properties of the surface. This graph shows also situations where errors increase drastically, namely: points for which a_{-1} and a_1 are close. In such cases, reconstruction is ill-conditioned. This corresponds, in this example (plane camera rotations), to regions of the surface where the three epipolar corresponding points are close. Hence, the estimated curvature is highly biased whereas the depth is still robust. However, for strongly non-linear camera motions, the depth computed is also highly noise sensitive in the vicinity of such points.

In addition, it appears in Fig. 10 that reconstruction errors increase when the rotation between two successive camera positions decreases. In fact, for a 2 degrees rotation the depth mean error is 3.53 mm and reaches 9 mm for a 1 degree rotation. Such noise sensitivity for small camera displacements is due to natural limitations of the reconstruction from images. Indeed, reconstruction algorithms (including stereo-vision) are based on viewing line intersections which are extremely noise sensitive if the viewing directions are close. Hence, very small camera displacements leads to highly biased results and therefore, robust surface reconstruction can not be achieved in such cases.

6.   Experimental Results

We present here results for three real image sequences. Occluding contours were tracked using *snakes* (Berger, 1994; Kass et al., 1988)—see Figs. 11(b), 12(b) and 13(b)—the first image sequence consists of 27 images of a tea pot which were taken with unknown camera motions. Therefore a preliminary calibration step is needed. This was done using a calibration pattern which is present in every image (see Fig. 11(a)). First,
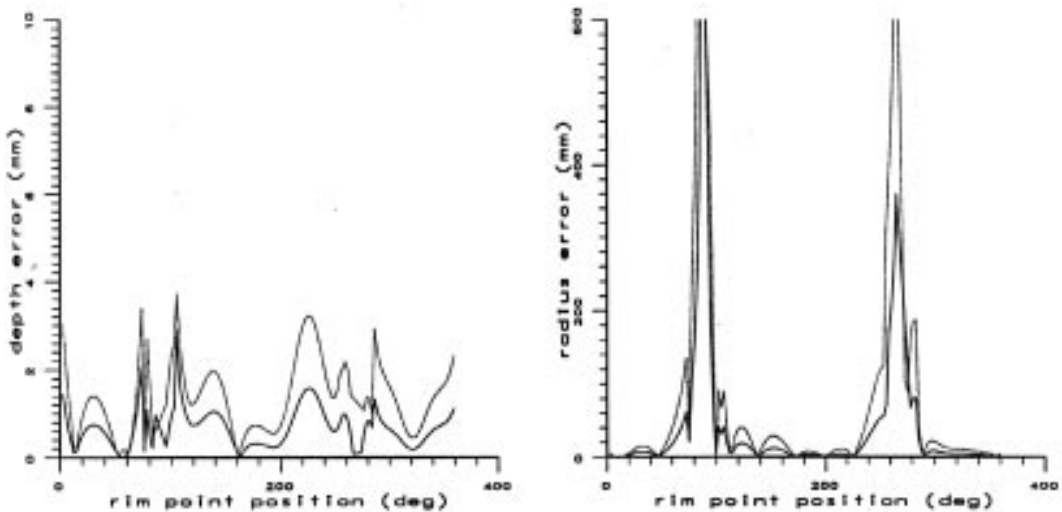

Figure 10.   Depth and radius errors for plane camera rotations. The abscissa represents the reconstructed point position on the rim in degrees, the position of the origin is shown in Fig. 9. Solid lines correspond to camera rotations of 10 degrees and dotted lines to camera rotations of 5 degrees. The two peaks in the radius error graph correspond to points where a_{-1} is equal to a_1.
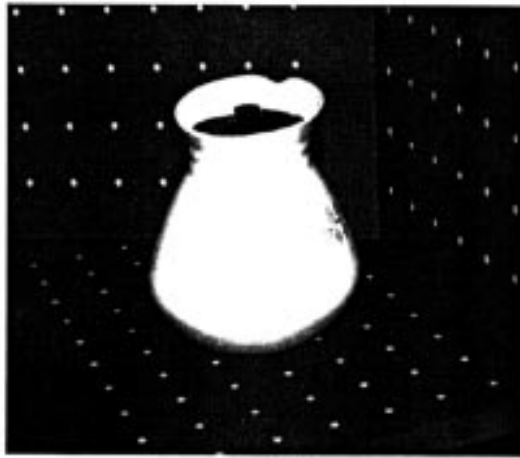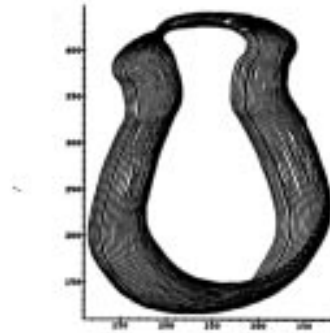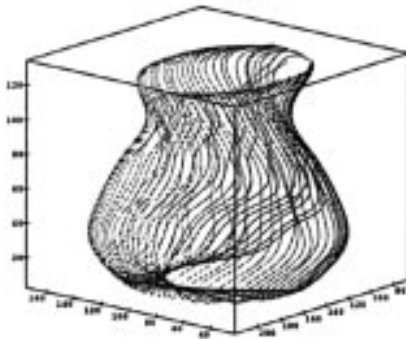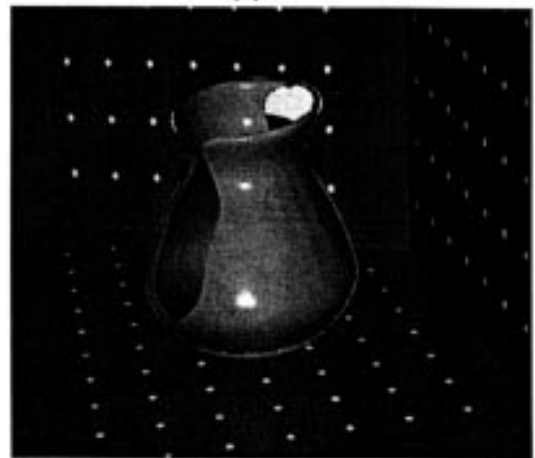
Figure 11.   (a) An image of the sequence, (b) tracked occluding contours, (c) reconstructed rims, (d) triangulated points, (e) rendered surface, (f) projection of the reconstructed surface in the original image.

Figure 12.   (a) An image of the sequence, (b) tracked occluding contours, (c) reconstructed rims, (d) triangulated points, (e) rendered surface, (f) projection of the reconstructed surface in the original image.
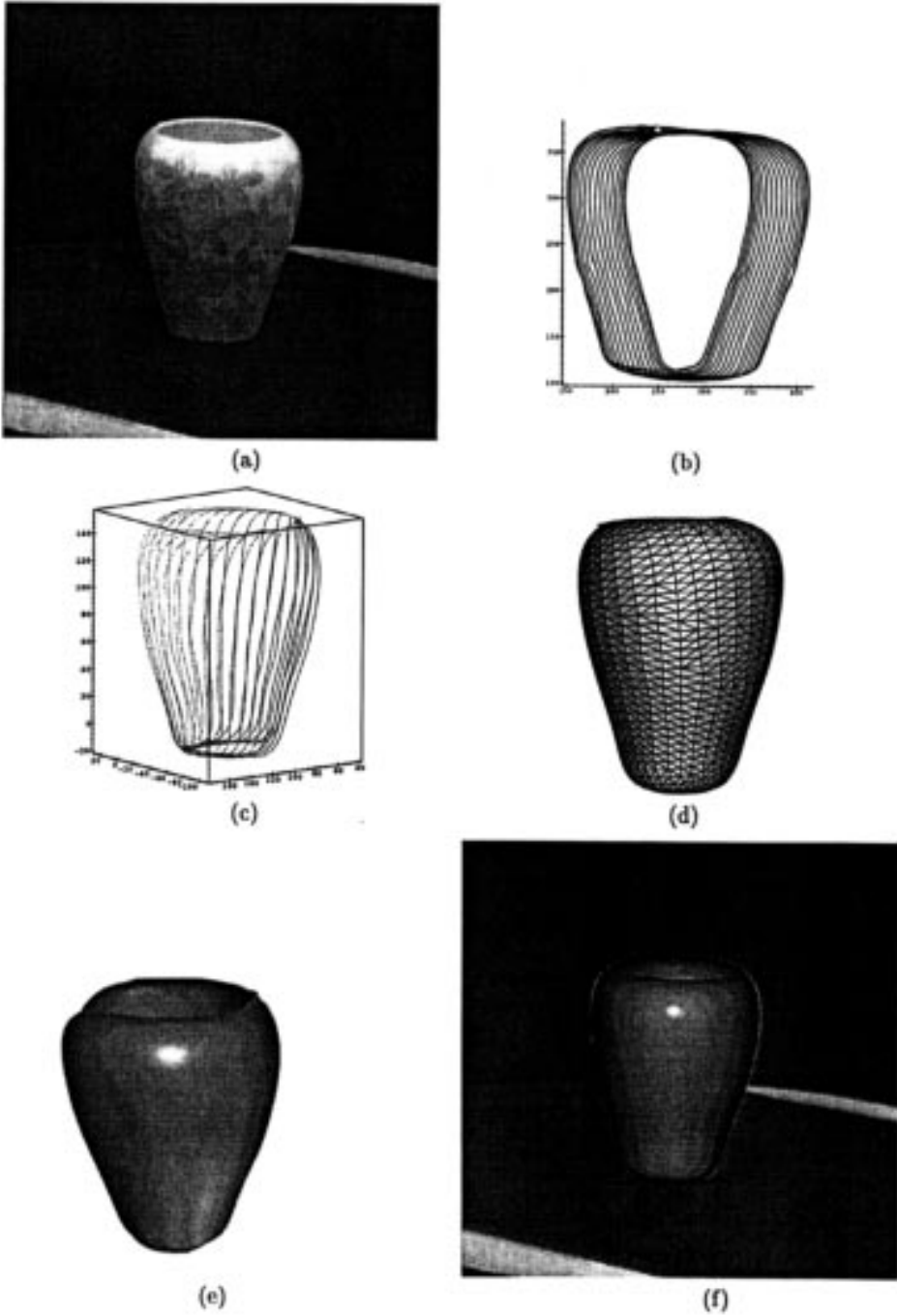
Figure 13.   (a) An image of the sequence, (b) tracked occluding contours, (c) reconstructed rims, (d) triangulated points, (e) rendered surface, (f) projection of the reconstructed surface in the original image.
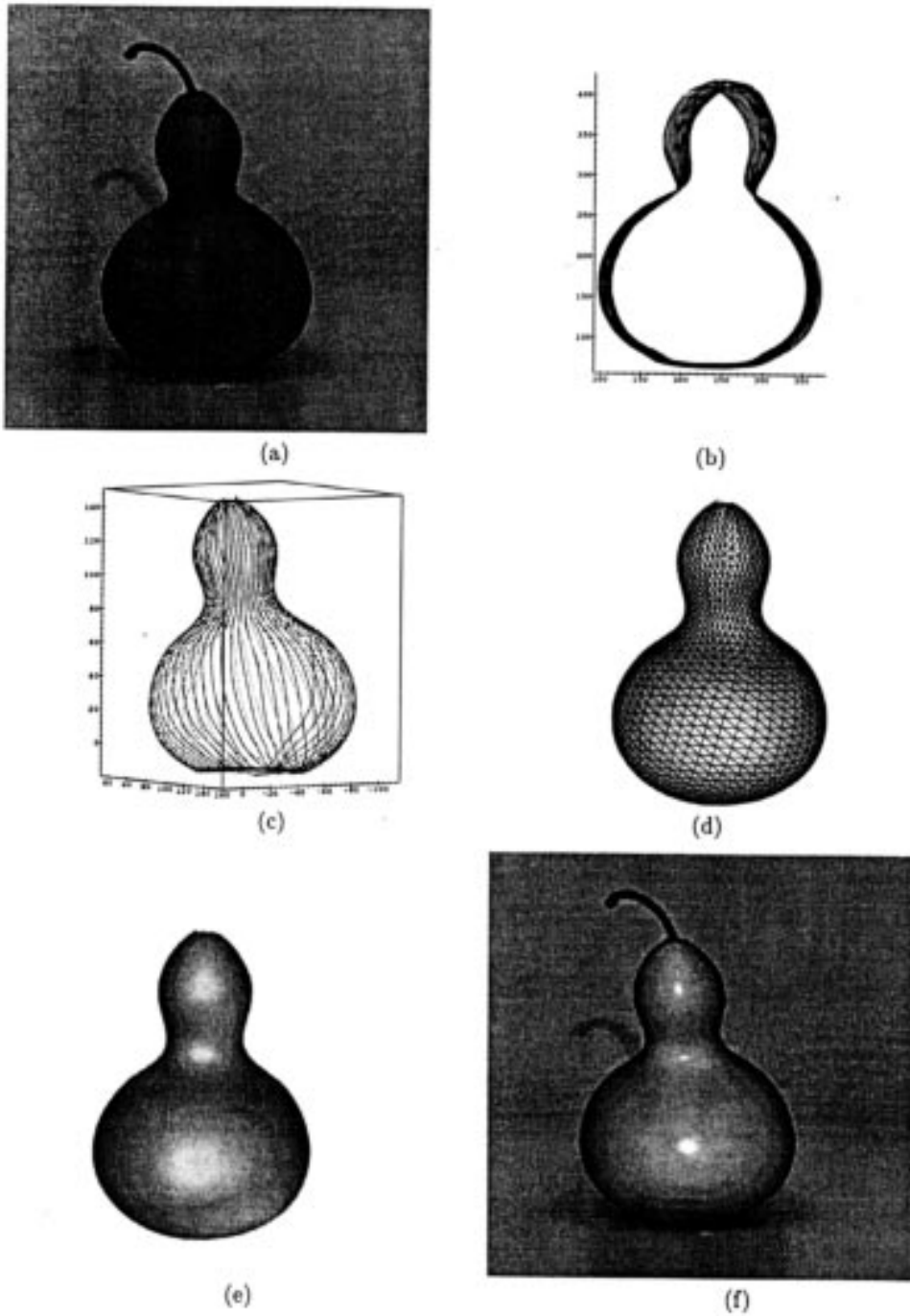
discrete points on rims are recovered if they are not close, according to a threshold, to points where a_1 and a_{-1} are equal. The resulting points are then used to construct a triangular mesh with respect to contour information: two 3D points are connected if and only if they are on two consecutive rims. A ray-tracer is then used to render the surface. Finally, in order to estimate the accuracy of the reconstruction, the rendered surface is projected onto the original images. This is done by use of the perspective projection matrix computed during the calibration step. The results of the different steps of the reconstruction process are shown in Fig. 11.

The reconstructed tea pot shown in this figure is incomplete. This results from the fact that the total camera rotation during the sequence is not sufficient to allow a global perception of the surface. In fact, we reconstruct only what we see and thus a partial surface description may be obtained, depending on the total amount of rotation. This points out that a surface model based on triangular facets is well adapted to modelise the reconstructed surface since it does not require any *a priori* information on the surface. Thus, it allows partial as well as complete surface representations without any parametric or topologic information.

In the next examples, real image sequences of a vase and a calabash were taken using a rotating turntable— see Figs. 12(a) and 13(a). The rotation angle between two successive images is of 10 degrees for the vase sequence and of 7 degrees for the calabash sequence. Except for the calibration step which was performed before the sequences were taken, the reconstruction process is the same as in the previous example. Results are shown in Figs. 12 and 13.

Figures 11(f), 12(f) and 13(f) show that reconstructed surfaces are coherent. It should also be noted that regions of the surfaces where a_{-1} is close to a_1 are still well reconstructed. This is due to the fact that even if these regions correspond to points on one rim for which the reconstruction is ill-conditioned, they may also correspond to points on other rims for which the reconstruction is well-conditioned. Such regions correspond, for the sequences presented, to points which are located at the top and the bottom of the surfaces. Figures 11(c), 12(c) and 13(c) show that the reconstructed rims cover also these regions.

7. Conclusion

We have established formulas for the depth and the normal curvature at a point on a rim of a surface. It shows that the computation of the local surface shape can be done with three consecutive occluding contours, even for a non-linear camera motion. Our work extends previous results obtained by others and leads to a general linear solution which is always defined except when the camera motion is in the viewing direction. Moreover, it gives a unique solution to the reconstruction problem without knowing on which side of the tangent lines the object lies. In addition, we have pointed out local properties of the epipolar correspondence and thus situations where this correspondence may fail, namely the multiple points of an image sequence.

This makes it possible to recover the global surface shape. We have tested our algorithm on real data and shown accurate and efficient results on simple objects (i.e., surfaces are at least C^2 and they are not locally plane). For more complex objects, occluding contours may not be sufficient to allow a complete perception of the object. This is the case, for example, if the object presents concavities. In addition discontinuities may appear on the observed rims and thus a local approximation of the surface by a quadric is not always adapted. See, for example, a composed object such as a coffee cup with a handle; epipolar correspondence may lead to points which belong to the cup and to the handle, which are two different surfaces. This shows that topological properties of the observed object must be taken into account in the reconstruction process. Our current work is concerned with rim discontinuities and we plan to integrate different reconstruction methods in order to handle cases such as concavities.

Appendix: Computing the Distance to the Intersection of Tangents

The distance d_1 from the camera centre position C to the intersection of tangents with direction T_1 and T verifies:

$$(C + d_1 T - C_1 - d' T_1) = 0,$$

where d' is the distance from C_1 to the intersection of tangents with direction T_1 and T. Since $T_1 \cdot N_1 = 0$ we can therefore write:

$$d_1 = \frac{-\Delta C \cdot N_1}{\Delta T \cdot N_1}, \qquad \Delta T \cdot N_1 \neq 0.$$

However, such a formulation does not hold for a multiple point. Since d_1 corresponds to the stereo

reconstruction of the tangents intersection, it should not depend on the normal N_1 to the surface. Thus, if we write instead:

$$d_1 = \frac{-\Delta C \cdot ((T \wedge T_1) \wedge T_1)}{\Delta T \cdot ((T \wedge T_1) \wedge T_1)}, \qquad \Delta T \neq 0,$$

d_1 is then defined at a multiple point except when the camera motion is along the line of sight ($T = T_1$). Furthermore, if T and T_1 are two different viewing directions of a multiple point, then d_1 is the depth of this point.

Acknowledgments

We acknowledge fruitful discussions with Sylvain Petitjean and André Gagalowicz. We also wish to thank Pascal Brand for providing the code for extracting reference points on the calibration pattern and the referees for their helpful comments.

Notes

1. Under the hypothesis we made on S (i.e., smooth and not locally a plane) and for a generic viewpoint, occluding contours are locally smooth except at a cusp or a T-junction (Koenderink, 1986).
2. See (do Carmo, 1976) for definitions.
3. As introduced in 2.1, the epipolar curve is defined for a continuous camera motion. Thus, the curves under considerations are the epipolar curves for linear camera motions going through two successive camera centre positions.

References

Berger, M.-O. 1994. How to track efficiently piecewise curved contours with a view to reconstructing 3D objects. In *Proceedings of the 12th International Conference on Pattern Recognition*, Jerusalem (Israel), 1:32–36.

Bruce, J. and Giblin, P. 1984. *Curves and Singularities*. Cambridge University Press: Cambridge.

Cipolla, R. and Blake, A. 1990. The dynamic analysis of apparent contours. In *Proceedings of 3rd International Conference on Computer Vision*, IEEE (Ed.), Osaka (Japan), pp. 616–623.

Cipolla, R., Åström, K., and Giblin, P. 1995. Motion from the frontier of curved surfaces. In *Proceedings of 5th International Conference on Computer Vision*, Boston (USA), pp. 269–275.

do Carmo, M. 1976. *Differential Geometry of Curves and Surfaces*. Prentice-Hall: Englewood Cliffs, New Jersey.

Faugeras, O. 1993. *Three-Dimensional Computer Vision: A Geometric Viewpoint*. Artificial Intelligence. MIT Press: Cambridge.

Forsyth, D., Mundy, J., Zisserman, A., and Rothwell, C. 1992. Recognising rotationally symmetric surfaces from their outlines. In *Proceedings of Second European Conference on Computer Vision*, Santa Margherita Ligure (Italy), pp. 639–647. Lecture Notes in Computer Science, Vol. 588.

Giblin, P. and Weiss, R. 1987. Reconstruction of surfaces from profiles. In *Proceedings of the First International Conference on Computer Vision*, London, pp. 136–144.

Giblin, P. and Weiss, R. 1995. Epipolar curves on surfaces. *Image and Vision Computing*, 13(1):33–44.

Glachet, R., Dhome, M., and Lapresté, J. 1992. Finding the pose of an object of revolution. In *Proceedings of Second European Conference on Computer Vision*, Santa Margherita Ligure (Italy), pp. 681–686. Lecture Notes in Computer Science, Vol. 588.

Joshi, T., Ahuja, N., and Ponce, J. 1995. Structure and motion estimation from dynamic silhouettes under perspective projection. In *Proceedings of 5th International Conference on Computer Vision*, Boston (USA), pp. 290–295.

Kass, M., Witkin, A., and Terzopoulos, D. 1988. Snakes: Active contour models. *International Journal of Computer Vision*, 1:321–331.

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

Koenderink, J. 1986. 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.), Ablex Publishing Corporation: Norwood, NJ, Chap. 9, pp. 257–285.

Kriegman, D. and Ponce, J. 1990. On recognizing and positioning curved 3D objects from image contours. *IEEE Transactions on PAMI*, 12(12):1127–1137.

Ponce, J., Chelberg, D., and Mann, W. 1989. Invariant properties of straight homogeneous generalized cylinders and their contours. *IEEE Transactions on PAMI*, 11(9):951–966.

Porrill, J. and Pollard, S. 1991. Curve matching and stereo calibration. *Image and Vision Computing*, 9(1):45–50.

Rieger, J. 1986. Three-dimensional motion from fixed points of a deforming profile curve. *Optics Letters*, 11:123–125.

Seales, W. and Faugeras, O. 1995. Building three-dimensional object models from image sequences. *Computer Vision and Image Understanding*, 61(3):308–324.

Szeliski, R. and Weiss, R. 1993. Robust shape recovery from occluding contours using a linear smoother. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, New York (USA), pp. 666–667.

Vaillant, R. and Faugeras, O. 1992. Using extremal boundaries for 3D object modeling. *IEEE Transactions on PAMI*, 14(2):157–173.

Zerroug, M. and Nevatia, R. 1993. Quasi-invariant properties and 3D shape recovery of non-straight, non-constant generalized cylinders. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, New York (USA), pp. 96–103.

Zhao, C. and Mohr, R. 1994. Relative 3D regularized B-spline surface reconstruction through image sequences. In *Proceedings of Third European Conference on Computer Vision* (Stockholm, Sweden), 2:417–426. Lecture Notes in Computer Science, Vol. 801.