

Affine Reconstruction of Curved Surfaces from Uncalibrated Views of Apparent Contours

Jun Sato and Roberto Cipolla

Abstract—In this paper, we consider uncalibrated reconstruction of curved surfaces from apparent contours. Since apparent contours are not fixed features (viewpoint independent), we cannot directly apply the recent results of the uncalibrated reconstruction from fixed features. We show that, nonetheless, curved surfaces can be reconstructed up to an affine ambiguity from their apparent contours viewed from uncalibrated cameras with unknown linear translations. Furthermore, we show that, even if the reconstruction is nonmetric (non-Euclidean), we can still extract useful information for many computer vision applications just from the apparent contours. We first show that if the camera motion is linear translation (but arbitrary direction and magnitude), the epipolar geometry can be recovered from the apparent contours without using any optimization process. The extracted epipolar geometry is next used for reconstructing curved surfaces from the deformations of the apparent contours viewed from uncalibrated cameras. The result is applied to distinguishing curved surfaces from fixed features in images. It is also shown that the time-to-contact to the curved surfaces can be computed from simple measurements of the apparent contours.

Index Terms—Curved surfaces, affine reconstruction, uncalibrated reconstruction, apparent contours, epipolar geometry, time-to-contact.

1 INTRODUCTION

FOR smooth curved surfaces, their apparent contours (or profiles) are dominant in images and they are rich sources of geometric information about the surfaces and motions [9], [6], [22]. These are the projection of the locus of points on the surface which separate the visible and occluded parts on the surface (see Fig. 1). Under perspective projection, this locus, the contour generator, can be constructed as the set of points on the surface which are touched by rays through the projection center.

The fundamental difficulty of recovering structure from apparent contours lies in the fact that the apparent contours are not fixed features. That is, the contour generator slips over the surface under viewer motion and the apparent contours observed from the different viewpoints do not have any correspondence in general (see Fig. 1). Giblin and Weiss [9] showed that if the camera motion is known and is coplanar, the curved surfaces can be recovered uniquely from their apparent contours. This result has been extended for general camera motion by Cipolla and Blake [6] and Vaillant and Faugeras [22]. For extracting curved surfaces reliably, these results are combined with B-spline surface patches [25], optimization processes [2], and with purposive viewpoint control [16]. Unfortunately, all these methods assume that the cameras are *calibrated* and their motions are

known or controlled for a specific position. To cope with unknown camera motion, Seales and Faugeras [21] derived a method for classifying the fixed features from apparent contours and estimating the camera motion from the extracted fixed features. The method assumes that the cameras are calibrated and some fixed features are visible together with the apparent contours. The question is that, if the camera motion is *unknown* and if the camera is *uncalibrated*, is it still possible to recover curved surfaces just from their apparent contours?

It has been shown recently that there exist some points on contour generators which are visible on both apparent contours before and after the camera motion. Such points are called *frontier points* [10], [20] and are visible from both viewpoints (see Fig. 2). The epipolar plane is tangent to the curved surface at the frontier point in 3D space and the epipolar line is tangent to the apparent contour at the projection of the frontier point in images, i.e., *epipolar tangency* [19]. Recent research [4], [1], [12] showed that, by iteratively searching for the frontier points, we can recover the epipolar geometry just from the apparent contours of curved surfaces. Although these works showed the possibility of recovering the epipolar geometry from apparent contours viewed under unknown arbitrary motions of a camera, the methods require nonlinear optimizations, which are sometimes unstable and fall into local minima.

Recent progress [7], [11] in nonmetric reconstruction showed that, if we have fixed features, their correspondences in two views from uncalibrated cameras enable us to recover 3D structures up to a 3D projective ambiguity. Furthermore, it has been shown [18], [17] that, if the camera has a fixed calibration matrix and its motions are limited to

• J. Sato is with the Department of Electrical and Computer Engineering, Nagoya Institute of Technology, Nagoya 466-8555, Japan.
E-mail: junsato@elcom.nitech.ac.jp.

• R. Cipolla is with the Department of Engineering, University of Cambridge, Cambridge CB2 1PZ, England.

Manuscript received 27 Nov. 1997; revised 17 Aug. 1999.

Recommended for acceptance by R. Szeliski.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 107756.

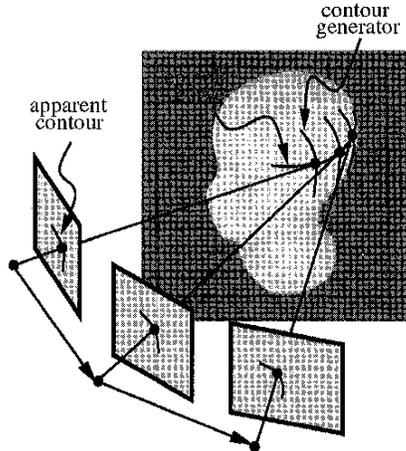


Fig. 1. Structure and motion of curved surfaces. The contour generators separate the visible and occluded parts on a surface. The projections of the contour generators in images are apparent contours. The contour generator slips over the surface under viewer motion and the apparent contours observed from the different viewpoints do not have any correspondence in general.

pure translations,¹ the structure can be recovered up to a 3D affine ambiguity. Unfortunately, we cannot apply these results to curved surfaces directly since apparent contours are not fixed features. To reconstruct curved surfaces from uncalibrated cameras, a purposive control of camera motions has been exploited under the assumption of orthographic projections [13].

In this paper, we consider the reconstruction of curved surfaces from uncalibrated views under perspective projections without knowing or purposively controlling camera motions. We assume that the internal parameters of the camera are unknown but fixed during the camera motions. We also assume that the camera motions are unknown but are limited to linear translations (pure translations in a fixed direction). We show that, under these conditions, the epipolar geometry can be obtained without using any optimization method and the curved surfaces are reconstructed up to an affine ambiguity just from the changes in apparent contours, as in the case of fixed features. The result is applied to labeling image curves as belonging to the projection of curved surfaces or fixed features. We also show that the partial reconstruction of curved surfaces allow us to extract an important cue for visual navigation, i.e., *time-to-contact* to the curved surfaces, just from the changes in apparent contours viewed from an uncalibrated camera.

In Section 2, we define a camera model considered in this paper. In Section 3, the epipolar geometry under pure translations is investigated. In Section 4, the computed epipolar geometry is used for recovering the structure of curved surfaces from uncalibrated views. It is also shown that the time-to-contact to the curved surfaces can be computed from the changes in apparent contours. In Section 5, a method for extracting epipolar geometry by

1. This condition is exactly the same as repeated structures observed in a single view [18].

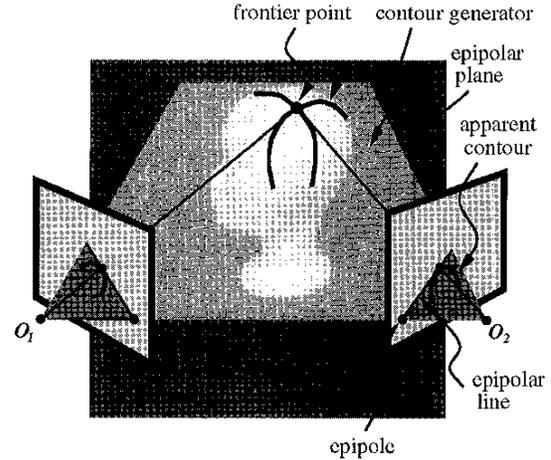


Fig. 2. The epipolar plane defined by the two projection centers touches to the surface at a frontier point. At this point, the contour generators from two viewpoints intersect each other. In the images, the epipolar lines are tangent to the apparent contours at the projections of the frontier point.

using the Hough space is shown. The results of some experiments are shown in Section 6.

2 CAMERA MODEL

We first define a camera model. Consider a point $\mathbf{X} \in \mathbb{R}^3$ in a 3D space observed with camera coordinates \mathbf{X}' so that \mathbf{X} can be described by $\mathbf{X}' \in \mathbb{R}^3$ with rotation, \mathbf{R} , and position, \mathbf{T} of camera with respect to the world coordinates:

$$\mathbf{X} = \mathbf{R}\mathbf{X}' + \mathbf{T}. \quad (1)$$

Suppose \mathbf{X}' is projected onto an image point $\mathbf{m} = [x, y, 1]^T$ by a pinhole camera whose internal parameters can be described by a 3×3 upper triangular matrix, i.e., calibration matrix, \mathbf{A} [8]:

$$\lambda \mathbf{m} = \mathbf{A}\mathbf{X}', \quad (2)$$

where \mathbf{A} is described by the focal length, f , horizontal and vertical scales, k_x and k_y , skew, s , and the principal point, x_0 and y_0 of the camera:

$$\mathbf{A} = \begin{bmatrix} fk_x & s & x_0 \\ 0 & fk_y & y_0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (3)$$

The scale factor, λ , in (2) coincides with the depth to the point with respect to the camera coordinates. In this paper, we assume that the camera is uncalibrated but fixed, i.e., the calibration matrix, \mathbf{A} , is unknown, but it does not change during camera motions.

3 EPIPOLAR GEOMETRY UNDER PURE TRANSLATION

In this section, we discuss the recovery of epipolar geometry from apparent contours viewed from uncalibrated cameras.

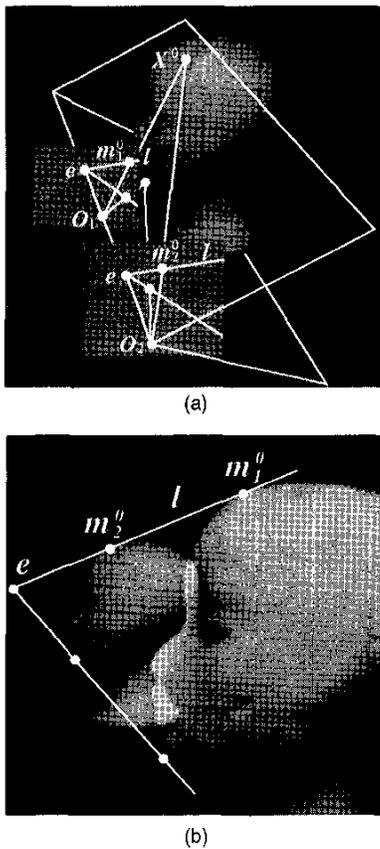


Fig. 3. Epipolar geometry under linear translations. If the motion of the camera is a linear translation, as shown in (a), the epipole, e , and the epipolar lines, l , in images from two viewpoints coincide as shown in (b).

Consider an uncalibrated camera, i.e., a camera whose calibration matrix A is unknown. It is known that if the camera motion is linear translation and the calibration matrix A is fixed during the motion, then the corresponding points observed from two different viewpoints are *auto-epipolar* [18], i.e., the epipoles and epipolar lines before and after the camera motion coincide. Since the frontier points on contour generators can be regarded as fixed features, the above fact can be extended for the apparent contours viewed from the same but uncalibrated cameras. Suppose m_1^0 and m_2^0 are the projections of a frontier point, X^0 , in two views (i.e., before and after a translation). As shown in Fig. 3, if the camera motion is linear translation and the calibration matrix A is fixed:

Property 1. *The projections of a frontier point, m_1^0 and m_2^0 , before and after the camera motion lie on the same epipolar line, l .*

Since the projections of a frontier point are observed as a tangent point of an epipolar line to the apparent contour in the image, we have the following properties:

Property 2 (Epipolar bitangency). *A bitangent line to the corresponding apparent contours coincides with an epipolar line, l , and the bitangent points coincide with the projections of a frontier point, m_1^0 and m_2^0 .*

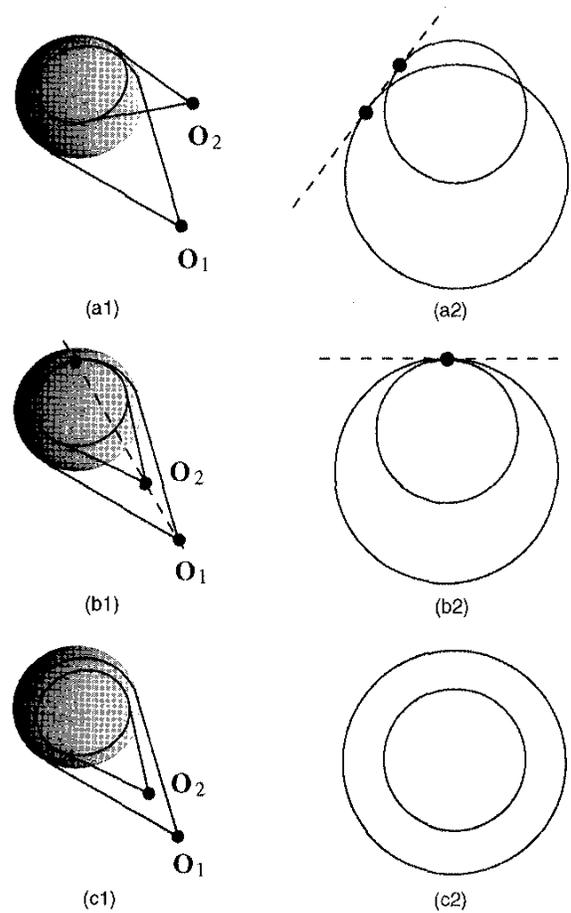


Fig. 4. Epipolar bitangency. (a1), (b1), and (c1) show the three cases of contour generators before and after a camera motion. (a2), (b2), and (c2) show the intersections of apparent contours in these cases. The dashed lines in (a2) and (b2) show bitangents. For finding bitangents which correspond to frontier points, contour generators at two viewpoints must intersect each other.

Thus, if the camera motions are limited to linear translation and if A is fixed, epipolar lines, l , and epipoles, e , can be computed uniquely from the bitangency of apparent contours.

However, bitangents which go through the projections of frontier points do not always exist in images. For example, when the camera is translating directly toward a sphere, contour generators do not intersect each other, and there is no bitangents on apparent contours (see Fig. 4c1 and Fig. 4c2). If there is no intersection of contour generators at two consecutive time instants, there is no frontier point on the surface. Hence, we have no bitangent which goes through the projections of a frontier point in images. On the other hand, if there is an intersection of contour generators at two consecutive time instants, a bitangent which goes through the projections of a frontier point always exists. The only critical case is when the contour generators intersect nontransversally (see Fig. 4b1). In this case, two apparent contours intersect nontransversally, as shown in Fig. 4b2. Nonetheless, it is possible to find a line which is tangent to both two apparent contours at the projections of a frontier

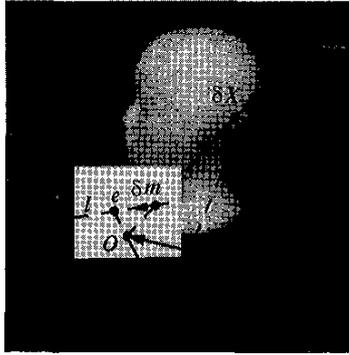


Fig. 5. The infinitesimal motion, δX , of a point X caused by a camera motion, δT , is observed as an infinitesimal motion, δm , of the point in the image.

point. This line is also considered as a bitangent, since it is tangent to both apparent contours (see Fig. 4b2). Thus, the bitangents which go through the projections of frontier points exist in images under linear translations, if and only if the following condition holds.

Condition 1. *The contour generators before and after a camera motion must intersect each other.*

As shown in Section 5, the bitangency can be extracted efficiently by using the Hough space. As we will see in the section, the above condition can also be characterized by using the Hough space.

In the next section, we use the extracted epipole, e , and the projections of frontier points, m^0 , for reconstruction.

4 AFFINE RECONSTRUCTION OF CURVED SURFACES

In this section, we show that, from given epipoles and frontier points, we can, up to an affine ambiguity, reconstruct curved surfaces from their apparent contours viewed from uncalibrated cameras.

4.1 Epipolar Parameterization

To analyze the 3D surface geometry from the changes in 2D image curves (apparent contours), we must identify the correspondences between the successive apparent contours and 3D surfaces. To do this, the epipolar geometry is useful.

Consider an instantaneous motion of a camera so that the view point $O(t_1)$ at time t_1 moves to $O(t_2)$ at time t_2 . Suppose an apparent contour, $C(t)$, in an image at time t corresponds to a contour generator, $\Gamma(t)$, on the surface. The two projection centers, $O(t_1)$ and $O(t_2)$, define a family of epipolar planes, $\Pi(s)$. Then, the contour generators, $\Gamma(t_1)$ and $\Gamma(t_2)$, at time t_1 and t_2 cross over an epipolar plane, $\Pi(s_1)$, at $X(s_1, t_1)$ and $X(s_1, t_2)$, respectively. Since $X(s_1, t_1)$ and $X(s_1, t_2)$ are on the same epipolar plane, $\Pi(s_1)$, their projections, $m(s_1, t_1)$ and $m(s_1, t_2)$, are on the corresponding epipolar lines in images. In the infinitesimal limit, this provides a natural spatio-temporal parameterization of the image and contour generators. This parameterization of curved surfaces and image sequence with respect to s and t parameters is called the *epipolar parameterization* [6], and the

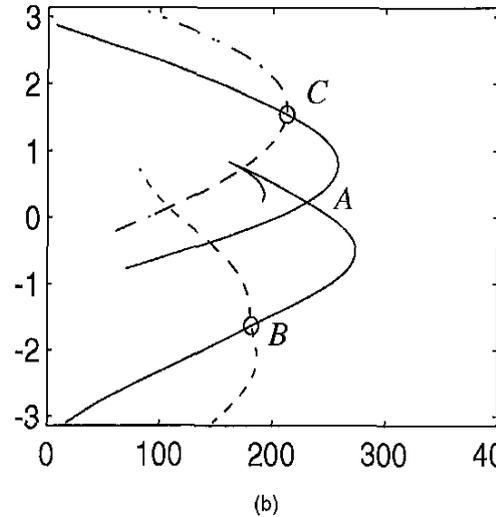
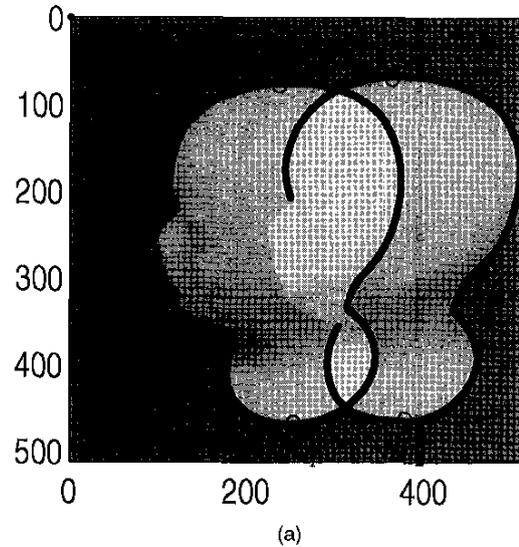


Fig. 6. Extraction of frontier points. The contour curves in images (a) taken before and after a camera motion are transformed into the Hough space in (b). The frontier points are found as intersection points in (b). The extracted frontier points are shown by circles in (a). The dashed lines in (a) show epipolar lines computed from the frontier points.

trajectory of a surface point, $X(s, t)$, with a fixed s parameter is called the *epipolar curve*. Since it enables us to identify correspondences between the changes in apparent contours and the changes in contour generators, the epipolar parameterization is very useful for recovering the surface geometry from apparent contours. The following analyses are based on the epipolar parameterization.

4.2 3D Motion and Changes in Images

Consider an infinitesimal motion, δX , of the 3D point X caused by an instantaneous motion, δT , of the camera (see Fig. 5). By differentiating (1) with respect to the time, t , the velocity and the acceleration of the point X along the epipolar curve caused by the camera motion are described as follows:

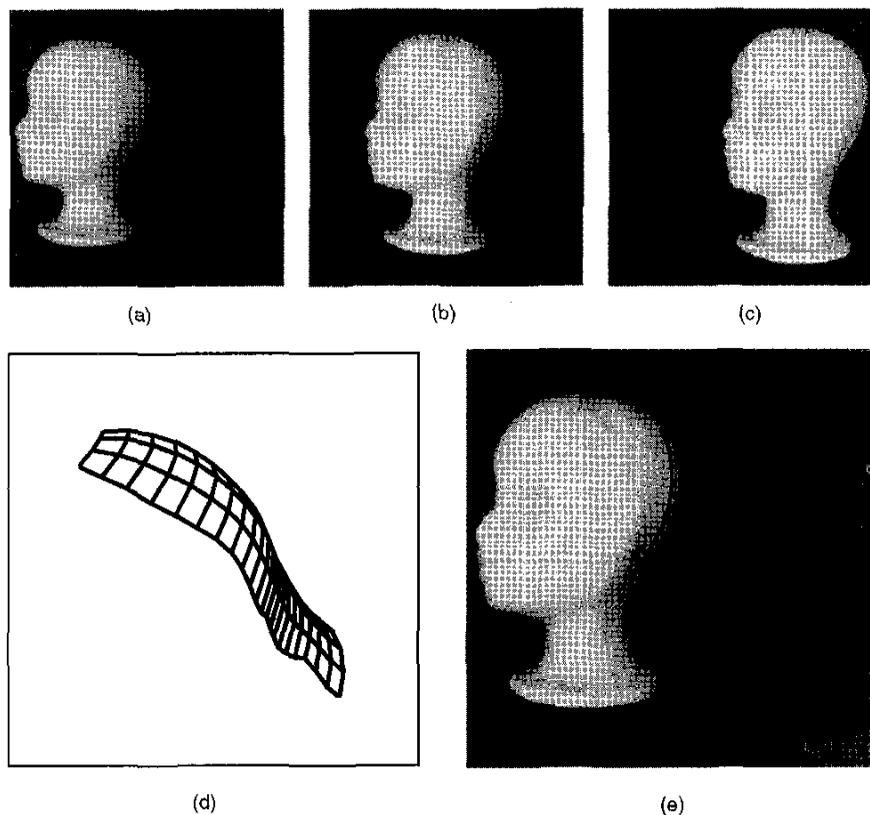


Fig. 7. Results of the affine reconstruction. (a), (b), and (c) show sequential images of a curved surface, i.e., head. These images are used for affine reconstruction of the curved surface. The back of the head is recovered and is shown in (d). Since we do not know the camera parameters, the recovered surface is ambiguous up to a 3D affine. (e) shows the projection of the recovered surface in the first image. Note there is no ambiguity in the projected surface in (e). (a) $t = 1$, (b) $t = 2$, (c) $t = 3$.

$$\mathbf{X}_t = \mathbf{R}_t \mathbf{X}' + \mathbf{R} \mathbf{X}'_t + \mathbf{U} \quad (4)$$

$$\mathbf{X}_{tt} = \mathbf{R}_{tt} \mathbf{X}' + 2\mathbf{R}_t \mathbf{X}'_t + \mathbf{R} \mathbf{X}'_{tt} + \mathbf{U}_t, \quad (5)$$

where the subscripts t and tt denote the first and the second derivatives with respect to time, t . Also, $\mathbf{U} = \mathbf{T}_t$ denotes the translational velocity of a camera and is described by the epipole, \mathbf{e} and the magnitude of a camera motion projected into the image, u , as follows:

$$u\mathbf{e} = \mathbf{A}\mathbf{R}^{-1}\mathbf{U}. \quad (6)$$

Substituting (2) and (6) into (4) and (5) and, since we assume that there is no rotational motion (i.e., $\mathbf{R}_t = \mathbf{0}$, $\mathbf{R}_{tt} = \mathbf{0}$), we have:

$$\mathbf{X}_t = \mathbf{R}\mathbf{A}^{-1}(\lambda_t \mathbf{m} + \lambda \mathbf{m}_t + u\mathbf{e}) \quad (7)$$

$$\mathbf{X}_{tt} = \mathbf{R}\mathbf{A}^{-1}(\lambda_{tt} \mathbf{m} + 2\lambda_t \mathbf{m}_t + \lambda \mathbf{m}_{tt} + u_t \mathbf{e} + u\mathbf{e}_t). \quad (8)$$

This equation shows how the changes in position of an apparent contour caused by the camera motions are observed in images.

4.3 Camera Motion from Frontier Points

We now consider how the camera motion can be computed from frontier points, \mathbf{X}^0 , on curved surfaces. Since a frontier point does not move under linear translations, \mathbf{X}_t^0 and \mathbf{X}_{tt}^0 ,

vanish. Substituting $\mathbf{X}_t^0 = \mathbf{0}$ and $\mathbf{X}_{tt}^0 = \mathbf{0}$ into (7) and (8), we find that the magnitude of motion, u , and its time derivative, u_t , can be computed from the projection of frontier points, \mathbf{m}^0 , \mathbf{m}_t^0 , \mathbf{m}_{tt}^0 , and epipoles, \mathbf{e} , up to the depth to the frontier point, λ^0 , (i.e., speed-scale ambiguity) as follows:

$$u = \lambda^0 F_1(\mathbf{m}^0, \mathbf{m}_t^0, \mathbf{e}) \quad (9)$$

$$u_t = \lambda^0 F_2(\mathbf{m}^0, \mathbf{m}_t^0, \mathbf{m}_{tt}^0, \mathbf{e}), \quad (10)$$

where F_1 and F_2 are computed from:

$$F_1 \mathbf{e} \wedge \mathbf{m}^0 = -\mathbf{m}_t^0 \wedge \mathbf{m}^0$$

$$F_2 \mathbf{e} \wedge \mathbf{m}^0 = -2F_1 \mathbf{e} \wedge \mathbf{m}_t^0 - \mathbf{m}_{tt}^0 \wedge \mathbf{m}^0,$$

where \wedge denotes the vector product. Note, epipoles are unchanged (i.e., $\mathbf{e}_t = \mathbf{0}$) under linear translations. From (9) and (10), we find that, even if the camera is uncalibrated, both u and u_t are computed from the changes in apparent contours in images up to the depth λ^0 and are independent of \mathbf{A} , \mathbf{R} , and \mathbf{T} .

4.4 Recovery of Surface Depth

The magnitude of camera motions computed from the frontier points is next used for recovering the depth to curved surfaces.

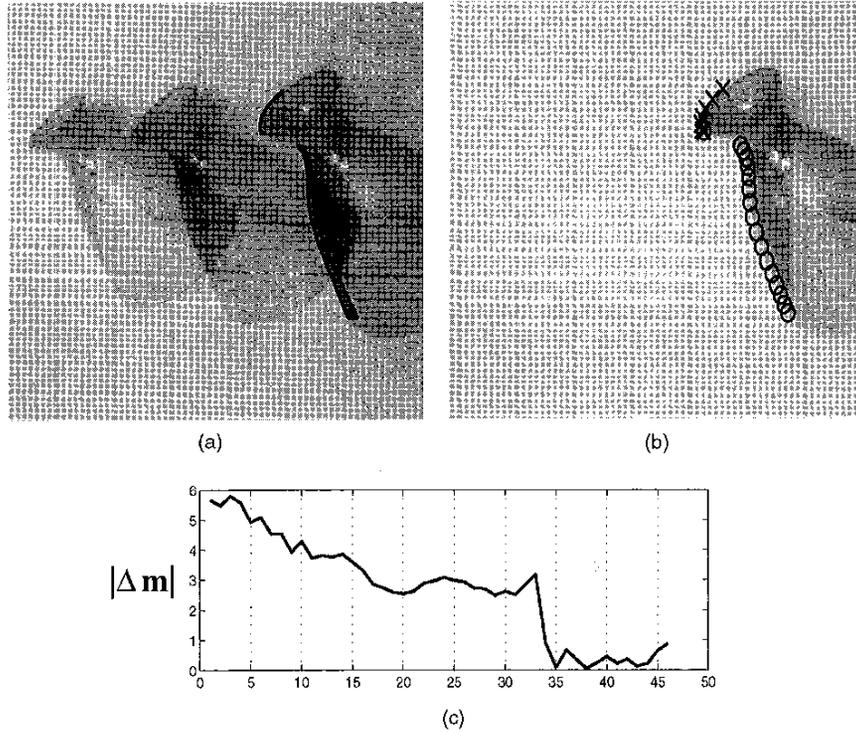


Fig. 8. Results of the distinction of apparent contours. (a) shows sequential images of a curved surface (vase) with fixed features (lip). the reconstructed surface is projected into the first image. (c) shows the distances, $|\Delta \mathbf{m}|$. The apparent contours are distinguished from the fixed features by using $|\Delta \mathbf{m}|$. The distinguished apparent contours and fixed features are shown by (o) and (x) in (b), respectively.

If the point \mathbf{X} on the contour generator is not a frontier point, \mathbf{X}_t and \mathbf{X}_{tt} in (7) and (8) do not vanish. As was shown in [6], since the viewing ray, $\mathbf{X} - \mathbf{T}$, is always tangent to the surface at a point \mathbf{X} , the change \mathbf{X}_t must be parallel to $\mathbf{X} - \mathbf{T}$:

$$\mathbf{X}_t \wedge (\mathbf{X} - \mathbf{T}) = \mathbf{0}. \quad (11)$$

Note, \mathbf{T} is the position of the camera. Substituting (1), (2), (7), and (9) into (11), we have:

$$\lambda = \lambda^0 F_3(\mathbf{m}^0, \mathbf{m}_t^0, \mathbf{m}, \mathbf{m}_t, \mathbf{e}), \quad (12)$$

where F_3 is computed from:

$$F_3 \mathbf{m}_t \wedge \mathbf{m} = -F_1 \mathbf{e} \wedge \mathbf{m}.$$

Thus, the following proposition holds:

Proposition 1. *The depth to a point on a contour generator can be computed from the changes in apparent contour viewed from an uncalibrated camera with linear translations, up to the depth to the frontier point λ^0 .*

4.5 Recovery of Curved Surfaces

We next consider how curved surfaces are recovered from apparent contours viewed from uncalibrated cameras.

Substituting (1), (2), and (7) into (11) and differentiating with respect to t , we find that the time derivative of the depth, λ_t , is computed from the changes in apparent contour and epipole as follows:

$$\lambda_t = \lambda^0 F_4(\mathbf{m}^0, \mathbf{m}_t^0, \mathbf{m}_{tt}^0, \mathbf{m}, \mathbf{m}_t, \mathbf{m}_{tt}, \mathbf{e}), \quad (13)$$

where F_4 is computed from:

$$F_4 \mathbf{m}_t \wedge \mathbf{m} = -F_3 \mathbf{m}_{tt} \wedge \mathbf{m} - F_2 \mathbf{e} \wedge \mathbf{m} - F_1 \mathbf{e} \wedge \mathbf{m}_t.$$

Substituting (12) into (2) and using (1), the contour generator at time $t = 0$ is reconstructed as follows:

$$\mathbf{X}(0) = \lambda^0(0) \mathbf{R} \mathbf{A}^{-1} F_3(0) \mathbf{m}(0) + \mathbf{T}. \quad (14)$$

Substituting (9), (12), and (13) into (7), we can compute the change in the contour generator at time t caused by the camera motion as follows:

$$\mathbf{X}_t(t) = \lambda^0(t) \mathbf{R} \mathbf{A}^{-1} (F_4(t) \mathbf{m}(t) + F_3(t) \mathbf{m}_t(t) + F_1(t) \mathbf{e}(t)).$$

From (9), it is clear that the the depth to the frontier point, $\lambda^0(t_1)$, at time t_1 is computed from the depth, $\lambda^0(0)$, at time $t = 0$ as follows:

$$\lambda^0(t_1) = \lambda^0(0) F_5(t_1), \quad (15)$$

where $F_5(t_1) = e^{-\int_0^{t_1} F_4(t) dt}$. Thus, by computing \mathbf{X}_t iteratively, the curved surface, $\mathbf{X}(s_1, t_1)$, can be reconstructed with respect to the epipolar parameterization, (s, t) , as follows:

$$\begin{bmatrix} \mathbf{X}(s_1, t_1) \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda^0(0) \mathbf{R} \mathbf{A}^{-1} & \mathbf{T} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{X}}(s_1, t_1) \\ 1 \end{bmatrix}, \quad (16)$$

where $\tilde{\mathbf{X}}$ is a 3×1 column vector whose components are computed from the image measurements as follows:

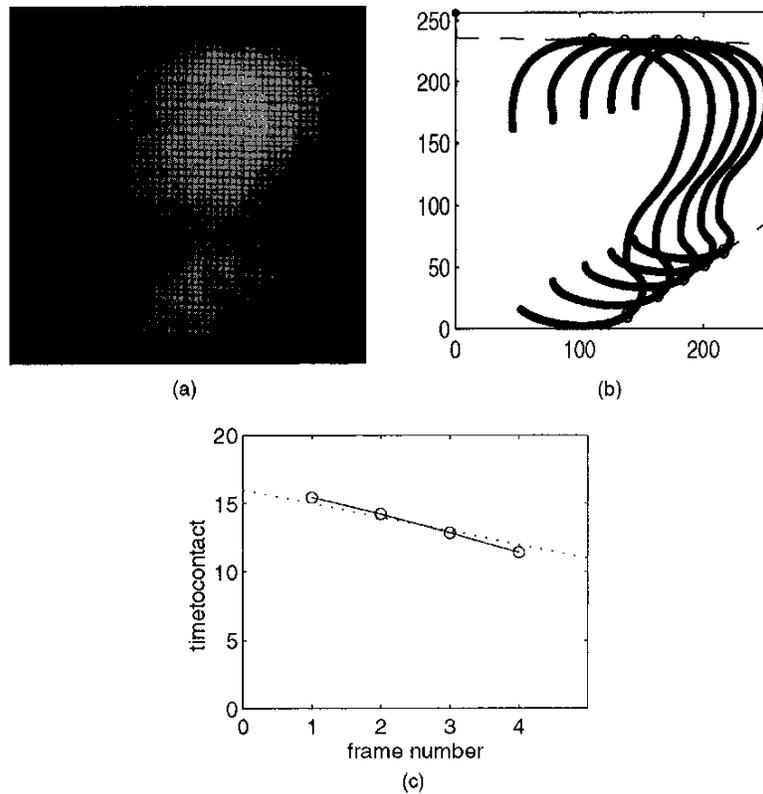


Fig. 9. Results of the computation of the time-to-contact. (a) Shows sequential images of a head. (b) Shows the extracted contour curves (solid lines), frontier points (small circles), and the epipolar lines (dashed lines). (c) Shows the computed time-to-contact to the frontier point (solid lines) together with the calibrated time-to-contact (dashed lines).

$$\tilde{\mathbf{X}}(s_1, t_1) = F_3(0)\mathbf{m}(0) + \int_0^{t_1} F_5(t)(F_4(t)\mathbf{m}(t) + F_3(t)\mathbf{m}_t(t) + F_1(t)\mathbf{e}(t))dt.$$

In (16), the 4×1 vector $[\tilde{\mathbf{X}}^T \ 1]^T$ is available from the image measurements and the 4×4 matrix

$$\begin{bmatrix} \lambda^0 \mathbf{R} \mathbf{A}^{-1} & \mathbf{T} \\ 0 & 1 \end{bmatrix}$$

represents an unknown 3D affine transformation. This means that, even if the camera is uncalibrated and its translational motion is unknown, the curved surfaces can be reconstructed up to a 3D affine transformation:

Theorem 1. *The curved surfaces can be reconstructed from the changes in apparent contour viewed from an uncalibrated camera with linear translations, up to a 3D affine ambiguity.*

4.6 Distinction of Apparent Contours

Up to now we have shown the affine reconstruction of curved surfaces. In this section, we show that the affine reconstruction can be used for distinguishing apparent contours from fixed features. This extends the results from orthographic views [15], [26].

Substituting (16) into (1) and using (12) and (15), we find that the reconstructed contour generator at time $t = t_1$ can be projected back into the original image at time $t = 0$ without any ambiguity as follows:

$$\tilde{\mathbf{m}}(s_1, t_1) = \frac{\tilde{\mathbf{X}}(s_1, t_1)}{F_5(t_1)F_3(t_1)}, \quad (17)$$

where $\tilde{\mathbf{m}}(s_1, t_1)$ denotes the projection of a reconstructed contour generator, $\tilde{\mathbf{X}}(s_1, t_1)$, on to the image at time $t = 0$. We now compute the differential component between the projection, $\tilde{\mathbf{m}}(s_1, t_1)$, and the original contour curve, $\mathbf{m}(s_1, 0)$, along the epipolar line:

$$\Delta \mathbf{m} = \tilde{\mathbf{m}}(s_1, t_1) - \mathbf{m}(s_1, 0). \quad (18)$$

If the image curve is of a fixed feature, the projection, $\tilde{\mathbf{m}}(s_1, t_1)$ must coincide with $\mathbf{m}(s_1, 0)$, and $\Delta \mathbf{m}$ vanishes. If the image curve is an apparent contour, $\tilde{\mathbf{m}}(s_1, t_1)$ does not coincide with $\mathbf{m}(s_1, 0)$, and $\Delta \mathbf{m}$ is not equal to zero. Thus, the magnitude $|\Delta \mathbf{m}|$ can be exploited for distinguishing apparent contours from fixed features even if the camera is uncalibrated.

4.7 Time-to-Contact to Curved Surfaces

As shown in some previous work [5], time-to-contact can provide a useful visual cue for robot navigation. It has been shown that, if the object is planar, the time-to-contact can be computed from the first order derivatives of image flow, i.e., image divergence and deformation [5]. Unfortunately, these analyses are limited to fixed features on planar surfaces. In this section, we show that it is also possible to compute time-to-contact to curved surfaces from apparent contours.

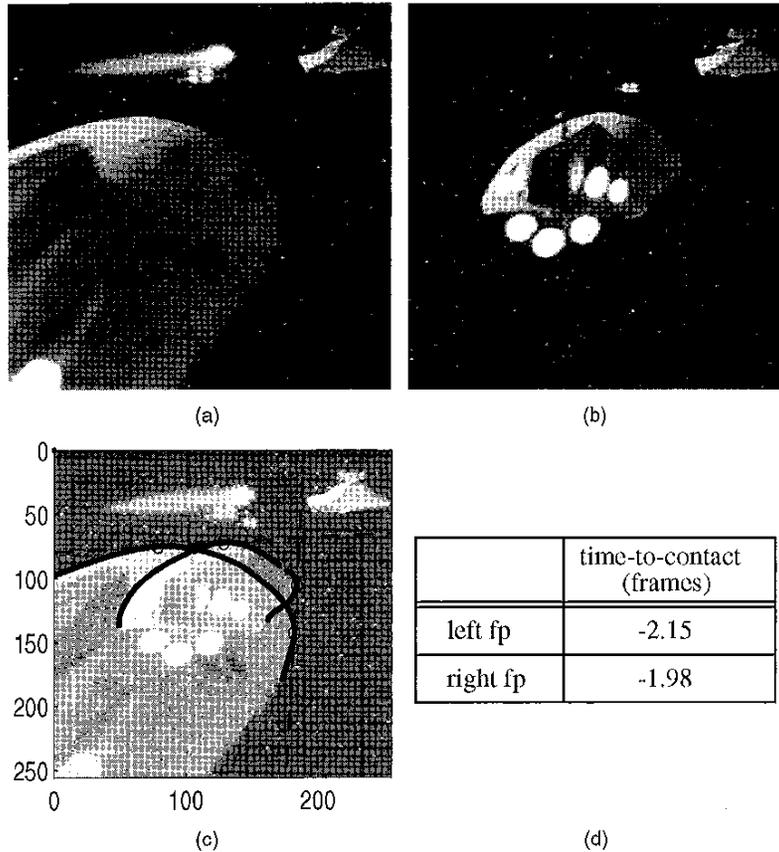


Fig. 10. Results of the computation of time-to-contact. (a) and (b) show two sequential images from the film *Star Wars*. The dashed lines and the circles in (c) show epipolar lines and frontier points extracted by using the Hough space. The time-to-contact to the left and the right frontier points is computed from the extracted epipolar geometry and is shown in (d).

From (9) and since $u = -\lambda_t^0$, we find that the time-to-contact, t_c , to the frontier point on a surface is computed simply from F_1 as follows:

$$t_c = \frac{\lambda^0}{\lambda_t^0} = -\frac{1}{F_1}. \quad (19)$$

Thus, t_c to frontier points is computed from the first derivatives of image curves with respect to time, t . Furthermore, from (12) and (13), we find that t_c to nonfrontier points on a surface can be computed from the second derivatives of image curves with respect to time, t , as follows:

$$t_c = \frac{\lambda}{\lambda_t} = \frac{F_3}{F_4}. \quad (20)$$

Note, since λ^0 cancels out, the time-to-contact to curve surfaces can be computed uniquely just from their apparent contours in images. The time-to-contact does not depend on camera calibration either. Thus, the following theorem holds:

Theorem 2. *The time-to-contact to a frontier point is computed from the first derivatives of apparent contours with respect to time and the time-to-contact to a nonfrontier point on a curved surface is computed from the second derivatives of apparent contours.*

Note, for computing the time-to-contact, Condition 1 must hold.

5 IMPLEMENTATIONS

In this section, we provide a method for extracting frontier points by using Hough transformations.

As we have seen, if the camera motions are linear translations, the frontier points are observed as bitangent points in images. We transform the image curves into the Hough space, that is, the space whose two coordinates are the orientation and the distance to the tangent line at every point on the curve (see Fig. 6). Then, the transformed curves produce intersections in the Hough space. It is known [24] that these intersections correspond to the bitangent in images and the envelope of curves in the Hough space corresponds to a convex hull of the image curves (for enveloping, the reference point must be in the convex hull). Thus, the curves in the Hough space produce an intersection if and only if the image curves have a bitangent at the corresponding points in images. Although not all bitangents correspond to frontier points, the following property of curves in the Hough space is very useful to distinguish frontier points from other bitangents.

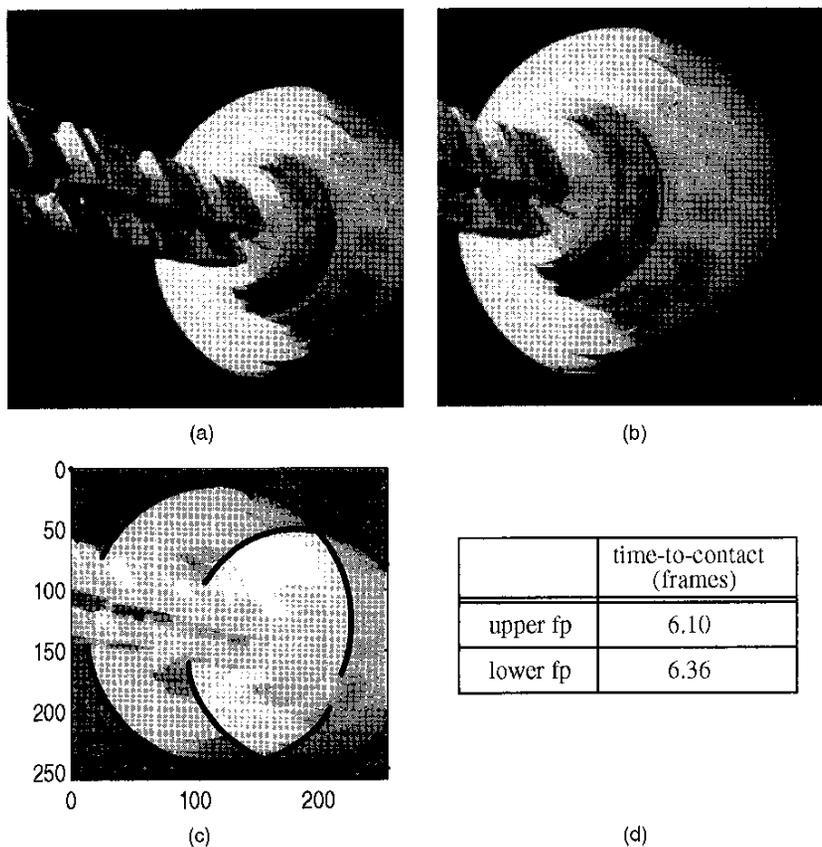


Fig. 11. Results of the computation of time-to-contact. (a) and (b) show two sequential images from the film *2001*. The dashed lines and the circles in (c) show epipolar lines and frontier points extracted by using the Hough space. The time-to-contact to the upper and the lower frontier points is computed from the extracted epipolar geometry and is shown in (d).

Property 3. *If the intersection point is on the envelope of curves in the Hough space, then it corresponds to a frontier point.*

This is because the bitangent points on convex hull of two consecutive image curves always correspond to frontier points, and the convex hull is observed as an envelope of curves in the Hough space.

The point *B* and *C* in Fig. 6b thus correspond to frontier points (although *A* is on the envelope, it does not correspond to a frontier point since this is a self-intersection and corresponds to a self-bitangent). The small circles in Fig. 6a show the extracted frontier points and the dashed lines show the computed epipolar lines. The epipole, *e*, is computed simply as the intersection of two or more epipolar lines.

6 EXPERIMENTS

6.1 Reconstruction of Curved Surfaces

We now show the results from reconstruction experiments. Fig. 7a, Fig. 7b, and Fig. 7c show the three sequential images of a head which are observed from a camera with unknown linear translations. The apparent contours are extracted by fitting B-spline curves [3] and the Hough space is used for computing frontier points. The curved surface is reconstructed from the changes in apparent contours and the

frontier points and is shown in Fig. 7d. Fig. 7e shows the projection of the reconstructed surface in the first image. Although the reconstructed surface, Fig. 7d, has 3D affine ambiguity, its projection in the image, Fig. 7e, has no ambiguity as described in Section 4.6.

6.2 Distinction of Apparent Contours

We next show the results from the distinction of apparent contours from fixed features. Fig. 8a shows a curved surface (vase) used in this experiment. The contour curve of a lip of the vase is a fixed feature, while the contour curve of the side of the vase is an apparent contour. We now distinguish the fixed features (lip of the vase) from the apparent contour (side of the vase) by using the proposed method.

The contour curves of the vase in sequential images are extracted and used for the affine reconstruction of contour generators. The reconstructed contour generators of the vase are projected back into the first image as shown in Fig. 8a. As shown in this image, the projected contour generators coincide with the original curves at the lip of the vase, and do not coincide at the side of the vase. The distances, $|\Delta m|$, between the projected contour generators and the original curves along epipolar lines are computed and plotted in Fig. 8c. The points on the apparent contour are distinguished from fixed features by using $|\Delta m|$ and are shown in Fig. 8b by (\circ) . (\times) in Fig. 8b shows extracted fixed features. The lip of the vase is clearly distinguished from the apparent contours.

6.3 Computation of Time-to-Contact

We next show the results from the computation of time-to-contact to curved surfaces. Fig. 9a shows five sequential images of a head observed from a camera with unknown linear translations and Fig. 9b shows their contour curves extracted by B-spline fitting. The frontier points and the epipolar lines are extracted by using the Hough space as before and are shown in Fig. 9b. The time-to-contact to the frontier point (upper frontier point) is computed and is shown in Fig. 9c by small circles and solid lines. The exact time-to-contact is shown by dotted lines in Fig. 9c. As shown in this figure, the computed time-to-contact is quite accurate and reliable.

Fig. 10 shows a result from another example. Fig. 10a and Fig. 10b are two sequential images from the film *Star Wars*. The motion of the spaceship is considered as a linear translation. The dashed lines in Fig. 10c show the computed epipolar lines and circles show the extracted frontier points. It is computed from the extracted geometry and frontier points that the time-to-contact to the left and the right frontier points on the spaceship is -2.15 frames and -1.98 frames, respectively (minus means that the point is going away from the observer). Fig. 11 shows a result from the film *2001*.

Since the spaceships in Fig. 10 and Fig. 11 are curved surfaces, we cannot compute the exact time-to-contact just from the changes in area of apparent contours. However, as we have seen, the exact time-to-contact is computed efficiently by using the proposed method.

7 CONCLUSIONS

We have shown that, from apparent contours of curved surfaces viewed from uncalibrated cameras, we can reconstruct the curved surfaces up to a 3D affine ambiguity.

We first showed that if the camera motion is a linear translation, the epipolar lines and the frontier points coincide with bitangent lines and bitangent points in sequential images. It has been shown that these epipolar lines and frontier points can be extracted efficiently by transforming image curves into the Hough space and finding intersection points on the envelope of curves in the Hough space. The epipolar geometry is thus recovered without any optimization process, unlike previous work [1], [4]. We next showed that, given the epipolar geometry, the curved surfaces can be reconstructed up to a 3D affine ambiguity from their apparent contours viewed from an uncalibrated same camera.

The result is used for distinguishing apparent contours from fixed features from uncalibrated views. It has also been shown that the time-to-contact to a curved surface can be computed just from apparent contours. For computing the time-to-contact to nonfrontier points, the second derivatives in spatio-temporal images are required and, for frontier points, the time-to-contact is computed just from the first derivatives in spatio-temporal images. These were implemented and tested on real images of curved surfaces. The results are promising.

ACKNOWLEDGMENTS

The authors acknowledge the support of the EPSRC, grant GR/K84202.

REFERENCES

- [1] K. Astrom, R. Cipolla, and P.J. Giblin, "Generalised Epipolar Constraints," *Proc. Fourth European Conf. Computer Vision*, B.F. Buxton and R. Cipolla, eds., vol. 2, pp. 97-108, Cambridge, Apr. 1996.
- [2] E. Boyer, "Object Models from Contour Sequences," *Proc. Fourth European Conf. Computer Vision*, B.F. Buxton and R. Cipolla, eds., vol. 2, pp. 109-118, Cambridge, Apr. 1996.
- [3] T.J. Cham and R. Cipolla, "Automated B-Spline Curve Representation with MDL-Based Active Contours," *Proc. Seventh British Machine Vision Conf.*, vol. 2, pp. 363-372, Edinburgh, Sept. 1996.
- [4] R. Cipolla, K. Astrom, and P.J. Giblin, "Motion from the Frontier of Curved Surfaces," *Proc. Fifth Int'l Conf. Computer Vision*, pp. 269-275, Cambridge, Mass., 1995.
- [5] R. Cipolla and A. Blake, "Surface Orientation and Time to Contact from Image Divergence and Deformation," *Proc. Second European Conf. Computer Vision*, G. Sandini, ed., pp. 187-202, Santa Margherita, Italy, 1992.
- [6] R. Cipolla and A. Blake, "Surface Shape from the Deformation of Apparent Contours," *Int'l J. Computer Vision*, vol. 9, no. 2, pp. 83-112, 1992.
- [7] O.D. Faugeras, "What Can Be Seen in Three Dimensions with an Uncalibrated Stereo Rig?" *Proc. Second European Conf. Computer Vision*, G. Sandini, ed., pp. 563-578, Santa Margherita, Italy, 1992.
- [8] O.D. Faugeras, *Three-Dimensional Computer Vision, A Geometric Viewpoint*. Cambridge, Mass.: MIT Press, 1993.
- [9] P.J. Giblin and R. Weiss, "Reconstruction of Surfaces from Profiles," *Proc. First Int'l Conf. Computer Vision*, pp. 136-144, London, 1987.
- [10] P.J. Giblin and R.S. Weiss, "Epipolar Fields on Surfaces," *Proc. Third European Conf. Computer Vision*, J.-O. Eklundh, ed., vol. 1, pp. 14-23, Stockholm, May 1994.
- [11] R.I. Hartley, R. Gupta, and T. Chang, "Stereo from Uncalibrated Cameras," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 761-764, 1992.
- [12] T. Joshi, N. Ahuja, and J. Ponce, "Structure and Motion Estimation from Dynamic Silhouettes under Perspective Projection," *Proc. Fifth Int'l Conf. Computer Vision*, pp. 290-295, Cambridge, Mass., 1995.
- [13] K.N. Kutulakos, "Affine Surface Reconstruction by Purposive Viewpoint Control," *Proc. Fifth Int'l Conf. Computer Vision*, pp. 894-901, 1995.
- [14] K.N. Kutulakos, "Shape from the Light Field Boundary," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 53-59, 1997.
- [15] K.N. Kutulakos and C.R. Dyer, "Occluding Contour Detection Using Affine Invariants and Purposive Viewpoint Control," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 323-330, 1994.
- [16] K.N. Kutulakos and C.R. Dyer, "Global Surface Reconstruction by Purposive Control of Observer Motion," *Artificial Intelligence*, vol. 78, pp. 147-177, 1995.
- [17] T. Moons, L.J. Van Gool, M.V. Diest, and E.J. Pauwels, "Affine Reconstruction from Perspective Image Pairs Obtained by a Translating Camera," *Applications of Invariance in Computer Vision*, J.L. Mundy, A. Zisserman, and D.A. Forsyth, eds., pp. 297-316, Springer-Verlag, 1994.
- [18] J.L. Mundy and A. Zisserman, "Repeated Structures: Image Correspondence Constraints and 3D Structure Recovery," *Applications of Invariance in Computer Vision*, J.L. Mundy, A. Zisserman, and D.A. Forsyth, eds., pp. 89-106, Springer-Verlag, 1994.
- [19] J. Porrill and S.B. Pollard, "Curve Matching and Stereo Calibration," *Image and Vision Computing*, vol. 9, no. 1, pp. 45-50, 1991.
- [20] J.H. Rieger, "Three Dimensional Motion from Fixed Points of a Deforming Profile Curve," *Optics Letters*, vol. 11, pp. 123-125, 1986.
- [21] W.B. Seales and O.D. Faugeras, "Building Three-Dimensional Object Models from Image Sequences," *Computer Vision and Image Understanding*, vol. 61, no. 3, pp. 308-324, 1995.
- [22] R. Vaillant and O.D. Faugeras, "Using Extremal Boundaries for 3D Object Modeling," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 157-173, 1992.

- [23] B. Vijayakumar, D.J. Kriegman, and J. Ponce, "Invariant-Based Recognition of Complex Curved 3D Objects from Image Contours," *Proc. Fifth Int'l Conf. Computer Vision*, pp. 508-514, 1995.
- [24] M. Wright, A. Fitzgibbon, P.J. Giblin, and R.B. Fisher, "Convex Hull, Occluding Contours, Aspect Graphs and the Hough Transform," *Image and Vision Computing*, vol. 14, pp. 627-634, 1996.
- [25] C.S. Zhao and R. Mohr, "Global Three-Dimensional Surface Reconstruction from Occluding Contours," *Computer Vision and Image Understanding*, vol. 64, no. 1, pp. 62-96, 1996.
- [26] A. Zisserman, A. Blake, C.A. Rothwell, L.J. Van Gool, and M. Van Diest, "Eliciting Qualitative Structure from Image Curve Deformations," *Proc. Fourth Int'l Conf. Computer Vision*, pp. 340-345, Berlin, 1993.



Jun Sato received the BE degree from Nagoya Institute of Technology, Japan, in 1984, and the PhD degree in information engineering from the University of Cambridge, England, in 1997. From 1996 to 1998, he was a research associate in the Department of Engineering, University of Cambridge. During 1997-1998, he was an invited researcher at the ATR Human Information Processing Research Laboratories, Kyoto, Japan. He is currently an associate professor in the Department of Electrical and Computer Engineering, Nagoya Institute of Technology, Japan. His research interests include computer vision, geometric invariants, visual navigation, visual interface, and mixed reality.



Roberto Cipolla obtained a BA (engineering) from the University of Cambridge, England, in 1984, and an MSEE (electrical engineering) from the University of Pennsylvania in 1985. From 1985 to 1988, he studied and worked in Japan at the Osaka University of Foreign Studies (Japanese language) and the Electro-technical Laboratory, Tsukuba (visiting scientist), and he obtained an MEng (robotics) from the University of Electro-Communications in Tokyo in 1988. In 1991, he was awarded a DPhil (computer vision) by the University of Oxford, England, and, from 1991-1992, he was a Toshiba fellow and engineer at the Toshiba Corporation Research and Development Center in Kawasaki, Japan. He joined the Department of Engineering, University of Cambridge, in 1992 as a lecturer and a fellow of Jesus College. He became a reader in information engineering in 1997. His research interests are in computer vision and robotics and include recovery of motion and 3D shape of visible surfaces from image sequences, visual tracking and navigation, robot hand-eye coordination, algebraic and geometric invariants for object recognition and perceptual grouping, and novel man-machine interfaces using visual gestures and visual inspection.