Motion Estimation in Image Sequences Using the Deformation of Apparent Contours

Kalle Åström, *Member, IEEE Computer Society*, and Fredrik Kahl

Abstract—The problem of determining the camera motion from apparent contours or silhouettes of a priori unknown curved three-dimensional surfaces is considered. In a sequence of images, it is shown how to use the generalized epipolar constraint on apparent contours. One such constraint is obtained for each epipolar tangency point in each image pair. An accurate algorithm for computing the motion is presented based on a maximum likelihood estimate. It is shown how to generate initial estimates on the camera motion using only the tracked contours. It is also shown that in theory the motion can be calculated from the deformation of a single contour. The algorithm has been tested on several real image sequences, for both Euclidean and projective reconstruction. The resulting motion estimate is compared to motion estimates calculated independently using standard feature-based methods. The motion estimate is also used to classify the silhouettes as curves or apparent contours. This is a strong indication that the motion estimate is of good quality. The statistical evaluation shows that the technique gives accurate and stable results.

Index Terms—Motion, surface geometry, silhouette, epipolar constraint.

◆

1 INTRODUCTION

The apparent contours of three-dimensional surfaces are known to be a rich source of information. Under *known* viewer motion, reliable descriptions of curved surfaces can be recovered from the apparent contours, cf. [19], [12], [10], [30], [28], [8]. Inference of 3D shape is even possible from a single image, cf. [29]. In this paper, we consider the problem of determining the camera motion from the deformation of apparent contours. The 3D structure is not assumed to be known a priori.

Motion estimates are needed in various situations, e.g., navigation, object grasping and, as mentioned, surface recovery. It is shown how special points on the apparent contour, called *frontier points* or *epipolar tangency points*, can be detected in image sequences and used to recover the viewer motion.

The special case of frontier points under orthographic projection and object rotation around a single axis was considered in [24], [18]. In [23], although primarily concerned with stereo calibration from 3D space curves, it was noted that the intersection of two contours from two discrete viewpoints generated a point, visible in both images. This is also mentioned in [30], where the apparent contours are used for object modeling. This constraint was exploited in [9] in the analysis of the visual motion of space curves. An approach for parallel projection has been presented in [31], [32]. Another approach using trinocular stereo has been presented in [20].

In [1], [11], it was shown how the viewer motion can be calculated from the constraints on the camera motion and the epipolar tangency points. However, only a pair of images was considered. In this situation the problem is often ill-conditioned and therefore it is difficult to accurately estimate the motion parameters. We extend this idea to treat each pair in an image sequence simultaneously to obtain more stable results and to recover the full motion of the camera. We show how an initial estimate of the camera motion can be obtained using only the tracked apparent contours. A fully automatic and practically working algorithm is developed. The algorithm is extensively tested on real image sequences and the results are validated in several independent ways. This is done for the case of an uncalibrated camera. For a camera with unknown and possibly varying intrinsic parameters, it is well-known that the motion can only be recovered up to a projective transformation, cf. [15], [26]. By assuming constant intrinsic parameters of the camera, Euclidean reconstruction of the motion is possible, cf. [16], [2]. The methodology in this paper can also be employed to calculate camera motion with calibrated cameras. A preliminary version of this work was presented in [21].

Given a sequence of images, the objective is to recover the motion of the camera from the deformation of apparent contours. In other words, we want to find the camera matrix for each image in the sequence. Considering a pair of camera matrices, the epipolar geometry can be expressed in terms of the fundamental matrix. All epipolar tangency points seen from these two views obey the epipolar constraint and this can be used to find the viewer motion that is compatible with all the fundamental matrices of the sequence. Additional insight to the problem is given by examining a dual formulation of the generalized epipolar constraint.

The paper is organized as follows. In Section 2, the viewing geometry of contours and the generalized epipolar constraint are described. An algorithm based on a maxi-

---

- *The authors are with the Centre for Mathematical Sciences, Lund University, Box 118, S-221 00 Lund, Sweden.*
  E-mail: {kalle, fredrik}@maths.lth.se.

Manuscript received 26 Sept. 1997; revised 2 Nov. 1998. 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 107582.

mum likelihood estimate of this constraint is given in Section 3. In Section 4, extensive experiments on real image sequences are presented together with a statistical evaluation. The limitations of the method are also described. Finally, Section 5 contains a brief conclusion.

2 THE VIEWING GEOMETRY OF CONTOURS

2.1 Understanding the Viewing Geometry

Introduce a coordinate system in the image and an object coordinate system in 3D space. It turns out that it is convenient to express points in so-called homogeneous coordinates. A point U in the scene with coordinates (X, Y, Z) is represented in homogeneous coordinates as $U = [X\ Y\ Z\ 1]^T$. Vectors are considered to be equal in homogeneous coordinates if they are a nonzero multiple of each other. Similarly in the image, a point with coordinates (x, y) is represented in homogeneous coordinates as $u = [x\ y\ 1]^T$. The projection of a point can then conveniently be represented by a 3 × 4 camera projection matrix P. A point U in the scene with homogeneous coordinates is projected to the point u in the image according to

$$\lambda u = \lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = PU, \quad \lambda \neq 0, \tag{1}$$

where λ is a scale factor. The camera projection matrix can be factorized as

$$P = K\big[R \quad (-Rc)\big], \tag{2}$$

where K contains information about the internal calibration of the camera, R is a rotation matrix representing camera orientation, and c is the camera center or focal point.

Consider a curved surface, for example the one illustrated in Fig. 1 and the image of that surface. For an arbitrary camera position, the *contour generator* of a surface is defined as the locus of the points on the surface for which the tangent planes contain the camera center. The contour generator divides the surface into the visible part and the occluded part. In turn, the *apparent contour* is defined as the image of the contour generator.

An abstract concept called the dual space gives additional insight. Consider a plane in the object world, which will be referred to as the *primal space*. Its equation can be written

$$aX + bY + cZ + d = 0. \tag{3}$$

Represent the plane by a homogeneous vector $\Pi = [a\ b\ c\ d]^T$. Notice that Π and 2Π represent the same plane since the equation is homogeneous. It is natural to think of each plane as being a point in $\mathbb{P}^3$. The set of all planes forms the *dual space*. We say that the plane in the primal space corresponds to a point in the dual space (or the dual of a plane is a point). The class of all points and the class of all planes are symmetrically related to each other. The dual of a point is the set of planes that go through the point. These planes form a plane in the dual space as can be seen by (3). Thus a point in the primal space corresponds to a plane in the dual (or the dual of a point is a plane). To every property of planes in the geometry of points, there corresponds a property of points in the geometry of planes. This is the Principle of Duality, cf. [25]. For smooth surfaces, the dual is defined as the set of planes that are tangent to the surface. Most often the tangent planes form a surface in the dual space. For special types of surfaces, like ruled surfaces the tangent planes form a curve in the dual space.

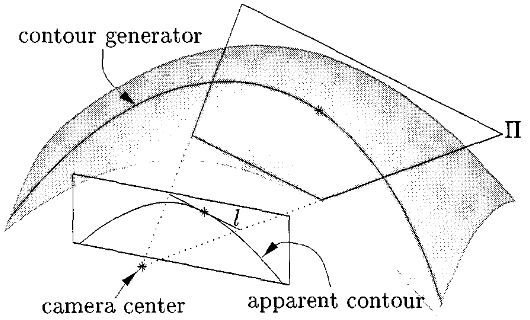


Fig. 1. The projection of a contour generator.

Consider again the viewing geometry of smooth surfaces. The set of tangency planes to the surface at the contour generator is precisely the set of planes that both are tangent to the surface and go through the camera center. In other words, it is the intersection of the dual of the surface with the dual of the camera center, cf. Fig. 2. Thus, this set of tangency planes in the primal space corresponds to a curve on the dual surface. The projection of such a tangency plane, Π, onto the image plane is a line, as can be seen in Fig. 1. If this line $ax + by + c = 0$ is represented by homogeneous coordinates $l = [a\ b\ c]^T$, then the constraint that the world point U lies in the plane Π is the same as saying that its projection PU must lie in the line l, i.e.,

$$\Pi^T U = l^T PU = 0.$$

Since this is valid for all points U, it follows that

$$\Pi = P^T l.$$

Let s be the curve parameter of the apparent contour. For each point $u(s)$ on the curve, calculate the tangent line to the curve $l(s)$. Then the curve in the dual space can be obtained as $\Pi(s) = P^T l(s)$. Thus a curve on the dual surface is easily obtained from the apparent contour in the image if the camera matrix is known.

2.2 The Generalized Epipolar Constraint

2.2.1 Two Images

Now consider the same surface seen from two viewpoints. As before, it is assumed that image and object coordinate systems have been chosen. Denote the projections by camera matrices P_1 for the first image and P_2 for the second image. The contour generator is viewpoint dependent. Therefore the contour generators are different for the two images. The points that belong to both contour generators can be seen in both images.

Consider the camera centers at two different time instants, $c(t_1)$, $c(t_2)$. A point U is projected onto the two images to image points u_1 and u_2 according to

$$\lambda_1 u_1 = P_1 U \text{ and } \lambda_2 u_2 = P_2 U.$$

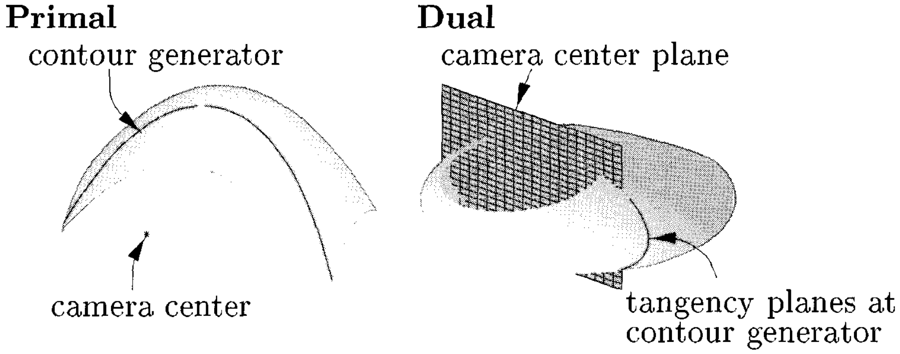


Fig. 2. The dual of the surface is a surface. The dual of the camera center is a plane. The intersection of the dual surface with the dual of the camera center is the dual contour generator.

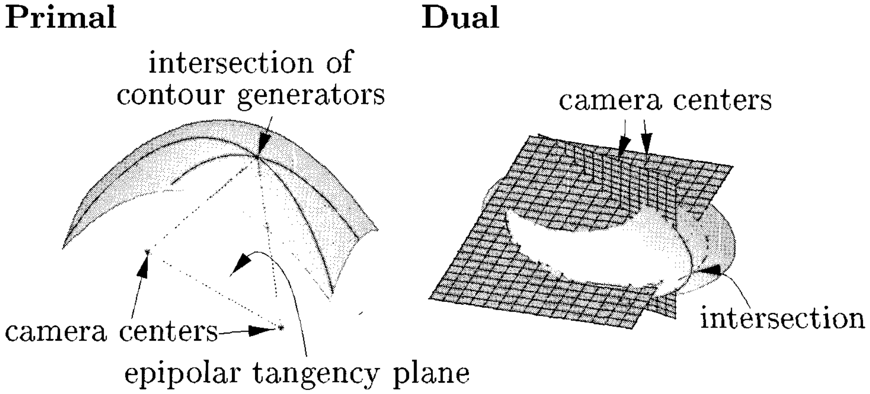


Fig. 3. The epipolar tangency plane is a surface tangent plane that also contains two camera centers. The point of tangency to the surface is the epipolar tangency point. The dual of an epipolar tangency plane is a point, i.e., the point in the intersection of the dual surface and the dual of the camera centers.

This can be rewritten as a matrix equation

$$M\begin{pmatrix} U \\ -\lambda_1 \\ -\lambda_2 \end{pmatrix} = \begin{pmatrix} P_1 & u_1 & 0 \\ P_2 & 0 & u_2 \end{pmatrix}\begin{pmatrix} U \\ -\lambda_1 \\ -\lambda_2 \end{pmatrix} = 0.$$

Notice that M is a 6×6 matrix that has a nontrivial null vector. Therefore the determinant must be zero. The determinant is bilinear in u_1 and u_2, and can therefore be written as

$$u_1^T F u_2 = 0.$$

This constraint is called the epipolar constraint and F is called the fundamental matrix. Since the constraint is linear in each image point, it is also called the bilinear constraint. The 3×3 matrix F can easily be calculated from the camera matrices P_1 and P_2. The element on position (i, j) of matrix F can be expressed as $(-1)^{i+j}$ times the determinant of the 4×4 matrix formed by removing row i from P_1 and row j from P_2 and then stacking the resulting 2×4 matrices on top of each other.

A similar constraint can be derived for the apparent contours of surfaces and curves. Consider all the tangent planes of the surface that go through the two camera centers, cf. Fig. 3. These are called the set of *epipolar tangency planes*. In each image, the epipolar tangency planes are projected to a set of lines, the *epipolar tangency lines*. They all go through a point, the *epipole e*, and each line is tangent to the apparent contour. The tangent points on the apparent contour are called *epipolar tangency points*. This leads to the following theorem, formulated in [1].

THEOREM 2.1. *Given two images, and the epipoles e_1 and e_2, the set of lines through e_1 in image one, which are tangent to an apparent contour, and the corresponding set of epipolar tangency lines in image two, are projectively related.*

This is called the *generalized epipolar constraint*. Notice that only a few points on the apparent contours, the epipolar tangency points, can be used in the constraint.

Returning to the dual formulation. The dual of the first camera center c_1 is a plane Π_1. Each point in the plane Π_1 corresponds to a plane in the primal space that goes through c_1. Similarly, the dual of the second camera center c_2 is a plane Π_2. The intersection $\Pi_{1,2}$ of Π_1 and Π_2 corresponds to all planes (the epipolar planes) in the primal that go through both camera centers c_1 and c_2.

Let l_1 be the dual of the apparent contour in the image, i.e., the tangency lines to the apparent contour in the image plane. Recall that $P_1^T l_1(s)$ is a curve in the dual space that lies both in the dual surface and in Π_1. Similarly, for the sec-

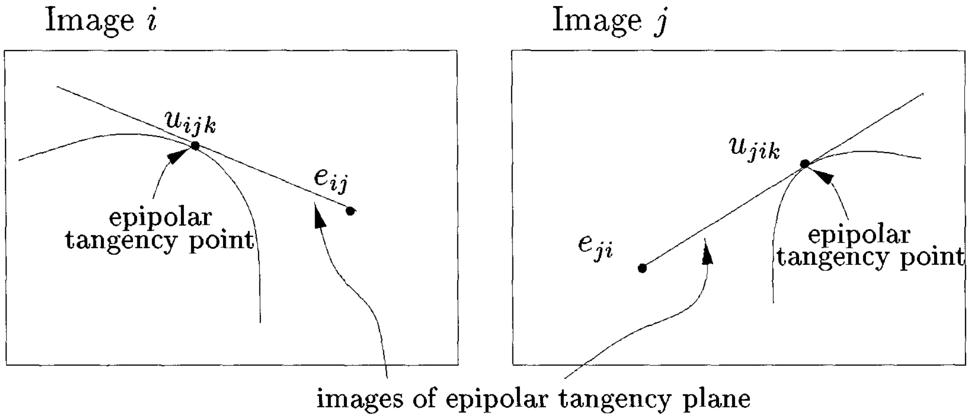


Fig. 4. Illustration of the generalized epipolar constraint.

ond image, $P_2^T l_2(s)$ lies in the dual to the second camera center Π_2. The two curves $P_1^T l_1(s)$ and $P_2^T l_2(s)$ should intersect on $\Pi_{1,2}$. This is the dual formulation of the generalized epipolar constraint.[1]

2.2.2 Many Images

Consider a sequence of m images. For each image pair we obtain several generalized epipolar constraints. In the dual space we obtain m curves $\left(P_1^T l_1(s), \ldots, P_m^T l_m(s)\right)$ that according to the generalized epipolar constraint should intersect pairwise. These curves all lie on the dual to the surface.

Let F_{ij} denote the fundamental matrix between image i and image j, where $1 \leq i < j \leq m$. Let K_{ij} denote the number of corresponding epipolar tangency points between image i and image j. Finally, let u_{ijk} denote the epipolar tangency point k in image i with respect to camera center j, where $1 \leq k \leq K_{ij}$. Notice that u_{ijk} and u_{jik} are corresponding points, that $K_{ij} = K_{ji}$ and that $F_{ij} = F_{ji}^T$. Then, the generalized epipolar constraint can be expressed by the classical epipolar equation

$$u_{ijk}^T F_{ij} u_{jik} = 0. \tag{4}$$

The points u_{ijk} and u_{jik} are in epipolar correspondence. An illustration of the situation is given in Fig. 4. Notice that the epipoles, e_{ij} of image i and e_{ji} of image j, can be obtained as the left and right null space of the fundamental matrix, i.e., $\mathbf{e}_{ij}^T F_{ij} = 0$ and $F_{ij}\mathbf{e}_{ji} = 0$. The tangency points can be obtained from the epipolar tangency constraint, i.e.,

$$\det\begin{bmatrix} \mathbf{e}_{ij} & u_{ijk} & \left(u_{ijk}\right)_s \end{bmatrix} = 0 \text{ and } \det\begin{bmatrix} e_{ji} & u_{jik} & \left(u_{jik}\right)_s \end{bmatrix} = 0 \tag{5}$$

where subscript s denotes differentiation with respect to the parameterization of the apparent contours. Thus, we are interested in computing two quantities, the fundamental matrices and the epipolar tangency points. If the fundamental matrix F_{ij} is known, the epipolar tangency points u_{ijk} and u_{jik}, $k = 1, \ldots, K_{ij}$, can be calculated, and vice versa.

REMARK. The generalized epipolar constraint holds, not only for apparent contours, but for three-dimensional curves as well. Therefore we do not have to know in advance whether a curve is the image of a contour generator or of a curve.

3 COMPUTATION OF CAMERA MOTION

After discussing the principles, we will now describe all steps involved in the computation of the camera motion. We show how to obtain both projective and Euclidean reconstruction of the motion from an uncalibrated camera. The process can be made fully automatic, i.e., no human intervention is needed. The steps involved are:

1) detection and tracking of the apparent contours in all images,
2) initial estimate of motion, and
3) iterative refinement of the motion estimate.

The refinements of the estimate are based on a maximum likelihood method which is a natural way to estimate parameters given noisy input data. The method is standard, cf. [13]. Each step in the iterative refinement involves the following computations:

1) calculation of the fundamental matrices from the camera matrices, and their derivatives with respect to the camera matrices,
2) calculation of the epipoles from the fundamental matrices, and their derivatives with respect to the camera matrices,
3) finding points u_{ijk} on the apparent contours where the tangents go through the epipoles, and their derivatives with respect to the camera matrices,
4) calculating the likelihood L and its derivative with respect to the camera matrices, and
5) updating the camera matrices.

In short, the following formula refines the motion estimate:

---

[1] Given many images, each curve $P_i^T l_i(s)$ lies on the dual surface. This gives a representation of the dual surface in terms of a family of curves. The primal surface is obtained by computing the dual to the dual surface.

$$\begin{tikzcd} P_i \arrow[r] & (F_{ij}, \frac{dF_{ij}}{dP_i}) \arrow[r] & (e_i, \frac{de_i}{dP_i}) \arrow[r] & (u_{ijk}, \frac{du_{ijk}}{dP_i}) \arrow[r] & (L, \frac{dL}{dP_i}) \arrow[llll, bend left, "\text{refinement}"] \end{tikzcd}$$

Each step in this procedure will now be described in a little more detail in the two following subsections. Section 3.4 gives a short summary of the algorithm for projective reconstruction, and Section 3.5 extends the algorithm to obtain a Euclidean reconstruction of the camera motion.

3.1 Extraction and Tracking of Contours

An important part of the calculation of motion from the deformation of apparent contour is the extraction and tracking of the apparent contours. This is a difficult practical problem which has received considerable attention, see [7]. The notion of *snake* has been used for this purpose, see [22]. Roughly speaking, a snake is a parameterized B-spline curve, whose parameters are changed dynamically to fit the apparent contour. The spline curve wriggles to adapt the image, thus resembling a snake. The curve is represented as a collection of B-spline segments, where each segment is represented by four control points (x_i, y_i), $i = 1, ..., 4$. These points generate a segment of the contour $u(s)$, according to

$$[0,1] \ni s \rightarrow u(s) =$$

$$\begin{bmatrix} 1 & s & s^2 & s^3 \end{bmatrix} \begin{bmatrix} 1 & 4 & 1 & 0 \\ -3 & 0 & 3 & 0 \\ 3 & -6 & 3 & 0 \\ -1 & 3 & -3 & 1 \end{bmatrix} \frac{1}{6} \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \\ x_4 & y_4 \end{bmatrix} \in \mathbb{R}^2$$

see [17, p. 493]. This representation has several nice properties. The contour obtained by joining the segment generated by control points (1, 2, 3, 4) and the segment generated by control points (2, 3, 4, 5) automatically has continuous second derivatives, unless some of the control points coincide. Closed contours are easily represented using the control points cyclically.

The snake is matched to the contour in two steps. Euclidean transformations are first applied. This ensures a fast, robust, but rough positioning of the snake in the new image, cf. Fig. 5a and Fig. 5b. The snake is then deformed to match the new image, cf. Fig. 5c. The procedure is explained in more detail in [14].

To deform the snakes, a subpixel edge detector is employed, that not only gives the location of the contour but also a confidence interval in the normal direction of the curve. This is done with the technique described in [3]. For clear, well-defined edges, like the ones in Fig. 5, the individual edge positions can be found with a standard deviation of about on 1/100th of a pixel. Additional errors are introduced in the B-spline fit and in the radial and tangential distortion of the camera. The errors caused by the B-spline fit are of order 1/10th of a pixel. This uncertainty measure is important and will be used in the maximum likelihood estimator. Different epipolar tangency points are weighted according to the uncertainty in their positioning. The bias caused by distortion was unknown prior to the experiment. This will be commented upon in the statistical evaluation later.

A rough estimate of point correspondences is obtained as a by-product of the snake type tracking. These correspondences are used to calculate an initial estimate of motion parameters as described in the next section.

3.2 Initial Hypothesis of Motion

If the epipoles are known, the epipolar tangency points can be computed and given the epipolar tangency points the epipolar geometry can be recovered. However, there is no analytic way of calculating the motion parameters from the apparent contours. An initial estimate of the motion parameters is needed in order to use the generalized epipolar constraints. In some situations, partial knowledge of the motion can be obtained by other means, e.g., using motion sensors or prediction from motion history.

In general, no a priori information is available. However, as a by-product of the snake tracker, approximate point correspondences are obtained for two consecutive images. They are not true correspondences in the sense that they are images of the same point, since the contour sweeps over the surface and because of the aperture problem for curves. The error caused by these effects is normally small compared to the motion between different apparent contours. This suggests that the approximation error may be negligible, and that the technique can be used to get an initial estimate of motion.

Thus the initial hypothesis of the camera matrices is obtained by assuming that points of the apparent contours are images of the same point in space. A point-based structure and motion algorithm is used to calculate the m camera matrices $P_1, ..., P_m$ from these approximate point correspondences, see [27].

3.3 Iterative Refinement of Camera Matrices

3.3.1 Calculation of Fundamental Matrices

Consider image i and image j and assume that an estimate of the two corresponding 3×4 camera matrices P_i and P_j are

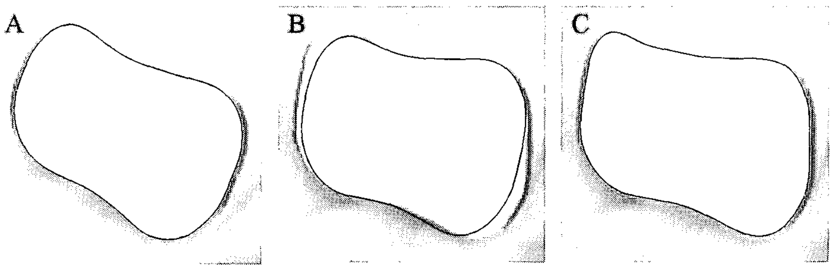


Fig. 5. The snake (A) is used as a template to track the contour in the next image. A rough positioning is found by allowing rigid motion of the snake (B). The new snake is then found by allowing the snake to deform (C).

given. The entries of the 3 × 3 fundamental matrix F_{ij} can be expressed as homogeneous polynomials of degree 4 in the entries of the camera matrices. Each entry is a 4 × 4 determinant formed by taking suitable rows of P_i and P_j. Thus it is straightforward to calculate F_{ij} and its derivatives with respect to P_i and P_j.

3.3.2 Calculation of Epipoles

Again with the specific image pair (i, j) in mind, the epipoles $\mathbf{e}_{ij}$ in image i with respect to image j and $\mathbf{e}_{ji}$ can be calculated as the left and right null spaces of F_{ij}, i.e., $\mathbf{e}_{ij}^T F_{ij} = 0$ and $F_{ij}\mathbf{e}_{ji} = 0$. Using the implicit function theorem, it is relatively straightforward to calculate the derivatives of the epipoles with respect to the fundamental matrix. The derivatives of the epipoles with respect to the camera matrices are then calculated using the chain rule.

3.3.3 Finding Corresponding Epipolar Tangency Points

The epipoles and the tangency constraint (5) determine the position of the tangency points on the apparent contours. The images of the epipolar tangency points are points on the apparent contours where the tangent goes through the epipole. In each image pair (i, j) there are K_{ij} such point pairs

$$\left(u_{ij1}, u_{ji1}\right), \dots \left(u_{ijK_{ij}}, u_{jiK_{ij}}\right).$$

Given the tangency points of a contour in image i,

$$\left\{u_{ijk}\right\}_{k=1,\dots,K_{ij}}$$

and image j,

$$\left\{u_{jil}\right\}_{l=1,\dots,K_{ji}},$$

we need also to determine the correspondence between the two sets of points. If the motion is correct, two corresponding points fulfill the generalized epipolar constraint (4). Since we have a rough estimate of the motion, the constraint can be used for matching. For u_{ijk} take as corresponding point

$$u_{jik} = \arg \min_{l=1,\dots,K_{ij}} \frac{\left\|u_{ijk}^T F_{ij} u_{jil}\right\|}{\sigma_{ijk}}, \text{ if } \frac{\left\|u_{ijk}^T F_{ij} u_{jik}\right\|}{\sigma_{ijk}} < D, \qquad (6)$$

where D is a threshold to avoid outliers. The number σ_{ijk} is a normalization factor described in the next section.

The derivative of u_{ijk} with respect to the epipole e_{ij} is obtained as follows:

$$\frac{du_{ijk}}{de_{ij}} = \frac{du_{ijk}}{ds}\frac{ds}{de_{ij}},$$

where s denotes differentiation with respect to the parameterization of the apparent contour. Using the implicit function theorem and (5), we get

$$\frac{ds}{de_{ij}} = \frac{\left(u_{ijk}\right)_s \times u_{ijk}}{\det\begin{bmatrix} u_{ijk} & \left(u_{ijk}\right)_{ss} & \mathbf{e}_{ij}\end{bmatrix}},$$

where × denotes cross product. The derivatives of the tangency point with respect to the curve parameter s, that is, $(u_{ijk})_s$ and $(u_{ijk})_{ss}$ are easily obtained from the B-spline representation. The chain rule gives the derivative of u_{ijk} with respect to the camera matrices.

3.3.4 The Likelihood Estimate

Let $\mathbf{P}$ be an abstract variable for the collection of the camera matrices $P_i = P(t_i)$, $t_1 < t_2 < \dots < t_m$, describing the motion.

As residuals α_{ijk} for the maximum likelihood estimate, we use the generalized epipolar constraint in (4), i.e.,

$$\alpha_{ijk} = u_{ijk}^T F_{ij} u_{ijk}. \qquad (7)$$

The error in α_{ijk} are due to localization errors of the tangency points. Since they are extracted from the apparent contours independently, it is a reasonable approximation that the residuals α_{ijk} are independent and Gaussian with zero-mean and standard deviation σ_{ijk},

$$\alpha_{ijk} \in \mathbf{N}(0, \sigma_{ijk}).$$

The assumption that the residuals are Gaussian is obviously an approximation. The assumption that the residuals are independent break down for epipolar tangency points close to each other. This effect has been neglected. The standard deviation σ_{ijk} is estimated using the first order approximation,

$$\sigma_{ijk} = \frac{\partial \alpha_{ijk}^T}{\partial u_{ijk}} \Lambda_{u_{ijk}} \frac{\partial \alpha_{ijk}}{\partial u_{ijk}} + \frac{\partial \alpha_{ijk}^T}{\partial u_{jik}} \Lambda_{u_{jik}} \frac{\partial \alpha_{ijk}}{\partial u_{jik}}, \qquad (8)$$

where $\Lambda_{u_{ijk}}$ and $\Lambda_{u_{jik}}$ are covariance matrices for the epipolar tangency points.[2] These covariance matrices are obtained from the edge detector as the contours are extracted, see Section 3.2. From (7), the derivatives

$$\frac{\partial \alpha_{ijk}}{\partial u_{ijk}} = F_{ij} u_{jik}$$

and

$$\frac{\partial \alpha_{ijk}}{\partial u_{jik}} = F_{ij}^T u_{ijk}$$

are obtained.

Then, for the normalized residuals, $\alpha_{ijk} / \sigma_{ijk} \in \mathbf{N}(0, 1)$, the likelihood function is

$$L = \prod_{\substack{1 \le i < j \le m \\ 1 \le k \le \overline{K}_{ij}}} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{\alpha_{ijk}}{\sigma_{ijk}}\right)^2 / 2} = \frac{1}{(2\pi)^{\frac{n}{2}}} e^{-\left(Y^T Y\right)/2},$$

where $n = \sum_{1 \le i < j \le m} K_{ij}$ is the number of elements in the normalized residual vector

$$Y = \left(\dots, Y_{ijk}, \dots\right)^T = \left(\dots, \alpha_{ijk} / \sigma_{ijk}, \dots\right)^T.$$

If we take,

---

[2] In [33], it was reported that the choice of normalized residuals $\alpha_{ijk} / \sigma_{ijk}$ according to (7) and (8) gives very accurate estimates of the fundamental matrix for points.

$$-\log L = Y^T Y / 2 + \frac{n}{2}\log(2\pi),$$

we see that maximizing the likelihood L is equivalent to minimizing

$$g(\mathbf{P}) = Y^T Y. \tag{9}$$

From the chain rule we get the derivative of α_{ijk} in (7) with respect to the camera matrices,

$$\frac{d\alpha_{ijk}}{d\mathbf{P}} = \frac{du_{ijk}^T}{d\mathbf{P}} F_{ij} u_{jik} + u_{ijk}^T \frac{dF_{ij}}{d\mathbf{P}} u_{jik} + u_{ijk}^T F_{ij} \frac{du_{jik}}{d\mathbf{P}}.$$

The derivatives of u_{ijk}, u_{jik}, and F_{ij} with respect $\mathbf{P}$ have already been calculated in the preceding sections. The derivative of σ_{ijk} is obtained by applying the chain rule to (8). Finally, the derivative of $g(\mathbf{P})$ is

$$\frac{dg}{d\mathbf{P}} = 2Y^T \frac{dY}{d\mathbf{P}} \text{ and } \frac{dY_{ijk}}{d\mathbf{P}} = \frac{\frac{d\alpha_{ijk}}{d\mathbf{P}}}{\sigma_{ijk}} - \frac{\alpha_{ijk}\frac{d\sigma_{ijk}}{d\mathbf{P}}}{\sigma_{ijk}^2}. \tag{10}$$

Now, all necessary quantities have been calculated in order to update the motion parameters.

3.3.5 Updating the Camera Matrices

The problem is clearly nonlinear and in order to minimize $g(\mathbf{P})$ in (9) over the motion parameter manifold, we have chosen the method of Levenberg-Marquardt. Then, the first and second derivatives of g with respect to $\mathbf{P}$ are needed. The calculation of the first derivative was described in the previous section and the second derivative (the Hessian matrix) can be approximated by

$$\frac{d^2 g}{d\mathbf{P}^2} = 2\frac{dY^T}{d\mathbf{P}}\frac{dY}{d\mathbf{P}} + 2\left(\sum \frac{d^2 Y_{ijk}}{d\mathbf{P}^2} Y_{ijk}\right) \approx 2\frac{dY^T}{d\mathbf{P}}\frac{dY}{d\mathbf{P}}. \tag{11}$$

The camera matrices are updated with the following formula:

$$\mathbf{P} = \mathbf{P} - \left(\frac{d^2 g}{d\mathbf{P}^2} + \epsilon I\right)^{-1} \frac{dg}{d\mathbf{P}}, \tag{12}$$

where ϵ is small, positive scalar to make sure that Hessian matrix is invertible.

Note that we can only use the constraints given at epipolar tangency points. The number of such constraints depends on the number of contours, their complexity but also on the motion of the camera. It is, thus, very difficult to give necessary and sufficient conditions for motion estimation from contours. One can, however, discuss this question in a loose sense by considering the number of constraints and the number of degrees of freedom in the motion.

For two images and totally uncalibrated cameras, the number of degrees of freedom is $2 \cdot 11 - 15 = 7$, i.e., 11 for each camera minus 15 for unknown projective coordinate system. Thus one would need at least 7 epipolar tangency points. This can be accomplished by having several convex contours or one complex contour. In the calibrated camera case one needs at least 5 tangency points. For three images the number of degrees of freedom is 18, i.e., $(3 \cdot 11 - 15)$. But now it is possible to get constraints for each image pair, between image 1-2, image 1-3 and between image 2-3. For each image pair, 6 epipolar tangency points are needed in average. As the number of images increases, the number of constraints increases quadratically, while the number of variables (i.e., motion parameters) increases linearly. With 22 images the number of degrees of freedom is 227 $(22 \cdot 11 - 15)$, the number of image pairs is 231 so one needs less than one epipolar tangency point in average in each image pair. This implies that, in theory, it could be sufficient to track one single contour in the sequence in order to compute the motion, as long as there are sufficiently many images. It should, however, be noted that this is impossible in practice. Most probably this problem would be extremely ill-conditioned.

Using image pair 1 and 2, we get constraints on camera matrices P_1 and P_2. Using image pair 2 and 3, we get constraints on camera matrices P_2 and P_3. Now using image pair 1 and 3, we get additional constraints on P_1 and P_3. Most often we get more constraints than needed for the determination of the camera matrices. The maximum likelihood estimate is the solution to an overdetermined nonlinear least squares problem. The fact that the problem is overdetermined gives us the extra noise reduction.

3.4 Algorithm

In summary, the following algorithm is proposed to compute the camera motion up to an projective transformation:

1) Detect the apparent contours in the image sequence (Section 3.1).
2) Compute an initial estimate of the camera motion $\mathbf{P}$ (Section 3.2).
3) Given $\mathbf{P}$, calculate all the fundamental matrices F_{ij}.
4) For each F_{ij}, calculate the epipoles e_{ij} and e_{ji}.
5) For each e_{ij}, calculate the epipolar tangency points u_{ijk} with constraint (5).
6) Determine corresponding tangency points with the matching constraint in (6).
7) Given F_{ij} and u_{ijk}, calculate $g(\mathbf{P})$ according to (9).
8) If $g(\mathbf{P})$ is sufficiently small, stop.
9) Calculate $\frac{dg}{d\mathbf{P}}$ in (10) and $\frac{d^2 g}{d\mathbf{P}^2}$ in (11) and update $\mathbf{P}$ according to (12).
10) Given the new estimate $\mathbf{P}$, calculate F_{ij}, e_{ij}, and e_{ji}.
11) Given e_{ij} and the old u_{ijk}, update u_{ijk} by looking in the neighborhood of the old point.
12) Go to 7.

Notice that the whole sequence of images is used to estimate the whole motion P_i, $1 \le i \le m$. Thus all epipolar constraints from $m(m-1)/2$ image pairs are used to estimate m camera matrices. After a few iterations, one may repeat step 6 in order to find new correspondences and to eliminate outliers.

3.5 Upgrading to Euclidean Motion

If the images were taken by the same camera, it is a reasonable assumption that the intrinsic parameters of the camera are constant. Then, the projective reconstruction of the camera motion can be upgraded to a Euclidean one. This is known as self-calibration, cf. [16]. It is accomplished with the following steps:

1) projective reconstruction of the motion,
2) estimation of the rotations and translations of the motion and the intrinsic parameters of the camera, and
3) iterative refinement of the Euclidean motion estimate.

The first step has already been described in detail in the previous sections. The second step involves estimation of the absolute conic from the camera matrices. The absolute conic encodes the Euclidean structure in the projective space. For further details, see [2]. The last step is performed similarly to the refinements in the projective case, but now the motion parameter manifold is different. In each refinement step, the camera motion should be updated over the manifold composed of rotations, translations and intrinsic camera parameters. Let **X** be an abstract variable for this manifold.

Still, we want to minimize the likelihood L and therefore we need the derivatives of L with respect to **X**. We have already calculated the derivatives of L with respect to the camera matrices, P_i. Differentiating (2) with respect to **X**, we get $\frac{dP_i}{d\mathbf{X}}$. The chain rule gives

$$\frac{dL}{d\mathbf{X}} = \frac{dL}{dP_i}\frac{dP_i}{d\mathbf{X}}.$$

In short, the following formula refines the Euclidean motion estimate:

$$\begin{tikzcd}\mathbf{X} =(R_i,t_i,K) \arrow[r] & (P_i, \frac{dP_i}{d\mathbf{X}}) \arrow[r] & (L, \frac{dL}{d\mathbf{X}}) \arrow[ll, bend left, "\text{refinement}"]\end{tikzcd}$$

All necessary components for Euclidean reconstruction are now available.

4 EXPERIMENTAL VALIDATION AND STATISTICAL EVALUATION

In this section, we will present several practical examples of motion determination, both indoor and outdoor scenes. The results are statistically analyzed and always validated by an independent method.

4.1 Validation of the Motion Estimate

In order to validate the motion estimates obtained from the generalized epipolar constraints above, it is important to have independent means of estimating the camera motion. Two such independent motion estimates were developed for comparison, which use standard feature-based methods.

The first method uses traditional camera calibration. We use corresponding point matches of accurate point targets. These points were identified and tracked over the whole image sequence. With these point matches it is straightforward to calculate the camera motion without using any information of the tracked contours.

The second independent method uses curve matches. In several experiments three-dimensional curves were added to the scene. These curves were extracted and then the structure of the curves and the motion of the camera was estimated using a shape or subspace method [5] followed by bundle adjustment for curves [6]. This gives the motion with high accuracy.

The motion estimate using the generalized epipolar constraint can then be compared to that of using points and curves. If the camera orbits are in close agreement (in an appropriate chosen coordinate system), it is a strong indication of the validity of the motion estimate.

A third way of assessing the quality of the motion estimate is to use the motion estimate to classify which image curves are apparent contours and which are images of a three-dimensional curves. In [30], [10], several methods are developed and compared for this purpose. In our experiment, we simply apply the epipolar constraint. Consider a point x on the image curve in the first image and its space point X. If X is a point on a three-dimensional curve, there should be corresponding points on all image curves in the sequence. The motion parameters, i.e., the camera matrices P_i, are considered to be known. In order to recover the space point X from x the only missing component is the depth, according to (1). So by first finding the depth that minimizes the distance between the projections of the point X and the image curve in all images, and then measuring this distance, we can determine if X belongs to a space curve or a contour. Experimental results of the classification are presented in Section 4.3.

4.2 Experiment 1: Fruit and Stones

Several objects (smooth stones, apples, and a banana) were placed on a dark background. Small point markers were also placed in the scene. A sequence of 30 images were then taken with an ordinary video camera, with wide angle lens, cf. Fig. 6. The radial and tangential distortion of the camera was measured offline using images of a calibration pattern. The remaining distortion after correction is less than 0.5 pixel. In the experiment, six contours were detected and tracked through the whole sequence. Each contour is described by a B-spline with around 40 control points, depending on the arc-length of the contour.

The initial estimate of camera motion was then calculated using approximate point correspondences as described in Section 3.2. This estimate was then refined with 15 iterations as described in Section 3.3. The residual error dropped rapidly after the first iterations and then stayed almost constant. Fifteen iterations have proved to be sufficient for all our experiments. The threshold in D in (6) has been set to 20. In the experiment the standard deviation or normalization factor σ_{ijk} in (6) is based upon the estimate of edge localization standard deviation. In the experiments, the edges were localized with a standard deviation of roughly 0.1 pixel in the normal direction. Thus the choice of D corresponds roughly to 2 pixels.

In Table 1, the number of corresponding epipolar tangency points between some of the 435 image pairs is shown. In the experiment the number of degrees of freedom for the projective camera motion was $30 \cdot 11 - 15 = 315$ and the number of constraints was 5,384. This means that there were 12.4 tangency points in average for the 435 image pairs. One silhouette is shown in Fig. 7 with epipolar tangency points relative to the other images in the sequence.

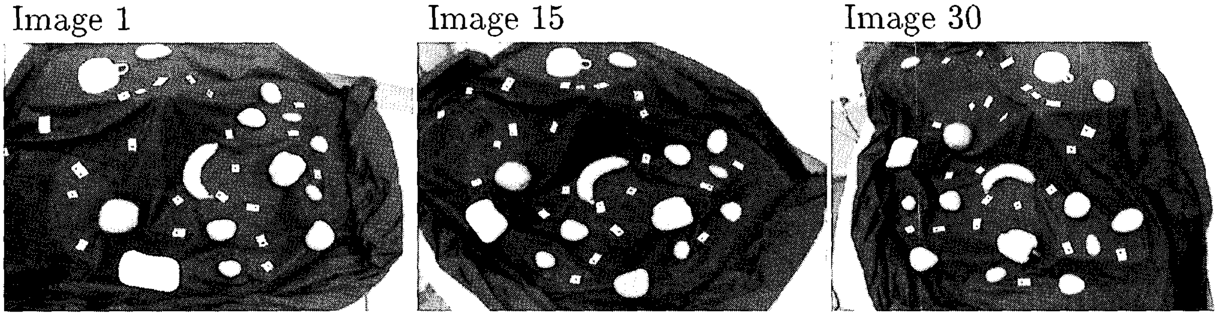


Fig. 6. Image sequence of fruit and small rocks. The six apparent contours employed in the experiment are marked in black.

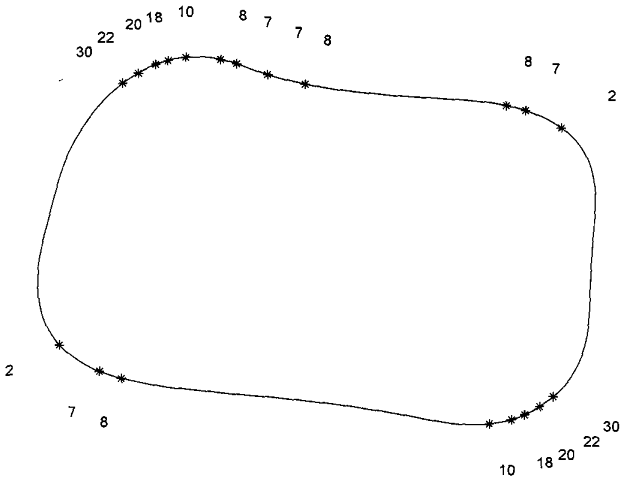


Fig. 7. The epipolar tangency points of one of the silhouettes in image 1 relative to some of the other 29 images. For each tangency point, it is also shown the image number *j* to which the epipolar tangency constraint is given.

The camera motion was also calculated *independently* using the special point markers in the sequence. These 24 points were detected and tracked through the whole sequence. A standard structure and motion routine was used to calculate the motion of the camera. This made it possible to verify the motion calculated from the apparent contours.

In Fig. 8, the camera centers calculated using the points and using the apparent contours are shown with (o) and (+), respectively. In the same figure, the initial estimate before optimization is also shown, marked with (x). The motion is reconstructed up to a projective transformation. Notice the good alignment between the two different approaches.

A statistical evaluation of the results has been performed in order to analyze the stability and accuracy of the solution. The normalized residuals are shown in Fig. 9. These residuals should in theory be approximately random with mean zero and standard deviation 1. As can be seen in the figure the residuals are slightly larger. However, the standard deviation has decreased considerably since the initial estimate, from 27.0 to 1.45. There are still systematic errors of the order 1/10 of a pixel. These errors are most probably due to the remaining distortion of the camera. It is possible to eliminate the errors caused by distortion by including these parameters in the maximum likelihood estimator. The errors caused by the imperfect B-spline fit can also be eliminated by tracking the tangency points directly from image data. Thus one could hope to achieve residuals corresponding to errors of the order 1/100 of a pixel.

TABLE 1
NUMBER OF CORRESPONDING EPIPOLAR TANGENCY POINTS BETWEEN DIFFERENT IMAGES

image	2	5	10	20	30
1	14	14	13	12	12
2		16	12	12	12
5			12	12	12
10				12	12
20					12

4.3 Experiment 2: Candle, Vase, and Curves

The next experiment was performed on an image sequence of a candle, a vase, and two three-dimensional curves, cf. Fig. 10. The generalized epipolar constraints hold for all of these features. The curves were produced by folding a piece of white paper. The sequence consists of a total of 20 images.

The candle and the vase are not so smooth objects as the fruit and stones in the previous experiment. The apparent contours are self-occluding, i.e., one part of the contour is occluding another. These effects are difficult to detect and

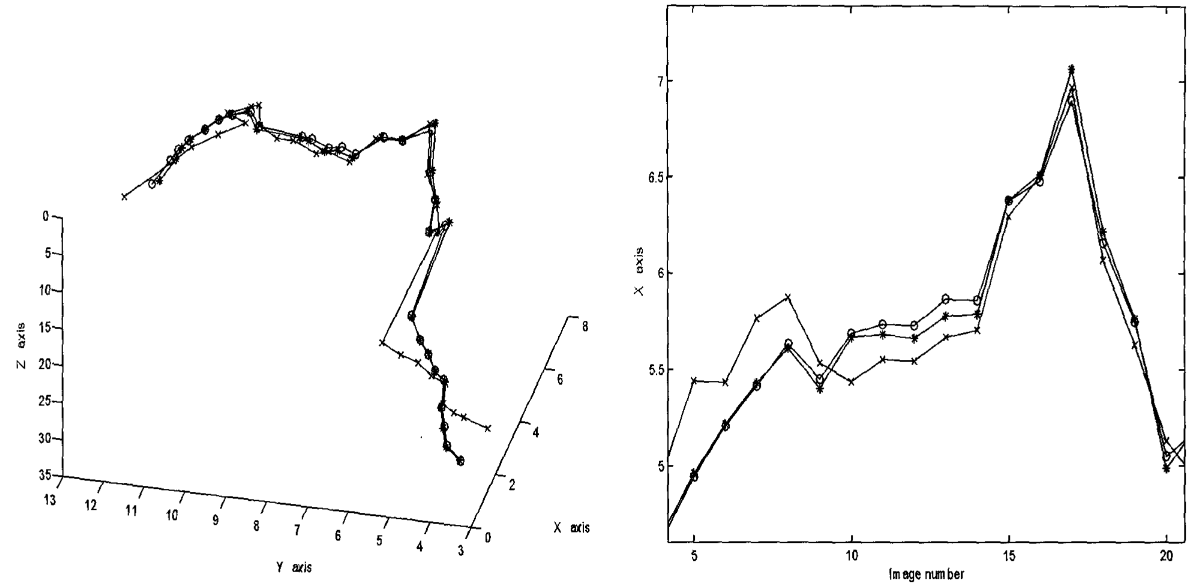


Fig. 8. *A comparison between the motion estimate before iterative refinement (marked with* ×), *the motion computed with the apparent contours after iterative refinement (marked with* *) and *the motion computed with special points (marked with* ∘). The left image shows the orbit of the camera focal points in a 3*D* graph and the right images show a zoom of the *X* coordinate for a subsequence of the image set.

they may introduce false epipolar tangency points. Tangency points may appear to be corresponding, but in reality they are not. However, it is possible to distinguish convex parts from concave parts of the apparent contours by measuring the curvature. The false matches due to self-occlusion only appear at concave parts of the contours and thus can be eliminated by utilizing only convex parts. There is, however, a price to pay. Less tangency points are obtained since only parts of the contour are used. This has proved useful in some of the experiments, but might be difficult to implement in a fully automatic system since it is difficult to distinguish convex from concave parts without knowing which side of the boundary the inner part of the object lies.

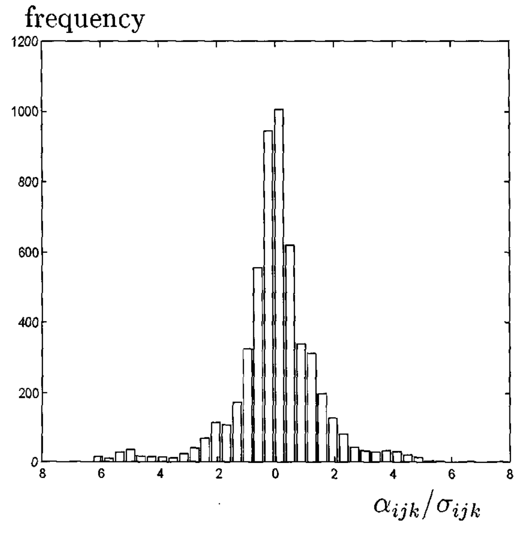


Fig. 9. Histogram of normalized residuals for the fruit and stones experiment. In the sequence of 30 images there are more than 5,000 epipolar tangency points.

In order to validate the motion estimate, the motion was calculated using only the two space curves, as described in Section 4.1. This provides us with an independent estimate of the motion. The reconstruction of the motion was computed up to an Euclidean transformation with a self-calibration technique.

First, we tried to perform a projective reconstruction of the motion, with exactly the same procedure as in the previous experiment. However, the result was not satisfactory. The normalized residuals are shown in Fig. 11a. As can be seen, the residuals are far larger than what can be expected. The alignment is very poor compared to that obtained using only curve matches.

Then, the same experiment was performed using only the convex parts of the apparent contours in order to avoid false epipolar tangency points. The normalized residuals after 15 iterations of iterative refinement for this experiment is shown in Fig. 11b. The errors are of the same order as in Experiment 1. After the projective reconstruction, the motion was upgraded to Euclidean motion, as described in Section 3.5. In this case the number of degrees of freedom for the camera motion is $20 \cdot 6 + 5 - 7 = 118$, i.e., 6 for the position and orientation of each camera, 5 for the intrinsic parameters which were constant minus 7 for the choice of Euclidean coordinate system and scale. The number of constraints was 1,929, which is slightly more than 10 tangency points per image pair.

A comparison was made of the motion obtained from the curve matches and the motion obtained from the generalized epipolar constraint. The camera centers of the two orbits are shown in Fig. 12 together with the reconstructed

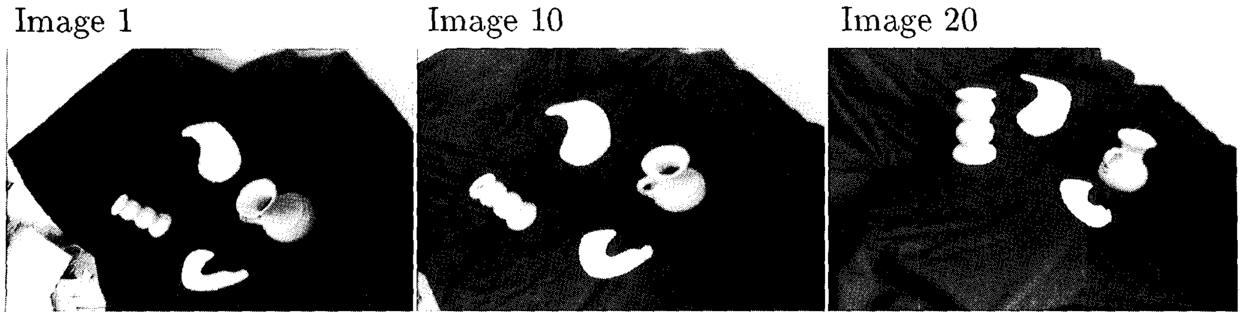
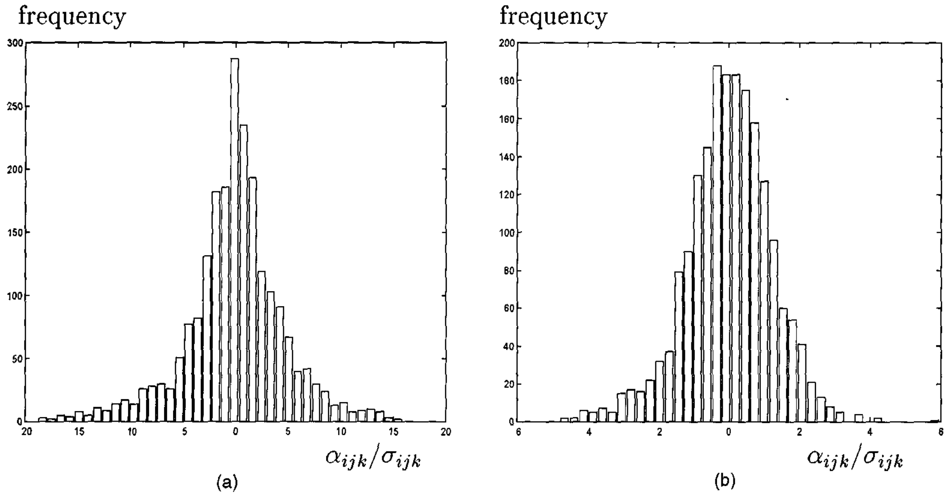


Fig. 10. Image sequence of a candle, a vase, and two 3D curves.



(a) (b)

Fig. 11. Histogram of normalized residuals for the candle and vase experiment. (a) Matching all contours. (b) Matching only convex contour segments.

space curves. Again, notice the good alignment. The reconstruction of the curves is fairly straightforward. The curve is estimated so that the reprojected errors are as small as possible. The initial estimate is relatively easy to obtain since the camera motion is known. This is known as *resection*, cf. [5], [4].

The motion estimate was also validated by trying to classify if the image curves are apparent contours or images of space curves, as described in Section 4.1. The results are presented in Table 2. The two curves show indeed low reprojected errors. For the candle and the vase, the errors are surprisingly small, but still they are far larger than for the curves. This indicates, again, the good quality of the motion estimate.

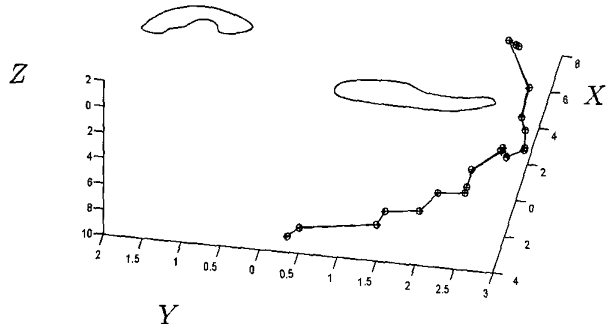


Fig. 12. Euclidean reconstruction of the space curves, the camera orbit from the apparent contours *and* and the camera orbit from the curve matches, where the camera centers are marked with crosses and rings, respectively.

4.4 Experiment 3: Statue in the Park

The final experiment presented in this paper was performed outdoors on a statue in a park. Seven images were taken by a hand-held camera from different viewpoints. Three of them are shown in Fig. 13 together with the extracted apparent contour segments.

This experiment is intended to show the applicability of the algorithm to more difficult scenarios than the previous experiments. Only a few contour segments were tracked in the image sequence and the accuracy of the extracted con-

TABLE 2
RMS OF REPROJECTION ERRORS IN PIXELS OF THE CLASSIFICATION TEST

Candle	1.60
Vase	2.02
Upper curve	0.32
Lower curve	0.39

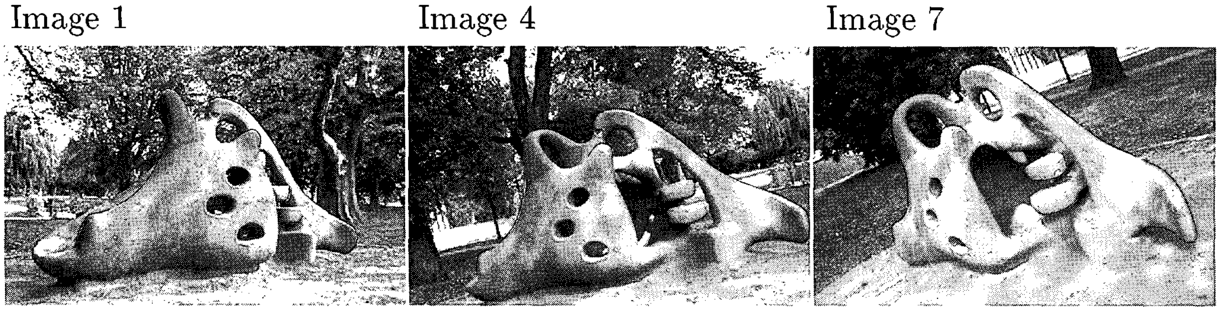


Fig. 13. Image sequence of a statue in a park. The extracted contour segments used in the experiment are marked in black.

tours is low due to bad lightning conditions and cluttered background. The depth variations of the tracked contours are small and this is known to make reconstruction harder. The radial and tangential distortion of the camera is unknown and not corrected for. Still, plausible results are obtained from the algorithm. The standard deviation of the normalized residuals (after 15 iterations) is 0.92 which is in agreement with theory. In this experiment, the number of degrees of freedom for the motion was 62 and the number of constraints was 159, i.e., roughly 7.6 constraints in average in the 21 image pairs.

In order to validate the computed camera motion, 12 corresponding points were picked out by hand in the seven images and these were, in turn, employed to compute the camera motion, without using any information of the tracked contours. A comparison of the two camera orbits is shown in Fig. 14 where camera centers are marked with (o) and (+), respectively. The alignment is not as good as in the previous experiments, but it is still satisfactory considering the accuracy of the measured quantities.

4.5 Limitations

The most notable limitation is that the generalized epipolar constraint only use information of the epipolar tangency points. For curves it is possible to get a constraint for each point of the image curve and not only at these special points. For surfaces, however, it is not obvious that there is any additional information.

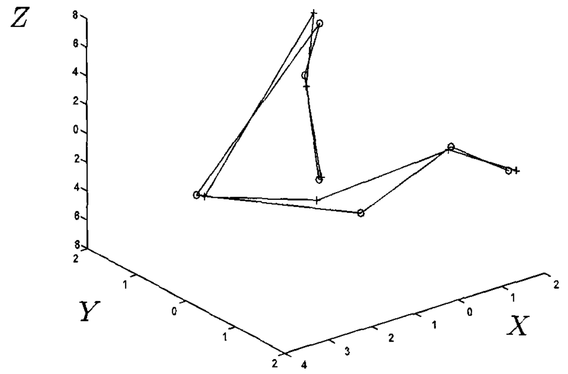


Fig. 14. Camera orbits of the motion obtained from the apparent contours, marked with crosses, and the motion obtained from 12 point matches, marked with rings, for the statue-in-park experiment.

Another practical difficulty is the problem of estimating the apparent contour. Parts of the apparent contour is impossible to see due to self-occlusions, other parts are difficult to find, particularly those close to a cusp.

As in the case of structure from motion for points there are practical difficulties with feature detection (edge detection) and tracking. We do however believe that the situation is easier for apparent contours than for points.

The main practical difficulty is, however, in finding a good initial estimate of camera motion. In our experiments this was obtained through the assumption that the images were taken in a smooth stream with small camera displacements. It is still an open question of how to obtain a good initial estimate of camera motion with a discrete set of images taken with large and random displacements between the different camera positions. As in all algorithms based on non-linear optimization there is the problem of determining whether the global minima has been found or not.

There is also the question of whether there is enough information (enough tangency points) in order to calculate camera motion, both from a theoretical and practical point of view. It is important to know how the motion estimates degrade if fewer feature points are available. This is often hard to tell in advance, since it depends on both the camera motion and the geometry of the scene. In practice one can use the singular values of the residual gradient to determine whether the constraints give enough information for motion estimation or not. If there is not sufficient number of tangency points then there will be additional zeroes among the singular values. If there is enough information but the problem is too ill-conditioned this will also be evident: Some singular values will be close to zero.

5 CONCLUSIONS

We have shown how the camera motion can be calculated from the deformation of apparent contours although nothing is known a priori about the 3D structure of the scene. The generalized epipolar constraint for a pair of images is extended to treat a sequence of images simultaneously. The proposed algorithm works with no human intervention. All steps involved have been described in detail for the case of an uncalibrated camera. The generalization to other cases is straightforward.

The concept of snakes has been employed to extract and track the apparent contours in the image sequence. As a by-product, we have solved the difficult initialization problem. We need not an initial estimate of motion obtained by other means. It is directly calculated from the extracted contours. A maximum likelihood based technique has then been used to refine the initial estimates. We have also shown, in theory, that it is possible to compute the motion with only one apparent contour in each image. By examining the dual formulation of the problem, additional insight has been gained. For example, a direct representation of the surface is obtained in the dual space.

The presented algorithm has been tested and evaluated on real image data, both in carefully controlled indoor scenes and under less favorable conditions in outdoor scenes. The calculated camera motion is verified with the independent calculation of camera motion using tracked point features and curve matches. The calculations are in close agreement. We have also been able to classify where an image curve is an apparent contour or an image of a three-dimensional curve. Unlike previous attempts that only use a pair of images, our method is robust and has good statistical properties.

ACKNOWLEDGMENTS

This work has been done within the ESPRIT Reactive LTR project 21914, CUMULI, and the Swedish Research Council for Engineering Sciences (TFR), project 95-64-222.

The authors thank Roberto Cipolla, Peter Giblin, Gunnar Sparr, Anders Heyden, and Rikard Berthilsson for valuable discussions.

REFERENCES

[1] K. Åström, R. Cipolla, and P.J. Giblin, "Generalised Epipolar Constraints," *Proc. Fourth European Conf. Computer Vision*, vol. 2, pp. 97–108, Cambridge, UK, 1996.

[2] K. Åström and A. Heyden, "Euclidean Reconstruction From Constant Intrinsic Parameters," *Proc. Int'l Conf. Pattern Recognition*, pp. 339–343, Vienna, 1996.

[3] K. Åström and A. Heyden, "Stochastic Analysis of Image Acquisition and Scale-Space Smoothing," J. Sporring, M. Nielsen, L. Florack, and P. Johansen, eds., *Gaussian Scale-Space Theory*. Kluwer Academic Publishers, 1997.

[4] K.B. Atkinson, *Close Range Photogrammetry and Machine Vision*. Whittles Publishing, 1996.

[5] R. Berthilsson and K. Åström, "Reconstruction of 3D-Curves From 2D-Images Using Affine Shape Methods for Curves," *Proc. Conf. Computer Vision and Pattern Recognition*, 1997.

[6] R. Berthilsson, K. Åström, and A. Heyden, "Projective Reconstruction of 3D-Curves From Its 2D-Images Using Error Models and Bundle Adjustments," *Proc. 10th Scandinavian Conf. Image Analysis*, pp. 581–588, Lappeenranta, Finland, 1997.

[7] A. Blake and A.L. Yuille, eds., *Active Vision*. Cambridge, Mass.: MIT Press, 1992.

[8] E. Boyer and M.-O. Berger, "3D Surface Reconstruction Using Occluding Contours," *Int'l J. Computer Vision*, vol. 22, no. 3, pp. 219–233, 1997.

[9] S. Carlsson, "Sufficient Image Structure for 3D Motion and Shape Estimation," J.-O. Eklundh, ed., *Proc. Third European Conf. Computer Vision*, vol. 1, pp. 83–91, Stockholm, 1994.

[10] R. Cipolla, *Active Visual Inference of Surface Shape*, PhD thesis, Univ. of Oxford, 1991. Also published by Springer-Verlag as LNCS 1016 (1995).

[11] R. Cipolla, K. Åström, and P.J. Giblin, "Motion From the Frontier of Curved Surfaces," *Proc. Fifth Int'l Conf. Computer Vision*, pp. 269–275, Boston, 1995.

[12] 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.

[13] H. Cramér, *Mathematical Methods of Statistics*. Princeton, N.J.: Princeton Univ. Press, 1946.

[14] R. Curwen and A. Blake, "Dynamic Contours: Real-Time Active Splines," A. Blake and A. Yuille, eds., *Active Vision*. Cambridge, Mass.: MIT Press, 1992, pp. 39–58.

[15] O.D. Faugeras, "What Can Be Seen in Three Dimensions With an Uncalibrated Stereo Rig?" G. Sandini, ed., *Proc. Second European Conf. Computer Vision*, pp. 563–578, Santa Margherita Ligure, Italy, 1992.

[16] O.D. Faugeras, Q.-T. Luong, and S. Maybank, "Camera Self-Calibration: Theory and Experiments," G. Sandini, ed., *Proc. Second European Conf. Computer Vision*, pp. 321–334, Santa Margherita Ligure, Italy, 1992.

[17] J.D. Foley, A. van Dam, S.K. Feiner, and J.F. Hughes, *Computer Graphics, Principles and Practice*. Reading, Mass.: Addison-Wesley Publishing Company, 1990.

[18] P.J. Giblin, F.E. Pollick, and J.E. Rycroft, "Recovery of an Unknown Axis or Rotation From the Profiles of a Rotating Surface," *J. Optical Soc. Amer.*, vol. 11A, pp. 1,976–1,984, 1994.

[19] P.J. Giblin and R. Weiss, "Reconstruction of Surfaces From Profiles," *Proc. First Int'l Conf. Computer Vision*, pp. 136–144, London, 1987.

[20] T. Joshi, N. Ahuja, and J. Ponce, "Structure and Motion Estimates From Dynamic Silhouettes Under Perspective Projection," *Proc. Fifth Int'l Conf. Computer Vision*, pp. 290–295, Boston, 1995.

[21] F. Kahl and K. Åström, "Motion Estimation in Image Sequences Using the Deformation of Apparent Contours," *Proc. Sixth Int'l Conf. Computer Vision*, pp. 934–944, Mumbai, India, 1998.

[22] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models," *Int'l J. Computer Vision*, vol. 1, no. 4, pp. 321–331, 1987.

[23] J. Porrill and S.B. Pollard, "Curve Matching and Stereo Calibration," *Image and Vision Computing*, vol. 9, no. 1, pp. 45–50, 1991.

[24] J.H. Rieger, "Three Dimensional Motion From Fixed Points of a Deforming Profile Curve," *Optics Letters*, vol. 11,k pp. 123–125, 1986.

[25] J.G. Semple and G.T. Kneebone, *Algebraic Projective Geometry*. Oxford, England: Clarendon Press, 1952.

[26] G. Sparr, "An Algebraic-Analytic Method for Affine Shapes of Point Configurations," *Proc. Seventh Scandinavian Conf. Image Analysis*, pp. 274–281, 1991.

[27] G. Sparr, "Simultaneous Reconstruction of Scene Structure and Camera Locations From Uncalibrated Image Sequences," *Proc. Int'l Conf. Pattern Recognition*, vol. A, pp. 328–333, Vienna, 1996.

[28] R. Szeliski and R. Weiss, "Robust Shape Recovery From Occluding Contours Using a Linear Smoother," T. Terzopoulis and C.M. Brown, eds., *Real-Time Vision*. Cambridge, England: Cambridge Univ. Press, 1994.

[29] F. Ulupinar and R. Nevatia, "Perception of 3-D Surfaces From 2-D Contours," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 1, pp. 3–18, Jan. 1993.

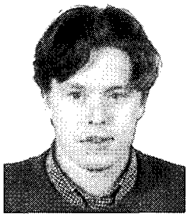
[30] R. Vaillant and O.D. Faugeras, "Using Extremal Boundaries for 3D Object Modelling," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 157–173, Feb. 1992.

[31] B. Vijayakumar, D. Kriegman, and J. Ponce, "Invariant-Based Recognition of Complex 3D Curved Objects From Image Contours," *Proc. Fifth Int'l Conf. Computer Vision*, pp. 508–514, Boston, 1995.

[32] B. Vijayakumar, D. Kriegman, and J. Ponce, "Structure and Motion of Curved 3D Objects From Monocular Silhouettes," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 327–334, 1996.

[33] Z. Zhang, "Determining the Epipolar Geometry and Its Uncertainty: A Review," *Int'l J. Computer Vision*, vol. 27, no. 2, pp. 161–195, 1998.

Kalle Åström received his MSc degree in engineering physics in 1991 and his PhD in mathematics in 1996 from Lund University, Sweden. His thesis was awarded Best Nordic PhD Thesis in pattern recognition and image analysis 1995–1996 at the *Scandinavian Conference in Image Analysis*, 1997. He has been a postdoctoral research fellow and subsequently research associate at the Centre for Mathematical Sciences, Lund University. He is the president of the Swedish Society for Automated Image Analysis. His teaching includes undergraduate and graduate courses in mathematics, image processing, and computer vision. His current research interests include stochastic analysis of low-level vision, computer vision for autonomous guided vehicles, and the geometry and algebra of multiple views of points, curves, and surfaces.



Fredrik Kahl received his MSc degree in computer science and technology in 1995 from Lund University, Sweden. He is currently pursuing a PhD degree in applied mathematics at the Centre for Mathematical Sciences, Lund University. He is the treasurer of the Swedish Society for Automated Image Analysis. His main research interest is computer vision, with emphasis on projective, affine, and Euclidean geometry, autocalibration, and image sequence analysis.