Generalised Epipolar Constraints

Kalle Åström[1] and Roberto Cipolla[2] and Peter J. Giblin[3]

[1] Dept of Mathematics, Lund University, Lund, Sweden
[2] Dept of Engineering, Univ. of Cambridge, Cambridge, UK
[3] Dept of Pure Mathematics, Univ. of Liverpool, Liverpool, UK

Abstract. The *frontier* of a curved surface is the envelope of contour generators showing the boundary, at least locally, of the visible region swept out under viewer motion. In general, the outlines of curved surfaces (apparent contours) from different viewpoints are generated by different contour generators on the surface and hence do not provide a constraint on viewer motion. Frontier points, however, have projections which correspond to a real point on the surface and can be used to constrain viewer motion by the epipolar constraint.

We show how to recover viewer motion from frontier points and generalise the ordinary epipolar constraint to deal with points, curves and apparent contours of surfaces. This is done for both continuous and discrete motion, known or unknown orientation, calibrated and uncalibrated, perspective, weak perspective and orthographic cameras. Results of an iterative scheme to recover the epipolar line structure from real image sequences using only the outlines of curved surfaces, is presented. A statistical evaluation is performed to estimate the stability of the solution. It is also shown how the full motion of the camera from a sequence of images can be obtained from the relative motion between image pairs.

1 Introduction

Structure and motion from the images of point features has attracted considerable attention and a large number of algorithms exist to recover both the spatial configuration of the points and the motion compatible with the views. Structure and motion from the outlines of curved surfaces, on the other hand, has been thought to be more difficult because of the *aperture problem*, i.e. it is not possible to get the correspondence of points between two images of the same curve.

For a smooth arbitrary curved surface an important image feature is the outline or apparent contour. This is the projection of the locus of points on the surface which separates the visible from the occluded parts (Fig. 1.A). Under perspective projection this locus – the critical set or contour generator, Σ – can be constructed as the set of points on the surface where rays through the projection centre $\mathbf{c}$ are tangent to the surface. Each viewpoint will generate a different contour generator with the contour generators 'slipping' over the visible surface under viewer motion (Fig. 1.B).

Under *known* viewer motion, the deformation of apparent contours can be used to recover the surface geometry (structure) [8, 5, 14]. This requires a spatio–

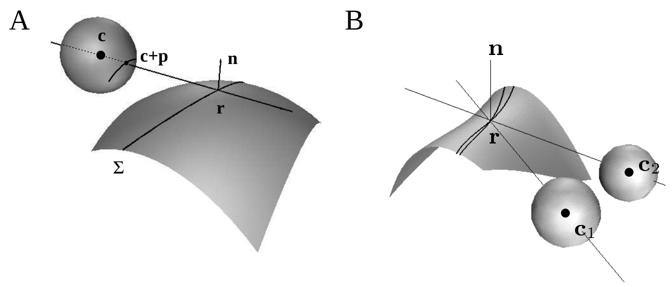

Fig. 1. Perspective projection (A): the contour generator Σ with a typical point $\mathbf{r}$, the image sphere with centre $\mathbf{c}$ and the corresponding apparent contour point $\mathbf{c}+\mathbf{p}$. Thus $\mathbf{p}$ is the unit vector joining the centre $\mathbf{c}$ to the apparent contour point. Degenerate case of epipolar parameterisation (B): The epipolar plane is a tangent plane of the surface at a frontier point. At a frontier point the contour generators from consecutive viewpoints intersect.

temporal parametrisation of the image-curve motion. The *epipolar* parametrisation is most naturally matched to the recovery of surface curvature, cf. [5].

In this paper we address the problem of recovering the viewer motion from the deformation of apparent contours. We show how frontier points can be detected in image sequences and used to recover viewer egomotion. The special case of frontier points under orthographic projection and object rotation about a single axis was considered by Rieger [13] and Giblin et al [7]. Porrill and Pollard [12], although primarily concerned with stereo calibration from 3D space curves, noted that the intersection of the two contour generators from two discrete viewpoints generated a real point visible in both images which could also be used to generate an epipolar constraint. This constraint was exploited by Carlsson [3] in the analysis of the visual motion of space-curves.

We analyse both the continuous (infinitesimal) and discrete viewer motion cases as well as considering calibrated and uncalibrated cameras, orthographic and perspective projection. We present preliminary experimental results obtained from real image sequences of curved surfaces from unknown viewpoints. An iterative technique is implemented which recovers the maximum likelihood estimate of the epipolar line structure (*fundamental and essential matrices*) from the image motion of frontier points. A statistical evaluation is also performed to estimate the stability of the solution. The work is a continuation of [4].

2 Generalised Epipolar Constraints

Consider a smooth surface M (closed, for example the boundary of a 3-dimensional object) which is viewed from a curve $\mathbf{c}(t)$ of camera centres. Each camera centre $\mathbf{c}(t)$ gives rise to a contour generator $\Gamma(t)$ consisting of points $\mathbf{r}$ in M

such that the viewline is perpendicular to the surface normal: $(\mathbf{r}-\mathbf{c}).\mathbf{n}=0$. As t changes the contour generator slides over the surface.

Now consider the camera center at two time instants $\mathbf{c}_1=\mathbf{c}(t_1)$ and $\mathbf{c}_2=\mathbf{c}(t_2)$, and consider all planes that go through these two camera centers which are tangent to the surface M. This pencil of planes will be called the pencil of *epipolar tangency planes* with respect to the camera motion. The image of the epipolar tangency planes is a star of lines through a point, the so called *epipole*, each line being tangent to an apparent contour. Since both pencils are the image of the same pencil of planes they must be related. Denote by $\mathbf{p}_1(s)$ the apparent contour in image 1 with curve parameter s and by $\mathbf{p}_2(s)$ corresponding apparent contour in image 2. The point $\mathbf{r}$ where the epipolar tangency point is tangent to the surface M must be on both contour generators as is illustrated in Fig. 1.B. The normal $\mathbf{n}$ to the surface at this point is orthogonal not only to $\mathbf{p}_1$ and $\mathbf{p}_2$ but also to their image tangents $(\mathbf{p}_1)'_s$ and $(\mathbf{p}_2)'_s$ and to the direction of motion $\Delta\mathbf{c}=\mathbf{c}_2-\mathbf{c}_1$. This can be written in a compact way as the *generalised epipolar constraints*,

$$\operatorname{rank}\left[\Delta\mathbf{c}\ \ \mathbf{p}_1(s_1)\ \ (\mathbf{p}_1)'_s(s_1)\ \ \mathbf{p}_2(s_2)\ \ (\mathbf{p}_2)'_s(s_2)\right]=2\ , \tag{1}$$

i.e. the five column vectors in the above matrix must lie in a plane perpendicular to the normal. Notice that the rank constraints involve both curve parameters s_1 and s_2 and motion parameters $\Delta\mathbf{c}$. This can be used in several ways. Firstly, once the direction of motion $\Delta\mathbf{c}$ is known or guessed the first image of the epipolar tangency point can be found by searching for the curve parameter s_1 so that $\det\left[\Delta\mathbf{c}\ \mathbf{p}_1(s_1)\ (\mathbf{p}_1)'_s(s_1)\right]=0$, and similarly for the second image. Secondly, when the image of the epipolar tangency points has been found it can be checked whether or not $\det\left[\Delta\mathbf{c}\ \mathbf{p}_1\ \mathbf{p}_2\right]=0$.

The above formulation of the generalised epipolar constraints depends on the specific choice of both object and image coordinate system. The motion parameters can therefore only be determined up to an unknown choice of coordinate system. The above formulation is convenient since it makes the generalisation to other camera models easy. As an example the infinitesimal version can be obtained as the limit when the time difference approaches zero.

$$\begin{aligned}\operatorname{rank}\left[\Delta\mathbf{c}\ \mathbf{p}_1\ (\mathbf{p}_1)_s\ \mathbf{p}_2\ (\mathbf{p}_2)_s\right] &= \operatorname{rank}\left[\Delta\mathbf{c}/\Delta t\ \ (\mathbf{p}_1)_s\ (\mathbf{p}_2-\mathbf{p}_1)/\Delta t\ (\mathbf{p}_2)_s\right]\rightarrow\\ &\rightarrow \operatorname{rank}\left[\mathbf{c}_t\ \mathbf{p}\ (\mathbf{p})_s\ \mathbf{p}_t\ (\mathbf{p})_s\right]=2, \qquad \text{as}\ \ t\rightarrow 0\ .\end{aligned}$$

Note that $\mathbf{p}$ is the orientation of the ray in the fixed reference/world frame for $\mathbb{R}^3$. For a *calibrated* camera it is determined by a spherical image position vector $\mathbf{q}$ (the orientation of the ray in a coordinate system attached to the viewer or camera) and the orientation of the camera co-ordinate system relative to the reference frame. For a moving observer the viewer coordinate system is continuously moving with respect to the reference frame. The relationship between $\mathbf{p}$ and $\mathbf{q}$ can be conveniently expressed in terms of a rotation operator, $\mathbf{R}(t)$, such that $\mathbf{p}=\mathbf{R}(t)\mathbf{q}$. The measurements in an *uncalibrated* camera, $\mathbf{w}$, are related to the spherical image position, $\mathbf{q}$, by an intrinsic calibration matrix (affine transformation), $\mathbf{A}$, such that $\mathbf{q}=\mathbf{A}(t)\mathbf{w}$. For simplicity the relationship between $\mathbf{w}$

and $\mathbf{p}$ will be expressed by a single matrix $\mathbf{S}$ representing both intrinsic calibration and orientation of the camera, i.e. $\mathbf{p} = \mathbf{R}(t)\mathbf{A}(t)\mathbf{w} = \mathbf{S}(t)\mathbf{w}$. A thorough treatment of the other cases can be found in [1]. The results are summarised in the following two tables. The direction of motion is represented with a three vector, $\Delta\mathbf{c}$, in the central projection cases and $\Delta\mathbf{k} = (\cos(\theta), \sin(\theta), 0)^T$ in the parallel projection case. In the tables $\Delta\mathbf{R}$ is a three by three rotation matrix, $\Delta\mathbf{S}$ is a general non-singular matrix, $\Delta\mathbf{B}$ is a matrix representing planar euclidean transformations and $\Delta\mathbf{C}$ is a matrix representing planar similarity transformations. The matrix $\mathbf{R}_t$ is the derivative of a rotation matrix and similarly for the other types.

Camera Model	Motion Params	Obs. d o f	Generalised Epipolar Constraints
Pure transl.	$\Delta\mathbf{c}$	2	rank $\left[\Delta\mathbf{c}\ \mathbf{p}_1\ (\mathbf{p}_1)_s\ \mathbf{p}_2\ (\mathbf{p}_2)_s\right] = 2$
Calibrated	$\Delta\mathbf{c}, \Delta\mathbf{R}$	5	rank $\left[\Delta\mathbf{c}\ \mathbf{q}_1\ (\mathbf{q}_1)_s\ \Delta\mathbf{R}\mathbf{q}_2\ \Delta\mathbf{R}(\mathbf{q}_2)_s\right] = 2$
Uncalibrated	$\Delta\mathbf{c}, \Delta\mathbf{S}$	7	rank $\left[\Delta\mathbf{c}\ \mathbf{w}_1\ (\mathbf{w}_1)_s\ \Delta\mathbf{S}\mathbf{w}_2\ \Delta\mathbf{S}(\mathbf{w}_2)_s\right] = 2$
Orthographic	$\Delta\mathbf{k}, \Delta\mathbf{B}$	3	rank $\left[\Delta\mathbf{k}\ \mathbf{m}\ (\mathbf{w}_1)_s\ (\Delta\mathbf{B}\mathbf{w}_2 - \mathbf{w}_1)\ \Delta\mathbf{B}(\mathbf{w}_2)_s\right] = 2$
Weak Persp.	$\Delta\mathbf{k}, \Delta\mathbf{C}$	4	rank $\left[\Delta\mathbf{k}\ \mathbf{m}\ (\mathbf{w}_1)_s\ (\Delta\mathbf{C}\mathbf{w}_2 - \mathbf{w}_1)\ \Delta\mathbf{C}(\mathbf{w}_2)_s\right] = 2$

Summary of relevant motion parameters, number of observable degrees of freedom and generalised epipolar constraints in the discrete case.

Camera model	Motion Params	Obs. d o f	Generalised Epipolar Constraints
Pure transl.	$\mathbf{c}_t$	2	rank $\left[\mathbf{c}_t\ \mathbf{p}\ \mathbf{p}_s\ \mathbf{p}_t\right] = 2$
Calibrated	$\mathbf{c}_t, \mathbf{R}_t$	5	rank $\left[\mathbf{c}_t\ \mathbf{q}\ \mathbf{q}_s\ \mathbf{R}_t\mathbf{q} + \mathbf{q}_t\right] = 2$
Uncalibrated	$\mathbf{c}_t, \mathbf{S}_t$	7	rank $\left[\mathbf{c}_t\ \mathbf{w}\ \mathbf{w}_s\ \mathbf{S}_t\mathbf{w} + \mathbf{w}_t\right] = 2$
Orthographic	$\mathbf{k}_t, \mathbf{B}_t$	3	rank $\left[\mathbf{k}_t\ \mathbf{m}\ \mathbf{w}_s\ \mathbf{B}_t\mathbf{w} + \mathbf{w}_t\right] = 2$
Weak Persp.	$\mathbf{k}_t, \mathbf{C}_t$	4	rank $\left[\mathbf{k}_t\ \mathbf{m}\ \mathbf{w}_s\ \mathbf{C}_t\mathbf{w} + \mathbf{w}_t\right] = 2$

Summary of relevant motion parameters, number of observable degrees of freedom and generalised epipolar constraints in the infinitesimal case.

3 Implementation

3.1 Extraction and Tracking of Apparent Contours

An important aspect in calculating motion from the deformation of apparent contours is the actual extraction and tracking of the contours. This can be achieved with B-spline snakes. In our implementation, the snake at time t_1 is used as a template to find corresponding contour in the image at the next time instant t_2. At first the snake is only allowed to move rigidly. This ensures a fast, robust, but rough positioning of the snake in the new image. The snake is then allowed to deform to match the new image. This procedure is explained in more detail in [6]. In the last steps of snake deformation we have used subpixel edge detectors that not only gives us the location of the contour but also a confidence interval in the normal direction of the curve. This is done with a new technique described in [2]. For clear well defined edges, the individual edge positions can be found with a standard deviation of roughly 1/5:th of a pixel in the normal direction. This uncertainty measure is important in estimating motion parameters later. Different frontier points are weighted according to the uncertainty in their positioning.

As a by-product of the snake type tracking, a rough guess of point correspondences are obtained. These approximate point correspondences can be used to calculate an initial estimate of motion parameters as described in the next section.

3.2 Initial Hypothesis of Motion

There are a number of different ways to obtain the initial hypothesis of the motion parameters, which are needed in order to use the generalised epipolar constraints.

Point matches can be used to estimate motion parameters using the linear eight point method [10] or non-linear methods [11].

Motion sensors: In some situations, partial knowledge of the motion can be obtained by other means, e.g. from motion sensors.

Prediction: If viewer motion is smooth it might be possible to predict motion parameters from estimates of motion history.

Simpler camera models: Approximate motion parameters can also be obtained using approximate camera models, e.g. the weak perspective model.

3.3 Maximum Likelihood Estimation

The maximum likelihood method is a natural way to estimate the motion parameters given noisy input data. It has several advantages and is relatively easy to apply. The main idea is the following. A residual, $\alpha_i = \alpha_i(m)$, as a function of motion parameters, m, is chosen. The joint distribution, $f_n(\alpha|m)$, of the residuals given motion parameters, m, is calculated. The maximum likelihood estimate, $\hat{m}$, is the motion parameter which maximises the likelihood function $L(m) = f_n(\alpha|m)$. To make this optimisation simple it is often assumed that the residuals, α_i, are independent and of Gaussian distribution with zero mean and standard deviation σ_i. This is a reasonable assumption if the images of the frontier points are not too close to each other in the image. The likelihood function is then

$$L = \prod \frac{1}{\sqrt{2\pi\sigma_i^2}} e^{-\alpha_i^2/2\sigma_i^2} \quad .$$

By taking the negative logarithm maximising the likelihood is approximately the same as minimising

$$g = \sum \frac{\alpha_i^2}{\sigma_i^2} \quad .$$

The estimate is the motion parameters that minimise this weighted sum of squared residuals. In our implementations minimisation is performed using either a modified Newton-Raphson method or Gauss-Newton method. For more details see [1].

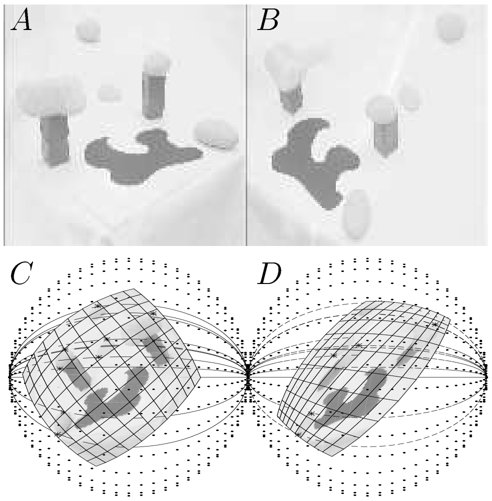

Fig. 2. Rectification of uncalibrated images. Two images (A-B) are projected onto the viewing sphere and rectified using the motion parameters (C-D). After rectification The epipolar tangency planes all intersect at the x-axis. The two sets of epipolar tangency planes should be equal. The angular difference is used as the residual.

Residuals in the discrete time, perspective projection case: After a standard rectification to parallel geometry the set of epipolar tangency planes in both images should be identical. The epipolar tangency points and the epipolar tangency planes are found using the epipolar tangency constraints

$$\left|\Delta\mathbf{c}\ \mathbf{w}_1\ (\mathbf{w}_1)_s\right| = 0\ , \tag{2}$$

$$\left|\Delta\mathbf{c}\ \Delta\mathbf{S}\mathbf{w}_2\ \Delta\mathbf{S}(\mathbf{w}_2)_s\right| = 0\ . \tag{3}$$

The angular difference, α_i, cf. Fig. 3, between the two representations of the same epipolar tangency planes, after rectification, is calculated as well as the standard deviation σ_i of this residual.

Residuals in the discrete time, parallel projection case: The orthographic and weak perspective cases are similar. After rectification the two sets of parallel epipolar tangency planes should be identical. The distance α_i, between the parallel epipolar tangency planes, is used as the residual. The residual is scaled with respect to its standard deviation σ_i. The residual variance, due to edge localisation error, is changed in these transformations. These effects are taken into account.

Residuals in the infinitesimal case: In the infinitesimal case, the direction of viewer motion $\mathbf{c}_t$ is used as an infinitesimal epipole or the focus of expansion.

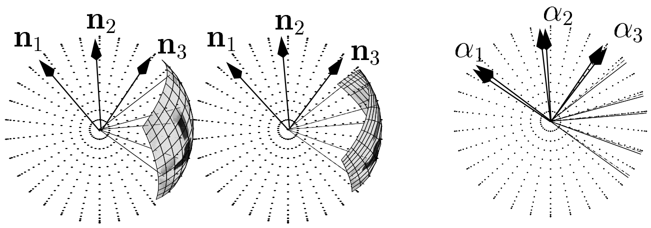

Fig. 3. Two sets of epipolar tangency planes are calculated from two images. These two sets should ideally be identical. The residual is defined as the angular difference α_i between corresponding epipolar tangency planes after rectification.

The tangency constraint is then used to find the epipolar tangency planes and corresponding frontier points. For example, in the calibrated case we have

$$\left| \mathbf{c}_t \; \mathbf{q} \; \mathbf{q}_s \right| = 0 \ .$$

Each plane defines a normal $\mathbf{n} = \mathbf{q} \times \mathbf{c}_t$. The motion constraint is then simply $\mathbf{n} \cdot (\mathbf{R}_t \mathbf{q} + \mathbf{q}_t) = 0$. It is reasonable to use

$$\alpha = \mathbf{n} \cdot (\mathbf{R}_t \mathbf{q} + \mathbf{q}_t)$$

as the residual. Errors in α are mostly due to the errors in $\mathbf{q}_t$, therefore the standard deviation, $\sigma[\alpha]$, can be approximated by $\sigma[\mathbf{n}_i \cdot \mathbf{q}_t]$, i.e. the standard deviation in estimated normal velocity. These standard deviations are obtained from the sub-pixel edge detector routines. The standard deviation is approximately constant around each frontier point, so g is in fact quadratic in $\mathbf{R}_t$ so that the minimum with respect to $\mathbf{R}_t$ can be found with linear methods. The same argument applies to the uncalibrated and parallel projection cases.

3.4 Statistical Evaluation

The maximum likelihood estimate has several good properties. One is that it is guaranteed to be asymptotically unbiased.

The residuals at the minimum can be used to estimate empirically the magnitude of edge localisation error. This can then be compared to the ones obtained from the edge detectors. The residuals can thus be used to automatically verify whether a reasonably low minimum has been found. Thus it may be possible to get out of local minima and also to remove outliers.

The second derivative matrix of g together with the variance of the scaled residuals gives us an estimate of the covariance of the estimated motion parameters, cf. [1].

4 Examples

4.1 Infinitesimal Motion

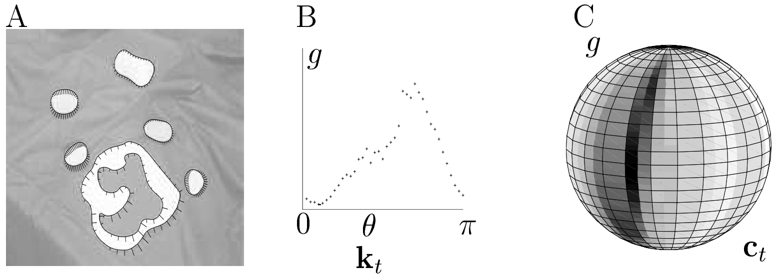

Fig. 4. Infinitesimal motion with a weak perspective camera. One image with apparent contours and the estimated normal velocity is used as input (A). The minimal error function g as a function of the tessalated line at infinity using the weak perspective camera model (B) and as a function of the sphere of directions using the uncalibrated camera model (C). Low values are dark in the figure.

One image, taken from a longer sequence, of the same scene is shown in Fig. 4.A. The deformation between this and the next image is used to estimate normal velocity. This is illustrated with small line segments.

For each of forty choices of focus of expansion $\mathbf{k}_t = [\cos(\theta), \sin(\theta), 0]^T$ with $\theta = 0, \pi/40, \ldots, 39\pi/40$, the epipolar tangency points $\mathbf{w}$ were calculated using the epipolar tangency constraint $\left|\mathbf{k}_t\ \mathbf{m}\ \mathbf{w}_s\right| = 0$. The normal velocity and its standard deviation was calculated at these points. For each $\mathbf{k}_t$ the weighted squared residual $g = \sum \alpha_i^2/\sigma_i^2$ with $\alpha = \mathbf{n} \cdot (\mathbf{B}_t\mathbf{w} + \mathbf{w}_t)$ was minimised with respect to $\mathbf{B}_t$. The minimal residual as a function of θ is shown in Fig. 4.B. Thus a rough estimate of the motion is obtained by finding the minimal g of these forty directions.

The same input was also used to illustrate the uncalibrated camera case. The same idea of tessellating the focus of expansion can be used. In the perspective projection case this means tessellating the sphere of directions. As in the previous example, for each choice of focus of expansion or direction of motion, $\mathbf{c}_t$, it is straightforward to find the epipolar tangency points. Minimising the weighted sum of squared residuals g with respect to $\mathbf{S}_t$ is a linear problem. The minimal residual as a function of $\mathbf{c}_t \in \mathbb{S}^2$ is shown in Fig. 4.C. Notice that the low values of, g, form a long dark valley on the sphere. We expect the direction of motion to be poorly located along that valley. This is confirmed by the statistical evaluation. Notice also that choosing the weak perspective model corresponds to searching this sphere along the equator only.

The minimum obtained from tessellating the sphere or the minimum obtained from the weak perspective case above can both be used as initial estimates in a

Gauss-Newton search of the minimum. This was done and about ten iterations were needed to find the minimum.

4.2 Discrete Motion, Known Rotation

Known rotation can easily be illustrated by incorporating and using a planar curve in the image. The image of a planar curve is deformed with a planar projective transformation. By detecting and aligning a planar feature in a sequence of images, *the sequence can be regarded as one of a purely translating camera.* The planar curve is then regarded as a curve on the plane at infinity and thus it has no apparent image motion. This idea has been used in [9] and is known as *projective reduction*, a generalisation of the *plane plus parallax method.*

This is illustrated in Fig. 6. Only the direction of motion $\Delta\mathbf{c}$ needs to be estimated. The sphere of possible directions can then be tessellated and the error function g can be calculated for each direction. The minimum obtained after tessellation is improved by local Newton-Raphson search (6 iterations were needed).

4.3 Discrete Motion, Uncalibrated Camera

The discrete motion case with uncalibrated camera is illustrated with a pair of images from the same sequence that was used to illustrate the infinitesimal case. The result from the infinitesimal case can be used as an initial estimates of the discrete motion parameters. A standard extrapolation is used:

$$\Delta\mathbf{S} = e^{\mathbf{S}_t \Delta t}, \qquad \Delta\mathbf{c} = \mathbf{c}_t \Delta t$$

This initial estimate is used as input in a gauss-newton search. At each iteration we use the current motion parameters to calculate the epipoles and the epipolar tangencies. Fig. 5.A and B illustrate the epipolar tangencies obtained at iteration 1. The two sets of epipolar tangency planes are then rectified using the current motion parameters, cf. Fig. 2. After rectification corresponding epipolar tangency planes should coincide. The difference, measured as the angle α is calculated. The estimated standard deviation σ of this angle due to edge localisation errors is also calculated and the weighted residual $Y = \alpha/\sigma$ is formed. The derivative of Y with respect to infinitesimal changes in motion parameters is then calculated. The weighted residual Y and its derivatives are then used to adjust the motion parameters according to the Gauss-Newton method, cf. [1]. Fig. 5 illustrates the epipolar tangencies and the rectified epipolar tangency planes of iteration 1, 4 and 10 of such a minimisation. Typically one needs a couple of iterations (6 in this example) to get close to the minima. During these iteration the error function $g = \sum Y^2$ decreases rapidly. Then a few more iterations are needed to localise the minima within machine accuracy. During these last iteration the norm of the gradient decreases rapidly, while the error function stays almost constant. This is illustrated in the table below, in which the error function g and the logarithm of the norm of the gradient of g is shown for the 10 iterations.

Iteration	1	2	3	4	5	6	7	8	9	10
g	110.5	66.54	34.24	16.45	7.695	6.482	6.476	6.476	6.476	6.476
$\log_{10}(\lvert\nabla g\rvert)$	-0.99	-1.11	-1.28	-1.48	-2.01	-3.63	-5.53	-7.62	-9.70	-9.70

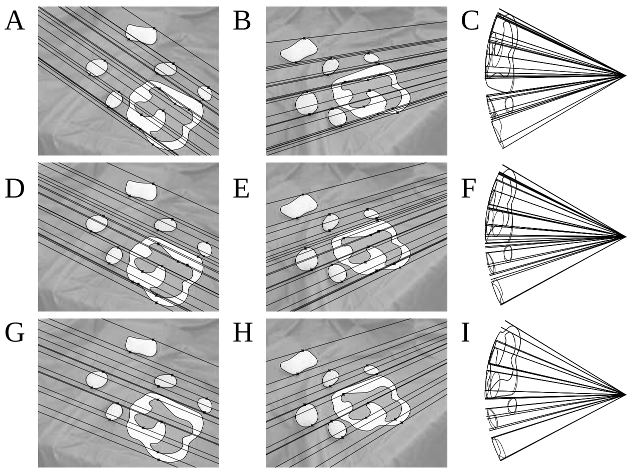

Fig. 5. Finding the uncalibrated motion parameters using the generalised epipolar constraints. Optimisation of the likelihood function. Iteration number 1 (first row), 4 (second row) and 10 (Third row). First and second columns illustrate the epipolar tangencies in the first and second image. The third column illustrate the rectified epipolar tangency planes, projected on the view-sphere and viewed along the direction of motion..

5 Extension to Image Sequences

The implementation briefly described above allows us to calculate motion parameters between pairs of images. A natural extension is to use the estimates of motion parameters for pairs of images in a sequence of images at times $(t_0, \ldots, t_n)$, to obtain the full motion of the camera. In the calibrated case the full motion of the camera as represented by camera positions $\mathbf{c}_i = \mathbf{c}(t_i)$ and camera orientation $\mathbf{R}_i = \mathbf{R}(t_i)$ must fulfill the following equations:

$$\mu_{ij} \Delta\mathbf{c}_{ij} = \mathbf{R}_i^{-1}(\mathbf{c}_j - \mathbf{c}_i)$$

$$\Delta\mathbf{R}_{ij} = \mathbf{R}_i^{-1}\mathbf{R}_j \ .$$

where $(\Delta\mathbf{c}_{ij}, \Delta\mathbf{R}_{ij})$ are the motion parameters from image i to image j. The overall coordinate system must be chosen, e.g. by choosing $\mathbf{c}_0 = 0$, $\mathbf{R}_0 = I$ and

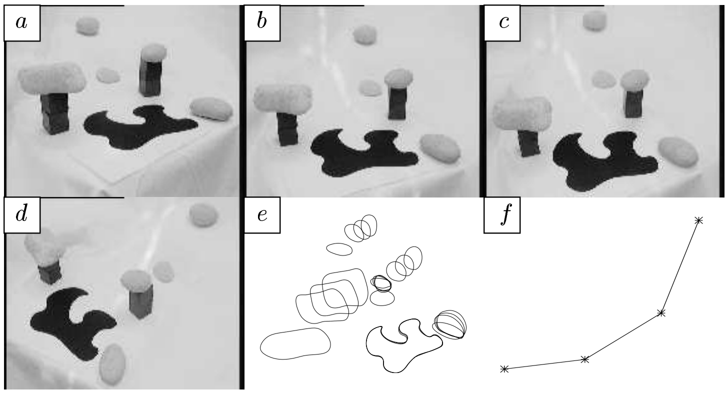

Fig. 6. The case of discrete motion with uncalibrated camera. Four images out of a longer sequence is shown in (A-D). By detecting and aligning the image of a planar feature the images can be thought of as coming from a purely translating camera. The apparent contours after alignement are shown in (E). This makes it relatively easy to extract motion parameter between each pair of images. These parameters can then be used to calculate the full motion of the camera (F).

$|\mathbf{c}_n| = 1$. The unknown scale factor μ_{ij} has to be found since it is only possible to determine $\Delta\mathbf{c}_{ij}$ up to an unknown scale factor. Similar equations apply to the other camera models. The idea is illustrated in Fig. 6. Four images of a short sequence is shown (A-D) and the camera motion are represented as the corresponding four camera positions (F).

6 Conclusions and Future Work

The apparent contour and its deformation under viewer motion is known to be a rich source of surface geometric information which can be used in visual navigation and object manipulation. Here we have shown how so called frontier points of apparent contours can be used to recover the viewer motion from the deformation of apparent contours. The epipolar constraint for points is generalised to points, curves and apparent contours, to both the continuous and discrete motion cases, to uncalibrated and calibrated cameras and to perspective and parallel camera models. An iterative method to obtain the maximum likelihood estimate of the motion parameters is presented and the problem of obtaining initial estimates is discussed. Statistical evaluation of the results are presented. These can be used to evaluate the validity of the solution but also to obtain estimates of the covariance of the estimated motion parameters. The theory is applied to real image sequences. It is also shown how motion between image pairs can be used to obtain full camera motion.

References

1. K. Åström, R. Cipolla, and P. J. Giblin. Generalised epipolar constraints. Technical report, Dept. of Mathematics, Lund University, 1996.
2. K. Åström and A. Heyden. Stochastic modelling and analysis of image acquisition and sub-pixel edge detection. Technical report, Dept. of Mathematics, Lund University, 1995.
3. S. Carlsson. Sufficient image structure for 3D motion and shape estimation. In J-O. Eklundh, editor, *Proc. ECCV'94*, volume I, pages 83–91. Springer–Verlag, 1994.
4. R. Cipolla, K. Åström, and P.J̃. Giblin. Motion from the frontier of curved surfaces. In *Proc. ICCV'95*, pages 269–275, 1995.
5. R. Cipolla and A. Blake. Surface shape from the deformation of apparent contours. *Int. Journal of Computer Vision*, 9(2):83–112, 1992.
6. R. Curwen and A. Blake. Dynamic contours: real-time active splines. In A. Blake and A. Yuille, editors, *Active Vision*. MIT Press, 1992.
7. 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. Opt. Soc. America*, A11:1976–1984, 1994.
8. P.J. Giblin and R. Weiss. Reconstruction of surfaces from profiles. In *Proc. ICCV'87*, pages 136–144, London, 1987.
9. A. Heyden and K. Åström. A canonical framework for sequences of images. In *Proc. IEEE workshop on Representation of Visual Scenes, MIT, USA*, 1995.
10. H.C. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293:133–135, 1981.
11. Q-T. Luong, R. Deriche, O. Faugeras, and T. Papadopoulo. On determining the fundamental matrix: analysis of different methods and experimental results. In *Israelian Conf. on Artificial Intelligence and Computer Vision*, Tel-Aviv, Israel, 1993. A longer version is INRIA Tech Report RR-1894.
12. J. Porrill and S.B. Pollard. Curve matching and stereo calibration. *Image and Vision Computing*, 9(1):45–50, 1991.
13. J.H. Rieger. Three dimensional motion from fixed points of a deforming profile curve. *Optics Letters*, 11:123–125, 1986.
14. R. Vaillant and O.D. Faugeras. Using extremal boundaries for 3D object modelling. *IEEE Trans. Pattern Analysis and Machine Intell.*, 14(2):157–173, 1992.

This article was processed using the LaTeX macro package with ECCV'96 style