Parallax geometry of smooth surfaces in multiple views

Geoff Cross Andrew W. Fitzgibbon Andrew Zisserman

Department of Engineering Science
University of Oxford
Oxford, OX1 3PJ, UK
{geoff,awf,az}@robots.ox.ac.uk
http://www.robots.ox.ac.uk/∼vgg

Abstract

This paper investigates the multiple view geometry of smooth surfaces and a plane, where the plane provides a planar homography mapping between the views. Innovations are made in three areas: first, new solutions are given for the computation of epipolar and trifocal geometry for this type of scene. In particular it is shown that the epipole may be determined from bitangents between the homography registered occluding contours, and a new minimal solution is given for computing the trifocal tensor. Second, algorithms are demonstrated for automatically estimating the fundamental matrix and trifocal tensor from images of such scenes. Third, a method is developed for estimating camera matrices for a sequence of images of these scenes.

These three areas are combined in a "freehand scanner" application where 3D texture-mapped graphical models of smooth objects are acquired directly from a video sequence of the object and plane.

1 Introduction

The computation of epipolar geometry and trifocal geometry from image point correspondences is now well established [10, 27, 29]. The case of scenes containing smooth curves and objects, where point correspondences are not available, is less thoroughly investigated although iterative search based algorithms are available for two views of smooth objects [6] (reviewed below) and for multiple views of a space curve [4] or surface [1]. However, as yet there are no closed form solutions in the case of smooth surfaces.

In this paper we investigate scenes containing smooth objects *and* a plane which may be used to register the views. The plane plus *points* configuration has received significant attention in the past, not least because it arises frequently in everyday scenes. The image motion is decomposed into a planar homographic transfer plus a residual image parallax vector [12, 14, 21, 23]. This decomposition has the advantage that it partially factors out dependence on the camera relative rotation and internal parameters. In this paper we show that similar advantages apply in the plane plus smooth object case. Furthermore there are additional advantages in that the multiple view geometry can be computed for smooth objects without entailing an iterative search.

For the plane plus smooth object configuration we derive new solutions and uniqueness results for the computation of the fundamental matrix (section 3), the trifocal tensor [11, 22, 24] (section 4) and cameras for multiple (i.e. more than 3) views (section 5). We then describe an application of these results to building graphical models of smooth objects from a sequence of images.

1.1 Planar parallax

The underlying parallax geometry is shown in figure 1. The world plane induces a homography $\mathtt{H}$ between the two images, so that the image of points on the plane are mapped as $\mathbf{x}' = \mathtt{H}\mathbf{x}$, where $\mathbf{x}$ and $\mathbf{x}'$ are the image points in the first and second views respectively. The homography can be determined from a minimum of four correspondences over the two views of points (or lines) on the plane [17]. Then the fundamental matrix for the views is given by [16]

$$\mathtt{F} = \mathtt{H}^{-\top}[\mathbf{e}]_\times = [\mathbf{e}']_\times \mathtt{H} \tag{1}$$

where $\mathbf{e}'$ is the epipole in the second image and $[\mathbf{e}']_\times$ is the 3×3 skew matrix for which $[\mathbf{e}']_\times \mathbf{x} = \mathbf{e}' \times \mathbf{x}$. Thus, in order to determine the fundamental matrix it is only necessary to determine the position of the epipole, i.e. only two parameters need be specified.

The parallax vector in the second view is the vector joining the image of a world point $\mathbf{x}'$ with the transferred image of that point from the first view $\tilde{\mathbf{x}}' = \mathtt{H}\mathbf{x}$. The line joining $\mathbf{x}'$ and $\tilde{\mathbf{x}}'$ contains the epipole $\mathbf{e}'$, so two such lines are sufficient to compute its position, and the full epipolar geometry follows from (1). The magnitude of the parallax vector is related to the distance of the world point and cameras from the world plane [12, 14, 21].

2 Parallax geometry for smooth surfaces

As illustrated in figure 2, at points at which the epipolar plane is tangent to a surface, the occluding contour is tan-

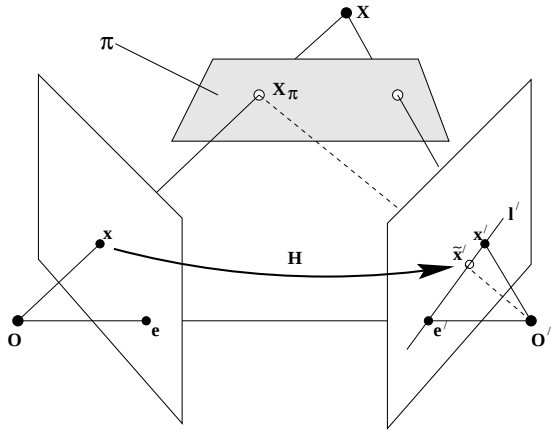


Figure 1: Plane induced parallax. The ray through $\mathbf{X}$ intersects the plane $\boldsymbol{\pi}$ at the point $\mathbf{X}_\pi$. The images of $\mathbf{X}$ and $\mathbf{X}_\pi$ are coincident points at $\mathbf{x}$ in the first view. In the second view the images are the points $\mathbf{x}'$ and $\tilde{\mathbf{x}}' = \mathtt{H}\mathbf{x}$ respectively. These points are not coincident (unless $\mathbf{X}$ is on $\boldsymbol{\pi}$), but both are on the epipolar line $\mathbf{l}'$ of $\mathbf{x}$.

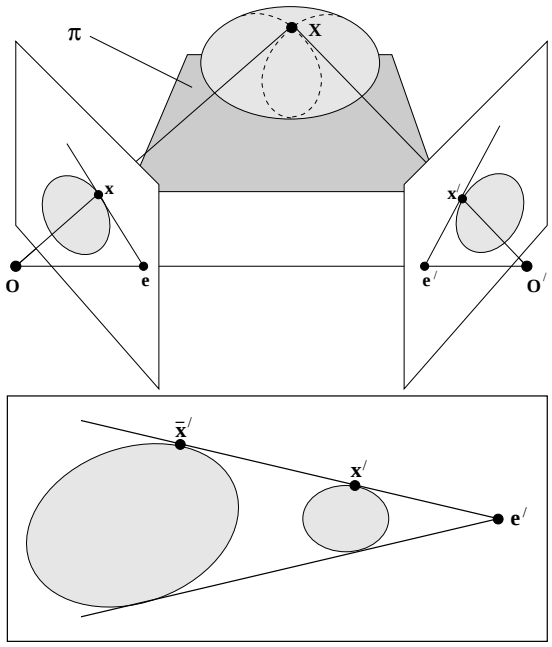


Figure 2: *Upper*: epipolar tangency. The epipolar plane $\mathbf{OO'X}$ is tangent to the surface at $\mathbf{X}$. The imaged outline is tangent to the epipolar lines at $\mathbf{x}$ and $\mathbf{x}'$ in the two views. The dashed curves on the surface are the contour generators. *Lower*: the first view is mapped to the second using the homography induced by the plane $\boldsymbol{\pi}$. The mapped point $\bar{\mathbf{x}}' = \mathtt{H}\mathbf{x}$. Imagine that the first camera is a light source and casts a shadow of the surface onto the plane, then the mapped outline is the image of this shadow boundary. The epipolar line is tangent to both the original and mapped outlines.

gent to the corresponding epipolar lines. Conversely, corresponding image epipolar tangent points backproject to a 3D point on the surface (at the point at which the epipolar plane is tangent to the surface). Thus, image epipolar tangent points are equivalent to a point correspondence and may be used to compute the epipolar geometry. A similar result holds in the case of epipolar tangents to an imaged space curve.

Porrill and Pollard [18] used epipolar tangent points to refine the epipolar geometry estimated from point correspondences. Cipolla *et al.* [6] then estimated epipolar geometry from epipolar tangencies alone. However, there are two problems in this case: first, the outline geometry must be sufficiently rich to provide at least 7 epipolar tangents; second, and more seriously, the solution involves a search over the 7 parameters of the fundamental matrix. This is a far worse situation than that of point correspondences where 7 correspondences directly determine either 1 or 3 solutions for the fundamental matrix, with no search involved.

If the outline from the first view is mapped to the second using the homography induced by the plane then, as illustrated in figure 2, an epipolar line which is tangent to the outline in the second view will also be tangent to the mapped outline. For brevity such lines will be referred to as *bitangents*, but it should be understood that the lines are only tangent to each contour once. Note that under the homography, the epipole $\mathbf{e}$ is mapped to $\mathbf{e}'$, and corresponding epipolar lines are mapped to each other. A real example showing the mapped outline and epipolar bitangents is given in figure 6. The situation is then completely equivalent to the point parallax case, with two bitangents (each the equivalent of a point correspondence) uniquely determining the epipole and thence the complete epipolar geometry using (1). Note that all epipolar tangents generate bitangents, but not all bitangents are epipolar tangents (just as not all putative point correspondences are correct).

To summarize: *Given the homography induced by the plane and two (epipolar) bitangents, then the epipolar geometry is uniquely determined.*

This result applies for a general motion of the camera including a change in the internal parameters. If the camera undergoes the particular motion of a pure translation (no rotation) with fixed internal parameters then the views may be registered by an identity homography (i.e. simply superimposed) as this homography corresponds to registering using the plane at infinity for this special motion. The bitangent result for this case was used in [20].

3 Computing the epipolar geometry

In this section we describe how the bitangent constraint described in section 2 may be used to automatically compute the fundamental matrix between two views.

We will require at least two epipolar tangents. For any finite smooth object, which does not intersect the baseline, there are at least two epipolar planes which are tangent to the surface ('below' and 'above'). There are thus at least two bitangents, provided the epipolar lines corresponding



Figure 3: Three images from different viewpoints. The surface of the jug is featureless other than a black mark that has been added to visually assess the computed geometry.

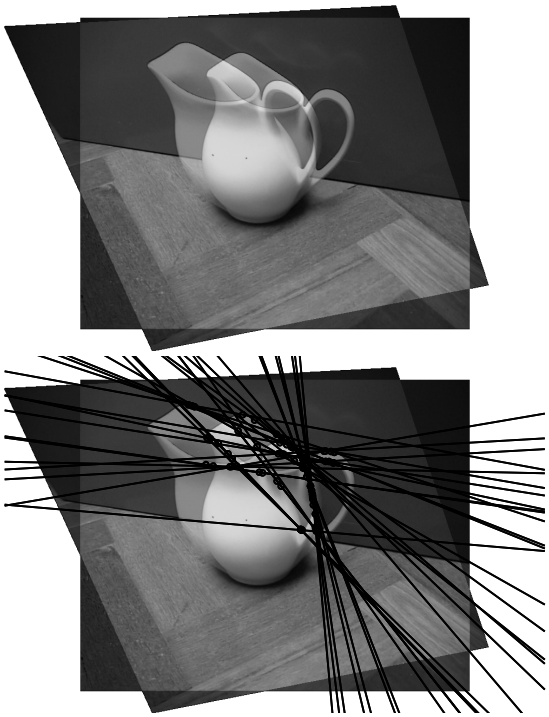


Figure 4: *Upper*: The first two images from figure 3 registered to the plane. *Lower*: The automatically detected bitangents are superimposed. In this case, 34 bitangents are detected. These figures are best viewed in the colour version of the paper.

to these planes lie within the finite image. However, for anything other than the most simple objects, it is normally the case that there are many epipolar tangents throughout the image.

3.1 Automatic computation of the epipole

It has been shown (section 2 above) that certain of the bitangents between corresponding curves in a pair of registered images contain the epipole. However, as not all bitangents contain the epipole, an algorithm is required to reject irrelevant bitangents and determine a best estimate of the epipole from those remaining.

One constraint that is satisfied by epipolar bitangents, but not by all bitangents, is that the corresponding outline curves have a consistent sign of curvature at epipolar tangents. This follows because the surface is on one side of

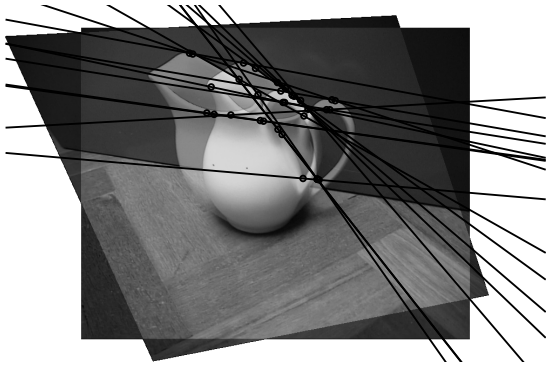


Figure 5: Many of the bitangents from figure 4 do not correspond to valid epipolar tangents from the curvature sign constraint (see text) and can be removed. This leaves 17 of the original 34 bitangents.

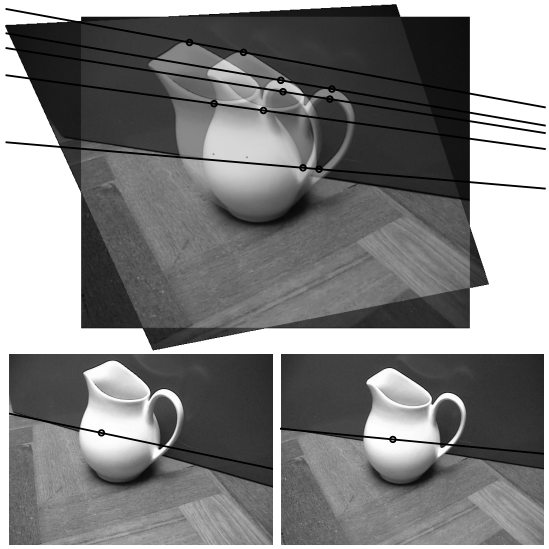


Figure 6: *Upper*: The epipole bitangent lines computed automatically from those of figure 5. The epipole lies on the intersection of these 5 bitangents. *Lower*: The accuracy of the computed epipolar geometry is demonstrated for a point on the jug. The residual error is sub-pixel.

the epipolar plane at surface tangent points (here sign includes zero in the case of tangency at outline inflections). Of course, the photometric constraint used to disambiguate interest point matches (based on the similarity of intensity neighbourhoods) may also be employed here.

A robust algorithm based on RANSAC [8] is then used to vote for the most likely epipole. The steps are: (1) Randomly select and intersect a pair of bitangents to produce a hypothesized epipole position; (2) Measure the support for this putative epipole position by the number of other lines through this epipole which are within a threshold distance of being bitangent; (3) Choose another pair of bitangents

and repeat. The selected epipole is the one with maximum support. Figures 4 – 6 illustrate the steps of this algorithm on a real example. In the implementation the bitangents are computed as described in [19]. It is generally found that provided there are 4 or more epipolar bitangents available the epipole is correctly determined.

The algorithm described above is able to compute the epipolar geometry using bitangents alone. However, it is often the case that a smooth surface contains surface markings which provide point correspondences. These point correspondences generate parallax vectors as described in section 1.1. Such point correspondences may be used simultaneously with bitangents to compute the epipolar geometry.

It is also worth noting that geometrically the epipolar tangents are necessarily in areas of the image where the surface is tangent to rays through the camera centre. Markings coinciding with these surface tangent points will therefore not be visible in the image, so epipolar tangents are a complementary source of correspondences to texture markings.

4 Computing the trifocal geometry

In this section a new minimal solution for computing the trifocal tenson is developed. Furthermore, the solution is suitable for use with epipolar tangencies. Previously it has been shown [7, 13] that the trifocal tensor is uniquely determined from the homography induced by a plane together with the correspondence of two points off the plane over all three views. However, this solution is not suitable in the case of epipolar tangencies because epipolar tangent points are a property of two views and for general motion the surface epipolar tangent points differ between the view pairs. Epipolar tangent points *are* common to three views for the special motion where the camera centres are collinear (so that epipolar planes are common to view pairs), or where the trifocal plane (the plane defined by the three camera centres) is tangent to the surface. It is shown here that the trifocal tensor may be determined uniquely from the plane homography together with two epipolar tangents between pairs of views.

The solution is obtained by constructing the three camera matrices (up to a common projective transformation of 3-space) and then computing the trifocal tensor from these matrices. The trifocal tensor is invariant to the 3-space projective ambiguity of the camera matrices and is computed from them as described by Hartley [11]. The views are numbered 1–3, with $\mathbf{e}_{ij}$ denoting the epipole in view i which is the image of the jth camera centre. It is supposed that the homographies induced by the world plane between the first and second view, $\mathtt{H}_{12}$, and first and third views, $\mathtt{H}_{13}$, are known. If the second view is mapped onto the first by $\mathtt{H}_{12}^{-1}$ and the third onto the first by $\mathtt{H}_{13}^{-1}$, then the computation of the trifocal tensor reduces to a particularly simple form. The 3×4 camera matrices of the three views may be chosen as [7]

$$\mathtt{P}^1 = [\mathtt{I} \mid \mathbf{0}], \quad \mathtt{P}^2 = [\mathtt{I} \mid \mathbf{e}_{12}], \quad \mathtt{P}^3 = [\mathtt{I} \mid \lambda \mathbf{e}_{13}]$$

up to a homography of 3-space, where λ is a scalar. As shown in section 2 the epipole $\mathbf{e}_{12}$ may then be computed from two epipolar tangents between views 1 and 2. Similarly, $\mathbf{e}_{13}$ may be computed from two epipolar tangents between views 1 and 3. It only then remains to determine λ (note that the relative scaling of the identity matrix and epipoles within the $\mathtt{P}$ matrices is fixed).

The scalar λ is computed from a consistency relationship on the epipoles over the three views. For simplicity assume the third camera centre is finite. Then it may be written as $\mathbf{C}^3 = (-\lambda \mathbf{e}_{13}{}^\top, 1)^\top$ (since $\mathtt{P}^3\mathbf{C}^3 = \mathbf{0}$). Then

$$\mathbf{e}_{23} = \mathtt{P}^2\mathbf{C}^3 = -\lambda \mathbf{e}_{13} + \mathbf{e}_{12} \tag{2}$$

The epipole $\mathbf{e}_{23}$ may be computed from two epipole tangents between views 2 and 3, and thence λ determined from the two inhomogeneous constraints of (2). However, provided the epipoles are not coincident, a single epipolar tangent line, $\mathbf{l}_{23}$, between views 2 and 3 determines λ because the epipole $\mathbf{e}_{23}$ lies on this line, and so $\mathbf{l}_{23}{}^\top \mathbf{e}_{23} = 0$ gives a linear equation in λ from (2). The epipoles in the registered views are coincident if, and only if, the camera centres are collinear.

To summarize, we have shown the following result in the case of general motion over three views: *Given the homography induced by the plane and two (epipolar) bitangents between two of the pairs of views (say views 1 & 2, and 1 & 3), and one epipolar bitangent between the other pair (say views 2 & 3), then the trifocal geometry is uniquely determined.*

4.1 Trifocal tensor computation algorithm

In summary, the trifocal tensor may be computed by the following four steps

1. **Register the views**: Compute the homography induced by the world plane between views 1 and 2, and transfer the outline from view 2 into view 1; similarly, compute the homography induced by the world plane between views 1 and 3, and transfer the outline from view 3 into view 1. From here on all computations are on outlines in the registered view.
2. **Compute epipolar bitangents:** Two bitangents are required between two of the outlines (between 1 and 2, and 1 and 3 say) and one, $\mathbf{l}_{23}$, between the others (2 and 3 say).
3. **Determine the epipoles and scalar λ:** The epipole $\mathbf{e}_{12}$ is the intersection of the two bitangents to outlines 1 and 2, and $\mathbf{e}_{13}$ is the intersection of the two bitangents to outlines 1 and 3. The scalar λ is determined as $\lambda = (\mathbf{l}_{23}{}^\top \mathbf{e}_{12})/(\mathbf{l}_{23}{}^\top \mathbf{e}_{13})$.
4. **Compute the trifocal tensor for the registered views from the camera matrices:** From $\mathtt{P}^1 = [\mathtt{I} \mid \mathbf{0}], \mathtt{P}^2 = [\mathtt{I} \mid \mathbf{e}_{12}], \mathtt{P}^3 = [\mathtt{I} \mid \lambda \mathbf{e}_{13}]$ determine the tensor as described in [11].

4.2 Automatic computation of the trifocal tensor

The solution described above has very undemanding requirements. There will always be at least one bitangent between outlines (at the imaged "top" of the object, if it is resting on the plane), and the solution only requires one more bitangent between two of the three view pairs. An automatic algorithm, based on the RANSAC paradigm, proceeds in much the same way as that for computing the fundamental matrix. Again epipolar tangents and point correspondences (off the plane) may be combined in the same algorithm. An example is shown in figures 7–9.

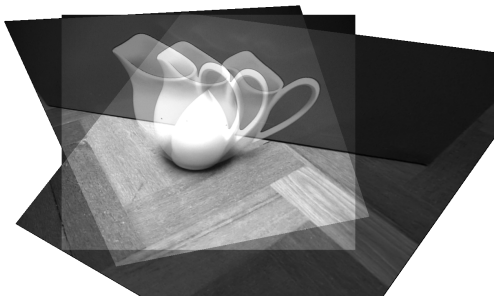


Figure 7: The three images of figure 3 registered to the plane.

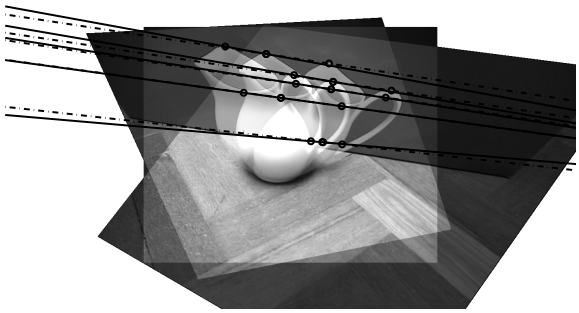


Figure 8: The bitangents (solid lines) between images 1 and 2 define the epipole e_{12}, whilst the bitangents (dashed lines) between images 2 and 3 define the epipole e_{23}. The remaining epipole, $e_{13} = e_{31}$, lies on the line joining e_{12} and e_{23}.



Figure 9: The trifocal tensor permits a single point in image 1 (shown here by a circle) to be transferred to an epipolar line in image 2. If the correspondence along the epipolar line in image 2 is known, a single point is determined in image 3. The black mark visible in figure 3 lies at the centre of the superimposed circle.

5 Computation of camera matrices for multiple views

In this section we describe how the 2 and 3 view results of the previous section may be extended to multiple views. The objective is to obtain a projectively consistent camera matrix $\mathtt{P}^i$ for each view of the sequence, such that if a point $\mathbf{X}_j$ is visible in several views then its image is given by $\mathbf{x}^i_j = \mathtt{P}^i\mathbf{X}_j$. We wish, of course, to make best use of the scene constraint that a plane is available.

Previous methods of computing consistent cameras for an uncalibrated image sequence, for example [3, 9, 15], have applied to general 3D scenes, and this has entailed extra difficulty. In our current case the homographies which are readily available throughout the sequence provide a firm basis for the camera computation.

Suppose we have obtained homographies $\mathtt{H}^i$ relative to a reference image, how does this constrain $\mathtt{P}^i$? A homography has 8 independent elements, and so only imposes 8 constraints on the 11 independent elements of the camera matrix. If the world coordinates are chosen such that the reference plane has $\mathrm{Z} = 0$, then points on that plane $\mathbf{X} = (\mathrm{X}, \mathrm{Y}, 0, 1)^\top$ are mapped by the projection matrix for a particular image as:

$$\begin{aligned}\mathbf{x} &= \mathtt{P}\mathbf{X} = [\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_3\ \mathbf{p}_4]\,(\mathrm{X}, \mathrm{Y}, 0, 1)^\top \\ &= [\mathbf{p}_1\ \mathbf{p}_2\ \mathbf{p}_4]\,(\mathrm{X}, \mathrm{Y}, 1)^\top\end{aligned}$$

Thus the 3 independent parameters of the third column of $\mathtt{P}$ are undetermined. This ambiguity corresponds to the position of the camera centre. However, if the internal parameter matrix $\mathtt{K}$ is known, then this ambiguity is reduced from a 3 parameter family to a finite number of solutions. Furthermore, the internal parameters can be computed from $\mathtt{H}^i$ if we assume that (some of) the camera parameters remain fixed throughout the sequence [28]. We now have a complete path to determining a metric set of camera matrices:

1. Compute the homographies $\mathtt{H}^i$ for the sequence.
2. Determine the internal calibration matrix $\mathtt{K}$.
3. Determine initial estimates of the camera matrices $\mathtt{P}^i$.
4. Use the pairwise epipolar tangent points to refine the estimated camera matrices $\mathtt{P}^i$.

These steps are now described in more detail.

5.1 Computation of homographies

The reference plane homographies are computed in a standard manner [5]. Point and line features are extracted in each image and the inter-image homographies are robustly estimated using RANSAC. Because the reference plane is generally visible in all views, each homography can then be directly related to the reference image, with a concomitant improvement in accuracy. The homographies are bundled adjusted to be projectively consistent throughout the sequence.

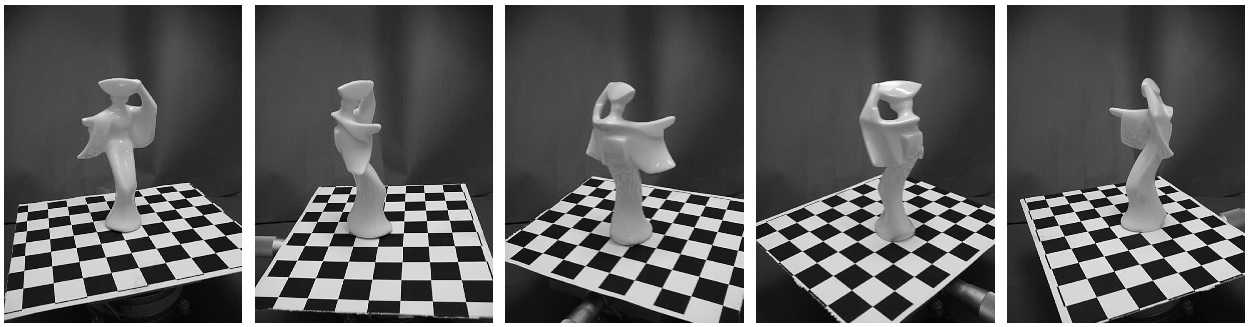


Figure 10: Example sequence 1. Although the object is resting on a calibration tile, the camera calibration is *not* determined from this.



Figure 11: Example sequence 2. *Reclining figure*, Henry Moore. These works are reproduced by permission of the Henry Moore Foundation.

5.2 Determination of camera calibration

Given the set of homographies, the camera calibration can be extracted as described by Triggs [28]. Briefly, each of the recovered homographies are of the form

$$\mathtt{H}^i = \mathtt{K}\left[\mathbf{r}_1^i \mathbf{r}_2^i \mathbf{t}^i\right]\mathtt{H}^0 \tag{3}$$

where $\mathtt{K}$ is the matrix of internal camera parameters, and $\mathtt{H}^0$ is the unknown projective transformation between the reference image and Euclidean coordinates on the plane $\mathtt{Z} = 0$. Writing

$$\mathtt{K}^{-1}\mathtt{H}^i(\mathtt{H}^0)^{-1} = \left[\mathbf{r}_1^i \mathbf{r}_2^i \mathbf{t}^i\right] \tag{4}$$

and requiring that $\mathbf{r}_1^i$ and $\mathbf{r}_2^i$ are proportional to the first two columns of a rotation matrix yields two constraints on the unknowns $\mathtt{K}$ and $\mathtt{H}^0$. Minimizing the reprojection error implied by imposing these constraints gives estimates of the unknown parameters.

5.3 Computation of projection matrices

The above procedure returns the internal camera parameters, and the first two columns of the rotation matrix for each camera. It is then a trivial matter to construct the third column as $\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$ and generate the full projection matrix for each view $\mathtt{P}^i = \mathtt{K}\left[\mathtt{R}^i \mid \mathbf{t}^i\right]$.

These projection matrices provide estimates of the epipoles that are sufficiently accurate to usefully prune the putative matches in the RANSAC for bitangents—only bitangents that pass within a threshold distance of the approximate epipoles are considered.

5.4 Refinement of projection matrices

The projection matrices produced by the above procedure are very accurate for points on the reference plane, but accuracy decreases sharply off the plane. Happily however, the pairwise epipolar tangencies typically provide corresponding points which are far from the plane.

The two sources of information are combined by entering the epipolar tangencies as 2-view point correspondences in a Euclidean bundle adjustment for the $\mathtt{P}^i$. This has the effect of stabilizing the camera computation, leading to very accurate results. Cameras computed for the example sequence of Figure 11 are shown in Figure 12.

6 Application — A freehand scanner for smooth objects

In this section we sketch how the computed cameras may be used to construct a 3D texture mapped graphical model of a smooth object on a planar surface. This technology is essentially a "freehand scanner" for smooth objects—a model of a smooth object can be acquired literally by walking around the object (seated on a suitable plane, such as a ground plane) with a video camera.

Given the refined camera matrices, a 3D surface model is computed by intersecting the cones back-projected from the object silhouettes [2, 25, 26]. Here, a volume-based approach is used, where each voxel is projected into each of the thresholded silhouette images, and only voxels which are inside the occluding contour in all images are retained. Figure 14 shows the volumetric models formed from the example sequences after this process.

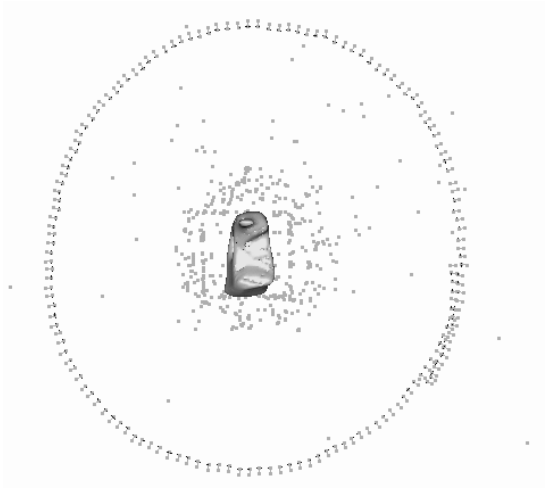


Figure 12: Camera positions, reconstructed ground plane points, and surface model computed for the example sequence of figure 11. Each camera is represented by its image plane and principal point.

7 Conclusions and Extensions

It has been demonstrated that given a registering homography from a plane and an occluding contour with sufficiently rich geometry (to support at least 2 epipolar tangents) the 2 and 3 view relations can be computed from this information alone, without requiring an iterative search. In practical circumstances the visual information is generally not so impoverished and both interest points and bitangents are available on the object. Both may be used interchangeably in the multiple view estimation algorithms.

This is a very pragmatic approach since planes are often available in the scene—indeed the plane need not be an actual plane, for example given sufficient (4 or more) vanishing points and their correspondences the plane at infinity may be used.

We are currently extending this approach to allow multiple sequences to be registered—again using the plane as the main registration device. In this manner sequences captured at a set of azimuths whilst circumnavigating an object may be fairly effortlessly combined. Much like peeling an apple in reverse.

References

[1] K. Åström and F. Kahl. Motion estimation in image sequences using the deformation of apparent contours. *IEEE T-PAMI*, 21(2):114–127, Feb 1999.

[2] B. G. Baumgart. *Geometric Modelling for Computer Vision.* PhD thesis, Stanford University, Palo Alto, 1974. AIM-249, STAN-CS-74-463.

[3] P. Beardsley, P. Torr, and A. Zisserman. 3D model acquisition from extended image sequences. In *Proc. ECCV*, LNCS 1064/1065, pages 683–695. Springer-Verlag, 1996.

[4] R. Berthilsson and K. Åström. Reconstruction of 3-d curves from 2-d images using affine shape methods for curves. In *Proc. CVPR*, 1997.

[5] D. Capel and A. Zisserman. Automated mosaicing with super-resolution zoom. In *Proc. CVPR*, pages 885–891, Jun 1998.

[6] R. Cipolla, K. Åström, and P. Giblin. Motion from the frontier of curved surfaces. In *Proc. ICCV*, pages 269–275, 1995.

[7] A. Criminisi, I. Reid, and A. Zisserman. Duality, rigidity and planar parallax. In *Proc. ECCV*, pages 846–861. Springer-Verlag, Jun 1998.

[8] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM*, 24(6):381–395, 1981.

[9] A. W. Fitzgibbon and A. Zisserman. Automatic camera recovery for closed or open image sequences. In *Proc. ECCV*, pages 311–326. Springer-Verlag, Jun 1998.

[10] R. I. Hartley. In defence of the 8-point algorithm. In *Proc. ICCV*, pages 1064–1070, 1995.

[11] R. I. Hartley. A linear method for reconstruction from lines and points. In *Proc. ICCV*, pages 882–887, 1995.

[12] M. Irani and P. Anandan. Parallax geometry of pairs of points for 3d scene analysis. In B. Buxton and R. Cipolla, editors, *Proc. ECCV*, pages 17–30. Springer, 1996.

[13] M. Irani, P. Anandan, and D. Weinshall. From reference frames to reference planes: Multi-view parallax geometry and applications. In *Proc. ECCV*, 1998.

[14] R. Kumar, P. Anandan, M. Irani, J. Bergen, and K. Hanna. Representation of scenes from collections of images. In *ICCV Workshop on the Representation of Visual Scenes*, 1995.

[15] S. Laveau. *Geometry of a system of N cameras. Theory, estimation, and applications.* PhD thesis, INRIA, 1996.

[16] Q.-T. Luong and T. Viéville. Canonic representations for the geometries of multiple projective views. In *Proc. ECCV*, pages 589–599, May 1994.

[17] J. Mundy and A. Zisserman. *Geometric Invariance in Computer Vision.* MIT Press, 1992.

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

[19] C. Rothwell, A. Zisserman, D. Forsyth, and J. Mundy. Planar object recognition using projective shape representation. *IJCV*, 16(2), 1995.

[20] J. Sato and R. Cipolla. Affine reconstruction of curved surfaces from uncalibrated views of apparent contours. In *Proc. ICCV*, pages 715–720, 1998.

[21] H. S. Sawhney. Simplifying motion and structure analysis using planar parallax and image warping. In *Proc. CVPR*, 1994.

[22] A. Shashua. Trilinearity in visual recognition by alignment. In *Proc. ECCV*, volume 1, pages 479–484, May 1994.

[23] A. Shashua and N. Navab. Relative affine structure: Canonical model for 3d from 2d geometry and applications. *IEEE Transaction on PAMI*, 18(9):873–883, 1996.

[24] M. E. Spetsakis and J. Aloimonos. Structure from motion using line correspondences. *IJCV*, 4(3):171–183, 1990.

[25] S. Sullivan and J. Ponce. Automatic model construction, pose estimation, and object recognition from photographs using triangular splines. In *Proc. ICCV*, 1998.

[26] R. Szeliski. Rapid octree construction from image sequences. *CVGIP*, 58(1):23–32, Jul 1993.

[27] P. H. S. Torr and A. Zisserman. Robust parameterization and computation of the trifocal tensor. *Image and Vision Computing*, 15:591–605, 1997.

[28] W. Triggs. Autocalibration from planar scenes. In *Proc. ECCV*, 1998.

[29] Z. Zhang. Determining the epipolar geometry and its uncertainty — a review. *IJCV*, 27(2):161–195, March 1998.

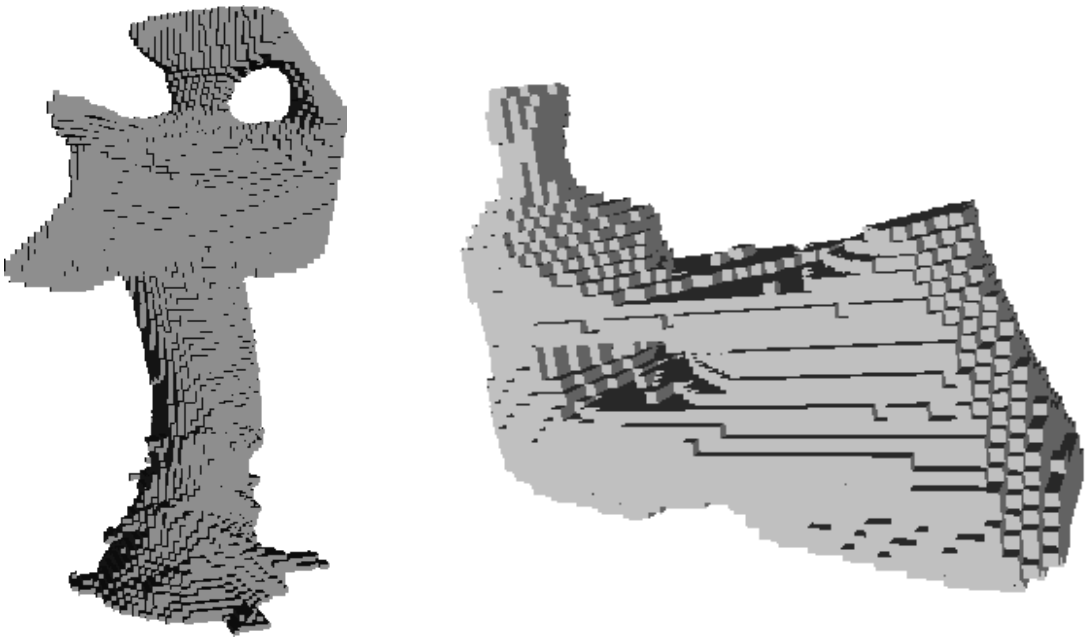


Figure 13: Volumetric occupancy grids, computed from silhouettes tracked in the original images, with camera positions computed as described in §5. This "visual hull" penetrates many, but not all, concavities.



Figure 14: Texture-mapped surface models, viewed from positions not in the original sequences. The black areas are not textured as they were not seen in any image.