

Euclidean Structure from Uncalibrated Images

Martin Armstrong, Andrew Zisserman and Paul Beardsley
Robotics Research Group, Department of Engineering Science
Oxford University, OX1 3PJ.

Abstract

A number of recent papers have demonstrated that camera "self-calibration" can be accomplished purely from image measurements, without requiring special calibration objects or known camera motion. We describe a method, based on self-calibration, for obtaining (scaled) Euclidean structure from multiple uncalibrated perspective images using only point matches between views.

The method is in two stages. First, using an uncalibrated camera, structure is recovered up to an affine ambiguity from two views. Second, from one or more further views of this affine structure the camera intrinsic parameters are determined, and the structure ambiguity reduced to scaled Euclidean. The technique is independent of how the affine structure is obtained. We analyse its limitations and degeneracies.

Results are given for images of real scenes. An application is described for active vision, where a Euclidean reconstruction is obtained during normal operation with an initially uncalibrated camera. Finally, it is demonstrated that Euclidean reconstruction can be obtained from a single perspective image of a repeated structure.

1 Introduction

If cameras are uncalibrated then, in the absence of other information, structure can only be recovered up to a projective ambiguity from an image pair or sequence [3, 9]. However, for special motions or unchanging intrinsic camera parameters, the projective ambiguity can be reduced to affine or scaled Euclidean (similarity). The latter approach - constant intrinsic parameters - was pioneered by Faugeras *et al.* [5] and termed "self-calibration". Visual tasks *can* be accomplished using only projective or affine structure. Examples include recognition [18] (projective) and path planning [2], grasping [10] and fixation point tracking [17] (all affine). However, there are clearly situations where having, at least approximate, calibration is useful. For example, within an active vision system, rotation about the optical centre and control of fixation angles requires approximate knowledge of the intrinsic parameters.

Self-calibration is important for a number of reasons. First, it removes the onerous task of calibrating cameras using special calibration blocks [20]. Second, calibration can be updated if, for example, focal length is altered during an image sequence. Third, if an incorrect intrinsic calibration is adopted this could conflict with veridical intrinsic parameters, which are implicit in the data during motion, and lead to inconsistencies or numerical instabilities.

The original self-calibration method of Faugeras, Luong and Maybank [5, 12, 14] required solving a system of polynomial equations obtained from three views. Hartley [8] developed a self calibration method for pure rotations with a *linear* solution for the intrinsic parameters. More recently stratified approaches have been developed [7, 13] for general viewer motions, where calibration is achieved in two stages. First, affine structure is recovered. Second, intrinsic parameters are determined based on affine relations, both stages generally involve *linear* computations. The stratified approach is adopted here, and described in detail in section 2.

All self-calibration methods require rotation of the image plane in order to constrain the intrinsic parameters. In essence, self-calibration is possible because the camera always carries its own calibration object - namely the image plane. The image plane acts as a gauge object defining angles in a projective space.

2 Theory

2.1 Camera Model and Notation

Here we define the camera matrix and the five intrinsic parameters. These results are based mainly on [3, 6, 16]. Perspective projection from 3D projective space \mathcal{P}^3 to the image plane \mathcal{P}^2 is modelled by a 3×4 matrix \mathbf{P} where $\mathbf{x} = \mathbf{P}\mathbf{X}$, $\mathbf{x} = (x, y, 1)^\top$ and $\mathbf{X} = (X, Y, Z, 1)^\top$ are homogeneous vectors. If $(X, Y, Z)^\top$ is measured in a Euclidean coordinate system, then \mathbf{P} can be decomposed as $\mathbf{P} = \mathbf{C}\mathbf{R}[\mathbf{I} \mid -\mathbf{t}_E]$ by QR decomposition¹ of the 3×3 matrix $\mathbf{C}\mathbf{R}$, where \mathbf{I} is the 3×3 identity, and $\{\mathbf{R}, \mathbf{t}_E\}$ is the Euclidean (rotation and translation) transformation between the world and camera coordinate frames, and \mathbf{C} the upper triangular matrix containing the intrinsic parameters of the camera. \mathbf{C} is the transformation between the camera coordinates of a ray and image (measured) homogeneous coordinates.

$$\mathbf{C} = \begin{bmatrix} \alpha_u & k & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

¹Any real matrix can be decomposed into the product of a triangular and orthogonal matrix.

where α_u and α_v are the focal length measured in pixels along the x and y directions respectively, (u_0, v_0) is the principal point, and $k = \alpha_u \cot \theta$ is the skew, with θ the angle between image axes.

2.2 Euclidean Reconstruction

Camera calibration, and hence Euclidean reconstruction, is achieved in three stages using point matches between a number of images of a static scene. More details can be found in [1].

1. Computing affine structure: An affine reconstruction is where point coordinates $\{\mathbf{X}^A\}$ are known up to an affine transformation of their Euclidean values $\mathbf{X}^E = (X, Y, Z, 1)^\top$, i.e.

$$\mathbf{X}^A = \mathbf{T}_A \mathbf{X}^E = \begin{bmatrix} \mathbf{A} & \mathbf{t}_A \\ \mathbf{0}^\top & 1 \end{bmatrix} \mathbf{X}^E$$

where \mathbf{A} is an invertible 3×3 matrix, and \mathbf{t}_A a 3-vector. $\{\mathbf{A}, \mathbf{t}_A\}$ are unknown but the same for all points.

Affine structure is recovered from two uncalibrated images with pure (though unknown) translation between the views (i.e. no rotation), and unchanging intrinsic parameters (again unknown), based on a result of Moons *et al.* [15]. The projective camera matrices for the two views are written $\mathbf{x}_1 = \mathbf{P}_1 \mathbf{X}$ and $\mathbf{x}_2 = \mathbf{P}_2 \mathbf{X}$, for the first and second images respectively. It can be shown that for this motion the matrices can be chosen as $\mathbf{P}_1 = [\mathbf{I} | \mathbf{0}]$, $\mathbf{P}_2 = [\mathbf{I} | -\mathbf{e}]$ where \mathbf{e} is the image epipole. Reconstruction proceeds in the normal fashion by backprojecting corresponding points in each view, and determining the 3D point as the ray's intersection.

Implementation details For this motion the fundamental matrix is skew symmetric with only three degrees of freedom, and epipoles have the same position in both images. \mathbf{F} is computed from point correspondences between images, and \mathbf{e} computed from $\mathbf{F}\mathbf{e} = \mathbf{0}$. The world point, \mathbf{X}^A , is computed by intersecting backprojected rays: from \mathbf{x}_1 , $\mathbf{X}^A = \lambda_1(\mathbf{x}_1^\top, 1)^\top$; and from \mathbf{x}_2 , $\mathbf{X}^A = \lambda_2(\mathbf{x}_2^\top, 1)^\top + \mathbf{e}$; where λ_i is an unknown scaling along each ray. Because of image localisation error the rays will be skew and a minimisation criteria is required. We minimise *image* error (as opposed to a 3D error) by orthogonally projecting points $\mathbf{x}_1, \mathbf{x}_2$ onto their "epipolar line", generating corrected points $\hat{\mathbf{x}}_i$. Note, for this special motion, an "epipolar line" for two corresponding points can be chosen as the line containing \mathbf{e} which minimises the perpendicular distance to the two points. The 3D position is the intersection of the rays corresponding to $\hat{\mathbf{x}}_i$.

2. Computing camera calibration: This method is an extension of that of Luong and Vieville [13] and Hartley [7]. \mathbf{P}_i can be partitioned as $\mathbf{P}_i = [\mathbf{M}_i | -\mathbf{M}_i \mathbf{t}_i]$. It can be shown [13] that in an affine coordinate frame, the *infinite homography*, \mathbf{H}_∞ , is given by $\mathbf{H}_\infty = \mathbf{M}_j \mathbf{M}_i^{-1}$. The infinite homography

is the point transformation matrix for vanishing points, i.e. $\mathbf{v}_j = \mathbf{H}_\infty \mathbf{v}_i$, where \mathbf{v}_i and \mathbf{v}_j are corresponding vanishing points in the i and j images respectively. Under the \mathbf{H}_∞ transformation $\mathbf{K}_j = \mathbf{H}_\infty \mathbf{K}_i \mathbf{H}_\infty^\top$ where $\mathbf{K}_i = \mathbf{C}_i \mathbf{C}_i^\top$. \mathbf{K}_i is the image in view i of the dual (i.e. the inverse) of the absolute conic. This equation provides a linear method for obtaining \mathbf{K} . Camera intrinsic parameters are fixed between views, so $\mathbf{K}_i = \mathbf{K}_j = \mathbf{K}$ and

$$\mathbf{K} = \mathbf{H}_\infty \mathbf{K} \mathbf{H}_\infty^\top \quad (2)$$

Once \mathbf{K} is determined (up to scale), \mathbf{C} can be obtained simply by a Choleski decomposition² of $\mathbf{K} = \mathbf{C} \mathbf{C}^\top$.

In the following we consider two methods for obtaining \mathbf{K} . Both depend on taking further views of the structure where the camera motion includes a rotation. If there is no rotation, equation (2) reduces to $\mathbf{K} = \mathbf{K}$ and there is no constraint on \mathbf{K} .

Single rotation: Equation (2) can be written in the form $\mathcal{H}\mathbf{k} = 0$, where \mathbf{k} is the six distinct elements of the symmetric matrix \mathbf{K} written as a vector, and \mathcal{H} is a 6×6 matrix. It can be shown [13], that \mathcal{H} is rank four, so there is a one parameter family of homogeneous solutions for \mathbf{k} . Following [13], this ambiguity can be removed by assuming $k = 0$ (k is the skew parameter in equation (1)) i.e. that image axes are perpendicular. This is a very good approximation in practice. This constraint results in a quadratic, $K_{12}K_{33} = K_{23}K_{13}$, giving two solutions for \mathbf{K} . Now \mathbf{K} can be decomposed into \mathbf{C} by Choleski factorisation. Often one of the solutions for \mathbf{K} is not positive definite and so can be eliminated.

Two or more rotations: Suppose there are two or more rotations about different axes. Each new view generates an equation, $\mathcal{H}\mathbf{k} = 0$, where \mathcal{H} is rank four. Combining these equations for n rotations generates a system, with \mathcal{H} a $6n \times 6$ matrix from n \mathbf{H}_∞ matrices, which will be rank five in general. Consequently a unique \mathbf{K} is obtained (up to scale). In practice \mathbf{k} is determined by minimising $E = \|\mathcal{H}\mathbf{k}\|^2 / \|\mathbf{k}\|^2$ using SVD to find the eigenvector corresponding to the minimum eigenvalue.

3. Computing Euclidean Structure: Once the camera intrinsic parameters are known, scaled Euclidean structure, \mathbf{X}^S , can be computed from the affine structure. It can be shown that $\mathbf{X}^S = \mathbf{C}^{-1} \mathbf{X}^A$. If measurements on the actual Euclidean structure are \mathbf{X}^E , then the remaining ambiguity is that \mathbf{X}^S is related to \mathbf{X}^E by a rotation, translation, and isotropic scaling.

2.3 Degeneracies

If camera rotation is only about one of the camera axes, then some of the intrinsic parameters are unconstrained by equation (2) and cannot be

²Any symmetric positive definite matrix \mathbf{A} can be decomposed into a unique upper triangular matrix \mathbf{U} such that $\mathbf{A} = \mathbf{U} \mathbf{U}^\top$.

determined. This can be analysed by examining the infinite homography H_∞ . It can be shown [13] that $H_\infty = CRC^{-1}$ where R is the rotation between images. When $k = 0$, a rotation of θ about the camera X axis gives

$$H_\infty(R(\theta)) = \begin{bmatrix} 1 & u_0 \sin \theta / \alpha_v & u_0(\cos \theta - 1) - u_0 v_0 \sin \theta / \alpha_v \\ 0 & \cos \theta + v_0 \sin \theta / \alpha_v & v_0(\cos \theta - 1) - v_0^2 \sin \theta / \alpha_v - \alpha_v \sin \theta \\ 0 & \sin \theta / \alpha_v & \cos \theta - v_0 \sin \theta / \alpha_v \end{bmatrix}$$

There are no α_u terms in H_∞ and consequently no possible constraint on α_u in equation (2) (the coefficient of α_u is zero). Consequently, α_u is unconstrained. Similarly, with a rotation about the Y axis, α_v is unconstrained. For rotations about the Z axis (the optical axis), α_u and α_v only appear in H_∞ as the ratio (α_v/α_u) , so their independent values cannot be determined.

Near degenerate situations occur if the magnitude of rotation about one axis is small. In this case the associated intrinsic parameter is poorly constrained.

3 Experimental results and assessment

Accuracy is assessed in two ways: first, by comparing the intrinsic parameters determined using self-calibration with “veridical” values obtained by conventional calibration [4] using a calibration object with known 3D structure; second, by comparing the recovered 3D Euclidean structure with that of the accurately known calibration object. The calibration object used is shown in figure 1.

3.1 Conventional Camera Calibration

The veridical intrinsic parameters are determined from a single perspective image of a calibration object (an orthogonal pair of Tsai grids, see figure 1a), with accurately known 3D Euclidean structure. The 3×4 projection matrix, P , is determined from the correspondence between image and 3D corners. C is extracted from P using QR decomposition as in section 2.1. Image corner positions are determined to sub-pixel accuracy by straight line intersections. The matrix P is computed by minimising the distance between measured and predicted image positions using a Levenberg-Marquardt scheme. Results are given in table 1 for averages computed over all available images. In general α_u , α_v , and the aspect ratio α_v/α_u are stable over views, varying by less than 0.5%. In contrast the principal point (u_0, v_0) varies over a 20×20 pixel region in a 512×512 image. This more substantial variation of the principal point commonly occurs in calibration [20].

3.2 Self Calibration

In the following, a number of examples are compared differing in the number of views and rotation between views. In all cases affine structure is recovered from two views related by a translation. Corner detection and matching are automatic. The results for the intrinsic parameters are given

in table 1. The tolerances given are the standard deviation over several sets of images. The number of sets used is also shown in the table.

3 Images There is a pure translation between the first two images, and a translation and rotation for the third. The translation between the second and third images is necessary only to ensure the grid is at the image centre, and hence reduce radial image distortion. H_∞ is determined between views two and three (it is identical to that obtained from views one and three). As described in section 2.2, the additional constraint that $k = 0$ is required in this single rotation case. Two sets of results are given. For the first (5°), the rotation magnitude is 5° about both the camera X and Y axes. For the second (10°), at least one of the rotations has magnitude larger than 5° : two images have rotations, $(R(X), R(Y))$, of $(5^\circ, 10^\circ)$, two have $(5^\circ, 15^\circ)$, one each of $(10^\circ, 10^\circ)$ and $(10^\circ, 15^\circ)$. The rotation is limited to 15° so that both sides of the calibration grid are still in view in the image. The larger rotations give more stable values, as discussed in section 2.3.

4 Images A second rotated view is added to the sequence, and the two H_∞ matrices, found from the mapping from the second to the third, and second to the fourth image, are used simultaneously to find K . This second H_∞ generally changes the rank of \mathcal{H} from four to five. The constraint $k = 0$ is not required, and the calculated value for k given. Note, H_∞ between the third and fourth images adds no additional information. These images are selected from the (5°) and (10°) sets above, such that rotations about corresponding camera axes (e.g. the X axis) in the third and fourth views differ in magnitude.

6 Images Four rotated views are used to give a 24×6 matrix \mathcal{H} of rank five, with the H_∞ mapping from the second image to each of the four rotated images. These give the most stable results, with intrinsic parameters approaching the veridical values.

Degenerate A three image sequence is used, but the rotation is here limited to either the camera X or Y axis. As shown in section 2.3 this results in α_u or α_v being unconstrained.

The stability of the intrinsic parameters increases with the number of views and the size of the rotations. α_u , α_v and α_u/α_v have an error of between 2% and 6%. Again (u_0, v_0) varies over a region in the centre of the image, the size of the region varying from 20×30 to 5×5 , decreasing with increasing number of views and rotation magnitude. In the six image case, the skew parameter $k = 21.8$ is significantly larger than the veridical value of $k = 1.5 \pm 0.4$, but this corresponds to only 2° off perpendicular.

Method (No. of Image sets)	α_u (pixels)	α_v (pixels)	α_v/α_u	u_0 (pixels)	v_0 (pixels)	k (pixels)
Known 3D Grid(10)	646.0 ± 3.5	968.7 ± 4.7	1.4996 ± 0.0016	246.5 ± 9.2	244.3 ± 11.4	1.5 ± 0.4
3 Images-5°(9)	625.2 ± 41.1	954.5 ± 35.5	1.53 ± 0.11	242.8 ± 20.1	225.6 ± 29.4	0.0
3 Images-10°(6)	684.4 ± 26.7	1034.0 ± 24.9	1.513 ± 0.081	261.0 ± 4.0	234.2 ± 32.0	0.0
4 Images (13)	661.0 ± 17.0	1018.0 ± 22.2	1.522 ± 0.028	258.0 ± 7.7	239.7 ± 11.6	1.2 ± 36.6
6 Images (35)	645.6 ± 13.1	1001.8 ± 16.9	1.552 ± 0.021	260.0 ± 4.6	236.9 ± 6.6	21.8 ± 14.4
Degenerate-X (6)	-	1064.0 ± 32.5	-	245.0 ± 4.2	208 ± 3	0.0
Degenerate-Y (2)	687.2 ± 34.9	-	-	260.2 ± 27.2	33.8 ± 13.7	0.0

Table 1: Mean and standard deviation of the intrinsic parameters determined from a varying number of views and rotations. See text for details.

3.3 Euclidean Reconstruction

Scaled Euclidean structure is recovered from the affine structure and camera calibration using the method of section 2.2. The accuracy of the reconstruction indicates how errors in camera calibration propagate through to errors in structure. For example, an error in α_u could have a more detrimental effect than one in u_0 . The accuracy of the scaled Euclidean reconstruction is assessed by comparing similarity invariants (angles, distance ratios) with those calculated from the known structure of the calibration object. Results are given in table 2 and figure 1 for the same image sets used to evaluate the intrinsic parameters. Three similarity invariants are measured, the first two measure local structure, and the third global.

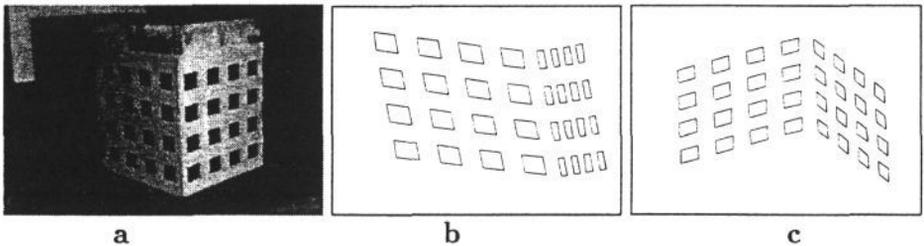


Figure 1: (a) A typical image of the calibration object used to assess accuracy. (b) Affine structure recovered from two perspective images of the calibration object, note the in plane skew. The angle between the two planes is 179.7° compared to 90.0° on the object. (c) Scaled Euclidean structure obtained from three perspective images, after determining intrinsic parameters. The planes are now orthogonal to high accuracy, see table 2.

Method	Distance Error	Pattern Error (degrees)	Plane Angle (degrees)
3 Images-5°	4.5%	2.51	91.03 ± 3.80
3 Images-10°	3.9%	1.66	91.52 ± 4.20
4 Images	3.9%	1.12	89.33 ± 2.21
6 Images	3.7%	1.00	90.35 ± 1.07

Table 2: Similarity invariants measured on recovered (scaled) Euclidean structure: The **Distance Error** is the ratio of standard deviation to mean over 196 measurements of distance between adjacent points; The **Pattern error** is the standard deviation of angles computed for all corners of grid squares; The **Plane angle** is the angle between the two planes of the calibration object, determined by orthogonal regression to all points on each plane of the grid. For ideal data the distance and pattern error would be zero, and plane angle 90°.

4 Applications

4.1 Active vision

For a camera mounted on a robot arm or AGV, it is trivial to perform pure translational motion. Consequently, the method of Moons *et al.* [15] adopted here for generating affine structure, is particularly appropriate to active vision tasks. Accuracy can be improved by servoing on the epipole position. The following results are for an implemented system [2], with automatic corner detection, matching and elimination of outliers. See [19] for details of the outlier rejection method. During the motion the camera is calibrated *on the fly* and similarity structure recovered. Note, corners are detected here using a corner detector, thus localisation is less accurate than the line intersection method used for the calibration grid.

Figure 2 shows two images from a sequence of 20 where the camera rotates overall by 20° about its X and Y axes. The calculated intrinsic parameters are $\{\alpha_u = 673.1, \alpha_v = 1005.4, \alpha_v/\alpha_u = 1.494, (u_0, v_0) = (249.3, 282.9)\}$, compare with values in table 1. The recovered structure is shown in figure 2. The angles between the roof, front and side of the house were calculated, by orthogonal regression to points on each plane, to be 53.9°, 94.7°, 87.5° compared to measured values of 51.4°, 90.0°, 90.0°.

4.2 Repeated structure

Structures that repeat in a single image of a scene are equivalent to multiple views of a single instance of the structure [11]. Thus, for example, a view of two identical objects related by a translation is equivalent to a stereo pair of images of one object, with the cameras related by a pure translation. If in a single image there are three identical objects, of which two are

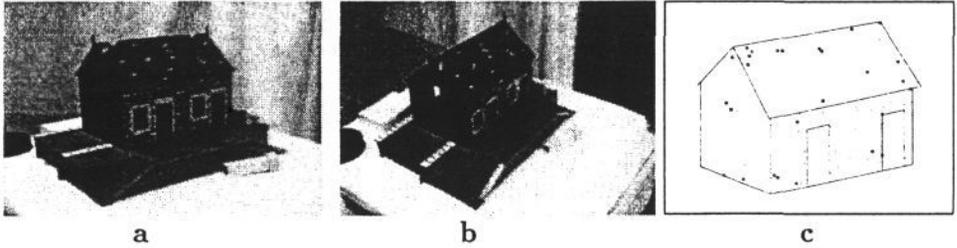


Figure 2: (a) and (b) Images from the sequence used for similarity reconstruction. (c) The reconstructed house projected from a different viewpoint, with lines added for clarity. The reconstruction is for 100 points matched between the first two images, of which 30 (bold) were tracked through all 20 images and used to compute calibration.

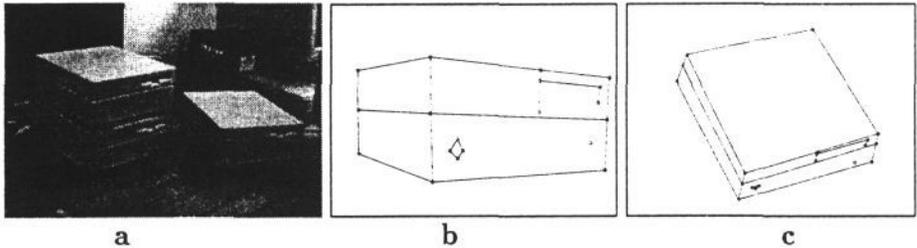


Figure 3: (a) A single image containing repeated structure. (b) and (c) Two views of the recovered scaled Euclidean structure.

related by a simple translation, then the camera can be calibrated and structure measured up to a similarity ambiguity. Figure 3 shows such an image and the reconstruction. The angles between the planar sides of the reconstruction, found by orthogonal regression to points on each plane, are 84° , 81° , 81° , compared to an actual value of 90° .

5 Discussion

We have demonstrated that numerically stable camera self-calibration can be achieved from images of real scenes. The accuracy approaches that of conventional calibration methods based on calibration objects. There are a number of interesting research questions to be addressed:

1. Are there other methods of achieving affine structure without scene knowledge (i.e. knowing affine quantities such as vanishing points or volume ratios) from two or more views.
2. Are perspective effects important? Will calibration accuracy degrade if the imaging geometry is near affine? (object depth small compared to the distance from the camera, and a small field of view).

Acknowledgements We are grateful for financial support from the SERC and ESPRIT BRA 6448 'VIVA'.

References

- [1] M. Armstrong. Euclidean structure and camera calibration from image sequences. First year report, Dept. of Engineering Science, University of Oxford, 1994.
- [2] P. Beardsley, A. Zisserman, and D. Murray. Navigation using affine structure and motion. In *Proc. ECCV*, LNCS 800/801. Springer-Verlag, 1994.
- [3] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proc. ECCV*, LNCS 588. Springer-Verlag, 1992.
- [4] O. Faugeras. *Three-dimensional computer vision: a geometric viewpoint*. MIT Press, 1993.
- [5] O. Faugeras, Q. Luong, and S. Maybank. Camera self-calibration: theory and experiments. In *Proc. ECCV*, LNCS 588. Springer-Verlag, 1992.
- [6] R. Hartley. Estimation of relative camera positions for uncalibrated cameras. In *Proc. ECCV*, LNCS 588. Springer-Verlag, 1992.
- [7] R. Hartley. Euclidean reconstruction from uncalibrated views. In J.L.Mundy, A.Zisserman, and D.A.Forsyth, editors, *Applications of Invariance in Computer Vision*, LNCS 825. Springer-Verlag, 1994.
- [8] R. Hartley. Self-calibration from multiple views with a rotating camera. In *Proc. ECCV*, LNCS 800/801. Springer-Verlag, 1994.
- [9] R. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. CVPR*, 1992.
- [10] N. Hollinghurst and R. Cipolla. Uncalibrated stereo hand-eye coordination. In *Proc. BMVC*. BMVA Press, 1993.
- [11] J. Liu, J. Mundy, and E. Walker. Recognizing arbitrary objects from multiple projections. In *Proc. Asian Conference on Computer Vision*, 1993.
- [12] Q. Luong. *Matrice Fondamentale et Autocalibration en Vision par Ordinateur*. PhD thesis, Université de Paris-Sud, France, 1992.
- [13] Q. Luong and T. Vieville. Canonic representations for the geometries of multiple projective views. Technical Report UCB/CSD-93-772, University of California, Berkeley, USA, 1993.
- [14] S. Maybank and O. Faugeras. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8, 1992.
- [15] T. Moons, L. Van Gool, M. Van Diest, and E. Pauwels. Affine reconstruction from perspective image pairs. In J.L. Mundy, A. Zisserman, and D.A. Forsyth, editors, *Applications of Invariance in Computer Vision*, LNCS 825. Springer-Verlag, 1994.
- [16] J. Mundy and A. Zisserman. *Geometric invariance in computer vision*. MIT Press, 1992.
- [17] I. Reid and D. Murray. Tracking foveated corner clusters using affine structure. In *Proc. ICCV*, 1993.
- [18] C. Rothwell, A. Zisserman, J. Mundy, and D. Forsyth. Efficient model library access by projectively invariant indexing functions. In *Proc. CVPR*, 1992.
- [19] P. Torr, P. Beardsley, and D. Murray. Robust vision. In *Proc. BMVC*, 1994.
- [20] R. Tsai. An efficient and accurate camera calibration technique for 3D machine vision. In *Proc. CVPR*, 1986.