Camera calibration and the search for infinity.

Richard I. Hartley,
G.E. Corporate Research and Development
1 Research Circle,
Niskayuna, NY 12309

*Eric Hayman, Lourdes de Agapito and Ian Reid
Department of Engineering,
Oxford University

Abstract

This paper considers the problem of self-calibration of a camera from an image sequence in the case where the camera's internal parameters (most notably focal length) may change. The problem of camera self-calibration from a sequence of images has proven to be a difficult one in practice, due to the need ultimately to resort to non-linear methods, which have often proven to be unreliable. In a stratified approach to self-calibration, a projective reconstruction is obtained first and this is successively refined first to an affine and then to a Euclidean (or metric) reconstruction. It has been observed that the difficult step is to obtain the affine reconstruction, or equivalently to locate the plane at infinity in the projective coordinate frame. The problem is inherently non-linear and requires iterative methods that risk not finding the optimal solution. The present paper overcomes this difficulty by imposing cheirality constraints to limit the search for the plane at infinity to a 3-dimensional cubic region of parameter space. It is then possible to carry out a dense search over this cube in reasonable time. For each hypothesised placement of the plane at infinity, the calibration problem is reduced to one of calibration of a non-translating camera, for which fast non-iterative algorithms exist. A cost function based on the result of the trial calibration is used to determine the best placement of the plane at infinity. Because of the simplicity of each trial, speeds of over 10,000 trials per second are achieved on a 256Mhz processor. It is shown that this dense search allows one to avoid areas of local minima effectively and find global minima of the cost function.

1 Introduction

Self calibration of a camera from image sequences has been the subject of much recent research since the seminal paper of Maybank and Faugeras [8]. Practical methods for computing the calibration have been given by [3, 11], but there have been remaining problems of convergence and stability. The basic assumption behind these papers is that the camera is the same for all views, which means that no zooming is allowed. More recently, the observation was made in [6, 7, 9] that self calibration is possible under much looser assumptions. Calibration is possible under a minimal assumption that the cameras have no skew, or the slightly tighter additional assumption that the pixels have a fixed or known aspect ratio. This extension of the theory of self-calibration allows calibration to be carried out on video sequences with a zooming camera.

In a parallel development, self-calibration methods have been given for non-translating cameras (that is ones that do not change their focal centre). The advantage of these methods is that more robust, simpler and often linear methods are available for self-calibration in this case, both for unchanging ([4]) and changing ([1, 2]) internal parameters. The theme of this paper is that these simple methods are applicable to the case of cameras undergoing general motion, once the plane at infinity has been determined. Indeed they may be used to guide the search for the plane at infinity in a projective reconstruction. The result is a stratified algorithm for self calibration, applicable to cameras undergoing general motion, with changing internal parameters, in which one proceeds from projective to quasi-affine to affine to Euclidean reconstruction.

The only other approaches to this calibration problem have been given in [6, 7, 9]. The method of Pollefeys et al. ([9]) applies a straight projective-to-Euclidean iterative approach in which it is necessary to make assumptions about internal parameters in order to initialize the iteration, and Heyden's approach ([6]) is similarly iterative. As this paper shows, iteration is quite chancy in the context of self-calibration under minimum assumptions on internal parameters. We replace the need for descent-based iteration by a quick, but exhaustive search for the best affine reconstruction. The method described is very effective at finding the global minimum of a calibration cost function, and lends it-

*Financial support has been provided by EPSRC (Advanced Research Fellowship to IDR and GR/L58668), Spanish Ministry of Education and Science (LA) and Research Coucil of Norway (EH).

self to generalization by the use of different costs and affine-to-euclidean reconstruction schemes.

2 Calibration of a non-translating camera

A method for computing the calibration of a rotating and zooming camera was given in [2] and is summarized here. The method given there relies on the fact that images obtained by such a camera are related by image-to-image homographies, otherwise known as 2D projective transformations. One selects a reference image J_0, and assigns homographies $\mathtt{H}_i$ to each of the other images. The homographies $\mathtt{H}_i$ are defined by the following condition. If $\mathbf{x}_0$ is any point in the image J_0 and x_j is the corresponding point in the image J_j, then $\mathbf{x}_j = \mathtt{H}_j \mathbf{x}_0$. The homographies may be computed by direct measurement of matching points in the set of images, as described in [4]. Since matching points are mapped to each other in this manner, the same is true for the points on the image of the absolute conic (the IAC). Denoting by $\boldsymbol{\omega}_i$ the IAC in the i-th image, the rule for transforming conics under a homography leads to an equation

$$\boldsymbol{\omega}_j = \mathtt{H}_j{}^{-\top} \boldsymbol{\omega}_0 \mathtt{H}_j^{-1} \quad (1)$$

The IAC is related to the calibration matrix of each camera by the formula ([4])

$$\boldsymbol{\omega} = \mathtt{K}^{-\top}\mathtt{K}^{-1} \quad (2)$$

where $\mathtt{K}$ is the calibration matrix of the camera. The entries of the matrix $\boldsymbol{\omega}$ are readily related to the entries of the calibration matrix $\mathtt{K}$ in the case where the skew parameter (that is $s = \mathtt{K}_{12}$) is zero[1].

In particular, if

$$\mathtt{K} = \begin{bmatrix} \alpha_x & 0 & x_0 \\ & \alpha_y & y_0 \\ & & 1 \end{bmatrix} \quad (3)$$

one may easily compute that

$$\begin{aligned} \boldsymbol{\omega} &= \mathtt{K}^{-\top}\mathtt{K}^{-1} \\ &= \begin{bmatrix} 1/\alpha_x^2 & 0 & -x_0/\alpha_x^2 \\ 0 & 1/\alpha_y^2 & -y_0/\alpha_y^2 \\ -x_0/\alpha_x^2 & -y_0/\alpha_y^2 & 1 + x_0^2/\alpha_x^2 + y_0^2/\alpha_y^2 \end{bmatrix} \end{aligned}$$

Assumptions about the calibration matrix are now easily related to conditions on the entries of $\boldsymbol{\omega}$. Specifically,

Proposition 1.

1. **Zero-skew :** *If* $s = \mathtt{K}_{12} = 0$, *then* $\boldsymbol{\omega}_{12} = 0$.
2. **Square-pixels :** *If* $s = 0$ *and* $\alpha_x = \alpha_y$, *then* $\boldsymbol{\omega}_{11} - \boldsymbol{\omega}_{22} = 0$.
3. **Known principal point :** *If* $s = 0$ *and* $x_0 = 0$, *then* $\boldsymbol{\omega}_{13} = 0$. *Similarly if* $y_0 = 0$ *then* $\boldsymbol{\omega}_{23} = 0$.

[1] This is not true of the dual of the IAC, used in [1, 4] and is the reason for the significant advantage of using the IAC instead of its dual.

Each equation of this type applied to $\boldsymbol{\omega}_j$, when combined with (1) gives a linear equation in the entries of $\boldsymbol{\omega}_0$. With at least five equations, one may solve for the five distinct entries of $\boldsymbol{\omega}_0$ up to scale. For instance, each image (including) J_0 gives one equation in the zero-skew case, and five images are required to find the calibration. If one assumes square pixels or known principal point, then fewer images are necessary. Finally, one computes each $\boldsymbol{\omega}_j$ using (1) and retrieves the calibration matrix from (2) using Cholesky factorization. Further details, and results of implementing this algorithm are reported in [2].

3 The general motion case

The purpose of this paper is to extend the techniques of [2] described above to the case of a moving camera, that is one undergoing translation as well as rotation. In addition, the camera may be zooming, which means that the internal parameters are changing. The problem is made considerably more difficult by the translation of the camera. In this case, there are no homography maps that map points directly from one image to the next. Instead, matching points are related by the fundamental matrix. However, the place of the inter-image homographies in the stationary camera case are taken by the so-called “infinite homographies” (as described later) in the case of moving cameras. Unfortunately, one can not compute the infinite homographies without knowing the position of the plane at infinity, which is not readily found. However, if only these infinite homographies could be discovered, then the theory and solution method given for the stationary camera case could be applied directly to find the calibration of the cameras. Thus, calibration is reduced to finding the plane at infinity. As has been noted previously ([4, 12]), the real difficulty in calibration is finding the plane at infinity.

Given a set of images of a scene, the first step in the calibration process is to compute a projective reconstruction of the scene, using the method described in [3], or any other method. The next step is to find the true plane at infinity in the coordinate frame of the projective reconstruction. Applying a projective transformation that takes this plane to infinity upgrades the projective reconstruction to an affine-reconstruction. After this, one can find the infinite homographies and use the method of [2] to compute the calibration of each camera, and ultimately compute a Euclidean (sometimes called “metric”) reconstruction of the scene.

The method of finding the plane at infinity proposed in this paper is to carry out a direct search over all possible planes to find the one that allows the best calibration. Thus, let $\mathtt{P}_j = [\mathtt{M}_j|\mathbf{t}_j]$ be a set of cameras and $\mathbf{X}_i$ be points together making up a projective reconstruction of a scene. Let $\mathbf{V}$ be a 4-vector representing a plane in the reconstructed scene, and let $\mathtt{G}$ be a projective transformation taking $\mathbf{V}$ to the plane at infinity. Any non-singular 4×4 matrix with 4-th row equal to $\mathbf{V}^\top$ has this property. To verify this, note that a point $\mathbf{X}$ lies on the plane represented by $\mathbf{V}$ if and only if $\mathbf{V}^\top\mathbf{X} = 0$, and this is equivalent to the condition that $\mathtt{G}\mathbf{X} = (X, Y, Z, 0)^\top$ lies on the plane at infinity. Now, we may apply $\mathtt{G}$ to the reconstruction, replacing each $\mathtt{P}_j$ by $\mathtt{P}_j\mathtt{G}^{-1}$ and each point $\mathbf{X}_i$ by $\mathtt{G}\mathbf{X}_i$. If by some chance $\mathbf{V}$ represented the true plane at infinity, then we now have an affine reconstruction, and one may proceed to calibrate the cameras, as described next.

For each j, let $[\mathtt{M}'_j|\mathbf{t}'_j] = \mathtt{P}'_j = \mathtt{P}_j\mathtt{G}$ be the camera matrices after transformation. The infinite homography for a pair of cameras is defined to be the homography between the two images relating the respective projections of points lying on the plane at infinity. For instance, a point $\mathbf{X} = (\mathbf{x}^\top, 0)^\top$ on the plane at infinity maps to points $\mathtt{M}'_i\mathbf{x}$ and $\mathtt{M}'_j\mathbf{x}$ in the i-th and j-th images. These points are related by the homography mapping $\mathtt{H}_{ij} = \mathtt{M}'_j\mathtt{M}'^{-1}_i$, which is the infinite homography for this pair of images. Since the absolute conic lies on the plane at infinity, its projections in two images are related via the infinite homography. From this it follows that (1) holds where $\mathtt{H}_j = \mathtt{H}_{0j}$ is the infinite homography for the image pair $(0, j)$.

For convenience, we introduce a reference camera represented by $\mathtt{P}'_0 = [\mathtt{I} \mid \mathbf{0}]$. Note that $\mathtt{P}'_0$ is **not** necessarily one of our cameras $\mathtt{P}'_j$. Let $\boldsymbol{\omega}_0$ be the IAC in the image taken with this camera. The infinite homography from image 0 to image j is then simply $\mathtt{M}'_j$, from which it follows that

$$\boldsymbol{\omega}_j = \mathtt{M}'^{-\top}_j \boldsymbol{\omega}_0 \mathtt{M}'^{-1}_j \tag{4}$$

As in the stationary camera case, each condition of the form given in Proposition 1 gives a linear equation on the entries of the symmetric matrix $\boldsymbol{\omega}_0$. One may solve this system to find $\boldsymbol{\omega}_0$, subsequently compute each $\boldsymbol{\omega}_j$ and obtain $\mathtt{K}_j$ by Cholesky factorization. If the computed $\boldsymbol{\omega}_j$ is not positive-definite, then this final step of Cholesky factorization is not possible. This factor will work in our favour, since it indicates that the supposed value of $\mathbf{V}$ representing the plane at infinity was in error. Note that each $\boldsymbol{\omega}_j$ is positive-definite if and only if $\boldsymbol{\omega}_0$ is.

The cost function. The complete set of equations derived from (4) may be written as $\mathtt{A}\mathbf{w} = \mathbf{0}$, where $\mathbf{w}$ is a 6-vector made up of the distinct entries of the symmetric matrix $\boldsymbol{\omega}_0$. If the plane at infinity (represented by $\mathbf{V}$) was correctly devined, and there is no noise, then this set of equations will have an exact non-zero solution, representing the matrix $\boldsymbol{\omega}_0$. One may find this solution by carrying out the Singular Value Decomposition (SVD) of $\mathtt{A}$, namely $\mathtt{A} = \mathtt{UDV}^\top$. Matrix $\mathtt{D}$ is a 6×6 diagonal matrix $\mathrm{diag}(d_1, \ldots, d_6)$, and we may arrange that d_6 is the smallest entry, equal to zero in the ideal case, and the solution is the last column of $\mathtt{V}$. In the absence of an exact solution, the last column of $\mathtt{V}$ represents a least-squares solution for $\mathbf{w}$, and the residual $\mathtt{A}\mathbf{w}$ is $d_6\mathbf{U}_6$, where $\mathbf{U}_6$ is the last column of $\mathtt{U}$. The magnitude of this residual vector is equal to d_6, the smallest singular value of $\mathtt{A}$.

Although measurement noise will lead to an inexact solution, and a non-zero residual, the most significant source of residual is the wrong placement of the plane at infinity. This suggests a strategy for finding the correct placement of the plane at infinity, as follows.

Algorithm 1.

1. For each 4-vector $\mathbf{V}$ representing a plane at infinity, compute a residual error $r(\mathbf{V})$ as follows
   (a) Form a 4×4 transformation matrix having last row $\mathbf{V}$ and compute transformation matrices $\mathtt{P}'_j = \mathtt{P}_j\mathtt{G}^{-1} = [\mathtt{M}'_j|\mathbf{t}'_j]$.
   (b) Form a linear equation set $\mathtt{A}\mathbf{w} = \mathbf{0}$ from (4) according to chosen constraints of the type given in Proposition 1
   (c) Compute the SVD of $\mathtt{A} = \mathtt{UDV}^\top$.
   (d) The last column of $\mathtt{V}$ is the least-squares solution containing the entries of $\boldsymbol{\omega}_0$.
   (e) If the computed $\boldsymbol{\omega}_0$ is not positive-definite, then reject the solution.
   (f) The magnitude of the residual $r(\mathbf{V})$ is the smallest singular value of $\mathtt{A}$.
2. Search for the value of $\mathbf{V}$ that minimizes the residual $r(\mathbf{V})$, and leads to a positive-definite solution for $\boldsymbol{\omega}_0$. Accept this value of $\mathbf{V}$ as the placement of the plane at infinity,
3. From the value of $\boldsymbol{\omega}_0$ computed as step 1, compute each $\boldsymbol{\omega}_j$ using (4), and compute $\mathtt{K}_j$ using Cholesky factorization $\boldsymbol{\omega}_j^{-1} = \mathtt{K}_j\mathtt{K}_j{}^\top$. Since $\boldsymbol{\omega}_0$ is positive definite, so is $\boldsymbol{\omega}_j$, and this factorization will succeed.

The cost function represented by the smallest singular value of $\mathtt{A}$ represents the residual error associated with the conditions given in Proposition 1. It is possible to use other cost functions in this context. In the zero-skew case, for instance, the vector of skew-angles for each of the different calibration matrices $\mathtt{K}_j$ has also been tried, and seems preferable. However, this represents a minimal modification to the algorithm.

4 Narrowing the Search for Infinity

The 4-vector $\mathbf{V}$ representing the plane at infinity is defined up to scale only, and hence the search for an optimum $\mathbf{V}$ may be carried out over a compact region by searching over a 3-sphere. However, it is possible to constrain the search even more effectively.

4.1 Obtaining a quasi-affine reconstruction.

The search for the best plane at infinity can be narrowed by making a preliminary transformation to a quasi-affine transformation ([5]). A quasi-affine reconstruction is computed by taking account of cheirality as described in [3]. In this paper, the technique is refined to give accurate bounds for a search for the plane at infinity.

As described in [3], the essence of cheirality is to use information about which points are visible in an image, and hence in front of the camera, to upgrade a projective reconstruction to a "quasi-affine" reconstruction. A quasi-affine reconstruction is a projective reconstruction in which the reconstructed scene is not split across the plane at infinity. A quasi-affine reconstruction may be computed from a projective-reconstruction by solving a linear programming problem. In particular, for a projective reconstruction consisting of points $\mathbf{X}_i$ and cameras $\mathtt{P}_j$, one finds a quasi-affine reconstruction in several steps as follows (for justification see [5, 3]) :

1. Multiply each $\mathtt{P}_j$ and $\mathbf{X}_i$ by ± 1 as necessary so that $\mathtt{P}_j\mathbf{X}_i = (x, y, w)^\top$ with $w > 0$. This is always possible ([5]).

2. For any camera matrix $\mathtt{P}$, let $\mathbf{C}^{\mathtt{P}} = (c_1, c_2, c_3, c_4)^\top$ be defined by $c_k = (-1)^k \det(\hat{\mathtt{P}}^k)$, where $\hat{\mathtt{P}}_k$ is obtained from $\mathtt{P}$ by removing the k-th column. The vector $\mathbf{C}^{\mathtt{P}}$ is a homogeneous representation of the camera centre, but the sign of $\mathbf{C}^{\mathtt{P}}$ is important in this context.

3. For each $\epsilon = \pm 1$, form the set of inequalities

$$\begin{aligned} {\mathbf{X}_i}^\top \mathbf{V} > 0 \quad &\text{for all points } \mathbf{X}_i \\ \epsilon {\mathbf{C}^{\mathtt{P}_j}}^\top \mathbf{V} > 0 \quad &\text{for all cameras } \mathtt{P}_j \end{aligned}$$

4. For each choice of ϵ solve the set of inequalities to find $\mathbf{V}$. There must be a solution for at least one of the choices of ϵ, perhaps both.

5. Choose a transformation matrix $\mathtt{G}$ with 4-th row equal to $\mathbf{V}$ and such that $\text{sign}(\det \mathtt{G}) = \epsilon$.

6. Replace each $\mathbf{X}_i$ by $\mathtt{G}\mathbf{X}_i$ and each $\mathtt{P}_j$ by $\mathtt{P}_j\mathtt{G}^{-1}$, and the resulting reconstruction will be a quasi-affine reconstruction.

The inequalities above are called the "cheiral inequalities". If solutions exist both for $\epsilon = 1$ and $\epsilon = -1$, then they have opposite orientation ([5]). As explained in [3] one adds extra inequalities $|v_i| <= 1$ for each component v_i of $\mathbf{V}$ to constrain a solution. In order to find a unique solution to the set of inequalities, the cheiral inequalities are modified by introducing a further variable δ. Then one forms the inequalities ${\mathbf{X}_i}^\top \mathbf{V} > \delta$ and $\epsilon {\mathbf{C}^{\mathtt{P}_j}}^\top \mathbf{V} > \delta$ and linear programming is used to find the solution that maximizes $\delta > 0$.

This method is used to obtain one or possibly two differently oriented quasi-affine reconstructions of the scene. One of these reconstructions differs from the true reconstruction by an orientation preserving (that is positive determinant) projective transformation in which the plane mapped to infinity does not cross the convex hull of the points and camera centres. Any further transformations applied to the reconstruction to achieve an affine or Euclidean reconstruction will have this form. Subsequent steps of the algorithm are carried out with the two differently oriented quasi-affine reconstructions (if both exist) until some time later a choice is made between them.

4.2 Translation to the origin

The next step is to translate the quasi-affine reconstruction (that is the 3D points and the camera centres) to the coordinate origin. At the same time, to make possible affine distortion more benign, anisotropic scaling is applied to make the reconstruction approximately round. More precisely a scaling is applied so that the principal moments of the point set are equal. This complete transformation is easily done in one step as follows. Let $S = 1/N \sum_1^N {\mathbf{Y}_i}^\top \mathbf{Y}_i$ be the scatter matrix of all the points $\mathbf{Y}_i$, which are the 3D points and the camera centres. Let $\mathtt{S} = \mathtt{H}\mathtt{H}^\top$, with $\mathtt{H}$ upper-triangular, be the Cholesky factorization of $\mathtt{S}$. Transforming the reconstruction by $\mathtt{H}^{-1}$, that is replacing $\mathtt{P}_j$ by $\mathtt{P}_j\mathtt{H}$ and points $\mathbf{X}_i$ by $\mathtt{H}^{-1}\mathbf{X}_i$, will carry out the desired translation and scaling of the data. Note that $\mathtt{H}$ represents an affine transformation.

4.3 Setting bounds on the plane at infinity

At the end of the previous step, one has a quasi-affine reconstruction centred on the origin. Next we want to upgrade to an affine reconstruction which requires the plane at infinity to be determined. In any further transformation to be applied to the reconstruction to achieve an affine reconstruction, the plane $\mathbf{V}$ mapped to infinity can not pass through the origin (since then it splits the point set). Hence one may assume that the vector $\mathbf{V}$ mapped to infinity is of the form $\mathbf{V} = (v_1, v_2, v_3, 1)^\top = (\mathbf{v}^\top, 1)^\top$. Furthermore, the transformation $\mathtt{G}_{\mathbf{V}}$ mapping $\mathbf{V}$ to infinity may be taken to have the form

$$\mathtt{G}_{\mathbf{v}} = \begin{bmatrix} \mathtt{I} & \mathbf{0} \\ \mathbf{v}^\top & 1 \end{bmatrix}$$

which has unit determinant, and is hence orientation preserving.

A search for the plane at infinity has thus been reduced to a search over the 3-dimensional space represented by the coordinates of the vector $\mathbf{v}$. Next it is shown how this search may be narrowed to a search over a rectangular region of parameter space, somewhat incorrectly here called a cube. First note that the cheiral inequalities may be written in terms of a matrix $\mathtt{C}$, and a vector $\mathbf{V} = (\mathbf{v}^\top, 1)^\top$ is a viable plane at infinity if and only if each component of the vector $\mathtt{C}\mathbf{V}$ is positive. We set $\epsilon = 1$ in forming these inequalities, since we are now interested only in orientation-preserving transforms. This condition gives a very rapid test for a proposed plane at infinity being acceptable.

The plane at infinity must lie outside of the convex hull of the scene, which is centred around the origin. This constraint places finite bounds on the coordinates of $\mathbf{v}$, since planes with unbounded coordinates lie arbitrarily close to the origin. One may determine the bounds for the coordinates of $\mathbf{v}$ by linear programming with the constraint matrix $\mathtt{C}$. One obtains upper and lower bounds for each v_i (six problems in total) by maximizing $\pm v_i$ subject to the constraints $\mathtt{C}\mathbf{V} > \mathbf{0}$. This limits the search for the plane at infinity to a search over a cube containg the origin. Typically, one finds bounds on v_i of the order of $-1.0 < v_i < 1.0$.

4.4 Searching for the plane at infinity

Searching over the complete cube for the best value of $\mathbf{v}$ is quite tractable, and is the preferred method. In our implementation, we take 50 samples in each coordinate direction, a total of $50^3 = 125,000$ trials in all. Each trial is represented by a vector $\mathbf{V} = (v_1, v_2, v_3, 1)^\top$. The following steps are carried for each such trial vector.

1. Cheirality test : If $\mathtt{C}\mathbf{V}$ is not a positive vector, then reject this trial.

2. Otherwise, for each camera matrix $\mathtt{P}_j = [\mathtt{M}_j | \mathbf{t}_j]$, compute $\mathtt{M}'_j = \mathtt{M}_j - \mathbf{t}_j \mathbf{v}^\top$. Note that this $\mathtt{M}'_j$ is the left-hand block of $\mathtt{P}'_j = \mathtt{P}_j \mathtt{G}_{\mathbf{v}}^{-1}$.

3. Form a linear equation set $\mathtt{A}\mathbf{w} = \mathbf{0}$ from (4) according to constraints as in Proposition 1 and solve to find $\boldsymbol{\omega}_0$.

4. IAC test : If $\boldsymbol{\omega}_0$ is not positive definite (determined using the Cholesky factorization), then reject this trial.

5. Otherwise, return a cost value (or vector) associated with the computed calibration.

The complete self-calibration algorithm is given by Algorithm 1 in which this search technique is used as step 2.

Although a blanket search may seem costly, in fact it is very fast, since the computational cost at each step is small. In fact, for a sequence of 19 images with over 1000 points, the search over 250,000 trials (both orientations) took only 23 seconds on a 256MHz Pentium machine. This time is insignificant, compared with the time taken for point tracking outlier detection and accurate bundle-adjusted projective reconstruction. Furthermore, one could probably reduce the density of search by half with little loss, thereby reducing the search time by a factor of 8.

Iterative search To obtain a more accurate estimate of the plane at infinity, one can carry out an iterative cost minimization to find the exact cost minimum. We used Levenberg-Marquardt starting at the minimum of the exhaustive search to minimize the cost vector with respect to $\mathbf{v}$. Since the minimum is very close, this search terminates very quickly (in a few milliseconds).

5 Convergence

The small dimension of the search space allows us to get some idea of how well-behaved the cost function is. Once the minimum of the cost function was found on the cube containing $\mathbf{v}$, values of the cost function on the axial planes passing through the minimum were plotted. The results are shown and discussed in Fig 1.

5.1 Experimental results with real images

The calibration algorithm was tested with generally satisfactory results on various different image sequences, including the LIFIA model house sequence used in [3]. In addition, a new image sequence was taken using a camera with a zoom lens mounted on a Yorick stereo head/eye platform [10]. To achieve general motion we used three of the four available degrees of freedom of the head using one of the two independent vergence axes, the common elevation axis and the pan axis to translate and rotate the camera. The servo lens provided ground truth data of the position of the zoom lens for each frame in the image sequence. The camera was then calibrated, using an accurately machined calibration grid and a classical calibration algorithm, to obtain ground truth values for the internal parameters at each of the different positions of the zoom lens. The focal length of the camera was set to increase linearly by a factor of approximately 1.4, using the controlled zoom lens. Figure 2 shows 6 of the 15 images of the sequence.

First, point correspondences were computed and the projective reconstruction was obtained using the method reported in [3]. The experiment was run using both the *zero*

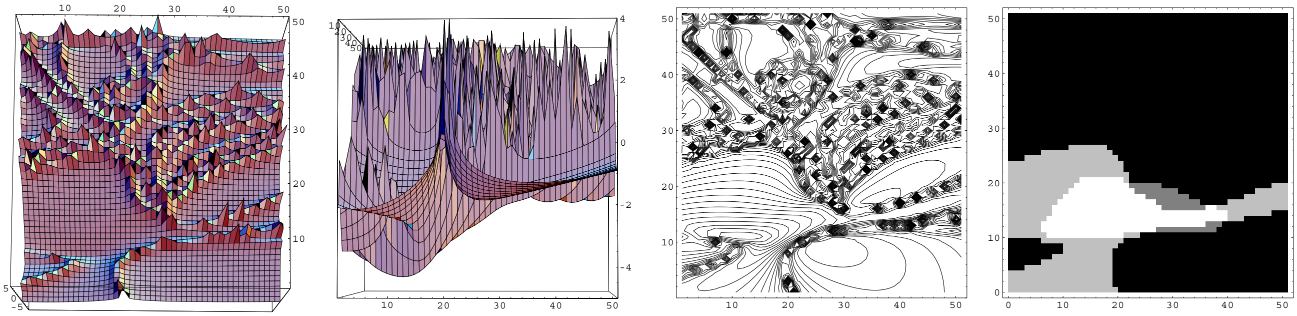

Figure 1. *These figures show the xy-plane cross section through the search cube. The first two plots show the shape of the cost surface (on a logarithmic scale) viewed from above and from the side, whereas the third plot shows a contour plot of the same slice. As may be seen, the surface is highly indented, and any attempt to converge to a global minimum through gradient-guided search from a random point in the cube would be doomed to failure. The last plot shows the shape of the regions defined by the various constraints on the vector* $\mathbf{v}$ *representing the plane at infinity. the dark-grey region shows the area in which only the cheirality constraint is satisfied. The light-grey region represents points satisfying the positive-definite IAC test constraint. The white region shows where both constraints are satisfied. The cost function needs to be computed only inside this region. Note also that the cost function is relatively well behaved inside the white region. Note that the region defined by the cheirality constraint is a convex polyhedral region inside the search cube, but the region defined by the positive definite IAC constraint is more complex.*

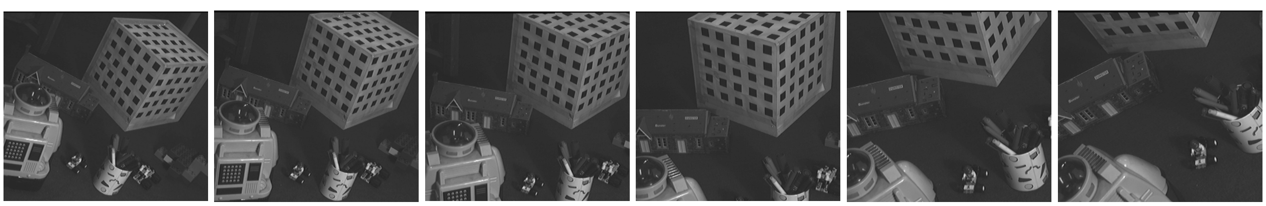

Figure 2. *Six of the images from a 19-image sequence used in these experiments.*

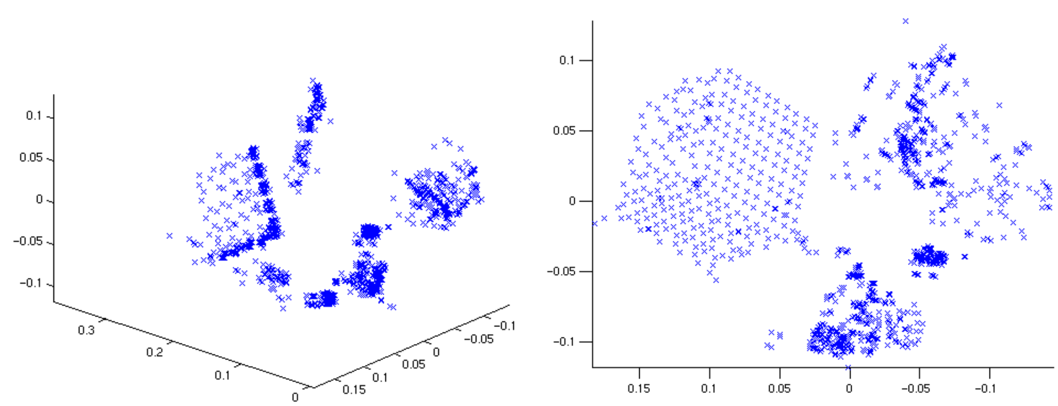

Figure 3. *Top view and side view of the reconstructed scene points as computed using the square-pixel constraint. These results show that the Euclidean structure of the scene is well captured, as evidenced by the square shape of the calibration cube shown in the scene. The reconstruction results using only the zero-skew constraint were less satisfactory as can be expected.*

skew and *square-pixel* constraints. Two differently oriented quasi-affine reconstructions were found in each case, but in each case only one of them lead to a Euclidean reconstruction.

Figure 4 shows the results obtained for some of the calibration parameters. The aspect ratio appears to be very well estimated and remains almost constant throughout the sequence. The estimate of the focal length was close to the ground truth value, the maximum error being around 5%. In this image sequence, the principal point (not shown) was badly conditioned and the results were inherently unstable, being correlated with camera rotation. However it was computed to lie always within the image.

Once the Euclidean calibration was computed the projective reconstruction of the scene points was upgraded to a Euclidean reconstruction. A top view and a side view of the reconstructed points is presented in figure 5.1. Note that the metric structure of the scene is well preserved.

6 Discussion and Conclusions

The shape of the cost surfaces arising from the self-calibration problem demonstrates the difficulties that are involved in cost-minimization to identify the plane at infinity, and hence make the step to an affine reconstruction. This suggests that it is imperative to take note of the constraints arising from considerations of cheirality, and also the positive-definiteness of the image of the absolute conic if one is to hope to find a robust affine, and subsequently Euclidean reconstruction of a scene. The technique of densely sampled search over the permissible range of the plane at infinity discussed in this paper has proven to be an effective way of robustly and rapidly finding a global minimum for the calibration cost.

In some cases, it has been observed, however that despite finding a cost function minimum, stable values of the camera calibration parameters are often not obtainable. This is especially true in the case where minimal constraints are applied, such as the zero-skew constraint alone. In this particular case, the calibration of the camera is obtained essentially independently of the other cameras, and one is subject to the usual ambiguities such as principal-point / rotation and focal-length / distance.

The method outlined in this paper is applicable with a large number of different cost functions and calibration methods, and a continuing research goal is to find which cost functions give the best results. As an example the goal function given here, which minimizes sum of squares of skew parameters seems natural for the case of assumed zero skew. However, other function, such as skew angle may give better results. The search method is not limited to non-iteratively computable cost functions such as those discussed here. Use of an iterative calibration algorithm for the search trials, such as that of [1] can allow other calibration constraints, such as fixed but unknown principal point.

References

[1] L. de Agapito, E. Hayman, and I. D. Reid Self-calibration of a rotating camera with varying intrinsic parameters. In *Proc. British Machine Vision Conference*, pages 105–114, 1998.

[2] L. de Agapito, R. I. Hartley and E. Hayman Linear calibration of a rotating and zooming camera. *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pages 15-20, 1999.

[3] R. I. Hartley. Euclidean reconstruction from uncalibrated views. *Applications of Invariance in Computer Vision: Proc. of the Second Joint European - US Workshop, Ponta Delgada, Azores – LNCS-Series Vol. 825, Springer Verlag*, pages 237–256, October, 1993.

[4] R. I. Hartley. Self-calibration of stationary cameras. *International Journal of Computer Vision*, volume 22, number 1, pages 5–23, February, 1997.

[5] R. I. Hartley. Chirality. *International Journal of Computer Vision*, volume 26, number 1, pages 41–61, 1998.

[6] A. Heyden and K. Åström. Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1997.

[7] A. Heyden and K. Åström. Minimal conditions on intrinsic parameters for euclidean reconstruction. In *Proc. Asian Conference on Computer Vision*, Hong Kong, 1998.

[8] S. Maybank and O. Faugeras. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, 1992.

[9] M. Pollefeys, R. Koch, and L. Van Gool. Self calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *Proc. 6th International Conf. on Computer Vision, Bombay*, pages 90–96, 1998.

[10] P. M. Sharkey, D. W. Murray, S. Vandevelde, I. D. Reid, and P. F. McLauchlan. A modular head/eye platform for real-time reactive vision. *Mechatronics*, 3(4):517–535, 1993.

[11] W. Triggs. Auto-Calibration and the absolute quadric. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 609–614, 1997.

[12] A. Zisserman, D. Liebowitz, and M. Armstrong. Resolving ambiguities in auto-calibration. Philosophical Transactions of the Royal Society of London. Vol 356(1740):1193–1211, 1998.

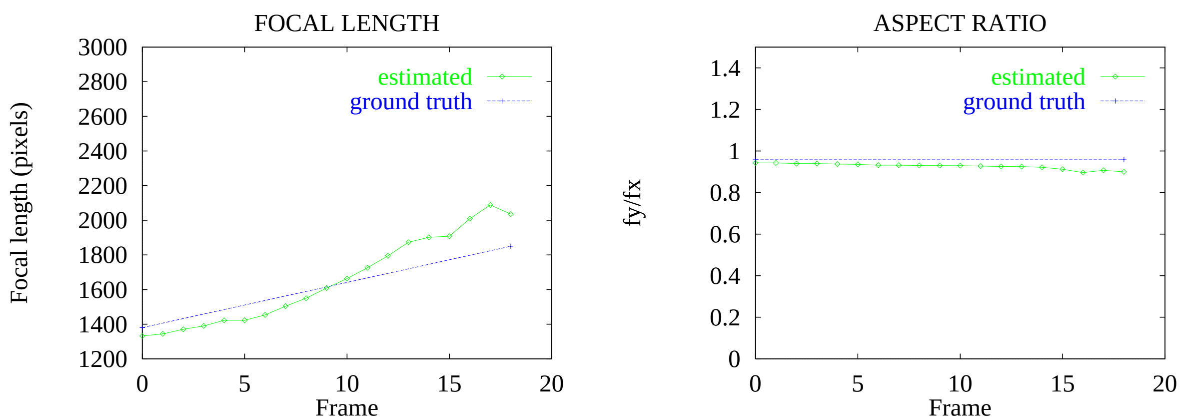

Figure 4. *Ground truth and computed values for the focal length and the aspect ratio of the camera. These results shown here were for the zero-skew case and show surprising accuracy, given the very minimal calibration assumptions being used. Note that the aspect ratio is very accurately estimated, although no assumption is made as to its value, or even to the fact that it is constant across all images. Other parameters, notably principal point position were not very accurately computed, confirming the known fact that the position of the principal point is not well conditioned.*