A Unification of Autocalibration Methods

Long Quan, Bill Triggs

A Unification of Autocalibration Methods

Long QUAN Bill TRIGGS
CNRS-GRAVIR-INRIA
ZIRST – 655 avenue de l'Europe
38330 Montbonnot, France

Abstract

This paper describes a unified theory for autocalibration of a moving camera. The camera models can be 2D projective, 2D affine camera and 1D projective projections. The scenes can be either non-planar scenes or planar scenes and the camera may also undergoes a pure rotation. All these cases are unified into the same framework based on the direction bases. Each formulation on the direction bases is also paralleled with that on projective geometry concepts, essentially encoded by the various forms of the absolute conic. This unification provides not only a common theoretical framework, but also suggests unified parameterisation schema for estimation procedures.

1 Introduction

Autocalibration is the recovery of metric information from only point correspondences of uncalibrated images, using geometric self-consistency constraints. The potential advantages of autocalibration are a reduced need for off-line calibration and greater on-line flexibility.

We will assume familiarity with the projective approach to vision: 2D and 3D projective spaces, homogeneous coordinates, projective scene reconstruction techniques . . . Projective reconstruction is actually the simplest type of autocalibration. Using only $m \geq 2$ uncalibrated projective images and linear algebra, it recovers the entire 3D scene and camera geometry modulo just 9 unknown parameters: 1 overall scale, 5 affine relative scalings and skewings, and 3 essentially projective displacements of the plane at infinity. However, this paper concentrates on the recovery of these last few 'metric' parameters, given (information equivalent to) a projective reconstruction. In fact, the overall scale is never recoverable. We will use the term 'Euclidean' to mean 'metric up to scale', *i.e.* modulo a Euclidean similarity transformation.

Since the seminal work of Maybank & Faugeras [15, 5], a number of different approaches to autocalibration have been developed [7, 8, 1, 33, 32, 4, 14, 10, 18, 17, 29, 11]. For the 'classical' problem of a single perspective camera with constant but unknown internal parameters moving with a general but unknown motion in a 3D scene, the original Kruppa equation based approach [15] seems to be being displaced by approaches based on the 'rectification' of an intermediate projective reconstruction [7, 10, 17, 29, 11, 8] of which Trigg's formulation based on absolute quadric [29] has been the most significant. More specialized methods exist for particular types of motion, particular scenes and simplified calibration models [8, 31, 1, 18, 30]. Affine cameras [22], 1D cameras [6] and Stereo heads [33, 12] can also be autocalibrated. Solutions are still — in theory — possible if some of the intrinsic parameters are allowed to vary [10, 17]. The numerical conditioning of classical autocalibration is historically delicate, although recent algorithms have improved the situation significantly [10, 17, 29]. The main problem is that classical autocalibration has some restrictive intrinsic degeneracies — classes of motion for which no algorithm can recover a full unique solution. Sturm [27, 28] has given a catalogue of these. In particular, at least 3 views, some translation and some rotation about at least two non-aligned axes are required. Further work on the degeneracies have also been studied in [2, 13]. Most of the materials presented in this paper come from [29, 30, 22, 6].

2 Preliminaries

2.1 Notation and Basics

Throughout the paper, vectors and matrices are respectively denoted by lower and upper case bold letters. We use P for image projections and H for inter-image homographies; K, for upper triangular camera calibration and $C = K^{-1}$ for its inverse; Λ for the absolute quadric (in plane coordinates) or the dual absolute conic Λ^* and $Q = K\,K^\top = P\,\Lambda\,P^\top$ for its images; Λ^* for the absolute conic (in point coordinates); and $Q^{-1} = C^\top C$ for its image; Λ_∞ for the dual absolute conic on the known plane at infinity and $\Lambda_{3\times 3}$ for the absolute quadric (in line coordinates) or the dual absolute points. The symbol $\sim$ is for homogeneous equality, *i.e.*equality up to a constant non-zero scale.

Introducing homogeneous Euclidean coordinates for k-dimensional Euclidean spaces, points, direction

Published in Asian Conference on Computer Vision — ACCV 2000.
Work supported by Esprit LTR project CUMULI.

vectors and hyperplanes are encoded respectively as homogeneous $k+1$ component column vectors $\boldsymbol{x} = (\boldsymbol{x}_e, 1)^\top$, $\boldsymbol{t} = (\boldsymbol{t}_e, 0)^\top$ and row vectors $\boldsymbol{p}^\top = (\boldsymbol{n}_e^\top, d)$. Points and directions on the plane satisfy respectively $\boldsymbol{p}^\top \boldsymbol{x} = \boldsymbol{n}_e^\top \boldsymbol{x}_e + d = 0$ and $\boldsymbol{p}^\top \boldsymbol{t} = \boldsymbol{n}_e^\top \boldsymbol{t}_e = 0$. Directions can be appended to the point space, as a 'hyperplane at infinity' $\boldsymbol{p}_\infty$ of points at infinity or vanishing points. Projective transformations indifferently mix finite and infinite points. Under a projective transformation encoded by an arbitrary nonsingular $(k+1) \times (k+1)$ matrix T, points and directions (column vectors) transform contravariantly, *i.e.* by T acting on the left: $\boldsymbol{x} \mapsto T\,\boldsymbol{x}$, $\boldsymbol{v} \mapsto T\,\boldsymbol{v}$. To preserve the point-on-plane incidence relation, hyperplanes (row vectors) transform covariantly, *i.e.* by $\boldsymbol{T}^{-1}$ acting on the right: $\boldsymbol{p} \mapsto \boldsymbol{p}\,T^{-1}$.

The usual Euclidean dot product between hyperplane normals is $\boldsymbol{n}_1 \cdot \boldsymbol{n}_2 = \boldsymbol{p}_1\,\Lambda\,\boldsymbol{p}_2^\top$ where the symmetric, rank k, positive semidefinite matrix $\Lambda = \left(\begin{smallmatrix} I_{k\times k} & \boldsymbol{0} \\ \boldsymbol{0} & 0 \end{smallmatrix}\right)$ is called the **absolute quadric**, which is the dual of the absolute conic in space—a space conic in plane coordinates [26, 20]. This quadric Λ encodes the Euclidean structure in projective coordinates. Under projective transformations it transforms contravariantly (*i.e.* like a point) in each of its two indices so that the dot product between plane normals is invariant: $\Lambda \mapsto T\,\Lambda\,T^\top$ and $\boldsymbol{p}_i \mapsto \boldsymbol{p}_i\,T^{-1}$, so $\boldsymbol{p}_1\,\Lambda\,\boldsymbol{p}_2^\top = \boldsymbol{n}_1\,\boldsymbol{n}_2$ is constant. Λ is invariant under Euclidean transformations, but in a general projective frame it loses its diagonal form and becomes an arbitrary symmetric positive semidefinite rank k matrix. When restricted to coordinates on $\boldsymbol{p}_\infty$, Λ becomes nonsingular Λ_∞ and dualizing this dualized form gives the $k \times k$ symmetric positive definite **absolute conic** Λ^*. This measures dot products between direction vectors, just as Λ measures them between hyperplane normals. Λ^* is defined *only* on direction vectors, not on finite points, and unlike Λ it has no unique canonical form in terms of the *un*restricted coordinates.

2.2 Direction bases, Absolute conics and Euclidean structure

In Euclidean coordinates, Λ can be decomposed as a sum of outer products of any orthonormal (in terms of Λ^*) basis of direction vectors: $\Lambda = \sum_{i=1}^k \boldsymbol{e}_i\,\boldsymbol{e}_i^\top$ where $\boldsymbol{e}_i\,\Lambda^*\,\boldsymbol{e}_j = \delta_{ij}$. For example in 2D $\Lambda = \left(\begin{smallmatrix} I_{2\times 2} & 0 \\ 0 & 0 \end{smallmatrix}\right) = \boldsymbol{e}_1\,\boldsymbol{e}_1^\top + \boldsymbol{e}_2\,\boldsymbol{e}_2^\top$ where $\boldsymbol{e}_1 = (1,0,0)$, $\boldsymbol{e}_2 = (0,1,0)$, are the usual unit direction vectors. Gathering the basis vectors into the columns of a $(k+1) \times k$ orthonormal rank k matrix U we have $\Lambda = U\,U^\top$, $\boldsymbol{p}_\infty\,U = \boldsymbol{0}$ and $U^\top \Lambda^* U = I_{k\times k}$. The columns of U span $\boldsymbol{p}_\infty$. All of these relations remain valid in an arbitrary projective frame T and with an arbitrary choice of representative for Λ^*, except that $U \mapsto T\,U$ ceases to be orthonormal.

U is defined only up to an arbitrary $k \times k$ orthogonal mixing of its columns (redefinition of the direction basis) $U \mapsto U\,R_{k\times k}$. Even in a projective frame where U itself is not orthonormal, this mixing freedom remains orthogonal. In a Euclidean frame $U = \left(\begin{smallmatrix} V \\ 0 \end{smallmatrix}\right)$ for some $k \times k$ rotation matrix V, so the effect of a Euclidean space transformation is $U \mapsto \left(\begin{smallmatrix} R & \boldsymbol{t} \\ \boldsymbol{0} & 1 \end{smallmatrix}\right) U = U\,R'$ where $R' = V^\top R V$ is the conjugate rotation: Euclidean transformations of direction bases (*i.e.* on the left) are equivalent to orthogonal re-mixings of them (*i.e.* on the right). This remains true in an arbitrary projective frame, even though U and the transformation no longer *look* Euclidean. This mixing freedom can be used to choose a direction basis in which U is orthonormal up to a diagonal rescaling: simply take the SVD $U'\,\boldsymbol{D}\,\boldsymbol{V}^\top$ of U and discard the mixing rotation $\boldsymbol{V}^\top$. Equivalently, the eigenvectors and square roots of eigenvalues of Λ can be used. Such orthogonal parametrizations of U make good numerical sense [30].

In 2D Euclidean space, given any two orthonormal direction vectors $\boldsymbol{u}, \boldsymbol{v}$, the complex conjugate vectors $\boldsymbol{x}_\pm \equiv \frac{1}{\sqrt{2}}(\boldsymbol{u} \pm i\boldsymbol{v})$ satisfy $\boldsymbol{x}_\pm\,\Lambda^*\,\boldsymbol{x}_\pm^\top = 0$. These complex directions algebraically lie on the absolute conic, and it is easy to check that any complex projective point which does so can be decomposed into two orthogonal direction vectors, its real and imaginary parts. The bottom line is that in the 2D case there is only one such conjugate pair up to complex phase, and these are called 'circular points' which characterize the Euclidean structure of the plane. However for numerical purposes, it is usually easier to avoid complex numbers by using the real and imaginary parts $\boldsymbol{u}$ and $\boldsymbol{v}$ rather than $\boldsymbol{x}_\pm$. The phase freedom in $\boldsymbol{x}_\pm$ corresponds to the 2×2 orthogonal mixing freedom of $\boldsymbol{u}$ and $\boldsymbol{v}$.

Theoretically, the above parametrizations of Euclidean structure are equivalent. Which is practically best depends on the problem. Λ is easy to use, except that constrained optimization is required to handle the rank k constraint $\det \Lambda = 0$. Direction bases U eliminate this constraint at the cost of numerical code to handle their $k \times k$ orthogonal gauge freedom. The absolute conic Λ^* has neither constraint nor gauge freedom, but has significantly more complicated image projection properties and can only be defined once the plane at infinity $\boldsymbol{p}_\infty$ is known and a projective coordinate system on it has been chosen (*e.g.* by induction from one of the images). It is also possible to parametrize Euclidean structure by non-orthogonal Choleski-like decompositions $\Lambda = \boldsymbol{L}\,\boldsymbol{L}^\top$ (*i.e.* the L part of the LQ decomposition of U), but this introduces singularities at maximally non-Euclidean frames unless pivoting is also used.

2.3 Projective images, Cameras and Autocalibration

To recover the metric information implicit in projective images, we need a projective encoding of Euclidean structure. The key to Euclidean structure is the dot product between direction vectors ("points at infinity"), or dually the dot product between (normals to) hyperplanes. The former leads to the stratified "hyperplane at infinity + absolute conic" formulation (affine + metric structure) ([3, 9, 16, 21] for projective structure and [19, 25] for affine structure), the latter to the "absolute quadric" one (cf. [29]). These are just dual ways of saying the same thing. The hyperplane formalism is preferable for 'pure' autocalibration or auto-metric-reconstruction where there is no *a priori* decomposition into affine and metric strata, while the point one is simpler if such a stratification is given.

Since the columns of a 3D direction basis matrix U are indeed 3D direction vectors, its image projection is simply $P\,U$, where P is the usual 3×4 point projection matrix. Hence, the projection of $\Lambda = U\,U^\top$ is the 3×3 symmetric positive definite contravariant image matrix $Q = P\,\Lambda\,P^\top$. Algebraically, this is the image of the absolute quadric, now in dual line coordiates which is dual to the image of the absolute conic. With the traditional Euclidean decomposition $K\,R\,(\boldsymbol{I}\,|-\boldsymbol{t})$ of P into an upper triangular **internal calibration matrix** K, a 3×3 **camera orientation** (rotation) R and an **optical centre** $\boldsymbol{t}$, Q becomes simply $K\,K^\top$. Since Λ is invariant under Euclidean motions, Q is invariant under camera displacements so long as K remains constant. K can be recovered from Q by Choleski decomposition, and similarly the Euclidean scene structure (in the form of a 'rectifying' projective transformation) can be recovered from Λ. The upper triangular **inverse calibration matrix** $C = K^{-1}$ converts homogeneous pixel coordinates to optical ray directions in the Euclidean camera frame. $Q^{-1} = C^\top C$ is the image of the absolute conic.

3 Unified Autocalibration Methods

3.1 Autocalibration of 2D projective camera

Given several images taken with projection matrices $P_i = K_i\,R_i(I|-\boldsymbol{t}_i)$, and (in the same Euclidean frame) an orthogonal direction basis $U = \left(\begin{smallmatrix} V \\ 0 \end{smallmatrix}\right)$, we find that $P_iU \sim K_iR_iV$, moving K_i to the left side gives *i.e.*

$$K_i^{-1}\,P_i\,U \;\sim\; R_i' \tag{1}$$

where $R_i' = R_i\,\boldsymbol{V}$ is a rotation matrix depending on the camera pose. This is the most basic form of the autocalibration constraint. It says that the calibrated images (*i.e.* 3D directions in the camera frame) of an orthogonal direction basis must remain orthogonal. It remains true in arbitrary projective 3D and image frames, as the projective deformations of U *vs.* P_i and P_i *vs.* C_i cancel each other out. As always, the direction basis U is defined only up to an arbitrary 3×3 orthogonal mixing $U \mapsto U\,R$.

The simplest approaches to autocalibration for non-planar scenes are based on this basic consistency equation (1), an intermediate projective reconstruction P_i, and some sort of knowledge about the K_i, *e.g.* classically that they are all the same: $K_i = K$ for some unknown K. Nonlinear optimization or algebraic elimination are used to estimate the Euclidean structure Λ or U, and the free parameters of the $C_i = K_i^{-1}$. Multiplying (1) either on the left or on the right by its transpose to eliminate the unknown rotation, and optionally moving the C's to the right hand side, gives the following constraints linking Λ to Q_i:

$$P_i\,\Lambda\,P_i^\top \;\sim\; Q_i \;=\; K_i\,K_i^\top \tag{2}$$

In each case there are 5 independent constraints per image on the 8 non-Euclidean d.o.f. of the 3D projective structure and the 5 (or fewer) d.o.f. of the internal calibration. For example, three images in general position suffice for classical constant-C autocalibration. In each case, the unknown scale factors can be eliminated by treating the symmetric 3×3 left and right hand side matrices as $3 \cdot 4/2 = 6$ component vectors, and either (*i*) projecting (say) the left hand sides orthogonally to the right hand ones (hence deleting the proportional components and focusing on the constraint-violating non-proportional ones), or (*ii*) cross-multiplying in the usual way:

$$\begin{cases} \boldsymbol{u}_i^\top\,\boldsymbol{v}_i = \boldsymbol{u}_i^\top\,\boldsymbol{w}_i = \boldsymbol{v}_i^\top\,\boldsymbol{w}_i = 0 \\ \|\boldsymbol{u}_i\|^2 = \|\boldsymbol{v}_i\|^2 = \|\boldsymbol{w}_i\|^2 \end{cases}$$
$$\text{where} \quad (\boldsymbol{u}_i, \boldsymbol{v}_i, \boldsymbol{w}_i) \;\equiv\; C_i\,P_i\,U$$

for the constraints on direction basis and

$$(P_i\,\Lambda\,P_i^\top)_{ij}\;(Q_i)_{mn} - (Q_i)_{ij}\;(P_i\,\Lambda\,P_i^\top)_{mn} = 0$$
$$\text{where} \quad i \le j, m \le n = 1\ldots3$$

for the constraints on the absolute conic.

Several recent autocalibration methods for 3D scenes (*e.g.* [29, 10]) are based implicitly on these constraints, parametrized by K or Q and by something equivalent to Λ or U. All of these methods seem to work well provided the intrinsic degeneracies of the autocalibration problem [27] are avoided.

In contrast, methods based on the Kruppa equations [15, 5, 32] can not be recommended for general use, because they add a serious additional singularity to the already-restrictive ones intrinsic to the problem [28]: if any 3D point projects to the same pixel and is viewed from the same distance in each image, a 'zoom' parameter can not be recovered from the Kruppa equations. Numerical experience suggests that Kruppa-based autocalibration remains ill-conditioned even quite far from this singularity. This is hardly surprising given that in any case the distinction between zooms and closes depends on fairly subtle 2^{nd}-order perspective effects, so that the recovery of focal lengths is never simple. (Conversely, the effects of an inaccurate zoom-close calibration on image measurements or local object-centred 3D ones are relatively minor).

3.2 Autocalibration of 2D affine camera

Autocalibration for affine cameras was first proposed in [22] which relies on the concept of the internal calibration matrix $K_{i,2\times2} \sim \left(\begin{smallmatrix}\alpha_u & 0\\ 0 & \alpha_v\end{smallmatrix}\right)$. This is also a stratified method but things became considerablely simplified for affine cameras as the intermediate 3D structure has already been upgraded to affine. It is therefore easy to recast this method into direction basis framework by noticing that the plane at infinity has been identifed and the uncalibrated translation vectors could be fixed. The direction basis is now restricted to the plane at infinity and therefore reduces to a 3×3 matrix U whose 3 columns are 3D direction vectors in affine coordinates. Projecting the direction basis U prallelly to the image plane involves only 2×3 part $M_i \sim P_{i,2\times3}$ as follows:

$$K^{-1}_{i,2\times2} M_i U_{3\times3} \sim R_{2\times3}$$

This is the basic affine camera autocalibration constraintes. Multiplying either side on the left or on the right by its transpose to eliminate the unknown rotation and moving K_i^{-1}'s to the right side gives

$$P_{i,2\times3} \Lambda_\infty P^\top_{i,2\times3} \sim Q_{2\times2} = K_i K_i^T,$$

where Λ_∞ is the dual absolute conic on the plane at infinity and $Q_{2\times2}$ is the line equation of the image of the absolute point-pair. Notice that the Λ_∞ is a full rank conic instead of singular Λ or $\Lambda_{3\times3}$.

The unknown scale factors can also be eliminated by treating the symmetric matrices as vectors. By cross-multiplying in the usual way to eliminate the unknown scale factors, we obtain the form of the constraints used by our affine calibration algorithm for the constant unknown aspect ratio:

$$(M_i \Lambda_\infty M_i^\top)_{ij} (Q_{2\times2})_{mn} - (Q_{2\times2})_{ij} (M_i \Lambda_\infty M_i^\top)_{mn} = 0$$
$$\text{where} \quad i \le j, m \le n = 1 \ldots 2.$$

Projecting the left hand sides orthogonally to the right hand ones gives

$$\begin{cases} \boldsymbol{u}_i^\top \boldsymbol{v}_i = 0 \\ \|\boldsymbol{u}_i\|^2 = \|\boldsymbol{v}_i\|^2 \end{cases}$$
$$\text{where} \qquad (\boldsymbol{u}_i, \boldsymbol{v}_i)^\top \;\equiv\; K^{-1}_{i,2\times2}\, P_{i,2\times3}\, U_{3\times3}$$

for the constraints on the direction basis.

Five independent parameters are required to specify the Euclidean structure from the affine structure: the 9 components of $U_{3\times3}$ modulo scale and the 3 d.o.f. of a rotation matrix; or just the 5 parameters of the absolute conic on the plane at infinity. Since each image gives 2 independent constraints, 5 images are necessary for the five intrinsic calibration parameters. Of course, further constraints on the intrinsic calibration parameters may reduce the number of images necessary. For constant intrinsic calibration parameters, i.e. the aspect ratio of the moving camera, three images are enough.

3.3 Autocalibration of 1D projective camera

For 1D projective camera [24], the scene to recover is 2D Euclidean space which is the usual Euclidean plane. The Euclidean structure of a plane is given by any one of

- a 3×3 rank 2 absolute line quadric $\Lambda_{3\times3}$;
- a 3 component line at infinity $\boldsymbol{l}_\infty$ and its associated 2×2 absolute conic matrix;
- a 3×2 direction basis matrix $U_{3\times2} = (\boldsymbol{u}\, \boldsymbol{v})$;
- a pair of absolute points $\boldsymbol{x}_\pm = \frac{1}{\sqrt{2}}(\boldsymbol{u} \pm i\boldsymbol{v})$ which are also the two roots of the absolute conic on $\boldsymbol{l}_\infty$.

In each case the structure is the natural restriction of the corresponding 3D one, re-expressed in 2D. Given 1D images taken with projection matrices $P_{i,2\times3} = K_{i,2\times2} R_{i,2\times2} (I| - \boldsymbol{t}_i)$ (the internal calibration matrix of a 1D projective camera is $K_{2\times2} = \left(\begin{smallmatrix}\alpha_u & u_0\\ 0 & 1\end{smallmatrix}\right)$), and an orthogonal direction basis $U_{3\times2}$, we find that

$$K^{-1}_{i,2\times2} P_{i,2\times3} U_{3\times2} \sim R_{2\times2}$$

Equivalently eliminating the rotation matrix gives the following constraints on the absolute conic:

$$P_{i,2\times3} \Lambda_{3\times3} P^T_{i,3\times2} \sim Q_{2\times2} = K_{i,2\times2} K^\top_{i,2\times2}$$

where $\Lambda_{3\times3}$ is the rank 2 absolute line quadric.

Four independent parameters are required to specify the Euclidean structure of a projective plane: the 6 components of U modulo scale and 1 d.o.f. of a rotation on the plane; or the 4 d.o.f. of the rank 2 absolute line quadric. In each case, the unknown scale factors can be eliminated by treating the symmetric 2×2 left and right hand side matrices as 3-vectors, and either projecting (say) the left hand sides orthogonally to the right hand ones (hence deleting the proportional components and focusing on the constraint-violating nonproportional ones),

$$\begin{cases} \boldsymbol{u}_i^\top \boldsymbol{v}_i = 0 \\ \|\boldsymbol{u}_i\|^2 = \|\boldsymbol{v}_i\|^2 \end{cases}$$
$$\text{where} \quad (\boldsymbol{u}_i, \boldsymbol{v}_i) \;\equiv\; K^{-1}_{i,2\times2}\, P_{i,2\times3}\, U_{3\times2}$$

for the constraints on direction basis, or cross-multiplying in the usual way:

$$(P_i \Lambda_{3\times3} P_i)_{ij} (Q_{2\times2})_{mn} - (Q_{2\times2})_{ij} (P_i \Lambda_{3\times3} P_i^\top)_{mn} = 0$$
$$\text{where} \quad i \le j, m \le n = 1 \ldots 2$$

for the constraints on the absolute conic.

The difference between 1D camera and 2D affine camera resides only on that of $\Lambda_{3\times3}$ and Λ_∞. Each image provides two independent constraints for two internal parameters of each camera and 4 for the Euclidean structre. For constant intrinsic parameters, three images are sufficiant.

The above developement assumes that intermediate 2D projective reconstruction is available. This is almost granted for 2D projective camera case, while it is peculiar for 1D camera. 2D projective reconstruction from 1D images has always the inherent 2-way ambiguity [23] while the trifocal tensor is unique. In [6], an elegant autocalibration method has been developed from the trifocal tensor of the 3 views for the constant unknown internal parameters.

3.4 Autocalibration of rotating camera

Now consider autocalibration from a rotating camera. This is a particular case in which the translation is zero. Obviously with zero-translation, no any 3D structure could be recovered, but inter-image homographies (here particular infinity homographies) still provide strong constraints on the internal calibration parameters. The key is that with zero-translation, the direction basis is restricted to the plane at infinity and reduces to a 3×3 matrix U. This direction basis is projected by the corresponding plane-at-infinity-to-image homography H. As the three columns of U represent 3D direction vectors in affine coordinates, their images still satisfy the autocalibration constraints (1):

$$K_i^{-1} H_i U_{3\times3} \sim R_{3\times3}$$

Multiplying on the left by its transpose and moving the K_i's to the right side gives the following constraints

$$H_i \Lambda_\infty H_i^\top \sim Q_i = K_i K_i^\top$$

on the absolute conic which says that indeed the image of the absolute conic on the image plane and the absolute conic on the plane at infinity are in homographical correspondence, expressed algebraically in dual form.

In each case, the unknown scale factors can be eliminated by treating the symmetric 3×3 matrices as 6-vectors, and either (*i*) projecting (say) the left hand sides orthogonally to the right hand ones

$$\begin{cases} \mathbf{u}_i^\top \mathbf{v}_i = \mathbf{u}_i^\top \mathbf{w}_i = \mathbf{v}_i^\top \mathbf{w}_i = 0 \\ \|\mathbf{u}_i\|^2 = \|\mathbf{v}_i\|^2 = \|\mathbf{w}_i\|^2 \end{cases}$$
$$\text{where} \quad (\mathbf{u}_i, \mathbf{v}_i, \mathbf{w}_i) \equiv K_i^{-1} H_i U$$

for the constraints on the direction basis, or (*ii*) cross-multiplying in the usual way:

$$(H_i \Lambda_\infty H_i^\top)_{ij} \ (Q)_{mn} - (Q)_{ij} \ (H_i \Lambda H_i^\top)_{mn} = 0$$
$$\text{where} \quad i \leq j, m \leq n = 1 \dots 3$$

for the constraints on the absolute conic on which many recent methods have been developed.

3.5 Autocalibration from Planar Scenes

Now consider autocalibration from *planar* scenes. Everything for non-planar scenes remains valid, except that no intermediate 3D projective reconstruction is available from which to bootstrap the process. However we will see that by using the inter-image homographies, autocalibration is still possible.

The Euclidean structure of the scene plane has been described in the previous section for 1D camera case. In each case the structure is the natural restriction of the corresponding 3D one, re-expressed in 2D. In each case it projects isomorphically into each image, either by the usual 3×4 3D projection matrix (using 3D coordinates), or by the corresponding 3×3 world-plane to image homography H (using scene plane coordinates). Hence, each image inherits a pair of circular points $H_i \, \mathbf{x}_\pm$ and the corresponding direction basis $H_i \, (\mathbf{u} \, \mathbf{v})$, line at infinity $\mathbf{l}_\infty \, H_i^{-1}$ and 3×3 rank 2 absolute line quadric $H_i \, \Lambda_{3\times3} \, H_i^\top$. As the columns of the planar $U_{3\times2}$ matrix represent 3D direction vectors (albeit expressed in the planar coordinate system), their images still satisfy the autocalibration constraints (1):

$$K_i^{-1} \, H_i \, U_{3\times2} \sim R_{3\times2} \tag{3}$$

where $R_{3\times2}$ contains the first two columns of a 3×3 rotation matrix.

Multiplying by its transpose on both sides to eliminate the unknown rotation gives:

$$H_i \Lambda_{3\times3} H_i^\top \sim Q_i = K_i K_i^\top$$

which says that the image of the dual absolute conic is the image of the rank 2 absolute line quadric.

The unknown scale factors can be eliminated by treating the symmetric 3×3 matrices as 6-vectors, and cross-multipling in the usual ways gives

$$(H_i \, \Lambda_{3\times3} \, H_i^\top)^{AB} \ (Q)^{CD} = (Q)^{AB} \ (H_i \, \Lambda_{3\times3} \, H_i^\top)^{CD}$$
$$\text{where} \quad A \leq B, C \leq D = 1 \dots 3$$

for the quadric based version.

Splitting the basic constraints into components gives the form of the constraints used by our planar autocalibration algorithm:

$$\begin{cases} \|\mathbf{u}_i\|^2 = \|\mathbf{v}_i\|^2 \\ \mathbf{u}_i \, \mathbf{v}_i = 0 \end{cases}$$
$$\text{where} \quad (\mathbf{u}_i, \mathbf{v}_i) \equiv C_i \, H_i \, (\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$$

Four independent parameters are required to specify the Euclidean structure of a projective plane: the 6 components of $U_{3\times2}$ modulo scale and the single d.o.f. of a 2×2 rotation; or the 6 components of a 3×3 absolute line quadric Q modulo scale and the rank 2 constraint $\det Q = 0$; or the 2 d.o.f. of the plane's line at infinity, plus the 2 d.o.f. of two circular points on it. Since the above constraint equations give two independent constraints for each image, $\lceil \frac{n+4}{2} \rceil$ images are required to estimate the Euclidean structure of the plane and n intrinsic calibration parameters. Two images suffice to recover the structure if the calibration is known, three are required if the focal length is also estimated, four for the perspective f, u_0, v_0 model, and five if all 5 intrinsic parameters are unknown.

4 Conclusion

In this paper, we have presented a unified approach to the autocalibration methods of a moving camera. The key is the introduction of direction bases which facilitates the comprehension of various autocalibration methods often presented in terms of the absolute conic from the pure projective geometry. The camera models we studied include 2D projective, 2D affine and 1D projective projections. The scenes may be either planar or non-planar. The moving camera can also be zero-translation.

From a practical point of view, autocalibration has some annoying fundamental limitations, so in practice some sort of compromise between auto- and conventional calibration is probably called for (this is an active research area!). In particular, many common types of camera motion are insufficient for classical autocalibration — the skew and aspect ratio parameters tend to be the hardest to estimate. Also, although the stability of autocalibration algorithms has improved immeasurably in the last few years, nonlinear equations must still be solved and multiple or false solutions still pose reliability problems. For both of these reasons, it definitely pays to keep the number of unknown parameters to be estimated to a minimum.

References

[1] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. ECCV'96.

[2] S. Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. ICCV'98.

[3] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? ECCV'92.

[4] O. Faugeras. Stratification of 3-d vision: Projective, affine, and metric representations. *J. Optical Society of America*, A 12(3):465–84, March 1995.

[5] O. Faugeras, Q.-T. Luong, and S. J. Maybank. Camera self calibration: Theory and experiments. ECCV'92.

[6] O. Faugeras, L. Quan, and P. Sturm. Self-calibration of a 1d projective camera and its application to the self-calibration of a 2d projective camera. ECCV'98.

[7] R. Hartley. Euclidean reconstruction from multiple views. In 2^{nd} *Europe-U.S. Workshop on Invariance*, pages 237–56, Ponta Delgada, Azores, October 1993.

[8] R. Hartley. Self-calibration from multiple views with a rotating camera. ECCV'94.

[9] R.I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. CVPR'92.

[10] A. Heyden and K. Åström. Euclidean reconstruction from constant intrinsic parameters. ICPR'96.

[11] A. Heyden and K. Åström. Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. CVPR'97.

[12] R. Horaud and G. Csurka. Self-calibration and euclidean reconstruction using motions of a stereo rig. ICCV'98.

[13] F. Kahl and B. Triggs. Critical motions in euclidean structure from motion. CVPR'99.

[14] Q.-T. Luong and T. Viéville. Canonic representations for the geometries of multiple projective views. TR-UCB/CSD-93-772, Dept. EECS, Berkeley, California, 1993.

[15] S. J. Maybank and O. Faugeras. A theory of self calibration of a moving camera. IJCV, 8(2):123–151, 1992.

[16] R. Mohr, F. Veillon, and L. Quan. Relative 3D reconstruction using multiple uncalibrated images. CVPR'93.

[17] M. Pollefeys and L Van Gool. A stratified approach to metric self-calibration. CVPR'97.

[18] M. Pollefeys, L Van Gool, and M. Proesmans. Euclidean 3d reconstruction from image sequences with variable focal length. ECCV'96.

[19] L. Quan. Affine stereo calibration for relative affine shape reconstruction. BMVC'93, pp. 659–668.

[20] L. Quan. Invariant of a pair of non-coplanar conics in space: Definition, geometric interpretation and computation. ICCV'95.

[21] L. Quan. Invariants of six points and projective reconstruction from three uncalibrated images. IEEE-TPAMI, 17(1):34–46, January 1995.

[22] L. Quan. Self-calibration of an affine camera from multiple views. IJCV, 19(1):93–105, May 1996.

[23] L. Quan. Inherent two-way ambiguity in 2d projective reconstruction from three uncalibrated 1d images. ICCV'99.

[24] L. Quan and T. Kanade. Affine structure from line correspondences with uncalibrated affine cameras. IEEE-TPAMI, 19(8):834–845, August 1997.

[25] L. Quan and R. Mohr. Towards structure from motion for linear features through reference points. IEEE Workshop on Visual Motion, Princeton, pp. 249–254, October 1991.

[26] J.G. Semple and G.T. Kneebone. *Algebraic Projective Geometry*. Oxford Science Publication, 1952.

[27] P. Sturm. Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction. CVPR'97.

[28] P. Sturm. *Vision 3D non calibrée*. Ph.D. Thesis, INP de Grenoble, 1997.

[29] B. Triggs. Autocalibration and the absolute quadric. CVPR'97.

[30] B. Triggs. Autocalibration from planar scenes. ECCV'98.

[31] T. Viéville and O. Faugeras. Motion analysis with a camera with unknown and possibly varying intrinsic parameters. ICCV'95.

[32] C. Zeller and O. Faugeras. Camera self-calibration from video sequences: the Kruppa equations revisited. TR2793, INRIA, 1996.

[33] Z. Zhang, Q.-T. Luong, and O. Faugeras. Motion of an uncalibrated stereo rig: Self-calibration and metric reconstruction. TR2079, INRIA, 1993.