Combining Scene and Auto-calibration Constraints

David Liebowitz and Andrew Zisserman*
Department of Engineering Science
University of Oxford
Oxford OX1 3PJ, UK
{dl,az}@robots.ox.ac.uk

Abstract

We present a simple approach to combining scene and auto-calibration constraints for the calibration of cameras from single views and stereo pairs. Calibration constraints are provided by imaged scene structure, such as vanishing points of orthogonal directions, or rectified planes. In addition, constraints are available from the nature of the cameras and the motion between views. We formulate these constraints in terms of the geometry of the imaged absolute conic and its relationship to pole-polar pairs and the imaged circular points of planes. Three significant advantages result: first, constraints from scene features, camera characteristics and auto-calibration constraints provide linear *equations in the elements of the image of the absolute conic. This means that constraints may easily be combined, and their solution is straightforward. Second, the degeneracies that occur when constraints are not independent may be easily identified. Lastly, the constraints from scene planes and image planes may be treated uniformly. Examples of various cases of constraint combination and degeneracy as well as computational techniques are presented.*

1 Introduction

Auto- (or self-) calibration is the computation of camera internal calibration and/or metric properties of the scene from a set of uncalibrated images. The original auto-calibration method based on Kruppa's equations [16] was restricted to cameras with fixed internal parameters, and early work in this area maintained this restriction [1, 9, 10, 18]. However, with a sufficiently large number of views auto-calibration has now been demonstrated for cases where several of the internal parameters, for example the focal length and principal point, are allowed to vary [17, 23].

More classically, cameras may be calibrated from scene points with known coordinates [21] or from sets of vanishing points corresponding to directions which are orthogonal in the scene [12]. In particular, Caprile and Torre [2] described a method to calibrate a camera with known aspect ratio and skew from a single view of three vanishing points corresponding to orthogonal directions.

We present here a method to combine image, scene and auto-calibration constraints for calibration of single views and multiple views. The method is formulated in terms of the geometry of the imaged absolute conic and the circular points of planes. There are three significant advantages of this formulation: first, it will be seen that both scene and auto-calibration constraints are *linear* equations. This means that constraints may easily be combined, and their solution is straightforward. Second, degeneracies, where there are insufficient independent constraints to determine a unique calibration, may be identified by simple geometric reasoning and the family of solutions parametrized. The method of Caprile and Torre [2], for example, is degenerate if one or more of the vanishing points is at infinity in the image. Third, the constraints from scene planes and multiple image planes are treated uniformly.

This is a very pragmatic approach for the case of a small number of views where auto-calibration based on motion alone is likely to be ill-conditioned. It is well suited to the case of building architectural models from a small number of views [4] using only calibration information from the images themselves [3, 13]. The ubiquitous presence on buildings and other man-made structures of planes and lines in orthogonal directions means that such scenes often provide sufficient constraints to determine a unique calibration.

2 Notation and background

Points on the image plane are represented by homogeneous 3 vectors $\mathbf{x} = (x_1,\ x_2,\ x_3)^\top$ with $(x, y) = (x_1/x_3, x_2/x_3)$ the Euclidean position on the plane. A line $\mathbf{l}$ is also represented by a homogeneous column 3-vector, such that if a point $\mathbf{x}$ lies on $\mathbf{l}$ then $\mathbf{l}^\top\mathbf{x} = 0$. There are two points on the plane that are of special importance in what follows: the circular points are a complex conjugate point pair with the property that their position is invariant

*This work was supported by the EU Esprit Project IMPROOFS. We are grateful for discussions with Stefan Carlsson.

to similarity transformations of the plane. In a metric coordinate frame on the plane the circular points' position is $\mathbf{x} = (1,\ \pm i,\ 0)^\top$, and it may be verified that the position is unchanged by a plane similarity transformation.

Conics are represented by a symmetric 3×3 homogeneous matrix $\mathtt{C}$ such that a point on the conic satisfies [20]

$$\mathbf{x}^\top \mathtt{C} \mathbf{x} = 0 \tag{1}$$

A point $\mathbf{x}$ and conic $\mathtt{C}$ define a line $\mathbf{l} = \mathtt{C}\mathbf{x}$. The line $\mathbf{l}$ is called the *polar* of $\mathbf{x}$ with respect to $\mathtt{C}$, and the point $\mathbf{x}$ is the *pole* of $\mathbf{l}$ with respect to $\mathtt{C}$. The polar line $\mathbf{l} = \mathtt{C}\mathbf{x}$ of the point $\mathbf{x}$ with respect to a conic $\mathtt{C}$ intersects the conic in two points at which tangent lines to $\mathtt{C}$ intersect at $\mathbf{x}$. If a point $\mathbf{v}_1$ lies on the polar of another point $\mathbf{v}_2$ then the two points are said to be conjugate with respect to the conic and satisfy

$$\mathbf{v}_1{}^\top \mathtt{C} \mathbf{v}_2 = 0 \tag{2}$$

The construction is shown in figure 1.

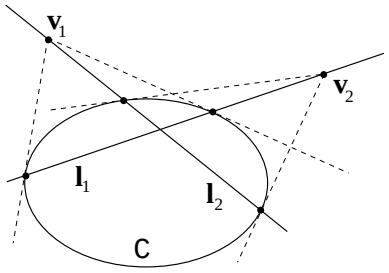


Figure 1: Points $\mathbf{v}_1$ and $\mathbf{v}_2$, with polars $\mathbf{l}_1$ and $\mathbf{l}_2$ respectively. The points $\mathbf{v}_1$ and $\mathbf{v}_2$ are conjugate with respect to the conic $\mathtt{C}$.

A 3D point is represented by a homogeneous 4 vector $\mathbf{X} = (X_1,\ X_2,\ X_3,\ X_4)^\top$, and is imaged at a point $\mathbf{x}$ as $\mathbf{x} = \mathtt{P}\mathbf{X}$. The 3×4 matrix $\mathtt{P} = \mathtt{K}[\mathtt{R}|\mathbf{t}]$ represents the projection from 3D to an image. The matrix $\mathtt{K}$ contains the five internal parameters of the camera

$$\mathtt{K} = \begin{bmatrix} f & k & u_0 \\ 0 & rf & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

where f is the focal length, r the aspect ratio, k the skew and (u_0, v_0) the principal point of the camera. $\mathtt{R}$ and $\mathbf{t}$ are the 3D rotation and translation of the camera relative to a world co-ordinate system.

The absolute conic lies on the plane at infinity, $\boldsymbol{\pi}_\infty$, in 3D and has the property that it is invariant to similarity transformations of space [20]. A plane in space intersects $\boldsymbol{\pi}_\infty$ in a line which intersects the absolute conic in two points. These two points are the circular points of that plane. The conic relevant to calibration is the image of the absolute conic $\boldsymbol{\omega}$. It is simply related to the camera calibration by $\boldsymbol{\omega} = \mathtt{K}^{-\top}\mathtt{K}^{-1}$[6]. The calibration matrix $\mathtt{K}$ may be computed from $\boldsymbol{\omega}$ by Cholesky decomposition [8], so determining $\boldsymbol{\omega}$ in an image is equivalent to knowing the camera internal parameters. The vanishing line of a space plane intersects $\boldsymbol{\omega}$ in two points. These points are the imaged circular points of the plane – they are the image of the intersection points of the plane with the absolute conic.

A further property of $\boldsymbol{\omega}$ is that it determines the angle between rays through the camera centre. The angle between the rays corresponding to the image points $\mathbf{v}_1$ and $\mathbf{v}_2$ is

$$\cos(\theta) = \frac{\mathbf{v}_1{}^\top \boldsymbol{\omega} \mathbf{v}_2}{(\mathbf{v}_1{}^\top \boldsymbol{\omega} \mathbf{v}_1)^{1/2} (\mathbf{v}_2{}^\top \boldsymbol{\omega} \mathbf{v}_2)^{1/2}} \tag{3}$$

When the rays are orthogonal (3) reduces to (2) and the points are conjugate with respect to $\boldsymbol{\omega}$, i.e. orthogonal rays satisfy

$$\mathbf{v}_1{}^\top \boldsymbol{\omega} \mathbf{v}_2 = 0 \tag{4}$$

Between two views of a scene there is a plane homography induced by the plane at infinity known as the infinite homography [15] and represented by a 3×3 homogeneous matrix $\mathtt{H}_\infty$. An image of a point on $\boldsymbol{\pi}_\infty$ is mapped as $\mathbf{x}' = \mathtt{H}_\infty \mathbf{x}$, where $\mathbf{x}$ and $\mathbf{x}'$ are the image points in the first and second views respectively. The infinite homography maps projections of entities at infinity in 3D, such as vanishing points and the image of the absolute conic, between views. In particular $\boldsymbol{\omega}$ is mapped as [15, 27]

$$\boldsymbol{\omega}' = \mathtt{H}_\infty^{-\top} \boldsymbol{\omega} \mathtt{H}_\infty^{-1} \tag{5}$$

where $\boldsymbol{\omega}'$ is the image of the absolute conic in the second view. If $\mathtt{H}_\infty$ is known then $\boldsymbol{\omega}$ determines $\boldsymbol{\omega}'$. Of particular importance here, where $\boldsymbol{\omega}$ is not known *a priori*, is that constraints on $\boldsymbol{\omega}$ may be transferred to constraints on $\boldsymbol{\omega}'$ and vice-versa.

3 Scene and motion constraints

In this section the *individual* constraints on $\boldsymbol{\omega}$ arising from the scene, motion and knowledge of $\mathtt{K}$ are described. The combination of these constraints and the solution for $\boldsymbol{\omega}$ is postponed until the following section. We will concentrate on the very common situation of a *natural* camera. This is a camera where there is no skew ($k = 0$) and the pixels are square (or equivalently the aspect ratio r is known). It will be shown that the constraint of a natural camera is equivalent to knowledge of the image plane circular points.

Orthogonal vanishing points: The vanishing points of orthogonal directions constrain $\boldsymbol{\omega}$ directly from the conjugacy property (4). Given orthogonal vanishing points $\mathbf{v}_1$ and $\mathbf{v}_2$, $\mathbf{v}_1{}^\top \boldsymbol{\omega} \mathbf{v}_2 = 0$ is an equation *linear* in the elements of $\boldsymbol{\omega}$.

In the common case that vanishing points in three orthogonal directions are available, the three vanishing points $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are all conjugate with respect to $\boldsymbol{\omega}$ and define a *self-polar triangle* [20] (see figure 3). There are three independent linear constraints on $\boldsymbol{\omega}$:

$$\mathbf{v}_1{}^\top \boldsymbol{\omega} \mathbf{v}_2 = \mathbf{v}_1{}^\top \boldsymbol{\omega} \mathbf{v}_3 = \mathbf{v}_2{}^\top \boldsymbol{\omega} \mathbf{v}_3 = 0 \qquad (6)$$

Rectified planes: A scene plane can be metric rectified if the homography which maps the world plane to the image plane is determined up to a similarity transformation. Suppose the homography is $\mathtt{H}$, then the imaged circular points are $\mathtt{H}(1,\ \pm i,\ 0)^\top$. These image points, which will be denoted as $\mathbf{I}$ and $\mathbf{J}$, are therefore known. These points lie on $\boldsymbol{\omega}$ and from (1) provide two linear constraints on $\boldsymbol{\omega}$:

$$\mathbf{I}^\top \boldsymbol{\omega} \mathbf{I} = 0 \quad \text{and} \quad \mathbf{J}^\top \boldsymbol{\omega} \mathbf{J} = 0 \qquad (7)$$

In practice $\mathbf{I}$ and $\mathbf{J}$ contain the same information, and the constraint pair follows from writing the real and imaginary components of either point constraint.

Planes may be metric rectified from a combinations of geometric constraints, such as known angles between lines, equal unknown angles and known ratios of lengths [14]. In particular a square uniquely determines the planar rectification homography, as does a rectangle with known aspect ratio. A circle on a plane directly determines the imaged circular points if the vanishing line of the plane is known.

Natural camera: Cameras commonly have a known aspect ratio, usually one, and negligible skew. The known aspect ratio and skew may be expressed as geometric constraints specifying the image plane circular points. For a natural camera, where the aspect ratio r is one and the skew k is zero, the image plane is a metric plane, and the circular points have the known position $(1,\ \pm i,\ 0)^\top$ on the line at infinity $\mathbf{l}_\infty$ in the image. The imaged circular points also have the position $(1,\ \pm i,\ 0)^\top$. That is, the constraint of a natural camera is equivalent to the constraint of two points on $\boldsymbol{\omega}$.

Suppose there are several views acquired by natural cameras. The image plane of a natural camera is equivalent to a metric scene plane in that the circular points are known. In exactly the same manner as a rectified plane (above) providing two linear constraints on $\boldsymbol{\omega}$ by its imaged circular points, each natural camera provides two linear constraints on $\boldsymbol{\omega}$. If the infinite homography between one of the views and the first is $\mathbf{x}' = \mathtt{H}_\infty \mathbf{x}$, then the circular points of the second image plane are imaged at $\mathbf{x} = (\mathtt{H}_\infty)^{-1}(1, \pm i, 0)^\top$ in the first view, and these points lie on $\boldsymbol{\omega}$ (and on the vanishing line of that image plane). Thus each additional view from a natural camera provides two points on $\boldsymbol{\omega}$ in the first view.

If the camera has zero skew but the aspect ratio is unknown, this is still expressible as a linear constraint on $\boldsymbol{\omega}$[27]: the axes are orthogonal and thus conjugate with respect to $\boldsymbol{\omega}$ from (4), i.e. $(1,0,0)\boldsymbol{\omega}(0,1,0)^\top = 0$. However, the zero skew constraint is not linear if applied to the dual image of the absolute conic [15].

Fixed camera: If the camera is fixed over two views, the infinite homography $\mathtt{H}_\infty$ between the views provides a set of auto-calibration constraints. In the case of a fixed camera (5) simplifies to $\boldsymbol{\omega} = \mathtt{H}_\infty{}^\top \boldsymbol{\omega} \mathtt{H}_\infty$. These are six linear equations in the elements of $\boldsymbol{\omega}$, of which only four are independent [15]. The four constraints may be written explicitly by considering the geometry defined by the infinite homography. In the case of a fixed camera $\mathtt{H}_\infty$ is conjugate to a rotation, $\mathtt{H}_\infty = \mathtt{K}\mathtt{R}\mathtt{K}^{-1}$, where $\mathtt{R}$ is the rotation between the views. The eigenvectors of $\mathtt{H}_\infty$ identify three fixed image points under the mapping $\mathbf{x} \rightarrow \mathtt{H}_\infty \mathbf{x}$[27]. One of the fixed points, $\mathbf{a}$, is the vanishing point of the rotation axis direction (and corresponds to the real eigenvector). The point $\mathbf{a}$ and the vanishing line $\mathbf{l}_\mathbf{a}$ of planes orthogonal to the axis of rotation are pole-polar with respect to $\boldsymbol{\omega}$. The two remaining eigenvectors of $\mathtt{H}_\infty$ are the imaged circular points $\mathbf{I}_\mathbf{a}$ and $\mathbf{J}_\mathbf{a}$ of these planes, and are the intersection points of $\boldsymbol{\omega}$ with $\mathbf{l}_\mathbf{a}$. The construction appears in figure 2. The circular points provide two constraints on $\boldsymbol{\omega}$, and the polarity of $\mathbf{a}$ and $\mathbf{l}_\mathbf{a}$ (a vector equation) two further constraints:

$$\mathbf{I}_\mathbf{a}{}^\top \boldsymbol{\omega} \mathbf{I}_\mathbf{a} = 0 \quad \mathbf{J}_\mathbf{a}{}^\top \boldsymbol{\omega} \mathbf{J}_\mathbf{a} = 0 \quad \mathbf{l}_\mathbf{a} = \mathbf{I}_\mathbf{a} \times \mathbf{J}_\mathbf{a} = \boldsymbol{\omega} \mathbf{a} \qquad (8)$$

The polarity $\mathbf{l}_\mathbf{a} = \boldsymbol{\omega}\mathbf{a}$ may also be written as a pair of conjugacies, and thus (8) is equivalent to the four linear constraints:

$$\mathbf{I}_\mathbf{a}{}^\top \boldsymbol{\omega} \mathbf{I}_\mathbf{a} = \mathbf{J}_\mathbf{a}{}^\top \boldsymbol{\omega} \mathbf{J}_\mathbf{a} = \mathbf{I}_\mathbf{a}{}^\top \boldsymbol{\omega} \mathbf{a} = \mathbf{J}_\mathbf{a}{}^\top \boldsymbol{\omega} \mathbf{a} = 0 \qquad (9)$$

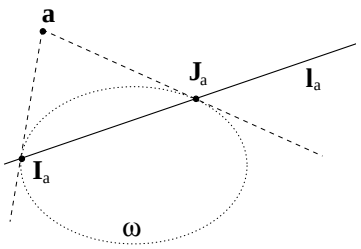


Figure 2: Constraints on $\boldsymbol{\omega}$ induced by $\mathtt{H}_\infty$ with a fixed camera.

Note, these linear constraints include the constraints arising from the Kruppa equations. However, the Kruppa equations (for which only the fundamental matrix $\mathtt{F}$, but not $\mathtt{H}_\infty$, is required) can not be written as a linear constraint on $\boldsymbol{\omega}$. Instead the constraint is linear on the dual image of the absolute conic $\boldsymbol{\omega}^* = \mathtt{K}\mathtt{K}^\top$, and may be written as $[\mathbf{e}']_\times \boldsymbol{\omega}^* [\mathbf{e}']_\times = \mathtt{F}\boldsymbol{\omega}^*\mathtt{F}^\top$ [25].

4 Constraint combination

In this section we describe the combination of constraints and their degeneracy. Identifying degeneracies is extremely important in practice, since if the number of independent constraints is insufficient then only a parametrized family of solutions for $\boldsymbol{\omega}$ (and hence $\mathtt{K}$) is obtained. Also, if conditions are near degenerate then the solution is ill-conditioned and the particular member of the family is determined by "noise".

Since all the above constraints are described algebraically as linear equations, it is a simple matter to combine them as rows of a constraint matrix. For example, for the orthogonality constraint $\mathbf{v}_1{}^\top \boldsymbol{\omega} \mathbf{v}_2 = 0$, writing

$$\boldsymbol{\omega} = \begin{bmatrix} \omega_1 & \omega_2 & \omega_4 \\ \omega_2 & \omega_3 & \omega_5 \\ \omega_4 & \omega_5 & \omega_6 \end{bmatrix}$$

and $\mathbf{v}_1 = (v_1,\ v_2,\ v_3)^\top$ and $\mathbf{v}_2 = (u_1,\ u_2,\ u_3)^\top$, the full expression is

$$v_1u_1w_1 + (v_1u_2 + v_2u_1)w_2 + v_2u_2w_3 + (v_1u_3 + v_3u_1)w_4 + (v_2u_3 + v_3u_2)w_5 + v_3u_3w_6 = \mathbf{g}^\top \mathbf{w} = 0$$

where $\mathbf{w}$ is a vector of the elements of $\boldsymbol{\omega}$ and $\mathbf{g}$ the vector of coefficients. All such constraints may be collected together so that for n constraints $\mathtt{A}\mathbf{w} = \mathbf{0}$, where $\mathtt{A}$ is a $n \times 6$ matrix. If there are five independent constraints the solution is obtained as the one-dimensional null-space of $\mathtt{A}$. If fewer constraints are available or the constraints are degenerate, the rank of $\mathtt{A}$ reflects the number of independent constraints (although noise may artificially elevate the rank). In practice, the rank is measured using the singular value decomposition [8]. If n is greater than 5, then as usual for over-constrained linear systems, a solution for $\mathbf{w}$ is obtained by least-squares as the singular vector of $\mathtt{A}$ associated with the smallest singular value (this vector minimizes $||\mathtt{A}\mathbf{w}||$ subject to $||\mathbf{w}|| = 1$). The rows of $\mathtt{A}$ should be normalized to unit magnitude to give all constraints the same weight. The solution for $\mathbf{w}$ will not exactly satisfy $\mathtt{A}\mathbf{w} = \mathbf{0}$. We return to the issue of solutions which exactly satisfy the constraints in section 7.

The geometric description of constraints has an important advantage. Degeneracies that occur when seemingly independent constraints are combined become clear when examined in terms of incidence and polarity relationships. The one and two view cases involving the self-polar triangle and natural or fixed cameras are discussed below, and summarized in table 1.

4.1 Single view calibration

If the camera is natural and three vanishing points corresponding to orthogonal directions in the scene are identified, then Caprile and Torre demonstrate that the principal point of the camera lies at the orthocentre of the triangle with vertices the vanishing points. In this section we examine the circumstances under which this construction will fail.

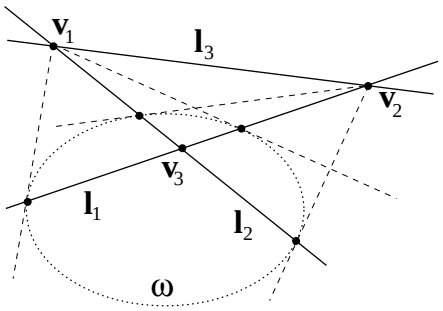


Figure 3: A self-polar triangle. A triangle is self-polar with respect to a conic if it has the property that each side of the triangle is the polar line of the opposite vertex.

We describe first the two parameter family of cameras defined by this triangle, without imposing the constraint that the camera is natural. A triad of orthogonal vanishing points is self-polar (see figure 3) with respect to $\boldsymbol{\omega}$, and since, from (6), it provides three constraints on $\boldsymbol{\omega}$, a two parameter family of conics consistent with these constraints results. The family of conics may be parametrized by considering the polar lines as degenerate (rank 1) conics of the form $\mathbf{l}_i\mathbf{l}_i{}^\top\ i = 1, 2, 3$. So, we write the two parameter family of conics $\mathtt{C}(\alpha, \beta, \gamma)$ as

$$\mathtt{C}(\alpha, \beta, \gamma) = \alpha\mathbf{l}_1\mathbf{l}_1{}^\top + \beta\mathbf{l}_2\mathbf{l}_2{}^\top + \gamma\mathbf{l}_3\mathbf{l}_3{}^\top \qquad (10)$$

and it may be verified that $\mathbf{v}_i{}^\top \mathtt{C}\mathbf{v}_i = 0, \forall\ \alpha, \beta, \gamma$. The parametrization vector $(\alpha,\ \beta,\ \gamma)^\top$ is homogeneous and represents two degrees of freedom.

In the general case, this two parameter ambiguity is resolved if the camera is natural since then the family must pass through the circular points $(1, \pm i, 0)^\top$ of the image plane, i.e.

$$\mathbf{I}^\top \mathtt{C}(\alpha, \beta, \gamma)\mathbf{I} = \mathbf{J}^\top \mathtt{C}(\alpha, \beta, \gamma)\mathbf{J} = 0 \qquad (11)$$

However, there are cases where these two constraints do not resolve the ambiguity. We show below that this occurs when the line containing the circular points is collinear with one of the vanishing points.

We wish to answer the following question: *under what circumstances does the constraint (11), that the conic passes through the circular points, uniquely determine a member of the conic family $\mathtt{C}(\alpha, \beta, \gamma)$ defined by the orthogonal triad of vanishing points?*

From (10) the constraint (11) may be written

$$\begin{bmatrix} (\mathbf{I}^\top \mathbf{l}_1)^2 & (\mathbf{I}^\top \mathbf{l}_2)^2 & (\mathbf{I}^\top \mathbf{l}_3)^2 \\ (\mathbf{J}^\top \mathbf{l}_1)^2 & (\mathbf{J}^\top \mathbf{l}_2)^2 & (\mathbf{J}^\top \mathbf{l}_3)^2 \end{bmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \mathbf{0}$$

In general this 2×3 matrix will have rank 2. However, if it drops rank, then the ambiguity is not resolved. The matrix drops rank if all the 2×2 sub-matrices are singular, and this is the case if the line, $\mathbf{I} \times \mathbf{J}$, through the circular points contains a triangle vertex.

We now examine three particular cases where the line through the circular points contains zero, one or two vertices. Since the vertices are vanishing points and the circular points are on the line at infinity in the image plane, these three cases correspond to three finite, two finite or one finite vanishing points respectively.

A natural camera and three finite (orthogonal) vanishing points: There are five independent constraints on $\boldsymbol{\omega}$ and it is defined uniquely. An example appears in figure 4. The principal point of the camera lies at the orthocentre of the triangle of vanishing points. The computed camera is

$$\mathtt{K} = \begin{bmatrix} 1081.03 & 0 & 396.65 \\ 0 & 1081.03 & 530.61 \\ 0 & 0 & 1.00 \end{bmatrix}$$

with an image size of 768×1024 pixels. The example also has an additional plane aligned at an angle to the parallel frontal planes. Having calibrated the camera, the orientation of this plane relative to others in the scene may be computed from

$$\cos(\theta) = \frac{\mathbf{l}_{\infty 1}{}^\top \boldsymbol{\omega}^{-1} \mathbf{l}_{\infty 2}}{(\mathbf{l}_{\infty 1}{}^\top \boldsymbol{\omega}^{-1} \mathbf{l}_{\infty 1})^{1/2} (\mathbf{l}_{\infty 2}{}^\top \boldsymbol{\omega}^{-1} \mathbf{l}_{\infty 2})^{1/2}}$$

where $\mathbf{l}_{\infty 1}$ and $\mathbf{l}_{\infty 2}$ are the vanishing lines of two planes and θ is the angle between them. The angled plane in the example is computed to be at 45.75^o to the frontal planes. The veridical value is 45.0^o.



(a) (b)

Figure 4: Calibration from a triad of finite orthogonal vanishing points and natural camera constraints. (a) Parallel line sets defining the three vanishing points. (b) The self-polar triangle and orthocentre.

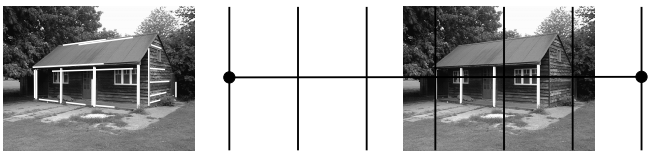
A natural camera and two finite, of three (orthogonal), vanishing points: One vertex of the self-polar triangle lies on $\mathbf{l}_\infty$ and is collinear with the circular points. The two circular points therefore only provide one independent constraint on $\boldsymbol{\omega}$ and a one parameter family of (natural) cameras remain. Suppose $\mathbf{v}_1$ is ideal, then the principal point lies on the line through the two finite vanishing points $\mathbf{v}_2$ and $\mathbf{v}_3$. In figure 5, for example, where one vanishing point is ideal, the 'orthocentre' of the triangle for the Caprile and Torre construction may be chosen anywhere on the line joining the finite vanishing points.

Figure 6 similarly has a near ideal vertical vanishing point. In this case, however, additional constraints are available since the frontal plane of the building may be rectified from its vanishing line together with the imaged circles (since circles intersect the line at infinity in the circular points). The imaged circular points of this plane lie on one of the sides of the self-polar triangle. This configuration provides five independent constraints.

The image is of size 1152×682 pixels, and the computed camera obtained by solving a linear system containing all the constraints is

$$\mathtt{K} = \begin{bmatrix} 1196.86 & 38.76 & 734.18 \\ 0 & 1253.42 & 426.26 \\ 0 & 0 & 1.00 \end{bmatrix}$$

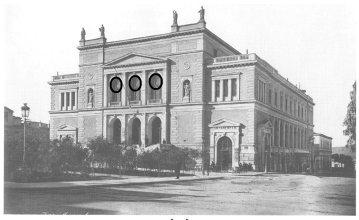
The computed aspect ratio is 1.05 and computed skew angle between axes is 88.15^o. This (non-natural camera) solution is a result of obtaining a least-squares solution for $\boldsymbol{\omega}$ from over-determined noisy measurements. Alternatively, the natural camera constraint may be enforced and a least-squares solution obtained for the remaining four homogeneous parameters of $\boldsymbol{\omega}$.



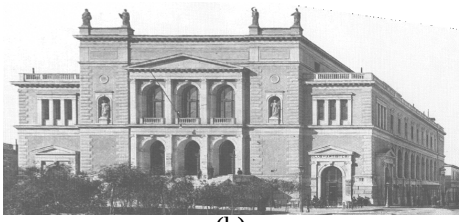
(a) (b)

Figure 5: Calibration with one of three orthogonal vanishing points ideal and natural camera constraints. (a) Parallel lines defining the vanishing points. (b) The vertical vanishing point is at infinity. The resultant camera is defined up to a 1 parameter ambiguity; the principal point lies on the line joining the finite vanishing points.

A natural camera and one finite, of three (orthogonal), vanishing points: Two vertices of the self-polar triangle lie on $\mathbf{l}_\infty$. Again $\boldsymbol{\omega}$ is determined up to a one parameter family of conics. The two ideal vanishing points are the vanishing points for directions parallel to the image plane. The third orthogonal direction must therefore be orthogonal to the image plane, parallel to the optical axis of the camera. The vanishing point of this axis is the principal point, which is therefore given by the finite vanishing point. The one



(a)



(b)

Figure 6: Calibration with one of three orthogonal vanishing near ideal, natural camera constraints and a rectified plane. (a) The vertical vanishing point is near ideal in this case, but the front plane may be rectified from the circles on the plane imaged as ellipses. (b) The rectified front plane.

parameter family of natural cameras is parametrized by the focal length.

Additional orthogonal vanishing points: In general, an additional pair of orthogonal vanishing points adds a conjugacy constraint. If, however, one of the additional vanishing points is a vertex of the self-polar triangle, the other lies on a polar of the triangle and the conjugacy is already satisfied, so no further independent constraints are added. Such a vanishing point pair may be found in figure 4, defined by the lines on the angled plane, and figure 5, by the bounding lines of the roof.

4.2 Calibration from two views

We consider the case of two views with non-coincident camera centres for which the infinite homography mapping $\mathtt{H}_\infty$ is known. This mapping can be computed from the corresponding of four or more vanishing points. However, we will focus here on the case that $\mathtt{H}_\infty$ is computed from the fundamental matrix $\mathtt{F}$ together with the correspondence of the vanishing points of the orthogonal triad. This computation is described in section 5.

In the absence of any other information each camera calibration has five unknown degrees of freedom, and this is reduced to a two parameter family by using the orthogonal triad independently in each view as described in section 4.1. The family in one view can be mapped onto the family in the other by (5), but no additional constraints are gained by this.

We will now examine how this ambiguity is reduced when additional information, such as that the cameras are natural, is introduced.

Two natural cameras: Consider the first camera. The two parameter family must contain the circular points $(1, \pm i, 0)^\top$, from the metric image plane of the first camera, and also the imaged circular points of the second camera which are transferred to the first camera via $\mathtt{H}_\infty$. The transferred circular points lie on $\mathbf{l}'_\infty$, the vanishing line of the second camera image plane as seen by the first camera. If neither this vanishing line nor the line at infinity of the image plane $\mathbf{l}_\infty$ intersect any of the vertices of the triangle of vanishing points, seven independent constraints on $\boldsymbol{\omega}$ are available. If either line passes through one of the vertices (as in the one view cases) six constraints are available. When both $\mathbf{l}_\infty$ and $\mathbf{l}'_\infty$ intersect a vertex, five constraints are available, i.e. the ambiguity is still resolved.



(a) (b)

Figure 7: A stereo pair imaged with the vertical vanishing point vertex of the self-polar triangle ideal, and rotation about the vertical axis.

The last case is common, for example when a building is imaged in a pair of views with the scene vertical axis parallel to the image plane and the rotation between views is about this axis. The vertical vanishing point is ideal, the line at infinity of the first image plane and the vanishing line of the second image plane both intersect the ideal vertex, as in figure 7.

The cameras are computed for this image pair using the fact that both cameras are natural. Image sizes are 1024×768 pixels and the computed cameras are

$$\mathtt{K}_1 = \begin{bmatrix} 1089.71 & 1.01 & 442.70 \\ 0 & 1089.66 & 380.32 \\ 0 & 0 & 1.00 \end{bmatrix}$$

$$\mathtt{K}_2 = \begin{bmatrix} 1015.97 & -0.39 & 493.78 \\ 0 & 1016.06 & 450.68 \\ 0 & 0 & 1.00 \end{bmatrix}$$

Fixed camera: A camera with the same parameters in both views introduces the four constraints of (9). Added to the three constraints from the self-polar triangle, the general case provides seven constraints on $\boldsymbol{\omega}$. In the case that $\mathbf{a}$ coincides with a vertex of the triangle of orthogonal vanishing points, the conjugacy and polarity are already satisfied and

the circular points lie on a polar, so only a single constraint is added, making four constraints on $\boldsymbol{\omega}$.

Vanishing points	Cameras	Independent constraints
3 finite	1 natural	5
1 ideal	1 natural	4
2 ideal	1 natural	4
3 finite $\mathbf{l}'_\infty{}^\top \mathbf{v}_i \neq 0$	2 natural	7
3 finite $\mathbf{l}'_\infty{}^\top \mathbf{v}_i = 0$	2 natural	6
1 or 2 ideal $\mathbf{l}'_\infty{}^\top \mathbf{v}_i \neq 0$	2 natural	6
1 or 2 ideal $\mathbf{l}'_\infty{}^\top \mathbf{v}_i = 0$	2 natural	5
$\mathbf{l}_\mathbf{a}{}^\top \mathbf{v}_i \neq 0$	2 fixed	7
$\mathbf{l}_\mathbf{a}{}^\top \mathbf{v}_i = 0$ and $\mathbf{l}_\mathbf{a} = \mathbf{l}_i$	2 fixed	4

Table 1: Constraints on $\boldsymbol{\omega}$ for various combinations of cameras and motion given the vanishing points $\mathbf{v}_i$ of three orthogonal directions in the scene. The first three cases describe a single view of the scene where additional constraints are obtained from the knowledge that the camera is natural. Where two views are available, the vanishing line of the second view image plane seen in the first view is $\mathbf{l}'_\infty$. In the case of a fixed camera, $\mathbf{l}_\mathbf{a}$ is the fixed line of $\mathtt{H}_\infty$ and $\mathbf{l}_i$ a side of the triangle of orthogonal vanishing points

5 Implementation

The calibration methods we have described rely heavily on the computation of vanishing points. The points are usually computed from imaged parallel line segments, but due to 'noise' a set of such segments will generally *not* intersect in a unique point. Often the vanishing point is then computed by finding the closest point to all the measured lines[2]. However, this is not optimal. The maximum likelihood estimate (MLE) of the vanishing point and line segments is found by computing a set of lines that do intersect in a single point, and which minimise the sum of squared orthogonal distances from the endpoints of the measured line segments [14], as shown in figure 8. The minimization is computed using the Levenberg-Marquardt algorithm [19].

When dealing with two views, vanishing point estimation is important both in defining consistent orthogonal triads in each view and in computing the infinite homography. For each vanishing point corresponding over two views, the MLE is obtained by varying a point $\mathbf{V}$ in 3-space which projects to the vanishing points $\mathbf{v}, \mathbf{v}'$ in the first and second views respectively [7]. The position of $\mathbf{V}$ is determined by minimizing a cost function of the sum of squared orthogonal distances from the end points of the measured (imaged parallel) lines to lines through $\mathbf{v}$ in the first view, and through $\mathbf{v}'$ in the second. See figure 8. Three such 3-space points determine the plane at infinity $\boldsymbol{\pi}_\infty$, and thence $\mathtt{H}_\infty$.

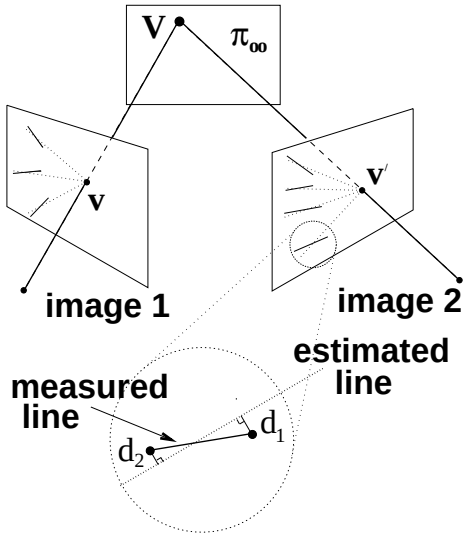


Figure 8: MLE Computation of $\boldsymbol{\pi}_\infty$ from vanishing points.

6 Further examples

A simple calibration object: Figure 9 shows a calibration object consisting of three planes imprinted with squares. Due to the presence of the squares, the planes are easily rectified. Each plane then provides two imaged circular points and thus two constraints on $\boldsymbol{\omega}$, giving a total of six, so three planes are sufficient to over-constrain the camera. If the camera is natural, so the image plane circular points are known, two scene planes are sufficient. As a calibration object, the planes have the advantages that no 3D co-ordinates are required and that the planes need not be orthogonal.



$$\mathtt{K} = \begin{bmatrix} 1096.91 & -24.43 & 510.5 \\ 0 & 1057.49 & 375.7 \\ 0 & 0 & 1 \end{bmatrix}$$

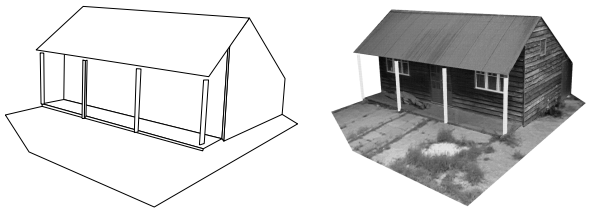
Figure 9: A simple calibration object and the computed camera. The image size is 1024×768 pixels.

This linear method of calibrating a single camera from multiple scene planes on which metric structure (such as

a square) is known was given in [14], and has subsequently been developed by [13, 22, 26]. Triggs [24] goes further and gives an iterative method in the case that metric structure is unknown.

Metric reconstruction from two views: Given two views of a scene with unknown cameras, a reconstruction from matched scene features is possible up to an unknown homography of 3D if the fundamental matrix is known [5, 11].

As an example, consider the image pair of figure 7. The infinite homography and π_∞ can be computed from the matched vanishing points for the three orthogonal directions as described in section 5. The structure may be affine rectified from π_∞ [7]. The calibration using the triad of orthogonal vanishing points, assuming both cameras are natural as in section 4.2, then defines the metric rectification of 3 space. The metric rectified model is shown in figure 10.



(a) (b)

Figure 10: Metric reconstruction computed from 2 views. (a) Wireframe (b) Texture mapped model.

7 Conclusion

We have developed a consistent calibration procedure that combines calibration information from scene constraints, knowledge of the camera, and the motion between views. The approach has the advantage of exposing conditions under which the calibration is likely to be unstable by identifying degenerate geometries. The linear solutions for ω provided by the calibration algorithms described may be used as a starting point for Maximum Likelihood Estimate optimisations. In the ML estimation the constraints (such as 6 circular points exactly lying on ω, as in the case of 3 rectified planes in figure 9) are enforced.

It is essential that some estimate of the accuracy of the computed cameras be made. We are currently investigating the estimation of confidence intervals for the calibration parameters based on first order uncertainty propagation.

References

[1] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. In *Proc. ECCV*, LNCS 1064/5, pages 3–16. Springer-Verlag, 1996.

[2] B. Caprile and V. Torre. Using vanishing points for camera calibration. *IJCV*, pages 127–140, 1990.

[3] R. Cipolla and E.G. Boyer. 3D model acquisition from uncalibrated images. In *Proc. IAPR Workshop on Machine Vision Applications, Chiba, Japan*, pages 559–568, November 1998.

[4] P. E. Debevec, C. J. Taylor, and J. Malik. Modeling and rendering architecture from photographs: A hybrid geometry- and image- based approach. In *Proceedings, ACM SIGGRAPH*, pages 11–20, 1996.

[5] O. Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? In *Proc. ECCV*, LNCS 588, pages 563–578. Springer-Verlag, 1992.

[6] O. D. Faugeras. Stratification of three-dimensional vision: projective, affine, and metric representation. *J. Opt. Soc. Am.*, A12:465–484, 1995.

[7] O.D. Faugeras, S. Laveau, L. Robert, G. Csurka, and C. Zeller. 3-D reconstruction of urban scenes from sequences of images. Tech. report, INRIA, 1995.

[8] G. H. Golub and C.F. Van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore, MD, second edition, 1989.

[9] R. I. Hartley. Euclidean reconstruction from uncalibrated views. In J. Mundy, A. Zisserman, and D. Forsyth, editors, *Applications of Invariance in Computer Vision*, LNCS 825, pages 237–256. Springer-Verlag, 1994.

[10] R. I. Hartley. Self-calibration from multiple views with a rotating camera. In *Proc. ECCV*, LNCS 800/801, pages 471–478. Springer-Verlag, 1994.

[11] R. I. Hartley, R. Gupta, and T. Chang. Stereo from uncalibrated cameras. In *Proc. CVPR*, 1992.

[12] G.E. Karras, P. Patias, and E. Petsa. Experiences with rectification of non-metric digital images when ground control is not available. In *CIPA XV International Symposium*, 1993.

[13] D. Liebowitz, A. Criminisi, and A. Zisserman. Creating architectural models from images. In *Proc. EuroGraphics*, Sep 1999.

[14] D. Liebowitz and A. Zisserman. Metric rectification for perspective images of planes. In *Proc. CVPR*, pages 482–488, Jun 1998.

[15] Q. T. Luong and T. Vieville. Canonical representations for the geometries of multiple projective views. *CVIU*, 64(2):193–229, Sep 1996.

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

[17] M. Pollefeys, R. Koch, and L. Van Gool. Self calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In *Proc. ICCV*, pages 90–96, 1998.

[18] M. Pollefeys, L. Van Gool, and A. Oosterlinck. The modulus constraint: a new constraint for self-calibration. In *Proc. ICPR*, pages 31–42, 1996.

[19] W. Press, B. Flannery, S. Teukolsky, and W. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1988.

[20] J. Semple and G. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, 1979.

[21] C. Slama. *Manual of Photogrammetry*. American Society of Photogrammetry, Falls Church, VA, USA, 4th edition, 1980.

[22] P. Sturm and S. Maybank. On plane based camera calibration: A general algorithm, singularities, applications. In *Proc. CVPR*, pages 432–437, June 1999.

[23] W. Triggs. Auto-calibration and the absolute quadric. In *Proc. CVPR*, pages 609–614, 1997.

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

[25] T. Vieville and D. Lingrand. Using singular displacements for uncalibrated monocular vision systems. Technical Report 2678, I.N.R.I.A., 1995.

[26] Z. Zhang. A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, Microsoft, 1998.

[27] A. Zisserman, D. Liebowitz, and M. Armstrong. Resolving ambiguities in auto-calibration. *Phil. Trans. R. Soc. Lond. A*, 356(1740):1193–1211, 1998.