Methods and Geometry for Plane-Based Self-Calibration

Pierre Gurdjos, Peter Sturm

HAL Id: inria-00525633

https://hal.inria.fr/inria-00525633

Submitted on 26 May 2011

Methods and Geometry for Plane-Based Self-Calibration[*]

Pierre Gurdjos

IRIT-TCI, 118 Route de Narbonne
31062 Toulouse Cedex 4, France
www.irit.fr/~Pierre.Gurdjos

Peter Sturm

INRIA Rhône-Alpes, 655 Avenue de l'Europe
38330 Montbonnot, France
www.inrialpes.fr/movi/people/Sturm

Abstract. We consider the problem of camera self-calibration, from images of a planar object with unknown Euclidean structure. The general case of possibly varying focal length is addressed. This problem is non-linear in general. One of our contributions is a non-linear approach, that makes abstraction of the (possibly varying) focal length, resulting in a computationally efficient algorithm. In addition, it does not require a good initial estimate of the focal length, unlike previous approaches. As for the initialization of other parameters, we propose a practical approach, that simply requires to take one image in roughly fronto-parallel position. Closed-form solutions for various configurations of unknown intrinsic parameters are provided. Our methods are evaluated and compared to previous approaches, using simulated and real images. Besides our practical contributions, we also provide a detailed geometrical interpretation of the principles underlying our approach.

1. Introduction

Calibration of a camera consists in recovering its metric properties, which are encoded as a set of so-called internal parameters. Two very opposite assumptions can be made: the Euclidean structure of the (observed) world is known or not. The term *calibration* usually applies to *known* 1D [13], 2D [8, 12, 2] and 3D structures. The prefix *self-* is added as soon as the world's Euclidean structure is *unknown*, which can be seen as a case of "0D" calibration.

In this paper, we focus on plane-based self-calibration [10, 5], *i.e.* from images of a rigid planar object lying on some plane – the "world plane" – whose Euclidean structure is unknown. The only prerequisite is to have at one's disposal the inter-image homographies (induced by the world plane), for which "direct" estimation algorithms, *i.e.* with closed-form solutions, exist [3, §3.1, pp. 71-76][1].

2. A Stratified Problem Formulation

The two goals of our work are to calibrate the camera and to recover the plane's Euclidean structure at the same time. On the one hand, it is well known that camera calibration is equivalent to computing the image of the absolute conic [3, §7.5, pp. 198-204]. On the other hand, the plane's structure can be characterized at different geometric levels: projective, affine or Euclidean by two (complex) points, called circular points (*cf.* §3). The circular points (CP) of the plane lie on the absolute conic (AC), and thus, in all views, the images of the CP (ICP) lie on the image of the AC (IAC). We now express the problems using this stratified framework.

We first consider the case where we have prior knowledge of the plane's Euclidean structure. The problem reduces to that of camera calibration and linear solutions exist [8, 12], even for varying intrinsics [8, 2]. Technically speaking, calibration can be seen as fitting an imaginary conic (the IAC) to all available ICP. This can be done by solving a linear equation system.

Consider now the general case, where the plane's Euclidean structure is unknown. We introduce unknowns to parameterize the ICP in one of our views, and compute the ICP in the other views using inter-image homographies. The estimation problem becomes non-linear, and iterative methods for its solution have been proposed [10, 5]. One of their drawbacks, common to non-linear problems, is the need for "good" initial estimates. Another problem is that the number m of unknowns may become relatively large in the case of varying intrinsics, increasing the sensitivity to the initial estimates and computation time (generally with $\mathcal{O}(m^3)$ complexity per iteration). One of our contributions is a parameterization that allows to solve the problem using a fixed number of unknowns (reducing the complexity to $\mathcal{O}(m)$), and that has a nice geometric interpretation (not shown completely in this paper due to lack of space).

Up to now, we have considered two extreme cases: completely known or completely unknown Euclidean plane structure. In the latter case, we know at least its projective structure (every image is a projective "model" of the plane). Consider now the obvious intermediate case: known affine structure. The problem remains non-linear, but can be expressed using fewer unknowns and simpler equations. Most importantly, closed-form solutions for interesting minimal cases are now possible (see §5.3). One way of recovering the plane's affine structure is *e.g.* to identify the projections of two sets of parallel lines on the plane. Another solution, that we use in this paper, is to simply take a fronto-parallel image of the plane. Taking an exactly fronto-parallel image is of course difficult. However, we show that in practice, a

[1]Interesting algorithms with "real-time" performance also exist [4].

roughly fronto-parallel image is sufficient to get good initial estimates using the closed-form solutions. We use them to start a non-linear optimization process, where the assumption of fronto-parallelism can be dropped.

We thus have established a stratification for plane-based calibration: *calibration* relies on the knowledge of the plane's Euclidean structure, whereas *self-calibration* only requires its projective structure. The intermediate case of known affine structure is analogous to using scene constraints in traditional 3D self-calibration [11], or, in the case of a fronto-parallel image, to self-calibration based on special motions (typically, pure translations [1, 7]).

3. Background

3.1. Calibration and the Absolute Conic

Our camera model is the usual specialization of the finite projective camera model [3, §5.1, pp. 143-144] with zero skew ; it has four d.o.f. linked to the (constant) principal point $(u_0, v_0)^\top$, the (constant) aspect ratio τ and the (possibly varying) focal length f. If A is the calibration matrix, then the IAC can be represented by $\boldsymbol{\omega} = \mathsf{A}^{-\top}\mathsf{A}^{-1}$, with:

$$\boldsymbol{\omega} \sim \begin{pmatrix} \tau^2 & 0 & -\tau^2 u_0 \\ 0 & 1 & -v_0 \\ -\tau^2 u_0 & -v_0 & \tau^2 u_0^2 + v_0^2 + \tau^2 f^2 \end{pmatrix}. \quad (1)$$

3.2. Plane's Structure and Circular Points

When some world plane Π is projected onto the image $\mathcal{I}$, the world-to-image homography that maps points $\mathbf{p}$ on Π onto pixels $\mathbf{m}$ on $\mathcal{I}$ is defined by a 3×3 matrix P such that $\mathbf{m} \sim \mathsf{P}\mathbf{p}$. We assume that some Euclidean coordinate system is attached to Π and P is decomposable into a product of three 3×3 matrices [3, §1.4.6, pp. 22-23]:

$$\mathsf{P} = \mathsf{P}_p \mathsf{P}_a \mathsf{P}_s, \quad (2)$$

where P_p (resp. P_a) is the projective (resp. affine) component of P, with four d.o.f. in all, and P_s is the Euclidean component of P with four additional d.o.f. (*i.e.* P_s is a 2D similarity in Π).

Since the matrix $(\mathsf{P}_p \mathsf{P}_a)^{-1}$ defines a *metric rectification* of the imaged plane, by mapping points on $\mathcal{I}$ onto Π w.r.t. some "arbitrary" Euclidean coordinate system, it is possible to ignore the component P_s in (2) so that:

$$\mathsf{P} = \begin{pmatrix} \beta & \alpha & 0 \\ 0 & 1 & 0 \\ \mu\beta & \mu\alpha + \lambda & 1 \end{pmatrix}, \quad (3)$$

where α, β, λ, μ are four scalars that encode the world plane's Euclidean structure.

Under P, the world plane's line at infinity is mapped to its vanishing line $\mathbf{v}$ according to:

$$\mathbf{v} = \mathsf{P}^{-\top} (0, 0, 1)^\top \sim (-\mu, -\lambda, 1)^\top. \quad (4)$$

Hence λ, μ represent the plane's affine structure (the meaning of α, β is discussed below).

In the following paragraphs, we establish links between the world-to-image homography P, as defined in (3), and the circular points.

Circular Points (CP). The circular points (CP) of a plane have the following properties: (1) They lie on the absolute conic. (2) Their coordinates $\mathbf{i}_\pm = (1, \pm i, 0)^\top$ are the same in every Euclidean coordinate system.

Image of Circular Points (ICP). Under P, the CP transform into the ICP (images of circular points) according to:

$$\mathsf{P}\mathbf{i}_\pm = (\beta, 0, \mu\beta)^\top \pm i (\alpha, 1, \mu\alpha + \lambda)^\top. \quad (5)$$

The ICP are another representation, besides P, of the world plane's Euclidean structure. Regarding the canonical representation $\mathbf{v} = (0, 0, 1)^\top$ in (4), it now can be readily seen that the vector $(\alpha \pm i\beta, 1)^\top$ represents the 1D homogeneous coordinates of the ICP on the line at infinity.

Conic Dual to the Circular Points (CDCP). The conic dual to the circular points is the symmetric rank-2 matrix defined by $\mathsf{D}_\infty^* = \mathbf{i}_+ \mathbf{i}_-^\top + \mathbf{i}_- \mathbf{i}_+^\top = \mathrm{diag}(1, 1, 0)$.

Under P, the CDCP transforms into:

$$\boldsymbol{\Sigma} = \mathsf{P}\mathsf{D}_\infty^*\mathsf{P}^\top. \quad (6)$$

The conic $\boldsymbol{\Sigma}$ is yet another representation of the world plane's Euclidean structure [3, §1.7.5, pp. 34-37], which will be used in our self-calibration approach. It is worth noting that $\boldsymbol{\Sigma}$ is linked to the vanishing line $\mathbf{v}$ by the property: $\mathrm{null}(\boldsymbol{\Sigma}) = \mathbf{v} = (-\mu, -\lambda, 1)^\top$.

Under A, the CDCP transforms into:

$$\boldsymbol{\Lambda} = \mathsf{A}\mathsf{D}_\infty^*\mathsf{A}^\top \sim \mathrm{diag}(1, \tau^2, 0). \quad (7)$$

with the key property, used later: image lines $\mathbf{m}_1$ and $\mathbf{m}_2$ are orthogonal (w.r.t. the image frame) iff $\mathbf{m}_1^\top \boldsymbol{\Lambda} \mathbf{m}_2 = 0$.

4. Plane-based Calibration

In this section, we review constraints on the IAC $\boldsymbol{\omega}$ that are used to solve the plane-based calibration problem. We first review the basic equations introduced in [8, 12]. In §4.2, we then describe an approach leading to equations that do not take into account the focal length. This allows the number of unknowns to remain constant, even in the case of a varying focal length [2]. This advantage may be very interesting for the non-linear self-calibration problem, and in §5.2, we accordingly extend the approach of §4.2.

4.1. Basic Equations

Plane-based calibration is based on two key ideas. First, the CP (encoding the plane's Euclidean structure) lie on the AC (encoding the camera's internal parameters). Second, the CP are mapped to their images (ICP) via the world-to-image homography P. Hence the fact that the ICP lie on the IAC can be expressed by: $(\mathsf{P}\mathbf{i}_\pm)^\top \boldsymbol{\omega}(\mathsf{P}\mathbf{i}_\pm) = 0$, or equivalently:

$$\mathbf{p}_1^\top \boldsymbol{\omega} \mathbf{p}_1 = \mathbf{p}_2^\top \boldsymbol{\omega} \mathbf{p}_2, \quad \mathbf{p}_1^\top \boldsymbol{\omega} \mathbf{p}_2 = 0, \quad (8)$$

where $\mathbf{p}_1$ and $\mathbf{p}_2$ are the first two columns of P. These constraints are linear in the elements of $\boldsymbol{\omega}$.

4.2. The Centre Line Constraint

Equations (8) include the focal length (contained in ω) which might be disadvantageous, as explained above. Thus, we now describe an alternative approach, based on a geometric constraint on the principal point, called *centre line constraint* [2], that is regardless of the focal length. The centre line constraint results from the projective properties of the central projection of a planar object. It expresses the fact that the locus of the (unknown) principal point of a camera viewing a (known) planar object is constrained to a (known) line segment, called *principal segment*. In other words, for any point taken as principal point on this segment, there exists a world plane on which the object can lie and yield *exactly* the same image. What it is mentioned here is best explained by the animation available at the URL www.irit.fr/~Pierre.Gurdjos/ECCV2002/.

The centre line. The image line that contains the principal segment is called *centre line*. In [2], we shown that:
(i) The centre line is orthogonal to the vanishing line.
(ii) The vector that represents its image coordinates only depends on the world-to-image homography P:

$$\boldsymbol{\phi} = (-\varphi_1, -\varphi_2, \varphi_3 + \varphi_4)^\top,$$

where $\varphi_i = \varphi_i(\mathsf{P})$ is a bilinear function of elements of P.

To put it more precisely, the property (ii) holds providing pixels are square, so one of our contributions is to define the locus of the centre line, in presence of a non-unit aspect ratio τ. Furthermore, we will give a matrix representation of the centre line constraint in terms of $\boldsymbol{\Lambda}$ and $\boldsymbol{\Sigma}$, more suitable for self-calibration as described in §5.2.

The pencil of centre lines. The following properties hold w.r.t. an arbitrary aspect ratio τ (proofs are omitted).
(iii) The vector that represents the image coordinates of the centre line can be written as:

$$\boldsymbol{\phi} = (0, -\varphi_2, \varphi_4)^\top + \tau^2 (-\varphi_1, 0, \varphi_3)^\top \sim \mathbf{d}_1 + \tau^2 \mathbf{d}_2,$$

where $\mathbf{d}_1$, $\mathbf{d}_2$ are the two vectors (written in terms of $\boldsymbol{\Sigma}$):

$$\mathbf{d}_1 = (\boldsymbol{\Sigma}\mathbf{e}_3) \times \mu\mathbf{e}_1, \quad \mathbf{d}_2 = (\boldsymbol{\Sigma}\mathbf{e}_3) \times \lambda\mathbf{e}_2,$$

in which $\mathbf{e}_1^\top = (1,0,0), \mathbf{e}_2^\top = (0,1,0), \mathbf{e}_3^\top = (0,0,1)$.
(iv) The locus of the centre line is then a linear family i.e. a line pencil, denoted by $\mathscr{P}$, irrespective of τ, with vertex:

$$\mathbf{d}_1 \times \mathbf{d}_2 \sim \boldsymbol{\Sigma}\mathbf{e}_3. \tag{9}$$

Basically, (iii) and (iv) mean that we solely inherit from the world-to-image homography P a pencil of (centre) lines.
(v) The following points are collinear: the vertex $\boldsymbol{\Sigma}\mathbf{e}_3$, defined in (9), the principal point $\mathbf{p}_0^\top = (u_0, v_0, 1)$ and the point given by $\boldsymbol{\Lambda}\mathbf{v}$; so we have:

$$\det (\boldsymbol{\Sigma}\mathbf{e}_3 \mid \boldsymbol{\Lambda}\mathbf{v} \mid \mathbf{p}_0) = 0. \tag{10}$$

Note that the point $\boldsymbol{\Lambda}\mathbf{v}$ is the point at infinity in the direction "orthogonal" to the vanishing line $\mathbf{v}$, *i.e.* in terms of $\boldsymbol{\Lambda}$, *cf.* (7). It is the point where all the lines parallel to ϕ meet (*cf.* Fig. 1). This is actually an alternative representation of

the centre line constraint. It links the aspect ratio τ (in $\boldsymbol{\Lambda}$), the principal point $(u_0, v_0)^\top$, and the parameters $\alpha, \beta, \lambda, \mu$ of the plane's Euclidean structure (in $\boldsymbol{\Sigma}$). This constraint is the basis for our self-calibration approach, *cf.* §5.2.
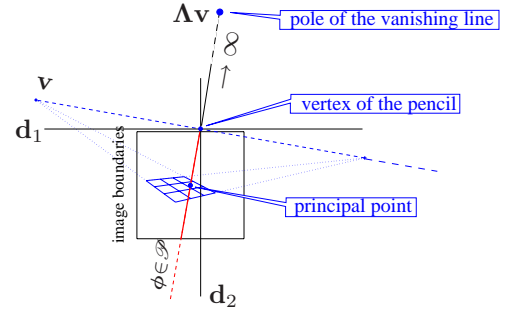

Figure 1: The pole of the vanishing line $\mathbf{v}$ w.r.t. $\boldsymbol{\Lambda}$, the vertex of the line pencil $\mathscr{P}$ (*i.e.* $\mathbf{d}_1 \times \mathbf{d}_2$) and the principal point are aligned.

Eventually, it is worth noting that:
(vi) From (9), it follows that the vertex of $\mathscr{P}$ is the pole of the image plane's line at infinity w.r.t. the imaged CDCP $\boldsymbol{\Sigma}$.
(vii) The intersection point of ϕ and the line at infinity is the pole of the vanishing line w.r.t. $\boldsymbol{\Lambda}$, i.e. $\boldsymbol{\phi} \times \mathbf{e}_3 = \boldsymbol{\Lambda}\mathbf{v}$.

5. Plane-Based Self-Calibration

5.1. Existing Non-Linear Solutions

"Basic" constraints on the IAC and ICP. In [10], a solution is given for plane-based self-calibration in the case of constant internal parameters, which has been extended in [3, §18.7, pp. 469-471] to the "varying focal length" case. The approach is basically the same as in calibration (*cf.* §4.1) with the difference that the ICP map from image to image via the inter-image homographies H_{ij}. If we denote by $\mathbf{p}_{1\pm} = \mathsf{P}\mathbf{i}_\pm$ the ICP in some key image (say image 1), then we have $\mathbf{p}_{j\pm} = \mathsf{H}_{1j}\mathbf{p}_{1\pm}$. Two equations are therefore provided by each image j:

$$(\mathsf{H}_{1j}\mathbf{p}_{1\pm})^\top \omega_j(\mathsf{H}_{1j}\mathbf{p}_{1\pm}) = 0. \tag{11}$$

Given n views, *i.e.* $(n-1)$ inter-image homographies H_{1j} ($1 < j \leq n$; $\mathsf{H}_{11} = \mathsf{I}$), the *self-calibration problem* of a camera with a possibly *varying focal length* is that of solving the system of $2n$ equations (11) for the $3 + m$ d.o.f. in all the ω_j and the 4 d.o.f. in $\mathbf{p}_{1\pm}$, where m is the number of unknown focal lengths. If $m = 1$, at least 4 views are required ; if $m = n$, at least 7 are required.

This problem is non-linear and can be solved using iterative methods. It requires initial values, in particular for the (possibly different) focal lengths. This critical issue, already mentioned in [10], motivated us to (1) seek a minimization criterion that is irrespective of f (see §5.2) ; (2) find closed-form solutions for minimal cases (see §5.3).

5.2. A New Non-Linear Solution

The centre line constraint (10) in §4.2 can be seen as somewhat of a dual formulation of Triggs's approach, described

in §5.1. Indeed, it involves the CDCP instead of the CP. If we denote by $\boldsymbol{\Sigma}$ the image of the CDCP in some key image (say again image 1), then we have $\boldsymbol{\Sigma}_j = \mathsf{H}_{1j}\boldsymbol{\Sigma}\mathsf{H}_{1j}^\top$, where H_{1j} is the inter-image homography from image 1 to image j. Moreover, under H_{1j}, the vanishing line $\mathbf{v}$ of image 1 transforms into $\mathbf{v}_j = \mathsf{H}_{1j}^{-\top}\mathbf{v}$.

It follows that one equation is provided by each image j:

$$\det\left(\mathsf{H}_{1j}\boldsymbol{\Sigma}\mathsf{H}_{1j}^\top\mathbf{e}_3 \mid \boldsymbol{\Lambda}\mathsf{H}_{1j}^{-\top}\mathbf{v} \mid \mathbf{p}_0\right) = 0. \qquad (12)$$

Problem 1 Given n views, *i.e.* $(n-1)$ inter-image homographies H_{1j} $(1 < j \le n$; $\mathsf{H}_{11} = \mathsf{I})$, the *self-calibration problem* of a camera with a possibly *varying focal length* is that of solving the system of n equations (12) for the 2 d.o.f. in $\mathbf{p}_0$, the single d.o.f. in $\boldsymbol{\Lambda}$ and the 4 d.o.f. in $\boldsymbol{\Sigma}$, under the condition $\mathrm{null}(\boldsymbol{\Sigma}) = \mathbf{v}$.

There is a *fixed* number of 7 unknowns $(\alpha, \beta, \lambda, \mu$ and $u_0, v_0, \tau)$ – *even with* a varying focal length – so that at least 7 views are required. Once the parameters are recovered, the different focal lengths can be computed using linear algorithms described in §4.

Implementation. Referring to (12), let us denote by M the matrix $\left(\mathsf{H}\boldsymbol{\Sigma}\mathsf{H}^\top\mathbf{e}_3 \mid \boldsymbol{\Lambda}\mathsf{H}^{-\top}\mathbf{v} \mid \mathbf{p}_0\right)$ (we omit the index for H). A solution to Problem 1 can be obtained by minimizing a cost function depending on $\det\mathsf{M}$. This is a non-linear problem but let us notice that since the condition $\mathrm{null}(\boldsymbol{\Sigma}) = \mathbf{v}$ is directly ensured from the definition of $\boldsymbol{\Sigma}$ in (6), no constrained optimization algorithm is required.

An interesting aspect of our formulation is that we can easily attach a geometric meaning to the algebraic quantity $\det\mathsf{M}$. Indeed, $\det\mathsf{M}$ is equal to the mixed triple product of its three column vectors: hence $\delta = \frac{1}{k}|\mathbf{m}_1^\top(\mathbf{m}_2 \times \mathbf{m}_3)|$, where $k = \sqrt{\boldsymbol{\phi}^\top\boldsymbol{\Lambda}\boldsymbol{\phi}}$, represents the distance from point $\mathbf{m}_1$ to the (centre) line $\boldsymbol{\phi} \sim \mathbf{m}_2 \times \mathbf{m}_3$.

In our experiments, we used the non-linear least squares implementation (Levenberg-Marquardt algorithm) available in the MATLAB Optimization Toolbox [6]. The Jacobian information for the objective function can be easily supplied, using the following properties:
(1) $\det\mathsf{M} = (H_{31} + \mu H_{33})\det\mathsf{N}_1 + (H_{32} + \lambda H_{33})\det\mathsf{N}_2$, where $\mathsf{N}_i = \left(\ \mathsf{H}\boldsymbol{\sigma}_i \mid \mathbf{m}_2 \mid \mathbf{m}_3\ \right)$; $\boldsymbol{\sigma}_i$ is the column i of $\boldsymbol{\Sigma}$.
(2) $\frac{d\det\mathsf{M}}{dx} = \mathrm{tr}\left\{\frac{d\mathsf{M}}{dx}\,\mathrm{adj}\,\mathsf{M}\right\}$, where $\mathrm{adj}\,\mathsf{M} = (\det\mathsf{M})\,\mathsf{M}^{-1}$ is the adjoint matrix of M.

5.3. Direct Solutions

We now develop closed-form solutions, based on the assumption that one image was taken in fronto-parallel position relative to the world plane (*i.e.* image and world planes are parallel). As mentioned in §2, we immediately have the world plane's affine structure; equivalently, we now have $\lambda = \mu = 0$ for the representations described in §3.2. Keep in mind that we use the assumption of fronto-parallelism *only for the algorithm initialization*. For the subsequent non-linear optimization, we drop this assumption.

In the fronto-parallel image, the ICP lie on both the IAC and plane's vanishing line (here, the image plane's line at infinity). Looking at (1), we see that the only points at infinity that lie on the IAC are the two points given by:

$$\left(\tau^{-1}, 0, 0\right)^\top \pm i\left(0, 1, 0\right)^\top \sim \left(1, \pm\tau i, 0\right)^\top. \qquad (13)$$

Consequently, using (5), we know that $\alpha = 0$ and thus, we can recover the world plane's Euclidean structure up to the single unknown $\beta = \tau^{-1}$.

We now sketch closed-form solutions for various scenarios, depending if the aspect ratio and/or principal point are known or not, and if the focal length is constant or varying. In the case of a *known aspect ratio*, the fronto-parallel image directly gives us the plane's Euclidean structure, and self-calibration reduces to calibration [8, 12, 2]. So, in the following, we only consider an *unknown aspect ratio*.

As shown in (13), the ICP in the fronto-parallel image can be parameterized by the unknown τ. Using inter-image homographies, we also parameterize the ICP in the other images using τ. Let H be the homography, mapping the fronto-parallel to some other image. The basic calibration equations (8) then become:

$$\mathbf{h}_1^\top\boldsymbol{\omega}\mathbf{h}_1 - \tau^2\mathbf{h}_2^\top\boldsymbol{\omega}\mathbf{h}_2 = 0, \qquad \mathbf{h}_1^\top\boldsymbol{\omega}\mathbf{h}_2 = 0. \qquad (14)$$

The second equation is linear and the same as in (8), hence with 5 or more inter-image homographies, the unknowns can be recovered linearly. As for the first equation, using the fact that $\tau^2 = \omega_{11}/\omega_{22}$, we may reformulate it as:

$$\omega_{22}\mathbf{h}_1^\top\boldsymbol{\omega}\mathbf{h}_1 - \omega_{11}\mathbf{h}_2^\top\boldsymbol{\omega}\mathbf{h}_2 = 0.$$

This is quadratic in the set of coefficients of $\boldsymbol{\omega}$, with only ω_{11} and ω_{22} appearing squared. In the following, we describe several minimal cases, but due to lack of space, without much detail. Note that the focal length of the fronto-parallel image can not be recovered [8], so we ignore it.

In the case of a *known principal point*, two images, the fronto-parallel and another one, are sufficient for self-calibration. The only unknowns are the aspect ratio and the focal length of the second view (that may be different from that of the fronto-parallel view). We suppose that the images are centered in the principal point, *i.e.* we have $\boldsymbol{\omega} = \mathrm{diag}(\tau^2, 1, \tau^2 f^2)$. Equations (14) thus become, after replacing the unknowns by $a = \tau^2$ and $b = \tau^2 f^2$:

$$H_{21}^2 + a(H_{11}^2 - H_{22}^2) + bH_{31}^2 - abH_{32}^2 - a^2H_{12}^2 = 0$$
$$H_{21}H_{22} + aH_{11}H_{12} + bH_{31}H_{32} = 0.$$

The two equations can be reduced a single quadratic one in b. Writing down explicit closed-form solutions for τ and f in terms of H is trivial.

In the case of an *unknown principal point* and *constant (resp. varying) focal length*, three (resp. four) images are sufficient and the problem can be written as a cubic (resp. quartic) polynomial in one variable. Hence, self-calibration has a closed-form solution.

6. Experiments

Synthetic data. Self-calibration using the centre line constraint of §5.2 ("CL-NONL-SELF") has first been tested using synthetic data. We compare it with the results of the algorithm using the basic constraints, see §5.1 ("BAS-NONL-SELF"). For each experiment, the camera has constant internal parameters with nominal values $u_0 = 255 \pm 50$ pixels, $v_0 = 255 \pm 50$ pixels, $\tau = 1 \pm 0.1$ (with normal distribution). For each camera c in each experiment, the inclination angle between the world and the image planes is set to $30° \pm 10°$ (except for the first for which it is set to $\pm 10°$); the angles for azimuth and rotation around the optical axis are set to $0° \pm 90°$ (normal distribution); the (varying) focal length is set to $f_c = 700 \pm 700$ pixels (normal distribution). 100 points are randomly generated in the first image, then transferred to the others with a of perturbation ± 1 pixel (Gaussian noise). The inter-image homographies have been estimated using the normalized DLT algorithm of [3, §3.1, pp. 71-73], from the perturbed points. We conducted 200 independent trials for a number of cameras varying from 8 to 24. In Fig. 2, we show the computed absolute errors for u_0, v_0, for world coordinates x, y (in mms), and relative errors for τ and the focal lengths. We also sought a threshold on the number of cameras for which the "CL-NONL-SELF" and "BAS-NONL-SELF" algorithms have similar accuracies. Regarding our tests, this threshold is about 15 views. Typically, the algorithm converges in 5 iterations. A "good" initialization of the parameters proved to be crucial. We used the direct solution given in §5.3, except in one case for "BAS-NONL-SELF": this case is plotted with the dashed line with marker '*' and corresponds to focal lengths initialized to 2000 (which is quite realistic) ; the "incoherent" behaviour of the algorithm is owing to the fact that, in a certain number of trials, it did not converge.

Real Images. To evaluate the performance of our non-linear self-calibration algorithm ("CL-NONL-SELF"), we compared the results with those obtained by both basic linear *calibration* ("BAS-LIN-CAL") and basic non-linear *self-calibration* ("BAS-NONL-SELF") algorithms. We used 15 images of a calibration checkerboard, taken from different orientations. A Nikon Coolpix 800 was used with resulting *jpeg* images of 640×480 pixels resolution in size. 48 corners have been extracted in order to compute the inter-image homographies. To avoid critical motions, we took care to apply significative rotations around the optical axis between successive shots. The (known) Euclidean structure of the calibration checkerboard was *only* used by the calibration algorithm BAS-LIN-CAL. The principal point and aspect ratio were assumed to be constant; their estimated values are: $(308, 250, 1.0083)$ for BAS-LIN-CAL, $(325, 253, 0.999)$ for BAS-NONL-SELF and $(325, 260, 0.999)$ for our CL-NONL-SELF algorithm. Note that the relative error between the different aspect ratios is less than 1%.
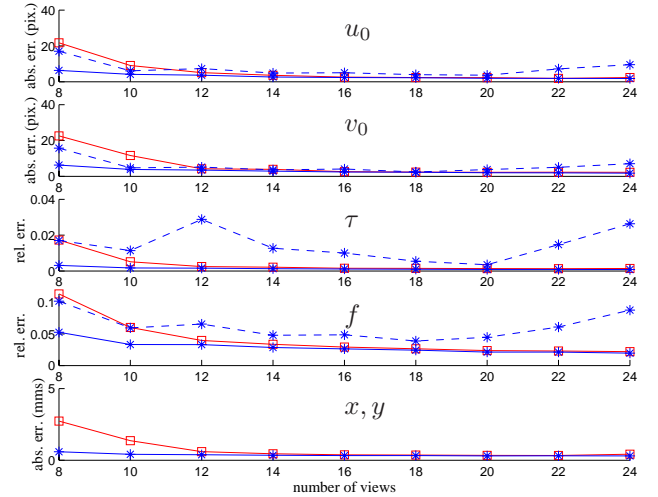

Figure 2: Self-calibration results. "CL-NONL-SELF" method: marker □. "BASIC-NONL-SELF" method: marker ∗ ; the dashed line highlights the "bad" initializations of f.

There could be variations of the focal length owing to the camera's auto-focus, so we assumed f to be varying. The different focal lengths recovered by the algorithm BAS-NONL-SELF are: 1368, 1390, 1383, 1352, 1357, 1357, 1371, 1322, 1346, 1352, 1358, 1345, 1390, 1394, 1387, with mean 1364 and standard deviation around 20 (1%). Table 1 gives the the relative difference (percent) with BAS-LIN-CAL obtained by CL-NONL-SELF. In brackets, the relative errors with BAS-NONL-SELF are shown (a negative value means "closer to BAS-LIN-CAL's estimates"). The relative errors between the calibration and self-calibration algorithms are very small (in most cases less than 1%) for all focal lengths.

0.5 (−0.1)	0.4 (−0.1)	1.2 (−0.1)	0.9 (−0.1)	0.7 (0.1)
0.0 (−0.1)	0.3 (0.1)	1.6 (1.0)	0.2 (0.1)	0.3 (0.2)
0.6 (0.3)	0.4 (0.2)	0.9 (0.1)	1.3 (0.0)	0.1 (−0.1)

Table 1: Focal length self-calibration obtained by the CL-NONL-SELF algorithm from 15 images of a checkerboard. Relative errors (percent) w.r.t. the BAS-LIN-CAL algorithm are shown. In brackets, the difference with the BAS-NONL-SELF algorithm.

Videos of a comic book. We acquired several videos of a comic book using a hand-held digital camcorder (Panasonic NV-MX 300). Acquisitions were started in roughly fronto-parallel position. The videos were processed automatically to extract and track interest points and to compute inter-image homographies (using a RANSAC-based method). This and the remaining processing was done on 1 out of every 10 frames. Figure 3 shows 8 of the 20 frames used for one of the sequences. We used the closed-form solution of §5.3 corresponding to an unknown aspect ratio but known principal point (image center). This gave the fo-

cal length for every frame but the first, and one estimate of the aspect ratio per frame. A single value for the aspect ratio was computed using robust statistics, and used to obtain an initial solution of the world plane's Euclidean structure from the fronto-parallel frame. Then, initial pose estimates for all frames were obtained [9].

These initial estimates were refined by a bundle adjustment (including position of points on the world plane and radial distortion). Bundle adjustment was implemented in the usual sparse way, and converged in 2 to 3 iterations, each iteration taking a few seconds.

The results were compared with calibration values obtained by filming a 3D calibration grid. The aspect ratio was estimated with 0.2% error (1.0893, compared to a "ground truth" of 1.0919). The principal point was estimated about 10 pixels off. As for the focal length, the results for the first half of the frames were bad as expected (frames are too close to fronto-parallel) whereas for the second half, the mean value was 1366, which means an error of 3.5% (ground truth was 1319). Similar results were obtained for other sequences of the same object.

Figure 3 shows a rectified image of the world plane, obtained from the first frame of the last row in the figure. The structure is well recovered, considering that the page of the comic book was not perfectly flat towards its left.

Figure 3: Left: input images. Right: rectified image.

7. Conclusion

We addressed the problem of camera self-calibration, from inter-image homographies induced by a single plane with unknown Euclidean structure. A non-linear solution, based on properties of circular points and the absolute conic, was previously proposed in [10]. This "basic" algorithm proved to be efficient, but requires a good initial estimate of the focal length. Our first contribution was to solve this issue by proposing a practical approach that simply requires to take a first image in roughly fronto-parallel position. Closed-form solutions for various configurations are obtained and discussed. The assumption of fronto-parallelism is only used for initialization and dropped for non-linear optimization.

Another contribution was to extend the so-called centre line constraint, already used for plane-based calibration in [2], to self-calibration. This constraint expresses the fact that the locus of the principal point is a line, called the centre line. With a non-unit aspect ratio, we showed that the locus of the centre line is a line pencil. The centre line constraint then equivalently expresses the alignment of three particular points, whose representations involve all intrinsic parameters *except the focal length* as well as, surprisingly enough, the images of the conic dual to the circular points (CDCP) ; technically speaking, self-calibration is based on the transformations of the CDCP via the inter-image homographies. Under the common assumption of supposedly known principal point and aspect ratio, the proposed constraint will not depend on the intrinsic parameters (*i.e.* only on the plane's structure), allowing metric "self-rectification" that generalizes [3, §1.7.5, pp. 34-37] ; this interesting point, not expanded here, had to be mentioned. We gave a comparative experimental evaluation for simulated and real data indicating good performances, even if we empirically noticed that the centre line-based algorithm had a higher number of critical configurations than the basic one.

References

[1] M. Armstrong, A. Zisserman, P. Beardsley. Euclidean Structure from Uncalibrated Images. BMVC, pp. 509-518, 1994.

[2] P. Gurdjos, A. Crouzil, R. Payrissat. Another Way of Looking at Plane-Based Calibration: the Centre Circle Constraint. ECCV, vol. 4, pp. 252-266, 2002.

[3] R. Hartley, A. Zisserman. Multiple View Geometry in Computer Vision. Cambridge University Press 2002.

[4] F. Jurie, M. Dhome. A simple and efficient template matching algorithm, ICCV, pp. 544-549, 2001.

[5] E. Malis, R. Cipolla. Multi-view Constraints between Collineations: Application to Self-Calibration from Unknown Planar Structures. ECCV, vol. 2, pp. 610-624, 2000.

[6] Optimization Toolbox User's Guide. Version 2. The MathWorks, Inc. September 2000.

[7] M. Pollefeys, L. Van Gool, M. Proesmans. Euclidean 3D Reconstruction from Image Sequences with Variable Focal Lengths. ECCV, pp. 31-42, 1996.

[8] P. Sturm, S. Maybank. On Plane-Based Camera Calibration: a General Algorithm, Singularities, Applications. CVPR, pp. 432-437, 1999.

[9] P. Sturm. Algorithms for Plane-Based Pose Estimation. CVPR, pp. 1010-1017, 2000.

[10] B. Triggs. Autocalibration from Planar Scenes. ECCV, pp. 89-105, 1998.

[11] C. Zeller. Calibration projective, affine et euclidienne en vision par ordinateur et application à la perception tridimensionnelle. PhD Thesis, École Polytechnique, 1996.

[12] Z. Zhang. A Flexible New Technique for Camera Calibration. PAMI, vol. 22, no. 11, pp. 1330-1334. 2000.

[13] Z. Zhang. Camera Calibration with One-Dimensional Objects. ECCV, vol. 4, pp. 161-174, 2002.