Towards a Guaranteed Solution to Plane-Based Self-calibration

Benoît Bocquillon, Pierre Gurdjos, and Alain Crouzil

Université Paul Sabatier, IRIT-TCI, 118, route de Narbonne,
31062 Toulouse, France
{bocquillon, gurdjos, crouzil}@irit.fr

Abstract. We investigate the problem of self-calibrating a camera, from multiple views of a planar scene. By self-calibrating, we refer to the problem of simultaneously estimate the camera intrinsic parameters and the Euclidean structure of one 3D plane. A solution is usually obtained by solving a non-linear system via local optimization, with the critical issue of parameter initialization, especially the focal length. Arguing that these five parameters are inter-dependent, we propose an alternate problem formulation, with only three d.o.f., corresponding to three parameters to estimate. In the light of this, we are concerned with global optimization in order to get a guaranteed solution, with the shortest response time. Interval analysis provides an efficient numerical framework, that reveals to be highly performant, with regard to both estimation accuracy and time-consuming.

1 Introduction

The self-calibration of a camera consists in determining, either partially or completely, the metric properties of the camera and/or the scene, from a set of uncalibrated views. The principle of self-calibration is to use "absolute entities" as targets, geometrically constrained by some prior information about the internal or external parameters of the camera. Absolute targets are abstract entities, located at infinity, encoding the Euclidean structure (ES) of the considered d-dimensional space, with the characteristic property of being left invariant under similarities[1] in d-space [1, 2]. In 3-space, the target is the *absolute conic* (AC), which is a circle of imaginary radius on the plane at infinity π_∞. The AC has the well-known property that its image (IAC) is globally invariant under camera motion, providing the camera internal parameters are constant. This is the starting point of numerous 3D self-calibration methods (see [1, chapter 19] for a review). On the basis of a projective reconstruction of the scene, 3D self-calibration determines the ES of the 3D space in terms of the AC *and* the plane at infinity, in projective coordinates. This can be achieved either separately, or simultaneously. In the latter case, the AC is treated as a rank-3 envelope of 3D planes, known as absolute dual quadric in the literature. Assuming that the focal length is the only unknown, closed-forms and linear solutions can be obtained e.g., as in [3].

The problem up to discussion in this work is the 2D (or plane-based) self-calibration of a camera i.e., by observing a 3D plane π, with regard to general camera motion. In

[1] i.e., transformations preserving angles and changing distances in the same ratio.

P.J. Narayanan et al. (Eds.): ACCV 2006, LNCS 3851, pp. 11–20, 2006.

2-space, the self-calibration targets are the *circular points* (CP) that are two conjugate complex points of π on the line at infinity, meeting all the circles of π and the AC of π_∞. Since Triggs' work [4], it is known that 2D self-calibration is possible, using the constraint that the image of the CP (ICP) lie on the IAC, only involving inter-view homographies induced by π. Because no other (general) invariance of the ICP can be exhibited, very few 2D self-calibration methods have been reported [5, 6, 4], except for some specific camera motion [7, 8]. Furthermore, contrary to 3D self-calibration, even with a simplified model of the camera, no closed-form or linear solution exist. Such a problem, consisting in determining simultaneously the CP and the AC, is non-linear in essence. As stated in [4], the problem parameterization requires 4 d.o.f. for the ICP plus 5 d.o.f. for the AC. A solution can be obtained via local optimization, from at least 5 views, with the critical issue of parameter initialization, especially the focal length.

Our starting point is to reduce the number of parameters to estimate by using the fact that, since the CP lie on the AC, there is a redundancy in the problem parameterization. This inter-dependence of parameters in Triggs' statement is a modeling constraint that has no reason not to be exactly ensured. Actually, Triggs initially treated it as an equation, which does not really make sense as we will argue later. That said, our contribution is to propose a new minimal parameterization of the 2D self-calibration problem, by introducing as target a degenerate conic envelope, consisting of the point-pair at which the AC meets the line at infinity i.e., consisting of the CP. Thanks to our propositions (1) and (2), we show that we only require to estimate the affine structure of the plane along with the internal parameters. This leads to a formulation with seven unknowns/d.o.f. instead of the nine initially mentioned in [4]. Assuming that the constant focal length is the sole unknown, only three parameters have to be estimated. This paves the way for finding a guaranteed solution to the problem as this small number of unknowns is well adapted to the use of interval analysis [9]. Interval analysis has been widely used in global optimization problems [10] and afford the guarantee that the global minimum has been found. Interval analysis has been succesfully used to the 3D self-calibration problem [11]. It provides an efficient numerical framework, that reveals to be highly performant, with regard to both estimation accuracy and time-consuming.

This paper is structured as follows. First, starting with the basic 2D self-calibration equations of [4], we explain how to obtain a minimal parameterization of the problem from which we derive a cost function. Second, we review the main rules of interval analysis and the global minimization scheme used here. Eventually, we give the results obtained with synthetic and real data and conclusions are drawn.

2 Minimal Parameterization of 2D Self-calibration

2.1 Foreword and Notations

Our problem is that of recovering the Euclidean structure (ES) of some 3D plane π, called world plane, seen in multiple views, for some *uncalibrated* camera. What is only assumed to be *known* is the inter-view homographies induced by π.

Without any additional knowledge, this problem cannot be separated from that of calibrating the camera i.e., of recovering its intrinsic parameters. Stated together, these are then referred to as the plane-based self-calibration problem [4]. [5] describes an alternative to [4]. We will give in §2.2 the link between these two constraints.

We use some MATLAB-like notations: $1 : n$ denotes the range $1, \dots, n$. $\mathtt{M}_{(1:r,1:c)}$ denotes the $r \times c$ submatrix of $\mathtt{M}$ selected by the row range $1 : r$ and the column range $1 : c$. The notation $\mathtt{M}_{(:,1:c)}$, resp. $\mathtt{M}_{(1:r,:)}$, selects the first c (resp. r) columns, resp. rows, of $\mathtt{M}$. We also define the canonical vectors:

$$\mathbf{e}_1 \equiv (1, 0, 0)^\top, \ \mathbf{e}_2 \equiv (0, 1, 0)^\top, \ \mathbf{e}_3 \equiv (0, 0, 1)^\top. \tag{1}$$

The matrix $[\mathbf{x}]_\times$ refers to the skew-symmetric, order-3, matrix, such that $[\mathbf{x}]_\times \mathbf{y} = \mathbf{x} \times \mathbf{y}$, $\mathbf{y} \in \mathbb{R}^3$. In this paper, we will make a heavily use of the equality $[\mathtt{T}\mathbf{x}]_\times = \det(\mathtt{T})\mathtt{T}^{-\top} [\mathbf{x}]_\times \mathtt{T}^{-1}$. The notation i *always* refers to the imaginary number $\sqrt{-1}$.

In the following we assume some basic results on projective geometry. These can be found in standard textbooks e.g., in [1, 2]. We remind the reader some essential notions and establish some novel properties relevant to our work.

The *image of the absolute conic* (IAC) matrix satisfies $\boldsymbol{\omega} = \mathbf{K}^{-\top}\mathbf{K}^{-1}$, where $\mathbf{K}$ is the calibration matrix [1, §5.1] that encodes the internal camera parameters, which is, in its more general form:

$$\mathbf{K} \equiv \begin{pmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{2}$$

where α_u, α_v represent the focal length in terms of pixel dimensions in the u, v direction respectively, (u_0, v_0) are the principal point pixel coordinates and γ is the skew factor.

2.2 Plane-Based Self-calibration Equations

Let P denote the *unknown* (Euclidean) world-to-image homography, mapping entities of π to their projections on the image plane $\tilde{\pi}$, and let H_j be the *known* inter-view homography, induced by π, from the current view to some view number j.

The (Regular) Plane-Based Self-calibration Equations. Rigorously, the ES of π is given in terms of its *imaged circular points* (ICP) $P(\mathtt{I}_\pm)$, whereas the circular points (CP) $\mathtt{I}_\pm$ are, by definition [1, pp. 52-53], conjugate complex points at infinity in π, common to all of its circles. In any Euclidean representation, the CP have canonical coordinates $\mathbf{e}_\pm \equiv \mathbf{e}_1 \pm i\mathbf{e}_2 = (1, \pm i, 0)^\top$, which are invariant under any 2D similarity S of π i.e., $\mathbf{e}_\pm \sim \mathtt{S}\mathbf{e}_\pm$, where $\mathtt{S} \in \mathbb{R}^{3\times3}$ is the matrix of S. In image representation, the coordinates of the ICP $P(\mathtt{I}_\pm)$, denoted by $\mathbf{x}_\pm \equiv \mathbf{x}_1 \pm i\mathbf{x}_2$, satisfy $\mathbf{x}_\pm \sim \mathtt{P}\mathtt{S}\mathbf{e}_\pm$, where $\mathtt{P} \in \mathbb{R}^{3\times3}$ is the matrix of P. Note that $\mathbf{x}_\pm$ only have four d.o.f., basically the eight d.o.f. of $\mathtt{P}$ minus the four d.o.f. of $\mathtt{S}$.

The ICP are, by projective invariance, on the vanishing line, common to all imaged circles, including the image $\boldsymbol{\omega}$ of the absolute conic [1, pp. 81-83] of the plane at infinity. The IAC $\boldsymbol{\omega}$ is the locus of all ICP (i.e., of all 3D planes) which entails that $\mathbf{x}_\pm^\top \boldsymbol{\omega} \mathbf{x}_\pm = 0$, or equivalently (see [1, p. 211] for more details):

$$\mathbf{x}_1^\top \boldsymbol{\omega} \mathbf{x}_2 = 0 \quad \text{and} \quad \mathbf{x}_1^\top \boldsymbol{\omega} \mathbf{x}_1 - \mathbf{x}_2^\top \boldsymbol{\omega} \mathbf{x}_2 = 0. \tag{3}$$

In view number j, the constraint is described by $\mathbf{x}_\pm^\top \mathtt{H}_j^\top \boldsymbol{\omega}_j \mathtt{H}_j \mathbf{x}_\pm = 0$, or:

$$\mathbf{x}_1^\top \mathtt{H}_j^\top \boldsymbol{\omega}_j \mathtt{H}_j \mathbf{x}_2 = 0 \quad \text{and} \quad \mathbf{x}_1^\top \mathtt{H}_j^\top \boldsymbol{\omega}_j \mathtt{H}_j \mathbf{x}_1 - \mathbf{x}_2^\top \mathtt{H}_j^\top \boldsymbol{\omega}_j \mathtt{H}_j \mathbf{x}_2 = 0, \tag{4}$$

where $\boldsymbol{\omega}_j$ is the matrix of the IAC in view number j and $\mathtt{H}_j$ is the matrix of H_j.

The Dual Plane-Based Self-calibration Equations. A (maybe) most intuitive parameterization of the ES can also be given in terms of any (Euclidean) world-to-image homography $P \circ S$, where S denotes an arbitrary 2D similarity. Indeed, by applying $(P \circ S)^{-1}$ to the image plane, we get an Euclidean reconstruction of π, $P \circ S$ being referred to as *rectifying homography*.

If we treat the ICP as a degenerate conic envelope i.e., as the assemblage of isotropic lines as tangents, we get a conic, referred to as the *image of the conic dual to circular points* (ICDCP) in [1, p.52], whose matrix is of the form:

$$\mathtt{C}^* \sim \mathbf{x}_-\mathbf{x}_+^\top + \mathbf{x}_+\mathbf{x}_-^\top \sim \mathtt{P}(\mathbf{e}_-\mathbf{e}_+^\top + \mathbf{e}_+\mathbf{e}_-^\top)\mathtt{P}^\top \sim \mathtt{PS}(\mathbf{e}_-\mathbf{e}_+^\top + \mathbf{e}_+\mathbf{e}_-^\top)\mathtt{S}^\top\mathtt{P}^\top, \tag{5}$$

where $\mathtt{S} \in \mathbb{R}^{3\times3}$ is the matrix of S. As $\mathbf{e}_-\mathbf{e}_+^\top + \mathbf{e}_+\mathbf{e}_-^\top \sim \mathrm{diag}(1,1,0)$, a rectifying homography can be obtained by the adequate factorization [1, pp.55-56] of $\mathtt{C}^*$ e.g., based on the singular value decomposition (SVD), with singular values $\sigma_1 \geq \sigma_2 > 0$ and $\sigma_3 = 0$:

$$\pm\mathtt{C}^* = \mathtt{U}\boldsymbol{\Sigma}\mathtt{U}^\top \equiv \mathtt{X}\,\mathrm{diag}(1,1,0)\mathtt{X}^\top = \begin{bmatrix}\mathbf{x}_1\ \mathbf{x}_2\end{bmatrix}\begin{bmatrix}\mathbf{x}_1\ \mathbf{x}_2\end{bmatrix}^\top. \tag{6}$$

Therefore, the ICP can be specified in the form of $\mathbf{x}_\pm = \mathtt{U}_{(:,1:2)}\sqrt{\boldsymbol{\Sigma}}_{(1:2,1:2)} \sim \mathtt{XS}\mathbf{e}_\pm$. Consequently, the constraints (4) can be put in the matrix form:

$$\begin{bmatrix}\mathbf{x}_1\ \mathbf{x}_2\end{bmatrix}^\top \mathtt{H}_j^\top \boldsymbol{\omega}_j \mathtt{H}_j \begin{bmatrix}\mathbf{x}_1\ \mathbf{x}_2\end{bmatrix} \sim \mathtt{I}_{2\times2}. \tag{7}$$

We now highlight an interesting decomposition of the ICDCP matrix $\mathtt{C}^*$. Basically, our aim is to put into equation the fact that the degenerate conic $\mathtt{C}^*$ consists of the two points at which the vanishing line $\mathbf{v}$ meets the IAC $\boldsymbol{\omega}$. Since the AC is a circle on the plane at infinity, these two points are the ICP.

Proposition 1. *The ICDCP matrix satisfies the following decomposition:*

$$\mathtt{C}^* \sim [\mathbf{v}]_\times \boldsymbol{\omega} [\mathbf{v}]_\times, \tag{8}$$

where $\boldsymbol{\omega}$ is the IAC matrix and $\mathbf{v}$ is the vanishing line vector.

Proof. Define $\Delta \equiv [\mathbf{v}]_\times \boldsymbol{\omega} [\mathbf{v}]_\times$. Clearly Δ is rank-2, so as a conic envelope, Δ consists of two distinct points $\mathbf{p}$, $\mathbf{q}$ i.e., $\Delta \sim \mathbf{pq}^\top + \mathbf{qp}^\top$. Let us show these are the ICP. On the one hand, we see that $\Delta\mathbf{v} = \mathbf{0}$ which implies that both $\mathbf{p}$, $\mathbf{q}$ are on the vanishing line $\mathbf{v}$. On the other hand, any line $\mathbf{w} \neq \mathbf{v}$, verifying $\mathbf{w}^\top\Delta\mathbf{w} = 0$, passes either through $\mathbf{p}$ or $\mathbf{q}$. Assume $\mathbf{w}$ contains $\mathbf{p}$: this entails that $\mathbf{v} \times \mathbf{w} \sim \mathbf{p}$ and so $\mathbf{p}^\top\boldsymbol{\omega}\mathbf{p} = 0$. As a result, since $\boldsymbol{\omega}$ is the locus of all ICP, $\mathbf{p}$ is one ICP of π and $\mathbf{q}$ its conjugate.

Minimal Parameterization. As explained above, the ICP can be specified from $\mathtt{C}^*$ in the form of $\mathbf{x}_1 \pm \mathbf{x}_2$, with $\left[\mathbf{x}_1\ \mathbf{x}_2\right] \equiv \mathtt{U}_{(:,1:2)}\sqrt{\boldsymbol{\Sigma}}_{(1:2,1:2)}$ obtained from (6).

In this work, we will need a formal expression of $\mathbf{x}_1$ and $\mathbf{x}_2$.

Proposition 2. *Vectors $\mathbf{x}_1$, $\mathbf{x}_2$ satisfying (6), and so (7), can be written in the form of:*

$$\left[\mathbf{x}_1\ \mathbf{x}_2\right] \sim \left[[\mathbf{v}]_\times \mathbf{e}_k\ \mu\,[\mathbf{v}]_\times \boldsymbol{\omega}\,[\mathbf{v}]_\times \mathbf{e}_k\right], \quad k \in \{1,2,3\}, \tag{9}$$

where $\mu \equiv \alpha_u\alpha_v / \left\|\mathtt{K}^\top \mathbf{v}\right\|$ and $\mathbf{e}_k$ is a canonical vector, as defined in (1).

The proof requires to remind the reader that the vanishing line can be written as $\mathbf{v} = \mathtt{K}^{-\top}\mathbf{n}$, where $\mathbf{n}$ is the unit normal to π in the camera frame. Let us also define the 'calibrated' ICDCP $\bar{\mathtt{C}}^* \equiv \xi\mathtt{K}^{-1}\mathtt{C}^*\mathtt{K}^{-\top}$, where ξ is a scalar such that $\bar{\mathtt{C}}^* = [\mathbf{n}]_\times^\top [\mathbf{n}]_\times$.

Proof. The singular values of $\bar{\mathtt{C}}^*$ are $\{1,1,0\}$ so $\operatorname{ran}(\bar{\mathtt{C}}^*) = \operatorname{ran}([\mathbf{n}]_\times)$ and $\operatorname{null}(\bar{\mathtt{C}}^*) = \operatorname{null}([\mathbf{n}]_\times)$. Thanks to the SVD theorem [12], we know that the matrix $\mathtt{W} \in \mathbb{R}^{3\times3}$, $\mathtt{W}\mathtt{W}^\top \sim \mathtt{I}_3$, such that $\bar{\mathtt{C}}^* \sim \mathtt{W}\operatorname{diag}(1,1,0)\mathtt{W}^\top$, has the properties that $\operatorname{ran}(\bar{\mathtt{C}}^*) = \operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2\}$ and $\operatorname{null}(\bar{\mathtt{C}}^*) = \operatorname{span}\{\mathbf{w}_3\}$. As a result, we can compute:

$$\mathbf{w}_1 = [\mathbf{n}]_\times \mathbf{e}_k, \quad \mathbf{w}_2 = [\mathbf{n}]_\times^2 \mathbf{e}_k, \quad \mathbf{w}_3 = \mathbf{n}, \tag{10}$$

where $\mathbf{w}_2 = \mathbf{w}_3 \times \mathbf{w}_1 = [\mathbf{w}_3]_\times \mathbf{w}_1$. Substituting $\mathtt{K}^\top\mathbf{v}$ to $\mathbf{n}$ into (10), after some normalizations, we obtain (9).

The proposed form (9) offers an obvious advantage of minimal parameterization of the self-calibration problem. Substituting (9) into (4), there are now seven d.o.f. instead of the nine in [4].

Link with Malis' Constraint [5]. Introducing $\bar{\mathtt{H}}_j \equiv \mathtt{K}_j^{-1}\mathtt{H}_j\mathtt{K}_j$, the 'calibrated' ICDCP, in the view number j, is $\bar{\mathtt{C}}_j^* \equiv [\mathbf{n}_j]_\times^\top [\mathbf{n}_j]_\times \sim \bar{\mathtt{H}}_j\,[\mathbf{n}]_\times^\top [\mathbf{n}]_\times \bar{\mathtt{H}}_j^\top$, where $\mathbf{n}_j$ is the unit normal to π in the camera frame number j. Interestingly enough, since the singular values of $\bar{\mathtt{C}}_j^*$ are $\{1,1,0\}$, those of $\bar{\mathtt{H}}_j\,[\mathbf{n}]_\times^\top$ are also $\{1,1,0\}$, up to a scale factor. This latter property is the theoretical foundation of the self-calibration constraints of [5].

2.3 Formulation of the Problem

Assume that the IAC is constant in the views i.e., $\boldsymbol{\omega} \sim \boldsymbol{\omega}_j$. Given N views, *i.e.* $(N-1)$ inter-view homographies H_j, $2 \le j \le N$, the *self-calibration problem* of a camera is that of solving the system consisting of two equations (3) and $2(N-1)$ equations (4) for the p d.o.f. in the IAC matrix plus q in the ICP vectors. This is a non-linear, possibly constrained, problem which has, until now, been solved using iterative methods. It requires initial values which is a critical issue, already mentioned in [4].

Because of the proposed form (9) of ICP, compared to [4], our problem modeling only exhibits seven unknowns instead of nine. However, there is no magic: With the proposed form, the equation (3), related to the key view, is implicitly satisfied, while, in [4], it is considered as an equation to be satisfied. We ask the question: do we have to consider (3) as a constraint or as an equation? Since no input data i.e., no estimated homography is involved in (3), there is no logical reason for this equation not to be exactly satisfied. Actually, the nine parameters of [4] are not independent and must satisfy the additional constraint (3). More generally, with regard to the estimation of the homography, from key view to some view number j, using feature correspondences, there is no logical reason for assigning any error to the positions of the (arbitrary) features in the key view.

As one can expected, there are no more than two constraints for the plane-based self-calibration problem, but several ways of expressing them.

Simplified Camera Model. We investigate now the minimal parameterization of ICP under the assumption of a simplified camera model. Let the calibration matrix be $\mathtt{K} = \mathrm{diag}(\alpha, \alpha, 1)$, where α represents the focal length in pixels. Let $\mathbf{v} \equiv (\cos\phi, \sin\phi, -\rho)^\top$, where ρ is the orthogonal distance from the principal point to the vanishing line in pixels. This means that (9) can also be written in the form of (7), with:

$$\begin{bmatrix} \mathbf{x}_1 \; \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} -\sqrt{\alpha^2+\rho^2}\sin\phi & \rho\cos\phi \\ \sqrt{\alpha^2+\rho^2}\cos\phi & \rho\sin\phi \\ 0 & 1 \end{bmatrix}. \tag{11}$$

3 Global Optimization Using Interval Analysis

3.1 Interval Analysis

Interval analysis (IA) is born about forty years ago [13]. Several good introductions to IA are available in [10, 9].

An interval is denoted by $\mathbf{x} = [\underline{x}, \overline{x}]$, where $\underline{x}$ and $\overline{x}$ are the lower bound and the upper bound of $\mathbf{x}$ respectively. Interval vectors are called boxes. If $\mathbf{x}$ and $\mathbf{y}$ are two intervals, then the four elementary operations are defined by $\mathbf{x}$ op $\mathbf{y} = \{x \text{ op } y \mid x \in \mathbf{x} \text{ and } y \in \mathbf{y}\}$ for op $\in \{+, -, \times, \div\}$. By composing these operations, we can compute an extension of the range of a function over an interval. For instance, if $f(x) = x(x-1)$, then an extension of f over $[-1, 1]$ is $\mathbf{f}([-1, 1]) = [-1, 1]([-1, 1] - 1) = [-1, 1][-2, 0] = [-2, 2]$, which necessarily contains the exact range $[-1/4, 2]$ of f.

3.2 IA-Based Global Optimization

The idea of using IA for global optimization has been investigated by many authors [10, 14], to cite a few. In recent years, IA-based global optimization has exhibited many successes in various domains. It has also been successfully applied to 3D self-calibration [11]. The problem is the following: Find the global minimum $f^\star$ of a smooth function f, $f^\star = \min\{f(x) \mid x \in \mathbf{D}\}$, as well as the set of points for which it is obtained, $\mathbf{X}^\star = \{x \in \mathbf{D} \mid f(x) = f^\star\}$, where $\mathbf{D}$ is a box. IA-based global optimization usually uses IA along with a branch and bound algorithm. Let $\mathbf{X}$ be the box representing the search region and $\mathcal{L}$ a list of boxes to be processed. The basic scheme of the method can be stated as follows:

1. Initialize $\mathcal{L}$ by placing the initial search region $\mathbf{X}_0$ in it.
2. While $\mathcal{L} \neq \varnothing$ do:
   a Remove a box $\mathbf{X}$ from $\mathcal{L}$.
   b Process $\mathbf{X}$ (rejecting, reducing, critical point existence, ...).
   c Subdivise $\mathbf{X}$ and insert the boxes derived from $\mathbf{X}$ onto $\mathcal{L}$.

Many details, such as stopping criteria or tolerances have been omitted here. We refer the reader to [10] for a complete description of the method.

3.3 Implementation

We give here pratical details about our implementation. We use the simplified camera model described in §2.3. The cost function we minimize is the sum of the two squared residuals of the equations (4) in which we use the simplified form of the ICP (11):

$$f(\alpha, \rho, \phi) = \sum_{j=2}^{N} (\mathbf{x}_1^T H_j^T \boldsymbol{\omega} H_j \mathbf{x}_1 - \mathbf{x}_2^T H_j^T \boldsymbol{\omega} H_j \mathbf{x}_2)^2 + (\mathbf{x}_1^T H_j^T \boldsymbol{\omega} H_j \mathbf{x}_2)^2.$$

We derived the symbolic expression of the residuals. At each function evaluation, the developped residuals are numerically evaluated. This choice seems to be a good compromise between the evaluation time and the quality of the function extension. We have implemented a C++ code based on the PROFIL/BIAS library[2].

4 Experimental Results

4.1 Synthetic Data

The experimental setup is the following: The world plane is a planar grid composed of 100 points, projected onto 720×576 images, adding a Gaussian noise with a standard deviation equal to σ pixels. In our simplified camera model, the principal point is fixed to the center of the image, aspect ratio is equal to 1 and skew is 0. The focal length α is set to 1024 pixels. The camera fixates the center of the plane from a varying distance of 1460 ± 570 pixels, from randomly generated orientations varying in $[10°, 70°]$ in the world plane X axis, by $\pm 30°$ in the Y axis and by $\pm 90°$ in the Z axis. The inter-view homographies are estimated using the normalized DLT method [1, chapter 4]. The homographies are then transformed such that the principal point coincides with the image frame origin and such that $\alpha \to \alpha/360$ and $\rho \to \rho/360$.

In order to assess the benefit of a global optimization method, we have minimized the cost function using an iterative method: We have performed tests with 5 images and $\sigma = 1$ pixel. For each test, the unknowns have been randomly initialized such that $\alpha = \alpha^\star \pm 30\%$, $\rho = \rho^\star \pm 30\%$ and $\phi = \phi^\star \pm 30\%$, where $(\alpha^\star, \rho^\star, \phi^\star)$ was the real global minimum. The method converged to the global minimum (within a 20% tolerance) in 38% of cases. The global optimization method we used found the solution in 100% of cases. In our experiments, we have taken an initial interval corresponding to $[300, 3000] \times [100, 12000] \times [0, 360]$. The initial interval has no effect on the accuracy

[2] `http://www.ti3.tu-harburg.de/Software/PROFILEnglisch.html`

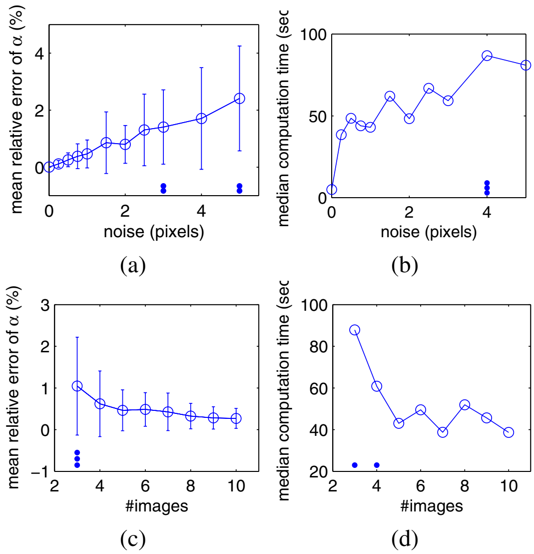

Fig. 1. Estimation error and computation time for varying noise level and varying image number. Each black point counts an excluded test (in (a) and (c)) or a failure (in (b) and (d)).

of the calculated minimum but only on the computation time. When a test cannot finish in a reasonable time, then we declare it as a *failure*. A few time-consuming tests will dramatically increase the mean computation time so we will present the median of computation time instead. Sometimes, the motion induced by the image sequence is closed to a degenerate configuration. In these cases, the global minimum is a bit far away from the real minimum. Therefore we exclude the tests such that the relative error is greater than 10% in the mean and standard deviation computations. The compuation times are given on a Pentium M 2GHz. Figure 1 gives some results about the estimation of α with respect to the noise level. Results for ρ and ϕ are very similar and are not presented here due to a lack of space. The accuracy remains quite good, even for σ=5 pixels. There are only a few excluded tests (there were 4 excluded tests over a total of 11000 tests). The estimation of α with respect to the image number is shown in the figure 1. The estimation error decreases when adding more images but the benefit is less relevant after 5-6 images. Figure 1 also shows the computation time with respect to noise and image number. In the first case, the presence of noise has a critical issue on the computation time whereas the noise level is less relevant. Indeed, the computation time does not dramatically increase with the noise level (the increasing looks linear). In the second case, the figure shows that after 5 images the computation time remains approximately the same: There are more terms to evaluate in the cost function whereas adding images makes the global minimum easier to enclose.

Since we have made hypotheses on τ, u_0 and v_0, we have assessed the tolerance of our method to a variation of these internal parameters. We first made the principal point varying randomly in a 50×50 pixels square. Second, we made τ varying in $[0.95, 1.05]$ (i.e. a variation of 5%). We used 5 images and $\sigma = 1$ pixel. The results are shown in table 4.1. Both varying principal point and varying τ have not critical consequences

Table 1. Estimation error and computation time for varying principal point and aspect ratio

Experiment	fixed (u_0, v_0), fixed τ	varying (u_0, v_0), fixed τ	fixed (u_0, v_0), varying τ
mean relative error of α (%)	0.5 ± 0.5	2.1 ± 1.8	2.7 ± 2.3
median computation time (sec)	43	92	83
number of failures	0	2	34

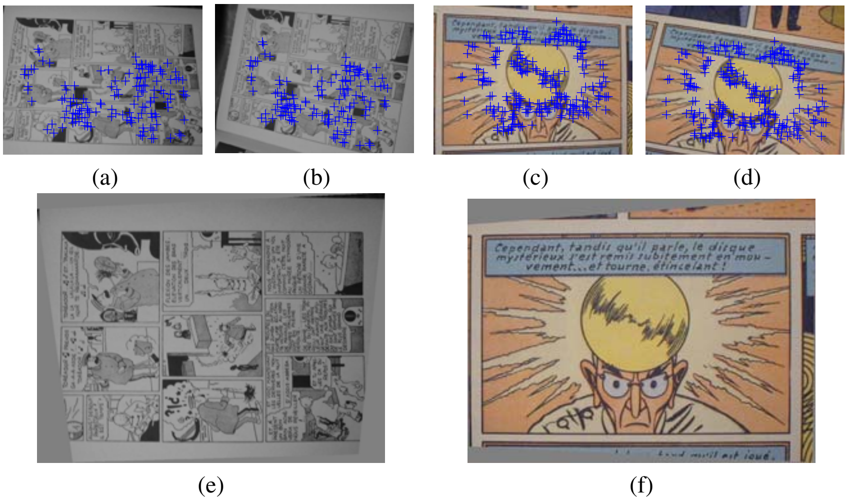

Fig. 2. Results for real images: (a) and (b) Key image and another image of the *tintin* sequence, along with the matched points. (e) The rectified image. (c), (d) and (f) Same results for *septimus*.

on the estimation of the parameters. However, we can see that varying τ has a critical impact on the number of failures. This is not a real limitation of the method since τ is usually very close to 1.

4.2 Real Data

We have tested the method on two sequences, *tintin* and *septimus*, composed of 7 and 4 640×480 images respectively. Homographies have been estimated by using the Kanatani optimal method [15], from automatically matched points. Then, we have applied a metric rectification of the key image. In the rectified image, the world plane should be parallel to the image plane and parallelism and angles should be recovered. Figure 2 shows two images of each sequence, including the matched points, and the rectified key image. We can see that the rectifications are quite good. The computation took 4 min 34 sec in the case of *tintin* and 8 min 14 sec in the case of *septimus*.

5 Conclusion

We proposed a minimal parameterization of the 2D self-calibration problem. Assuming the focal length is the only unknown, there are 3 parameters that can be estimated using a global minimization method, providing a guaranteed solution. A guaranteed solution to 3D self-calibration has been recently proposed by [11], for which closed-form solutions exist [3], contrary to 2D self-calibration. Although our constraint is more complex than Triggs' one, this not a real problem for the global optimization method we have used. Only the number of unknowns is relevant here. Our simplified camera model does not reveal some real limitation since we showed that the method is tolerant to consequent variations in the principal point position and since the hypothesis of a unit aspect ratio is quite realistic regarding recent digital cameras. The proposed method seems to work well provided the degeneracies of the problem are avoided. The study of such degeneracies is outside the scope of the current work, but reveals to be essential if we are aimed at using a minimal set of images [16]. We also aim at improving the performances of the algorithm we used.

Acknowledgement. The authors wish to acknowledge Frédéric Messine for fruitful discussions.

References

1. Hartley, R., Zisserman, A. In: Multiple View Geometry. Second ed. Cambridge University Press (2003)
2. Semple, J., Kneebone, G. In: Algebraic Projective Geometry. Oxford Classic Series, Clarendon Press (1952, reprinted, 1998)
3. Bougnoux, S.: From Projective to Euclidean Space under any Practical Situation, a Criticism of Self-Calibration. In: Proc. of the ICCV. Volume 1., Bombay, India (1998) 790–796
4. Triggs, B.: Autocalibration from Planar Scenes. In: Proc. of the ECCV. Volume 2., Freiburg, Germany (1998) 89–105
5. Malis, E., Cipolla, R.: Multi-view Constraints between Collineations: Application to Self-Calibration from Unknown Planar Structures. In: Proc. of the ECCV. Volume 2., Dublin, Ireland (2000) 610–624
6. Gurdjos, P., Sturm, P.: Methods and Geometry for Plane-Based Self-Calibration. In: Proc. of the CVPR. Volume 1., Madison, Wisconsin, USA (2003) 491–496
7. Jiang, G., Tsui, H., Quan, L.: Circular Motion Geometry Using Minimal Data. IEEE Trans. on PAMI **26** (2004) 721–731
8. Knight, J., Zisserman, A., Reid, I.: Linear Auto-Calibration for Ground Plane Motion. In: Proc. of the CVPR, Madison, Wisconsin, USA (2003) 503–510
9. Neumaier, A. In: Introduction to Numerical Analysis. Cambridge University Press (2001)
10. Hansen, E.R., Walster, G.W. In: Global Optimization Using Interval Analysis. Second ed. Marcel Dekker (2003)
11. Fusiello, A., Benedetti, A., Farenzena, M., Busti, A.: Globally Convergent Autocalibration Using Interval Analysis. IEEE Trans. on PAMI **26** (2004) 1633–1638
12. Golub, G., Loan, C.V. In: Matrix computations. Third ed. John Hopkins University Press (1996)
13. Moore, R.E. In: Interval Analysis. Prentice-Hall (1966)
14. Kearfott, R.B. In: Rigorous Global Search: Continous Problems. Kluwer Academic Publishers (1996)
15. Kanatani, K., Ohta, N.: Accuracy Bounds and Optimal Computation of Homography for Image Mosaicing Applications. In: Proc. of the ICCV. Volume 1., Kerkyra, Greece (1999) 73–78
16. Sturm, P., Maybank, S.J.: On plane-based camera calibration: A general algorithm, singularities, applications. In: Proc. of the CVPR. Volume 1., Fort Collins, Colorado, USA (1999) 432–437