On Constant Focal Length Self-Calibration From Multiple Views

Benoît Bocquillon* Adrien Bartoli† Pierre Gurdjos* Alain Crouzil*

Abstract

We investigate the problem of finding the metric structure of a general 3D scene viewed by a moving camera with square pixels and constant unknown focal length. While the problem has a concise and well-understood formulation in the stratified framework thanks to the absolute dual quadric, two open issues remain.

The first issue concerns the generic Critical Motion Sequences, i.e. camera motions for which self-calibration is ambiguous. Most of the previous work focuses on the varying focal length case. We provide a thorough study of the constant focal length case.

The second issue is to solve the nonlinear set of equations in four unknowns arising from the dual quadric formulation. Most of the previous work either does local nonlinear optimization, thereby requiring an initial solution, or linearizes the problem, which introduces artificial degeneracies, most of which likely to arise in practice. We use interval analysis to solve this problem. The resulting algorithm is guaranteed to find the solution and is not subject to artificial degeneracies. Directly using interval analysis usually results in computationally expensive algorithms. We propose a carefully chosen set of inclusion functions, making it possible to find the solution within few seconds.

Comparisons of the proposed algorithm with existing ones are reported for simulated and real data.

1. Introduction

Structure-from-Motion, the recovery of a metric reconstruction, i.e. cameras and points, from images is a fundamental computer vision problem. Its extensive study over the last few decades led to clear geometrical formulations and numerous solution methods. One of the key results is that a projective reconstruction can be computed from uncalibrated images, providing that neither the cameras nor the points lie on a critical surface. The projective reconstruction is equivalent to the sought after metric one up to an unknown upgrading homography of the projective space. Computing this homography from assumptions on the cameras is a *self-calibration* problem, and is equivalent to recovering the unknown intrinsic parameters of the cameras.

Three main approaches can be distinguished from the literature. *(i)* The Kruppa equations [2, 10], requiring pairwise epipolar geometry. *(ii)* The stratified approach, relying on upgrading a projective reconstruction to affine, and linearly solving for the affine-to-metric transformation. The former is solved using the modulus constraint [12] or exhaustive search [5]. *(iii)* Directly computing the projective-to-metric upgrading homography [7]. Triggs [21] proposed a convenient model with the so-called absolute dual quadric, encapsulating the plane at infinity and camera intrinsics in a compact manner. [11] drew on this model for proposing a linear algorithm dealing with the varying focal length case.

We tackle the self-calibration problem for a camera with constant intrinsic parameters in Triggs' absolute dual quadric framework. More precisely, we assume that the camera has square pixels and known principal point. Only the constant focal length thus remains to be computed. Our contributions deal with two important aspects of this problem.

First, section 3, we give a thorough study of the Critical Motion Sequences (CMS) which are generic for this problem. These are camera motions for which self-calibration is ambiguous, i.e. that defeat any self-calibration algorithm. Sturm [15] gave a complete classification of the CMS for constant, all unknown intrinsics. Note that CMS for which the calibration can be partly performed only lead to specific methods, see [6, chap. 19]. An example is the purely translational case for which the affine to metric upgrade is not uniquely defined. The case of a varying focal length with all other intrinsics known has been extensively studied [9, 13, 18], while, except for two cameras [19], the case of a constant focal length remained open.

Second, section 4, we propose a deterministic method for solving the nonlinear self-calibration problem which does not introduce artificial CMS and does not require an initial solution. Previous methods [11, 21] linearize the problem to find an initial solution and refine it through iterative nonlinear optimization. The basic step towards linearization is to neglect the rank deficiency of the dual absolute quadric. Appart from being suboptimal, linearizing the problem introduces artificial CMS, most of which likely to appear in

*UPS - IRIT, France, {bocquillon,gurdjos,crouzil}@irit.fr
†CNRS - LASMEA, France, Adrien.Bartoli@gmail.com

1-4244-1180-7/07/$25.00 ©2007 IEEE

practice. For instance, a fixating camera is not a generic CMS but defeats linear algorithms. Iterative nonlinear optimization is prone to falling in local minima, in particular if the initial solution lies "too far" from the optima, sought after one. Our algorithm is based on Interval Analysis Global Optimization. It finds the global minimum of the nonlinear error function over the four unknown parameters – the focal length and the plane at infinity. Computational time is reasonable, the order of which ranges from few seconds to a minute. Experimental results show that it scales linearly with the number of images while being almost unaffected by the level of noise on the data. We note that self-calibration based on Interval Analysis Global Optimization and the Kruppa equations is proposed in [3]. This approach is however different from ours, requiring hours to find the solution, and subject to important singularities related to the Kruppa equations.

Section 5 reports experimental evaluation and comparison with other algorithms. Section 6 concludes and discusses the paper. Next section reviews some background.

2. Background

In some metric coordinate frame, the camera projection matrices are written as $\mathtt{P}^i_M = \mathtt{K}^i\mathtt{R}^{i\top}\left(\mathtt{I}|-\mathbf{t}^i\right)$ with $i = 1, ..., n$, where $\mathtt{K}^i$ is the upper triangular calibration matrix which encodes the intrinsic parameters, $\mathtt{R}^i$ represents the orientation of the camera and $\mathbf{t}^i$ the camera centre in the world coordinate frame. The reconstruction problem is equivalent to finding a 4×4 rectifying homography $\mathtt{H}$ to upgrade the projective cameras $\mathtt{P}^i$ to $\mathtt{P}^i_M$, such that $\mathtt{P}^i_M = \mathtt{P}^i\mathtt{H}$ for $i = 1, ..., n$. The absolute dual quadric $\mathtt{Q}^*_\infty$ can be represented by a 4×4, semi-definite, rank-3 matrix which projects to the dual absolute conic $\boldsymbol{\omega}^* = \mathtt{K}\mathtt{K}^\top$:

$$\boldsymbol{\omega}^{*i} = \mathtt{P}^i\mathtt{Q}^*_\infty\mathtt{P}^{i\top}. \tag{1}$$

This quadric encodes both the absolute conic $\boldsymbol{\Omega}_\infty$ and the plane at infinity $\boldsymbol{\Pi}_\infty$. In a metric coordinate frame, $\mathtt{Q}^*_\infty$ has the form $\hat{\mathtt{I}} = \begin{pmatrix} \mathtt{I} & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{pmatrix}$ and in a projective frame we have $\mathtt{Q}^*_\infty = \mathtt{H}\hat{\mathtt{I}}\mathtt{H}^\top$. If intrinsic parameters are constant, $\boldsymbol{\omega}^{*i} \sim \boldsymbol{\omega}^{*j}$, where $\sim$ denotes equality up to a scale factor. For all known intrinsic parameters except the focal length α and writing $\mathtt{P}^1$ as $\mathtt{P}^1 = (\mathtt{I}|\mathbf{0})$, we have $\boldsymbol{\omega}^{*1} = \mathrm{diag}(\alpha^2, \alpha^2, 1)$. Each image pair $(1, j)$ gives us a set of five equations:

$$\begin{array}{ccc} \boldsymbol{\omega}^{*j}_{12} = 0 & \boldsymbol{\omega}^{*j}_{13} = 0 & \boldsymbol{\omega}^{*j}_{23} = 0 \\ \boldsymbol{\omega}^{*j}_{11} = \boldsymbol{\omega}^{*j}_{22} & \alpha^2\boldsymbol{\omega}^{*j}_{33} = \boldsymbol{\omega}^{*j}_{22} & \text{for } j = 2, ..., n. \end{array} \tag{2}$$

Solving for the 4 unknowns $\gamma = \alpha^2$ and $\boldsymbol{\Pi}_\infty$ is possible for $n \geq 2$ images. There are mainly three approaches: *(i) Linear method:* It consists in keeping the first four equations in (2), which are linear. *(ii) Quasi-linear method:* In [21], Triggs linearizes the problem by introducing additional unknowns and by solving for the entries of $\mathtt{Q}^*_\infty$ and $\boldsymbol{\omega}^{*j}$. *(iii) Nonlinear method:* This is the direct approach. Local iterative minimization methods are usually used. They require a good initialization. The first two methods have two major problems: First, the rank-3 constraint on $\mathtt{Q}^*_\infty$ is not automatically ensured (generally, this is imposed in a further step). This means that the obtained plane at infinity cannot be the supporting plane of the obtained absolute conic. Second, they do not ensure either the positive semi-definiteness condition on $\boldsymbol{\omega}^*$, preventing the Cholesky decomposition of this latter. Notice that in the two view case, the problem is equivalent to solve a quadratic equation.

In this work, we propose an algorithm based on the nonlinear formulation and on a global optimization method.

3. Generic Singularities

3.1. Previous Works

We distinguish the problems of recovering the intrinsics from those of recovering the affine and Euclidean structures, all three problems being usually packaged under the single term self-calibration. These problems carry theoretical singularities i.e., degenerate configurations at which problems have no solution or ambiguous solutions, which are generally due to (so-called critical) special camera motions. Theoretical singularities are generic in the sense that they cannot be overcome by *any* algorithm. In contrast, singularities introduced by algorithms are called artificial.

As suggested in [17], self-calibration algorithms may be classified into at least three groups, according to increasing levels of singularities. The first includes algorithms which only have the generic singularities. The second, which is that of Pollefeys [11] and Triggs' [21] linear algorithms, add singularities by not enforcing the sought-after dual absolute quadric to be rank-3. The last group – including algorithms based on Kruppa equations – add more singularities by not enforcing such quadrics to be the same for any pair of views. Remind that our aim is to design a self-calibration algorithm that falls into the first group i.e., whose implementation does not introduce additional/artificial singularities.

The first thorough study of critical motions for $n \geq 2$ views is due to Sturm in [16], for constant intrinsics. A generalisation progressively incorporating the assumptions of known skew, aspect ratio and principal point has been given by Kahl [8] even if the case of all known intrinsics except a varying focal length has also been studied by Sturm in [18]. Both authors have reported[1] the following (generic) critical motions for $n > 2$:

1. Optical axes are parallel i.e., cameras are translating.

[1] Even if the most complete analysis can be found in [18, §5].

2. Camera centres are aligned. All camera optical axes coincide except at, at most, two positions at which they can be arbitrarily oriented.
3. Camera centres move on two conics, one ellipse $\mathtt{E}$ and one hyperbola $\mathtt{H}$, lying on orthogonal planes that meet in the focal axis of $\mathtt{E}$, such that $\mathtt{E}$ and $\mathtt{H}$ have, as principal vertices, the foci of $\mathtt{H}$ and $\mathtt{E}$, respectively. Optical axes lie on the supporting planes and are tangent to the supported conic at each position.

Before going further, we set up some general background i.e., for any camera model.

In the sequel, the term *virtual conic* will refer to some proper purely imaginary conic i.e., whose real order-3 matrix is definite in its supporting plane. If $\mathbf{\Phi} \in \mathbb{R}^{3\times3}$ represents a conic, then $\mathbf{\Phi}^* \in \mathbb{R}^{3\times3}$ represents its dual (envelope) and $\mathtt{Q}^*_\Phi \in \mathbb{R}^{4\times4}$ represents the associated rank-3 (disk) quadric in 3-space [14].

$\mathcal{D} \subset \mathcal{K}$ will denote some subset of the group $\mathcal{K}$ of upper triangular order-3 matrices, determined by the constraints on intrinsics at our disposal. In accordance with [8, 18], a motion sequence M is *critical for the Euclidean reconstruction* (or simply *critical*) if there exists a virtual conic $\mathbf{\Phi}$ on some plane $\mathbf{\Pi}$ such that its dual image $\boldsymbol{\phi}^{*i} \sim \mathtt{P}^i\mathtt{Q}^*_\Phi\mathtt{P}^{i\top}$ satisfies:

$$\boldsymbol{\phi}^{*i} = \mathtt{D}^i\mathtt{D}^{i\top}, \quad \mathtt{D}^i \in \mathcal{D} \tag{3}$$

$$\text{and } \mathtt{Q}^*_\Phi \nsim \mathtt{Q}^*_\infty. \tag{4}$$

M is *critical for the affine reconstruction* if M is critical and $\mathbf{\Pi} \nsim \mathbf{\Pi}_\infty$. M is *critical for recovering the intrinsics* if M is critical and $\boldsymbol{\phi}^{*i} \nsim \boldsymbol{\omega}^{*i}$. Note that M is *critical for the Euclidean reconstruction* if it is critical for the affine reconstruction or for recovering the intrinsics.

3.2. Unknown *Constant* Focal Length

We now describe the generic singularities for $n \geq 2$ views of a camera, with all known intrinsics except a constant focal length α. We will refer to this case as the *unknown constant focal length* case or simply "our case".

For an unknown (possibly varying) focal length α, it has been shown [8, 18] that projection matrices can be considered as "calibrated" i.e., $\mathtt{P}^i = \mathtt{R}^{i\top}\left(\mathtt{I} \mid -\mathbf{t}^i\right)$, requiring that, in (3), $\mathcal{D}$ be the subgroup of diagonal matrices subject to condition (H1) below. Now, in our case, as α is supposedly constant, we will additionally require the relation (H2) to be satisfied:

- (H1) $D_{11} = D_{22}, \quad \mathtt{D} \in \mathcal{D}$;
- (H2) $\phi^{*i}_{11}/\phi^{*i}_{33} = \phi^{*j}_{22}/\phi^{*j}_{33}, \quad 1 \leq i, j \leq n.$

Clearly, critical motions for an unknown constant focal length belong to the set of critical motions for an unknown varying focal length as (H2) is simply an equivalence relation on $\mathcal{D}$.

Only virtual conics on finite planes have to be considered. For virtual conics on $\mathbf{\Pi}_\infty$, criticality is independent of camera positions and only depends on camera orientations [8, 18]. It has been shown that, under (H1), a motion is critical for the intrinsics but not for the affine structure if and only if all optical axes are parallel. Of course, it also holds under (H2).

Proposition 1 (Critical Camera Positions) *Let $\mathbf{\Phi}$ be a virtual conic on some* finite *plane $\mathbf{\Pi}$. In the unknown constant focal length case, a necessary condition for a motion sequence to be critical w.r.t. $\mathbf{\Phi}$ is that the camera centres be:*

- *(P1) at two different positions;*
- *(P2) at three/four distinct positions corresponding to the vertices of a right triangle/of a rectangle.*

Proof. As in [8, §8.3], we choose a Euclidean frame in which $\mathbf{\Phi}$ has, as supporting plane, $\mathbf{\Pi}$ with equation $z = 0$ and associated quadric $\mathtt{Q}^*_\Phi = \mathrm{diag}(d_1, d_2, 0, d_3)$, assuming $d_1 \geq d_2$ and $d_1, d_2, d_3 > 0$. Thus, $\mathtt{Q}^*_\Phi$ projects to:

$$\boldsymbol{\phi}^{i*} \sim \mathtt{P}^i\mathtt{Q}^*_\Phi\mathtt{P}^{i\top} = \mathtt{R}^{i\top}\mathbf{\Phi}^{*i}_\infty\mathtt{R}^i, \tag{5}$$
$$\text{where } \mathbf{\Phi}^{*i}_\infty = \mathrm{diag}(d_1, d_2, 0) + d_3\mathbf{t}^i\mathbf{t}^{i\top};$$

$\mathbf{\Phi}^i_\infty$ is the conic generated by intersecting $\mathbf{\Pi}_\infty$ with the ith projection cone of $\mathbf{\Phi}$, whose vertex is the camera centre $\mathbf{t}^i$.

Assume the position $\mathbf{t}^i$ to be critical w.r.t. $\mathbf{\Phi}$ i.e., equation (5) holds with (H1) and (H2) being satisfied, cf. §3.2. Hence, thanks to the spectral decomposition of $\mathbf{\Phi}^{*i}_\infty$, one infer from (H1) that two of its eigenvalues $\lambda^i_1, \lambda^i_2, \lambda^i_3$ are equal and, from (H2), the constraint $\lambda^i_1\lambda^i_2/(\lambda^i_3)^2 = \kappa^2$ for some fixed $\kappa > 0$. As said earlier, by (only) using the constraint resulting from (H1), the authors of [8, 18] obtain, as locus $\mathscr{L}_{(\mathrm{H1})}$ of critical positions, the union of two real central conics, one ellipse in the yz-plane and one hyperbola in the xz-plane. In the yz-plane, the obtained ellipse is represented by matrix $\mathtt{L}_1$, cf. (6); it is centered at the origin O with Oy and Oz as symmetry axes.

Now, only assume that (H2) holds. Let us determine what is the locus $\mathscr{L}_{(\mathrm{H2})}$ in the yz-plane, induced by applying the constraint resulting from (H2). We easily establish[2], that $\mathscr{L}_{(\mathrm{H2})}$ is a degenerate conic, with matrix $\mathtt{L}_2$, cf. (7), consisting of two distinct real parallel lines, being symmetric about the axis Oz.

$$\mathtt{L}_1 = \mathrm{diag}\left(d_1d_3, (d_1 - d_2)d_3, (d_2 - d_1)d_1\right), \tag{6}$$

$$\mathtt{L}_2 = \mathrm{diag}\left(0, d_2d_3, -d_1^2\kappa^2\right). \tag{7}$$

Since two conics meet at four points, we infer that the common points of $\mathtt{L}_1$ and $\mathtt{L}_2$ are, in our case, the four vertices of a rectangle, whose homogeneous yz-coordinates are:

$$[\pm\sqrt{(d_1 - d_2)(d_2 - d_1\kappa^2)}, \pm d_1\kappa, \sqrt{d_2d_3}]^\top. \tag{8}$$

[2] Due to lack of space, the details are omitted.

	Critical motions in the unknown constant focal length case	Self-calibration ambiguity
1	*All cameras* having parallel axes.	(A)
2	*Four cameras*, with centres defining a rectangle ; optical axes lie in the supporting plane of the rectangle and for each pair of cameras with adjacent centres, axes are symmetric about the symmetry axis of the rectangle separating the two centres.	(P)
3	*Three cameras*, out of the four cameras of case 2 (so defining a right triangle).	(P)
4	*Two cameras*, out of the four cameras of case 2, providing they have adjacent centres (optical axes intersect in a real finite point, with the centres being equidistant from this point).	(P)
5	*Two cameras* having coinciding axes.	(P)

Table 1. List of all critical motions for known skew, aspect ratio and principal point and unknown constant focal length α. Ambiguities are classified as (A) if $\mathbf{\Pi}_\infty$ can be recovered and (P) if neither α nor $\mathbf{\Pi}_\infty$ can be recovered.

As $d_1 \geq d_2$, all these are real providing $d_1 \geq d_2 \geq d_1\kappa^2$.

Similar results are obtained in the xz-plane, except that the four vertices must satisfy $d_1\kappa^2 \geq d_1 \geq d_2$ to be real. As the two above inequalities cannot simultaneously hold for non-zero coordinates, there are at most four (real) positions that can be critical.

A first special case (C1) must be considered if $d_1 = d_2$ i.e., if $\mathbf{\Phi}$ is a circle. In this case, $\mathtt{L}_1$ "degenerates"to the (repeated) Oz axis while $\mathtt{L}_2$ is unchanged: their intersection yields two distinct points, being symmetric about Oy. A second case (C2) is when $d_2/d_1 = \kappa^2$, for which (8) also reduces to the same two points. Similar results can be obtained in the xz-plane for exactly the *same* conditions. ■

3.2.1 Critical Camera Orientations

From each of the critical positions (P1) and (P2) listed in proposition 1, we now determine the critical camera orientations i.e., sufficient conditions for the motions to be critical.

Proposition 2 *In the case of all known intrinsics except a focal length, the critical orientations are the same whether the focal length be* varying *or* constant*.*

Proof. Only the "unsigned" directions of optical axes have to be considered. Indeed, if a camera orientation is critical, then it is also critical for any rotation around the optical axis or reversed direction. The study of critical motions in [8, 18] gives rise to two cases, according to whether $\mathbf{\Phi}$ is a circle, corresponding to the previously mentioned case (C1) in proof of proposition 1, or an ellipse. If $\mathbf{\Phi}$ is a circle, a direction is critical if and only if it is orthogonal to the supporting plane of $\mathbf{\Phi}$ for all $n - m$ cameras, with $m \in \{0, 1, 2\}$, and arbitrary for the m others. The configuration $m > 0$ happens when $\mathbf{\Phi}^{*i}_\infty \sim \mathtt{I}$, from which we infer that (H2) is also satisfied, cf. equation (5). If $\mathbf{\Phi}$ is an ellipse, the direction is critical if and only if it is parallel to tangents of any conic of $\mathscr{L}_{(\mathrm{H1})}$ lying on its supporting plane. Thus, the fact that directions are uniquely determined or in some configuration such that (H2) holds ends the proof. ■

By proposition 2, we inherit the results of [8, 18] and can adjust them to the unknown constant focal length case.

- Given critical positions (P1), critical orientations are:
  - either (O1a) arbitrary or
  - (O1b) such that optical axes coincide.
- Given (P2), they are (O2) such that the optical axes of any pair of adjacent cameras i.e., whose centres correspond to adjacent vertices of the rectangle, are symmetric about the symmetry axis of the rectangle separating their two centres while not being parallel; all four axes cannot intersect in a single real point.

Critical orientation (O1a) only applies to case (C2) given in the proof of proposition 1, which entails criticality only for affine structure.

3.2.2 Describing All Critical Camera Motions

All the critical motions in the unknown constant focal length case are listed in Table 1. Except for pure translations, the only motions that are likely to occur in practice are when centres correspond to only $p = 2$ different positions. At $2 < p \leq 4$ critical positions (coinciding with vertices of some right triangle or rectangle), some ambiguities can be removed by using cheirality invariants [6, chap. 21]. This is reassuring regarding our motivation to design a self-calibration algorithm with the fewest singularities.

4. A Guaranteed Solution Method

We describe how Interval Analysis (IA) and particularly IA-based Global Optimization finds the global optimum of nonlinear objective functions arising in computer vision problems. We do not give an extensive review of the background, which can be found in [4]. IA is an arithmetic of intervals such as $\mathbf{x} = [\underline{x}, \overline{x}]$, defined by two real bounds. One interest of IA is its ability of bounding the range of a function. An inclusion function $\mathbf{f}$ of a function f over a multidimensional interval $\mathbf{X} = ([\underline{x}_1, \overline{x}_1], ..., [\underline{x}_n, \overline{x}_n])$, called a box, is an interval $\mathbf{f} = [\underline{f}, \overline{f}]$ containing the range of f

over $\mathbf{X}$: $f(\mathbf{X}) \subseteq \mathbf{f}(\mathbf{X})$. By good "inclusion functions", we mean those with tight bounds around the range of f.

IA-based Global Optimization is a global deterministic guaranteed optimization method. From a general (nonlinear, non convex) objective function f to be minimized and from an initial box $\mathbf{X}_0$, it returns a box which encloses the global minimum of f, if it exists. It consists of: *(i)* A Branch and Bound algorithm to subdivide $\mathbf{X}_0$ into subboxes $\mathbf{X}_i$; *(ii)* An interval inclusion function $\mathbf{f}$ for any $\mathbf{X}_i$; *(iii)* Several tests to reject $\mathbf{X}_i$ if it does not contain the global minimum. Taking into account constraints in the parameter space is possible. Exhaustive examination of the whole solution space is a NP-hard problem. Therefore, good inclusion functions, allowing to reject boxes as early as possible in the optimization process, are required to achieve acceptable computational time. The latter varies from a few seconds to several days depending on $\mathbf{f}$.

Finding good inclusion functions is difficult. The difficulties arise from properties of IA such as: *(i) Subdistributivity:* If $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are three intervals, then $\mathbf{x} \times (\mathbf{y} + \mathbf{z}) \subseteq \mathbf{x} \times \mathbf{y} + \mathbf{x} \times \mathbf{z}$; *(ii) Variable dependencies:* $\mathbf{x} - \mathbf{x} \neq 0$. If each variable appears only once in an inclusion function, then the thighest bounds can be easily formulated. Unfortunately, this case is unlikely to happen in real problems. *(iii) Wrapping:* A box is always mapped into a box, introducing pessimism on the bounds.

One possible inclusion function for a rational function f is the so-called natural extension. This is obtained by replacing the real variables by intervals in f. It is known that this solution gives pessimistic bounds. A better solution consists in rewriting this natural extension by factorizing variables and by avoiding repetitions to improve the bounds. Better inclusion functions can be computed using decompositions such as Taylor expansion but the related implementation is less straightforward.

Computational time may be reduced with several tricks. One of them consists in reducing the size of the initial box by propagating constraints, scaling the unknowns or using analytical derivatives of the objective function. Two major packages are freely available: GlobSol[3] and ALIAS[4]. GlobSol only needs a well rewritten natural extension and automatically differentiates it whereas ALIAS requires analytical bounds for the function and possibly for its derivatives.

The use of IA-based Global Optimization in computer vision is somewhat marginal. The self-calibration problem (five unknowns), based on the Kruppa equations, is treated in [3]. Computational times over one hour for a few images are reported. In [1], a guaranteed solution for the plane-based self-calibration problem, using a simplified camera model (three unknowns), is obtained in less than one minute.

The objective function we use is the sum of squares of the residuals corresponding to equations (2). Four unknowns are involved, namely the focal length and the plane at infinity. The initial box is set to a large domain (see section 5). Note that we could reduce this initial box by propagating cheirality constraints [6, chap. 21]. As an inclusion function, we used the natural extension of the equations and applied automatic factorizations with the Maple software. Experience has shown that factorization has to be done first by the focal length, then by the other unknowns. As an example, the factorized expression of the residual $\alpha^2\boldsymbol{\omega}_{33}^{*j} - \boldsymbol{\omega}_{22}^{*j}$ is:

$$\begin{gathered}\alpha^2(\alpha^2(-(\mathrm{P}_{32}^j)^2 + p_1(-(\mathrm{P}_{34}^j)^2 p_1 + 2\mathrm{P}_{31}^j\mathrm{P}_{34}^j) + \\ p_2(-(\mathrm{P}_{34}^j)^2 p_2 + 2\mathrm{P}_{32}^j\mathrm{P}_{34}^j) - (\mathrm{P}_{31}^j)^2) + ((\mathrm{P}_{22}^j)^2 + \\ p_1((\mathrm{P}_{24}^j)^2 p_1 - 2\mathrm{P}_{21}^j\mathrm{P}_{24}^j) + p_2((\mathrm{P}_{24}^j)^2 p_2 - 2\mathrm{P}_{22}^j\mathrm{P}_{24}^j) + \\ (\mathrm{P}_{21}^j)^2 - (\mathrm{P}_{33}^j)^2 + p_3(-(\mathrm{P}_{34}^j)^2 p_3 + 2\mathrm{P}_{33}^j\mathrm{P}_{34}^j))) + \\ (\mathrm{P}_{23}^j)^2 + p_3((\mathrm{P}_{24}^j)^2 p_3 - 2\mathrm{P}_{23}^j\mathrm{P}_{24}^j),\end{gathered} \tag{9}$$

where $\boldsymbol{\Pi}_\infty = (p_1, p_2, p_3, 1)$. To assess the benefit of the rewriting, figure 1 shows the evaluation of two inclusion functions of this residual: the matricial form $\alpha^2\boldsymbol{\omega}_{33}^{*j} - \boldsymbol{\omega}_{22}^{*j}$ and the expression (9). In this figure, the inclusion functions are evaluated on boxes which enclose exactly the true plane at infinity on boxes which enclose exactly the true plane at infinity and whose intervals for the focal length are 100 pixels wide. The bounds obtained from the factorized inclusion function are significally tighter than the bounds without rewriting. This effect is amplified when we consider the square of the residuals and when the parameters of the plane at infinity are also considered as variables in the box. Our implementation uses the GlobSol package with automatic differentiation.

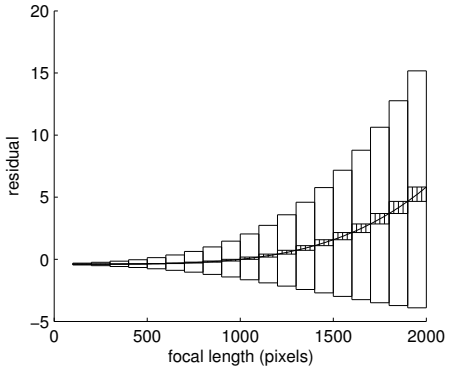

Figure 1. Examples of bounds of two inclusion functions of a residual used in the objective function, without factorization (empty boxes) and with factorization (hatched boxes).

[3] http://interval.louisiana.edu/GlobSol

[4] http://www-sop.inria.fr/coprin/logiciels/ALIAS

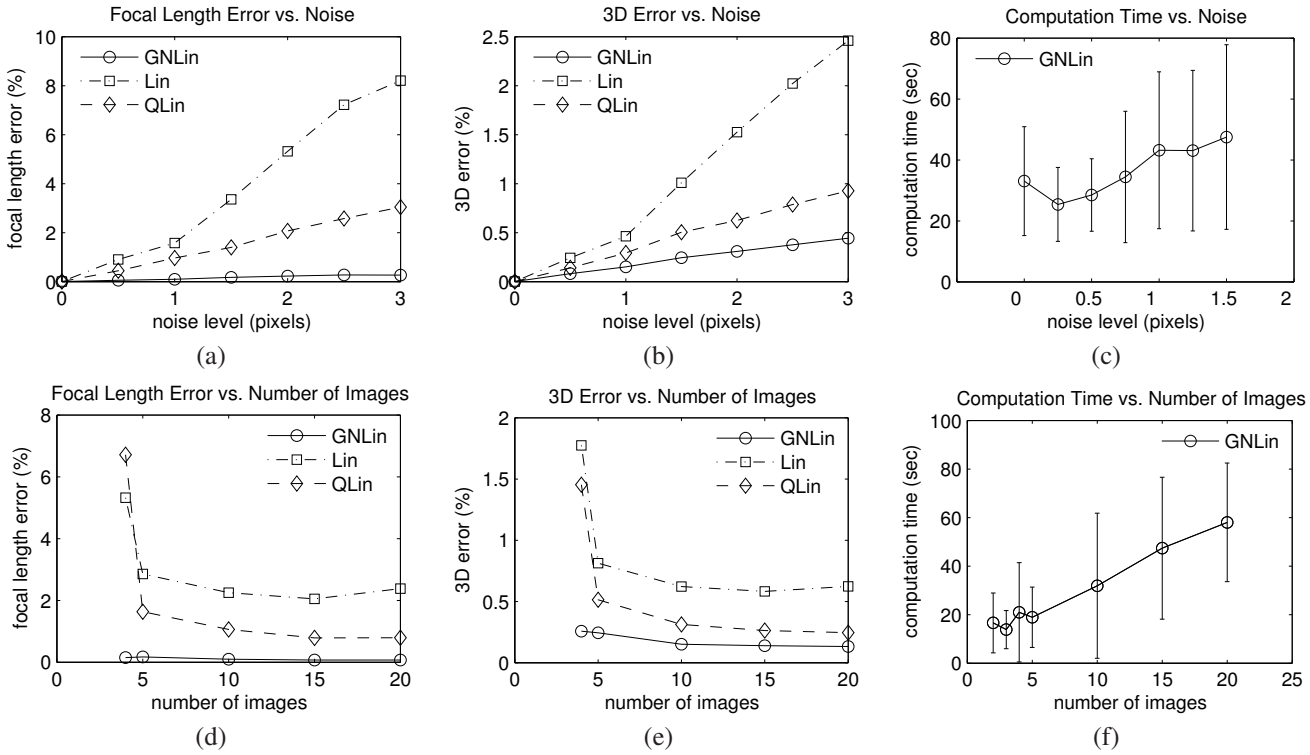

Figure 2. Mean relative error on the focal length (a), mean 3D error (b) and computation time (mean and standard deviation) (c) for varying noise level and 10 images. (d), (e), (f): Same criteria for varying number of images and a 1 pixel noise level. Squares stand for Lin, diamonds for QLin and circles for GNLin.

5. Results

Our approach has been validated on synthetic and real data. A software implementation is available online[5].

5.1. Synthetic Data

The experimental setup is composed of 100 points randomly distributed in a 3D sphere of unit radius. The 3D points are seen in n images with side length 256 pixels. Cameras have a fixed focal length of 1000 pixels, square pixels and principal point fixed at the image centre. They are randomly placed at a mean distance of 2.5 ± 0.25 units from the scene origin. Each camera fixates a random point located in a sphere of 0.1 unit radius centred at the origin. Gaussian noise is added to the 2D projected points. Projective bundle adjustment is used to compute noise contaminated cameras. Cameras are standardized so that the focal length is scaled around unity. We compared the linear method (Lin), the quasi-linear method (QLin), described in section 2, and our approach, the guaranteed nonlinear method (GNLin). The initial search box is set to $[100, 10000]$ for the focal length and $[-10^9, 10^9]^3$ for the plane at infinity. We measure the mean relative error on the focal length and the mean 3D error. The 3D error is the mean distance between the true 3D points and the recontructed 3D points, obtained from Euclidean rectification and alignment. It is expressed as a percentage of the scene size. All results are averages over 50 trials.

In the first experiment, n is fixed to 10 images and the noise level is varied from 0 to 3 pixels. As expected, the 3D error and the error on the focal length (figures 2.a and 2.b) are lower for GNLin than for Lin and QLin. The reason is that the rank and the positive semi-definiteness condition are implicitly enforced in GNLin. For all methods, the error increases linearly with the noise level. For clarity reasons, error bars are not displayed on the figures. The standard deviation on the focal length error, for a 3 pixel noise, are 5% for Lin, 2.5% for QLin and 0.3% for GNLin.

In the second experiment, the noise level is fixed to 1 pixel and n is varied from 4 to 20 images. Again, GNLin has the lowest error (figures 2.d and 2.e). We can also see that a few images are sufficient to obtain a good result. Although the computation time (figures 2.c and 2.f) required to find the global minimum is significantly larger than for the other methods (less than a second), it is reasonable. It increases linearly against the number of images while being weakly dependent on noise.

In order to see what happens near artificial degeneracies arising in Lin and GLin, we tested the following camera

[5] http://www.irit.fr/~Benoit.Bocquillon

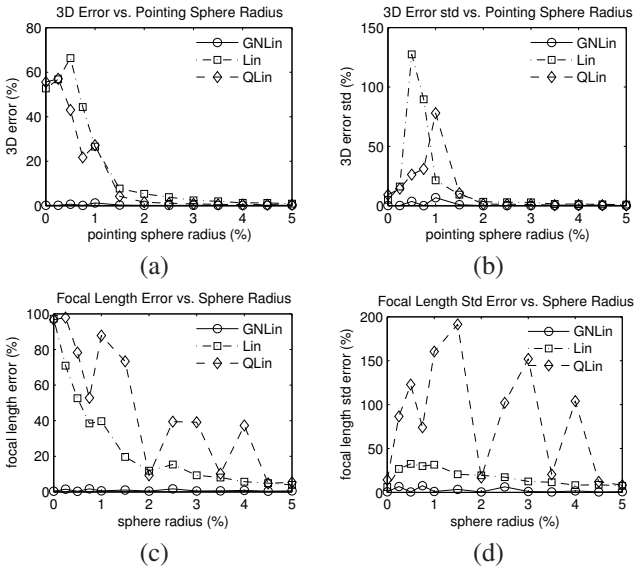

Figure 3. (a) and (b): Mean relative 3D error and standard deviation for a motion near the "fixating cameras" singularity. (c) and (d): Mean relative error on the focal length and standard deviation for a motion near the "aligned optical axes" singularity. All the distances are expressed as a percentage of the scene size. Squares stand for Lin, diamonds for QLin and circles for GNLin.

motions.

Fixating cameras. Cameras are placed as previously described except that all their optical centres lie on a plane. Each camera fixates a random point in a sphere of radius r centred at the origin. The singularity occurs at $r = 0$. Figures 3.a and 3.b show that the 3D error is important for $r < 2\%$ of the scene size. Our algorithm is not affected by the singularity, finding the correct answer in all cases.

Aligned optical axes. All the optical centres are aligned at different positions. All cameras look at the origin, except for one camera whose orientation is arbitrary. To move away from this singularity, cameras look in a sphere of radius r_1 and the optical centres are located in a sphere of radius r_2 centred on the critical position. Due to lack of space, we show results for $r_1 = r_2 = r$ only. Mean relative error on the focal length and standard deviation are reported in figures 3.c and 3.d. The error for Lin and QLin are quite important below 5% of the scene size. QLin is very unstable. GNLin is not affected by the singularity.

5.2. Real data

We took 4 images, shown on figure 4, by moving around a building, so that the motion happens to be near the "fixating camera", a critical motion for Lin and QLin. We semi-automatically detected and matched 63 interest points lying in the two dominant planes of the scene. To get the projective cameras, we used Sturm-Triggs projective factorization [20], followed by projective bundle adjustment. Lin and QLin failed in the sense that they gave meaningless solutions. GNLin converged within 17 seconds on a 1.7GHz Core Duo computer and found a 3604 pixels focal length. From this focal length and the retrieved plane at infinity, a Euclidean rectification was made on the 3D points obtained from the projective reconstruction. Figure 4.e shows a top view of the 3D point cloud. The ratio between the two wall lengths was evaluated to 1.36, obtained from real measurements, whereas our reconstruction has a ratio of 1.35. In the same way, the angle between the two reconstructed planes is equal to 91.7 degrees. We tried to add other images to the sequence to get off the singularity: Lin and QLin have remained very unstable, often failing or giving bad reconstructions in terms of distance ratio and angle (see figure 4.e). Therefore, these methods cannot be used in practice for this kind of motion. On the contrary, GNLin has always given an acceptable solution.

6. Conclusion and Future Work

This paper has addressed two issues associated with metric reconstruction in the case of all known intrinsic parameters except a constant focal length. First, we described the critical motion sequences for that case. Second, we proposed a self-calibration algorithm which does not introduce artificial singularities. It is based on Interval Analysis Global Optimization and is able to find the guaranteed solution within few seconds. We also noticed that the linear or quasi-linear methods are very unstable, even away from singular cases. These results show that such methods are sometimes difficult to use in practice.

Acknowledgment. The authors would like to thank Peter Sturm for very fruitful discussions.

References

[1] B. Bocquillon, P. Gurdjos, and A. Crouzil. Towards a guaranteed solution to plane-based self-calibration. In *ACCV*, volume 1, pages 11–20, Hyderabad, India, Jan. 2006.

[2] O. Faugeras, Q.-T. Luong, and S. J. Maybank. Camera Self-Calibration: Theory and Experiments. In *ECCV*, pages 321–334, Santa Margherita Ligure, Italy, May 1992.

[3] A. Fusiello, A. Benedetti, M. Farenzena, and A. Busti. Globally Convergent Autocalibration Using Interval Analysis. *PAMI*, 26(12):1633–1638, Dec. 2004.

[4] E. R. Hansen and G. W. Walster. *Global Optimization Using Interval Analysis*. Second ed. Marcel Dekker, 2003.

[5] R. Hartley, E. Hayman, L. de Agapito, and I. Reid. Camera calibration and the search for infinity. In *ICCV*, volume 1, pages 510–517, Kerkyra, Greece, Sept. 1999.

[6] R. Hartley and A. Zisserman. *Multiple View Geometry*. Second ed. Cambridge University Press, 2003.

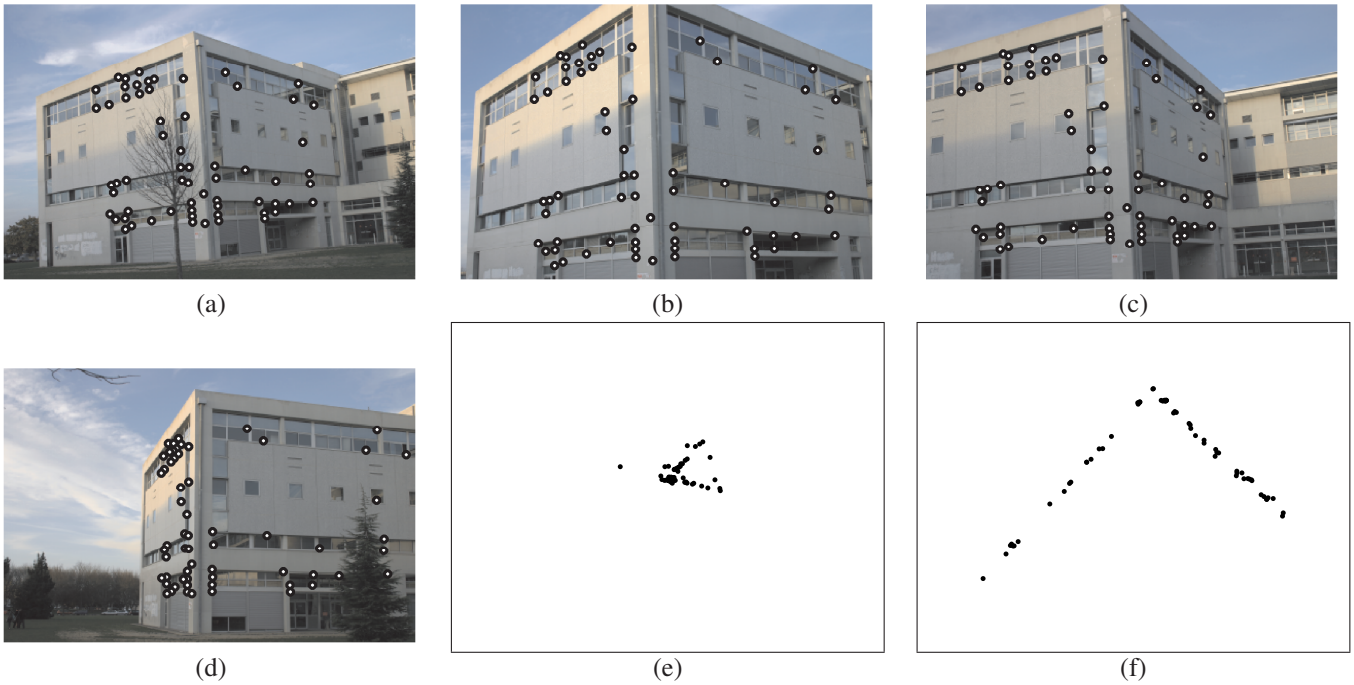

Figure 4. (a) to (d): Four images of a building with interest points. (e): An example of bad reconstruction, obtained with Lin. (f): A top view of the 63 3D points reconstructed with GNLin.

[7] A. Heyden and K. Åström. Euclidean Reconstruction from Constant Intrinsic Parameters. In *ICPR*, volume 1, pages 339–343, Vienna, Austria, Aug. 1996.

[8] F. Kahl. *Geometry and Critical Configurations of Multiple Views*. PhD thesis, Lund Institute of Technology, Sweden, 2001.

[9] F. Kahl and B. Triggs. Critical Motions in Euclidean Structure from Motion. In *CVPR*, volume 2, pages 2366–2372, Fort Collins, Colorado, USA, June 1999.

[10] Q.-T. Luong and O. Faugeras. Self-calibration of a moving camera from point correspondences and fundamental matrices. *IJCV*, 22(3):261–289, Mar. 1997.

[11] M. Pollefeys, R. Koch, and L. Van Gool. Self-Calibration and Metric Reconstruction in spite of Varying and Unknown Internal Camera Parameters. In *ICCV*, pages 90–95, Bombay, India, Jan. 1998.

[12] M. Pollefeys and L. Van Gool. Stratified Self-Calibration with the Modulus Constraint. *PAMI*, 21(8):707–724, Jan. 1999.

[13] M. Pollefeys and L. Van Gool. Some issues on self-calibration and critical motion sequences. In *ACCV*, pages 893–898, Taipei, Taiwan, Jan. 2000.

[14] J. Semple and G. Kneebone. *Algebraic Projective Geometry*. Oxford Classic Series, Clarendon Press, 1952, reprinted, 1998.

[15] P. Sturm. Critical Motion Sequences for Monocular Self-Calibration and Uncalibrated Euclidean Reconstruction. In *CVPR*, pages 1100–1105, San Juan, Puerto Rico, June 1997.

[16] P. Sturm. *Vision 3D non calibrée : contributions à la reconstruction projective et étude des mouvements critiques pour l'auto-calibrage*. PhD thesis, Institut National Polytechnique de Grenoble, France, Dec. 1997.

[17] P. Sturm. A case against Kruppa's equations for camera self-calibration. *PAMI*, 22(10):1199–1204, Sept. 2000.

[18] P. Sturm. Critical Motion Sequences for the Self-Calibration of Cameras and Stereo Systems with Variable Focal Length. *IVC*, 20(5-6):415–426, Mar. 2002.

[19] P. Sturm, Z. Cheng, P. C. Y. Chen, and A. N. Poo. Focal length calibration from two views: Method and analysis of singular cases. *CVIU*, 99(1):58–95, July 2005.

[20] B. Triggs. Factorization Methods for Projective Structure and Motion. In *CVPR*, pages 845–851, San Francisco, California, USA, June 1996.

[21] B. Triggs. Autocalibration and the absolute quadric. In *CVPR*, pages 609–614, San Juan, Puerto Rico, June 1997.