Stratified Autocalibration of Cameras with Euclidean Image Plane

Devesh Adlakha[1,3]
adlakha@unistra.fr
Adlane Habed[1]
habed@unistra.fr
Fabio Morbidi[2]
fabio.morbidi@u-picardie.fr
Cédric Demonceaux[3]
cedric.demonceaux@u-bourgogne.fr
Michel de Mathelin[1]
demathelin@unistra.fr

[1] ICube, CNRS, University of Strasbourg, France
[2] MIS, University of Picardie Jules Verne, France
[3] ImViA, VIBOT ERL CNRS, University of Burgundy Franche-Comté, France

Abstract

This paper tackles the problem of stratified autocalibration of a moving camera with Euclidean image plane (i.e. zero skew and unit aspect ratio) and constant intrinsic parameters. We show that with these assumptions, in addition to the polynomial derived from the so-called modulus constraint, each image pair provides a new quartic polynomial in the unknown plane at infinity. For three or more images, the plane at infinity estimation is stated as a constrained polynomial optimization problem that can efficiently be solved using Lasserre's hierarchy of semidefinite relaxations. The calibration parameters and thus a metric reconstruction are subsequently obtained by solving a system of linear equations. Synthetic data and real image experiments show that the new polynomial in our proposed algorithm leads to a more reliable performance than existing methods.

1 Introduction

Retrieving the camera calibration parameters from feature correspondences across images, i.e. camera autocalibration, is a prerequisite to recover the metric structure of an unknown scene imaged by uncalibrated perspective cameras. Autocalibration methods rely on some assumptions on the calibration parameters, such as constant [15, 20, 26, 33] or partially known intrinsic parameters [9, 17, 27, 28, 34]. When the images are captured by the same moving camera, its internal geometry remains unchanged in the absence of zooming and focusing. The sensor's aspect ratio and skew factor also remain quite stable despite a change in focus or zoom. Moreover, modern cameras commonly have square pixels, i.e. zero skew and unit aspect ratio. Such cameras are said to have a *Euclidean Image Plane* (EIP) [16].

The EIP assumption has often been exploited in direct autocalibration methods, which simultaneously estimate the plane at infinity (π_∞) and the intrinsic parameters. Direct methods rely mainly on either the *Dual Absolute Quadric* (DAQ) [33] or *Absolute Line Quadric* [28,

34] formulations that encode both π_∞ and the intrinsic parameters. A practical difficulty with these virtual quadrics is to enforce the nonlinear rank-3 constraint in their estimation. Linearization, on the other hand, results in artificial degeneracies [10]. Unlike direct methods, a stratified approach first tackles the more challenging problem of estimating π_∞. Once π_∞ is located, the intrinsic parameters can be retrieved by solving linear equations for the *Dual Image of the Absolute Conic* (DIAC) [13]. The advantage of this approach over a direct quadric-based one is that the nonlinearity is confined to fewer unknowns and a rank condition is not required. Most stratified methods assume constant intrinsic parameters [1, 6, 11, 26, 35] and rely on the polynomial derived from the modulus constraint [26]. The zero skew assumption was exploited in [11, 35] using the *Infinite Cayley Transform* (ICT) [35] to derive quartic polynomials in π_∞ for image triplets. However, using triplets introduces several unknown projective scale factors that render these polynomials not very practical to use. To the best of our knowledge, the assumption of a camera with EIP and constant intrinsic parameters has not been exploited so far in stratified autocalibration.

The autocalibration problem is inherently nonlinear and methods have traditionally relied on local optimization to obtain the calibration parameters. More recent work has also investigated globally optimal optimization approaches, either based on Branch-and-Bound algorithms [2, 6, 8, 12] or on polynomial optimization [5] using Lasserre's hierarchy of semidefinite relaxations [14, 19]. In [5], Lasserre's hierarchy has been used to estimate the DAQ under rank, semi-definiteness, and chirality [13] constraints.

In this paper, we present a stratified autocalibration method for a moving camera with EIP and constant intrinsic parameters. Our key contribution is in the formulation of a new quartic polynomial in the unknown π_∞, that, in addition to the polynomial from the modulus constraint, is obtained for each image pair with the assumed camera model. This polynomial is derived using a yet unexploited property of the ICT. For three or more images, estimating π_∞ is stated as a constrained polynomial optimization problem that is solved using Lasserre's hierarchy. The estimated π_∞ is refined using local optimization of a normalized cost and the intrinsic parameters are recovered subsequently by estimating the DIAC. Experiments with synthetic data and real images show that the new polynomial in our proposed algorithm leads to a more reliable performance than existing methods, especially for short sequences.

2 Background and related work

We consider a scene embedded in a projective 3-space and imaged n times by a moving perspective camera. The 3×4 uncalibrated projection matrices [13] P_i are of the form, $\mathsf{P}_i = [\ \mathsf{A}_i \mid \mathsf{a}_i\ ],\ \ i = 1,2,\ldots,n$, where A_i is a 3×3 matrix and a_i is a 3-vector. The world frame is attached to camera 1 such that $\mathsf{A}_1 = \mathsf{I}$ (identity) and $\mathsf{a}_1 = 0$.

Inter-image homography: An inter-image homography maps image projections of coplanar scene points from one image to another. It was shown in [11] that given projection matrices P_i, all inter-image homographies induced by planes not containing the origin of the world frame are linear functions of a real 3-vector π,

$$\mathsf{H}_{ij} = \mathsf{A}_j\mathsf{A}_i^* - \mathsf{A}_j[\ \pi\ ]_\times \mathsf{A}_i^\mathsf{T}[\ \mathsf{a}_i\ ]_\times^\mathsf{T} - \mathsf{a}_j\pi^\mathsf{T}\mathsf{A}_i^* \ \text{ for all } \ i\neq j, \tag{1}$$

where A^* is the adjoint matrix of A (the transpose of the cofactor matrix) and $[\ \pi\ ]_\times$ denotes the skew-symmetric matrix associated with vector π. The more usual forward, $\mathsf{H}_{1i} = \mathsf{A}_i - \mathsf{a}_i\pi^\mathsf{T}$, and inverse, $\mathsf{H}_{i1} = \mathsf{A}_i^* - [\ \pi\ ]_\times\mathsf{A}_i^\mathsf{T}[\ \mathsf{a}_i\ ]_\times^\mathsf{T}$, mappings relating any view i and the reference

image can be extracted from (1). Note that,

$$\mathsf{H}_{i1} = \mathsf{H}_{1i}^{*} = \det(\mathsf{H}_{1i})\mathsf{H}_{1i}^{-1}, \tag{2}$$

where $\det(\mathsf{H}_{1i})$ is the determinant of H_{1i}, and that H_{ij} in (1) is obtained through $\mathsf{H}_{ij} = \mathsf{H}_{1j}\mathsf{H}_{i1}$.

Infinite homography: For a fixed π, all H_{ij} represent inter-image homographies induced by the same plane. In particular, for some appropriate $\pi = \pi_\infty$, the inter-image homograhies expressed by (1) and denoted hereafter by $\mathsf{H}_{\infty ij}$, are those induced by the plane at infinity. Such a homography $\mathsf{H}_{\infty ij}$, referred to as an infinite homography, is distinctively independent of the camera translation. In particular, when the camera intrinsic parameters are constant,

$$\mathsf{H}_{\infty ij} = \mathsf{A}_j\mathsf{A}_i^{*} - \mathsf{A}_j[\,\pi_\infty\,]_\times \mathsf{A}_i^{\mathsf{T}}[\,\mathsf{a}_i\,]_\times^{\mathsf{T}} - \mathsf{a}_j\pi_\infty^{\mathsf{T}}\mathsf{A}_i^{*}, \tag{3a}$$

$$= \lambda_i^2\lambda_j \mathsf{K}\mathsf{R}_{ij}\mathsf{K}^{-1}, \tag{3b}$$

for all $i \neq j$, where R_{ij} is the rotation matrix relating cameras i and j. The scalar λ_i is such that $\mathsf{H}_{\infty 1i} = \lambda_i \mathsf{K}\mathsf{R}_{1i}\mathsf{K}^{-1}$ with $\lambda_1 = 1$. Matrix K is the intrinsic parameters matrix of the form,

$$\mathsf{K} = \begin{bmatrix} f_x & \gamma & u \\ 0 & f_y & v \\ 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

encapsulating the focal lengths (f_x, f_y), the principal point coordinates (u, v), and the skew γ.

Modulus constraint: When the camera intrinsic parameters are constant, $\mathsf{H}_{\infty ij}$ is similar to a scaled rotation matrix (3b). Since a rotation matrix has eigenvalues with unit modulus, those of $\mathsf{H}_{\infty ij}$ necessarily have equal moduli. A necessary condition on π_∞ (the so-called modulus constraint) for a given $\mathsf{H}_{\infty ij}$ to carry this property was derived and exploited in [26]. For any two views i and j, this condition involves the coefficients of the characteristic polynomial,

$$\det\big(\mathsf{H}_{\infty 1j} - \lambda\mathsf{H}_{\infty 1i}\big) = -c_i(\pi_\infty)\lambda^3 + t_{ij}(\pi_\infty)\lambda^2 - t_{ji}(\pi_\infty)\lambda + c_j(\pi_\infty) = 0, \tag{5}$$

where c_i and t_{ij} are affine functions of π_∞. For $\mathsf{H}_{\infty ij}$ to satisfy the modulus constraint, it was shown in [26] that π_∞ must satisfy the following quartic polynomial equation,

$$m_{ij}(\pi_\infty) = c_i(\pi_\infty)t_{ji}^3(\pi_\infty) - c_j(\pi_\infty)t_{ij}^3(\pi_\infty) = 0 \;\text{ for all }\; i \neq j. \tag{6}$$

This polynomial can also be derived from (3a) and (3b) by noticing that, on the one hand,

$$c_i(\pi_\infty) = \det(\mathsf{H}_{\infty 1i}),\; i = 1,2,\ldots,n \;\text{ and }\; t_{ij}(\pi_\infty) = \mathrm{tr}(\mathsf{H}_{\infty ij}) \;\text{ for all }\; i \neq j, \tag{7}$$

and, on the other hand, that these are related to the scaling of the infinite homographies,

$$\det(\mathsf{H}_{\infty 1i}) = \lambda_i^3, \qquad \mathrm{tr}(\mathsf{H}_{\infty ij})/\mathrm{tr}(\mathsf{H}_{\infty ji}) = \lambda_i/\lambda_j, \tag{8}$$

where $\mathrm{tr}(\mathsf{H}_{\infty ij})$ denotes the trace of $\mathsf{H}_{\infty ij}$. Using (6), a finite number of candidate solutions for π_∞ can be obtained when at least 3 such polynomials (i.e. as many images) are available.

Infinite Cayley Transform: The matrix $\mathsf{Q}_{\infty ij}$, defined as,

$$\mathsf{Q}_{\infty ij} = \lambda_j\mathsf{H}_{\infty ij} - \lambda_i\mathsf{H}_{\infty ji} = \lambda_i^2\lambda_j^2\,\mathsf{K}[\,\mathsf{r}_{ij}\,]_\times\mathsf{K}^{-1}, \tag{9}$$

was introduced in [11] and [35] in the context of stratified autocalibration. This matrix, referred to as the *Infinite Cayley Transform* (ICT) in [35], is similar to the skew-symmetric

matrix $[\,\mathsf{r}_{ij}\,]_\times = \mathsf{R}_{ij} - \mathsf{R}_{ij}^\mathsf{T}$. As such, it carries interesting properties that allow the derivation of constraints on π_∞ that are complementary to the modulus constraint (6). For instance,

$$\mathrm{tr}(\mathsf{Q}^*_{\infty ij}) > 0, \tag{10}$$

combined with the modulus constraint, are necessary and sufficient conditions for $\mathsf{Q}_{\infty ij}$ to be similar to a skew-symmetric matrix [35]. Note that, using (3a) and (8), inequality (10) is a polynomial in π_∞. Furthermore, in [11] the authors showed that for cameras with zero skew i.e. $\gamma = 0$, the coordinates (u, v) of the principal point can be expressed as follows,

$$u = (\mathsf{Q}_{\infty ij})_{11}/(\mathsf{Q}_{\infty ij})_{31}, \quad v = (\mathsf{Q}_{\infty ij})_{22}/(\mathsf{Q}_{\infty ij})_{32}, \tag{11}$$

where $(\cdot)_{hk}$ is the element in the hth row and kth column of a matrix. New polynomials in π_∞, enforcing the constancy of (u, v) across images, were derived from image triplets in [11].

Polynomial optimization: Consider the, generally nonconvex, optimization problem,

$$\begin{aligned} \min_{\mathsf{x}} \quad & f(\mathsf{x}) \\ \text{s.t.} \quad & g_i(\mathsf{x}) \geq 0, \ i = 1, 2, \ldots, \ell, \end{aligned} \tag{12}$$

where $f(\mathsf{x})$ and all $g_i(\mathsf{x})$ are multivariate scalar polynomials in x, an m-vector. Lasserre, in [19], has shown that (12) can be solved through a hierarchy of convex LMI (Linear Matrix Inequality) relaxations of increasing order $d = 1, 2, \ldots,$ yielding monotonically non-increasing lower bounds on the original problem and converging to its global minimum. At each order d, the problem is linearized and a surrogate *Semidefinite Program* (SDP) is solved. Linearization is possible at a starting relaxation order d, in which no monomial in the problem is of a degree higher than $2d$. Lasserre's method, implemented in GloptiPoly [14], has been used to solve several polynomial optimization problems in computer vision [4, 18, 22]. The interested reader may refer to [14, 19] for more details on this method.

3 EIP-based polynomial constraint

We consider a camera with a Euclidean image plane, EIP (i.e. zero skew, $\gamma = 0$, and unit aspect ratio, $f_x/f_y = 1$) whose focal length and principal point coordinates are constant but unknown. We first show that the Infinite Cayley Transform (ICT) satisfies a yet unexploited property under these assumptions (Proposition 1). Using this property, we then derive a new quartic polynomial constraint on π_∞ (Proposition 2).

Definition 1. *Given a* 3×3 *matrix* B*, we define the matrix operator* $\Phi(\cdot)$ *as,*

$$\Phi(\mathsf{B}) = (\mathsf{B}^* \circ \mathsf{B})_{31} + (\mathsf{B}^* \circ \mathsf{B})_{32}, \tag{13}$$

where $\circ$ *denotes the Hadamard (elementwise) product, i.e.* $(\mathsf{B} \circ \mathsf{C})_{hk} = (\mathsf{B})_{hk}(\mathsf{C})_{hk}$ *for any two matrices* B *and* C *of the same dimensions.*

Proposition 1. *Consider two images* i *and* j *captured by a moving camera with EIP and constant intrinsic parameters. The ICT* $\mathsf{Q}_{\infty ij}$ *of these images satisfies,*

$$\Phi(\mathsf{Q}_{\infty ij}) = 0. \tag{14}$$

Proof. The ICT $\mathsf{Q}_{\infty ij}$, as given by (9), is similar to a skew-symmetric matrix. As such, the matrix $\mathsf{K}^{-1}\mathsf{Q}_{\infty ij}\mathsf{K}$ is skew-symmetric. Given a camera with EIP, the 2×2 matrix,

$$\begin{bmatrix} (\mathsf{Q}_{\infty ij})_{11} - u(\mathsf{Q}_{\infty ij})_{31} & (\mathsf{Q}_{\infty ij})_{12} - u(\mathsf{Q}_{\infty ij})_{32} \\ (\mathsf{Q}_{\infty ij})_{21} - v(\mathsf{Q}_{\infty ij})_{31} & (\mathsf{Q}_{\infty ij})_{22} - v(\mathsf{Q}_{\infty ij})_{32} \end{bmatrix}, \tag{15}$$

obtained by eliminating the 3rd row and 3rd column of $\mathsf{K}^{-1}\mathsf{Q}_{\infty ij}\mathsf{K}$ is also skew-symmetric. Enforcing the diagonal entries of (15) to be zero leads to the expressions of u and v given in (11), as obtained in [□] for zero-skew cameras. Furthermore, the sum of the off-diagonal elements of (15) also being zero yields: $(\mathsf{Q}_{\infty ij})_{12} + (\mathsf{Q}_{\infty ij})_{21} - u(\mathsf{Q}_{\infty ij})_{32} - v(\mathsf{Q}_{\infty ij})_{31} = 0$. Substituting the expressions of u and v in (11) in this equation leads exactly to (14). □

Expressing the ICT in terms of π_∞ using (9) and (3a), we observe that $\Phi(\mathsf{Q}_{\infty ij})$ expands as,

$$\Phi(\mathsf{Q}_{\infty ij}) = a_{ij}(\pi_\infty)\lambda_j^3 - b_{ij}(\pi_\infty)\lambda_i\lambda_j^2 + b_{ji}(\pi_\infty)\lambda_i^2\lambda_j - a_{ji}(\pi_\infty)\lambda_i^3, \tag{16}$$

where the coefficients a_{ij} and b_{ij}, for any combination of i and j, are cubic polynomials in π_∞. Substituting for λ_i and λ_j using (8) leads to a polynomial of degree 6 at best. We show, however, that equation (14) can be used to derive a quartic polynomial in π_∞. Although a_{ij}, b_{ij} are fully defined through expansion (16), it is interesting to note that $a_{ij}(\pi_\infty) = \Phi(\mathsf{H}_{\infty ij})$ and $a_{ji}(\pi_\infty) = \Phi(\mathsf{H}_{\infty ji})$. As such, we show that the terms involving $a_{ij}(\pi_\infty)$ and $a_{ji}(\pi_\infty)$ can be eliminated from (16), so long as the modulus constraint is satisfied. To show this, we recall here some properties of the Hadamard product and adjoint matrices.

Property 1. *Let* B *and* C *be two* 3×3 *matrices and* λ *a scalar. We have,*

$$\mathsf{B}\circ\mathsf{C} = \mathsf{C}\circ\mathsf{B}, \tag{17a}$$
$$(\lambda\mathsf{B})\circ\mathsf{C} = \mathsf{B}\circ(\lambda\mathsf{C}) = \lambda(\mathsf{B}\circ\mathsf{C}), \tag{17b}$$
$$(\mathsf{BC})^* = \mathsf{C}^*\mathsf{B}^*, \tag{18a}$$
$$(\mathsf{B}^*)^* = \det(\mathsf{B})\mathsf{B}. \tag{18b}$$

Using the adjoint matrix properties, one can deduce that,

$$\mathsf{H}^*_{\infty ij} = \lambda_i^3\,\mathsf{H}_{\infty ji}. \tag{19}$$

This is because, with property (18a), $\mathsf{H}^*_{\infty ij} = (\mathsf{H}_{\infty 1j}\mathsf{H}_{\infty i1})^* = \mathsf{H}^*_{\infty i1}\mathsf{H}^*_{\infty 1j}$. In addition, using (2), $\mathsf{H}^*_{\infty ij} = (\mathsf{H}^*_{\infty 1i})^*\mathsf{H}_{\infty j1}$. Now, property (18b) leads to (19). Note that the same procedure applies to obtain, $\mathsf{H}^*_{\infty ji} = \lambda_j^3\,\mathsf{H}_{\infty ij}$. We can deduce, using (19) and property (17b), that $\mathsf{H}^*_{\infty ij}\circ\mathsf{H}_{\infty ij} = \lambda_i^3(\mathsf{H}_{\infty ji}\circ\mathsf{H}_{\infty ij})$ and also that $\mathsf{H}^*_{\infty ji}\circ\mathsf{H}_{\infty ji} = \lambda_j^3(\mathsf{H}_{\infty ij}\circ\mathsf{H}_{\infty ji})$. It must be clear now, with property (17a), that $\lambda_j^3(\mathsf{H}^*_{\infty ij}\circ\mathsf{H}_{\infty ij}) = \lambda_i^3(\mathsf{H}^*_{\infty ji}\circ\mathsf{H}_{\infty ji})$ and, as a consequence, $\lambda_j^3\,a_{ij}(\pi_\infty) = \lambda_i^3\,a_{ji}(\pi_\infty)$. This constitutes the proof to our main proposition, which we state below.

Proposition 2. *Consider two images* i *and* j *captured by a moving camera with EIP and constant intrinsic parameters. The plane at infinity* π_∞ *satisfies the quartic polynomial equation,*

$$p_{ij}(\pi_\infty) = -b_{ij}(\pi_\infty)t_{ji}(\pi_\infty) + b_{ji}(\pi_\infty)t_{ij}(\pi_\infty) = 0, \tag{20}$$

for all $i \neq j$*. The expressions* b_{ij} *and* b_{ji} *are cubic polynomials in* π_∞ *defined by the expansion in* (16) *while* t_{ij} *and* t_{ji} *are linear functions of* π_∞ *defined by* (5)*.*

We refer to the polynomial p_{ij} as the EIP polynomial. It is obtained from (14) by substituting λ_i and λ_j in (16) using (8) and eliminating the a_{ij} and a_{ji} terms.

4 Stratified autocalibration

Our stratified method relies on polynomial optimization using Lasserre's hierarchy to estimate π_∞. The modulus constraint m_{ij} (6) and the EIP polynomial p_{ij} (20) are used to define a suitable cost function. Since these two polynomials are obtained for each image pair, our method can be used with three or more images. Additional polynomial inequality constraints, such as those derived from the ICT (based on (10) and (11)) as well as Hartley's chirality inequalities [13], can be easily incorporated in Lasserre's method. These constraints are especially useful when few images are used as multiple solutions may persist. The estimated π_∞ is refined using local optimization to fit a normalized cost. The camera intrisics are subsequently recovered by solving a linear system of equations. The steps of our algorithm are summarized at the end of this section. We first detail the inequality constraints on π_∞, followed by a description of the polynomial and local optimization problems to obtain π_∞.

ICT-based inequalities: Inequality (10), a necessary condition for the ICT $\mathsf{Q}_{\infty ij}$ to be similar to a skew-symmetric matrix, can be expressed in terms of π_∞. This can be done in different ways. For instance, denoting by $\simeq$ the equality up to scale, one can consider (8) to obtain,

$$\mathsf{Q}_{\infty ij} \simeq \widetilde{\mathsf{Q}}_{\infty ij} = \mathrm{tr}(\mathsf{H}_{\infty ji})\mathsf{H}_{\infty ij} - \mathrm{tr}(\mathsf{H}_{\infty ij})\mathsf{H}_{\infty ji}, \tag{21}$$

and the constraint $q_{ij}(\pi_\infty) = \mathrm{tr}(\widetilde{\mathsf{Q}}^*_{\infty ij}) > 0$. Note that the sign of the unknown scale in (21) does not affect the sign of $q_{ij}(\pi_\infty)$ since it is squared in the adjoint matrix. One may also exploit (11) to restrict the principal point to lie within the image bounds. Assuming a $2\overline{u} \times 2\overline{v}$ image and an image-centered frame, the principal point (u, v) is within the image bounds if $u_{ij}(\pi_\infty) = \overline{u}^2(\mathsf{Q}_{\infty ij})^2_{31} - (\mathsf{Q}_{\infty ij})^2_{11} \geq 0$ and $v_{ij}(\pi_\infty) = \overline{v}^2(\mathsf{Q}_{\infty ij})^2_{32} - (\mathsf{Q}_{\infty ij})^2_{22} \geq 0$.

Chirality inequalities: Hartley's chirality inequalities [13] can be used to preserve the convex hull of camera centers. They impose that all $c_i(\pi_\infty)$, $i = 1, 2, \ldots, n$, as defined in (7), carry the same sign, provided that all P_i matrices are sign-corrected [13, 24]. Although the same can be done for scene points, these are generally considered to be less reliable [24].

Polynomial optimization with Lasserre's hierarchy: A normalized cost function is preferred in uncalibrated vision problems to eliminate the effect of projective scale factors. This usually leads to a cost in the form of a sum of rational functions. This problem is difficult to solve globally and optimally [3] and is not handled well by Lasserre's method. To reduce the effect of the scaling with an unnormalized cost, we propose to use *homogenized polynomials* [7]. Doing so allows us to include an additional constraint to impose some global scaling. Though this is not equivalent to the normalized case, it works well in practice. The homogeneous counterpart of a polynomial f of degree d in π , denoted ${}^h f$, is defined by introducing an additional variable π_4 such that ${}^h f(\pi, \pi_4) = \pi_4^d f(\pi/\pi_4)$. Note that the degree of ${}^h f$ remains the same as that of f. We solve the following problem to estimate π_∞,

$$\min_{\pi,\, \pi_4} \quad \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} {}^h m_{ij}^2(\pi, \pi_4) + {}^h p_{ij}^2(\pi, \pi_4) \tag{22a}$$

$$\text{s.t.} \quad {}^h c_i(\pi, \pi_4) > 0, \qquad i = 1, \ldots, n, \tag{22b}$$

$${}^h q_{ij}(\pi, \pi_4) > 0, \qquad i = 1, \ldots, n-1, \quad j = i+1, \ldots, n, \tag{22c}$$

$${}^h u_{ij}(\pi, \pi_4) \geq 0, \quad {}^h v_{ij}(\pi, \pi_4) \geq 0, \quad i = 1, \ldots, n-1, \quad j = i+1, \ldots, n, \tag{22d}$$

$${}^h c_1(\pi, \pi_4)\, {}^h c_n(\pi, \pi_4) + \frac{1}{n-1} \sum_{i=1}^{n-1} {}^h c_i(\pi, \pi_4)\, {}^h c_{i+1}(\pi, \pi_4) = 1, \tag{22e}$$

where (22e) is the global scaling constraint that we have found most appropriate. We have observed that this constraint improves the numerical stability and leads more often to a certified optimal solution with the minimal relaxation order of 4. We suggest using the pairwise constraints (22c)-(22d) only between consecutive views so that their number grows linearly and not quadratically with increasing views. For short sequences, constraints between all image pairs can be included. With long sequences, all but (22e) are optional and can be dropped. Even if the EIP assumption is not exactly satisfied in practice, problem (22) is still suitable since the EIP polynomial is minimized in the cost (22a), allowing for residual errors.

Refinement: We locally refine the π_∞ estimated from (22) using the normalized cost,

$$\min_{\pi} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{m_{ij}^2(\pi) + p_{ij}^2(\pi)}{(c_i(\pi)\, c_j(\pi))^4}. \tag{23}$$

Though the constraint (22e) scales the polynomials suitably, it is recommended, particularly with high levels of noise, to refine the solution to fit a normalized cost.

Algorithm: Given a projective reconstruction, our autocalibration algorithm proceeds as follows: (*i*) estimate π_∞ by solving problem (22), (*ii*) refine the estimated π_∞ using (23), and (*iii*) solve a system of linear equations for the *Dual Image of the Absolute Conic* (DIAC) and extract the intrinsic parameters matrix K through Cholesky factorization (see [13, p. 479]).

5 Experimental results

We tested our autocalibration method using both synthetic data and real images. We computed the 3D RMS error and the following calibration error metrics to assess our results,

$$\Delta f = \sqrt{\frac{(f_x - \widehat{f}_x)^2 + (f_y - \widehat{f}_y)^2}{{f_x}^2 + {f_y}^2}}, \quad \Delta uv = \sqrt{\frac{(u-\widehat{u})^2 + (v-\widehat{v})^2}{u^2+v^2}}, \quad \Delta\gamma = |\gamma - \widehat{\gamma}|,$$

where $(\widehat{f}_x, \widehat{f}_y)$, $(\widehat{u}, \widehat{v})$, and $\widehat{\gamma}$ are the estimated focal lengths, principal point, and skew, respectively. The 3D RMS error was computed after aligning the estimated metric point cloud to the ground truth Euclidean point cloud by a best-fit similarity transformation in the least-squares sense. Our algorithm was implemented in MATLAB R2018b. We used GloptiPoly [14] for problem (22) and set a relaxation order of $d = 4$ in all the experiments. We used MOSEK [23] as the SDP solver and set `MSK_DPAR_INTPNT_CO_TOL_{P|D}FEAS` $= 10^{-20}$. All the experiments were conducted on an i7 3.10 GHz 32 GB RAM computer.

We denote our algorithm by EIP*, and the same algorithm without the EIP polynomial, which then relies only on the modulus constraint in (22a) and (23), by MODULUS*. In addition, the two approaches EIP* and MODULUS* excluding the inequality constraints in problem (22) are denoted by EIP and MODULUS, respectively. In these experiments, we used the inequalities in problem (22) only between consecutive views. Furthermore, we estimated all five intrinsic parameters in step (*iii*) of our algorithm. This ensured a fair comparison with the selected existing methods, all of which solve for five intrinsic parameters.

5.1 Synthetic data experiments

Each synthetic scene consisted of 200 points sampled randomly from the surface of the unit sphere. The cameras were positioned at a distance of 3.5–4 units from the sphere center,

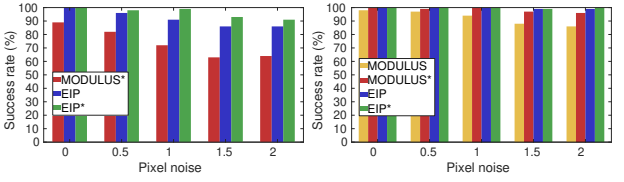

Figure 1: Benefits of EIP polynomial. Success rate of EIP and EIP* compared with MODULUS and MODULUS* using 3 views (left) and 4 views (right).

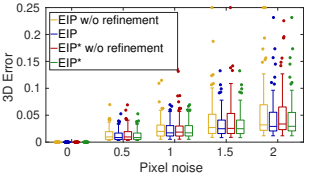

Figure 2: Effect of refinement. 3D errors with and without refinement using 4 views.

and oriented such that their optical axes passed close to the sphere center. All cameras were simulated to have an EIP with focal length $f_x = f_y = 800$, and an image-centered principal point, $(u, v) = (256, 256)$, in pixels. Noise, modeled as a zero-mean Gaussian distribution with standard deviation in the range $[0, 2]$ pixels, was added to the pixel coordinates in increments of 0.5 pixels. Projective reconstructions were obtained using [25] implemented in [29]. We report the statistics collected over 100 generated scenes.

Benefits of EIP polynomial: We assessed the contribution of the EIP polynomial (20) in our algorithm's performance by comparing the reliability of EIP(*) in obtaining a metric reconstruction with that of MODULUS(*). We focused on short sequences and considered a 3D error above 0.25 as a failed metric upgrade. Figure 1 shows the success rate using 3 and 4 views for varying noise levels. With 3 views, MODULUS failed most of the time (result not shown) as multiple solutions exist using the modulus constraint alone. A higher success rate was obtained using MODULUS* due to the inequality constraints, but it declined considerably with increasing noise. On the other hand, EIP led to a reliable metric upgrade even with high noise levels, and the inequalities in EIP* further improved the success rate. With 4 views, there are sufficient polynomials from the modulus constraint to obtain a unique solution. Even so, the success rate of EIP was significantly higher than that of MODULUS. With additional views, all the approaches performed reliably. For the successful trials, the estimated plane at infinity and consequently the 3D errors were similar using all the approaches. These results show that the EIP polynomial is especially useful for short sequences.

Effect of refinement: We analyzed the impact of the refinement step (step (*ii*)) in our algorithm. Figure 2 shows the 3D error distribution using EIP and EIP* with and without refinement for 4 views. The errors decreased overall after refinement, particularly for high levels of noise. Moreover, a few reconstructions that failed in the metric upgrade without refinement were recovered after refinement. We clipped errors above the 0.25 threshold to the axis limit in Figure 2. The observations were similar when varying the number of views. Thus, refinement using a normalized cost improves the accuracy of our algorithm.

Comparisons with the state of the art: We compared EIP and EIP* with two stratified methods, GO-Stratified [6] and QUARCH*M [1], and a DAQ-based method, GO-DAQ [5]. For GO-Stratified, we computed solutions for both signs of chirality and retained the one with lower calibration error (the authors' implementation was used). For GO-DAQ, we set a relaxation order of $d = 2$. The rotation angle assumption of QUARCH*M is satisfied in our simulations. Figure 3 shows the success rate (top row) of the tested methods. With three views, the success rate of GO-DAQ was lower than that of EIP* and it dropped below that of EIP with increasing noise. Although GO-DAQ uses additional priors on the principal point location, the results are inferior because our simulated cameras are close to an

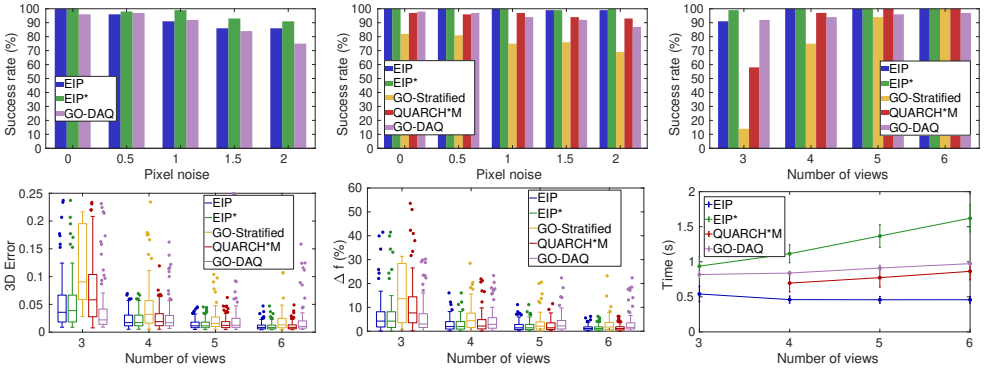

Figure 3: Comparisons with the state of the art. Top: success rate using 3 views (left) and 4 views (middle) with varying noise levels, and using 3-6 views with 1 pixel noise level (right). Bottom: distribution of 3D error (left) and focal length error (middle), and runtime results (right) using 3-6 views with 1 pixel noise level.

artificial degenerate configuration for the DAQ estimation (when all optical axes intersect in one point) [10]. GO-DAQ then fails when its rank-3 constraint is not well enforced due to numerical scaling issues, as has also been reported in [1]. With four views, EIP and EIP* outperformed GO-Stratified and QUARCH*M as well due to the additional EIP constraint. With more views, all the methods succeeded most of the time. The 3D and focal length error distributions as well as the runtime results are shown in Figure 3 (bottom row). GO-Stratified obtained relatively higher 3D and calibration errors and was also two orders of magnitude slower (not shown) than the other methods. For our algorithm, we report the computation time excluding the problem modeling overhead in GloptiPoly. The time complexity of EIP is constant with respect to the number of images as it does not use the pairwise inequalities. From 4 views onward, EIP can thus be used instead of EIP* for a speedup.

5.2 Real image experiments

We used four sequences, *fountain-P11*, *Herz-Jesu-P8*, *Herz-Jesu-P25* [32], and *City hall Leuven* [31], with known ground truth calibration, to quantitatively compare our algorithm with existing methods. We also qualitatively assessed the metric reconstructions obtained with our algorithm using different sequences (see supplementary material). The projective reconstructions were obtained using P2SfM [21] and feature matches using COLMAP [30].

Quantitative assessment: Table 1 reports the calibration errors from EIP, MODULUS, and state-of-the-art methods on three tested sequences. With *fountain-P11*, MODULUS led to large calibration errors and thus it failed to obtain a metric upgrade. While MODULUS* provided a calibration similar to that from EIP, it required 10 times the computation time. MODULUS and EIP otherwise yielded the same calibration as MODULUS* and EIP*, respectively. With *Herz-Jesu-P8*, GO-Stratified failed to obtain a metric upgrade. Moreover, while the results improved with the longer *Herz-Jesu-P25* sequence (not shown) for all the methods, GO-Stratified still led to an erroneous calibration. This is due to the method relying on scene points that prove unreliable with noise and outliers. With *City hall Leuven*, the reference calibration parameters do not fit the assumptions of GO-DAQ as closely as those of the previous sequences. The principal point is farther from the image center and the skew is not null.

Sequence	Method	$\Delta f(\%)$	$\Delta uv(\%)$	$\Delta\gamma$	Time (s)
fountain-P11	MODULUS	62.90	66.39	2172.12	0.91
	EIP	0.08	0.25	1.06	0.59
	GO-Stratified	0.10	0.19	1.08	302.90
	QUARCH*M	0.05	0.23	1.05	2.44
	GO-DAQ	0.36	1.26	0.01	1.49
Herz-Jesu-P8	MODULUS	0.89	3.12	2.16	0.82
	EIP	0.55	2.84	3.98	0.57
	GO-Stratified	43.86	31.13	157.31	243.18
	QUARCH*M	0.88	3.11	2.03	1.26
	GO-DAQ	1.43	1.27	0.05	1.53
City hall Leuven	MODULUS	2.96	6.73	5.90	0.62
	EIP	0.78	0.72	2.80	0.56
	GO-Stratified	7.09	10.10	25.85	169.21
	QUARCH*M	2.94	6.70	5.81	1.02
	GO-DAQ	9.93	7.68	9.70	1.38

Table 1: Quantitative assessment. Autocalibration results on the real image sequences from [31, 32].

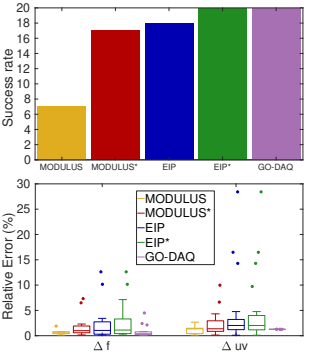

Figure 4: Quantitative assessment using 3 views. Success rate (top), and error distribution (bottom) over 20 triplets sampled from *Herz-Jesu-P25*.

As a result, the errors are larger using GO-DAQ. In contrast, EIP provided an accurate calibration for all the sequences and required only half a second of computation time. The errors from MODULUS and QUARCH*M are similar as they rely on the same cost.

Quantitative assessment using three views: To test the minimal case of three views, we sampled image triplets sequentially from the *Herz-Jesu-P25* sequence, discarding those with insufficient feature matches, leaving a set of 20 triplets. Figure 4 shows the results from EIP(*), MODULUS(*), and GO-DAQ on this set. We considered a focal length error above 25% as a failure in this experiment. From our tests, the quality of the metric reconstruction was mostly influenced by the estimated focal length, and errors above this threshold corresponded to distorted reconstructions. The results in Figure 4 are consistent with those on the synthetic data as MODULUS failed most of the time and EIP performed reliably. Both EIP* and GO-DAQ succeeded with all the triplets. GO-DAQ also consistently provided an accurate calibration as its assumptions on the calibration parameters are closely satisfied in this sequence. The skew parameter was also accurately estimated by all the methods (not shown), resulting in less than 1° deviation from a rectangular image plane on average.

6 Conclusion

We presented a stratified autocalibration method for a moving camera with a Euclidean image plane and constant intrinsic parameters. Our method relies on a new quartic polynomial in the plane at infinity that is obtained for each image pair with these assumptions. For three or more images, estimating the plane at infinity is formulated as a constrained polynomial optimization problem that is solved using Lasserre's hierarchy of semidefinite relaxations. Experiments with synthetic data and real images showed that our method performs more reliably than existing ones, especially for short sequences.

Acknowledgements: This research was supported by the ANR SUMUM project under grant number ANR-17-CE38-0004 and by the ICube lab, CNRS.

References

[1] D. Adlakha, A. Habed, F. Morbidi, C. Demonceaux, and M. de Mathelin. QUARCH: A New Quasi-Affine Reconstruction Stratum From Vague Relative Camera Orientation Knowledge. In *Proc. IEEE International Conference on Computer Vision*, pages 1082–1090, 2019.

[2] B. Bocquillon, A. Bartoli, P. Gurdjos, and A. Crouzil. On Constant Focal Length Self-Calibration From Multiple Views. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2007.

[3] F. Bugarin, D. Henrion, and J.B. Lasserre. Minimizing the sum of many rational functions. *Mathematical Programming Computation*, 8(1):83–111, 2015.

[4] F. Bugarin, A. Bartoli, D. Henrion, J.B. Lasserre, J.-J. Orteu, and T. Sentenac. Rank-Constrained Fundamental Matrix Estimation by Polynomial Global Optimization Versus the Eight-Point Algorithm. *Journal of Mathematical Imaging and Vision*, 53(1): 42–60, 2016.

[5] M. Chandraker, S. Agarwal, F. Kahl, D. Nistér, and D. Kriegman. Autocalibration via Rank-Constrained Estimation of the Absolute Quadric. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, 2007.

[6] M. Chandraker, S. Agarwal, D. Kriegman, and S. Belongie. Globally Optimal Algorithms for Stratified Autocalibration. *International Journal of Computer Vision*, 90(2): 236–254, 2010.

[7] D.A. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry*. Graduate Texts in Mathematics. Springer New York, 2005.

[8] A. Fusiello, A. Benedetti, M. Farenzena, and A. Busti. Globally convergent autocalibration using interval analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(12):1633–1638, 2004.

[9] R. Gherardi and A. Fusiello. Practical Autocalibration. In *Proc. European Conference on Computer Vision*, pages 790–801, 2010.

[10] P. Gurdjos, A. Bartoli, and P.F. Sturm. Is dual linear self-calibration artificially ambiguous? In *Proc. IEEE International Conference on Computer Vision*, pages 88–95, 2009.

[11] A. Habed, K. Al Ismaeil, and D. Fofi. A New Set of Quartic Trivariate Polynomial Equations for Stratified Camera Self-calibration under Zero-Skew and Constant Parameters Assumptions. In *Proc. European Conference on Computer Vision*, pages 710–723, 2012.

[12] A. Habed, D.P. Paudel, C. Demonceaux, and D. Fofi. Efficient Pruning LMI Conditions for Branch-and-Prune Rank and Chirality-Constrained Estimation of the Dual Absolute Quadric. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 493–500, 2014.

[13] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2003.

[14] D. Henrion, J.B. Lasserre, and J. Löfberg. GloptiPoly 3: Moments, Optimization and Semidefinite Programming. *Optimization Methods and Software*, 24(4–5):761–779, 2009.

[15] A. Heyden and K. Åström. Euclidean reconstruction from constant intrinsic parameters. In *Proc. IEEE International Conference on Pattern Recognition*, volume 1, pages 339–343, 1996.

[16] A. Heyden and K. Åström. Euclidean Reconstruction from Image Sequences with Varying and Unknown Focal Length and Principal Point. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 438–443, 1997.

[17] A. Heyden and K. Åström. Flexible Calibration: Minimal Cases for Auto-calibration. In *Proc. IEEE International Conference on Computer Vision*, 1999.

[18] F. Kahl and D. Henrion. Globally Optimal Estimates for Geometric Reconstruction Problems. *International Journal of Computer Vision*, 74(1):3–15, 2007.

[19] J.B. Lasserre. Global Optimization with Polynomials and the Problem of Moments. *SIAM Journal on Optimization*, 11(3):796–817, 2001.

[20] Q. Luong and O. Faugeras. Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices. *International Journal of Computer Vision*, 22 (3):261–289, 1997.

[21] L. Magerand and A. Del Bue. Revisiting Projective Structure from Motion: A Robust and Efficient Incremental Solution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2):430–443, 2018.

[22] L. Magerand, A. Bartoli, O. Ait-Aider, and D. Pizarro. Global Optimization of Object Pose and Motion from a Single Rolling Shutter Image with Automatic 2D-3D Matching. In *Proc. European Conference on Computer Vision*, pages 456–469, 2012.

[23] MOSEK ApS. Manual of the MOSEK optimization toolbox for MATLAB. Version 8.1. http://docs.mosek.com/8.1/toolbox/index.html, 2017.

[24] D. Nistér. Untwisting a Projective Reconstruction. *International Journal of Computer Vision*, 60(2):165–183, 2004.

[25] J. Oliensis and R. Hartley. Iterative Extensions of the Sturm/Triggs Algorithm: Convergence and Nonconvergence. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2217–2233, 2007.

[26] M. Pollefeys and L. Van Gool. Stratified self-calibration with the modulus constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):707–724, 1999.

[27] M. Pollefeys, R. Koch, and L. Van Gool. Self-Calibration and Metric Reconstruction Inspite of Varying and Unknown Intrinsic Camera Parameters. *International Journal of Computer Vision*, 32(1):7–25, 1999.

[28] J. Ponce, K. McHenry, T. Papadopoulo, M. Teillaud, and B. Triggs. On the absolute quadratic complex and its application to autocalibration. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, 2005.

[29] V. Rabaud. Vincent's Structure from Motion Toolbox. http://github.com/vrabaud/sfm_toolbox.

[30] J. Schönberger and J.-M. Frahm. Structure-from-Motion Revisited. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 4104–4113, 2016.

[31] C. Strecha, T. Tuytelaars, and L. Van Gool. Dense Matching of Multiple Wide-baseline Views. In *Proc. IEEE International Conference on Computer Vision*, volume 2, pages 1194–1201, 2003.

[32] C. Strecha, W. von Hansen, L. Van Gool, P. Fua, and U. Thoennessen. On Benchmarking Camera Calibration and Multi-View Stereo for High Resolution Imagery. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, 2008.

[33] B. Triggs. Autocalibration and the absolute quadric. In *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pages 609–614, 1997.

[34] A. Valdés, J.I. Ronda, and G. Gallego. The absolute line quadric and camera autocalibration. *International Journal of Computer Vision*, 66(3):283–303, 2006.

[35] F.C. Wu, M. Zhang, and Z.Y. Hu. Self-Calibration Under the Cayley Framework. *International Journal of Computer Vision*, 103:372–398, 2013.