

Self-Calibration of Cameras with Euclidean Image Plane in Case of Two Views and Known Relative Rotation Angle

Evgeniy Martynushev*^[0000-0002-6892-079X]

South Ural State University, 454080 Chelyabinsk, Russia
mev@susu.ac.ru

Abstract. The internal calibration of a pinhole camera is given by five parameters that are combined into an upper-triangular 3×3 calibration matrix. If the skew parameter is zero and the aspect ratio is equal to one, then the camera is said to have Euclidean image plane. In this paper, we propose a non-iterative self-calibration algorithm for a camera with Euclidean image plane in case the remaining three internal parameters — the focal length and the principal point coordinates — are fixed but unknown. The algorithm requires a set of $N \geq 7$ point correspondences in two views and also the measured relative rotation angle between the views. We show that the problem generically has six solutions (including complex ones).

The algorithm has been implemented and tested both on synthetic data and on publicly available real dataset. The experiments demonstrate that the method is correct, numerically stable and robust.

Keywords: Multiview geometry · Self-calibration · Essential matrix · Euclidean image plane · Relative rotation angle · Gröbner basis

1 Introduction

The problem of camera calibration is an essential part of numerous computer vision applications including 3d reconstruction, visual odometry, medical imaging, etc. At present, a number of calibration algorithms and techniques have been developed. Some of them require to observe a planar pattern viewed at several different positions [5, 9, 25]. Other methods use 3d calibration objects consisting of two or three pairwise orthogonal planes, which geometry is known with good accuracy [24]. Also, there are calibration algorithms assuming that a scene involves the pairs of mutually orthogonal directions [3, 14]. In contrast with the just mentioned methods, *self-calibration* does not require any special calibration objects or scene restrictions [6, 16, 18, 23], so only image feature correspondences in several uncalibrated views are required. This provides the self-calibration approach with a great flexibility and makes it indispensable in some real-time applications.

* The work was supported by Act 211 Government of the Russian Federation, contract No. 02.A03.21.0011.

In two views, camera calibration is given by ten parameters — five internal and five external, whereas the fundamental matrix describing the epipolar geometry in two views has only seven degrees of freedom [6]. This means that self-calibration in two views is only possible under at least three further assumptions on the calibration parameters. For example, we can assume that the skew parameter is zero and so is the translation vector, i.e. the motion is a pure rotation. Then, the three orientation angles and the remaining four internals can be self-calibrated from at least seven point matches [7]. Another possibility is that all the internal parameters except common focal length are known. Then, there exist a minimal self-calibration solution operating with six matched points [20, 2, 11, 13].

The five internal calibration parameters have different interpretation. The skew parameter and the aspect ratio describe the pixel's shape. In most situations, e.g. under zooming, these internals do not change. Moreover, for modern cameras the pixel's shape is very close to a square and hence the skew and aspect ratio can be assumed as given and equal to 0 and 1 respectively. Following [10], we say that a camera in this case has *Euclidean image plane*. On the other hand, the focal length and the principal point coordinates describe the relative placement of the camera center and the image plane. The focal length is the distance between the camera center and the image plane, whereas the principal point is an orthographic projection of the center onto the plane. All these internals should be considered as unknown, since even for modern cameras the principal point can be relatively far from the geometric image center. Besides, it is well known that the focal length and the principal point always vary together by zooming [22].

The aim of this paper is to propose an efficient solution to the self-calibration problem of a camera with Euclidean image plane. As it was mentioned above, in two views at most seven calibration parameters can be self-calibrated. Since a camera with Euclidean image plane has eight parameters, we conclude that one additional assumption should be made. In this paper we reduce the number of external parameters assuming that the *relative rotation angle* between the views is known. The problem thus becomes minimally constrained from seven point correspondences in two views. In practice, the relative rotation angle can be reliably found from e.g. the readings of an inertial measurement unit (IMU) sensor. The possibility of using such additional data in structure-from-motion has been demonstrated in [12].

In general, the joint usage of a camera and an IMU requires *external calibration* between the devices, i.e. we have to know the transformation matrix between their coordinate frames. However, if only relative rotation angle is used, then the external calibration is unnecessary, provided that both devices are fixed on some rigid platform. Thus, the rotation angle of the IMU can be directly used as the rotation angle of the camera [12]. This fact makes the presented self-calibration method more convenient and flexible for practical use.

To summarize, we propose a new non-iterative solution to the self-calibration problem in case of at least seven matched points in two views, provided the following assumptions:

- the camera intrinsic parameters are the same for both views;
- the camera has Euclidean image plane;
- the relative rotation angle between the views is known.

Our self-calibration method is based on the ten quartic equations. Nine of them are well-known and follow from the famous cubic constraint on the essential matrix. The novel last one (see Eq. (13)) arises from the condition that the relative angle is known.

Finally, throughout the paper it is assumed that the cameras and scene points are in *sufficiently general position*. There are critical camera motions for which self-calibration is impossible, unless some further assumptions on the internal parameters or the motion are made [21]. Also there exist degenerate configurations of scene points. However, in this paper we restrict ourselves to the generic case of camera motions and points configurations.

The rest of the paper is organized as follows. In Section 2, we recall some definitions and results from multiview geometry and deduce our self-calibration constraints. In Section 3, we describe in detail the algorithm. In Section 4 and Section 5, the algorithm is validated in a series of experiments on synthetic and real data. In Section 6, we discuss the results and make conclusions.

2 Preliminaries

2.1 Notation

We preferably use α, β, \dots for scalars, a, b, \dots for column 3-vectors or polynomials, and A, B, \dots both for matrices and column 4-vectors. For a matrix A the entries are $(A)_{ij}$, the transpose is A^T , the determinant is $\det A$, and the trace is $\text{tr } A$. For two 3-vectors a and b the cross product is $a \times b$. For a vector a the notation $[a]_{\times}$ stands for the skew-symmetric matrix such that $[a]_{\times} b = a \times b$ for any vector b . We use I for the identity matrix.

2.2 Fundamental and essential matrices

Let there be given two cameras $P = [I \ 0]$ and $P' = [A \ a]$, where A is a 3×3 matrix and a is a 3-vector. Let Q be a 4-vector representing homogeneous coordinates of a point in 3-space, q and q' be its images, that is

$$q \sim PQ, \quad q' \sim P'Q, \tag{1}$$

where \sim means an equality up to non-zero scale. The *coplanarity constraint* for a pair (q, q') says

$$q'^T F q = 0, \tag{2}$$

where matrix $F = [a]_{\times} A$ is called the *fundamental matrix*.

It follows from the definition of matrix F that $\det F = 0$. This condition is also sufficient. Thus we have

Theorem 1 ([8]). *A non-zero 3×3 matrix F is a fundamental matrix if and only if*

$$\det F = 0. \quad (3)$$

The *essential matrix* E is the fundamental matrix for *calibrated cameras* $\hat{P} = [I \ 0]$ and $\hat{P}' = [R \ t]$, where $R \in \text{SO}(3)$ is called the *rotation matrix* and t is called the *translation vector*. Hence, $E = [t]_{\times} R$. Matrices F and E are related by

$$E \sim K'^T F K, \quad (4)$$

where K and K' are the upper-triangular *calibration matrices* of the first and second camera respectively.

The fundamental matrix has seven degrees of freedom, whereas the essential matrix has only five degrees of freedom. It is translated into extra constraints on the essential matrix. The following theorem gives one possible form of such constraints.

Theorem 2 ([15]). *A 3×3 matrix E of rank two is an essential matrix if and only if*

$$\frac{1}{2} \text{tr}(EE^T)E - EE^T E = \mathbf{0}_{3 \times 3}. \quad (5)$$

2.3 Self-Calibration Constraints

Let θ be the angle of rotation between two calibrated camera frames. In case θ is known, the trace of rotation matrix $\tau = 2 \cos \theta + 1$ is known too. This leads to an additional quadratic constraint on the essential matrix.

Proposition 1. *Let $E = [t]_{\times} R$ be a real non-zero essential matrix and $\text{tr} R = \tau$. Then E satisfies the equation*

$$\frac{1}{2} (\tau^2 - 1) \text{tr}(EE^T) + (\tau + 1) \text{tr}(E^2) - \tau \text{tr}^2 E = 0. \quad (6)$$

Proof. Let $U \in \text{SO}(3)$ be such that $Ut = [0 \ 0 \ 1]^T$. Then,

$$\hat{E} = UEU^T = U[t]_{\times} U^T U R U^T = [Ut]_{\times} \hat{R} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \hat{R}, \quad (7)$$

where $\hat{R} = URU^T$. It is clear that if E satisfies Eq. (6), then so does \hat{E} and vice versa. Let us represent \hat{R} in terms of a unit quaternion $s + u\mathbf{i} + v\mathbf{j} + w\mathbf{k}$, i.e.

$$\hat{R} = \begin{bmatrix} 1 - 2v^2 - 2w^2 & 2uv - 2ws & 2uw + 2vs \\ 2uv + 2ws & 1 - 2u^2 - 2w^2 & 2vw - 2us \\ 2uw - 2vs & 2vw + 2us & 1 - 2u^2 - 2v^2 \end{bmatrix}, \quad (8)$$

where $s^2 + u^2 + v^2 + w^2 = 1$. Then,

$$\hat{E} = \begin{bmatrix} 2uv + 2ws & 1 - 2u^2 - 2w^2 & 2vw - 2us \\ -1 + 2v^2 + 2w^2 & -2uv + 2ws & -2uw - 2vs \\ 0 & 0 & 0 \end{bmatrix}. \quad (9)$$

Substituting this into (6), after some computation, we get

$$\text{l.h.s. of (6)} = (\tau + 1 - 4s^2)(\tau + 1 - 4w^2). \quad (10)$$

This completes the proof, since $\tau = \text{tr } R = \text{tr } \hat{R} = 4s^2 - 1$.

It is well known [8, 15] that for a given essential matrix E there is a “twisted pair” of rotations R_a and R_b so that $E \sim [\pm t]_{\times} R_a \sim [\pm t]_{\times} R_b$. By Proposition 1, $\tau_a = \text{tr } R_a$ and $\tau_b = \text{tr } R_b$ must be roots of Eq. (6). Since the equation is quadratic in τ there are no other roots. Thus we have

Proposition 2. *Let E be a real non-zero essential matrix satisfying Eq. (6) for a certain $\tau \in [-1, 3]$. Then, either $\tau = \text{tr } R_a$ or $\tau = \text{tr } R_b$, where (R_a, R_b) is the twisted pair of rotations for E .*

Now suppose that we are given two cameras with unknown but identical calibration matrices K and $K' = K$. Then we have

$$E \sim K^T F K, \quad (11)$$

where F is the fundamental matrix. Substituting this into Eqs. (5)–(6), we get the following ten equations:

$$\frac{1}{2} \text{tr}(F\omega^* F^T \omega^*) F - F\omega^* F^T \omega^* F = 0_{3 \times 3}, \quad (12)$$

$$\frac{1}{2} (\tau^2 - 1) \text{tr}(F\omega^* F^T \omega^*) + (\tau + 1) \text{tr}(\omega^* F \omega^* F) - \tau \text{tr}^2(\omega^* F) = 0, \quad (13)$$

where $\omega^* = K K^T$. Constraints (12)–(13) involve the internal parameters of a camera and hence can be used for its self-calibration. We notice that not all of these constraints are necessarily linearly independent.

Proposition 3. *If the fundamental matrix F is known, then Eq. (12) gives at most three linearly independent constraints on the entries of ω^* .*

Proof. Recall that matrix F is generically of rank two. Let the right and left null vectors of F be e and e' respectively. Denote by G the l.h.s. of Eq. (12). Then it is clear that

$$Ge = G^T e' = 0_{3 \times 1}. \quad (14)$$

It follows that, given F , at least six of $(G)_{ij}$ are linearly dependent. Proposition 3 is proved.

3 Description of the Algorithm

The initial data for our algorithm is $N \geq 7$ point correspondences $q_i \leftrightarrow q'_i$, $i = 1, \dots, N$, and also the trace τ of the rotation matrix R .

3.1 Data pre-normalization

To significantly improve the numerical stability of our algorithm, points q_i and q'_i are first normalized as follows, adapted from [8]. We construct a 3×3 matrix S of the form

$$S = \begin{bmatrix} \gamma & 0 & \alpha \\ 0 & \gamma & \beta \\ 0 & 0 & 1 \end{bmatrix} \quad (15)$$

so that the $2N$ new points, represented by the columns of matrix

$$S [q_1 \dots q_N \ q'_1 \dots q'_N], \quad (16)$$

satisfy the following conditions:

- their centroid is at the coordinate origin;
- their average distance from the origin is $\sqrt{2}$.

From now on we assume that q_i and q'_i are normalized.

3.2 Polynomial equations

The fundamental matrix F is estimated from $N \geq 7$ point correspondences in two views. The algorithm is well known, see [8] for details. In the minimal case $N = 7$ there are either one or three real solutions. Otherwise, the solution is generically unique.

Let both cameras be identically calibrated and have Euclidean image planes, i.e.

$$K = K' = \begin{bmatrix} f & 0 & a \\ 0 & f & b \\ 0 & 0 & 1 \end{bmatrix}, \quad (17)$$

where f is the focal length and (a, b) is the principal point. It follows that

$$\omega^* = KK^T = \begin{bmatrix} a^2 + p & ab & a \\ ab & b^2 + p & b \\ a & b & 1 \end{bmatrix}, \quad (18)$$

where we introduce a new variable $p = f^2$.

Substituting F and ω^* into constraints (12) and (13), we get ten quartic equations in variables a, b and p . Let G be the l.h.s. of (12). By Proposition 3, up to three of $(G)_{ij}$ are linearly independent. Let $f_1 = (G)_{11}$, $f_2 = (G)_{22}$,

$f_3 = (G)_{33}$ and $f_4 = \text{l.h.s. of (13)}$. The objective is to find all feasible solutions of the following system of polynomial equations

$$f_1 = f_2 = f_3 = f_4 = 0. \quad (19)$$

Let us define the ideal $J = \langle f_1, f_2, f_3, f_4 \rangle \subset \mathbb{C}[a, b, p]$. Unfortunately, ideal J is not zero-dimensional. There is a one-parametric family of solutions of system (19) corresponding to the unfeasible case $p = 0$. We state without proof the decomposition

$$\sqrt{J} = J' \cap \langle p, \text{tr}(F\omega^*) \rangle, \quad (20)$$

where \sqrt{J} is the radical of J and $J' = J/\langle p \rangle$ is the quotient ideal, which is already zero-dimensional. By (20), the affine variety of J is the union of a finite set of points in \mathbb{C}^3 and a conic in the plane $p = 0$.

3.3 Gröbner basis

In this subsection, the Gröbner basis of ideal J' will be constructed. We start from rewriting equations (19) in form

$$B_0 y_0 = 0, \quad (21)$$

where B_0 is a 4×22 coefficient matrix, and

$$y_0 = [a^3 b \quad a^2 b^2 \quad ab^3 \quad a^2 b \quad a^4 \quad b^4 \quad a^3 \quad ab^2 \quad b^3 \quad a^2 p \quad abp \quad b^2 p \quad a^2 \quad ab \quad b^2 \quad ap \quad bp \quad p^2 \quad a \quad b \quad p \quad 1]^T \quad (22)$$

is a monomial vector. Let us consider the following sequence of transformations:

$$(B_i, y_i) \rightarrow (\tilde{B}_i, y_i) \rightarrow (B_{i+1}, y_{i+1}), \quad i = 0, \dots, 4. \quad (23)$$

Here each \tilde{B}_i is the *reduced row echelon form* of B_i . The monomials in y_i are ordered so that the left of matrix \tilde{B}_i is an identity matrix for each i .

We exploit below some properties of the intermediate polynomials in sequence (23), e.g. they can be factorized or have degree lower than one expects from the corresponding monomial vector. All of these properties have been verified in Maple by using randomly generated instances of the problem over the field of rationals.

Let us denote by $(A)_i$ the i th row of matrix A . Now we describe in detail each transformation $(\tilde{B}_i, y_i) \rightarrow (B_{i+1}, y_{i+1})$ of sequence (23).

- The last row of \tilde{B}_0 corresponds to a 3rd degree polynomial. Matrix B_1 of size 7×32 is obtained from \tilde{B}_0 by appending 3 new rows and 10 new columns. The rows are: $a(\tilde{B}_0)_4$, $b(\tilde{B}_0)_4$ and $p(\tilde{B}_0)_4$. Monomial vector

$$y_1 = [a^3 b \quad a^2 b^2 \quad ab^3 \quad a^2 b \quad \underline{a^3 p} \quad \underline{a^2 bp} \quad \underline{ab^2 p} \quad a^4 \quad b^4 \quad a^3 \quad ab^2 \quad b^3 \quad a^2 p \quad \underline{b^2 p^2} \quad \underline{ap^2} \quad abp \quad b^2 p \quad \underline{b^3 p} \quad a^2 \quad \underline{abp^2} \quad \underline{a^2 p^2} \quad ab \quad b^2 \quad \underline{bp^2} \quad \underline{p^3} \quad ap \quad bp \quad p^2 \quad a \quad b \quad p \quad 1]^T, \quad (24)$$

where we underlined the new monomials (columns) of matrix B_1 .

- The polynomials corresponding to the last three rows of \tilde{B}_1 are divisible by p . Matrix B_2 of size 13×32 is obtained as follows. We append 6 new rows to \tilde{B}_1 , which are $(\tilde{B}_1)_i/p$, $a(\tilde{B}_1)_i/p$ and $b(\tilde{B}_1)_i/p$ for $i = 6, 7$. Monomial vector $y_2 = y_1$.
- Matrix B_3 of size 19×32 is obtained from \tilde{B}_2 by appending 6 new rows: $a(\tilde{B}_2)_i$, $b(\tilde{B}_2)_i$ and $p(\tilde{B}_2)_i$ for $i = 12, 13$. Monomial vector $y_3 = y_1$.
- The last row of \tilde{B}_3 corresponds to a 2nd degree polynomial. Thus we proceed with the polynomials of degree up to 3. We eliminate from \tilde{B}_3 rows and columns corresponding to all 4th degree polynomials and monomials respectively. Matrix B_4 of size 11×20 is obtained from \tilde{B}_3 as follows. We hold the rows of \tilde{B}_3 with numbers 4, 10, 11, 12, 13, 16, 17, 19, and append 3 new rows: $a(\tilde{B}_3)_{19}$, $b(\tilde{B}_3)_{19}$ and $p(\tilde{B}_3)_{19}$. Monomial vector

$$y_4 = [a^2b \quad a^3 \quad ab^2 \quad b^3 \quad a^2p \quad ap^2 \quad abp \quad b^2p \quad a^2 \quad ab \quad b^2 \quad bp^2 \quad p^3 \quad ap \quad bp \quad p^2 \quad a \quad b \quad p \quad 1]^T. \quad (25)$$

- Finally, matrix B_5 of size 14×20 is obtained from \tilde{B}_4 by appending 3 new rows: $a(\tilde{B}_4)_{11}$, $b(\tilde{B}_4)_{11}$ and $p(\tilde{B}_4)_{11}$. Monomial vector $y_5 = y_4$.

The last six rows of matrix \tilde{B}_5 constitute the (reduced) Gröbner basis of ideal J' w.r.t. the graded reverse lexicographic order with $a > b > p$.

3.4 Internal and external parameters

Given the Gröbner basis of J' , the 6×6 action matrix M_p for multiplication by p in the quotient ring $\mathbb{C}[a, b, p]/J'$ can be easily constructed as follows. We denote by C the 6×6 right lower submatrix of \tilde{B}_5 . Then the first three rows of M_p are the last three rows of $(-C)$. The rest of M_p consists of almost all zeros except

$$(M_p)_{41} = (M_p)_{52} = (M_p)_{65} = 1. \quad (26)$$

The six solutions are then found from the eigenvectors of matrix M_p , see [4] for details. Complex solutions and the ones with $p < 0$ are excluded.

Having found the calibration matrix K , we compute the essential matrix E from formula (11) and then the externals R and t using the standard procedure, see e.g. [17]. Note that, due to Proposition 2 and the *cheirality constraint* [8, 17], the trace of the estimated matrix R must equal τ .

Finally, the denormalized entities (see subsection 3.1 and the definition of matrix S) are found as follows:

- fundamental matrix is $S^T F S$;
- calibration matrix is $S^{-1} K$;
- essential matrix is unchanged, as $K^T S^{-T} S^T F S S^{-1} K = K^T F K \sim E$;
- externals R and t are also unchanged.

4 Experiments on Synthetic Data

In this section we test the algorithm on synthetic image data. The default data setup is given in the following table:

Distance to the scene	1
Scene depth	0.5
Baseline length	0.1
Image dimensions	1280 × 720

4.1 Numerical Accuracy

First we verify the numerical accuracy of our method. The numerical error is defined as the minimal relative error in the calibration matrix, that is

$$\min_i \left(\frac{\|K_i - K_{\text{gt}}\|}{\|K_{\text{gt}}\|} \right). \quad (27)$$

Here $\|\cdot\|$ stands for the Frobenius norm, i counts all real solutions, and

$$K_{\text{gt}} = \begin{bmatrix} 1000 & 0 & 640 \\ 0 & 1000 & 360 \\ 0 & 0 & 1 \end{bmatrix} \quad (28)$$

is the ground truth calibration matrix. The numerical error distribution is reported in Fig. 1.

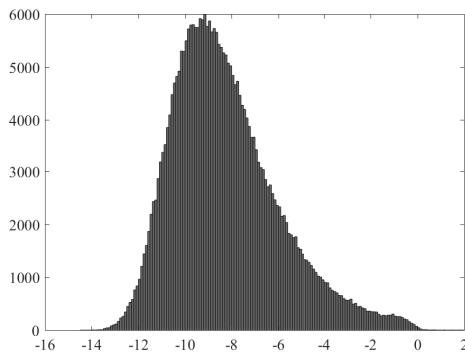


Fig. 1. \log_{10} of numerical error for noise free data. The median error is 2.5×10^{-9}

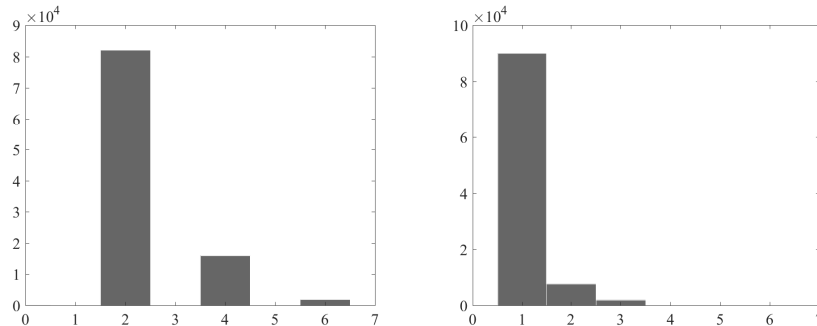


Fig. 2. The number of real (left) and feasible (right) solutions for the internal calibration. Noise free data

4.2 Number of solutions

In general, the algorithm outputs six solutions for the calibration matrix, counting both real and complex ones. The number of real solutions is usually two or four. The number of real solutions with positive p (we call such solutions feasible) is equal to one in most cases. The corresponding distributions are demonstrated in Fig. 2.

4.3 Behaviour under noise

To evaluate the robustness of our solver, we add two types of noise to the initial data. First, the image noise which is modelled as zero-mean, Gaussian distributed with a standard deviation varying from 0 to 1 pixel in a 1280×720 image. Second, the angle noise resulting from inaccurately found relative rotation angle θ . In practice, θ is computed by integrating the angular velocity measurements obtained from a 3d gyroscope. Therefore, a realistic model for the angle noise is quite complicated and depends on a number of factors such as the value of θ , the measurement rates, the inertial sensor noise model, etc. In a very simplified manner, the angle noise can be modelled as θs [12], where s has the Gaussian distribution with zero mean and standard deviation σ . In our experiments σ ranges from 0 to 0.09.

Fig. 3 demonstrates the behaviour of the algorithm under increasing the image and angle noise. Here and below each point on the diagram is a median of 10^4 trials.

4.4 Comparison with the existing solvers

We compare our algorithm for the minimal number of points ($N = 7$) with the 6-point solver from [20] and the 5-point solver from [19].

Fig. 4 depicts the relative focal length error $|f - 1000|/1000$ at varying levels of image noise. Here and below $\alpha = 0.1$ means a 10%-miscalibration in the

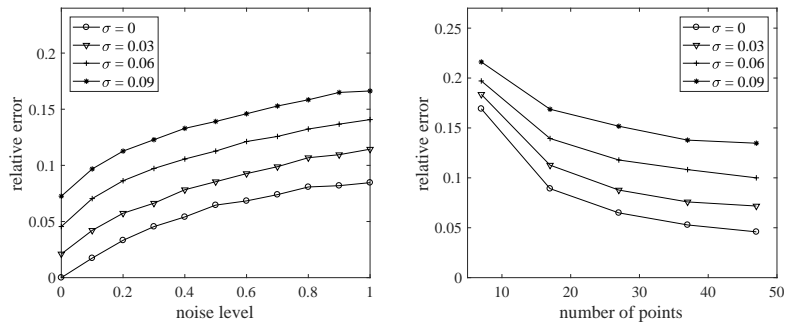


Fig. 3. The relative error in the calibration matrix at varying level of noise (left) and at varying number of points (right). The number of image points on the left figure is $N = 20$. The standard deviation for the image noise on the right figure is 1 pixel

parameters a and b , i.e. the data was generated using the ground truth calibration matrix K_{gt} , whereas the solutions were found assuming that

$$K = \begin{bmatrix} 1000 & 0 & (1 + \alpha)640 \\ 0 & 1000 & (1 + \alpha)360 \\ 0 & 0 & 1 \end{bmatrix}. \quad (29)$$

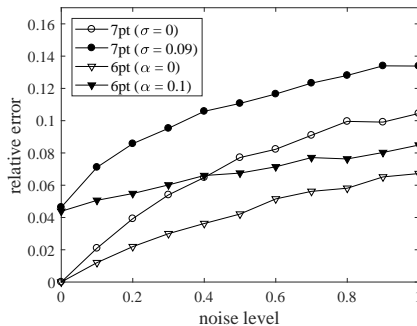


Fig. 4. The relative focal length error against noise standard deviation in pixels

In Fig. 5, the rotational and translational errors of the algorithms are reported. As it can be seen, for realistic levels of image noise, the 5-point algorithm expectedly outperforms as our solution as the 6-point solver. However, it is worth emphasizing that our solution is more suitable for the internal self-calibration of a camera rather than for its pose estimation. Once the self-calibration is done, it is more efficient to switch to the 5- or even 4-point solvers from [12, 17].

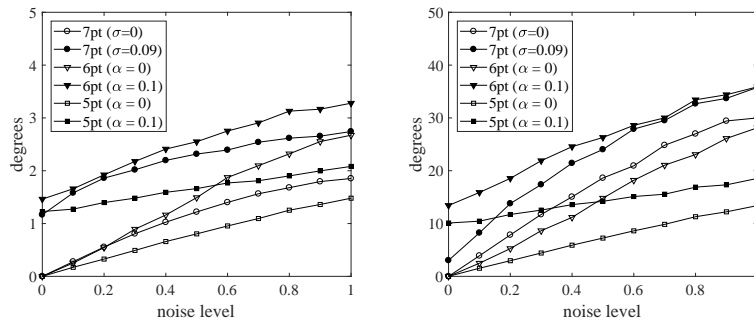


Fig. 5. The rotational (left) and translational (right) errors in degrees against noise standard deviation in pixels

We also compared the speed of the algorithms. The average running times over 10^5 trials are 2.0 ms (our 7pt), 1.6 ms (6pt) and 0.4 ms (5pt) on a system with 2.3 GHz processor. The most expensive step of our algorithm is the computation of the five reduced row echelon forms in sequence (23).

5 Experiments on Real Data

In this section we validate the algorithm by using the publicly available EuRoC dataset [1]. This dataset contains sequences of the data recorded from an IMU and two cameras on board a micro-aerial vehicle and also a ground truth. Specifically, we used the “Machine Hall 01” dataset (easy conditions) and only the images taken by camera “0”. The image size is 752×480 (WVGA) and the ground truth calibration matrix is

$$K_{\text{gt}} = \begin{bmatrix} 458.654 & 0 & 367.215 \\ 0 & 457.296 & 248.375 \\ 0 & 0 & 1 \end{bmatrix}. \quad (30)$$

The sequence of 68 image pairs was derived from the dataset and the algorithm was applied to every image pair, see example in Fig. 6. Here it is necessary to make a few remarks.

- Since the algorithm assumes the pinhole camera model, every image was first undistorted using the ground truth parameters.
- As it was mentioned in Subsection 4.2, the feasible solution is almost always unique. However in rare cases multiple solutions are possible. To reduce the probability of getting such solutions we additionally assumed that the principal point is sufficiently close to the geometric image center. More precisely, the solutions with

$$(a, b) \notin \{|x - 376| < 50, |y - 240| < 50\} \quad (31)$$

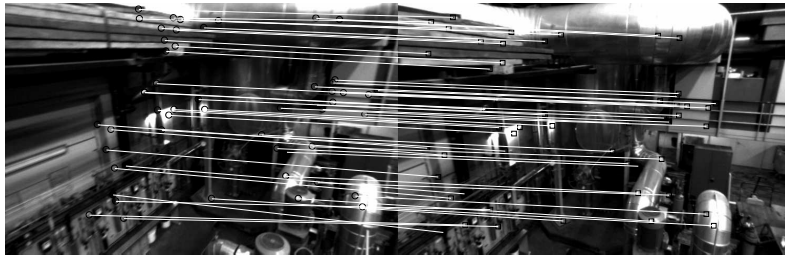


Fig. 6. The pair of undistorted images with the time stamps “140363664626355584” and “1403636646613555456” from the EuRoC dataset and the matched points

were marked as unfeasible and hence discarded. This condition almost guarantees that the algorithm outputs at most one solution.

- Given the readings of a triple-axis gyroscope the relative rotation angle can be computed as follows. The gyroscope reading at time ξ_i is an angular rate 3-vector w_i . Let $\Delta\xi_i = \xi_i - \xi_{i-1}$, where $i = 1, \dots, n$. Then the relative rotation matrix R_n between the 0th and n th frames is approximately found from the recursion

$$R_i = \exp([w_i]_{\times} \Delta\xi_i) R_{i-1}, \quad (32)$$

where $R_0 = I$ and the matrix exponential is computed by the Rodrigues formula

$$\exp([v]_{\times}) = I + \frac{\sin \|v\|}{\|v\|} [v]_{\times} + \frac{1 - \cos \|v\|}{\|v\|^2} [v]_{\times}^2. \quad (33)$$

The relative rotation angle is then derived from $\text{tr } R_n$.

- The image pairs with the relative rotation angle less than 5 degrees were discarded, since the motion in this case is close to a pure translation and self-calibration becomes unstable.

The estimated internal parameters for each image pair are shown in Fig. 7. The calibration matrix averaged over the entire sequence is given by

$$K = \begin{bmatrix} 457.574 & 0 & 366.040 \\ 0 & 457.574 & 243.616 \\ 0 & 0 & 1 \end{bmatrix}. \quad (34)$$

Hence the relative error in the calibration matrix is about 0.6 %.

6 Conclusions

We have presented a new practical solution to the problem of self-calibration of a camera with Euclidean image plane. The solution operates with at least seven

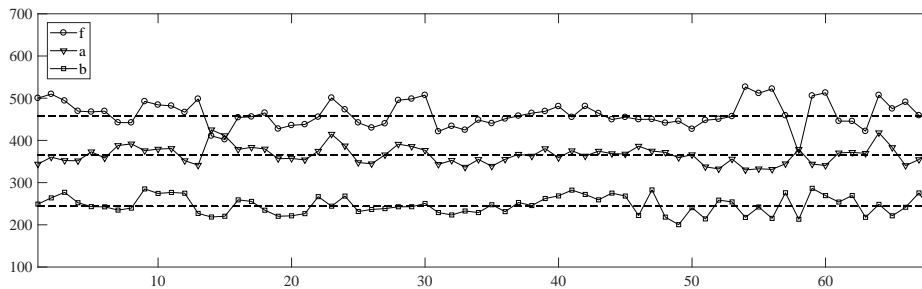


Fig. 7. The estimated focal length and principal point coordinates for the 68 image pairs. The dashed horizontal lines represent the average values

point correspondences in two views and also with the known value of the relative rotation angle.

Our method is based on a novel quadratic constraint on the essential matrix, see Eq. (6). We expect that there could be other applications of that constraint. In particular, it can be used to obtain an alternative solution of the problem from paper [12]. The investigation of that possibility is left for further work.

We have validated the solution in a series of experiments on synthetic and real data. Under the assumption of generic camera motion and points configuration, it is shown that the algorithm is numerically stable and demonstrates a good performance in the presence of noise. It is also shown that the algorithm is fast enough and in most cases it produces a unique feasible solution for camera calibration.

References

1. Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M., Siegwart, R.: The EuRoC micro aerial vehicle datasets. *The International Journal of Robotics Research* **35**(10), 1157–1163 (2016)
2. Byröd, M., Josephson, K., Åström, K.: Improving numerical accuracy of Gröbner basis polynomial equation solver. In: *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. pp. 1–8. IEEE (2007)
3. Caprile, B., Torre, V.: Using vanishing points for camera calibration. *International journal of computer vision* **4**(2), 127–139 (1990)
4. Cox, D., Little, J., O’Shea, D.: *Ideals, Varieties, and Algorithms*, vol. 3. Springer (2007)
5. Faugeras, O.: *Three-Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press (1993)
6. Hartley, R.: Estimation of relative camera positions for uncalibrated cameras. In: *European Conference on Computer Vision*. pp. 579–587. Springer (1992)
7. Hartley, R.: Self-calibration from multiple views with a rotating camera. In: *European Conference on Computer Vision*. pp. 471–478. Springer (1994)
8. Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*. Cambridge University Press (2003)

9. Heikkilä, J.: Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(10), 1066–1077 (2000)
10. Heyden, A., Åström, K.: Euclidean reconstruction from image sequences with varying and unknown focal length and principal point. In: *Conference on Computer Vision and Pattern Recognition*. pp. 438–443. IEEE (1997)
11. Kukulova, Z., Bujnak, M., Pajdla, T.: Polynomial eigenvalue solutions to the 5-pt and 6-pt relative pose problems. In: *British Machine Vision Conference*. vol. 2 (2008)
12. Li, B., Heng, L., Lee, G., Pollefeys, M.: A 4-point algorithm for relative pose estimation of a calibrated camera with a known relative rotation angle. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems*. pp. 1595–1601. IEEE (2013)
13. Li, H.: A simple solution to the six-point two-view focal-length problem. In: *European Conference on Computer Vision*. pp. 200–213. Springer (2006)
14. Liebowitz, D., Zisserman, A.: Combining scene and auto-calibration constraints. In: *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*. vol. 1, pp. 293–300. IEEE (1999)
15. Maybank, S.: *Theory of Reconstruction from Image Motion*. Springer-Verlag (1993)
16. Maybank, S., Faugeras, O.: A theory of self calibration of a moving camera. *International Journal of Computer Vision* **8**(2), 123–151 (1992)
17. Nistér, D.: An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26**(6), 756–770 (2004)
18. Quan, L., Triggs, B.: A unification of autocalibration methods. In: *Asian Conference on Computer Vision*. pp. 917–922 (2000)
19. Stewénius, H., Engels, C., Nistér, D.: Recent developments on direct relative orientation. *ISPRS Journal of Photogrammetry and Remote Sensing* **60**(4), 284–294 (2006)
20. Stewénius, H., Nistér, D., Kahl, F., Schaffalitzky, F.: A minimal solution for relative pose with unknown focal length. *Image and Vision Computing* **26**(7), 871–877 (2008)
21. Sturm, P.: Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction. In: *Conference on Computer Vision and Pattern Recognition*. pp. 1100–1105. IEEE (1997)
22. Sturm, P.: Self-calibration of a moving zoom-lens camera by pre-calibration. *Image and Vision Computing* **15**(8), 583–589 (1997)
23. Triggs, B.: Autocalibration and the absolute quadric. In: *Conference on Computer Vision and Pattern Recognition*. pp. 609–614. IEEE (1997)
24. Tsai, R.: A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal on Robotics and Automation* **3**(4), 323–344 (1987)
25. Zhang, Z.: A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence* **22**(11), 1330–1334 (2000)