Six-Point Method for Multi-Camera Systems with Reduced Solution Space

Banglei Guan[1⋆], Ji Zhao[2⋆(✉)], and Laurent Kneip[3]

[1] College of Aerospace Science and Engineering, National University of Defense Technology, China. guanbanglei12@nudt.edu.cn
[2] Independent Researcher. Beijing, China. zhaoji84@gmail.com
[3] Mobile Perception Lab, ShanghaiTech University, China. lkneip@shanghaitech.edu.cn

Abstract. Relative pose estimation using point correspondences (PC) is a widely used technique. A minimal configuration of six PCs is required for two views of generalized cameras. In this paper, we present several minimal solvers that use six PCs to compute the 6DOF relative pose of multi-camera systems, including a minimal solver for the generalized camera and two minimal solvers for the practical configuration of two-camera rigs. The equation construction is based on the decoupling of rotation and translation. Rotation is represented by Cayley or quaternion parametrization, and translation can be eliminated by using the hidden variable technique. Ray bundle constraints are found and proven when a subset of PCs relate the same cameras across two views. This is the key to reducing the number of solutions and generating numerically stable solvers. Moreover, all configurations of six-point problems for multi-camera systems are enumerated. Extensive experiments demonstrate the superior accuracy and efficiency of our solvers compared to state-of-the-art six-point methods. The code is available at https://github.com/jizhaox/relpose-6pt.

Keywords: Minimal solver · Relative pose estimation · Point correspondence · Multi-camera system · Ray bundle constraint

1 Introduction

Relative pose estimation utilizing feature correspondences is a fundamental problem in geometric computer vision. It plays a crucial role in numerous tasks such as autonomous driving, augmented reality, simultaneous localization and mapping, etc. Despite having a long history, the research on relative pose estimation remains active. These efforts focus on enhancing the efficiency, stability, and accuracy of algorithms [3, 7, 13, 14, 16].

Camera models have a significant impact on computing the relative pose. A pinhole or perspective camera model is used to model a single camera [18], and more complicated cameras like multi-camera systems necessitate the use of

⋆ Equal contribution. ✉ Corresponding author.

a generalized camera model [12, 47]. A generalized camera encapsulates various imaging systems by representing the landmark observations as spatial rays, which do not necessarily require originating from the projection center [41]. This paper is principally concerned with a multi-camera system comprising several cameras that have been installed rigidly. As we shall demonstrate in this paper, n point correspondences (PCs) for a generalized camera can be represented similarly by using $2n$ single cameras in a multi-camera system. It is a proven fact that the standard epipolar geometry using five PCs is incapable of recovering the scale of translation [18]. Conversely, the translation scale of multi-camera systems can be uniquely determined, and the minimum requirement for solving the relative pose increases from five to six PCs across both views [46].

Due to the presence of outliers in PCs, a robust estimator is essential for accurate relative pose estimation and outlier rejection. The random sample consensus (RANSAC) framework [9] and its various adaptations [1,2,30,43] are widely employed in computer vision community. A minimal solver is the core component in the RANSAC framework. Using the epipolar geometry corresponding to each PC, a constraint can be derived to solve for the relative pose [18]. Various methods for computing the relative pose of a single camera are known as five-point methods [8,25,26,34,40,44]. The six-point method [46] is the first minimal solver proposed for computing the relative pose of a multi-camera system. Numerous methods have been introduced subsequently, such as the enhanced version of the six-point method [5], the seventeen-point linear solvers [21, 35], an iterative optimization-based solver [23], and a global optimization-based solver [52].

This paper utilizes PCs to estimate the full DOF relative pose for multi-camera systems. The main contributions of this work are summarized as follows:

- We estimate 6DOF relative pose from a minimal number of six PCs for multi-camera systems. By decoupling rotation and translation, a generic minimal solver for the generalized camera and two minimal solvers for popular configurations of two-camera rigs are proposed.
- For multi-camera systems, when a subset of PCs relate the same cameras across two views, ray bundle constraints are found and proven. It can be seen that using ray bundle constraints reduces the number of solutions and generates numerically stable solvers for relative pose estimation.
- All configurations of minimal six-point problems for multi-camera systems are first enumerated using graph enumeration and the Pólya enumeration theorem. Totally there are 5953 cases. Moreover, we enumerate all the distinct graphs in a recursive way with 50 lines of Matlab code only.

2 Related Work

The research on relative pose estimation remains active in geometric vision, with a lot of classical solvers existing in this area. First, these solvers can be divided into the relative pose estimation for single cameras [8,18,25,26,34,40,44] and generalized cameras [21, 23, 35, 46, 52]. These cameras can be calibrated or partially uncalibrated with unknown focal length or radial distortion.

Second, the relative pose estimation solvers can be categorized as minimal solvers [40, 46], non-minimal solvers [21, 23, 24, 50, 52] and linear solvers [19, 35]. The minimal solvers aim to estimate relative pose using the minimum number of geometric primitives. The non-minimal solvers leverage all the feature correspondences to compute the relative pose. The linear solvers demand a higher number of feature correspondences compared to minimal solvers and provide a straightforward and efficient solution. To achieve computational efficiency, the linear solvers often ignore implicit constraints on unknown parameters. Conversely, the minimal and non-minimal solvers usually exploit implicit constraints while solving for unknown parameters. When dealing with feature correspondences that may contain outliers, minimal solvers play a crucial role in ensuring robust estimation. For instance, RANSAC and its variants depend heavily on efficient minimal solvers.

The integration of these two aspects yields numerous subcategories. This paper specifically investigates minimal solvers for multi-camera systems using PCs. A minimal solver with 64 solutions was first proposed to address the relative pose of multi-camera systems using 6 PCs [46]. They also propose a special case with 56 solutions when the extrinsic parameters satisfy a certain condition [46]. Then, a linear solver taking 17 PCs was introduced in [21,35]. Several solvers were proposed for applying to structure-from-motion with special configurations [20, 53]. Moreover, a few cases for semi-generalized cameras were proposed in [53], which is a rather small subset of our enumerated cases. Several non-minimal solvers were proposed which leverage either local optimization [23] or global optimization [52] to determine optimal relative poses. Some solvers required the motion priors of multi-camera systems, including known rotation axis [32, 37, 48] and Ackermann motion [31]. Moreover, an efficient solver was achieved by implementing a first-order approximation of relative rotation [49]. By exploiting the additional affine parameters besides PCs, a minimal solver with two affine correspondences was proposed in [14]. Three significant differences between [14] and our method are clarified in the supplementary material.

3 Relative Pose Estimation for Generalized Cameras

This section presents a novel six-point method for multi-camera systems and generalized cameras by decoupling rotation and translation. The ray bundle constraints are proven and exploited for solution space reduction. In addition, all configurations of minimal six-point problems for multi-camera systems are enumerated in this paper.

3.1 Geometric Constraints

Fig. 1 illustrates a multi-camera system consisting of multiple perspective cameras, assuming that both the intrinsic and extrinsic parameters of these cameras have been calibrated. The extrinsic parameters of the camera C_i are represented as $\{\mathbf{Q}_i, \mathbf{s}_i\}$, where $\mathbf{Q}_i$ represents the rotation and $\mathbf{s}_i$ represents the translation

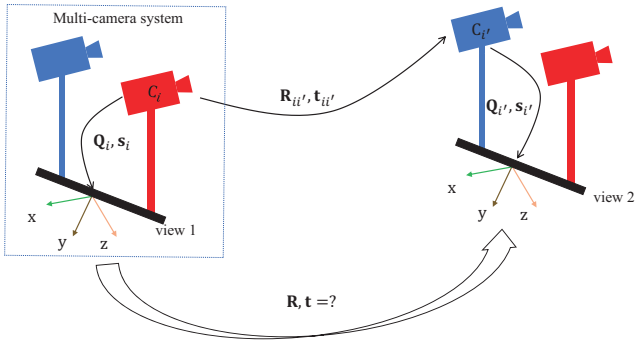


Fig. 1: Relative pose estimation for a multi-camera system. A point is observed by perspective camera C_i in view 1 and by camera $C_{i'}$ in view 2. $\{\mathbf{Q}_i, \mathbf{s}_i\}$ and $\{\mathbf{Q}_{i'}, \mathbf{s}_{i'}\}$ are extrinsic parameters for C_i and $C_{i'}$, respectively. The related point correspondence is described by two-view epipolar geometry of cameras C_i and $C_{i'}$.

with respect to the multi-camera system's reference. Let $\mathbf{R}$ denote the rotation and $\mathbf{t}$ denote the translation between the first and second views.

A PC in a multi-camera system establishes the relationship of a point captured by two cameras across different views. Let the k-th PC be denoted by $(\mathbf{x}_k, \mathbf{x}'_k, i_k, i'_k)$. This indicates that the i_k-th camera observes a point in view 1, which is represented by its homogeneous coordinate $\mathbf{x}_k$ in the normalized image plane. Furthermore, this same point is also observed by the i'_k-th camera in view 2, which is represented by its homogeneous coordinate as $\mathbf{x}'_k$. For simplicity, we omit the subscript k from camera indices i and i' to simplify the notation. In a multi-camera system, essential matrices vary for different PCs, distinguishing them from the two-view epipolar geometry of single cameras. Therefore, one constraint of epipolar geometry [18] induced by the k-th PC is

$$\mathbf{x}_{k'}^T \mathbf{E}_k \mathbf{x}_k = 0, \tag{1}$$

where the essential matrix is represented as

$$\mathbf{E}_k = [\mathbf{t}_{ii'}]_\times \mathbf{R}_{ii'}. \tag{2}$$

Here $\{\mathbf{R}_{ii'}, \mathbf{t}_{ii'}\}$ are the relative rotation and translation from camera i in the first view to camera i' in the second view. According to Fig. 1, it is obtained by a composition of three spatial transformations as

$$\begin{bmatrix} \mathbf{R}_{ii'} & \mathbf{t}_{ii'} \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{i'} & \mathbf{s}_{i'} \\ \mathbf{0} & 1 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{Q}_i & \mathbf{s}_i \\ \mathbf{0} & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_{i'}^T \mathbf{R} \mathbf{Q}_i & \mathbf{Q}_{i'}^T (\mathbf{R}\mathbf{s}_i + \mathbf{t} - \mathbf{s}_{i'}) \\ \mathbf{0} & 1 \end{bmatrix}. \tag{3}$$

By substituting $\mathbf{R}_{ii'}$ and $\mathbf{t}_{ii'}$ into Eq. (2), the essential matrix $\mathbf{E}_k$ can be reformulated as

$$\mathbf{E}_k = \mathbf{Q}_{i'}^T (\mathbf{R}[\mathbf{s}_i]_\times + [\mathbf{t} - \mathbf{s}_{i'}]_\times \mathbf{R}) \mathbf{Q}_i. \tag{4}$$

Based on the above equation, it can be seen that Eq. (1) is bilinear in the relative pose $\{\mathbf{R}, \mathbf{t}\}$.

3.2 Relative Pose Parameterization

We need to parametrize the relative pose of multi-camera systems. Rotation can be parameterized using Cayley parametrization, quaternions, Euler angles, direction cosine matrix (DCM), etc. Cayley and quaternion parameterizations have demonstrated superiority in minimal problems [51]. Rotation matrix $\mathbf{R}$ using Cayley parameterization can be expressed as

$$\mathbf{R}_{\text{cayl}} = \frac{1}{q_x^2 + q_y^2 + q_z^2 + 1} \begin{bmatrix} 1 + q_x^2 - q_y^2 - q_z^2 & 2q_xq_y - 2q_z & 2q_xq_z + 2q_y \\ 2q_xq_y + 2q_z & 1 - q_x^2 + q_y^2 - q_z^2 & 2q_yq_z - 2q_x \\ 2q_xq_z - 2q_y & 2q_yq_z + 2q_x & 1 - q_x^2 - q_y^2 + q_z^2 \end{bmatrix}, \quad (5)$$

where $[1, q_x, q_y, q_z]^T$ is a homogeneous quaternion vector. Rotation matrix $\mathbf{R}$ using quaternion parameterization can be written as

$$\mathbf{R}_{\text{quat}} = \begin{bmatrix} q_w^2 + q_x^2 - q_y^2 - q_z^2 & 2q_xq_y - 2q_wq_z & 2q_xq_z + 2q_wq_y \\ 2q_xq_y + 2q_wq_z & q_w^2 - q_x^2 + q_y^2 - q_z^2 & 2q_yq_z - 2q_wq_x \\ 2q_xq_z - 2q_wq_y & 2q_yq_z + 2q_wq_x & q_w^2 - q_x^2 - q_y^2 + q_z^2 \end{bmatrix}, \quad (6)$$

where $[q_w, q_x, q_y, q_z]^T$ is a quaternion vector satisfying the normalization constraint $q_w^2 + q_x^2 + q_y^2 + q_z^2 = 1$.

Note that 180° rotations are not allowed in Cayley parameterization, although this is uncommon for typical image pairs. In practical applications, Cayley parameterization has been extensively utilized in minimal problems [23, 45, 51, 53]. By contrast, quaternion parameterization does not have any degeneracy. However, quaternion introduces more variables than Cayley, so its solvers are usually less efficient. We construct solvers using both of these two parameterizations for completeness. In the following, we introduce the solver generation procedure based on Cayley parameterization. This approach can also be directly extended to quaternion parameterization.

The translation $\mathbf{t}$ can be parametrized as

$$\mathbf{t} = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T. \quad (7)$$

3.3 Equation System Construction

For multi-camera systems, the relative pose between two views is 6DOF. Thus, the relative pose estimation of a multi-camera system requires a minimal number of six PCs. Using Cayley parameterization, we obtain six polynomials for six unknowns $\{q_x, q_y, q_z, t_x, t_y, t_z\}$ from Eq. (1) by substituting the essential matrix (4) into them. After separating q_x, q_y, q_z from t_x, t_y, t_z, we arrive at an equation system

$$\underbrace{\mathbf{M}(q_x, q_y, q_z)}_{6\times 4} \begin{bmatrix} t_x \\ t_y \\ t_z \\ 1 \end{bmatrix} = \mathbf{0}, \quad (8)$$

where the entries of $\mathbf{M}$ are quadratic polynomials in three unknowns q_x, q_y, q_z. The i-th row corresponds to the constraint associated with the i-th PC. It can be observed that $\mathbf{M}$ has a null vector. Hence, the determinants of all the 4×4 submatrices of $\mathbf{M}$ must be zero.

Moreover, when a subset of PCs relates to the same perspective cameras across two views, there exists a specific property for this scenario. This property can be used to reduce the number of solutions and generate numerically stable solvers. Taking the number of a subset of PCs equal to 3 as an introductory example. When the number of a subset of PCs is greater than or equal to 3, it can be directly derived from Theorem 1.

Theorem 1. *Denote $\mathcal{S}$ as a matrix set with elements satisfying $\mathbf{N} = \mathbf{M}([k_1, k_2, k_3], 1:3)$, where $\mathbf{N}$ is formed from rows $\{k_1, k_2, k_3\}$ and columns $\{1, 2, 3\}$ of $\mathbf{M}$. In addition, k_1-th, k_2-th, and k_3-th PCs are captured by the same perspective camera in each view and $k_1 < k_2 < k_3$. Then $\mathrm{rank}(\mathbf{N}) = 2$, $\forall \mathbf{N} \in \mathcal{S}$ holds for non-degenerate cases.*

Proof. Let's proceed with the investigation of an arbitrary element in $\mathcal{S}$. We denote $\mathbf{N}_k$ as the k-th element in $\mathcal{S}$. The extrinsic parameters of the corresponding perspective camera in view 1 are denoted by $\{\mathbf{Q}_i, \mathbf{s}_i\}$, and for view 2, they are denoted by $\{\mathbf{Q}_{i'}, \mathbf{s}_{i'}\}$.

We begin by proving that $\mathrm{rank}(\mathbf{N}_k) \leq 2$. To accomplish this objective, we must prove that the null space of $\mathbf{N}_k$ is non-empty. Given that the k_1-th, k_2-th, and k_3-th PCs are observed by the same perspective camera in each view, their associated essential matrices remain unchanged. Referring to Eq. (4), the essential matrix can be parametrized as

$$\mathbf{E}_k = \mathbf{Q}_{i'}^T [\mathbf{t} + \mathbf{R}\mathbf{s}_i - \mathbf{s}_{i'}]_\times \mathbf{R}\mathbf{Q}_i. \tag{9}$$

Denote $\bar{\mathbf{t}} \triangleq \mathbf{t} + \mathbf{R}\mathbf{s}_i - \mathbf{s}_{i'}$, then we have

$$\mathbf{E}_k = \mathbf{Q}_{i'}^T [\bar{\mathbf{t}}]_\times \mathbf{R}\mathbf{Q}_i. \tag{10}$$

Substituting Eq. (10) into Eq. (1), we obtain three equations for the three PCs. Each monomial in the three equations is linear with respect to one entry of vector $\bar{\mathbf{t}}$, and there is no constant term. Hence, these equations can be expressed as

$$\frac{1}{q_x^2 + q_y^2 + q_z^2 + 1} \mathbf{A}_k \bar{\mathbf{t}} = \mathbf{0} \;\Rightarrow \mathbf{A}_k(\mathbf{t} + \mathbf{R}\mathbf{s}_i - \mathbf{s}_{i'}) = \mathbf{0}, \tag{11}$$

$$\Rightarrow \begin{bmatrix} \mathbf{A}_k & \mathbf{A}_k(\mathbf{R}\mathbf{s}_i - \mathbf{s}_{i'}) \end{bmatrix} \begin{bmatrix} \mathbf{t} \\ 1 \end{bmatrix} = \mathbf{0}. \tag{12}$$

By comparing the construction procedure of Eq. (8) and Eq. (12), we can see that

$$\mathbf{A}_k = \mathbf{M}([k_1, k_2, k_3], 1:3) = \mathbf{N}_k. \tag{13}$$

Substituting this equation into Eq. (11), we can observe that the null space of $\mathbf{N}_k$ is indeed not empty.

Next we aim to prove that $\text{rank}(\mathbf{N}_k) \geq 2$. This goal can be achieved using proof by contradiction. If $\text{rank}(\mathbf{N}_k) \leq 1$, then $\text{rank}(\mathbf{M}([k_1, k_2, k_3], 1:4)) \leq 2$ given that $\mathbf{M}([k_1, k_2, k_3], 1:4)$ includes an extra column compared to $\mathbf{N}_k$. This degenerate case implies that the three PCs offer no more than two independent constraints for the relative pose estimation. This does not hold true for non-degenerate cases, hence invalidating the assumption that $\text{rank}(\mathbf{N}_k) \leq 1$. It can be seen that the rank of $\mathbf{N}$ is 2 in non-degenerate cases. □

We call the constraints in Theorem 1 as the *ray bundle constraints*. In computer graphics, a ray bundle is a collection of light rays that share a common origin and propagate in different directions. In addition, a factor $q_x^2 + q_y^2 + q_z^2 + 1$ can be factored out to simplify the equation system. Please see the supplementary material for proof. It leads to the generation of more efficient solvers while also potentially avoiding extraneous roots. In summary, the whole polynomial equation system consists of two types of constraints, which can be written as

$$\mathcal{E}_1 \triangleq \{\text{quot}(\det(\mathbf{N}), q_x^2 + q_y^2 + q_z^2 + 1) = 0 \mid \mathbf{N} \in 4 \times 4 \text{ submatrices of } \mathbf{M}\}, \quad (14)$$

and

$$\mathcal{E}_2 \triangleq \{\text{quot}(\det(\mathbf{N}), q_x^2 + q_y^2 + q_z^2 + 1) = 0 \mid \mathbf{N} \in \mathcal{S}\}, \quad (15)$$

where $\text{quot}(a, b)$ denotes the division quotient of a by b, and $\det(\cdot)$ represents the determinant operator.

In $\mathcal{E}_1$, there are 15 equations of degree 6, while $\mathcal{E}_2$ contains some equations of degree 4. It should be noted that the number of equations in $\mathcal{E}_2$ varies depending on the specific configurations of the PCs. For certain PC configurations of multi-camera systems, $\mathcal{E}_2$ may be empty.

After obtaining the rotation parameters $\{q_x, q_y, q_z\}$, the translation $[t_x, t_y, t_z]^T$ can be determined by initially calculating a vector within the null space of $\mathbf{M}$, followed by normalization where the last entry of the vector is divided.

3.4 Polynomial System Solving

Based on the polynomial equation system Eqs. (14) and (15), we propose a minimal generic solver for the generalized camera and two minimal solvers for standard configurations of two-camera rigs. The Gröbner basis technique can be applied to discover algebraic solutions for the polynomial equation system [27, 39]. Firstly, we construct a random instance of the original equation system in either a finite prime field $\mathbb{Z}_p$ [36] or a rational field. This strategy helps to maintain numerical stability and avoid arithmetic with large numbers when computing Gröbner basis. Secondly, `Macaulay 2` [11] is utilized for computing Gröbner basis. Finally, we use an automatic Gröbner basis solver [27] to find the solution of

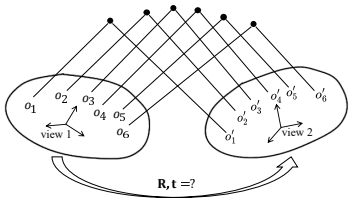


Fig. 2: Relative pose estimation for generalized cameras. Note that points o_i and o'_i do not necessarily correspond to the same physical point of the generalized camera.

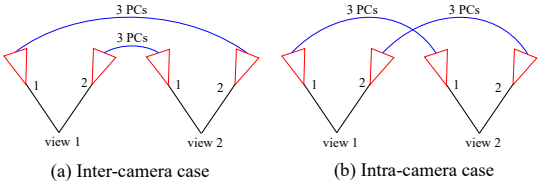


Fig. 3: Relative pose estimation for two-camera rigs. Specifically, our goal is to determine the 6DOF relative pose while six PCs are observable by two views of a two-camera rig. (a) inter-camera case, (b) intra-camera case.

the polynomial equation system. It should be noted that the polynomial equations $\mathcal{E}_1$ and $\mathcal{E}_2$ can be extended to deal with relative pose estimation for other configurations of multi-camera systems, such as with partially uncalibrated cameras and known rotation angles.

Minimal Solver for Generalized Cameras We propose a generic solver for the generalized camera given six PCs. As shown in Fig. 2, there are two views of a generalized camera, and there are 6 PCs across two views. Specifically, two-view geometry for 6 PCs and a generalized camera can be described by a 12-camera rig. We can define 12 virtual perspective cameras by the following method. The origins $o_1, \cdots, o_6$ and $o'_1, \cdots, o'_6$ of the virtual cameras are the positions of PCs. The orientations of these virtual cameras are consistent with the generalized camera's reference. The 6 PCs can be equivalently captured by a virtual 12-camera rig. Moreover, we can calculate the extrinsic parameters of these 12 virtual perspective cameras and the image coordinates of PCs in these virtual cameras. Please see the supplementary material for details. As a result, we can construct a minimal generic solver to recover the relative pose of generalized cameras.

Tab. 1 shows the statistics of the proposed minimal solvers for generalized cameras. Here, `#sym` indicates the number of symmetries, `#sol` indicates the number of solutions, and 1-`dim` indicates one-dimensional extraneous roots. The Cayley parameterization and quaternion parameterization solvers are named as `6pt+cayl+generic` and `6pt+quat+generic`, respectively. The observations are summarized as follows: (1) $\mathcal{E}_2$ is an empty set, and $\mathcal{E}_1$ is sufficient to solve the relative pose. (2) Due to one-fold symmetry in quaternion, Cayley parameterization results in fewer solutions than quaternion parameterization. The number of complex solutions yielded by the `6pt+cayl+generic` solver and `6pt+quat+generic` solver is 64 and 128, respectively. (3) Elimination templates of the `6pt+cayl+generic` solver and the `6pt+quat+generic` solver are 99×163 and 342×406, respectively. Given that the solvers using Cayley parameterization yields smaller eliminate templates compared to the solvers using quaternion parameterization, we adopt the former as our default choice in this paper.

Table 1: Minimal solvers for relative pose estimation. cayl: Cayley parameterization; quat: quaternion parameterization; inter: inter-camera PCs; intra: intra-camera PCs. For the generic case, $\mathcal{E}_2$ is an empty set and $\mathcal{E}_1 \cup \mathcal{E}_2 = \mathcal{E}_1$.

configuration	equations $\mathcal{E}_1$			equations $\mathcal{E}_1 \cup \mathcal{E}_2$		
	#sym	#sol	template	#sym	#sol	template
6pt+cayl+generic	0	64	99×163	0	64	99×163
6pt+cayl+inter	0	56	56×120	0	48	64×120
6pt+cayl+intra	0	1-dim	–	0	48	72×120
6pt+quat+generic	1	128	342×406	1	128	342×406
6pt+quat+inter	1	112	174×243	1	96	152×200
6pt+quat+intra	1	1-dim	–	1	96	152×200

Moreover, we enumerate all configurations of minimal six-point problems for multi-camera systems. The Pólya enumeration theorem can be applied to solve this problem [17], and a combinatorics solution shows that there are 5953 cases totally in this problem. Most of these cases can be solved by the generic solver. Please see the supplementary material for details.

Minimal Solvers for Two-camera Rigs Two minimal solvers are proposed for two practical configurations of two-camera rigs in Fig. 3. These solvers comprise an inter-camera solver and an intra-camera solver, and both configurations offer two ray bundle constraints within $\mathcal{E}_2$. The inter-camera solver utilizes inter-camera PCs that are observable to different cameras across two views. This solver is appropriate for multi-camera systems characterized by significant overlap between views. Conversely, the intra-camera solver utilizes intra-camera PCs that are observable to the same camera across two views. This solver is appropriate for multi-camera systems characterized by small or no overlap between views.

For inter-camera case, $\mathcal{E}_1$ is enough to compute the relative pose of two-camera rigs. The number of solutions can be reduced by employing both $\mathcal{E}_1$ and $\mathcal{E}_2$. For intra-camera case, one-dimensional families of extraneous roots exist when only $\mathcal{E}_1$ is employed. Combining $\mathcal{E}_1$ and $\mathcal{E}_2$ can solve the relative pose in the intra-camera case. The solvers for these two cases using Cayley parametrization are named as `6pt+cayl+inter` and `6pt+cayl+intra`, respectively. The solvers for these two cases using quaternion parameterization are named as `6pt+quat+inter` and `6pt+quat+intra`, respectively.

Tab. 1 also shows the statistics of our minimal solvers for two-camera rigs. The observations are summarized as follows: (1) When $\mathcal{E}_1$ is employed, the inter-camera solver has up to 56 complex solutions, and there exist one-dimensional families of extraneous roots in the intra-camera case. (2) By employing both $\mathcal{E}_1$ and $\mathcal{E}_2$, both the inter-camera and intra-camera solvers yield a total of 48 complex solutions. (3) Compared to the solvers using quaternion parameterization, the solvers using Cayley parameterization yield smaller eliminate templates. We adopt Cayley parameterization as our default choice. (4) When only $\mathcal{E}_1$ is used in the `6pt+quat+inter` configuration, it is necessary to explicitly consider the inequality $q_w \neq 0$. Otherwise, one-dimensional extraneous roots exist. We ac-

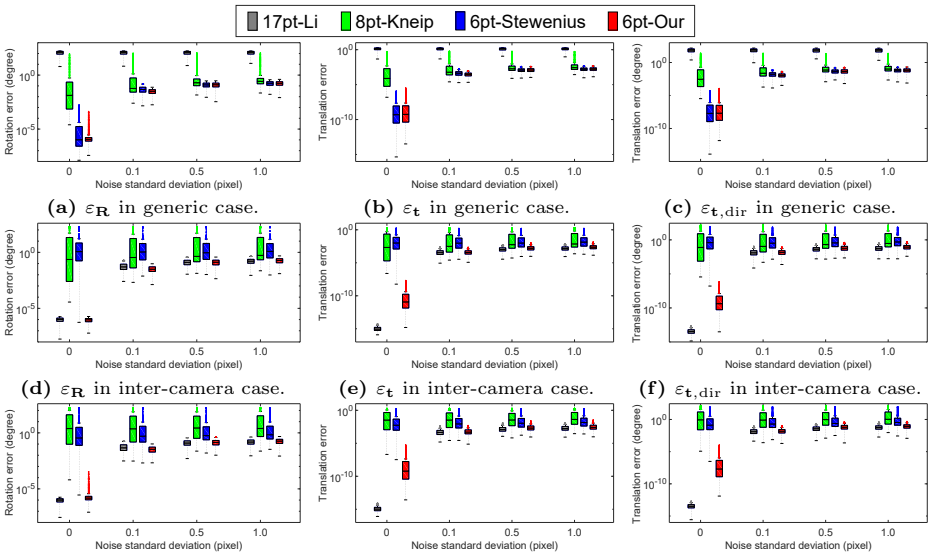
count for this inequality utilizing the saturation method [28], which yields 112 solutions exhibiting one-fold symmetry. (5) For the `6pt+cayl+inter` configuration, compared to the solvers derived from both $\mathcal{E}_1$ and $\mathcal{E}_2$, the solvers solely derived from $\mathcal{E}_1$ yield smaller eliminate templates and exhibit better numerical stability. This phenomenon indicates that the number of bases might impact the numerical stability of the solvers, which has been previously observed in the literature [5, 14].

4 Experiments

This section presents a series of experiments performed on synthetic and real-world datasets to assess the performance of our solvers. All the proposed solvers employ Cayley parameterization for their implementations. The solver for the generalized camera is named as the `6pt-Our-generic` method. The solvers aim to handle inter-camera and intra-camera cases, are named as the `6pt-Our-inter` and `6pt-Our-intra` solvers, respectively. To further differentiate between different solvers for `6pt-Our-inter`, we denote the solvers resulting from $\mathcal{E}_1$ and $\mathcal{E}_1 \cup \mathcal{E}_2$ as `6pt-Our-inter56` and `6pt-Our-inter48`, respectively. Following [14, 23], our solvers are evaluated against state-of-the-art solvers using PCs, including `17pt-Li` [35], `8pt-Kneip` [23], and `6pt-Stewénius` [46]. The paper does not evaluate solvers that exploit the additional affine parameters besides PCs [14, 15] or utilize a prior for relative pose estimation [48, 49]. The proposed minimal solvers are implemented in C++. The source codes for `17pt-Li` [35] and `8pt-Kneip` [23] are adopted from the OpenGV library [22]. The source codes for `6pt-Stewénius` [46] is adopted from the PoseLib library [29].

In the experiments presented in Sec. 4.1 and Sec. 4.2, each solver is independently integrated into RANSAC [9] to reject outliers and obtain the estimated relative pose with the highest number of inliers. To ensure the fairness of the experiment, all the different solvers are compared over the same data and within the same RANSAC implementation. In addition, we do not apply a non-minimal solver or perform optimization for the estimated relative pose with all inliers. The angular re-projection error [22, 33] induced by PCs is used to classify inliers. We follow the default parameters of OpenGV to set the inlier threshold angle as 0.1° [22]. During RANSAC iterations, the outlier ratio is from the current best model. The stopping criterion of RANSAC iterations is that at least one outlier-free set is sampled with a probability 0.99, or the maximum number $20,000$ of iterations is reached. We compute the rotation error as the angular difference between the ground truth rotation $\mathbf{R}_{\text{gt}}$ and the estimated rotation $\mathbf{R}$: $\varepsilon_{\mathbf{R}} = \arccos((\text{tr}(\mathbf{R}_{\text{gt}}\mathbf{R}^T) - 1)/2)$. To calculate the translation error, we employ the definition introduced in [32,42]: $\varepsilon_{\mathbf{t}} = 2\,\|\mathbf{t}_{\text{gt}} - \mathbf{t}\|\,/(\|\mathbf{t}_{\text{gt}}\| + \|\mathbf{t}\|)$, where $\mathbf{t}_{\text{gt}}$ and $\mathbf{t}$ are the ground truth translation and the estimated translation, respectively. We also compute the translation direction error: $\varepsilon_{\mathbf{t},\text{dir}} = \arccos(\mathbf{t}_{\text{gt}}^T\mathbf{t}/(\|\mathbf{t}_{\text{gt}}\| \cdot \|\mathbf{t}\|))$.

The principles of minimal solver choice are listed below. First, when `6pt-Our-inter` or `6pt-Our-intra` are applicable, we apply them with higher priority than `6pt-Our-generic`. The reason is that the solvers of specific con-



(a) $\varepsilon_{\mathbf{R}}$ in generic case. **(b)** $\varepsilon_{\mathbf{t}}$ in generic case. **(c)** $\varepsilon_{\mathbf{t},\mathrm{dir}}$ in generic case.

(d) $\varepsilon_{\mathbf{R}}$ in inter-camera case. **(e)** $\varepsilon_{\mathbf{t}}$ in inter-camera case. **(f)** $\varepsilon_{\mathbf{t},\mathrm{dir}}$ in inter-camera case.

(g) $\varepsilon_{\mathbf{R}}$ in intra-camera case. **(h)** $\varepsilon_{\mathbf{t}}$ in intra-camera case. **(i)** $\varepsilon_{\mathbf{t},\mathrm{dir}}$ in intra-camera case.

Fig. 4: Relative pose estimation error with varying image noise for a generalized camera. We design a simulated multi-camera system comprising 12 omnidirectional cameras. The extrinsic parameters, including orientation and position, are totally random. The three rows correspond to the generic, inter-camera, and intra-camera cases, respectively. The `6pt-Our` solver in the three rows represent `6pt-Our-generic`, `6pt-Our-inter`, and `6pt-Our-intra`, respectively.

figurations usually have better performance than the generic solver. Second, we recommend `6pt-Our-intra` as the default solver for real-world image sequences. The relative pose is usually estimated for consecutive image pairs captured in a small time interval. Thus, intra-camera PCs can be built up with a higher probability than inter-camera PCs. The results for efficiency and numerical stability are shown in the supplementary material.

4.1 Experiments on Synthetic Data

Two simulated scenarios are designed and tested for the synthetic experiments. First, we design a simulated two-camera rig composed of two perspective cameras. The orientations of the two perspective cameras are roughly forward-facing with random perturbation. This setting is practical for autonomous driving with two front-facing cameras. Second, we design a simulated generalized camera composed of 12 omnidirectional cameras. The extrinsic parameters, including orientation and position, are totally random. The results for two-camera rigs under the first scenario are shown in the supplementary material.

Under the second scenario, the simulated scenario for a generalized camera is described as below. For the generic case, a generalized camera comprises

12 omnidirectional cameras. For the inter-camera and intra-camera cases, we only use 2 omnidirectional cameras of the generalized camera. The extrinsic parameters of each omnidirectional camera are generated randomly. Each scene point is generated randomly. In addition, the relative poses between two views are also generated randomly. Omnidirectional cameras are selected in our settings, because usually there is no overlap for pinhole cameras with random extrinsic parameters and relative poses. We test the accuracy of pose estimation for all the proposed solvers, including `6pt-Our-generic`, `6pt-Our-inter`, and `6pt-Our-intra` solvers.

Some of the comparison solvers use more PCs than the proposed solvers. The strategy of correspondence selection has a significant influence on their performance. We design a rule to select matches for different cases to guarantee fairness. Please see the supplementary material for details. In the synthetic experiments, we conduct 1000 trials for each case and a specific noise level combined with the RANSAC framework. For each trial, we randomly generate 100 PCs and select correspondences randomly for the different solvers. Fig. 4 illustrates the performance of various solvers against image noise for generalized cameras. The observations are summarized as follows: (1) `17pt-Li` has good overall accuracy for inter-camera and intra-camera cases. However, it fails for the generic case due to rank deficiency, and the essential matrix with scale ambiguity cannot be uniquely recovered. (2) `8pt-Kneip` has acceptable results according to the median metric. However, the error variance is large for all the cases. (3) `6pt-Stewénius` has good overall accuracy for the generic case. It does not work well for inter-camera and intra-camera cases, and the error variance is large for both cases. This phenomenon is consistent with [21], which observes this solver does not work for most axial cameras where every bearing vector intersects a line in 3D. (4) The proposed `6pt-Our` solver works for all the cases and has satisfactory overall accuracy.

4.2 Experiments on Real-World Data

To assess the performance of our solvers in practical applications, three datasets `KITTI` [10], `nuScenes` [6], and `EuRoc` [4] are used in the experiments. Specifically, the `KITTI` and `nuScenes` datasets are collected in an autonomous driving environments, while the `EuRoc` dataset is collected in a micro aerial vehicle environment. These datasets contain challenging image pairs with highly dynamic scenes, such as significant motion, moving objects, and varying illumination. The proposed solvers are compared against state-of-the-art solvers. The accuracy of relative pose estimation is evaluated using the rotation error $\varepsilon_{\mathbf{R}}$ and the translation direction error $\varepsilon_{\mathbf{t},\text{dir}}$ [23, 37]. Our evaluation is based on approximately $30,000$ image pairs, and we report the final estimation results by integrating the minimal solver with RANSAC. The relative pose estimation results for the `nuScenes` dataset are provided in the supplementary material.

Experiments on KITTI Dataset We evaluate the proposed solvers on `KITTI` dataset [10], which is collected using outdoor autonomous vehicles installed with

Table 2: Rotation and translation error on `KITTI` dataset (unit: degree).

Seq.	17pt-Li [35]		8pt-Kneip [23]		6pt-Stew. [46]		6pt-Our-intra	
	$\varepsilon_{\mathbf{R}}$	$\varepsilon_{\mathbf{t},\text{dir}}$	$\varepsilon_{\mathbf{R}}$	$\varepsilon_{\mathbf{t},\text{dir}}$	$\varepsilon_{\mathbf{R}}$	$\varepsilon_{\mathbf{t},\text{dir}}$	$\varepsilon_{\mathbf{R}}$	$\varepsilon_{\mathbf{t},\text{dir}}$
00	0.147	2.537	0.148	2.496	0.243	4.521	**0.136**	**2.415**
01	0.178	4.407	0.182	3.485	0.293	7.187	**0.175**	**3.323**
02	0.142	1.988	0.147	2.094	0.227	3.315	**0.139**	**1.897**
03	**0.126**	2.762	0.139	2.833	0.314	6.254	0.143	**2.740**
04	0.113	1.733	0.123	1.829	0.262	3.670	**0.101**	**1.677**
05	0.132	2.663	0.130	2.461	0.216	4.212	**0.128**	**2.342**
06	0.139	2.146	0.151	2.145	0.197	3.240	**0.121**	**2.064**
07	0.131	3.085	0.172	3.259	0.259	6.664	**0.129**	**2.904**
08	**0.133**	2.705	0.135	2.762	0.217	4.590	0.140	**2.620**
09	0.144	2.022	0.138	**1.974**	0.210	3.204	**0.126**	2.002
10	0.142	2.398	0.141	2.393	0.246	3.849	**0.137**	**2.314**

Table 3: Average runtime of RANSAC on `KITTI` dataset (unit: s).

Methods	17pt-Li [35]	8pt-Kneip [23]	6pt-Stew. [46]	6pt-Our-intra
Mean time	3.157	0.648	4.161	1.546
Std. deviation	0.119	0.009	0.145	0.069

forward-facing stereo cameras. It is treated as a general multi-camera system, disregarding overlapping overlapping fields of view for cameras. The `6pt-Our-intra` solver is tested on 11 available sequences containing 23,000 image pairs. The ground truth is obtained directly from the GPS/IMU localization unit [10]. To establish PCs for consecutive views of each camera, the SIFT method [38] is used. Additionally, all the solvers have been integrated into RANSAC to remove mismatches in the experiments.

Tab. 2 illustrates the rotation and translation error of the `6pt-Our-intra` solver on the `KITTI` dataset. We use median error to evaluate the performance of solvers. The proposed `6pt-Our-intra` solver outperforms the comparative solvers in overall performance. Since `8pt-Kneip` uses an identity matrix to initialize the rotation, it has acceptable results for the forward motion in autonomous driving environment. To further compare computational efficiency, Tab. 3 illustrates the corresponding average runtime of RANSAC on the `KITTI` dataset. Although the runtime of `17pt-Li` is lower than the proposed `6pt-Our-intra`, the proposed solver demonstrates better efficiency when each is integrated separately into the RANSAC framework.

Experiments on EuRoC Dataset The proposed solvers are further validated within the context of an unmanned aerial vehicle (UAV) environment, leveraging the `EuRoC` MAV dataset [4] for the evaluation of 6DOF relative pose estimation. This dataset records data using a stereo camera mounted on a micro

Table 4: Rotation and translation error on `EuRoC` dataset (unit: degree).

Seq.	17pt-Li [35]		8pt-Kneip [23]		6pt-Stew. [46]		6pt-Our-intra	
	$\varepsilon_{\mathbf{R}}$	$\varepsilon_{\mathbf{t},\text{dir}}$	$\varepsilon_{\mathbf{R}}$	$\varepsilon_{\mathbf{t},\text{dir}}$	$\varepsilon_{\mathbf{R}}$	$\varepsilon_{\mathbf{t},\text{dir}}$	$\varepsilon_{\mathbf{R}}$	$\varepsilon_{\mathbf{t},\text{dir}}$
MH01	0.136	3.055	0.156	3.214	0.186	4.085	**0.130**	**2.961**
MH02	0.129	2.806	0.132	2.796	0.180	3.828	**0.127**	**2.579**
MH03	0.199	2.422	0.187	2.517	0.222	3.576	**0.181**	**2.376**
MH04	0.195	3.159	0.178	3.237	0.213	5.371	**0.193**	**3.105**
MH05	0.186	3.124	0.163	2.940	0.235	4.094	**0.158**	**2.892**

aerial vehicle, with sequences labeled from MH01 to MH05 collected within a large industrial machine hall. The ground truth for the relative pose is established through a nonlinear least-squares batch solution utilizing Leica position and IMU measurements. Estimating relative pose estimation in these sequences is challenging due to the unstructured and cluttered nature of the industrial environment. Consecutive image pairs with small movement are selectively thinned out, retaining only one out of every four consecutive images for relative pose estimation. Moreover, the image pairs exhibiting insufficient motion are excluded from the experiment. The proposed solvers are evaluated against state-of-the-art solvers, including `17pt-Li` [35], `8pt-Kneip` [23], and `6pt-Stewénius` [46]. PCs in the image pair are established using the SIFT detector [38]. All solvers are integrated into the RANSAC framework to eliminate mismatches, ensuring a more robust estimation process.

Tab. 4 illustrates the rotation and translation error of the `6pt-Our-intra` solver on `EuRoC` dataset. We use median error to assess the performance of different solvers. The experiment results demonstrate that the `6pt-Our-intra` solver surpasses comparative solvers, including `17pt-Li`, `8pt-Kneip`, and `6pt-Stewénius`. This experiment confirms the suitability of the proposed `6pt-Our-intra` solver for achieving accurate 6DOF relative pose estimation in the context of unmanned aerial vehicle environments.

5 Conclusion

We proposed a series of minimal solvers to compute the 6DOF relative pose of multi-camera systems using a minimal number of six PCs. We also exploit ray bundle constraints that allow for a reduction of the solution space and the development of more stable solvers. A generic solver is proposed for the relative pose estimation of generalized cameras. All configurations of minimal six-point problems for multi-camera systems are enumerated. Moreover, two minimal solvers, including an inter-camera solver and an intra-camera solver, are proposed for practical configurations of two-camera rigs. Based on both synthetic and real-world experiments, we demonstrate that the proposed solvers offer an efficient solution for estimating ego-motion of multi-camera systems, surpassing state-of-the-art solvers in terms of accuracy.

Acknowledgments

This research has been supported in part by the Hunan Provincial Natural Science Foundation for Excellent Young Scholars under Grant 2023JJ20045, and the National Natural Science Foundation of China under Grant 12372189. Further funding support is provided by project 62250610225 by the Natural Science Foundation of China, as well as projects 22DZ1201900, 22ZR1441300, and dfycbj-1 by the Natural Science Foundation of Shanghai. J. Zhao would like to thank Saibal Mitra, from the Netherlands, for the valuable discussions on graph enumeration.

References

1. Barath, D., Matas, J.: Graph-cut RANSAC: Local optimization on spatially coherent structures. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(9), 4961–4974 (2022)
2. Barath, D., Noskova, J., Matas, J.: Marginalizing sample consensus. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(11), 8420–8432 (2022)
3. Barath, D., Polic, M., FÃűrstner, W., Sattler, T., Pajdla, T., Kukelova, Z.: Making affine correspondences work in camera geometry computation. In: European Conference on Computer Vision. pp. 723–740 (2020)
4. Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M.W., Siegwart, R.: The EuRoC micro aerial vehicle datasets. The International Journal of Robotics Research **35**(10), 1157–1163 (2016)
5. Byröd, M., Josephson, K., Aström, K.: Fast and stable polynomial equation solving and its application to computer vision. International Journal of Computer Vision **84**(3), 237–256 (2009)
6. Caesar, H., Bankiti, V., Lang, A.H., Vora, S., Liong, V.E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O.: nuScenes: A multimodal dataset for autonomous driving. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 11621–11631 (2020)
7. Eichhardt, I., Barath, D.: Relative pose from deep learned depth and a single affine correspondence. In: European Conference on Computer Vision. pp. 627–644 (2020)
8. Fathian, K., Ramirez-Paredes, J.P., Doucette, E.A., Curtis, J.W., Gans, N.R.: QuEst: A quaternion-based approach for camera motion estimation from minimal feature points. IEEE Robotics and Automation Letters **3**(2), 857–864 (2018)
9. Fischler, M.A., Bolles, R.C.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM **24**(6), 381–395 (1981)
10. Geiger, A., Lenz, P., Stiller, C., Urtasun, R.: Vision meets robotics: The KITTI dataset. The International Journal of Robotics Research **32**(11), 1231–1237 (2013)
11. Grayson, D.R., Stillman, M.E.: Macaulay 2, a software system for research in algebraic geometry. `https://faculty.math.illinois.edu/Macaulay2/` (2002)
12. Grossberg, M.D., Nayar, S.K.: A general imaging model and a method for finding its parameters. In: IEEE International Conference on Computer Vision. vol. 2, pp. 108–115. IEEE (2001)
13. Guan, B., Vasseur, P., Demonceaux, C., Fraundorfer, F.: Visual odometry using a homography formulation with decoupled rotation and translation estimation using minimal solutions. In: IEEE International Conference on Robotics and Automation. pp. 2320–2327 (2018)

14. Guan, B., Zhao, J.: Affine correspondences between multi-camera systems for 6DOF relative pose estimation. In: European Conference on Computer Vision. pp. 634–650. Springer Nature Switzerland, Cham (2022)
15. Guan, B., Zhao, J., Barath, D., Fraundorfer, F.: Efficient recovery of multi-camera motion from two affine correspondences. In: IEEE International Conference on Robotics and Automation. pp. 1305–1311 (2021)
16. Guan, B., Zhao, J., Barath, D., Fraundorfer, F.: Minimal solvers for relative pose estimation of multi-camera systems using affine correspondences. International Journal of Computer Vision **131**(1), 324–345 (2023)
17. Guichard, D.: Combinatorics and Graph Theory. LibreTexts (2023)
18. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge University Press (2003)
19. Hartley, R.I.: In defense of the eight-point algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence **19**(6), 580–593 (1997)
20. Kasten, Y., Galun, M., Basri, R.: Resultant based incremental recovery of camera pose from pairwise matches. In: IEEE Winter Conference on Applications of Computer Vision. pp. 1080–1088 (2019)
21. Kim, J.H., Li, H., Hartley, R.: Motion estimation for nonoverlapping multicamera rigs: Linear algebraic and L_∞ geometric solutions. IEEE Transactions on Pattern Analysis and Machine Intelligence **32**(6), 1044–1059 (2009)
22. Kneip, L., Furgale, P.: OpenGV: A unified and generalized approach to real-time calibrated geometric vision. In: IEEE International Conference on Robotics and Automation. pp. 12034–12043 (2014)
23. Kneip, L., Li, H.: Efficient computation of relative pose for multi-camera systems. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 446–453 (2014)
24. Kneip, L., Lynen, S.: Direct optimization of frame-to-frame rotation. In: IEEE International Conference on Computer Vision. pp. 2352–2359 (2013)
25. Kneip, L., Siegwart, R., Pollefeys, M.: Finding the exact rotation between two images independently of the translation. In: European Conference on Computer Vision. pp. 696–709. Springer (2012)
26. Kukelova, Z., Bujnak, M., Pajdla, T.: Polynomial eigenvalue solutions to minimal problems in computer vision. IEEE Transactions on Pattern Analysis and Machine Intelligence **34**(7), 1381–1393 (2012)
27. Larsson, V., Aström, K., Oskarsson, M.: Efficient solvers for minimal problems by syzygy-based reduction. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 820–828 (2017)
28. Larsson, V., Aström, K., Oskarsson, M.: Polynomial solvers for saturated ideals. In: IEEE International Conference on Computer Vision. pp. 2288–2297 (2017)
29. Larsson, V., contributors: PoseLib - Minimal Solvers for Camera Pose Estimation (2020), `https://github.com/vlarsson/PoseLib`
30. Lebeda, K., Matas, J., Chum, O.: Fixing the locally optimized RANSAC. In: British Machine Vision Conference (2012)
31. Lee, G.H., Faundorfer, F., Pollefeys, M.: Motion estimation for self-driving cars with a generalized camera. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 2746–2753 (2013)
32. Lee, G.H., Pollefeys, M., Fraundorfer, F.: Relative pose estimation for a multi-camera system with known vertical direction. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 540–547 (2014)

33. Lee, S.H., Civera, J.: Closed-form optimal two-view triangulation based on angular errors. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2681–2689 (2019)
34. Li, H., Hartley, R.: Five-point motion estimation made easy. In: International Conference on Pattern Recognition. pp. 630–633 (2006)
35. Li, H., Hartley, R., Kim, J.h.: A linear approach to motion estimation using generalized camera models. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–8 (2008)
36. Lidl, R., Niederreiter, H.: Finite Fields. Cambridge University Press (1997)
37. Liu, L., Li, H., Dai, Y., Pan, Q.: Robust and efficient relative pose with a multi-camera system for autonomous driving in highly dynamic environments. IEEE Transactions on Intelligent Transportation Systems **19**(8), 2432–2444 (2017)
38. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60**(2), 91–110 (2004)
39. Martyushev, E., Vráblíková, J., Pajdla, T.: Optimizing elimination templates by greedy parameter search. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 15754–15764 (2022)
40. Nistér, D.: An efficient solution to the five-point relative pose problem. IEEE Transactions on Pattern Analysis and Machine Intelligence **26**(6), 756–777 (2004)
41. Pless, R.: Using many cameras as one. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–7 (2003)
42. Quan, L., Lan, Z.: Linear n-point camera pose determination. IEEE Transactions on Pattern Analysis and Machine Intelligence **21**(8), 774–780 (1999)
43. Raguram, R., Chum, O., Pollefeys, M., Matas, J., Frahm, J.M.: USAC: A universal framework for random sample consensus. IEEE Transactions on Pattern Analysis and Machine Intelligence **35**(8), 2022–2038 (2012)
44. Stewénius, H., Engels, C., Nistér, D.: Recent developments on direct relative orientation. ISPRS Journal of Photogrammetry and Remote Sensing **60**(4), 284–294 (2006)
45. Stewénius, H., Nistér, D., Kahl, F., Schaffalitzky, F.: A minimal solution for relative pose with unknown focal length. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 789–794 (2005)
46. Stewénius, H., Oskarsson, M., Aström, K., Nistér, D.: Solutions to minimal generalized relative pose problems. In: Workshop on Omnidirectional Vision in conjunction with ICCV. pp. 1–8 (2005)
47. Sturm, P., Ramalingam, S.: A generic concept for camera calibration. In: European Conference on Computer Vision. vol. 3022, pp. 1–13. Springer-Verlag (2004)
48. Sweeney, C., Flynn, J., Turk, M.: Solving for relative pose with a partially known rotation is a quadratic eigenvalue problem. In: International Conference on 3D Vision. pp. 483–490 (2014)
49. Ventura, J., Arth, C., Lepetit, V.: An efficient minimal solution for multi-camera motion. In: IEEE International Conference on Computer Vision. pp. 747–755 (2015)
50. Zhao, J.: An efficient solution to non-minimal case essential matrix estimation. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(4), 1777–1792 (2022)
51. Zhao, J., Kneip, L., He, Y., Ma, J.: Minimal case relative pose computation using ray-point-ray features. IEEE Transactions on Pattern Analysis and Machine Intelligence **42**(5), 1176–1190 (2020)

52. Zhao, J., Xu, W., Kneip, L.: A certifiably globally optimal solution to generalized essential matrix estimation. In: IEEE Conference on Computer Vision and Pattern Recognition. pp. 12034–12043 (2020)
53. Zheng, E., Wu, C.: Structure from motion using structure-less resection. In: IEEE International Conference on Computer Vision. pp. 2075–2083 (2015)