This CVPR 2020 paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

A Certifiably Globally Optimal Solution to Generalized Essential Matrix Estimation

Ji Zhao
TuSimple
zhaoji84@gmail.com

Wanting Xu
ShanghaiTech University
xuwt@shanghaitech.edu.cn

Laurent Kneip
ShanghaiTech University
lkneip@shanghaitech.edu.cn

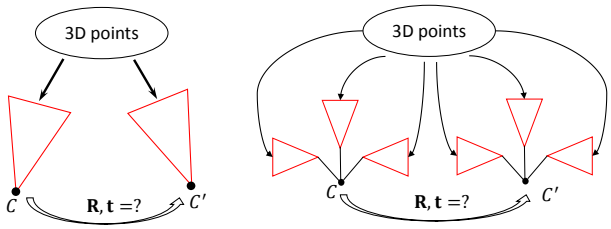
Abstract

We present a convex optimization approach for generalized essential matrix (GEM) estimation. The six-point minimal solver for the GEM has poor numerical stability and applies only for a minimal number of points. Existing non-minimal solvers for GEM estimation rely on either local optimization or relinearization techniques, which impedes high accuracy in common scenarios. Our proposed non-minimal solver minimizes the sum of squared residuals by reformulating the problem as a quadratically constrained quadratic program. The globally optimal solution is thus obtained by a semidefinite relaxation. The algorithm retrieves certifiably globally optimal solutions to the original non-convex problem in polynomial time. We also provide the necessary and sufficient conditions to recover the optimal GEM from the relaxed problems. The improved performance is demonstrated over experiments on both synthetic and real multi-camera systems.

1. Introduction

Relative pose estimation from images plays an important role in many geometric vision tasks, such as structure-from-motion (SfM) and simultaneous localization and mapping (SLAM). While *central cameras* can be modeled by the pinhole or perspective camera model [11], more general *non-central cameras* such as multi-camera arrays are modelled by the generalized camera model [32]. This paper presents a new method to estimate the generalized essential matrix (GEM) or relative pose for non-central cameras.

The essential matrix encodes the relative pose for pinhole cameras and is well understood [11, 30, 4]. GEM estimation is more involved. A generalized camera is formed by abstracting landmark observations into spatial rays that are no longer required to originate from a common point (i.e. the focal point). Figure 1 demonstrates the difference between central and non-central cameras. As illustrated, the generalized camera model allows us to describe the measurements of a number of interesting camera systems, such as a multi-camera rig of rigidly attached cameras.



(a) central camera (b) non-central camera

Figure 1. Relative pose estimation. Red triangles represent perspective cameras. Solid arrows pointing from 3D points to cameras depict the imaging process. In the non-central camera scenario, three rigidly attached central cameras constitute a non-central camera.

From a more abstract and geometric point of view, a generalized camera consists of a Euclidean reference frame in which measurements are represented by rays in space, described by a suitable parameterization such as Plücker line vectors. In contrast to the standard essential matrix for which there exists an unobservability in the norm of the translation, the translation extracted from a GEM is generally unique. The down-side of this scale observability is that the minimal solution of the GEM requires at least 6 instead of only 5 correspondences across the two views (i.e. one correspondence per degree of freedom in the problem).

There are both linear [24] and non-linear solutions [34, 18, 5] to GEM estimation. The linear solver—also known as the 17-point algorithm—takes 17 correspondences to derive the relative pose of the generalized camera. This method can be easily applied to an arbitrarily large number of points. However, its solution is not globally optimal, as the linearization ignores side-constraints on the GEM and the contained essential and rotation matrices. The most closely related works to ours are the non-minimal solvers by Kneip and Li [18] and Campos *et al.* [5], which use many correspondences to calculate a potentially accurate relative pose, but rely on local optimization methods and therefore

may depend on a sufficiently accurate initial guess. They do not guarantee global optimality.

By contrast, the present paper leverages convex optimization to—for the first time—come up with a fast and certifiably globally optimal solution to the non-minimal generalized relative pose problem, whose optimality may be certified *a-posteriori*. In summary, the contribution of this paper is two-fold:

- **Formulation**. We propose a novel formulation for GEM estimation of generalized cameras, and discuss its relation to the previously proposed eigenvalue-based formulation in [18].

- **Optimization**. We provide a certifiably globally optimal solution by semidefinite relaxation (SDR) of the original formulation. We also provide a sufficient and necessary condition to recover the optimal GEM from the relaxed problem.

As demonstrated in Section 5, our method sets a new state-of-the-art in terms of both accuracy and robustness while at the same time remaining computationally efficient.

2. Related Work

Using a non-central camera rig has attracted much attention from both academic and industrial communities. The most common case is that of a set of cameras—often with non-overlapping views—attached to a headset, micro air vehicle (MAV) or ground vehicle. Our work is particularly relevant for real-time visual localization [7, 14] and autonomous driving [22].

Relative Pose of a Generalized Camera: The minimal solver is based on algebraic geometry, and uses 6 correspondences in order to come up with 64 solutions [34]. However, its large elimination template leads to poor numerical stability. Kim *et al*. later proposed alternative approaches for relative displacement estimation with non-overlapping multi-camera systems using second-order cone programming (SOCP) [15] or branch-and-bound over the space of all rotations [16]. [7] furthermore derived a 5+1 point algorithm, and [25] proposed the antipodal epipolar constraint. A minimal solution for the case of non-holonomic motion was proposed in [22]. Minimal solutions for motions with a common direction were proposed in [23, 26]. An eigenvalue-based formulation for GEM estimation together with efficient local optimization was proposed in [18]. Very recently, another local optimization method for GEM estimation was proposed using an alternating minimization method [5].

Generalized Relative Pose and Scale: There is a further generalization of GEM estimation, the *generalized relative pose and scale* problem. It introduces a further unknown: a relative scale factor between the ray origins in both generalized camera frames. It has an important application in structure from motion with central cameras. While existing work already introduces specialized solvers for this problem [35, 21], the method proposed in this paper could easily be extended to provide a more general solution as well.

Relative Pose of a Central Camera: Essential or fundamental matrix estimation by algebraic error minimization has been extensively studied in previous literature [29, 10, 19, 4]. For both the essential and the fundamental matrix, pose estimation by algebraic error minimization can be formulated as a polynomial optimization problem [28]. A polynomial optimization problem can be reformulated as a quadratically constrained quadratic program (QCQP), which has numerous off-the-shelf solvers. In multiple view geometry, semidefinite relaxation (SDR) for polynomial optimization problems was first studied by Kahl and Henrion in [13]. Recent work [4, 40] has successfully applied it to globally optimal, non-minimal central relative pose computation, which serves as a further motivation of our work.

3. Non-Minimal Solver for GEM Estimation

Relative pose consists of a translation $\mathbf{t}$ — expressed in the first frame and denoting the position of the second frame w.r.t. the first one — and a rotation $\mathbf{R}$ — transforming vectors from the second into the first frame[1]. The translation $\mathbf{t} = [t_1, t_2, t_3]^\top$ is thus identical with a point in $\mathbb{R}^3$. The 3D rotation $\mathbf{R}$ is a 3×3 orthogonal matrix with determinant 1 and belonging to the Special Orthogonal group $\mathrm{SO}(3)$, i.e.,

$$\mathrm{SO}(3) \triangleq \{\mathbf{R} \in \mathbb{R}^{3\times 3} | \mathbf{R}^\top \mathbf{R} = \mathbf{I}_3, \det(\mathbf{R}) = 1\}, \quad (1)$$

where $\mathbf{I}_3$ is the 3×3 identity matrix.

The essential matrix $\mathbf{E}$ is defined as

$$\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}, \quad (2)$$

where $[\cdot]_\times$ constructs the corresponding skew-symmetric matrix of a 3-dimensional vector [11]. The elements of the essential matrix $\mathbf{E}$ and the rotation matrix $\mathbf{R}$ are denoted by e_{ij} and r_{ij}, respectively, where i represents the row index and j the column index. We furthermore define the vectors

$$\mathbf{e} \triangleq \mathrm{vec}(\mathbf{E}) = [e_{11}\ e_{21}\ \ldots\ e_{33}]^T, \text{ and} \quad (3)$$

$$\mathbf{r} \triangleq \mathrm{vec}(\mathbf{R}) = [r_{11}\ r_{21}\ \ldots\ r_{33}]^T, \quad (4)$$

where $\mathrm{vec}(\cdot)$ stacks matrix entries by column-first order.

We define the essential matrix set as

$$\mathcal{M}_\mathbf{E} \triangleq \{\mathbf{E} \mid \mathbf{E} = [\mathbf{t}]_\times \mathbf{R}, \exists\, \mathbf{R} \in \mathrm{SO}(3)\}. \quad (5)$$

[1] Bold capital letters denote matrices (e.g., $\mathbf{E}$ and $\mathbf{R}$); bold lower-case letters denote column vectors (e.g., $\mathbf{e}$, $\mathbf{r}$, and $\mathbf{t}$); non-bold lower-case letters represent scalars (e.g., e and r). $\mathbf{X}_{[a:b,c:d]}$ stands for the submatrix of $\mathbf{X}$ constructed by rows $a \sim b$ and columns $c \sim d$; $\mathbf{x}_{[a:b]}$ stands for the entries of vector $\mathbf{x}$ indexed from a to b.

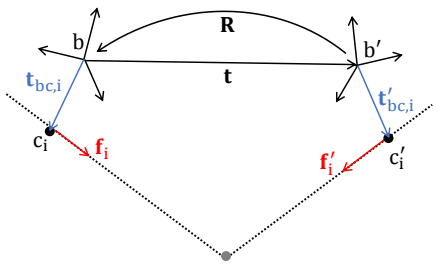


Figure 2. Geometry of the generalized relative pose problem for multi-camera systems.

This essential matrix set is called the *essential matrix manifold* [12]. It is worth mentioning that scale-ambiguity does not exist in GEM estimation, which is why $\mathcal{M}_{\mathbf{E}}$ does not contain any constraints on $\mathbf{t}$. By contrast, there is a scale-ambiguity for standard relative pose estimation, and the translation $\mathbf{t}$ is typically restricted to length 1.

3.1. Generalized Essential Matrix

We now review the GEM describing the relative pose geometry for generalized cameras [32, 24, 18]. As outlined in [32], the transformation rule and the intersection-constraint of Plücker line-vectors easily leads to the epipolar constraint

$$\mathbf{l}_i^\top \begin{bmatrix} \mathbf{E} & \mathbf{R} \\ \mathbf{R} & \mathbf{0} \end{bmatrix} \mathbf{l}'_i = 0, \tag{6}$$

where $(\mathbf{l}_i, \mathbf{l}'_i)$ denotes a pair of corresponding Plücker line-vectors pointing at the i-th 3D point from two different generalized cameras.

Figure 2 illustrates a multi-camera system, which is a common special case of a generalized camera. A point on each Plücker-line is easily given by the capturing camera's center c_i, seen from the origin of the multi-camera system b. If denoting this displacement by $\mathbf{t}_{bc,i}$, we obtain

$$\mathbf{l}_i = \begin{bmatrix} \mathbf{f}_i \\ \mathbf{t}_{bc,i} \times \mathbf{f}_i \end{bmatrix}. \tag{7}$$

Note that we assume that—without loss of generality—c and b have identical orientation. The generalized epipolar constraint thus becomes

$$\mathbf{f}_i^\top \mathbf{E} \mathbf{f}'_i + \mathbf{f}_i^\top \mathbf{R} \mathbf{h}'_i + \mathbf{h}_i^\top \mathbf{R} \mathbf{f}'_i = 0, \tag{8}$$

where

$$\mathbf{h}_i \triangleq \mathbf{t}_{bc,i} \times \mathbf{f}_i; \quad \mathbf{h}'_i \triangleq \mathbf{t}'_{bc,i} \times \mathbf{f}'_i.$$

3.2. Optimizing the GEM by Minimizing the Algebraic Error

Due to the existence of measurement noise, the generalized epipolar constraint will not be strictly satisfied. Denoting the residual for i-th correspondence as

$$\varepsilon_i = \mathbf{f}_i^\top \mathbf{E} \mathbf{f}'_i + \mathbf{f}_i^\top \mathbf{R} \mathbf{h}'_i + \mathbf{h}_i^\top \mathbf{R} \mathbf{f}'_i, \tag{9}$$

the summation of squared residuals for N correspondences $\{(\mathbf{l}_i, \mathbf{l}'_i)\}_{i=1}^N$ becomes a quadratic function in $\mathbf{e}$ and $\mathbf{r}$

$$\varepsilon \triangleq \sum_{i=1}^N \varepsilon_i^2 = [\mathbf{e}^\top, \mathbf{r}^\top] \mathbf{C} \begin{bmatrix} \mathbf{e} \\ \mathbf{r} \end{bmatrix}. \tag{10}$$

$\mathbf{C}$ can be expressed explicitly by

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_4 + \mathbf{C}_5 \\ \mathbf{C}_4^\top + \mathbf{C}_5^\top & \mathbf{C}_2 + \mathbf{C}_3 + \mathbf{C}_6 + \mathbf{C}_6^\top \end{bmatrix}, \tag{11}$$

where

$$\begin{cases} \mathbf{C}_1 = \sum_{i=1}^N (\mathbf{f}'_i \otimes \mathbf{f}_i)(\mathbf{f}'_i \otimes \mathbf{f}_i)^\top \\ \mathbf{C}_2 = \sum_{i=1}^N (\mathbf{h}'_i \otimes \mathbf{f}_i)(\mathbf{h}'_i \otimes \mathbf{f}_i)^\top \\ \mathbf{C}_3 = \sum_{i=1}^N (\mathbf{f}'_i \otimes \mathbf{h}_i)(\mathbf{f}'_i \otimes \mathbf{h}_i)^\top \\ \mathbf{C}_4 = \sum_{i=1}^N (\mathbf{f}'_i \otimes \mathbf{f}_i)(\mathbf{h}'_i \otimes \mathbf{f}_i)^\top \\ \mathbf{C}_5 = \sum_{i=1}^N (\mathbf{f}'_i \otimes \mathbf{f}_i)(\mathbf{f}'_i \otimes \mathbf{h}_i)^\top \\ \mathbf{C}_6 = \sum_{i=1}^N (\mathbf{h}'_i \otimes \mathbf{f}_i)(\mathbf{f}'_i \otimes \mathbf{h}_i)^\top. \end{cases}$$

Note that $\{\mathbf{C}_j\}_{j=1}^6$ are Gram matrices, so they are positive semidefinite (PSD) and symmetric (and so does $\mathbf{C}$). In practice, $\mathbf{C}$ is positive definite for non-minimal GEM estimation scenario.

3.3. A QCQP Formulation

The problem of minimizing the algebraic error on the manifold $\mathcal{M}_{\mathbf{E}}$ can be formulated as

$$\begin{aligned} \min_{\mathbf{E},\mathbf{R},\mathbf{t}} \quad & [\mathbf{e}^\top, \mathbf{r}^\top] \mathbf{C} \begin{bmatrix} \mathbf{e} \\ \mathbf{r} \end{bmatrix} \\ \text{s.t.} \quad & \mathbf{E} = [\mathbf{t}]_\times \mathbf{R}, \quad \mathbf{R} \in \mathrm{SO}(3). \end{aligned} \tag{12}$$

This problem is a QCQP: The objective is a sum of squares, which are PSD quadratic polynomials; the largest set of independent quadratic constraints to define SO(3) is 20 [20, 3]; and, lastly, the constraint between $\mathbf{E}$, $\mathbf{R}$ and $\mathbf{t}$, meaning $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$, is also quadratic. The problem has 21 variables and 29 constraints.

There are some interesting examples in the literature on how the introduction of linearly independent redundant constraints into a QCQP formulation may significantly improve the tightness of the subsequent semidefinite relaxation [1, 3, 4, 38]. For the 20 quadratic constraints considered for SO(3), more than half of them are also redundant and added only for the sake of better tightness [3]. Inspired by this idea, we introduce redundant constraints for problem (12). The below equalities are easily verified:

$$\begin{cases} \mathbf{t}^\top \mathbf{E} = \mathbf{t}^\top ([\mathbf{t}]_\times \mathbf{R}) = \mathbf{0} \\ \mathbf{E}\mathbf{E}^\top = ([\mathbf{t}]_\times \mathbf{R})([\mathbf{t}]_\times \mathbf{R})^\top = [\mathbf{t}]_\times [\mathbf{t}]_\times^\top \\ \mathbf{E}\mathbf{R}^\top = ([\mathbf{t}]_\times \mathbf{R})\mathbf{R}^\top = [\mathbf{t}]_\times \end{cases} \tag{13}$$

These 3 equalities introduce 3, 6 and 9 additional constraints, respectively.

3.4. Relations between Algebraic-Error-Based and Eigenvalue-Based Formulations

In [18], an eigenvalue-based formulation was proposed. Here we demonstrate the close relation between the algebraic-error-based and the eigenvalue-based formulation. By substituting (2) into (8) and applying the permutation rule for triple scalar products, we obtain

$$-(\mathbf{f}_i \times \mathbf{R}\mathbf{f}_i')^\top \mathbf{t} + \left(\mathbf{f}_i^\top \mathbf{R}\mathbf{h}_i' + \mathbf{h}_i^\top \mathbf{R}\mathbf{f}_i'\right) = 0, \tag{14}$$

which can obviously be rewritten as

$$\mathbf{g}_i^\top \tilde{\mathbf{t}} = 0, \text{ with} \tag{15}$$
$$\mathbf{g}_i = \begin{bmatrix} \mathbf{f}_i \times \mathbf{R}\mathbf{f}_i' \\ \mathbf{f}_i^\top \mathbf{R}\mathbf{h}_i' + \mathbf{h}_i^\top \mathbf{R}\mathbf{f}_i' \end{bmatrix} \text{ and } \tilde{\mathbf{t}} = \begin{bmatrix} -w\mathbf{t} \\ w \end{bmatrix}.$$

$\mathbf{g}_i$ here is called a *generalized epipolar plane normal vector*, and $\tilde{\mathbf{t}}$ the *homogeneous translation vector*, which has arbitrary scale [18]. We set w as 1 without loss of generality. Denote

$$\mathbf{G} = [\mathbf{g}_1, \cdots, \mathbf{g}_k], \tag{16}$$

$$\mathbf{H} = \mathbf{G}\mathbf{G}^\top = \sum_{i=1}^{N} \mathbf{g}_i \mathbf{g}_i^\top. \tag{17}$$

Then we can express the summation of residuals by this new parameterization

$$\varepsilon = \sum_{i=1}^{N} \varepsilon_i^2 = \sum_{i=1}^{N} (\mathbf{g}_i^\top \tilde{\mathbf{t}})^2 = \left\|\mathbf{G}^\top \tilde{\mathbf{t}}\right\|_2^2. \tag{18}$$

Thus the algebraic-error-based formulation (12) is equivalent to the following problem

$$\min_{\mathbf{R}, \tilde{\mathbf{t}}} \left\|\mathbf{G}^\top \tilde{\mathbf{t}}\right\|_2^2 \quad \text{s.t. } \mathbf{R} \in \mathrm{SO}(3), \ \ \tilde{\mathbf{t}}_{[4]} = 1. \tag{19}$$

This problem can be further reformulated as

$$\min_{\mathbf{R}} \ J(\mathbf{R}) \quad \text{s.t. } \mathbf{R} \in \mathrm{SO}(3), \tag{20}$$

where

$$J(\mathbf{R}) = \min_{\tilde{\mathbf{t}}} \ \left\|\mathbf{G}^\top \tilde{\mathbf{t}}\right\| \quad \text{s.t. } \tilde{\mathbf{t}}_{[4]} = 1. \tag{21}$$

If we replace the constraint in problem (21) by $\|\tilde{\mathbf{t}}\| = 1$, $J(\mathbf{R})$ can be viewed as finding the optimal $\tilde{\mathbf{t}}$ to minimize $\|\mathbf{G}\tilde{\mathbf{t}}\|$ subject to the condition $\|\tilde{\mathbf{t}}\| = 1$. The solution is the unit eigenvector corresponding to the smallest eigenvalue of the matrix $\mathbf{H} = \mathbf{G}^\top \mathbf{G}$. Let $\sigma_{\mathbf{H},\min}$ denote the smallest eigenvalue of $\mathbf{H}$, thus the optimization problem becomes

$$\min_{\mathbf{R}} \ \sigma_{\mathbf{H},\min} \quad \text{s.t. } \mathbf{R} \in \mathrm{SO}(3). \tag{22}$$

which is exactly the eigenvalue-based formulation that was proposed in [18].

From the previous analysis, it can be seen that the algebraic error formulation and the eigenvalue-based formulation differ only by the domain of the translation vector. The algebraic error method implicitly assumes that the optimal translation is never infinite, as otherwise we can not assume that the homogeneous coordinate of $\tilde{\mathbf{t}}$ is 1. Fortunately, infinite translations in relative pose estimation are not a practical concern.

4. Semidefinite Relaxation and Optimization

We use semidefinite relaxation (SDR) to solve QCQP problem (12). Let us rewrite it in more general form as

$$\begin{aligned} \min_{\mathbf{x} \in \mathbb{R}^n} \ & \mathbf{x}^\top \mathbf{C}_0 \mathbf{x} \\ \text{s.t. } & \mathbf{x}^\top \mathbf{A}_i \mathbf{x} = 0, \ \ i = 1, \cdots, m \\ & \mathbf{x}^\top \mathbf{L} \mathbf{x} = 1, \end{aligned} \tag{23}$$

where

$$\mathbf{x} = [\mathrm{vec}(\mathbf{E}); \mathrm{vec}(\mathbf{R}); \mathbf{t}; y], \tag{24}$$

is a vector stacking all variables. Note that we add an additional variable y that makes the objective and constraints purely quadratic (i.e., no linear or constant term in the objective and no linear term in the equality constraints). This trick is called *homogenization* [27, 3], and introduces the constraint $\mathbf{x}_{[n]}^2 = 1$. By introducing a matrix $\mathbf{L} = \mathrm{diag}([0, \cdots, 0, 1])$, this constraint can be reformulated as $\mathbf{x}^\top \mathbf{L} \mathbf{x} = 1$. Matrices $\mathbf{C}_0, \mathbf{A}_1, \cdots, \mathbf{A}_m \in \mathbb{S}^n$ are determined by the original problem (12), where $\mathbb{S}^n$ denotes the set of all real symmetric $n \times n$ matrices.

In our problem, $n = 22$; $\mathbf{C}_0 = \begin{bmatrix} \mathbf{C} & \mathbf{0}_{18\times 4} \\ \mathbf{0}_{4\times 18} & \mathbf{0}_{4\times 4} \end{bmatrix}$. A crucial first step in deriving an SDR of problem (23) is to observe that

$$\mathbf{x}^\top \mathbf{C}_0 \mathbf{x} = \mathrm{trace}(\mathbf{x}^\top \mathbf{C}_0 \mathbf{x}) = \mathrm{trace}(\mathbf{C}_0 \mathbf{x}\mathbf{x}^\top), \tag{25}$$

$$\mathbf{x}^\top \mathbf{A}_i \mathbf{x} = \mathrm{trace}(\mathbf{x}^\top \mathbf{A}_i \mathbf{x}) = \mathrm{trace}(\mathbf{A}_i \mathbf{x}\mathbf{x}^\top). \tag{26}$$

In particular, both the objective function and constraints in problem (23) are linear in the matrix $\mathbf{x}\mathbf{x}^\top$. Thus, by introducing a new variable $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ and noting that $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ is equivalent to $\mathbf{X}$ being a rank one symmetric PSD matrix, we obtain the following equivalent form of problem (23):

$$\begin{aligned} \min_{\mathbf{X} \in \mathbb{S}^n} \ & \mathrm{trace}(\mathbf{C}_0 \mathbf{X}) \\ \text{s.t. } & \mathrm{trace}(\mathbf{A}_i \mathbf{X}) = 0, \ \ i = 1, \cdots, m, \\ & \mathrm{trace}(\mathbf{L}\mathbf{X}) = 1, \quad \mathbf{X} \succeq \mathbf{0}, \quad \mathrm{rank}(\mathbf{X}) = 1. \end{aligned} \tag{27}$$

Here, $\mathbf{X} \succeq \mathbf{0}$ means that $\mathbf{X}$ is PSD. Solving rank constrained semidefinite programs is NP-hard [36]. SDR drops

the rank constraint $\text{rank}(\mathbf{X}) = 1$ to obtain the following relaxed version of problem (27)

$$\begin{aligned} \min_{\mathbf{X}\in\mathbb{S}^n} \quad & \text{trace}(\mathbf{C}_0\mathbf{X}) \\ \text{s.t.} \quad & \text{trace}(\mathbf{A}_i\mathbf{X}) = 0, \;\; i = 1, \cdots, m, \\ & \text{trace}(\mathbf{L}\mathbf{X}) = 1, \;\; \mathbf{X} \succeq \mathbf{0}. \end{aligned} \tag{28}$$

Problem (28) turns out to be an instance of a semidefinite program (SDP) [36, 27], which may be solved using convex optimization. Modern solvers for SDP are based on primal-dual interior point methods. Its dual problem is

$$\begin{aligned} \max_{\boldsymbol{\lambda},\rho} \quad & \rho \\ \text{s.t.} \quad & \mathbf{Q}(\boldsymbol{\lambda},\rho) = \mathbf{C}_0 - \sum_{i=1}^{m} \lambda_i \mathbf{A}_i - \rho\mathbf{L} \succeq 0, \end{aligned} \tag{29}$$

where $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_m]^\top \in \mathbb{R}^m$. Problem (29) is called the *Lagrangian dual problem* of problem (23), and $\mathbf{Q}(\boldsymbol{\lambda},\rho)$ is the Hessian of the Lagrangian. In summary, the relations between the main formulations are demonstrated in Fig. 3.

We now prove that there is no duality gap between (28) and (29). Thus the problem can be readily solved using off-the-shelf primal-dual interior point methods [39].

Theorem 4.1. *For QCQP problem (12), there is no duality gap between the primal SDP problem (28) and its dual problem (29).*

Proof. Denote the optimal value for problem (28) and its dual problem (29) as f_{primal} and f_{dual}. The inequality $f_{\text{primal}} \geq f_{\text{dual}}$ follows from weak duality. Equality, and the existence of $\mathbf{X}^\star$ and $\boldsymbol{\lambda}^\star$ which attain the optimal values follow if we can show that the feasible regions of both the primal and dual problems have nonempty interiors [36, Theorem 3.1] (also known as Slater's constraint qualification [2].)

For the primal problem (28), let $\mathbf{E}_0$ be an arbitrary point on the essential matrix manifold $\mathcal{M}_{\mathbf{E}}$: $\mathbf{E}_0 = [\mathbf{t}_0]_\times \mathbf{R}_0$. Denote $\mathbf{x}_0 = [\text{vec}(\mathbf{E}_0); \text{vec}(\mathbf{R}_0); \mathbf{t_0}; 1]$. It can be verified that $\mathbf{X}_0 = \mathbf{x}_0\mathbf{x}_0^\top$ is an interior in the feasible domain of the primal problem. For the dual problem (29), we first list part of the constraints as follows

$$\begin{cases} h_1 : e_{11}^2 + e_{12}^2 + e_{13}^2 - (t_2^2 + t_3^2) = 0, & \text{(30a)} \\ h_2 : e_{21}^2 + e_{22}^2 + e_{23}^2 - (t_1^2 + t_3^2) = 0, & \text{(30b)} \\ h_3 : e_{31}^2 + e_{32}^2 + e_{33}^2 - (t_1^2 + t_2^2) = 0, & \text{(30c)} \\ h_4 : r_{11}^2 + r_{12}^2 + r_{13}^2 - y^2 = 0, & \text{(30d)} \\ h_5 : r_{21}^2 + r_{22}^2 + r_{23}^2 - y^2 = 0, & \text{(30e)} \\ h_6 : r_{31}^2 + r_{32}^2 + r_{33}^2 - y^2 = 0, & \text{(30f)} \\ h_7 : r_{11}^2 + r_{21}^2 + r_{31}^2 - y^2 = 0, & \text{(30g)} \\ h_8 : r_{12}^2 + r_{22}^2 + r_{32}^2 - y^2 = 0, & \text{(30h)} \\ h_9 : r_{13}^2 + r_{23}^2 + r_{33}^2 - y^2 = 0, & \text{(30i)} \end{cases}$$

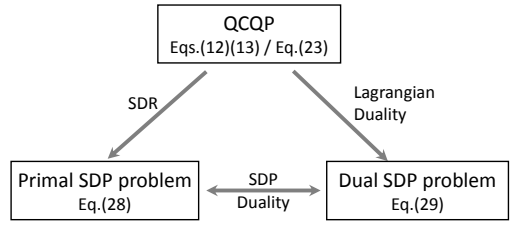


Figure 3. Relations between the main formulations in this work.

where $h_1 \sim h_3$ follows from the constraint $\mathbf{E}\mathbf{E}^\top = [\mathbf{t}]_\times[\mathbf{t}]_\times^\top$, and $h_4 \sim h_9$ originates from the constraints $\mathbf{R}\mathbf{R}^\top = \mathbf{R}^\top\mathbf{R} = \mathbf{I}_3$. Recall that $\mathbf{C} \succ 0$, thus its minimal eigenvector σ_{min} is positive. Let $\lambda_1 \sim \lambda_9$ correspond to the Lagrangian of $h_1 \sim h_9$ respectively. Let the first 9 entries in $\boldsymbol{\lambda}_0$ satisfy $\boldsymbol{\lambda}_{0[1:9]} = -\epsilon[1,1,1,1,1,1,1,1,1]^\top$, and other entries in $\boldsymbol{\lambda}_0$ and ρ_0 be zero. It can be verified that $\mathbf{Q}(\boldsymbol{\lambda}_0, \rho_0) = \begin{bmatrix} \mathbf{C} - \epsilon\mathbf{I}_{18} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & 2\epsilon\mathbf{I}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & 6\epsilon \end{bmatrix} \succ 0$, $\forall \epsilon \in (0, \sigma_{\text{min}})$. That means $\{\boldsymbol{\lambda}_0, \rho_0\}$ is an interior point in the feasible domain of the dual problem. □

4.1. Further Redundant Constraints

To improve tightness of the SDR, we add further redundant constraints on our SDP. The redundant constraint is taken from the SO(3) orbitope.

Definition 4.1 (Orbitope [33])**.** *An orbitope is the convex hull of an orbit of a compact algebraic group that acts linearly on a real vector space. The orbit has the structure of a real algebraic variety, and the orbitope is a convex semi-algebraic set.*

Theorem 4.2 (SO(3) Orbitope, Proposition 4.1 in [33])**.** *The tautological orbitope conv(SO(3)) is a spectrahedron whose boundary is a quartic hypersurface. In fact, a 3×3 matrix $\mathbf{R}$ lies in conv(SO(3)) if and only if*

$$\mathcal{L}(\mathbf{R}) + \mathbf{I}_4 \succeq 0 \tag{31}$$

where $\mathcal{L}(\mathbf{R}) =$

$$\begin{bmatrix} r_{11}+r_{22}+r_{33} & r_{32}-r_{23} & r_{13}-r_{31} & r_{21}-r_{12} \\ r_{32}-r_{23} & r_{11}-r_{22}-r_{33} & r_{21}+r_{12} & r_{13}+r_{31} \\ r_{13}-r_{31} & r_{21}+r_{12} & r_{22}-r_{11}-r_{33} & r_{32}+r_{23} \\ r_{21}-r_{12} & r_{13}+r_{31} & r_{32}+r_{23} & r_{33}-r_{11}-r_{22} \end{bmatrix}.$$

Inequality (31) provides an additional linear matrix inequality for our optimization problem. Note that $\{r_{ij}\}_{i,j=1}^3$ in $\mathbf{R}$ are also entries in $\mathbf{X}$ since $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ and $\mathbf{x} = [\text{vec}(\mathbf{E}); \text{vec}(\mathbf{R}); \mathbf{t}; 1]$. Therefore (31) can be reformulated in terms of $\mathbf{X}$.

4.2. Recovery of Essential Matrix and Relative Pose

Once the optimal $\mathbf{X}^\star$ of the SDP primal problem (28) has been calculated by an SDP solver, we are left with the

task to recover the optimal essential matrix $\mathbf{E}^\star$. Let us denote $\mathbf{X}_e^\star = \mathbf{X}_{[1:9,1:9]}^\star$, $\mathbf{X}_r^\star = \mathbf{X}_{[10:18,10:18]}^\star$ and $\mathbf{X}_t^\star = \mathbf{X}_{[19:21,19:21]}^\star$. Empirically, we found that $\operatorname{rank}(\mathbf{X}_e^\star) = 1$. Denoting the eigenvector that corresponds to the nonzero eigenvalue of $\mathbf{X}_e^\star$ as $\mathbf{e}^\star$, the optimal essential matrix is

$$\mathbf{E}^\star = \operatorname{mat}(\mathbf{e}^\star, [3,3]), \tag{32}$$

where $\operatorname{mat}(\mathbf{e}, [r,c])$ reshapes the vector $\mathbf{e}$ to an $r \times c$ matrix by column-first order.

Once the essential matrix has been obtained, we can recover rotation $\mathbf{R}^\star$ and translation $\mathbf{t}^\star$ by the standard textbook method [11]. However, $\mathbf{E}^\star$ and its derived translation $\mathbf{t}^\star$ do not have the proper scale. To recover the proper scale, we denote the unknown scale factor as s and substitute $s\mathbf{E}^\star$ and $\mathbf{R}^\star$ into the generalized epipolar constraint (8). We then calculate the scale s by solving the least squares problem

$$s = -\frac{\sum_{i=1}^N (\mathbf{f}_i^\top \mathbf{E}\mathbf{f}_i') \cdot (\mathbf{f}_i^\top \mathbf{R}\mathbf{h}_i' + \mathbf{h}_i^\top \mathbf{R}\mathbf{f}_i')}{\sum_{i=1}^N (\mathbf{f}_i^\top \mathbf{E}\mathbf{f}_i')^2}. \tag{33}$$

If the denominator in Eq. (33) is (near) zero, the problem is in a (near) degenerate configuration in which the scale is (nearly) unobservable. Known degenerate configurations correspond to a generalized camera moving along a straight line or—in some cases—a circular arc. In such a scenario, the real scale can not be recovered, while rotation and translation direction can still be found.

We empirically verified that $\operatorname{rank}(\mathbf{X}_e^\star)$ and $\operatorname{rank}(\mathbf{X}_t^\star)$ remain 1, while $\operatorname{rank}(\mathbf{X}_r^\star)$ may be 2. Since $\mathbf{X}_r^\star$ does not satisfy the rank-1 constraint, we can no longer recover the rotation from it. Fortunately, we do not require $\mathbf{X}_r^\star$, and may recover the translation directly from $\mathbf{X}_t^\star$. Similarly, Section 4.3 introduces Theorem 4.3, a sufficient and necessary condition for global optimality, which again does not depend on $\operatorname{rank}(\mathbf{X}_r^\star)$, which is why global optimality is not influenced by an eventual unobservability of scale. The outline of our method is shown in Algorithm 1.

4.3. A Sufficient and Necessary Condition for Global Optimality

Since SDR drops the rank-1 constraint, a sufficient condition for global optimality is that the optimal $\mathbf{X}^\star$ satisfies the rank-1 constraint. However, the rank-1 constraint of $\mathbf{X}^\star$ may not be necessary to guarantee global optimality. The following theorem provides a sufficient and necessary condition, which provides a theoretical foundation for the practical pose recovery method described in Section 4.2.

Theorem 4.3. *For QCQP problem (12) with constraints (13), its SDR problem is tight if and only if: the optimal solution $\mathbf{X}^\star$ to its primal SDP problem (28) satisfies* $\operatorname{rank}(\mathbf{X}_e^\star) = \operatorname{rank}(\mathbf{X}_t^\star) = 1$.

Algorithm 1: Generalized Essential Matrix Estimation by SDP Optimization

Input: observations $\{(\mathbf{l}_i, \mathbf{l}_i')\}_{i=1}^N$
Output: Essential matrix $\mathbf{E}^\star$, rotation $\mathbf{R}^\star$, and translation $\mathbf{t}^\star$

Construct $\mathbf{C}$ by Eq. (11); $\mathbf{C}_0 = \begin{bmatrix} \mathbf{C} & \mathbf{0}_{18\times 4} \\ \mathbf{0}_{4\times 18} & \mathbf{0}_{4\times 4} \end{bmatrix}$;
Construct $\{\mathbf{A}_i\}_{i=1}^m$ and $\mathbf{L}$ in problem (23) which are independent of input;
Obtain $\mathbf{X}^\star$ by solving SDP problem (28) or its dual (29) with redundant constraints;
Assert that $\operatorname{rank}(\mathbf{X}_e^\star) = \operatorname{rank}(\mathbf{X}_t^\star) = 1$;
$\mathbf{E}^\star = \operatorname{mat}(\mathbf{e}^\star, [3,3])$, where $\mathbf{e}^\star$ is the eigenvector corresponding to the largest eigenvalue of $\mathbf{X}_e^\star$;
Decompose $\mathbf{E}^\star$ to obtain rotation $\mathbf{R}^\star$ and normalized translation $\mathbf{t}^\star$;
if $\sum_{i=1}^N (\mathbf{f}_i^\top \mathbf{E}\mathbf{f}_i')^2$ *is larger than a threshold* **then**
  Calculate scale s by Eq. (33);
  $\mathbf{t}^\star \leftarrow s\mathbf{t}^\star$
else
  $\mathbf{t}^\star$ can only be determined up to scale.
end

Proof. First, we prove the *if* part. Note that $\mathbf{X}_e^\star$ and $\mathbf{X}_t^\star$ are real symmetric matrices because they are in the feasible region of the primal SDP. Besides it is given that $\operatorname{rank}(\mathbf{X}_e^\star) = \operatorname{rank}(\mathbf{X}_t^\star) = 1$, thus there exist two vectors $\mathbf{e}^\star$ and $\mathbf{t}^\star$ satisfying $\mathbf{e}^\star(\mathbf{e}^\star)^\top = \mathbf{X}_e^\star$ and $\mathbf{t}^\star(\mathbf{t}^\star)^\top = \mathbf{X}_t^\star$.

According to Theorem 1 in [40], a real 3×3 matrix $\mathbf{E}$ is an essential matrix if and only if there exists a vector $\mathbf{t}$ satisfying $\mathbf{E}\mathbf{E}^\top = [\mathbf{t}]_\times [\mathbf{t}]_\times$. Note that $\mathbf{X}_e^\star = \mathbf{e}^\star(\mathbf{e}^\star)^\top$ and $\mathbf{X}_t^\star = \mathbf{t}^\star(\mathbf{t}^\star)^\top$ satisfy the constraints in problem (28) since they are sub-matrices of a valid solution $\mathbf{X}^\star$. By algebraic derivation based on these constraints, it can be proven that $\mathbf{E}^\star = \operatorname{mat}(\mathbf{e}^\star, [3,3])$ and $\mathbf{t}^\star$ satisfy $\mathbf{E}^\star\mathbf{E}^{\star\top} = [\mathbf{t}^\star]_\times [\mathbf{t}^\star]_\times$. Thus $\mathbf{E}^\star$ is a valid essential matrix.

Next we prove the *only if* part. Since SDR is tight, it means we can uniquely recover a valid relative pose from matrix $\mathbf{X}^\star$. According to Theorem 1 in [40], the minimal requirement to define a valid relative pose is constraints about $\mathbf{E}$ and $\mathbf{t}$. To ensure valid $\mathbf{E}$ and $\mathbf{t}$ can be recovered from $\mathbf{X}^\star$, it should satisfy that $\operatorname{rank}(\mathbf{X}_e^\star) \leq 1$ and $\operatorname{rank}(\mathbf{X}_t^\star) \leq 1$. Since $\mathbf{X}_e^\star$ and $\mathbf{X}_t^\star$ cannot be zero matrices (otherwise $\mathbf{X}^\star$ is not in the feasible region), the equalities should hold. □

Theorem 4.3 provides a sufficient and necessary global optimality condition to recover the optimal solution for the original QCQP. It also provides a method to verify global optimality. Empirically, the optimal $\mathbf{X}^\star$ obtained by the SDP problem always satisfies this condition. Finding the essential conditions to guarantee tightness however remains an open problem [6].

5. Experimental Results

We choose SDPA [37] as the interior point method (IPM) solver, and use the default parameters in all experiments. Our method is implemented in MATLAB, and all experiments are performed on an Intel Core i7 CPU with 1.7 Hz. To improve efficiency, we use the results of the 17-point solver [24] for initialization when more than 17 inliers are available. In this paper, only experiments for synthetic data use this initialization. To improve accuracy, we follow the *suggest-and-improve* framework for general QCQPs [31]. We furthermore use a local optimization method [5] to refine the results provided by SDPA. The complete method takes an average of only 15 ms.

We compared our method against several state-of-the-art methods on both synthetic and real data. Specifically, we compare our method against: (1) the minimal solver 6pt [34]; (2) the linear solver 17pt [24]; (3) the generalized eigenvalue solver ge [18]; and (4) an alternating minimization method (AMM), denoted 17pt-amm [5]. Methods ge and 17pt-amm are both initialized by 17pt. Our own methods are referred to as sdp (without any refinement) and sdp-amm (with AMM refinement).

Among these methods, the implementation of 17pt-amm was provided by the authors, and other comparison methods were taken from OpenGV [17]. Note furthermore that we always ensure a balanced number of samples in each camera, independently of the experiment and number of cameras.

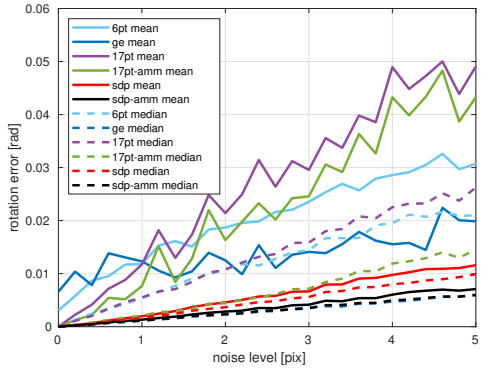
5.1. Results on Synthetic Data

Noise Resilience: The setup of our experiments is similar to the one proposed in [17]. We first test image noise resilience. Each method is evaluated for various noise levels reaching from 0 to 5 pixels and over 1000 random experiments per noise level. The rotation errors of all method is shown in Fig. 4(a). Translation errors follow a similar trend, but are omitted here for the sake of space limitations.

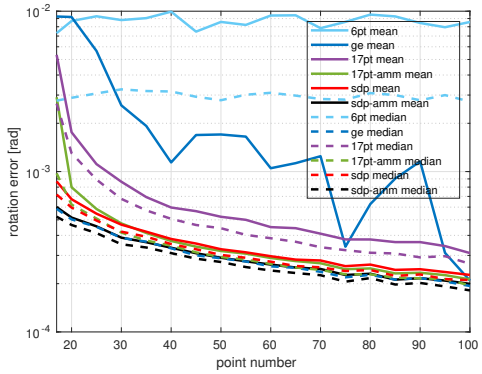
Looking at Fig. 4(a), we make the following observations: (1) sdp-amm degrades the least, and has a relatively obvious advantage over other methods in terms of both accuracy and robustness. This is partially due to the fact that our method does not depend on any initialization and can always find the global optimum. By contrast, ge strongly depends on a good initial value. (2) sdp-amm consistently performs better than sdp, which underlines the effectiveness of the *suggest-and-improve* framework for general QCQPs [31]. (3) sdp still has smaller error than previous state-of-the-art methods.

Number of correspondences: In our next experiment, we fix the image noise level to 0.5 pixel in standard deviation and vary the number N of point correspondences. 6pt can only take a subset of the point correspondences, while other methods utilize all point correspondences. To make the comparison more fair, we randomly sample 20 minimal sets of point correspondences for 6pt, and take the best result in each experiment. The best here is defined as the result that leads to the smallest algebraic error over all simulated correspondences. We again show the rotation error in Fig. 4(b), and make the following observations: (1) As expected, the errors of 17pt, 17pt-amm, sdp, and sdp-amm all decrease as the number of point correspondences is increased. (2) ge still depends on a good initialisation. (3) sdp-amm still leads to the smallest error among all methods.

Optimality Gap: We compare the residuals of our method against those of 17pt and 17pt-amm, respectively. The corresponding scatter plots are indicated in

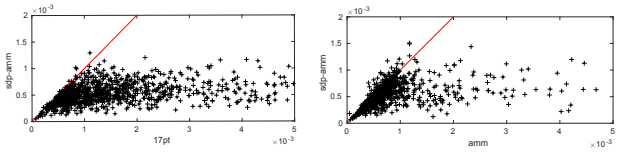


(a) Noise level variations



(b) Number of correspondences

Figure 4. Mean and median of rotation errors with respect to (a) noise level variations and (b) the number of point correspondences.



(a) ours vs. 17pt [24] (b) ours vs. 17pt-amm [5]

Figure 5. Scatter plot comparing the residuals between two methods. A point lying below the red line indicates that our method outperforms in terms of a smaller residual.

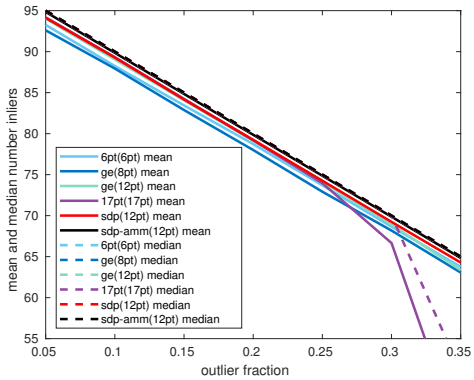


Figure 6. Mean and median number of identified inliers over outlier fraction.

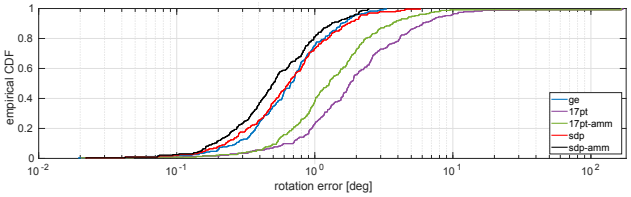
Fig. 5. As can be observed `sdp-amm` has smaller residuals for most of the experiments. The residual of our method typically remains below 1.5×10^{-3}. By contrast, `17pt` and `17pt-amm` may have residuals as large as 5×10^{-3}.

Performance within RANSAC: The most relevant performance measure consists of testing all algorithms as part of a hypothesize-and-test framework. We use the classical RANSAC framework [8], and the same model verification for all methods. For `6pt`, we use an additional 3 points per hypothesis to disambiguate the solution multiplicity. This has no effect on the cost of the disambiguation, and is safer than disambiguation with only one point, especially regarding the high number of solutions and the cost of hypothesis generation. For `17pt`, `ge`, and our methods, we sample 17, 12, and 12 points in each iteration, respectively.

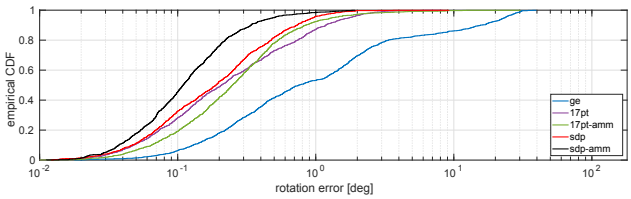
The noise is kept at 0.5 pixel. The total number of point correspondences is fixed to 100, and we vary the outlier fraction. For each outlier fraction, we generate 2000 synthetic scenes and report the mean and median number of identified inliers. Figure 6 reports the number of inliers found by the different methods when integrated into RANSAC. As can be observed, the median of the methods is nearly ideal for all methods except `17pt`. However, `sdp-amm` obtains the largest mean number of identified inliers. In fact, `sdp-amm` is the only method that consistently finds all inliers in each experiment.

5.2. Results on Real Data

To conclude our evaluation, we perform experiments on real data and demonstrate that the advantage of our proposed method applies here as well. We evaluate two datasets. The first one is captured by a custom-made, synchronized 4-camera system mounted on a small-scale automated guided vehicle (AGV), and ground truth is provided by an external motion tracking system. The cameras have a 1216×1936 resolution, are equipped with 48° field-of-view lenses, and are pointing forward, left, right, and backward.



(a) Results on omni-directional 4-cam dataset



(b) Results on forward-facing 2-cam dataset

Figure 7. Empirical cumulative errors distributions for (a) a 4-camera dataset with roughly omni-directional measurement distributions (b) a 2-camera dataset with forward facing cameras.

The second dataset is taken from the KITTI [9] benchmark, which only has a forward facing stereo camera. We ignore the overlap in their fields of view, and treat it as a general multi-camera array. Ground truth is provided by a Velodyne LiDAR and a differential GPS. Figures 7(a) and 7(b) show the cumulative distribution functions (CDFs) of respective rotation errors, demonstrating how `sdp-amm` remains the most accurate method. The difference to alternative methods is particularly important on the KITTI sequence. In this sequence, the bearings of the landmark measurements do not have an omni-directional distribution, which is known to be a challenging case for relative pose estimation with generalized cameras.

6. Conclusions

We introduced the first certifiably globally optimal solution to the non-minimal generalized relative pose estimation problem. Extensive experiments on both synthetic and real data demonstrate clearly improved accuracy and robustness over the previous state-of-the-art, including the ability to handle the difficult scenario of a limited combined field of view of all cameras. Furthermore, by including the essential matrix in our parameterization, the dimensionality of our formulation turns out to be even smaller than the one of a previous SDR based method for central cameras. Even without further polishing of our implementation, this technique already enables real-time processing.

Acknowledgments

This work has been supported by the Natural Science Foundation of Shanghai (grant number: 19ZR1434000) and NSFC (grant number: 61950410612).

References

[1] Kurt Anstreicher and Henry Wolkowicz. On Lagrangian relaxation of quadratic matrix constraints. *SIAM Journal on Matrix Analysis and Applications*, 22(1):41–55, 2000.

[2] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[3] Jesus Briales and Javier Gonzalez-Jimenez. Convex global 3D registration with Lagrangian duality. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 4960–4969, 2017.

[4] Jesus Briales, Laurent Kneip, and Javier Gonzalez-Jimenez. A certifiably globally optimal solution to the non-minimal relative pose problem. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 145–154, 2018.

[5] João Campos, João R. Cardoso, and Pedro Miraldo. POSEAMM: A unified framework for solving pose problems using an alternating minimization method. In *IEEE International Conference on Robotics and Automation*, pages 3493–3499, 2019.

[6] Diego Cifuentes, Sameer Agarwal, Pablo A. Parrilo, and Rekha R. Thomas. On the local stability of semidefinite relaxations. *arXiv:1710.04287v2*, 2018.

[7] Brian Clipp, Jae-Hak Kim, Jan-Michael Frahm, Marc Pollefeys, and Richard Hartley. Robust 6DOF motion estimation for non-overlapping, multi-camera systems. In *IEEE Workshop on Applications of Computer Vision*, pages 1–8, 2008.

[8] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[9] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3354–3361, 2012.

[10] Richard Hartley. Minimizing algebraic error in geometric estimation problem. In *IEEE International Conference on Computer Vision*, pages 469–476, 1998.

[11] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.

[12] Uwe Helmke, Knut Hüper, Pei Yean Lee, and John Moore. Essential matrix estimation using Gauss-Newton iterations on a manifold. *International Journal of Computer Vision*, 74(2):117–136, 2007.

[13] Fredrik Kahl and Didier Henrion. Globally optimal estimates for geometric reconstruction problems. *International Journal of Computer Vision*, 74(1):3–15, 2007.

[14] Tim Kazik, Laurent Kneip, Janosch Nikolic, Marc Pollefeys, and Roland Siegwart. Real-time 6D stereo visual odometry with non-overlapping fields of view. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1529–1536, 2012.

[15] Jae Hak Kim, Richard Hartley, Jan Michael Frahm, and Marc Pollefeys. Visual odometry for non-overlapping views using second-order cone programming. In *Asian Conference on Computer Vision*, pages 353–362, 2007.

[16] Jae Hak Kim, Hongdong Li, and Richard Hartley. Motion estimation for nonoverlapping multicamera rigs: Linear algebraic and L_∞ geometric solutions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(6):1044–1059, 2010.

[17] Laurent Kneip and Paul Furgale. OpenGV: A unified and generalized approach to real-time calibrated geometric vision. In *IEEE International Conference on Robotics and Automation*, pages 1–8, 2014.

[18] Laurent Kneip and Hongdong Li. Efficient computation of relative pose for multi-camera systems. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 446–453, 2014.

[19] Laurent Kneip and Simon Lynen. Direct optimization of frame-to-frame rotation. In *IEEE International Conference on Computer Vision*, pages 2352–2359, 2013.

[20] Laurent Kneip, Roland Siegwart, and Marc Pollefeys. Finding the exact rotation between two images independently of the translation. In *European Conference on Computer Vision*, pages 696–709. Springer, 2012.

[21] Laurent Kneip, Chris Sweeney, and Richard Hartley. The generalized relative pose and scale problem: View-graph fusion via 2D-2D registration. In *IEEE Winter Conference on Applications of Computer Vision*, pages 1–9, 2016.

[22] Gim Hee Lee, Friedrich Faundorfer, and Marc Pollefeys. Motion estimation for self-driving cars with a generalized camera. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 2746–2753, 2013.

[23] Gim Hee Lee, Marc Pollefeys, and Friedrich Fraundorfer. Relative pose estimation for a multi-camera system with known vertical direction. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 540–547, 2014.

[24] Hongdong Li, Richard Hartley, and Jae-hak Kim. A linear approach to motion estimation using generalized camera models. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008.

[25] John Lim, Nick Barnes, and Hongdong Li. Estimating relative camera motion from the antipodal-epipolar constraint. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(10):1907–1914, 2010.

[26] Liu Liu, Hongdong Li, Yuchao Dai, and Quan Pan. 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, 2018.

[27] Zhi-Quan Luo, Wing-Kin Ma, Anthony Man-Cho So, Yinyu Ye, and Shuzhong Zhang. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Processing Magazine*, 27(3):20–34, 2010.

[28] Martin Mevissen and Masakazu Kojima. SDP relaxations for quadratic optimization problems derived from polynomial optimization problems. *Asia-Pacific Journal of Operational Research*, 27(1):15–38, 2010.

[29] Tsuyoshi Migita and Takeshi Shakunaga. Evaluation of epipole estimation methods with/without rank-2 constraint across algebraic/geometric error functions. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–7, 2007.

[30] David Nistér. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):756–770, 2004.

[31] Jaehyun Park and Stephen Boyd. General heuristics for nonconvex quadratically constrained quadratic programming. *arXiv preprint arXiv:1703.07870*, 2017.

[32] Robert Pless. Using many cameras as one. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2003.

[33] Raman Sanyal, Frank Sottile, and Bernd Sturmfels. Orbitopes. *Mathematika*, 57(2):275–314, 2011.

[34] Henrik Stewénius, Magnus Oskarsson, , Kalle Aström, and David Nistér. Solutions to minimal generalized relative pose problems. In *Workshop on Omnidirectional Vision in conjunction with ICCV*, 2005.

[35] Chris Sweeney, Laurent Kneip, Tobias Höllerer, and Matthew Turk. Computing similarity transformations from only image correspondences. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 3305–3313, 2015.

[36] Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, 1996.

[37] Makoto Yamashita, Katsuki Fujisawa, Mituhiro Fukuda, Kazuhiro Kobayashi, Kazuhide Nakata, and Maho Nakata. Latest developments in the SDPA family for solving large-scale SDPs. In *Handbook on semidefinite, conic and polynomial optimization*, pages 687–713. Springer, 2012.

[38] Heng Yang and Luca Carlone. A quaternion-based certifiably optimal solution to the Wahba problem with outliers. In *International Conference on Computer Vision*, pages 1665–1674, 2019.

[39] Yinyu Ye. *Interior Point Algorithms: Theory and Analysis*. Wiley & Sons, 1997.

[40] Ji Zhao. An efficient solution to non-minimal case essential matrix estimation. *arXiv:1903.09067*, 2019.