A Branch and Contract Algorithm For Globally Optimal Fundamental Matrix Estimation

Yinqiang Zheng  Shigeki Sugimoto  Masatoshi Okutomi
Department of Mechanical and Control Engineering, Tokyo Institute of Technology
http://www.ok.ctrl.titech.ac.jp/

Abstract

We propose a unified branch and contract method to estimate the fundamental matrix with guaranteed global optimality, by minimizing either the Sampson error or the point to epipolar line distance, and explicitly handling the rank-2 constraint and scale ambiguity. Based on a novel denominator linearization strategy, the fundamental matrix estimation problem can be transformed into an equivalent problem that involves 9 squared univariate, 12 bilinear and 6 trilinear terms. We build tight convex and concave relaxations for these nonconvex terms and solve the problem deterministically under the branch and bound framework. For acceleration, a bound contraction mechanism is introduced to reduce the size of the branching region at the root node. Given high-quality correspondences and proper data normalization, our experiments show that the state-of-the-art locally optimal methods generally converge to the globally optimal solution. However, they indeed have the risk of being trapped into local minimum in case of noise. As another important experimental result, we also demonstrate, from the viewpoint of global optimization, that the point to epipolar line distance is slightly inferior to the Sampson error in case of drastically varying object scales across two views.

1. Introduction

Two uncalibrated perspective views of a rigid scene are related by the epipolar geometry, which can be algebraically represented by a $3 \times 3$ rank-2 singular matrix, named as the fundamental matrix [9]. Accurate estimation of the fundamental matrix is crucial to many computer vision tasks, like structure from motion and self-calibration. Numerous algorithms have been proposed in the past two decades, and readers can refer to [16, 24, 1] for a review. To be consistent with [1], we classify existing methods into three groups:

**Linear methods**, the normalized 8-point algorithm [10] as a representative, minimizing an algebraic error and usually enforcing the rank constraint through posterior correction. Since these methods generally generate non-optimal solutions, they are frequently used to initialize iterative methods or incorporated into robust methods.

**Iterative methods**, minimizing the reprojection error, the Sampson error or the point to epipolar line distance with/without the rank constraint. As in [2, 14], it is desirable to minimize the reprojective error, referred to as the Gold-Standard in [9], since it is the maximum likelihood estimation (MLE) in case of independently and identically distributed (i.i.d) Gaussian noise. However, the resulting optimization problem is quite complex, due to a large number of subsidiary variables [24, 9].

Quite recently, Kanatani and Sugaya [14] proposed a compact algorithm to minimize the reprojection error without introducing subsidiary variables. Interestingly, their results, together with many other previous works, e.g. [24, 13, 12], demonstrated that the Sampson error, also known as gradient-weighted epipolar error [24], has comparable performance without involving subsidiary variables. This might explain the popularity of the Sampson error in fundamental matrix estimation. Typical algorithms include the FNS and HEIV methods [6] without the rank constraint. Migita and Shakunaga [19] underlined the importance of the rank-2 constraint, and two recent methods by Kanatani and Sugaya, named as EFNS [12] and CLM [13], enforced it through exterior or interior correction. Due to their excellent performance, we regard them as the state-of-the-art methods for Sampson error minimization.

Another well-known criterion is the point to epipolar line distance, named as P2ELD in this paper. However, only a few works, e.g. [24], aimed at minimizing the P2ELD with rank constraint. This is perhaps due to the conclusion by Zhang [25] that the P2ELD criterion is inferior to the Sampson error, especially when the object scale varies drastically across two views.

These three geometrically meaningful criteria can tolerate certain amount of noise, but work poorly in the presence of mismatches, i.e., outliers.

**Robust methods**, applicable for noisy point correspon-
ences with potential outliers. A range of robust methods have been proposed, and readers can refer to the comprehensive work [22], in which Torr and Murray developed and compared various robust estimators for epipolar geometry, such as the well-known RANSAC.

Assuming that potential outliers have been removed using certain robust method, the remaining challenge is to obtain better estimates by minimizing a geometrically meaningful criterion. However, the convergence of existing iterative methods relies on initialization, and there lacks a guaranty of global optimality. The recent trend is to solve this problem globally, but with limited success in somewhat simplified cases so far. For example, Li and Tsin [15] estimated the affine fundamental matrix (fundamental matrix under affine projection) from apparent contours with guaranteed optimality. In [8], Hartley and Kahl proposed to search the rotation space and deterministically solve the essential matrix (calibrated fundamental matrix) estimation. Chesi et al. [5] solved the fundamental matrix estimation by using linear matrix inequality (LMI), but they simply minimized the algebraic error. To the best of our knowledge, the only work to estimate the rank-constrained perspective fundamental matrix by globally minimizing the Sampson error is [11], in which a semidefinite programming (SDP) relaxation was used. Unfortunately, no asymptotic convergence to the global optimum can be guaranteed due to the partial SDP relaxation strategy.

In this paper, we deterministically solve the fundamental matrix estimation under the branch and bound framework. We minimize the Sampson error or the P2ELD, and directly handle the rank-2 constraint. Our method generates a guaranteed \(\epsilon\)-suboptimal solution, for an arbitrary tolerance \(\epsilon\). Note that the optimal solution for Sampson error (or P2ELD) minimization is actually not optimal in the sense of maximum likelihood estimation. Fortunately, Kanatani and Sugaya [14] recently pointed out that the reprojection error can be globally minimized through recursive Sampson error-like minimization. Therefore, it is in principle possible to minimize the reprojection error by embedding our method into the algorithm in [14].

Using a novel denominator linearization method, the nonlinear optimization of fundamental matrix estimation can be transformed into an equivalent, in which the nonconvexity lies merely in 9 squared univariate, 12 bilinear and 6 trilinear terms. We build tight convex and concave relaxations for all these nonconvex terms. We also introduce a bound contraction mechanism to reduce the size of the branching region at the root node, and solve the problem deterministically with reasonable computational burden. Our contribution can be summarized as follows:

1. Leveraging on the novel denominator linearization method, we are able to solve this extremely challenging problem for the first time with guaranteed optimality. Our method is theoretically meaningful, since it can be used to evaluate the performance of existing locally optimal methods. Our conclusion is that the two state-of-the-art methods EFNS [12] and CLM [13] usually converge to the globally optimal solution, when fed with high-quality point correspondences. However, they indeed have a risk of being trapped in local minimum or even un convergence.

2. As an important experimental result, we present a comprehensive comparison of the two widely used criteria, the Sampson error and the point to epipolar line distance (P2ELD), from the viewpoint of global optimization. Our result further validates the conclusion in [25], based otherwise on local optimization, that the latter is comparable to the former, except when the object scale varies significantly in two views.

3. In the aspect of computation, a bound contraction mechanism is introduced, which can be regarded as an augmentation to the usual branch and bound method widely used for computer vision problems ([3, 20] to cite a few).

4. We also introduce the convex and concave envelopes for trilinear programming into computer vision. Compared with the recursive use of bilinear relaxation, the trilinear envelopes generate tighter relaxation, but with smaller problem size. We believe that they would benefit many other vision problems involving trilinearity, like the open nonrigid structure from motion problem mentioned in [4].

2. Problem Formulation

2.1. Notations and Parameterization

Given a set of point correspondences in homogeneous coordinates \(\{x_i \leftrightarrow x'_i\}, i = 1, 2, \ldots, N, N \geq 7\), between two uncalibrated perspective views, ideally, the fundamental matrix \(F\) is a \(3 \times 3\) rank-2 singular matrix satisfying

\[
x'_i^T F x_i = 0, i = 1, 2, \ldots, N.
\]

Because there are only seven independent variables in \(F\), it is desirable to parameterize the fundamental matrix \(F\) properly in 7-D space, e.g., the epipole or bi-epipole parameterizations. Unfortunately, it usually arouses additional singularities, and makes the optimization problem uneasy to handle [9].

In this work, we use a plain 9-D parametrization. Let

\[
F = \begin{bmatrix}
    f_1 & f_2 & f_3 \\
    f_4 & f_5 & f_6 \\
    f_7 & f_8 & f_9
\end{bmatrix},
\]

and the rank-2 constraint is enforced through the determinant

\[
\det(F) = 0.
\]

To eliminate the scale ambiguity, we let

\[
\|f\| = 1,
\]

\[2954\]
where \( \mathbf{f} = (f_1, f_2, \ldots, f_9) \), and \( \| \cdot \| \) denotes the \( L_2 \)-norm of a vector. Note that there still remains a sign ambiguity, since both \( \mathbf{F} \) and \(-\mathbf{F}\) satisfy eq.(1). In our implementation, we eliminate it by enforcing \( f_9 \geq 0 \).

In a branch and bound method, it is crucial to find variable bounds that should be narrow enough to tighten the relaxation, while wide enough to contain the globally optimal solution. We are in favor of this plain parametrization, since we can obtain natural variable bounds from eq.(4) as

\[
-1 \leq f_n \leq 1, \ n = 1, 2, \cdots, 8, 0 \leq f_9 \leq 1. \tag{5}
\]

**2.2. Two Criteria - Sampson Error and Point To Epipolar Line Distance**

According to [9], the Sampson error \( E_{\text{samp}} \) reads:

\[
E_{\text{samp}} = \sum_{i=1}^{N} \frac{(x_i^T \mathbf{F} x_i)^2}{(\mathbf{F} x_i)^2 + (\mathbf{F}^T x_i)^2 + (\mathbf{F} x_i)^2 + (\mathbf{F}^T x_i)^2}, \tag{6}
\]

where \( (\mathbf{F} x_i)^2, j = 1, 2, \) means the square of the \( j \)th entry of the vector \( \mathbf{F} x_i \).

Another geometrically meaningful criterion is the point to epipolar line distance \( E_{\text{p2eld}} \). Considering symmetry, the \( E_{\text{p2eld}} \) can be calculated as follows

\[
E_{\text{p2eld}} = \sum_{i=1}^{N} \frac{(x_i^T \mathbf{F} x_i)^2}{(\mathbf{F} x_i)^2 + (\mathbf{F} x_i)^2 + (\mathbf{F} x_i)^2 + (\mathbf{F}^T x_i)^2}. \tag{7}
\]

Comparing eq.(6) and eq.(7), it is easy to recognize that these two criteria \( E_{\text{samp}} \) and \( E_{\text{p2eld}} \) assume quite similar formation. Due to their similarity, in the following of this paper, we shall derive our algorithm for the Sampson error \( E_{\text{samp}} \) only. Readers can come up with the version for \( E_{\text{p2eld}} \) effortlessly.

Recalling the constraints in eq.(3-5), the fundamental matrix estimation is to solve the following nonlinear optimization problem (P):

\[
\begin{align*}
\min_{\mathbf{f}} & \quad E_{\text{samp}} \\
\text{s.t.} & \quad \det(\mathbf{F}) = 0, \|\mathbf{f}\| = 1, \\
& \quad -1 \leq f_n \leq 1, \ n = 1, 2, \cdots, 8, 0 \leq f_9 \leq 1. \tag{8}
\end{align*}
\]

**2.3. Equivalent Problem**

The original problem (P) is extremely challenging to handle directly due to \( N \) fractional terms in the objective function. Now we present a novel denominator linearization technique and transform it into an equivalent problem (EP) without fractional terms.

First of all, introduce 9 auxiliary variables \( u_k \) and 12 auxiliary variables \( v_{pq} \) such that

\[
u_k = f_k^2, \ k = 1, 2, \cdots, 9, \quad \tag{9}
\]

\[
v_{pq} = f_pf_q, \quad \tag{10}
\]

where \( \{p, q\} \in \Omega_1, \Omega_1 = \{(1, 2), (1, 3), (2, 3), (4, 5), (4, 6), (5, 6), (1, 4), (1, 7), (4, 7), (2, 5), (2, 8), (5, 8)\}. \)

Now expand the denominators of the objective function in (P) for each \( i \), and linearize them by replacing the squared univariate terms and the bilinear terms using eq.(9) and eq.(10). The linearized denominator for each \( i \), denoted by \( L_i(\mathbf{u}, \mathbf{v}, \mathbf{f}) \) for short, is a linear function in \( \mathbf{u}, \mathbf{v} \) and \( \mathbf{f} \).

After introducing \( N \) additional nonnegative variables \( \lambda_i \), the objective function in (P) can be equivalently written as

\[
\min_{\mathbf{f}} \sum_{i=1}^{N} \lambda_i, \tag{11}
\]

\[
s.t., \quad (x_i^T \mathbf{F} x_i)^2 \leq \lambda_i L_i(\mathbf{u}, \mathbf{v}, \mathbf{f}), \ i = 1, 2, \cdots, N. \]

It is easily to recognize that each constraint in eq.(11) defines a rotated second order cone, since \( \lambda_i \geq 0 \) and \( L_i(\mathbf{u}, \mathbf{v}, \mathbf{f}) > 0 \).

Similarly, expanding the determinant constraint in eq.(3), we obtain

\[
f_1f_5f_9 - f_1f_6f_8 - f_2f_4f_9 + f_2f_6f_7 + f_3f_4f_8 - f_3f_5f_7 = 0. \tag{12}
\]

We can introduce 6 variables \( w_{rst} \) such that

\[
w_{rst} = f_rf_sf_t, \tag{13}
\]

where \( \{r, s, t\} \in \Omega_2, \Omega_2 = \{(1, 5, 9), (1, 6, 8), (2, 4, 9), (2, 6, 7), (3, 4, 8), (3, 5, 7)\} \), and linearize the determinant constraint as follows

\[
w_{159} - w_{168} + w_{249} + w_{267} + w_{348} - w_{357} = 0. \tag{14}
\]

The scale constraint in eq.(4) can also be linearized as

\[
\sum_{k=1}^{9} u_k = 1. \tag{15}
\]

Now we obtain the equivalent optimization problem (EP) composed of eq.(11), eq.(9-10), eq.(13-15) and eq.(5).

Interestingly, the nonconvexity of (EP) lies solely in 9 squared univariate (eq.(9)), 12 bilinear (eq.(10), and 6 trilinear (eq.(13)) terms, independent of the number of correspondence \( N \). Undoubtedly, the denominator linearization method is crucial and enables our method to solve large-scale problems with hundreds of point correspondences.

Note that (EP) is still a highly nonconvex problem. In the following we build tight relaxations for these nonconvex terms, and develop a branch and contract algorithm to solve (EP), thus equivalently (P), with global optimality.
3. Globally Optimal Solution to Fundamental Matrix Estimation

3.1. Tight Convex and Concave Relaxation

Given a region $Q$ defined by

$$f_R \leq f_n \leq \overline{f}_n, n = 1, 2, \cdots, 9,$$

(16)

where $f_n$ and $\overline{f}_n$ represent the lower and upper variable bounds of $f_n$, respectively, it is crucial to build high-quality relaxation over $Q$, in which the nonconvex equality constraints should be substituted by their corresponding convex and concave relaxations [3]. The relaxation should be as tight as possible, hopefully the tightest possible one, i.e. the envelope. Specifically, we should replace the nonconvex equalities in eq.(9), eq.(10) and eq.(13) by their corresponding convex and concave relaxations $\text{conv}()$ and $\text{conc}()$ as follows

$$\text{conv}(u_k) \leq u_k \leq \text{conc}(u_k), k = 1, 2, \cdots, 9,$$

$$\text{conv}(v_{pq}) \leq v_{pq} \leq \text{conc}(v_{pq}), (p, q) \in \Omega_1,$$

$$\text{conv}(w_{rst}) \leq w_{rst} \leq \text{conc}(w_{rst}), (r, s, t) \in \Omega_2,$$

(17)

3.1.1 Squared Univariate Terms

For a squared univariate term $u_k = f_k^2, f_k \leq f_k \leq \overline{f}_k, k = 1, 2, \cdots, 9$, the convex relaxation $\text{conv}(u_k)$ is trivial, since $u_k \geq f_k^2$ is a second order cone, thus convex. The tightest concave relaxation $\text{conc}(u_k)$ for a square term is the chord connecting the two endpoints $(\overline{f}_k, \overline{f}_k^2)$ and $(\overline{f}_k, \overline{f}_k^2)$ [20], i.e.

$$u_k \leq (\overline{f}_k + f_k)f_k - f_k \overline{f}_k.$$

(18)

3.1.2 Bilinear Terms

According to [3, 20, 4], the convex relaxation (envelope) $\text{conv}(v_{pq})$ and concave relaxation (envelope) $\text{conc}(v_{pq})$ for a bilinear term $v_{pq} = f_p f_q, f_p \leq f_p \leq \overline{f}_p, f_q \leq f_q \leq \overline{f}_q, (p, q) \in \Omega_1$, are defined by the following two sets of linear inequalities, respectively,

$$v_{pq} \geq \overline{f}_p f_q + \overline{f}_q f_p - \overline{f}_p \overline{f}_q,$$

$$v_{pq} \geq f_p \overline{f}_q + f_q \overline{f}_p - f_p f_q,$$

(19)

and

$$v_{pq} \leq \overline{f}_p f_q + f_q \overline{f}_p - f_p \overline{f}_q,$$

$$v_{pq} \leq f_p f_q + \overline{f}_q f_p - \overline{f}_p \overline{f}_q,$$

(20)

3.1.3 Trilinear Terms

We show two different relaxation methods here.

**Recursive Bilinear Relaxation** For a trilinear term $w_{rst} = f_r f_s f_t, f_r \leq f_r \leq \overline{f}_r, f_s \leq f_s \leq \overline{f}_s, f_t \leq \overline{f}_t, (r, s, t) \in \Omega_2$, it is intuitive to build the relaxation by recursively using bilinear relaxation [18, 4]. Introducing three auxiliary variables $w_{rs, s, w_{rst}, w_{r, s, t} = f_r f_s f_t, w_{r, s, t} = f_r f_s f_t, w_{r, s, t} = f_r f_s f_t$ such that $w_{rs} = f_r f_s, w_{rt} = f_r f_t, w_{st} = f_s f_t$, we actually obtain 6 bilinear equalities, including $w_{rst} = w_{rs} f_t, w_{rst} = w_{rt} f_s, w_{rst} = w_{st} f_r$. The bounds on $w_{rs}, w_{rt}, w_{st}$ can be simply determined through interval analysis, and the relaxation $\text{conv}(w_{rst})$ and $\text{conc}(w_{rst})$ can be obtained by using the bilinear relaxation in previous subsection repeatedly. The detailed derivation is omitted due to lack of space.

**Trilinear Convex and Concave Envelope** In spite of its straightforwardness, the recursive use of bilinear relaxation rarely generates relaxation that coincides with the envelope of a trilinear term, except the special cases pointed out in [18]. Another problem is that there auxiliary variables should be introduced and many linear inequalities involved, both of which increase the computational burden. Fortunately, Meyer and Floudas [18, 17] recently presented the convex and concave envelopes for trilinear terms, offering the tightest possible relaxation without introducing additional variables. Note that the trilinear envelopes rely on the sign of the variable bounds, and there are up to 15 different permutations. Readers can refer to [18, 17] for all these cases.

In our experiments, we found the trilinear envelopes are indeed superior over the recursive use of bilinear relaxation in terms of the quality of relaxation and the problem size. We believe that they could benefit some other computer vision problems involving trilinearity, such as the open question on nonrigid structure from motion in [4].

By embedding the above mentioned relaxations into eq.(17), respectively and correspondingly, we obtain the relaxation problem on $Q$, denoted by (REP), which is composed of eq.(11) and eq.(14-17). The relaxation problem (REP) is actually a second order cone program (SOCP), and can be easily solved by existing convex optimization softwares.

3.2. Bound Contraction Strategy

As mentioned in section 2.1, we obtain natural variable bounds in eq.(5) with moderate tightness through proper parametrization. Usually, the branch and bound algorithm should start from the region defined by these variable bounds, denoted as $Q_0$. However, our test results showed that the convergence speed was quite slow. To accelerate convergence, it is desirable to contract $Q_0$ further and obtain a smaller region $Q_0'$, before the branch and bound method starts. We are inspired by the work [23]. Through any local method, we can obtain a locally optimal solution $E_{local}$ with Sampson error $E_{local}$. Obviously, the Sampson error at the globally optimal solution should not be greater than $E_{samp}$. We conduct bound contraction by solving the following SOCP convex optimization problem for each vari-
able $f_n, n = 1, 2, \cdots, 9$.

\[ \forall n, \min f_n, \text{ and } \max f_n, \]
\[ \text{s.t., } \sum_{i=1}^{N} \lambda_i \leq E_{\text{local}}^{\text{samp}}, \]
\[ \text{All constraints in (REP).} \]

After the bounds are contracted, the convex relaxations in eq.(17) usually become tighter. Therefore, it is possible to contract the bounds further by iteratively solving the problem in eq.(21), until no further contraction can be achieved. Our experiments showed that contracting the bounds once is enough to remove the majority of the region.

### 3.3. Branch and Bound Algorithm

#### Initial Branching Box

The contracted 9D space $Q_0^9$ is the initial branching box.

#### Lower Bound and Upper Bound

To solve the relaxation problem (REP), we get the optimal objective and the optimizer. We use the optimal objective as the lower bound for the current box. In general, the optimizer does not satisfy the rank-2 constraint, until the globally optimal solution is reached. However, it is easy to get an approximation through SVD posterior rank-2 correction [2957]. We evaluate the Sampson error at this corrected solution and regard it as the upper bound.

#### Branching Strategy

We heuristically choose to branch on the box with the lowest lower bound. The dimension with the largest interval is branched at the middle point. This simple branching strategy works well for our application.

#### Convergence Criterion

We terminate the branch and bound iteration when the relative gap between the current best upper bound and lower bound is less than $\varepsilon$. In all experiments, we set $\varepsilon$ to be 5%. Note that such a 5% tolerance level does not necessarily imply that the solution is always 5% away from the real optimal solution. According to our experience, our branch and bound usually reaches the optimal solution in the first one third of the total iterations.

Considering that the computational framework of branch-and-bound is not too new in the computer vision community, we omit the detailed procedures here. Readers can refer to many other works, like [3, 20, 4], and easily come up with the details based on our explanation mentioned above.

### 4. Experimental Results

In this section, we have two independent objectives: (i) to evaluate the performance of locally optimal methods, and (ii) to compare the performance of the Sampson error and the P2ELD in the sense of global optimization.

#### 4.1. Evaluation Criteria and Evaluated Methods

We use three evaluation criteria, including the Sampson error in eq.(6), the P2ELD in eq.(7) and the reprojection error. Given an estimate of $F$, the reprojection error is measured through optimal two-view triangulation [9], so as to avoid any potential local minimum. Considering that our branch and contract method globally minimizes the Sampson error and the P2ELD, we can compare the performance of these two criteria by comparing the reprojection error obtained from optimal triangulation.

There are a huge number of locally optimal methods minimizing the Sampson error. Here, we only consider the two state-of-the-art algorithms EFNS [12] and CLM [13], denoted in short as Samp-EFNS and Samp-CLM. We use the nonlinear optimization tool SNOPT [7] to minimize the Sampson error with rank constraint (Samp-SNOPT), initialized by the normalized 8-point method (N8P) [10]. In recent literature, little attention was paid to the P2ELD minimization. Therefore, we directly use SNOPT to minimize the P2ELD with rank-2 constraint (P2ELD-SNOPT). We compare these locally optimal methods against our global optimization tool SNOPT initialized by the N8P.

In [10], Hartley pointed out the importance of data normalization. We therefore normalize all the point correspondences according to [10]. In addition, all the errors are evaluated in the normalized image coordinates.

#### 4.2. Synthetic Data

For synthetic data, we evaluate the average error over many runs, including the average Sampson error ($\#C1\text{-ASampE}$), the average P2ELD ($\#C2\text{-AP2ELD}$), and the average (squared) reprojection error ($\#C3\text{-ARprojE}$). We synthesize data on the basis of real images to make sure that our experimental settings reflect the practical situation. We establish point correspondences by matching SIFT points, and manually choose 50 correspondences without any outlier. Then we add i.i.d. Gaussian noise with zero mean and $\sigma$ standard deviation on the pixel coordinates. We vary $\sigma$ from 0.8 to 4 pixels. At each noise level, we repeat 100 times for all methods and present the average results.

\[ \sigma \text{ from 0.8 to 4 pixels. At each noise level, we repeat 100 times for all methods and present the average results.} \]
To be consistent with [25], we use three test cases from the well-known Oxford database. In case 1, we use two images with 600×600 resolution from the Aerial View 1 sequence with almost infinite epipoles. In case 2, we use two 512×512 images from the Corridor sequence, in which the epipoles lie in the image due to forward motion. In case 3, we use the 1st and 10th images with 768×512 resolution from the Valbonne Church sequence. The object scale in the first one is almost twice as large as that in the second image. Fig.1 shows the results for all three cases. From the 1st column, we can see that the locally optimal methods EFNS and CLM are likely to be trapped into local minimum, as the noise level increases. Compared with CLM, EFNS is not so stable. As for the performance of the Sampson error and the P2ELD, they have almost the same performance in case 1 and case 2 (see the 3rd column of the first two rows), justifying the conclusion in [25] that these two criteria are comparable when the object scale across two images keep almost the same, despite the position of the epipoles. However, from the result for case 3 (see the last row of Fig.1), we observe that the P2ELD criterion is slightly inferior to the Sampson error when the object scale varies wildly.

4.3. Real Images

In this section, the fundamental matrix is computed automatically by combing SIFT and RANSAC. We use three well-known indoor image sequences: the Dinosaur, Corridor and House sequences. We test all the neighboring pairs, i.e. 36 pairs for Dinosaur, 10 for Corridor and 9 for House sequence, respectively. Only a few outliers exist in the tentative correspondences by matching SIFT points, which might be further filtered out by RANSAC. Considering the randomness of RANSAC, for each image pair, we run 100 times. Among 5500 tests, EFNS fails to reach the global optimum about 30 times, while CLM 10 times. Given high-quality images, the two-state-of-the-art methods actually converge to the globally optimal solution with a very high probability (>99%). In Fig 2, we show one example, with 341 point correspondences after RANSAC, in which both EFNS and CLM find the globally optimal solu-
The convergence process and performance statistics of our optimal method are also presented. Although we need to branch in 9D space, our globally optimal method can converge reasonably fast (e.g., the optimal Sampson error minimization using trilinear envelopes takes 63 seconds for convergence with 24 iterations). When comparing the trilinear envelope and the recursive use of bilinear relaxation, it seems that they almost have the same relaxation quality due to similar convergence process. This is because of the special structure of (EP), whose relaxation quality relies more on the relaxations for square terms and bilinear terms. In spite of that, the algorithm using trilinear envelopes needs less time due to smaller problem size in (REP).

In contrast, Fig.3 shows another example, in which both EFNS and CLM fails to find the global minimum. Note that there is no obvious outlier in the 44 correspondences after RANSAC. We present the convergence process and performance statistics in Fig.4. For this case, our method takes much longer time for convergence. However, we still regard it as efficient, especially considering the difficulty of our problem. We emphasize that the bound contraction strategy is critical. Take the test case shown in Fig.3 as an example, even the upper bound used for contraction $E_{local}$ is 0.0175 and quite far away from the optimal solution 0.0036, the volume of $Q_0$ is reduced from 256 to 0.0067 after solving eq.(21) once. Only a very small region in the 9D space remains valid after contraction. It only takes 3 seconds to solve 18 SOCP problems in eq.(21). Without this contraction step, the method (using trilinear envelopes) would take 3047 iterations in 1684 seconds, almost three times slower than its counterpart with contraction.

5. Conclusions

We have presented a globally optimal solution to fundamental matrix estimation. Our method can handle the Sampson error and the point to epipolar line distance, while explicitly enforcing the rank-2 constraint. According to our best knowledge, no existing method could solve this extremely nonconvex problem deterministically. The key is to transform the original fractional optimization problem into an equivalent, whose nonconvexity lies in 9 squared univariate, 12 bilinear and 6 trilinear terms only. We intro-
produced the envelopes for trilinear terms, which have better performance than the recursive use of bilinear relaxation. A bound contraction strategy was used to reduce the size of the branching box at the root node.

Our method cannot directly minimize the reprojection error, the MLE in case of i.i.d. Gaussian noise. A meaningful future work is to embed our method into the iterative algorithm in [14] so as to minimize the reprojection error.

Acknowledgement. This work was partly supported by Grant-in-Aid for Scientific Research (21240015) from the Japan Society for the Promotion of Science.

References


Figure 4. The convergence process and performance statistics of our globally optimal method for Sampson-error and P2ELD minimization, involving either trilinear envelopes or recursive use of bilinear relaxation.