

A Versatile Approach for Solving PnP, PnPf, and PnPfr Problems

Gaku Nakano

Central Research Labs, NEC Corporation, Japan
g-nakano@cq.jp.nec.com

Abstract. This paper proposes a versatile approach for solving three kinds of absolute camera pose estimation problem: PnP problem for calibrated cameras, PnPf problem for cameras with unknown focal length, and PnPfr problem for cameras with unknown focal length and unknown radial distortion. This is not only the first least squares solution to PnPfr problem, but also the first approach formulating three problems in the same theoretical manner. We show that all problems have a common subproblem represented as multivariate polynomial equations. Solving these equations by Gröbner basis method, we derive a linear form for the remaining parameters of each problem. Finally, we apply root polishing to strictly satisfy the original KKT condition. The proposed PnP and PnPf solvers have comparable performance to the state-of-the-art methods on synthetic distortion-free data. Moreover, the novel PnPfr solver gives the best result on distorted point data and demonstrates real image rectification against significant distortion.

Keywords: Absolute Camera Pose Estimation, PnP Problem, Focal Length, Radial Distortion

1 Introduction

Camera parameter estimation from n pairs of 2D-3D point correspondence in a single image has been a fundamental problem in computer vision and photogrammetry community. The camera parameters consist of two kinds of parameters. One is the extrinsic parameters which determine the position and the orientation of the camera, i.e., 3D rotation and translation. The other is the intrinsic parameters which are optical properties of the camera unaffected by the extrinsic parameters, i.e., focal length, skew, principal point, aspect ratio, lens distortion, etc. The name of the parameter estimation problem is different depending on unknown parameters: Perspective- n -Point (PnP) problem when the extrinsic parameters are unknown and all the intrinsic parameters are calibrated in advance, PnPf problem for partially calibrated cameras when only focal length is known, PnPfr problem when radial distortion of the lens is additionally unknown.

It is well discussed that $n = 3$ is the minimal number of the points required to solve PnP problem [1–3]. The trend of the latest PnP solvers is to find the global optimal solution for $n \geq 3$ case in linear complexity $O(n)$ without considering

planar or non-planar scene. The first $O(n)$ method is EPnP [4], but it does not assure the global optimality. Hesch and Roumeliotis [5] proposed Direct Least Square method (DLS) which finds all stationary points of the first optimality condition, also known as the Karush-Kuhn-Tucker (KKT) condition, by solving a system of nonlinear multivariate polynomial equations. The DLS approach has been improved for more stability and efficiency [6, 7], and extended to generalized camera model which has multiple focal points [8]. However, applications of PnP problem are limited due to the strict assumption that the intrinsic parameters are never changed during shooting a scene. Prior full calibration is mandatory, but it is difficult for cameras having a zoom lens.

PnPf problem deals with a relaxed assumption where the intrinsic parameters are known except for focal length. Since principal point, skew, and aspect ratio are invariant to zoom change, focal length is the only varying parameter. Moreover, for recent digital cameras, we can assume that zero skew (square pixels), one aspect ratio (parallel mount of lens and camera), and principal point is at the image center (center aligned lens and camera). P4Pf [9–11] and PnPf [12–14] solvers have been proposed, which use $n = 4$ for the minimal case and $n \geq 4$ for the least square case, respectively. Kanaeva et al. [14] extended EPnP to PnPf problem by improving EPnP’s drawbacks, and pointed out that Zheng et al.’s PnPf solver [13] sometimes fails to calculate focal length on real data. However, PnPf problem’s assumption ignores the fact that lens distortion is also changeable according to zoom variation. Similarly to PnP problem, complete prior lens distortion correction is difficult for zooming cameras. Therefore, PnPf solvers can handle only slight zoom change where lens distortion can be ignored or approximated by fixed parameters.

To deal with lens distortion, P4Pfr [15, 16] and P5Pfr [17] solvers have been developed. They modeled radial lens distortion by Fitzgibbon’s division model [18] for simple formulations. Kukulova et al. [17] showed that the three-parameter division model is practically sufficient for 3D shape reconstruction from real images even with significant distortion. Since these solvers are designed for the minimal case, they cannot improve the parameter accuracy for n points without a costly reprojection error minimization. Addition to the P4Pfr and P5Pfr solvers, some methods correcting lens distortion from a single image have been proposed [19, 20]. However, those methods are not sufficiently fast for real-time applications, such as Visual SLAM and augmented reality.

This paper proposes three solvers for PnP, PnPf, and PnPfr problems which are derived from the same theoretical formulation. Inspired by Kukulova et al.’s P5Pfr solver [17], the key is to find a common subproblem among the three problems. The common subproblem is expressed by only a part of the extrinsic parameters, therefore, this subproblem can be solved by Gröbner basis method similarly to the existing PnP solvers [6–8]. Regarding the solutions of the common subproblem as known parameters, we show that estimation of the remaining parameters can be formulated as a linear problem. This part slightly differs depending on each problem but can be solved in the same manner. Finally, for easy implementation of root polishing, we derive new equations without La-

Table 1. Comparison of PnP, PnPf, and PnPfr problems. Numbers marked with \dagger and $*$ indicate the case of the one- and the three-parameter division model for radial distortion, respectively. This paper discusses only the latter case

Parameters	PnP	PnPf	PnPfr
Rotation	✓	✓	✓
Translation	✓	✓	✓
Focal Length		✓	✓
Radial Distortion			✓
# of parameters	6	7	$8^\dagger, 10^*$
Required points	3	4	$4^\dagger, 5^*$

grange multipliers, which are equivalent to the original KKT condition. This is an extension of Nakano’s approach for PnP problem [7].

Synthetic data experiments show that the proposed PnP and PnPf solvers have the same accuracy and efficiency as the state-of-the-art methods on PnP and PnPf problems without lens distortion. For lens distortion data, the proposed PnPfr solver is the only method that is able to improve the parameter accuracy with increasing the number of the points. Moreover, we show that the PnPfr solver successfully corrects significant lens distortion on real images taken by an ultra-wide zoom camera.

2 Problem formulation

This section describes mathematical formulations of PnP, PnPf, and PnPfr problems. In this paper, we assume the standard pinhole camera model for the projection between 2D-3D point correspondences and the three-parameter division model for radial distortion [17].

The projection of a 3D point $\mathbf{p}_i = [x_i, y_i, z_i]^\top$ onto a 2D image point $\mathbf{m}_i = [u_i, v_i, w_i]^\top$ represented by the homogeneous coordinates can be written as

$$\mathbf{m}_i \sim \mathbf{K}(\mathbf{R}\mathbf{p}_i + \mathbf{t}), \quad (1)$$

where \sim denotes equality up to scale, \mathbf{R} is a 3×3 rotation matrix, $\mathbf{t} = [t_x, t_y, t_z]^\top$ is a translation vector, and $\mathbf{K} = \text{diag}([1, 1, f^{-1}])$ is the calibration matrix of the camera with focal length f . As mentioned in Section 1, we assume zero skew, one aspect ratio, and principal point corresponding to the image center.

The common unknowns among PnP, PnPf, and PnPfr problems are the extrinsic parameters, \mathbf{R} and \mathbf{t} . The homogeneous term w_i and intrinsic parameters to be estimated are different in each problem. In PnP problem, $w_i = 1$ and f is known. In PnPf problem, $w_i = 1$ but f is unknown. In PnPfr problem, f is also unknown and $w_i = 1 + \mathbf{k}^\top \mathbf{d}_i$, where $\mathbf{d}_i = [u_i^2 + v_i^2, (u_i^2 + v_i^2)^2, (u_i^2 + v_i^2)^3]^\top$ and $\mathbf{k} = [k_1, k_2, k_3]^\top$ is a 3×1 vector containing the unknown radial distortion coefficients.

Note that the image coordinates u_i and v_i represent undistorted points in PnP and PnPf problems but distorted points in PnPfr problem. Hereafter, for simple notations, this paper does not distinguish the description of distorted or undistorted points.

Now we formulate PnPfr problem. PnP and PnPf problems can be similarly derived by regarding f or \mathbf{k} as the knowns. Given n point correspondences, PnPfr problem can be written as a constrained nonlinear optimization,

$$\begin{aligned} \min_{\mathbf{R}, \mathbf{t}, f, \mathbf{k}} \quad & \sum_{i=1}^n \left\| [\mathbf{m}_i]_{\times} \mathbf{K}(\mathbf{R}\mathbf{p}_i + \mathbf{t}) \right\|^2 \\ \text{s.t.} \quad & \mathbf{R}^{\top} \mathbf{R} = \mathbf{I}, \quad \det(\mathbf{R}) = 1, \end{aligned} \quad (2)$$

where $[\]_{\times}$ denotes a matrix representation of the vector cross product, i.e.,

$$[\mathbf{m}_i]_{\times} = \begin{bmatrix} 0 & -w_i & v_i \\ w_i & 0 & -u_i \\ -v_i & u_i & 0 \end{bmatrix}. \quad (3)$$

This operator is introduced to eliminate the scale ambiguity of Eq. (1).

In Eq. (2), the total number of the unknowns is 10, of which three from \mathbf{R} , three from \mathbf{t} , one from f , and three from \mathbf{k} . Since $[\mathbf{m}_i]_{\times}$ is of rank two, Eq. (1) gives us two equations for each point correspondences. Therefore, Eq. (2) can be solved by $n \geq 5$ point correspondence. PnP and PnPf problems are also solvable because they have totally six and seven unknowns, which are less than 10, respectively. Table 1 summarizes the unknown parameters and the number of the unknowns in each problem.

3 Proposed method

This section describes the derivation of the proposed method. We begin with an overview of the key idea, which divides the least squares problem into two subproblems. Then, we derive efficient solutions to the subproblems based on Gröbner basis method and a linear method, respectively. Finally, we introduce a root polishing technique to satisfy the KKT condition of the original problem.

3.1 Overview

Define three row vectors \mathbf{a}_i^{\top} , \mathbf{b}_i^{\top} , and \mathbf{c}_i^{\top} corresponding to the first, second and third row of $[\mathbf{m}_i]_{\times}$, respectively. Note that those terms contain the unknown radial distortion \mathbf{k} in w_i . By introducing them into Eq. (2), the cost function can be rewritten by

$$\min_{\mathbf{R}, \mathbf{t}, f, \mathbf{k}} \sum_{i=1}^n \left\| [\mathbf{m}_i]_{\times} \mathbf{q}_i \right\|^2 = \sum_{i=1}^n (\mathbf{a}_i^{\top} \mathbf{q}_i)^2 + (\mathbf{b}_i^{\top} \mathbf{q}_i)^2 + (\mathbf{c}_i^{\top} \mathbf{q}_i)^2, \quad (4)$$

where $\mathbf{q}_i = \mathbf{K}(\mathbf{R}\mathbf{p}_i + \mathbf{t})$. The rotation matrix constraints are omitted here.

Interpreting Eq. (4) from the point of view of algebraic geometry, the two vectors, \mathbf{m}_i and \mathbf{q}_i , are collinear without noise in data. In other words, minimizing Eq. (4) with noisy data is equivalent to finding the optimal parameters so that the three terms are closed to zeros. Therefore, if we minimized each term as an independent subproblem and obtained solutions from them, we can expect that one of the solutions is closed to the global optimum. This is the key idea of the proposed method.

Let us move on how to build the subproblems. Expanding $\mathbf{a}_i^\top \mathbf{q}_i$, $\mathbf{b}_i^\top \mathbf{q}_i$ and $\mathbf{c}_i^\top \mathbf{q}_i$, we obtain

$$\mathbf{a}_i^\top \mathbf{q}_i = -w_i(\mathbf{r}_2^\top \mathbf{p}_i + t_y) + v_i f^{-1}(\mathbf{r}_3^\top \mathbf{p}_i + t_z), \quad (5)$$

$$\mathbf{b}_i^\top \mathbf{q}_i = w_i(\mathbf{r}_1^\top \mathbf{p}_i + t_x) - u_i f^{-1}(\mathbf{r}_3^\top \mathbf{p}_i + t_z), \quad (6)$$

$$\mathbf{c}_i^\top \mathbf{q}_i = -v_i(\mathbf{r}_1^\top \mathbf{p}_i + t_x) + u_i(\mathbf{r}_2^\top \mathbf{p}_i + t_y), \quad (7)$$

where \mathbf{r}_j denotes the j -th row of \mathbf{R} . Interestingly, Eq. (7) does not have w_i , a function of \mathbf{k} , and is expressed by only a part of the extrinsic parameters, \mathbf{r}_1 , \mathbf{r}_2 , t_x , and t_y whereas Eqs. (5) and (6) consist of all the unknown parameters.

Thus, we can define the first subproblem by

$$\begin{aligned} \min_{\mathbf{r}_1, \mathbf{r}_2, t_x, t_y} & \sum_{i=1}^n (\mathbf{c}_i^\top \mathbf{q}_i)^2 \\ \text{s.t.} & \quad \|\mathbf{r}_1\|^2 = 1, \quad \|\mathbf{r}_1\|^2 - \|\mathbf{r}_2\|^2 = 0, \quad \mathbf{r}_1^\top \mathbf{r}_2 = 0. \end{aligned} \quad (8)$$

There seems to be eight unknowns in Eq. (8). However, actual degrees of freedom is five due to the three constraints for \mathbf{r}_1 and \mathbf{r}_2 . Therefore, Eq. (8) can be solved by $n \geq 5$ point correspondences. After finding \mathbf{r}_1 and \mathbf{r}_2 , we can recover \mathbf{R} by calculating the third row, $\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2$.

Plugging \mathbf{R} , t_x , and t_y into Eqs. (5) and (6), we still have five unknowns, t_z , f , and \mathbf{k} . Since the rotation matrix has been already estimated, the remaining unknowns do not have any constraints.

Therefore, we can build the second subproblem as

$$\min_{t_z, f, \mathbf{k}} \sum_{i=1}^n (\mathbf{a}_i^\top \mathbf{q}_i)^2 + (\mathbf{b}_i^\top \mathbf{q}_i)^2. \quad (9)$$

Given $n \geq 5$ point correspondences, we can solve Eq. (9) because $2n$ equations are available for the five unknowns.

The estimated parameters from the above two subproblems are not the optimal solution to the original problem, Eq. (4). Therefore, we finally refine the parameters by conducting a root polishing to get more accuracy and optimality.

From Sections 3.2 to 3.4, we will discuss the details of specific methods for each step.

3.2 Solving the first subproblem

From Eq. (7), the cost function of Eq. (8) can be rewritten by

$$\min_{\mathbf{r}_1, \mathbf{r}_2, t_x, t_y} \sum_{i=1}^n (\mathbf{c}_i^\top \mathbf{q}_i)^2 = \|\mathbf{A}\hat{\mathbf{r}} + \mathbf{B}\hat{\mathbf{t}}\|^2, \quad (10)$$

where

$$\mathbf{A} = \begin{bmatrix} -v_1 \mathbf{p}_1^\top & u_1 \mathbf{p}_1^\top \\ \vdots & \vdots \\ -v_n \mathbf{p}_n^\top & u_n \mathbf{p}_n^\top \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -v_1 & u_1 \\ \vdots & \vdots \\ -v_n & u_n \end{bmatrix}, \quad \hat{\mathbf{r}} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \end{bmatrix}, \quad \hat{\mathbf{t}} = \begin{bmatrix} t_x \\ t_y \end{bmatrix}. \quad (11)$$

Since there are no constraints about $\hat{\mathbf{t}}$, we can express $\hat{\mathbf{t}}$ as a function of $\hat{\mathbf{r}}$, i.e.,

$$\hat{\mathbf{t}} = -(\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{A} \hat{\mathbf{r}}. \quad (12)$$

Substituting this into Eq. (10), we obtain a new constrained problem as follows:

$$\begin{aligned} \min_{\mathbf{r}_1, \mathbf{r}_2} \quad & \hat{\mathbf{r}}^\top \mathbf{M} \hat{\mathbf{r}} \\ \text{s.t.} \quad & \|\mathbf{r}_1\|^2 = 1, \quad \|\mathbf{r}_1\|^2 - \|\mathbf{r}_2\|^2 = 0, \quad \mathbf{r}_1^\top \mathbf{r}_2 = 0, \end{aligned} \quad (13)$$

where

$$\mathbf{M} = \mathbf{A}^\top \mathbf{A} - \mathbf{A}^\top \mathbf{B} (\mathbf{B}^\top \mathbf{B})^{-1} \mathbf{B}^\top \mathbf{A}. \quad (14)$$

Since Eq. (13) has a similar form in the existing PnP solvers [6–8], we can use same Gröbner basis technique for solving the optimal $\hat{\mathbf{r}}$. If we introduce a quaternion based parameterization for representing the rotation matrix as in [6, 8], we obtain up to 40 solutions, which is exactly the same number of the solutions to [6–8]. However, there is a sign ambiguity for \mathbf{r}_1 and \mathbf{r}_2 , that means $-\mathbf{r}_1$ and $-\mathbf{r}_2$ also give the minimum error with satisfying the constraints. Therefore, the number of the solutions is actually 20, not 40. Any quaternion based parameterizations cannot distinguish the sign ambiguity of $\pm \hat{\mathbf{r}}$ because quaternion has a sign ambiguity in itself. To obtain 20 solutions by Gröbner basis method, we need to derive new equations independent to the norm definition of \mathbf{r}_1 and \mathbf{r}_2 .

Let M_{ij} be a (i, j) entry of 3×3 block matrix which partitions the 6×6 matrix \mathbf{M} into 2×2 blocks. The Lagrange function of Eq. (13) can be written by

$$\begin{aligned} L = & \mathbf{r}_1^\top \mathbf{M}_{11} \mathbf{r}_1 + 2\mathbf{r}_1^\top \mathbf{M}_{12} \mathbf{r}_2 + \mathbf{r}_2^\top \mathbf{M}_{22} \mathbf{r}_2 \\ & + \lambda_1 (1 - \|\mathbf{r}_1\|^2) + \lambda_2 (\|\mathbf{r}_1\|^2 - \|\mathbf{r}_2\|^2) + 2\lambda_3 \mathbf{r}_1^\top \mathbf{r}_2, \end{aligned} \quad (15)$$

where λ_i is a Lagrange multiplier and the multiplier 2 for λ_3 is merely for convenience. The KKT condition of Eq. (15) is given by

$$\frac{\partial L}{\partial \mathbf{r}_1} = \mathbf{M}_{11}\mathbf{r}_1 + \mathbf{M}_{12}\mathbf{r}_2 + (\lambda_2 - \lambda_1)\mathbf{r}_1 + \lambda_3\mathbf{r}_2 = \mathbf{0}, \quad (16)$$

$$\frac{\partial L}{\partial \mathbf{r}_2} = \mathbf{M}_{22}\mathbf{r}_2 + \mathbf{M}_{12}^\top\mathbf{r}_1 - \lambda_2\mathbf{r}_2 + \lambda_3\mathbf{r}_1 = \mathbf{0}, \quad (17)$$

$$\frac{\partial L}{\partial \lambda_1} = 1 - \|\mathbf{r}_1\|^2 = 0, \quad (18)$$

$$\frac{\partial L}{\partial \lambda_2} = \|\mathbf{r}_1\|^2 - \|\mathbf{r}_2\|^2 = 0, \quad (19)$$

$$\frac{\partial L}{\partial \lambda_3} = \mathbf{r}_1^\top\mathbf{r}_2 = 0, \quad (20)$$

Multiplying $[\mathbf{r}_1]_\times$ and $[\mathbf{r}_2]_\times$ to Eqs. (16) and (17), respectively, we obtain

$$[\mathbf{r}_1]_\times (\mathbf{M}_{11}\mathbf{r}_1 + \mathbf{M}_{12}\mathbf{r}_2) + \lambda_3 [\mathbf{r}_1]_\times \mathbf{r}_2 = \mathbf{0}, \quad (21)$$

$$[\mathbf{r}_2]_\times (\mathbf{M}_{22}\mathbf{r}_2 + \mathbf{M}_{12}^\top\mathbf{r}_1) + \lambda_3 [\mathbf{r}_2]_\times \mathbf{r}_1 = \mathbf{0}. \quad (22)$$

Using the relation $[\mathbf{r}_1]_\times \mathbf{r}_2 = -[\mathbf{r}_2]_\times \mathbf{r}_1$, we can eliminate λ_3 by adding Eqs. (21) and (22). Thus, we obtain

$$[\mathbf{r}_1]_\times (\mathbf{M}_{11}\mathbf{r}_1 + \mathbf{M}_{12}\mathbf{r}_2) + [\mathbf{r}_2]_\times (\mathbf{M}_{22}\mathbf{r}_2 + \mathbf{M}_{12}^\top\mathbf{r}_1) = \mathbf{0}. \quad (23)$$

Moreover, multiplying \mathbf{r}_2^\top and \mathbf{r}_1^\top to Eqs. (16) and (17), respectively, we obtain

$$\mathbf{r}_2^\top (\mathbf{M}_{11}\mathbf{r}_1 + \mathbf{M}_{12}\mathbf{r}_2) + \lambda_3 \|\mathbf{r}_2\|^2 = 0, \quad (24)$$

$$\mathbf{r}_1^\top (\mathbf{M}_{22}\mathbf{r}_2 + \mathbf{M}_{12}^\top\mathbf{r}_1) + \lambda_3 \|\mathbf{r}_1\|^2 = 0. \quad (25)$$

Since the norm of \mathbf{r}_1 and \mathbf{r}_2 are equal to each other as in Eq. (19), we can also eliminate λ_3 by subtracting Eq. (24) from Eq. (25),

$$\mathbf{r}_2^\top (\mathbf{M}_{11}\mathbf{r}_1 + \mathbf{M}_{12}\mathbf{r}_2) - \mathbf{r}_1^\top (\mathbf{M}_{22}\mathbf{r}_2 + \mathbf{M}_{12}^\top\mathbf{r}_1) = 0. \quad (26)$$

It should be noted that Eqs. (23) and (26) hold for any types of normalization of \mathbf{r}_1 as long as the other constraints, Eqs. (19) and (20), are satisfied. Hence, instead of Eq. (18), we can use a linear constraint for eliminating the sign ambiguity of \mathbf{r}_1 and \mathbf{r}_2 , e.g., $r_{11} = 1$ or $r_{11} + r_{12} + r_{13} = 1$, where r_{ij} is the (i, j) element of \mathbf{R} . Therefore, we can obtain \mathbf{r}_1 and \mathbf{r}_2 by solving Eqs. (19), (20), (23), and (26) together with the new linear constraint for \mathbf{r}_1 .

Since the above equations can be represented by a system of nonlinear polynomial equations, the solution can be obtained by using Gröbner basis method. A simple way is to use an automatic generator of Gröbner basis solvers developed by Kukelova et al. [21]. In our case, the automatic generator gives a 105×125 template matrix for Gauss-Jordan elimination and a 20×20 action matrix for

the eigenvalue computation. We obtained a further optimized 97×117 template matrix by sorting all equations before starting necessary equation extraction in the automatic solver. This solver gives at most 20 pairs of \mathbf{r}_1 and \mathbf{r}_2 , from which we can recover $\hat{\mathbf{t}}$ by Eq. (12) and two rotation matrices by considering the sign ambiguity:

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1^\top \\ \mathbf{r}_2^\top \\ (\mathbf{r}_1 \times \mathbf{r}_2)^\top \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} -\mathbf{r}_1^\top \\ -\mathbf{r}_2^\top \\ (\mathbf{r}_1 \times \mathbf{r}_2)^\top \end{bmatrix}. \quad (27)$$

An example code for the automatic generator of this subproblem is shown in Appendix A in the supplemental material.

3.3 Solving the second subproblem

The solution to the second subproblem slightly differs on PnP, PnPf, PnPfr problems. Due to limitations of space, we show a solution to PnPfr problem only. Solutions to PnP and PnPf problems are described in Appendix B in the supplemental material.

Regarding \mathbf{R} , t_x , and t_y from Section 3.2 as known parameters, we can rewrite Eq. (9) as

$$\min_{t_z, f, \mathbf{k}} \sum_{i=1}^n \left\| \begin{bmatrix} \mathbf{a}_i^\top \\ \mathbf{b}_i^\top \end{bmatrix} \mathbf{q}_i \right\|^2 = \|\mathbf{L}\mathbf{x} + \mathbf{g}\|^2, \quad (28)$$

where

$$\mathbf{L} = \begin{bmatrix} v_1 & v_1 z_1^c & -y_1^c \mathbf{d}_1^\top \\ -u_1 & -u_1 z_1^c & x_1^c \mathbf{d}_1^\top \\ & \vdots & \\ v_n & v_n z_1^c & -y_n^c \mathbf{d}_n^\top \\ -u_n & -u_n z_1^c & x_n^c \mathbf{d}_n^\top \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} f^{-1} t_z \\ f^{-1} \\ \mathbf{k} \end{bmatrix}, \quad \mathbf{g} = \begin{bmatrix} -y_1^c \\ x_1^c \\ \vdots \\ -y_n^c \\ x_n^c \end{bmatrix}, \quad (29)$$

$$x_i^c = \mathbf{r}_1^\top \mathbf{p}_i + t_x, \quad y_i^c = \mathbf{r}_2^\top \mathbf{p}_i + t_y, \quad z_i^c = \mathbf{r}_3^\top \mathbf{p}_i.$$

This is a linear form for the unknown vector \mathbf{x} , therefore, the solution can be obtained by solving a normal equation $\mathbf{x} = -(\mathbf{L}^\top \mathbf{L})^{-1} \mathbf{L}^\top \mathbf{g}$. Then, t_z can be recovered by dividing the first element by the second element in \mathbf{x} .

3.4 Root polishing

As a result from Sections 3.2 and 3.3, we can obtain all the unknown parameters. However, these parameters are not the optimal solution because the subproblems are kinds of approximation of the original problem. In order to increase the accuracy and optimality, we introduce a root polishing technique so that the parameters strictly satisfies the KKT condition.

Let us recall the original PnPfr problem, Eq. (4). Since \mathbf{a}_i , \mathbf{b}_i , and \mathbf{c}_i are linearly independent, we can equivalently rewrite the cost function of Eq. (4) as

$$\min_{\mathbf{R}, \mathbf{t}, f, \mathbf{k}} \sum_{i=1}^n (\mathbf{a}_i^\top \mathbf{q}_i)^2 + (\mathbf{b}_i^\top \mathbf{q}_i)^2 = \|\mathbf{C}_{(f, \mathbf{k})} \mathbf{r} + \mathbf{D}_{(f, \mathbf{k})} \mathbf{t}\|^2, \quad (30)$$

where $\mathbf{C}_{(f, \mathbf{k})}$ and $\mathbf{D}_{(f, \mathbf{k})}$ are $n \times 9$ and $n \times 3$ coefficient matrices containing the unknowns f and \mathbf{k} , respectively. Due to limitations of space, we describe the details of the formulation in Appendix C in the supplemental material.

Similarly to the first subproblem in Section 3.2, we can express \mathbf{t} as a function of the other unknowns,

$$\mathbf{t}_{(r, f, \mathbf{k})} = -(\mathbf{D}^\top \mathbf{D})^{-1} \mathbf{D}^\top \mathbf{C} \mathbf{r}. \quad (31)$$

Here, we omitted the subscript (f, \mathbf{k}) of \mathbf{C} and \mathbf{D} for simple notations. Then, plugging Eq. (31) into Eq. (30), we obtain a new constrained problem

$$\begin{aligned} \min_{\mathbf{r}, f, \mathbf{k}} \quad & \mathbf{r}^\top \mathbf{G}_{(f, \mathbf{k})} \mathbf{r} \\ \text{s.t.} \quad & \mathbf{R}^\top \mathbf{R} = \mathbf{I}, \quad \det(\mathbf{R}) = 1, \end{aligned} \quad (32)$$

where

$$\mathbf{G}_{(f, \mathbf{k})} = \mathbf{C}^\top \mathbf{C} - \mathbf{C}^\top \mathbf{D} (\mathbf{D}^\top \mathbf{D})^{-1} \mathbf{D}^\top \mathbf{C}. \quad (33)$$

Root polishing is performed to find the optimal solution of Eq. (32), a constrained problem, with initial guess from the first and the second subproblems. A typical and easy way to solve Eq. (32) is convert the constrained problem into an unconstrained problem by expressing the rotation matrix with Euler angle or Cayley transform. However, those representations cannot be uniquely determined in the singularity case, which often happens in real camera motions. Alternative way is to solve new equations, which are equivalent to the original KKT condition without any Lagrange multipliers. Introducing Nakano's approach [7] for PnP problem, we obtain such new equations as follows:

$$\mathbf{R}^\top \text{mat}(\mathbf{G} \mathbf{r}) - \text{mat}(\mathbf{G} \mathbf{r})^\top \mathbf{R} = \mathbf{0}_{3 \times 3}, \quad (34)$$

$$\text{mat}(\mathbf{G} \mathbf{r}) \mathbf{R}^\top - \mathbf{R} \text{mat}(\mathbf{G} \mathbf{r})^\top = \mathbf{0}_{3 \times 3}, \quad (35)$$

$$\mathbf{R}^\top \mathbf{R} - \mathbf{I} = \mathbf{0}_{3 \times 3}, \quad (36)$$

$$\det(\mathbf{R}) - 1 = 0, \quad (37)$$

$$\frac{\partial}{\partial f} \mathbf{r}^\top \mathbf{G} \mathbf{r} = 0, \quad (38)$$

$$\frac{\partial}{\partial \mathbf{k}} \mathbf{r}^\top \mathbf{G} \mathbf{r} = \mathbf{0}_{3 \times 1}. \quad (39)$$

Here, $mat(\)$ is a reshaping operator from a 9×1 vector to a 3×3 square matrix. As Nakano proved in [7], the above equations hold for any types of rotation parameterization instead of Eqs. (36) and (37), e.g., quaternion.

We can solve the system of nonlinear equations, Eq. (34) through Eq. (39), by a simple Gauss-Newton method. An important thing to note here is that numerical differentiation is required in the Gauss-Newton iteration for PnPfr problem because $\mathbf{G}_{(f,\mathbf{k})}$, Eq. (33), cannot be analytically represented by the two unknowns, f and \mathbf{k} . On the other hand, in the case of PnP and PnPf problems, we can compute \mathbf{C} and \mathbf{D} without the unknowns and do not need to update \mathbf{G} in the iteration. The details of the formulation are also described in Appendix C in the supplemental material. This procedure takes less than 10 iterations in almost all cases as long as we have tested. After the convergence, we can recover \mathbf{t} according to Eq. (31).

4 Experiments on synthetic data

Using synthetic data, we have evaluated the proposed PnP, PnPf and PnPfr solvers on accuracy with respect to varying the number of the points n and varying zero-mean Gaussian noise with standard deviation σ on image points. All tests were executed on Core i7-6700 with 16GB RAM on MATLAB 2015b.

In this section, we call our PnP, PnPf, and PnPfr solvers as VPnP, VPnPf, and VPnPfr. The proposed solvers were compared with the following existing methods; EPnP+GN [4], OPnP [6], UPnP [8] for PnP problem, and DLT [22], GPnPf+GN [13], EPnPfR [14]¹ for PnPf and PnPfr problems. We used the original MATLAB code available on the web except for UPnP written in C++.

Due to limitations of space, we will discuss non-planar scene only. However, the conclusion and the tendency of the methods would not change if tested with planar scene. We generated randomly distributed 3D points in the x-, y-, and z-range of $[-2, 2] \times [-2, 2] \times [4, 8]$. Then, those points are projected onto a virtual camera with image resolution 640×480 [pixels], focal length 800 [pixels], principal point at the coordinate $[320, 240]$. For evaluating VPnPfr, we distorted image points by small radial distortion $[k_1, k_2, k_3] = [-0.1, 0, 0]$, and compared with the conventional PnPf solvers assuming zero distortion. In the case of PnPf and PnPfr problems, as suggested in [15], we scaled image points with a factor of $2/\max(\text{width}, \text{height})$ so that all points have normalized coordinates between ± 1 . The ground-truth rotation and translation of the camera are randomly generated. We measured the relative error of estimated parameters except for the rotation matrix. The rotation error was the absolute error given by $\max_{k \in \{1,2,3\}} \cos^{-1}(\mathbf{r}_k^T \mathbf{r}_{k,true})$ [degrees], where \mathbf{r}_k and $\mathbf{r}_{k,true}$ are k -th column of the estimated and the ground-truth rotation matrices, respectively. We performed 500 independent trials for each test.

¹ EPnPfR is not a method for PnPfr problem but an extension of EPnP to PnPf problem. Although the meaning of "fR" is not mentioned in the paper, it is inferred "focal length" and "Regularization" for EPnP.

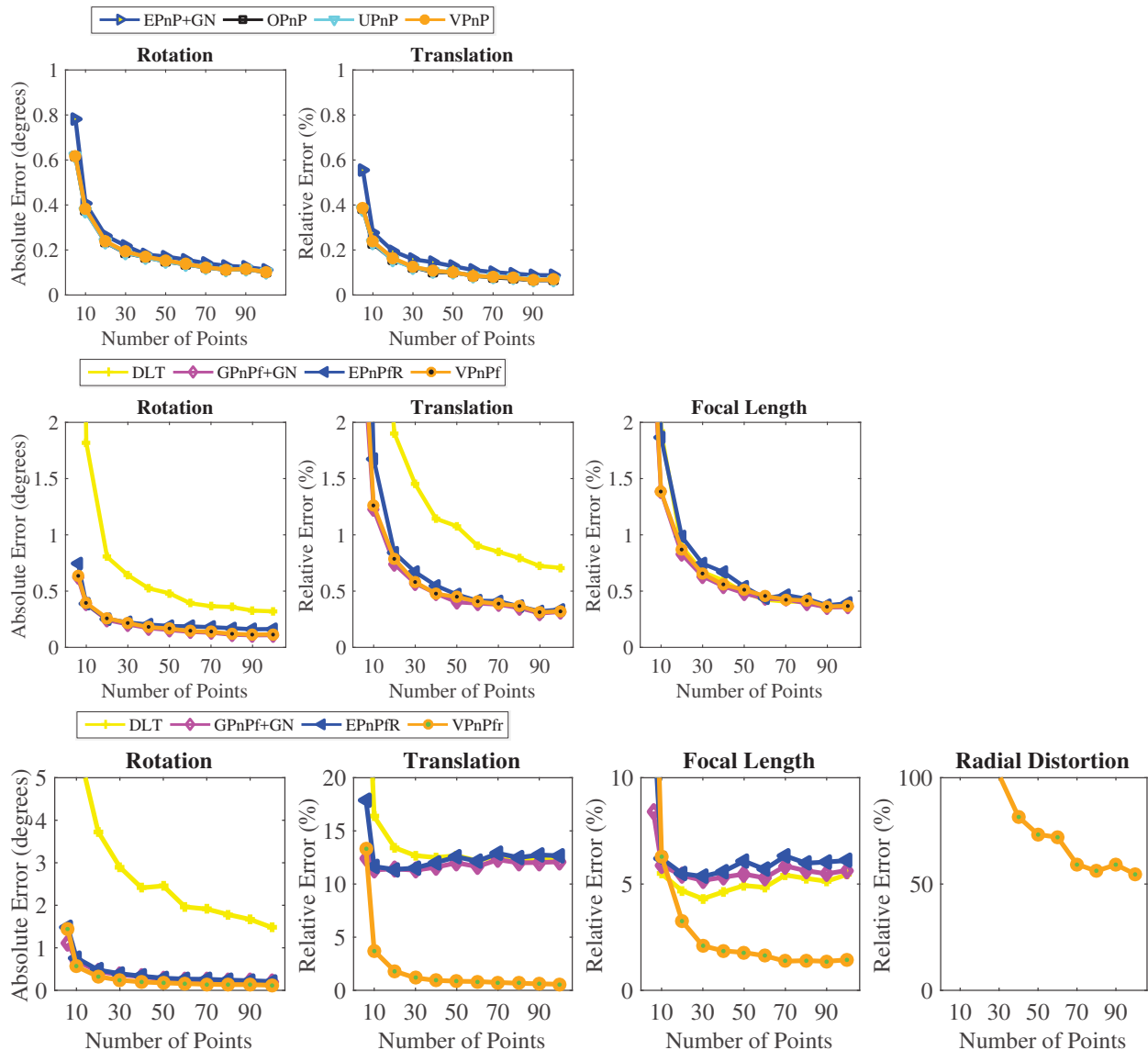


Fig. 1. Median error w.r.t. varying number of points ($6 \leq n \leq 100$) with fixed image noise ($\sigma = 2$). *Top:* PnP problem. *Middle:* PnPf problem. *Bottom:* PnPfr problem

4.1 Accuracy w.r.t. varying number of points

We configured $6 \leq n \leq 100$ and $\sigma = 2$ in this experiment. The reason for starting by $n = 6$ is that DLT and EPNP+GN cannot work on $n = 5$. Fig. 1 shows the median errors of PnP, PnPf, and PnPfr solvers.

In the case of PnP and PnPf problems (top and middle in Fig. 1), most of the solvers except for DLT have same performance. This result shows that the proposed approach, which sequentially solves subproblems, gives globally optimal solution as the existing methods do.

As shown in the bottom plots in Fig. 1, VPnPfr outperforms the other methods in the case of distorted image points. Interestingly, the existing PnPf solvers cannot improve the accuracy of translation and focal length with increasing n , whereas the rotation error becomes lower. The result of the intrinsic parameters implies that VPnPfr requires $n \geq 100$ for focal length and radial distortion estimation to converge the optimal solution on $\sigma = 2$.

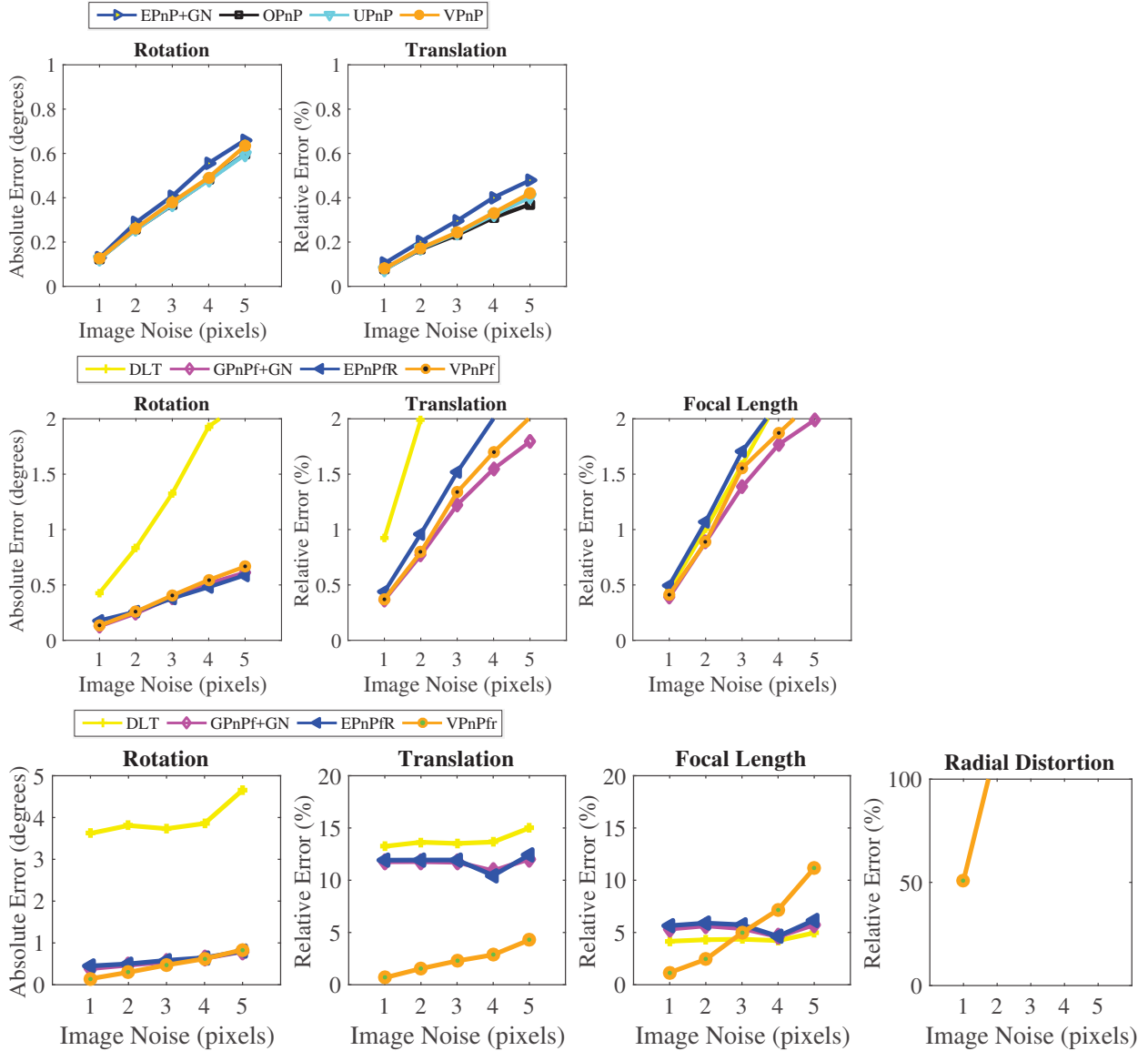


Fig. 2. Median error w.r.t. varying image noise ($1 \leq \sigma \leq 5$) with fixed number of points ($n = 20$). *Top:* PnP problem. *Middle:* PnPf problem. *Bottom:* PnPfr problem

4.2 Accuracy w.r.t. varying image noise

In the next experiment, we have studied the accuracy with respect to varying $1 \leq \sigma \leq 5$ in the case of fixed $n = 20$. The median errors of PnP, PnPf, and PnPfr solvers are shown on top, middle, and bottom, in Fig. 2, respectively.

Similarly to the previous experiment in Section 4.1, our VPnP and VPnPf have comparable performance to the state-of-the-art solvers in PnP and PnPf problems. In addition to that, we can also observe an interesting result in the case of PnPfr problem.

As the image noise increases, focal length estimation of VPnPfr becomes worse than that of PnPf solvers, which assume zero distortion. The image noise seems to affect translation rather than focal length for PnPf solvers. From this, if we need only focal length, PnPf solvers might be more suitable than VPnPfr when the image points have potentially large errors. However, VPnPfr is still the best method for estimating the intrinsic and extrinsic parameter simultaneously.

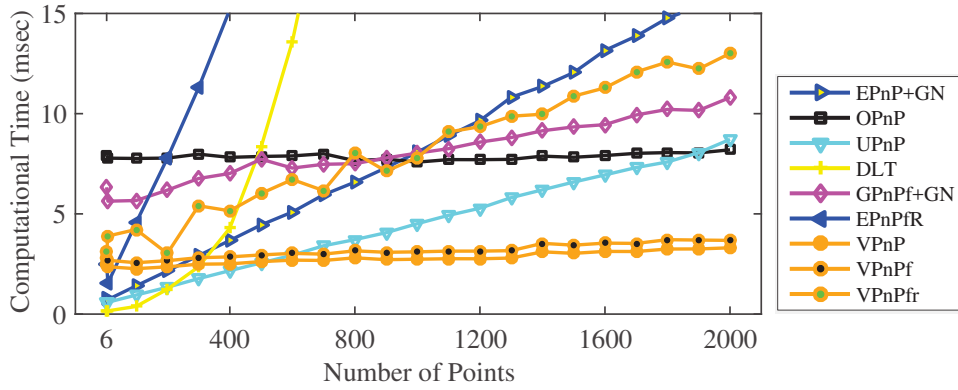


Fig. 3. Computational time w.r.t. varying number of points ($6 \leq n \leq 2000$) with fixed image noise ($\sigma = 2$)

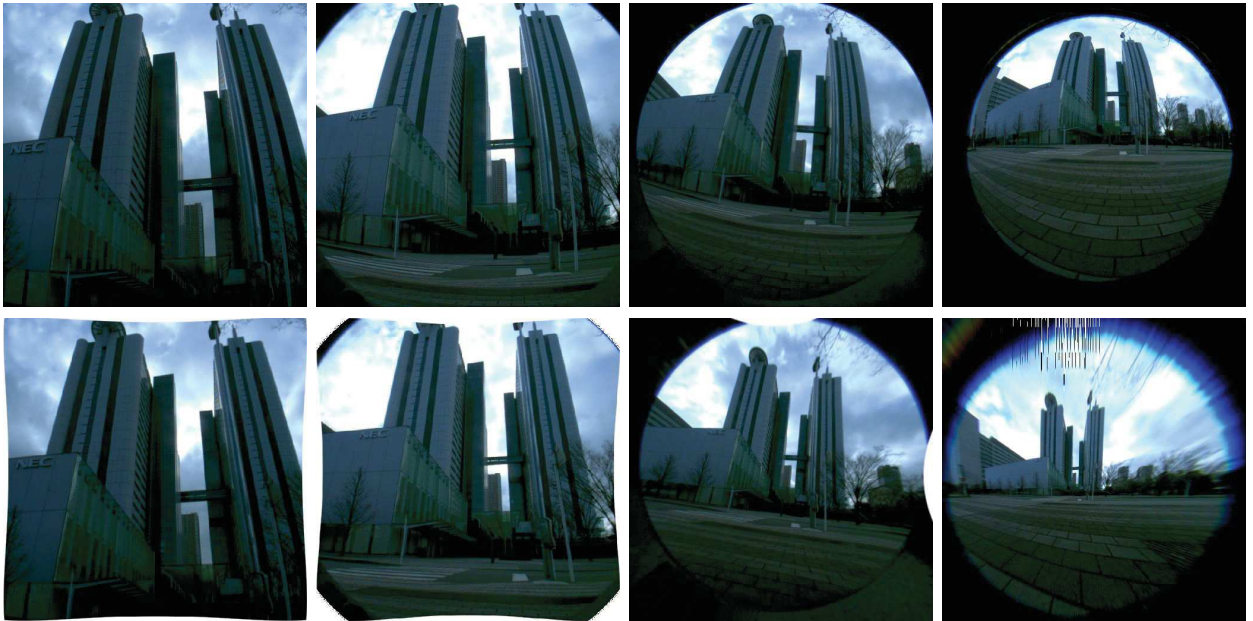


Fig. 4. Results of the proposed PnPfr solver on real images. *Top:* Original images with small (1st column, 58° HFOV) to significant (4th column, 118° HFOV) distortion. *Bottom:* Undistorted images corresponding to the original images on the top row

4.3 Computational time w.r.t. varying number of points

We measured the computational time with $6 \leq n \leq 2000$ and $\sigma = 2$. Fig. 3 shows the average time. Note that UPnP is a mex implementation.

The proposed VPnP and VPnPf take less than 3 msec which is sufficiently fast for real-time applications. Moreover, these runtime increases moderately, almost in $O(1)$, even in the thousands of the points. This is the fastest for large number of points, $n \geq 400$. In contrast, the runtime of VPnPfr grows $O(n)$. The solver for the first subproblem is completely same, therefore, the difference is caused by the root polishing. In the current implementation, updating the matrix \mathbf{G} is required in every Gauss-Newton iteration for VPnPfr, but only once for VPnP and VPnPf as shown in Appendix C in the supplemental material. We can expect that the runtime of VPnPfr becomes closed to that of VPnP and VPnPf if we introduce a more optimized implementation on the root polishing.

5 Experiments on real data

We have tested the proposed PnPfr solver to remove lens distortion on real images. We mounted an ultra-wide vari-focal lens, TAMRON 12VM412ASIR, on a USB 3.0 camera, iDS UI-3370CP-C-HQ. The focal length of the lens is manually changeable from 4.0 mm to 12 mm. This is equivalent to the horizontal field of view from 58° to 118° for the camera. To obtain 2D-3D point correspondences for calculating the camera parameters, we took a single image of a 6×9 checkerboard pattern for each scene and detected corners by libcbdetect [23].

Fig. 4 shows the original distorted and the undistorted images on the top and the bottom rows, respectively. Straight lines of buildings and brick patterns on the road are successfully corrected even with a significant distortion.

6 Conclusions

In this paper, we have proposed a versatile approach for solving PnP, PnPf, and PnPfr problems from $n \geq 5$ point correspondences. The proposed PnPfr solver is the first method for PnPfr problem in the least-squares sense. Based on the derivation of the PnPfr solver, we also have formulated PnP and PnPf solvers in the same theoretical manner, which can be implemented with slight changes from the PnPfr solver. By evaluating the proposed methods on synthetic data, we have shown that the PnP and PnPf solvers have the same performance with the-state-of-the-art methods for undistorted points. Moreover, the PnP and PnPf solvers are the fastest for large point set, $n \geq 400$. On a real image experiment using an ultra-wide zoom camera, the novel PnPfr solver have corrected significant lens distortion corresponding to 118° HFOV. Future works of the PnPfr solver are to improve the runtime of root polishing and accuracy of the distortion coefficients for high image noise.

References

1. 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) (1981) 381–395
2. Haralick, B.M., Lee, C.N., Ottenberg, K., Nölle, M.: Review and analysis of solutions of the three point perspective pose estimation problem. *International journal of computer vision* **13**(3) (1994) 331–356
3. Wu, Y., Hu, Z.: Pnp problem revisited. *Journal of Mathematical Imaging and Vision* **24**(1) (2006) 131–141
4. Lepetit, V., Moreno-Noguer, F., Fua, P.: Epnp: An accurate $o(n)$ solution to the pnp problem. *International journal of computer vision* **81**(2) (2009) 155–166
5. Hesch, J.A., Roumeliotis, S.I.: A direct least-squares (dls) method for pnp. In: *Computer Vision (ICCV), 2011 IEEE International Conference on, IEEE (2011)* 383–390
6. Zheng, Y., Kuang, Y., Sugimoto, S., Astrom, K., Okutomi, M.: Revisiting the pnp problem: A fast, general and optimal solution. In: *Computer Vision (ICCV), 2013 IEEE International Conference on, IEEE (2013)* 2344–2351

7. Nakano, G.: Globally optimal dls method for pnp problem with cayley parameterization. In: Proceedings of the British Machine Vision Conference (BMVC), BMVA Press (September 2015) 78.1–78.11
8. Kneip, L., Li, H., Seo, Y.: Upnp: An optimal $\mathcal{O}(n)$ solution to the absolute pose problem with universal applicability. In: Computer Vision–ECCV 2014. Springer (2014) 127–142
9. Abidi, M.A., Chandra, T.: A new efficient and direct solution for pose estimation using quadrangular targets: Algorithm and evaluation. Pattern Analysis and Machine Intelligence, IEEE Transactions on **17**(5) (1995) 534–538
10. Bujnak, M., Kukulova, Z., Pajdla, T.: A general solution to the p4p problem for camera with unknown focal length. In: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, IEEE (2008) 1–8
11. Wu, C.: P3.5p: Pose estimation with unknown focal length. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 2440–2448
12. Penate-Sanchez, A., Andrade-Cetto, J., Moreno-Noguer, F.: Exhaustive linearization for robust camera pose and focal length estimation. Pattern Analysis and Machine Intelligence, IEEE Transactions on **35**(10) (2013) 2387–2400
13. Zheng, Y., Sugimoto, S., Sato, I., Okutomi, M.: A general and simple method for camera pose and focal length determination. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2014) 430–437
14. Kanaeva, E., Gurevich, L., Vakhitov, A.: Camera pose and focal length estimation using regularized distance constraints. In: Proceedings of the British Machine Vision Conference (BMVC), BMVA Press (September 2015) 162.1–162.12
15. Josephson, K., Byrod, M.: Pose estimation with radial distortion and unknown focal length. In: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE (2009) 2419–2426
16. Bujnak, M., Kukulova, Z., Pajdla, T.: New efficient solution to the absolute pose problem for camera with unknown focal length and radial distortion. In: Computer Vision–ACCV 2010. Springer (2010) 11–24
17. Kukulova, Z., Bujnak, M., Pajdla, T.: Real-time solution to the absolute pose problem with unknown radial distortion and focal length. In: Proceedings of the IEEE International Conference on Computer Vision. (2013) 2816–2823
18. Fitzgibbon, A.W.: Simultaneous linear estimation of multiple view geometry and lens distortion. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. Volume 1., IEEE (2001) I–125
19. Bukhari, F., Dailey, M.N.: Automatic radial distortion estimation from a single image. Journal of Mathematical Imaging and Vision **45**(1) (2013) 31–45
20. Wildenauer, H., Micusik, B.: Closed form solution for radial distortion estimation from a single vanishing point. In: Proceedings of the British Machine Vision Conference (BMVC), BMVA Press (2013)
21. Kukulova, Z., Bujnak, M., Pajdla, T.: Automatic generator of minimal problem solvers. In: Computer Vision–ECCV 2008. Springer (2008) 302–315
22. Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Second edn. Cambridge University Press, ISBN: 0521540518 (2004)
23. Geiger, A., Moosmann, F., Car, O., Schuster, B.: Automatic camera and range sensor calibration using a single shot. In: International Conference on Robotics and Automation (ICRA), St. Paul, USA (May 2012)