



Efficient Radial Distortion Correction for Planar Motion

Marcus Valtonen Örnhag^(✉) 

Centre for Mathematical Sciences, Lund University, Lund, Sweden
`marcus.valtonen_ornhag@math.lth.se`

Abstract. In this paper we investigate simultaneous radial distortion calibration and motion estimation for vehicles travelling parallel to planar surfaces. This is done by estimating the inter-image homography between two poses, as well as the distortion parameter. Radial distortion correction is often performed as a pre-calibration step; however, accurately estimating the distortion profile without special scene requirements may make such procedures obsolete. As many modern day consumer cameras are affected by radial distortion to some degree, there is a great potential to reduce production time, if properly implemented.

We devise two polynomial solvers, for radially distorted homographies compatible with different models of planar motion. We show that the algorithms are numerically stable, and sufficiently fast to be incorporated in a real-time frameworks. Furthermore, we show on both synthetic and real data, that the proposed solvers perform well compared to competing methods.

Keywords: Radial distortion correction · Homography · Visual odometry · Polynomial solvers

1 Introduction

When designing a Visual Odometry (VO) pipeline it is beneficial to integrate any prior knowledge of the intended environment or known motion model parameters. One particular instance, that will be further investigated in this paper, is the planar motion model, in which a vehicle travels on—or parallel to—a planar surface. Such a scenario is common in man-made environments, but can also accurately approximate outdoor scenarios under certain conditions, such as cars travelling on a highway. In the current literature we find several papers on planar motion models, restricted to fit particular use cases or pre-calibrated parameters [3, 6, 20]. The general case, however, was first introduced in [17] which

The author gratefully acknowledges Mårten Wadenbäck and Martin Karlsson for providing the data for the planar motion compatible sequences, and Magnus Oskarsson for fruitful discussions regarding the basis selection heuristic which made the proposed solvers faster. This work has been funded by the Swedish Research Council through grant no. 2015-05639 ‘Visual SLAM based on Planar Homographies’.

incorporates two unknown overhead tilt angles, which are assumed to be constant throughout the trajectory of the vehicle. They assumed the floor is in the field of view of the camera, allowing them to compute the motion parameters through inter-image homographies. Another approach utilizing the floor to compute the motions was done by [6]. More recent development was done by [26, 27], and is the first to accurately recover the complete set of motion parameters using inter-image homographies. Other notable approaches include that of [30], in which a planar VO pipeline using a dense matching scheme was proposed.

In the most general setting, assuming the camera is rigidly mounted on the vehicle, the number of motion parameters are reduced to five, which should be compared to the general homography, which has eight degrees of freedom [7]. These parameters consist of the two overhead tilt parameters, and three non-constant parameters: one rotational angle (about the floor normal), and two translational components.

In order to obtain a homography, keypoints are extracted and matched. These keypoints then serve as input to the homography estimation algorithm. Since the extraction and matching steps are imperfect for any realistic image sequence, outliers and noise are prone to exist. Typical steps taken to resolve this issue include the use of robust estimation frameworks, e.g. RANSAC. It is at this point the benefit of working with fewer motion parameters come to light. Since fewer motion parameters demand fewer point correspondences in order to be estimated, one can select a minimal amount of points in the RANSAC framework. The fewer points you are able to select, the greater the probability of selecting only inliers. By doing so, one can reduce the number of RANSAC iterations.

In the general case, with four point correspondences, one may linearly extract the homography; however, if any of the motion parameters are known or constrained, this may no longer be the case, as the resulting systems of equations often are nonlinear. This poses a new type of problem—can we solve these equations sufficiently fast and accurate? Luckily, many methods from computational algebraic geometry [4] has been used in many computer vision problems, and certain frameworks already exist for how to proceed. One of the earliest, and still used today, was [9].

This paper is a revised journal version of [25], where we will consider the general planar motion model with unknown radial distortion, and devise a polynomial solver that can accurately recover the motion and distortion parameters in real-time applications. Furthermore, we propose a planar motion compatible minimal two point solver with radial distortion when the tilt angles are known. This situation arises when the tilt is pre-calibrated or can be accessed using external sensors, such as an IMU to extract the gravity direction. For indoor scenarios, assuming that the gravity direction is aligned with the floor normal—which often is a valid approximation—is equivalent to knowing the tilt angles. There are situations where similar assumptions can be made, without significant loss in accuracy, e.g. aerial imagery. Regardless of the situation, radial distortion is necessary to account for in any accurate VO pipeline, and is often done in a pre-calibration step, where the distortion parameters are obtained. By incorporating the parameter in the homography estimation process, we hope to eliminate this pre-processing step.

2 Related Work

2.1 Homography Estimation

The *Direct Linear Transform* (DLT) equations is a linear system of equations to extract a homography \mathbf{H} given a number of point correspondences. In the general setting, with eight degrees of freedom, the minimal case requires four point correspondences. To see this, let us consider a single pair of point correspondences on a common scene plane, denoted by $\mathbf{x} \leftrightarrow \hat{\mathbf{x}}$. They are related by a homography \mathbf{H} as $\lambda \hat{\mathbf{x}} = \mathbf{H}\mathbf{x}$, for some scalar $\lambda \neq 0$. Equivalently, we may express this as $\hat{\mathbf{x}} \times \mathbf{H}\mathbf{x} = \mathbf{0}$, thus eliminating the scale parameter λ , or,

$$\begin{bmatrix} \mathbf{0} & -\hat{w}\mathbf{x}^T & \hat{y}\mathbf{x}^T \\ \hat{w}\mathbf{x}^T & \mathbf{0} & -\hat{x}\mathbf{x}^T \\ -\hat{y}\mathbf{x}^T & \hat{x}\mathbf{x}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \end{bmatrix} = \mathbf{0}, \quad (1)$$

assuming homogeneous coordinates, i.e. $\mathbf{x} = [x, y, w]^T$ and $\hat{\mathbf{x}} = [\hat{x}, \hat{y}, \hat{w}]^T$, respectively. Here \mathbf{h}_k^T is the k :th row of the homography matrix \mathbf{H} . As the cross product introduces a linear dependence, only two of the equations are necessary, hence explaining why four point correspondences are minimal in the general case. Thus, using four point correspondences the problem can be transformed into finding the one-dimensional null space $\mathbf{h} = [\mathbf{h}_1^T \ \mathbf{h}_2^T \ \mathbf{h}_3^T]^T$, which is typically obtained using SVD of the coefficient matrix.

In the general planar motion model there are only five motion parameters, hence the minimal case requires but 2.5 point correspondences. Similarly, we may construct a system of equations, by using three point correspondences and discard the last equation in the corresponding DLT system. This can be written as $\mathbf{C}\mathbf{h} = \mathbf{0}$ where $\mathbf{C} \in \mathbb{R}^{5 \times 9}$ is the coefficient matrix. Again, this is a problem of finding the null space of \mathbf{C} ; however, the null space is now four-dimensional. As an additional step, one must now find the null space coefficients which makes \mathbf{H} a homography compatible with the general planar motion model. It was shown in [24, 29] that there are eleven quartic constraints (as well as a sextic constraint) in the elements of \mathbf{H} that has to be fulfilled in order to guarantee compatibility.

2.2 Modelling Radial Distortion

In order to compensate for the radial distortion, several models have been proposed. A classic method, still in use today, is the Brown–Conrady model [2], in which also tangential distortion is corrected. The division model introduced in [5], has gained attention as it provides accurate approximations of the distortion profile with fewer parameters. For this reason, we will only consider the distortion model, and restrict ourselves to a single distortion parameter, as this allows us to use fewer point correspondences.

Let λ denote the distortion parameter. Then the distorted (or measured) image points can be expressed as $\mathbf{x}_i = [x_i, y_i, 1]^T$

$$\mathbf{x}_i^u = f(\mathbf{x}_i, \lambda) = \begin{bmatrix} x_i \\ y_i \\ 1 + \lambda(x_i^2 + y_i^2) \end{bmatrix}, \quad (2)$$

where \mathbf{x}_i^u are the undistorted image points, assuming the distortion center is aligned to the center of the image. Furthermore, we select the coordinate system such that the origin is aligned with the distortion center.

We may now modify the DLT equations (1) as the distortion parameter only appears in the homogeneous coordinates. Consider two point correspondences $\mathbf{x}_i \leftrightarrow \hat{\mathbf{x}}_i$, then

$$f(\hat{\mathbf{x}}_i, \lambda) \times \mathbf{H}f(\mathbf{x}_i, \lambda) = \mathbf{0}. \quad (3)$$

This approach has been used for the general case of radially distorted homographies [11], conjugate translations with radial distortion [21], and the case of jointly estimating lens distortion and affine rectification from coplanar features [22]. The last two use an explicit parameterization of the motion parameters, instead of trying to parameterize the null space of the DLT system. In common for all methods is that the resulting problem is a polynomial system of equations, and is solved by further reduction to an eigenvalue problem [4]. Automatic solvers for polynomial systems have been proposed, primarily using Gröbner bases, such as [9, 12–14, 16], or resultant based methods [1]. Alternative approaches include considering the problem as a Quadratic Eigenvalue Problem (QEP) [5, 8, 10].

3 The General Planar Motion Model

Consider a camera mounted rigidly on a vehicle travelling on a planar surface. We model this scenario by assuming that the camera moves in the plane $z = 0$, parallel to the surface on which the vehicle moves, located in $z = 1$. This parameterization also fixes the scale of the global coordinate system.

Consider two consecutive views, A and B , with the corresponding camera matrices

$$\begin{aligned} \mathbf{P}_A &= \mathbf{R}_{\psi\theta}[\mathbf{I} \mid \mathbf{0}], \\ \mathbf{P}_B &= \mathbf{R}_{\psi\theta}\mathbf{R}_\phi[\mathbf{I} \mid -\mathbf{t}], \end{aligned} \quad (4)$$

where the constant overhead tilt is modeled by $\mathbf{R}_{\psi\theta}$, and consists of a rotation θ about the y -axis followed by a rotation of ψ about the x -axis. Furthermore, we allow the vehicle to rotate an angle ϕ about the z -axis, which may vary. As the camera is assumed to be mounted rigidly on the vehicle, the height above the floor is constant, hence we may assume that it travels in the plane $z = 0$, leaving two translation components t_x and t_y , see Fig. 1. From this, one may derive the corresponding inter-image homography

$$\mathbf{H} = \lambda\mathbf{R}_{\psi\theta}\mathbf{R}_\phi\mathbf{T}_t\mathbf{R}_{\psi\theta}^T, \quad (5)$$

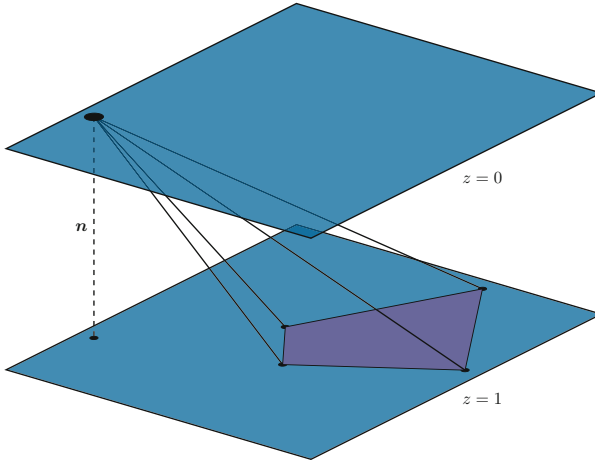


Fig. 1. Illustration of the problem geometry considered in the paper. The camera is mounted rigidly on a mobile platform, thus travelling parallel to the ground floor in the plane $z = 0$. We allow a constant, but generally unknown, overhead tilt to be present, which is modelled by the angles ψ (about the x -axis and θ (about the y -axis). Furthermore, the camera can rotate about the z -axis (by an angle ϕ) and translate in the plane $z = 0$, i.e. there are in total five degrees of freedom—three rotations and two translations. Figure reproduced from [25].

where $\mathbf{T}_t = \mathbf{I} - t\mathbf{n}^T$ is a translation matrix, for the translation $\mathbf{t} = [t_x, t_y, 0]^T$, relative the plane normal $\mathbf{n} = [0, 0, 1]^T$. The homography matrix can be made unique by e.g. imposing $\det(\mathbf{H}) = 1$.

In addition to the DLT constraints, the elements of a homography compatible with the general planar motion model must satisfy a number of polynomial constraints. Such constraints were numerically derived in [29], where it was shown that there are at least eleven quartic constraints. The novel theoretical framework used in [24], showed that these constraints were necessary, but not sufficient; however, by adding a sextic constraint, it was shown that they are sufficient.

4 Polynomial Solvers

4.1 A Non-minimal Relaxation (4 Point)

In theory, one would be able to construct a minimal solver with three point correspondences, as there are six degrees of freedom—the five motion parameters discussed in Sect. 3, and the distortion parameter. In practice, however, this problem is hard, and we have yet to find a tractable solution which is numerically stable and sufficiently fast for real-time applications. Consequently, we have opted for a non-minimal four point relaxation. We do believe this is an acceptable compromise, as a general homography with a single distortion parameter requires 4.5 point correspondences for the minimal configuration. This effectively

means one has to sample five point pairs to estimate a hypothesis. This section is largely reproduced from [25].

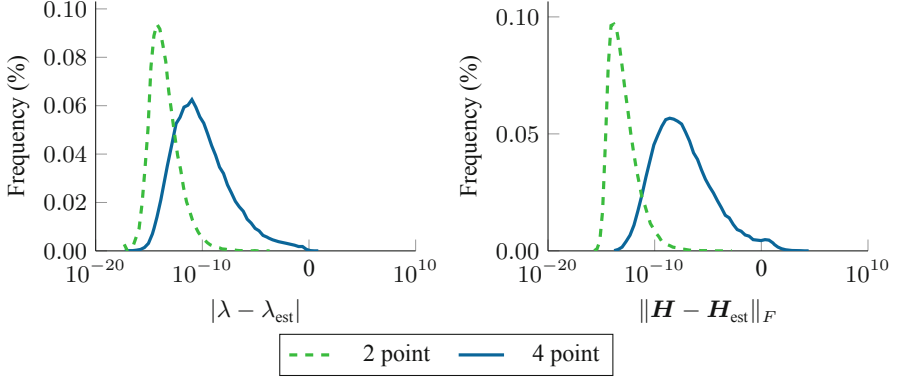


Fig. 2. Error histogram of the estimated distortion parameter λ (left) and the homography \mathbf{H} for 100,000 random instances, for both of the proposed methods.

Similarly, to the approach in [11] we expand the third row of (3); however, we consider using only four point correspondences. This results in the following equation

$$(-\hat{y}_i h_{11} + \hat{x}_i h_{21})x_i + (-\hat{y}_i h_{12} + \hat{x}_i h_{22})y_i + (-\hat{y}_i h_{13} + \hat{x}_i h_{23})w_i = 0, \quad (6)$$

where $w_i = 1 + \lambda(x_i^2 + y_i^2)$ and $\hat{w}_i = 1 + \lambda(\hat{x}_i^2 + \hat{y}_i^2)$ are functions of the radial distortion parameter λ . There are eight monomials involved in this expression, namely

$$\mathbf{v}_1 = [h_{11} \ h_{12} \ h_{13} \ h_{21} \ h_{22} \ h_{23} \ \lambda h_{13} \ \lambda h_{23}]^T. \quad (7)$$

Using four point correspondences results in a system of equations, which can be written as

$$\mathbf{M}_1 \mathbf{v}_1 = \mathbf{0}, \quad (8)$$

where \mathbf{M}_1 is a 4×8 matrix. For non-degenerate configurations the null space of \mathbf{M}_1 is four-dimensional. Consequently, we may parameterize \mathbf{v}_1 as

$$\mathbf{v}_1 = \sum_{i=1}^4 \gamma_i \mathbf{n}_i, \quad (9)$$

where γ_i are unknown basis coefficients. Since the last two monomials of \mathbf{v}_1 depend on the previous elements, this relation has to be enforced when computing the basis coefficients γ_i . These give rise to two equations

$$v_8 = \lambda v_6 \quad \text{and} \quad v_7 = \lambda v_3. \quad (10)$$

Furthermore, we proceed to fix the scale by letting $\gamma_4 = 1$.

We will now use the second row of (3). Similarly, we may write this as

$$\mathbf{M}_2 \mathbf{v}_2 = \mathbf{0}, \tag{11}$$

where $\mathbf{M}_2 \in \mathbb{R}^{4 \times 16}$. Here the null space vector \mathbf{v}_2 consists of seven variables, and 16 monomials: $h_{31}, h_{32}, h_{33}, \lambda h_{33}$ and $\lambda^2 \gamma_i, \lambda \gamma_i, \gamma_i$ for $i = 1, 2, 3$ and

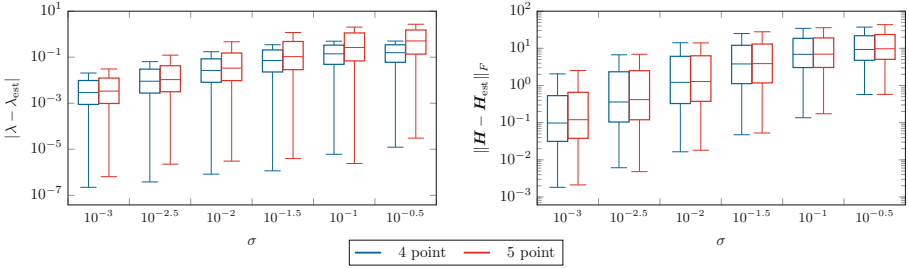


Fig. 3. Distribution of estimation error in the distortion parameter λ , and the homography \mathbf{H} (measured in the Frobenius norm) for different noise levels σ and unknown tilt. The proposed solver is compared to the five point solver [5]. Figure and caption reproduced from [25].

$\lambda^2, \lambda, 1$. We may now proceed to eliminate the first three variables— h_{31}, h_{32} and h_{33} —as they are only present in four monomials. As we are using four point correspondences, yielding four equations, Gauss–Jordan elimination can be used. We obtain the following upon performing the elimination

$$\hat{\mathbf{M}}_2 = \begin{bmatrix} h_{31} & h_{32} & \lambda h_{33} & h_{33} & \lambda^2 \gamma_1 & \lambda \gamma_1 & \gamma_1 & \dots & \lambda^2 & \lambda & 1 \\ 1 & & & & \bullet & \bullet & \bullet & \dots & \bullet & \bullet & \bullet \\ & 1 & & & \bullet & \bullet & \bullet & \dots & \bullet & \bullet & \bullet \\ & & 1 & & \bullet & \bullet & \bullet & \dots & \bullet & \bullet & \bullet \\ & & & 1 & \bullet & \bullet & \bullet & \dots & \bullet & \bullet & \bullet \end{bmatrix}. \tag{12}$$

It turns out that the columns of the right 4×12 submatrix are not independent. In order to generate a correct solver, it is important to generate integer instances satisfying these dependencies.

From the eliminated system $\hat{\mathbf{M}}_2 \mathbf{v}_2 = \mathbf{0}$ we get the four equations

$$\begin{aligned} h_{31} + f_1(\gamma_1, \gamma_2, \gamma_3, \lambda) &= 0, \\ h_{32} + f_2(\gamma_1, \gamma_2, \gamma_3, \lambda) &= 0, \\ \lambda h_{33} + f_3(\gamma_1, \gamma_2, \gamma_3, \lambda) &= 0, \\ h_{33} + f_4(\gamma_1, \gamma_2, \gamma_3, \lambda) &= 0, \end{aligned} \tag{13}$$

where $f_i(\gamma_1, \gamma_2, \gamma_3, \lambda)$ are polynomials in the variables $\gamma_1, \gamma_2, \gamma_3, \lambda$. Exploiting the relations between the last two equations of (13), an additional constraint is obtained

$$\lambda f_4(\gamma_1, \gamma_2, \gamma_3, \lambda) = f_3(\gamma_1, \gamma_2, \gamma_3, \lambda). \tag{14}$$

The eliminated variables h_{31} , h_{32} and h_{33} are polynomials of degree three, thus making (14) of degree four. Together with (10) we have three equations in four unknowns. Since we are able to express all elements of the homography \mathbf{H} as a function of four variables, we can enforce one of the 11 quartic constraints originally found in [29]. Evaluating these constraints using \mathbf{H} it turns out that ten of the constraints are of degree 12 and one of degree 10 due to cancellation of higher order terms. We choose the smallest one to build the polynomial solver.

Using the automatic generator [12] we find that there are 18 solutions to the problem in general, and by sampling a basis based on the heuristic presented in [15] an elimination template of size 177×195 could be created.

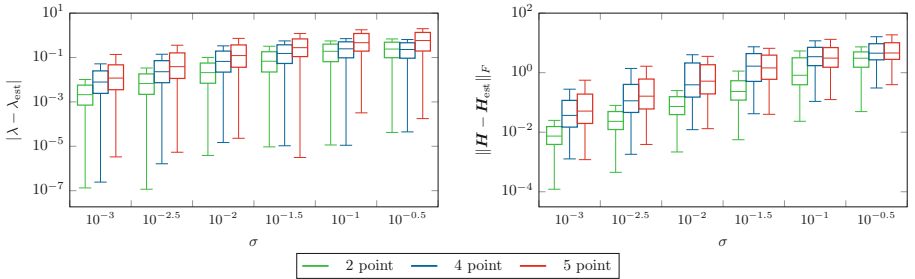


Fig. 4. Distribution of estimation error in the distortion parameter λ , and the homography \mathbf{H} (measured in the Frobenius norm) for different noise levels σ and known tilt (assumed to be compensated for). The proposed two point solver is compared to the four point and five point solver [5]. Distribution of estimation error in the distortion parameter λ , and the homography \mathbf{H} (measured in the Frobenius norm) for different noise levels σ and known tilt (assumed to be compensated for). The proposed two point solver is compared to the four point and five point solver [5].

4.2 Minimal Solver with Known Tilt (2 Point)

If the tilt angles are known, we can treat the planar motion case with radial distortion. In this case there are four degrees of freedom, and thus the minimal configuration requires two point correspondences. In this section, we will derive a novel solver for this case. Using a different approach than in the previous section, we may explicitly parameterize the homography. Let us use the following parameterization for the rotation matrix

$$\mathbf{R}_z = \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (15)$$

where $c^2 + s^2 = 1$, hence the sought homography is given by $\mathbf{H} \sim \mathbf{R}_z + \mathbf{t}\mathbf{n}^T$, where $\mathbf{t} = [t_x, t_y, 0]^T$ is a translation vector and $\mathbf{n} = [0, 0, 1]^T$ is a floor normal. Let us consider the modified DLT equations (3) again, but this time using two point correspondences. Using the first and third rows, we note that there are

in total five unknowns— c , s , t_x , t_y and the radial distortion parameter λ —and in total eleven monomials, hence we may write the system as $\mathbf{M}\mathbf{v} = \mathbf{0}$, where \mathbf{M} is a 4×11 matrix and \mathbf{v} is the vector of monomials. Furthermore, of these eleven monomials, we find only four which contain the variables c and s . Therefore, it is possible to use Gauss–Jordan elimination to eliminate these variables. The corresponding system, after elimination, is on the form

$$\hat{\mathbf{M}} = \begin{bmatrix} \lambda c & c & \lambda s & s & \lambda t_x & t_x & \lambda^2 t_y & \lambda t_y & t_y & \lambda & 1 \\ 1 & & & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & 1 & & & \bullet & \bullet & 0 & \bullet & \bullet & 0 & 0 \\ & & 1 & & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ & & & 1 & \bullet & \bullet & 0 & \bullet & \bullet & 0 & 0 \end{bmatrix}. \quad (16)$$

Notice the pattern of zeros emerging in the eliminated system. This, and other more intricate relations, between the coefficients are necessary to account for in order to create an accurate polynomial solver.

From the above system we may introduce the functions g_i , such that

$$\begin{aligned} \lambda c + g_1(t_x, t_y, \lambda) &= 0, \\ c + g_2(t_x, t_y, \lambda) &= 0, \\ \lambda s + g_3(t_x, t_y, \lambda) &= 0, \\ s + g_4(t_x, t_y, \lambda) &= 0. \end{aligned} \quad (17)$$

where $g_i(t_x, t_y, \lambda)$ are polynomials in the variables t_x , t_y and λ . Furthermore, we utilize the two relations

$$\begin{aligned} g_1(t_x, t_y, \lambda) &= \lambda g_2(t_x, t_y, \lambda), \\ g_3(t_x, t_y, \lambda) &= \lambda g_4(t_x, t_y, \lambda). \end{aligned} \quad (18)$$

The constraint $c^2 + s^2 = 1$ translates into

$$g_2^2(t_x, t_y, \lambda) + g_4^2(t_x, t_y, \lambda) = 1. \quad (19)$$

Now, we have a reduced system with three unknowns— t_x , t_y and λ —given by (18) and (19). It turns out that (18) are cubic and (19) are quartic, and by analyzing the dimension of the corresponding quotient ring, we find that the system has six solutions in total (it can be verified that the original system has six solutions as well). Using [15] an elimination template of size 18×24 was constructed.

5 Experiments

5.1 Synthetic Data

In this section we investigate the numerical stability and noise sensitivity of the proposed solver. We generate synthetic homographies, compatible with the

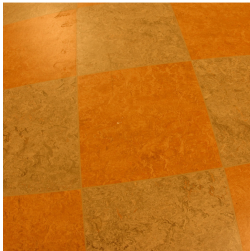
general planar motion model (with and without tilt), as well as distortion parameters. Random scene points are generated using the homography and subsequently distorted using the division model.

The polynomial solvers were generated according to Sect. 4 in C++, and the mean runtime for the 4 point solver is 730 μs and for the 2 point solver 13 μs (measured over 100,000 instances on a standard desktop computer).

5.2 Numerical Stability

By using the described method, we generate noise-free problem instances. Similarly to [11], we use physically reasonable parameters, and cover a wide range of distortions by allowing the distortion parameter λ to be chosen at random in the interval $[-0.7, 0]$. In Fig. 2 we show the error histogram for 100,000 random problem instances. When measuring the Frobenius norm error, the homographies have been normalized to $h_{33} = 1$.

Input



Output



Fig. 5. Two radially distorted images (left) and the rectified and stitched panorama. The distortion parameter and homography was obtained using the proposed solver in a RANSAC framework. Blue border added for visualization. Figure and caption reproduced from [25]. (Color figure online)

From the histogram of the four point solver, we conclude that most parameters are estimated accurately, with an error in the range of 10^{-10} . Such an error is acceptable for most applications; however, some errors are higher,

reaching an error around 10^{-2} . After careful analysis, we attribute this to the ten degree polynomial, which was added to conform with one of the original quartic constraints necessary for making the proposed solver compatible with the general planar motion model. Luckily, errors of the higher magnitude is less frequently occurring, and can be efficiently discarded in a robust framework, such as RANSAC. We will show that this is the case in the coming sections.

For the two point solver the errors are negligible for most computer vision applications, and is also a strong candidate for a robust framework, given that the assumptions of known tilt are met.

5.3 Noise Sensitivity

Similar to the previous section we generate synthetic problem instances, but corrupt the radially distorted image coordinates with Gaussian noise with a variance σ^2 . The noise is varied from mild to severe and at every noise level 10,000 problem instances were generated and the corresponding error measured. As a comparison, the five point method based on the QEP approach [5] was used.

The result is shown in Fig. 3. Note that the mean error for both quantities are lower for the proposed method compared to the five point method, for all noise levels. Analogously, but with known overhead tilt, we compare the two point solver to the other methods, see Fig. 4. Here we see a clear benefit over the other, more general, methods.



Fig. 6. Setup used in the panorama stitching experiment. Figure and caption reproduced from [25].

5.4 Image Stitching

In this section, we use the proposed four point solver in a classic stitching pipeline based on a standard approach for estimating a homography. The pipeline consists of first detecting and extracting SURF keypoints, followed by nearest neighbor matching. From all matched keypoints we select four at random and feed to the proposed solver in a RANSAC framework. The input images are taken using a digital camera with a fish-eye lens mounted on a tripod, overlooking a textured floor, see Fig. 6. The camera tilt was fixed during the experiment, and only the tripod itself was moved, hence generating a motion compatible with the general planar motion model.

The output from the experiment is shown in Fig. 5. Bundle adjustment or other non-linear refinements of the obtained homography was not performed. Apart from being aligned with the correct edges we also note that lines that are straight in reality also appear straight in the final panorama, thus indicating that the radial distortion parameter was correctly estimated.

In terms of the efficiency of the robust framework, we use the same input images and compare the five point solver [5] with the proposed solver. This is done by recording the number of inliers as a function of the number of RANSAC iterations. We repeat the experiment 500 times, and show the average result in Fig. 7, which shows that the proposed method consistently has a higher number of inliers.

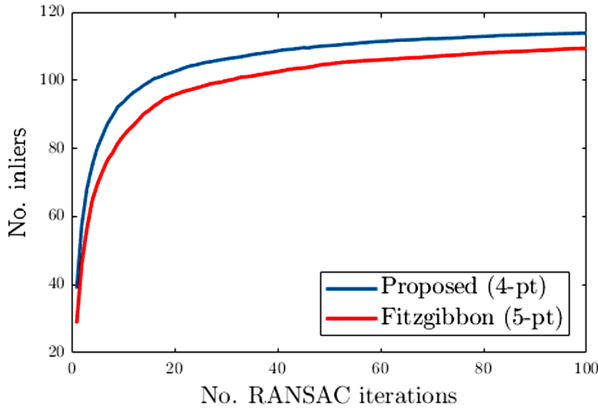


Fig. 7. Number of inliers vs. number of RANSAC iterations for the images in Fig. 5. The data has been averaged over 500 test instances. Figure and caption reproduced from [25].

5.5 Application to Visual Odometry

In this section we use real data from a mobile robot of model Fraunhofer IPA rob@work. The sequence was originally used in [28], but the radial distortion

profile was pre-calibrated. On the mobile robot a camera is mounted rigidly, with an unknown overhead tilt, which excludes the application of the two point solver. The distortion is clearly noticeable and the field of view is almost entirely of the textured floor upon which the robot travels. Furthermore, the robot is equipped with omni-directional wheels, which allows for pure rotations. A reference system with an absolute accuracy of $100\ \mu\text{m}$ tracks the robot as it moves about, and the resulting data is used as ground truth.

We consider three sequences:

- Line.** Forward motion in a straight line with a constant orientation (320 images),
- Turn.** Forward motion while rotating, resulting in a slight turn (344 images),
- Parallel Parking.** Forward motion followed by a sharp turn, while keeping constant rotation (325 images).

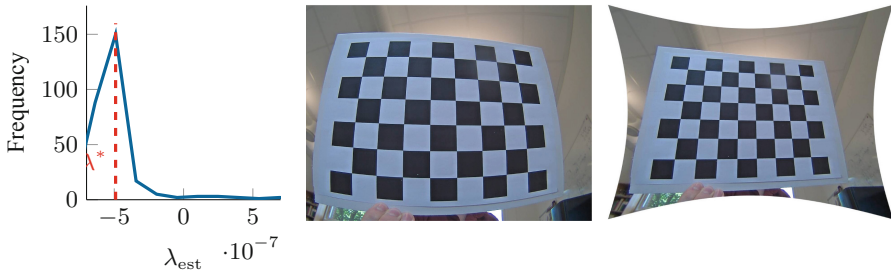


Fig. 8. (Left) Histogram of estimated distortion parameters for the proposed method evaluating during the *parallel parking* sequence. The selected parameter λ^* is marked with a dashed line. (Middle) Undistorted image of a calibration chart, not part of the sequence. (Right) Rectified image using the estimated parameter λ^* . Figure and caption reproduced from [25].

We consider a standard VO pipeline, including an initial solution via homography estimation, from which the initial camera poses are estimated (both intrinsic and extrinsic parameters) and finally a non-linear refinement step using bundle adjustment. Both the proposed method and the five point method [5] are capable of producing an initial estimation through inter-image homographies. Given a pair of consecutive images we may estimate the distortion parameter as well as the homography, using either solver, in a RANSAC framework. To extract the full set of motion parameters, we use the method in [27], hence establishing the initial poses. The estimated robot trajectory can then be extracted and compared to the ground truth. Note that in a complete VO pipeline, the initial position is important in order to avoid excessive amounts of bundle adjustment iterations, as these typically become large-scale optimization problems. Therefore, it is of interest to decrease the number of necessary iterations, by supplying a good initial guess.

The methods are comparable in terms of accuracy, as can be seen in Fig. 9, with a slight preference for the proposed method. As noted in [23, 24], there is no significant boost in performance by pre-optimizing early on in the VO pipeline. One of the main issues is that the constant overhead tilt, due to the camera being rigidly mounted onto the robot, is not enforced throughout the entire trajectory by only considering a single homography. For consistency, one must consider an entire sequence of homographies. Nevertheless, the proposed method benefits from the same performance gain as was noted in Sect. 5.4; namely, that the number of RANSAC iterations required are fewer than for the five point method.

The problem with considering only a single homography also affects the estimation of the radial distortion coefficient. In return, every pair of consecutive images yields a new estimate; however, we know a priori that it is constant through the trajectory. We propose using histogram voting as a robust way to obtain an initial guess. To evaluate the performance we use previously unseen images of calibration charts, that were acquired during the creation of the robot test sequences. We proceed by considering the *parallel parking* test sequence, and use the estimated parameters as a basis for the histogram voting experiment, see Fig. 8. As can be seen, the chosen parameter λ^* , yields an acceptable initial solution, to be refined in a bundle adjustment framework.

5.6 Application to Aerial Imagery

In this final section we test the novel two point solver for aerial imagery. We use the TNT Aerial VideoTestset (TAVT) [19]. In this dataset, video sequences from a UAV have been recorded, at varying flight heights. The onboard global shutter camera is recording in full HDTV resolution at 30 fps, and suffers from mild radial distortion. Although the distortion is not severe, it was shown in [18] that failure to compensate for it results in severely distorted mosaicing attempts.

We use the sequences recorded at higher altitudes, in this case 1000 m and 1500 m above ground, as these are not affected as much by potential non-zero and non-constant tilt, making the two point solver suitable. The solver is incorporated in a RANSAC framework, and the pipeline is identical to previous setups for real images. The sequences are subsampled to include every tenth image of the original sequences, hence contain 117 and 158 images each. The resulting mosaics are shown in Fig. 10. Note that no non-linear optimization has been performed, nor histogram voting to determine the distortion profile. Yet, even for this simple pipeline, we manage to produce visually acceptable results, similar to those of the original articles [18, 19]. Perhaps the only noticeable difference is the lack of blending, seam-finding and other processing involved; however, these artifacts do not stem from the solver.

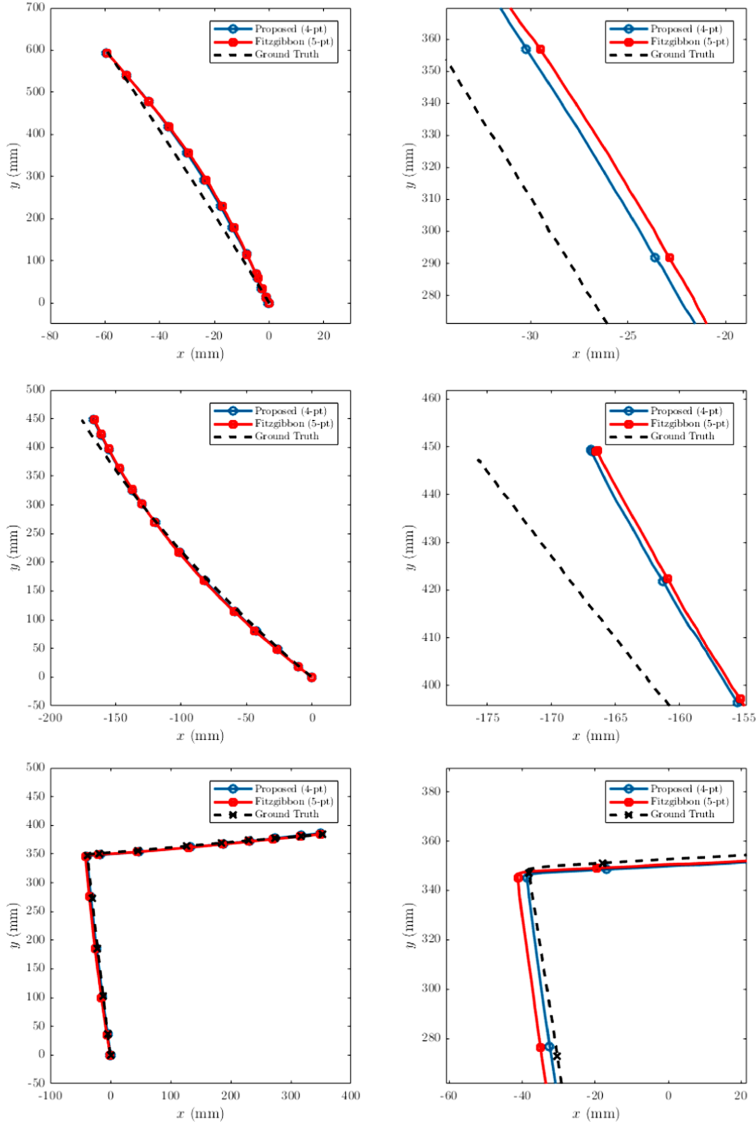


Fig. 9. Estimated trajectories for *line*, *turn* and *parallel parking* of the VO experiment in Sect. 5.5. Images to the left show the entire trajectory, and the ones to the right are zoomed in on a region of interest. Figure and caption reproduced from [25].



Fig. 10. Mosaics from the 1000 m sequence (top) and 1500 m sequence of the TAVT dataset [19] obtained using the proposed two point solver.

6 Conclusions

In this paper, we studied simultaneous radial distortion correction and motion estimation for planar motion. We proposed two polynomial solvers for estimating the homography and distortion parameter, and showed that they are sufficiently numerically robust and fast to be incorporated in a real-time VO pipeline. The proposed solvers were tested rigorously on both synthetic and real data, and were shown to be on par or superior to competing methods.

References

1. Bhayani, S., Kukulova, Z., Heikkila, J.: A sparse resultant based method for efficient minimal solvers. In: Computer Vision and Pattern Recognition (CVPR), June 2020
2. Brown, D.C.: Decentering distortion of lenses. *Photogram. Eng.* **32**, 444–462 (1966)
3. Chen, T., Liu, Y.H.: A robust approach for structure from planar motion by stereo image sequences. *Mach. Vis. Appl. (MVA)* **17**(3), 197–209 (2006)
4. Cox, D.A., Little, J., O’Shea, D.: *Using Algebraic Geometry*. Graduate Texts in Mathematics. Springer, New York (2005)
5. Fitzgibbon, A.W.: Simultaneous linear estimation of multiple view geometry and lens distortion. In: Conference on Computer Vision and Pattern Recognition (CVPR), December 2001
6. Hajjdiab, H., Laganière, R.: Vision-based multi-robot simultaneous localization and mapping. In: Canadian Conference on Computer and Robot Vision (CRV), London, ON, Canada, pp. 155–162, May 2004
7. Hartley, R.I., Zisserman, A.: *Multiple View Geometry in Computer Vision*, 2nd edn. Cambridge University Press, Cambridge (2004)

8. Kayumbi, G., Cavallaro, A.: Multiview trajectory mapping using homography with lens distortion correction. *EURASIP J. Image Video Process.* **2008**(1), 1–11 (2008). <https://doi.org/10.1155/2008/145715>
9. Kukulova, Z., Bujnak, M., Pajdla, T.: Automatic generator of minimal problem solvers. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) *ECCV 2008*. LNCS, vol. 5304, pp. 302–315. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-88690-7_23
10. Kukulova, Z., Bujnak, M., Pajdla, T.: Polynomial eigenvalue solutions to the 5-pt and 6-pt relative pose problems. In: *British Machine Vision Conference (BMVC)* (2008)
11. Kukulova, Z., Heller, J., Bujnak, M., Pajdla, T.: Radial distortion homography. In: *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 639–647, June 2015
12. Larsson, V., Åström, K.: Uncovering symmetries in polynomial systems. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds.) *ECCV 2016*. LNCS, vol. 9907, pp. 252–267. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46487-9_16
13. Larsson, V., Åström, K., Oskarsson, M.: Efficient solvers for minimal problems by syzygy-based reduction. In: *Computer Vision and Pattern Recognition (CVPR)*, pp. 2383–2392, July 2017
14. Larsson, V., Åström, K., Oskarsson, M.: Polynomial solvers for saturated ideals. In: *International Conference on Computer Vision (ICCV)*, pp. 2307–2316, October 2017
15. Larsson, V., Kukulova, Z., Zheng, Y.: Camera pose estimation with unknown principal point. In: *Computer Vision and Pattern Recognition (CVPR)*, pp. 2984–2992 (2018)
16. Larsson, V., Oskarsson, M., Åström, K., Wallis, A., Kukulova, Z., Pajdla, T.: Beyond gröbner bases: Basis selection for minimal solvers. In: *Computer Vision and Pattern Recognition (CVPR)*, pp. 3945–3954 (2018)
17. Liang, B., Pears, N.: Visual navigation using planar homographies. In: *International Conference on Robotics and Automation (ICRA)*, Washington, DC, USA, pp. 205–210, May 2002
18. Meuel, H., Ferenz, S., Munderloh, M., Ackermann, H., Ostermann, J.: In-loop radial distortion compensation for long-term mosaicking of aerial videos. In: *Proceedings of the 23rd IEEE International Conference on Image Processing (ICIP)*, pp. 2961–2965, September 2016
19. Meuel, H., Munderloh, M., Reso, M., Ostermann, J.: Mesh-based piecewise planar motion compensation and optical flow clustering for ROI coding. *APSIPA Trans. Sign. Inf. Process.* **4**, e13 (2015)
20. Ortín, D., Montiel, J.M.M.: Indoor robot motion based on monocular images. *Robotica* **19**(3), 331–342 (2001)
21. Pritts, J., Kukulova, Z., Larsson, V., Chum, O.: Radially-distorted conjugate translations. In: *Conference on Computer Vision and Pattern Recognition (CVPR)* (2018)
22. Pritts, J., Kukulova, Z., Larsson, V., Chum, O.: Rectification from radially-distorted scales. In: Jawahar, C.V., Li, H., Mori, G., Schindler, K. (eds.) *ACCV 2018*. LNCS, vol. 11365, pp. 36–52. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-20873-8_3
23. Valtonen Örnghag, M., Wadenbäck, M.: Enforcing the general planar motion model: bundle adjustment for planar scenes. In: De Marsico, M., Sanniti di Baja, G., Fred, A. (eds.) *ICPRAM 2019*. LNCS, vol. 11996, pp. 119–135. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-40014-9_6

24. Valtonen Örn hag, M.: Fast non-minimal solvers for planar motion compatible homographies. In: International Conference on Pattern Recognition Applications and Methods (ICPRAM), Prague, Czech Republic, pp. 40–51, February 2019
25. Valtonen Örn hag, M.: Radially distorted planar motion compatible homographies. In: International Conference on Pattern Recognition Applications and Methods (ICPRAM), pp. 280–288 (2020)
26. Wadenbäck, M., Heyden, A.: Planar motion and hand-eye calibration using inter-image homographies from a planar scene. In: International Conference on Computer Vision Theory and Applications (VISAPP), pp. 164–168 (2013)
27. Wadenbäck, M., Heyden, A.: Ego-motion recovery and robust tilt estimation for planar motion using several homographies. In: International Conference on Computer Vision Theory and Applications (VISAPP), pp. 635–639, January 2014
28. Wadenbäck, M., Karlsson, M., Heyden, A., Robertsson, A., Johansson, R.: Visual odometry from two point correspondences and initial automatic tilt calibration. In: International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017), pp. 340–346 (2017)
29. Wadenbäck, M., Åström, K., Heyden, A.: Recovering planar motion from homographies obtained using a 2.5-point solver for a polynomial system. In: International Conference on Image Processing (ICIP), pp. 2966–2970 (2016)
30. Zienkiewicz, J., Davison, A.J.: Extrinsic autocalibration for dense planar visual odometry. *J. Field Robot.* (JFR) **32**(5), 803–825 (2015)