Radially Distorted Homographies, Revisited

Mårten Wadenbäck[a] Marcus Valtonen Örnhag[b] Johan Edstedt[a]
[a]Linköping University [b]Ericsson Research

Abstract

Homographies are among the most prevalent transformations occurring in geometric computer vision and projective geometry, and homography estimation is consequently a crucial step in a wide assortment of computer vision tasks. When working with real images, which are often afflicted with geometric distortions caused by the camera lens, it may be necessary to determine both the homography and the lens distortion—particularly the radial component, called radial distortion—simultaneously *to obtain anything resembling useful estimates. When considering a homography with radial distortion between two images, there are three conceptually distinct configurations for the radial distortion; (i) distortion in only one image, (ii) identical distortion in the two images, and (iii) independent distortion in the two images. While these cases have been addressed separately in the past, the present paper provides a novel and unified approach to solve all three cases. We demonstrate how the proposed approach can be used to construct new fast, stable, and accurate minimal solvers for radially distorted homographies. In all three cases, our proposed solvers are faster than the existing state-of-the-art solvers while maintaining similar accuracy. The solvers are tested on well-established benchmarks including images taken with fisheye cameras. A reference implementation of the proposed solvers is made available as part of HomLib*[1].

1. Introduction

Homographies constitute an important class of transformations that are often used in algorithms for solving computer vision tasks such as precalibration [23, 52, 66], autocalibration [56], stereo and multi-camera calibration [20, 27, 59, 67], colour calibration [15], metric rectification [13, 37], stereo rectification [22, 25, 30, 39], ego-motion estimation [36, 60–62, 64, 68], panoramic stitching and mosaicing [8, 9, 19, 42, 53, 65], scene understanding [3, 18, 44], visual servoing [7, 26], augmented reality [46, 50, 51], and more.

[1] https://github.com/marcusvaltonen/HomLib

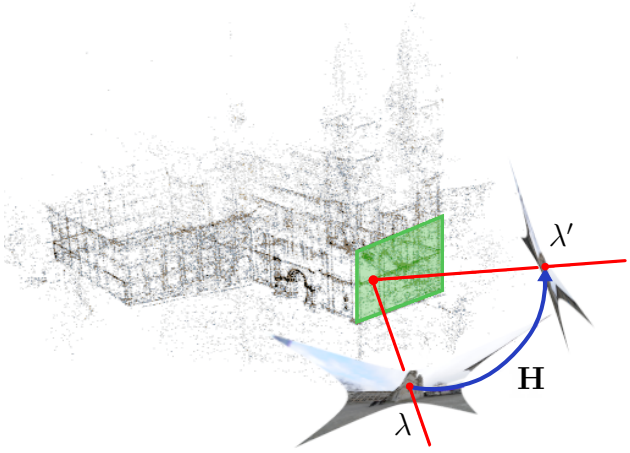

Figure 1. We propose three novel homography solvers simultaneously estimating a homography between two views and radial distortion coefficients. It is well-known that a homography $\mathbf{H}$ maps image point correspondences which lie on a planar surface (*cf.* the green plane in the scene); however, physical cameras may deviate significantly from the pinhole camera model due to *e.g.* radial distortion. In this figure of the *Grossmünster church* the input images are significantly distorted due to the camera being equipped with a fisheye lens. The reference 3D reconstruction was obtained using RadialSfM [33].

In most of these situations, the invocation of a homography is justified by the underlying geometry being considered, as a homography describes central projection of one projective plane onto another [24]. For example, consider a perspective pinhole camera viewing a plane in the scene. In this case, the mapping of homogeneous coordinates from the scene plane to the image plane will be a homography, *cf.* [66]. Alternatively, considering two such views of a plane in the scene, the mapping between the two views will, again, be a homography, for the image region that contains the plane [24].

Algebraically, a homography is represented by a (unique up to scale) non-singular matrix $\mathbf{H} \in \mathbb{R}^{3\times 3}$ that maps homogeneous coordinates $\mathbf{x}$ in one plane to $\mathbf{x}'$ in the other plane according to $\mathbf{x}' \sim \mathbf{H}\mathbf{x}$, where '$\sim$' denotes equality up to scale [24, 48]. To fixate the eight degrees of free-

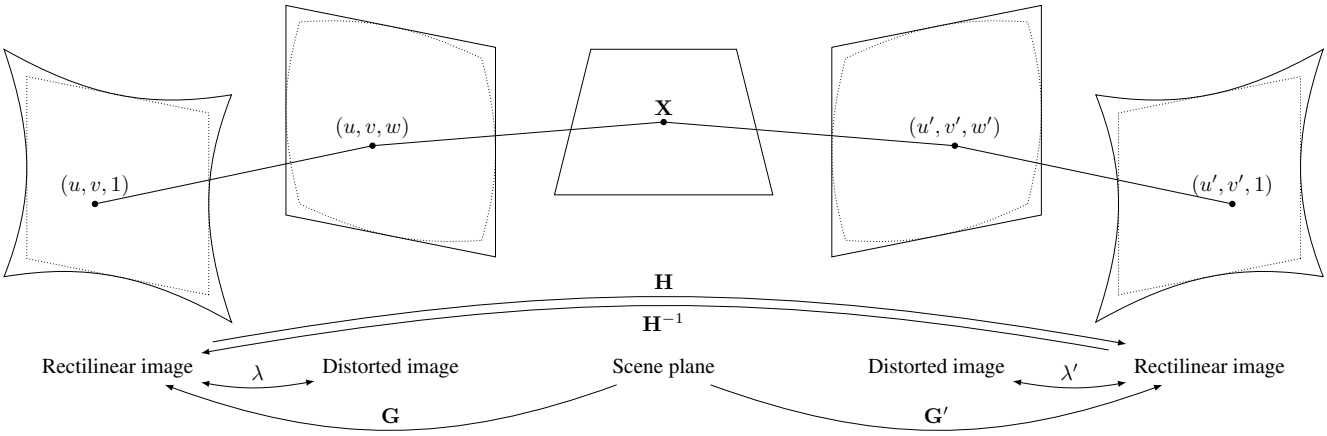

Figure 2. To each of the rectilinear image planes, there is a homography ($\mathbf{G}$ and $\mathbf{G}'$, respectively) from the scene plane (which itself is inherently rectilinear). There is also a homography between the two rectified image planes. For the distortion, three distinct cases are therefore of interest here: (i) the one-sided case $\lambda' = 0$, (ii) the two-sided equal case $\lambda' = \lambda$, and (iii) the two-sided independent case (where λ and λ' are independent). The one-sided case is equivalent to just viewing the image plane with one camera.

dom of $\mathbf{H}$, it is sufficient to specify four point correspondences, *i.e*. four points and their respective images under $\mathbf{H}$, as long as no three points (or their images under $\mathbf{H}$) are collinear. If this non-collinearity condition is met, the points are said to be in *general position*. Given at least four such correspondences, one can use the well-known *Direct Linear Transformation* (DLT) to solve for $\mathbf{H}$ linearly [24]. Since the pinhole perspective camera does not model lenses, it is typically used together with a separate distortion model such as the Brown–Conrady model (see *e.g*. [54]) or the one-parameter division model [2, 17, 35] when applied to real images. For this reason, homography estimation methods based on DLT but which also consider radial distortion have been proposed in the literature [17, 29, 41].

Methods for automatically establishing correspondences generally cannot be guaranteed to only produce correct matches, so in practice the set of correspondences will contain outliers in the form of spurious correlations. Such outliers have the potential to completely invalidate any subsequent model estimation, and it is important to reject them before proceeding. This is often done using RANSAC [16] or one of its more refined incarnations [11, 12, 43, 47, 55], where a consensus set of inliers is found by repeatedly fitting the model to small random samples and then checking which of the other data also support this model. The fewer data that are used for fitting the model in RANSAC, the higher is the chance of drawing an uncontaminated sample that only contains inliers, and thus hopefully in the process gains support from most of the other inliers. For this reason, considerable effort has been directed to create so called *minimal solvers* for different geometric estimation problems, *i.e*. algorithms that use as few data as possible while ensuring a finite number of solutions. Minimal solvers can provide both likely inlier sets and reasonable initialisations for further refinement such as bundle adjustment [57].

Our contributions. Inspired primarily by earlier work by Kukelova et al. [29] and more recent work by Nakano [41], we revisit the important problem of estimating radially distorted homographies, and demonstrate that by replacing the DLT with a closed-form expression for $\mathbf{H}$ [21, 48, 49, 63], we obtain numerically stable minimal solvers that are faster than the current state-of-the-art solvers for all possible configurations, *i.e*. (i) the *one-sided case*, (ii) the *two-sided equal case*, and (iii) the *two-sided independent case*. An illustration of the problem geometry for the three cases is shown in Fig. 2.

Organisation of the paper. The rest of the paper is organised as follows. In Sec. 2 we discuss relevant related work, including current state-of-the-art methods for radially distorted homographies. We formulate the problem mathematically and recount the classical closed-form formula for homographies in Sec. 3. Our proposed solvers are derived in Sec. 4. The solvers are evaluated in Sec. 5, and Sec. 6 concludes the paper.

2. Related Work

As mentioned in the introduction, homographies are a crucial component of many camera calibration algorithms. For example, Sturm and Maybank [52] use homographies from scene planes to provide constraints on the intrinsic camera parameters. Similarly, Zhang [66] computes homographies from a planar calibration target to the image plane using DLT, but in contrast to [52] also incorporates radial distor-

tion in a final non-linear refinement step, which forms the basis for the widely used calibration method in OpenCV.

The first work to include radial distortion already at the homography computation step is due to Fitzgibbon [17], who extended the DLT equations into a quadratic eigenvalue problem to capture lens distortion modelled by the one-parameter division model [17, 35].

Subsequently, solvers that make use of additional assumptions have been proposed for some notable special cases, *e.g.* for panoramic stitching [9, 10, 28] where images are assumed to be captured with coinciding camera centres. Instead of eliminating the translation, another option is to assume that the rotation is known, which can make sense in drone applications where an IMU can provide sufficiently accurate rotation estimates [62]. The present paper, however, does not make any such simplifying assumptions for the scene or pose parameters.

We consider three different cases for radially distorted homographies, which we will refer to as the *one-sided case*, the *two-sided equal case*, and the *two-sided independent case* (see Fig. 2). Each of these cases has been considered separately in previous work, but we take a novel unified approach for all three cases. Starting chronologically, Fitzgibbon [17] solved the *two-sided equal case* by deriving a quadratic eigenvalue problem with up to 18 real solutions. Kukelova et al. [29] proposed two solvers for the *two-sided independent case*, *i.e.* for radially distorted homographies with independent distortion coefficients λ and λ'. Their minimal 5 point solver uses a Gröbner basis approach, and requires performing Gauss–Jordan elimination on a 16×21 template matrix. Their non-minimal approach uses 6 points, and is computationally significantly cheaper, but while less sensitive to noise is more sensitive to outliers due to the inclusion of the additional point. Recently, Nakano [41] proposed a DLT-based solution to the *one-sided case* $\lambda' = 0$, finding the parameters of the transformation from the nullspace to a 10×12 design matrix. For an overview comparison of some of the key properties (including timings) of the existing solvers and our proposed solvers, see Tab. 1.

3. Problem formulation

Let us start by considering the problem without radial distortion. Given a number of point correspondences, $\mathbf{x}'_j \leftrightarrow \mathbf{x}_j$ for $j = 1, \ldots, n$, expressed using homogeneous coordinates $\mathbf{x}_j = (u_j, v_j, w_j)$ and $\mathbf{x}'_j = (u'_j, v'_j, w'_j)$, we want to find a non-singular $\mathbf{H} \in \mathbb{R}^{3\times 3}$ such that

$$\alpha_j \mathbf{x}'_j = \mathbf{H}\mathbf{x}_j, \quad \alpha_j \in \mathbb{R}, \quad j = 1, \ldots, n. \tag{1}$$

In the computer vision literature, this problem is traditionally solved using the *Direct Linear Transformation* (DLT) [24, 54], and this route was used as a foundation for the solvers with radial distortion introduced in [17], [29] and [41]. We will instead use a much older approach, which we shall discuss next, as a springboard to constructing our proposed solvers.

3.1. The classical closed-form solution

While largely overlooked by the computer vision community, the classical closed-form solution for a homography has a long history and is well established in the literature [48, 49, 63]. Its core idea is the fact that it is easy to find the mapping to (or from) the four intermediary points $\{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\}$. More specifically, it works as follows. Let $n = 4$ and assume that the points are in general position. This means that we can find non-zero scalars $\gamma_1, \gamma_2, \gamma_3$ such that

$$\mathbf{x}_4 = \gamma_1 \mathbf{x}_1 + \gamma_2 \mathbf{x}_2 + \gamma_3 \mathbf{x}_3, \tag{2}$$

and likewise, non-zero scalars $\gamma'_1, \gamma'_2, \gamma'_3$ such that

$$\mathbf{x}'_4 = \gamma'_1 \mathbf{x}'_1 + \gamma'_2 \mathbf{x}'_2 + \gamma'_3 \mathbf{x}'_3. \tag{3}$$

Indeed, if one of these scalars were to be zero, that would contravene our assumption that the points are in general position, *i.e.*, three of the points would need to be collinear on at least one side of the transformation.

Now, the homography $\mathbf{U} = \begin{pmatrix}\gamma_1\mathbf{x}_1 & \gamma_2\mathbf{x}_2 & \gamma_3\mathbf{x}_3\end{pmatrix}$ maps the intermediary points to $\mathbf{x}_1, \ldots, \mathbf{x}_4$, and in the same way, $\mathbf{U}' = \begin{pmatrix}\gamma'_1\mathbf{x}'_1 & \gamma'_2\mathbf{x}'_2 & \gamma'_3\mathbf{x}'_3\end{pmatrix}$ maps the intermediary points to $\mathbf{x}'_1, \ldots, \mathbf{x}'_4$. Hence, the sought homography is obtained by reversing the first mapping and applying the mappings in succession, yielding the composite mapping $\mathbf{H} = \mathbf{U}'\mathbf{U}^{-1}$.

We may write this procedure more compactly by denoting $\mathbf{\Gamma} = (\gamma_1, \gamma_2, \gamma_3)$ and $\mathbf{\Gamma}' = (\gamma'_1, \gamma'_2, \gamma'_3)$, as well as

$$\mathbf{\Xi} = \begin{pmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3\end{pmatrix} \quad \text{and} \quad \mathbf{\Xi}' = \begin{pmatrix}\mathbf{x}'_1 & \mathbf{x}'_2 & \mathbf{x}'_3\end{pmatrix}. \tag{4}$$

With this notation, we now have

$$\mathbf{\Gamma} = \mathbf{\Xi}^{-1}\mathbf{x}_4 \quad \text{and} \quad \mathbf{\Gamma}' = (\mathbf{\Xi}')^{-1}\mathbf{x}'_4. \tag{5}$$

Then, finally, $\mathbf{H}$ can be computed in closed-form as

$$\mathbf{H} = \mathbf{\Xi}'(\operatorname{diag}\mathbf{\Gamma}')(\operatorname{diag}\mathbf{\Gamma})^{-1}\mathbf{\Xi}^{-1}. \tag{6}$$

There are some small and obvious improvements that we shall wish to apply to the basic method described above. First of all, since

$$\mathbf{M}\operatorname{adj}\mathbf{M} = \det(\mathbf{M})\mathbf{I} \tag{7}$$

for any square matrix $\mathbf{M}$ [1], the non-diagonal inverses in (5) and (6) can be replaced with adjugates, as the scalar factor is irrelevant when working with homogeneous coordinates. Recalling the explicit formula for 3×3 adjugates (see *e.g.* Ch. 3 in [1]), we may write

$$\operatorname{adj}\mathbf{\Xi} = \begin{pmatrix} v_2w_3 - v_3w_2 & u_3w_2 - u_2w_3 & u_2v_3 - u_3v_2 \\ v_3w_1 - v_1w_3 & u_1w_3 - u_3w_1 & u_3v_1 - u_1v_3 \\ v_1w_2 - v_2w_1 & u_2w_1 - u_1w_2 & u_1v_2 - u_2v_1 \end{pmatrix}, \tag{8}$$

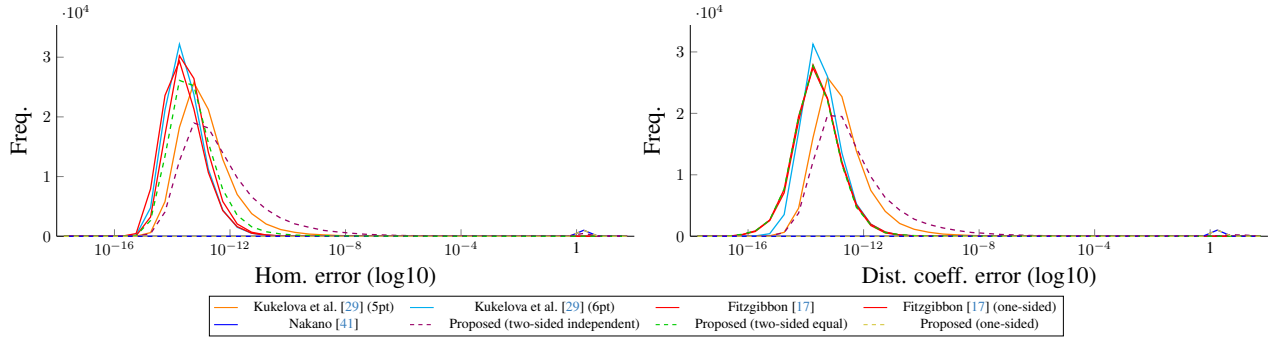

Figure 3. *Numerical stability*. Homography errors (left) and distortion coefficient error (right) for 10 000 noise-free randomly generated problem instances.

and similarly for $\mathrm{adj}(\boldsymbol{\Xi}')$. These are now plugged into (5) and (6) to give

$$\boldsymbol{\Gamma} = \mathrm{adj}(\boldsymbol{\Xi})\mathbf{x}_4, \quad \boldsymbol{\Gamma}' = \mathrm{adj}(\boldsymbol{\Xi}')\mathbf{x}'_4, \tag{9}$$

and

$$\mathbf{H} = \boldsymbol{\Xi}'(\mathrm{diag}\,\boldsymbol{\Gamma}')(\mathrm{diag}\,\boldsymbol{\Gamma})^{-1}\,\mathrm{adj}(\boldsymbol{\Xi}). \tag{10}$$

Secondly, if all points are finite and normalised to have their last entry equal to one, the resulting expression (10) becomes a little simpler and requires fewer arithmetic operations to evaluate [21].

3.2. Including radial distortion

Assuming the standard one-parameter division model for radial distortion [17, 35], we have

$$\begin{aligned} w_j &= 1 + \lambda(u_j^2 + v_j^2), \quad \text{and} \\ w'_j &= 1 + \lambda'\big((u'_j)^2 + (v'_j)^2\big). \end{aligned} \tag{11}$$

Three distinct cases are of interest here (*cf*. Fig. 2): (i) the one-sided case $\lambda' = 0$, (ii) the two-sided equal case $\lambda' = \lambda$, and (iii) the two-sided independent case (where λ and λ' are independent).

4. Deriving the solvers

For the sake of generality, let us initially assume that λ and λ' are independent and potentially both non-zero. Then, for a fifth point correspondence $\mathbf{x}'_5 \leftrightarrow \mathbf{x}_5$, we can use (10) to write

$$\mathrm{adj}(\mathrm{diag}\,\boldsymbol{\Gamma}')\,\mathrm{adj}(\boldsymbol{\Xi}')\mathbf{x}'_5 \sim \mathrm{adj}(\mathrm{diag}\,\boldsymbol{\Gamma})\,\mathrm{adj}(\boldsymbol{\Xi})\mathbf{x}_5, \tag{12}$$

where the left hand side is a vector valued cubic polynomial $\mathbf{N}'(\lambda') \in \mathbb{R}^3[\lambda']$ and the right hand side will be a cubic polynomial $\mathbf{N}(\lambda) \in \mathbb{R}^3[\lambda]$ (the expressions for the coefficients will be identical up to toggling of the primes).

Using the explicit formula for the adjugate of a 3×3 matrix again, we can expand both sides as necessary and read off the coefficients of $\mathbf{N}(\lambda)$ and $\mathbf{N}'(\lambda')$.

4.1. The one-sided case $\lambda' = 0$

This is the case considered by Nakano [41], and it arises naturally in connection with camera calibration. Here, the goal is to find the homography from the scene plane (*e.g*. a planar calibration target) to the rectified scene plane, together with the distortion coefficient λ. An alternative, equivalent interpretation, is that of finding a homography between a distorted and a rectified image.

Setting $\lambda' = 0$ in (12) and taking the cross product yields

$$\mathbf{N}(\lambda) \sim \mathbf{N}'(0) \iff \mathbf{N}(\lambda) \times \mathbf{N}'(0) = \mathbf{0}, \tag{13}$$

which corresponds to three scalar cubic equations in λ, which will have up to three real-valued solutions. We can solve any one of them using standard techniques, *e.g*. Cardano's formula [58], the trigonometric method [58], the companion matrix method [4] or Sturm sequences [58]. For its speed, and to avoid complex numbers, we use the trigonometric method.

4.2. The two-sided equal case

This is the case originally considered by Fitzgibbon [17], and arises naturally in situations with two images taken by the same camera (and without changing the optical configuration), for example when taking shots for a panorama. Setting $\lambda' = \lambda$ in (12) yields

$$\mathbf{N}(\lambda) \sim \mathbf{N}'(\lambda) \iff \mathbf{N}(\lambda) \times \mathbf{N}'(\lambda) = \mathbf{0}, \tag{14}$$

which corresponds to three scalar sextic equations in λ. We can again solve any one of them using a standard technique of our choice, and it will result in up to six real-valued solutions. To avoid computations involving complex numbers, we use Sturm sequences [58].

4.3. The two-sided independent case

This is the case considered by Kukelova et al. [29], and arises naturally *e.g*. when stitching images coming from different cameras (or if the optical configuration has changed,

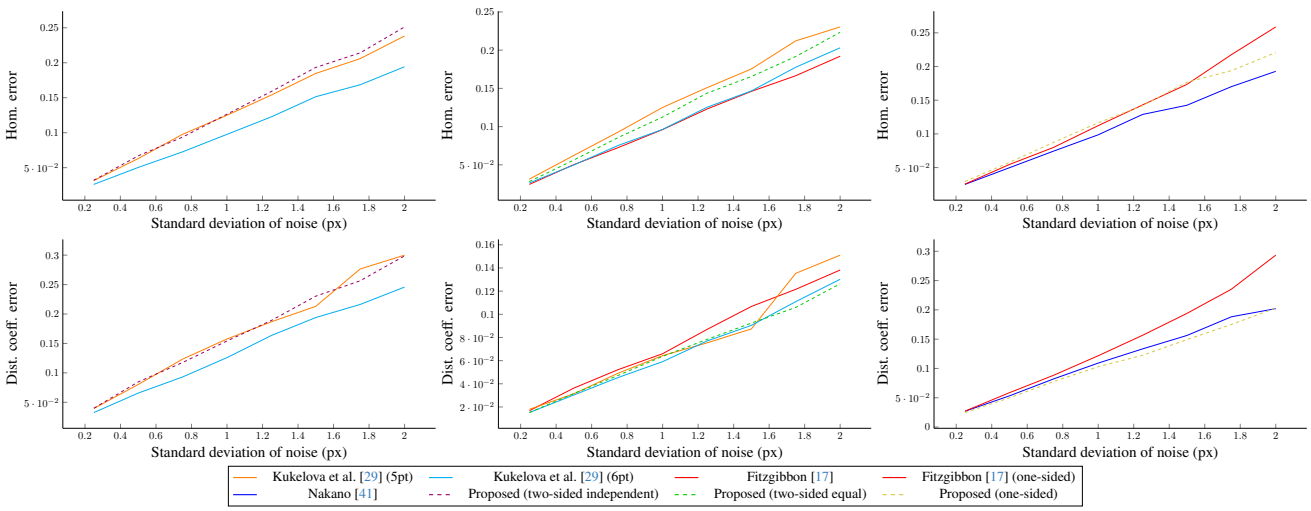

Figure 4. *Sensitivity to noise*. Homography errors (top) and distortion coefficient errors (bottom) for different noise levels. Each point reports the median error of 10 000 randomly generated problem instances (per noise level). The left column shows the noise sensitivity for the two-sided independent solvers, the middle column shows the noise sensitivity for the two-sided equal solvers, and the right column shows the noise sensitivity for the one-sided solvers.

e.g. due to zoom or focus adjustments). From (12) we get

$$\mathbf{N}(\lambda) \sim \mathbf{N}'(\lambda') \iff \mathbf{N}(\lambda) \times \mathbf{N}'(\lambda') = \mathbf{0}, \qquad (15)$$

which corresponds to three scalar polynomial equations of degree six in λ and λ'.

A polynomial system of equations of this type can be solved *e.g.* using resultant-based methods or Gröbner basis methods. We use the automatic Gröbner basis generator proposed in [31] to create a numerical solver. The system has in general nine complex-valued solutions. Attentive readers note that the solver by Kukelova et al. [29] only had five potential solutions—the four extra solutions obtained from our solver are introduced due to the adjugate matrix computations and can therefore be discarded by checking the determinant of the corresponding matrix. By using the basis heuristic proposed in [32], an elimination template of size 9×18 was found, which combined with Sturm sequence root finding generates a numerically stable and computationally efficient solver.

5. Experiments

We compare our solvers against the existing state-of-the-art solvers, *i.e.* Nakano [41] for the one-sided case, Fitzgibbon [17] for the two-sided equal case, and Kukelova et al. [29] for the two-sided independent case. All solvers are implemented in C++ using Eigen, and timings are shown in Tab. 1. The results show the median value of 100 000 random problem instances, executed on a standard laptop equipped with an 11th Gen Intel® Core™ i5-1145G7 @2.6 GHz CPU. While this reporting of the runtime gives some general sense of the speed of the solver, additional aspects, *e.g.* the sample size and the typical number of solutions, will influence the effective runtime.

Table 1. Properties of the evaluated state-of-the-art solvers including the median runtime for 100 000 random problem instances.

Solver	Pts.	Sols.	Min.	Time
Two-sided with independent λ and λ'				
Kukelova et al. [29]	5	5	✓	34.8 µs
Kukelova et al. [29]	6	2		12.3 µs
Proposed	5	5	✓	**8.1 µs**
Two-sided with $\lambda' = \lambda$				
Fitzgibbon [17]	5	18		64.0 µs
Kukelova et al. [29]	5*	5		25.6 µs
Kukelova et al. [29]	6*	2		11.2 µs
Proposed	4.5	6	✓	**3.3 µs**
One-sided, *i.e.* $\lambda' = 0$				
Fitzgibbon [17]	5**	6		17.6 µs
Nakano [41]	4.5	3	✓	4.5 µs
Proposed	4.5	3	✓	**0.6 µs**

*: The solver is terminated early if distortion coefficients are not approximately equal, and the geometric mean is returned in case they are.
**: Only the two-sided case was originally treated by Fitzgibbon [17]; however, Nakano [41] showed how to generalise the approach to the one-sided case.

5.1. Numerical stability

To evaluate the numerical stability of the solvers, we have generated 10 000 scenes with 3D points distributed randomly on an unknown plane. A pair of cameras with a depth of 0.1–10 to the scene plane, focal length = 1000,

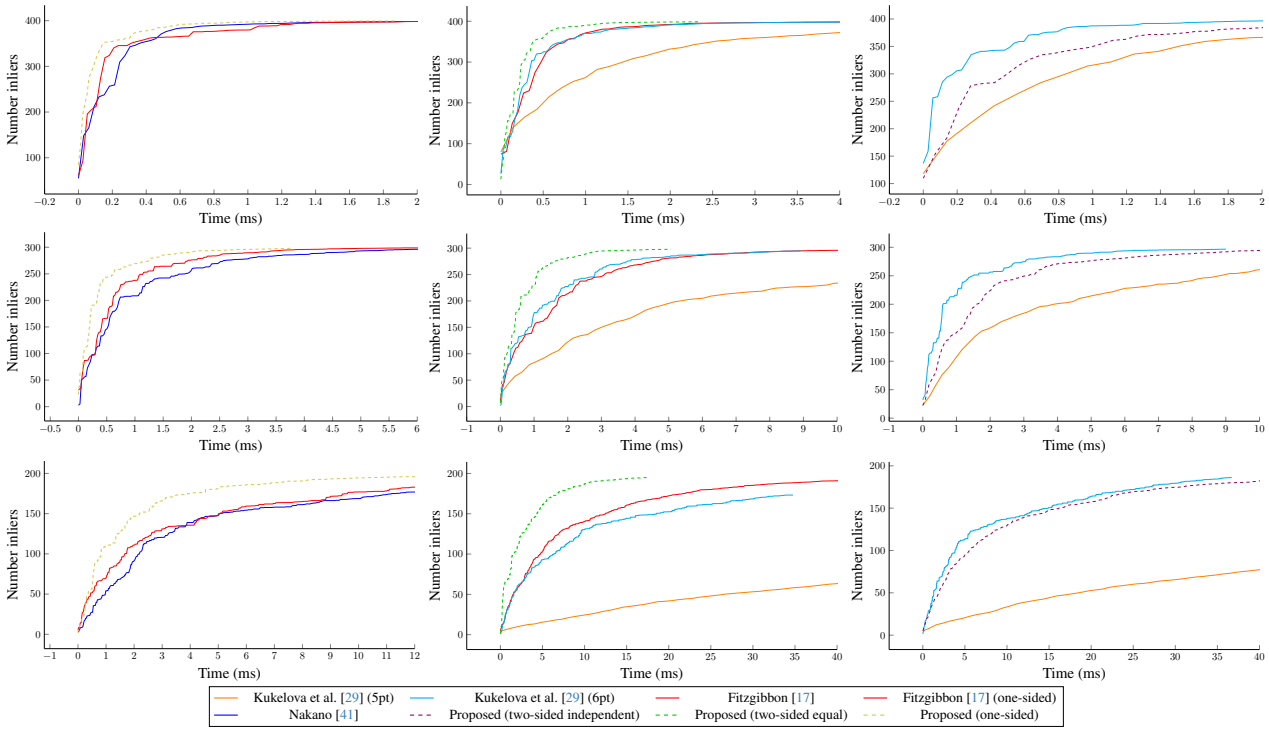

Figure 5. *RANSAC integration*. Here we plot the cumulative number of inliers versus time for all the evaluated solver, for 20 % outliers (top row), 40 % outliers (mid row), and 60 % outliers (bottom row). The left column shows the results for the one-sided solvers, the middle column shows the results for the two-sided equal solvers, and the right column shows the results for the two-sided independent solvers.

and distortion coefficients in the range $[-0.2, -0.01]$ were generated such that the field-of-view for both cameras is 70 degrees and the scene points are in front of the cameras. This captures a low to medium level of distortion, for which the one-parameter division model is known to be a good approximation for typical physical cameras.

As can be seen in Fig. 3, all the considered methods are numerically stable. For the solvers with two distortion coefficients, the algebraic error is computed as the geometric mean of the individual ones, *i.e.* $k_{\text{err}} = \sqrt{k_{1,\text{err}} k_{2,\text{err}}}$.

5.2. Sensitivity to noise

In addition to the numerical stability, we want to investigate the sensitivity of the proposed solvers with respect to Gaussian noise affecting the data. Gaussian noise of varying standard deviation is therefore applied to the distorted (normalised) point correspondences and the corresponding median errors over 10 000 random problem instances are reported, see Fig. 4.

Overall, the evaluated solvers all have comparable noise sensitivity to the other solvers for the same case. The strongest deviations are seen for Fitzgibbon [17] which is somewhat more sensitive with respect to the distortion coefficient in the one-sided case, and for the 6 point version of Kukelova et al. [29] in the two-sided independent case, which is less sensitive (this is somewhat expected, as this method uses an additional point, *cf.* [45]). In situations with both high noise and a high inlier ratio, this could be advantageous, but conversely, if the noise and inlier ratio are low, the inclusion of the extra point may instead be an encumbrance.

5.3. Integration in a robust framework

We integrate the polynomial solvers under evaluation into a RANSAC framework to get a better understanding of the performance. While all solvers are fast, RANSAC also requires evaluating a hypothesis for each model estimated, which affects the total time. We study the cumulative number of of inliers found versus execution time in Fig. 5, where each measurement shows the average over 1000 random problem instances. Settings in previous sections apply, and the standard deviation of the noise is set to 0.5 px, which is realistic for traditional descriptors such as SIFT [40], and an inlier threshold of 5 px is used. We find that our proposed solvers for the one-sided and two-sided equal cases consistently find more inliers for different outlier ratios. For the two-sided independent case, the 6 point solver by Kukelova et al. [29] has a clear advantage due to it only having two solutions to evaluate, even though the margin decreases for higher outlier ratios.

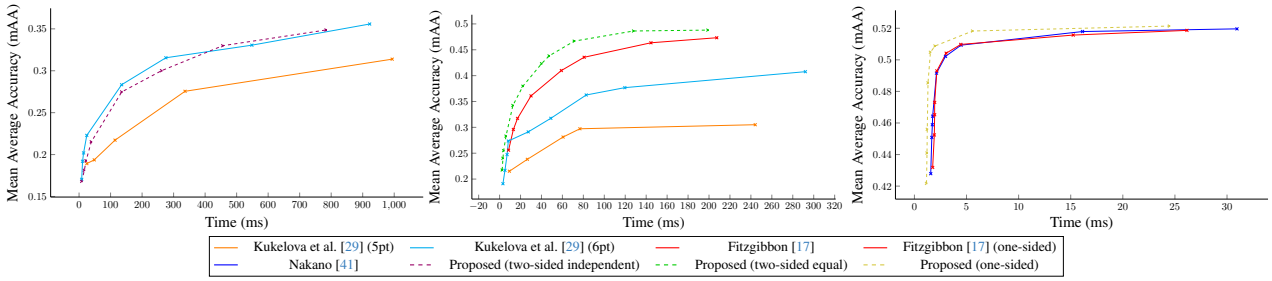

Figure 6. *Results on HPatches*. The left column shows the results for one-sided solvers, the middle column shows the results for the two-sided equal solvers, and the right column shows the results for the two-sided independent solvers.

5.4. Evaluation on real data

Next, we evaluate the proposed solvers on real images by integrating them in a LOMSAC-framework [34] applying a local optimisation step to the so-far-the-best model and non-linear refinement.

Table 2. *Results on HPatches*. mAA (AUC score) when allowing the solvers to run up to 1000 ms (two-sided independent), 250 ms (two-sided equal), 30 ms (one-sided). Best results are marked in bold.

Solver	mAA (AUC)		
	@1 px	@3 px	@5 px
Two-sided with independent λ and λ'			
Kukelova et al. [29] (5pt)	0.0983	0.3343	0.5079
Kukelova et al. [29] (6pt)	**0.1704**	0.3693	0.5276
Proposed (5pt)	0.1261	**0.3820**	**0.5379**
Two-sided with $\lambda' = \lambda$			
Fitzgibbon [17] (5pt)	0.3092	0.5133	0.6310
Kukelova et al. [29] (5pt)*	0.1127	0.3756	0.5379
Kukelova et al. [29] (6pt)*	0.2186	0.4415	0.5867
Proposed (4.5pt)	**0.3227**	**0.5192**	**0.6329**
One-sided, *i.e.* $\lambda' = 0$			
Fitzgibbon [17] (5pt)**	0.2542	0.5996	0.7285
Nakano [41] (4.5pt)	0.2537	0.5992	0.7269
Proposed (4.5pt)	**0.2566**	**0.6019**	**0.7298**

Evaluation on HPatches. For homography estimation, it is common to evaluate on the HPatches dataset [5], is a Homography benchmark consisting of 116 planar scenes and comes with an accurate reference in the form of manually annotated correspondences that are often used as ground truth. Unfortunately, HPatches does not have noticeable radial distortion, and the dataset does not contain a reference estimate for this. As a sanity check, we nonetheless report the mAA (AUC scores) on HPatches for thresholds of 1 px, 3 px and 5 px in Tab. 2.

We follow the evaluation settings in LightGlue [38] and resize the shorter side of the image to 480 pixels, and use a total of 1024 keypoints with SuperPoint [14]. SuperPoint + LightGlue is used for establishing the correspondences.

Detailed scores for the maximum runtime is shown in Tab. 2. Here we allowed the solvers to run up to 1 s (two-sided independent), 250 ms (two-sided equal), and 30 ms (one-sided), respectively. mAA (AUC score) as a function of runtime is shown in Fig. 6. In general our proposed solvers yield higher mAA than previous solvers.

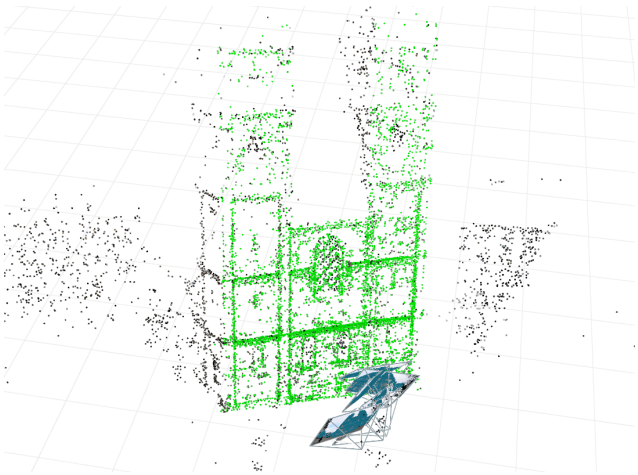

Figure 7. *Grossmünster Church*. 3D reconstruction containing the planar points (green) and the seven camera positions.

Evaluation on Grossmünster Church. Due to the lack of real-life homography benchmarks with radial distortion we use the Grossmünster Church dataset [33] which contains real images with fisheye distortion. In contrast to the HPatches dataset distortion is present, and is not perfectly approximated by the one-parameter division model. This dataset was originally used in RadialSfM [33], and we use it to generate a 3D reconstruction of the scene, see Fig. 1. In order to extract ground truth homographies, we use the acquired poses, and compute the relative pose $\mathbf{R} = \mathbf{R}_2\mathbf{R}_1^\top$ and relative translation $\mathbf{t} = \mathbf{t}_2 - \mathbf{R}\mathbf{t}_1$. As suggested in [6], we select only images with a single dominant plane—in our

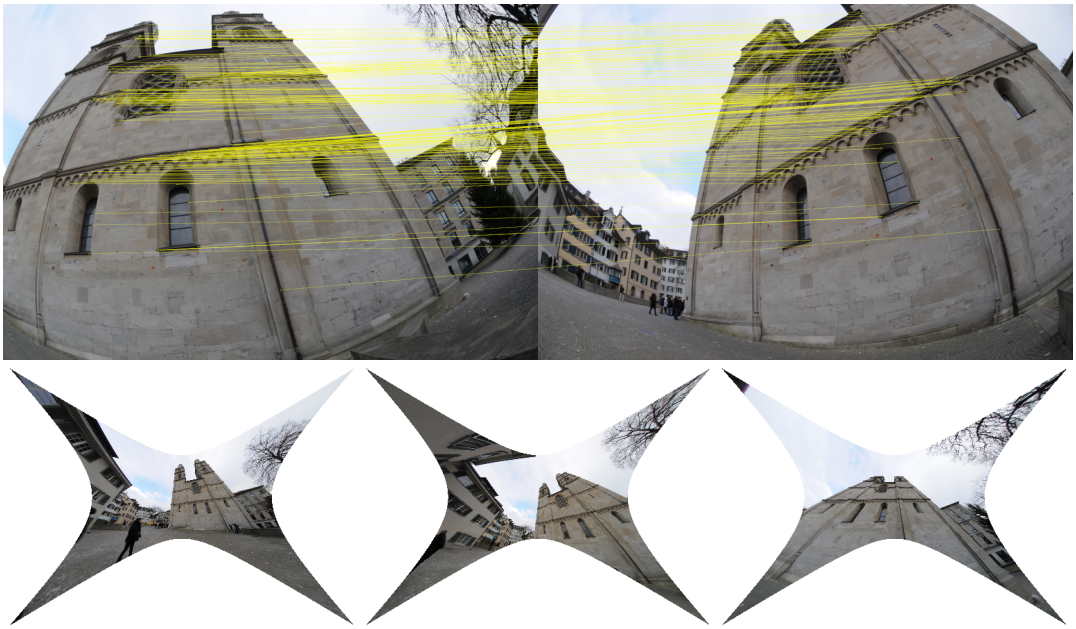

Figure 8. *Grossmünster Church*. (Top) Inliers obtained using the proposed two-sided equal solver. (Bottom) Examples of rectified images obtained using the proposed algorithm. Note that the edges of the building appear straight, indicating a successful correction.

case we chose the wall containing the two church towers—and fitted a plane directly on the 3D model. Among the images in the dataset, seven images were selected for which it was possible to extract six consecutive image pairs with more than 100 tentative image correspondences (including outliers), see Fig. 7. The ground truth homography was then computed as $\mathbf{H} = \mathbf{K}(\mathbf{R} - \mathbf{t}(\mathbf{n}')^\top/d)\mathbf{K}^{-1}$, where $\mathbf{K}$ is the shared intrinsic parameters across all views, $\mathbf{n}' = \mathbf{R}_1\mathbf{n}$ is the relative plane normal, and d the corresponding depth.

For the experiment, we extracted SIFT keypoints and applied Lowe's ratio test. The final correspondences for one image pair are shown in Fig. 8. Since radial distortion is present in both images, only the two-sided solvers are used in this comparison. The homography error is computed as $\epsilon = \|\mathbf{H} - \mathbf{H}_{\text{est}}\|_{\text{F}}$, where both $\mathbf{H}$ and $\mathbf{H}_{\text{est}}$ are normalised to have unit Frobenius norm and positive determinant.

All solvers are configured to run LOMSAC [34] with a minimum of 500 iterations and final non-linear least squares fitting is applied. The average homography error is almost identical for all solvers; however, the execution times differ, see Tab. 3.

Table 3. *Results on Grossmünster Church*. Mean error and mean execution time of the solvers in a LOMSAC framework. Best results are marked in bold (across both categories).

Solver	Error	Exec. time
Two-sided with independent λ and λ'		
Kukelova et al. [29] (5pt)	0.039	2657 ms
Kukelova et al. [29] (6pt)	0.040	330 ms
Proposed (5pt)	**0.038**	150 ms
Two-sided with $\lambda' = \lambda$		
Fitzgibbon [17] (5pt)	0.040	233 ms
Kukelova et al. [29] (5pt)*	0.040	2053 ms
Kukelova et al. [29] (6pt)*	0.039	321 ms
Proposed (4.5pt)	0.039	**84 ms**

6. Conclusion

In this paper, we have revisited the fundamental problem of simultaneously estimating a homography and radial distortion. Basing our analysis on a classical closed-form solution instead of the DLT, we have taken a unified approach to derive three new, fast, stable and accurate solvers for the possible cases of radially distorted homographies, *i.e.* (i) the one-sided case, (ii) the two-sided equal case, and (iii) the two-sided independent case. We have evaluated our proposed solvers against existing state-of-the-art solvers for each of the three cases, and our new solvers either outperform or compare favourably to the state-of-the-art in these evaluations.

Acknowledgements

This work was supported by the *Wallenberg Artificial Intelligence, Autonomous Systems and Software Program* (WASP), funded by the Knut and Alice Wallenberg Foundation, as well as by the strategic research environment ELLIIT, funded by the Swedish government.

References

[1] A. C. Aitken. *Determinants and Matrices*. Oliver and Boyd, Edinburgh, UK, 3 edition, 1944. 3

[2] Miguel Alemán-Flores, Luis Alvarez, Luis Gomez, and Daniel Santana-Cedrés. Automatic lens distortion correction using one-parameter division models. *Image Processing On Line*, 4:327–343, 2014. 2

[3] J. Arróspide, L. Salgado, M. Nieto, and R. Mohedano. Homography-based ground plane detection using a single on-board camera. *IET Intelligent Transport Systems*, 4(2): 149–160, 2010. 1

[4] Jared L. Aurentz, Thomas Mach, Raf Vandebril, and David S. Watkins. Fast and backward stable computation of roots of polynomials. *SIAM Journal on Matrix Analysis and Applications*, 36(3):942–973, 2015. 4

[5] Vassileios Balntas, Karel Lenc, and Andrea Vedaldi. HPatches: A benchmark and evaluation of handcrafted and learned local descriptors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3852–3861, Honolulu, HI, USA, 2017. IEEE. 7

[6] Daniel Barath, Dmytro Mishkin, Michal Polic, Wolfgang Förstner, and Jiri Matas. A large-scale homography benchmark. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 21360–21370, Vancouver, BC, Canada, 2023. IEEE. 7

[7] S. Benhimane and E. Malis. Homography-based 2D visual tracking and servoing. *The International Journal of Robotics Research*, 26(7):661–676, 2007. 1

[8] Matthew Brown and David G. Lowe. Automatic panoramic image stitching using invariant features. *International Journal of Computer Vision (IJCV)*, 72(1):59–73, 2007. 1

[9] Matthew Brown, Richard I. Hartley, and David Nistér. Minimal solutions for panoramic stitching. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, MN, USA, 2007. IEEE. 1, 3

[10] Martin Byröd, Matthew Brown, and Kalle Åström. Minimal solutions for panoramic stitching with radial distortion. In *Proceedings of the British Machine Vision Conference (BMVC)*, pages 41.1–41.11, London, UK, 2009. BMVA Press. 3

[11] Ondřej Chum and Jiří Matas. Matching with PROSAC – Progressive sample consensus. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 220–226, San Diego, CA, USA, 2005. IEEE. 2

[12] Ondřej Chum, Jiří Matas, and Josef Kittler. Locally optimized RANSAC. In *Proceedings of the 25th DAGM Symposium on Pattern Recognition*, pages 236–243, Magdeburg, Germany, 2003. Springer-Verlag. 2

[13] Antonio Criminisi, Ian D. Reid, and Andrew Zisserman. Single view metrology. *International Journal of Computer Vision (IJCV)*, 40(2):123–148, 2000. 1

[14] Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. SuperPoint: Self-supervised interest point detection and description. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 337–349, Salt Lake City, UT, USA, 2018. IEEE. 7

[15] Graham Finlayson, Han Gong, and Robert B. Fisher. Color homography: Theory and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 41(1): 20–33, 2019. 1

[16] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. 2

[17] Andrew W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 125–132, Kauai, HI, USA, 2001. IEEE. 2, 3, 4, 5, 6, 7, 8

[18] Alex Flint, David Murray, and Ian Reid. Manhattan scene understanding using monocular, stereo, and 3D features. In *Proceedings of the International Conference on Computer Vision (ICCV)*, pages 2228–2235, Barcelona, Spain, 2011. IEEE. 1

[19] Junhong Gao, Seon Joo Kim, and Michael S. Brown. Constructing image panoramas using dual-homography warping. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 49–56, Colorado Springs, CO, USA, 2011. IEEE. 1

[20] Banglei Guan, Yang Shang, and Qifeng Yu. Planar self-calibration for stereo cameras with radial distortion. *Applied Optics*, 56(33):9257–9267, 2017. 1

[21] Yang Guo. A minimal solution for plane motion and structure from two perspective views. *Pattern Recognition Letters*, 120:96–103, 2019. 2, 4

[22] Richard Hartley and Rajiv Gupta. Computing matched-epipolar projections. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 549–555, New York, NY, USA, 1993. IEEE. 1

[23] Richard Hartley and Sing Bing Kang. Parameter-free radial distortion correction with center of distortion estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 29(8):1309–1321, 2007. 1

[24] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2 edition, 2004. 1, 2, 3

[25] Richard I. Hartley. Theory and practice of projective rectification. *International Journal of Computer Vision (IJCV)*, 35 (2):151–173, 1999. 1

[26] G. Hu, W. MacKunis, N. Gans, W. E. Dixon, J. Chen, A. Behal, and D. Dawson. Homography-based visual servo control with imperfect camera calibration. *IEEE Transactions on Automatic Control*, 54(6):1318–1324, 2009. 1

[27] Zhaozheng Hu and Zheng Tan. Calibration of stereo cameras from two perpendicular planes. *Applied Optics*, 44(24): 5086–5090, 2005. 1

[28] Hailin Jin. A three-point minimal solution for panoramic stitching with lens distortion. In *Proceedings of the IEEE*

Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Anchorage, AK, USA, 2008. IEEE. 3

[29] Zuzana Kukelova, Jan Heller, Martin Bujnak, and Tomas Pajdla. Radial distortion homography. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 639–647, Boston, MA, USA, 2015. IEEE. 2, 3, 4, 5, 6, 7, 8

[30] S. Kumar, C. Micheloni, C. Piciarelli, and G. L. Foresti. Stereo rectification of uncalibrated and heterogeneous images. *Pattern Recognition Letters*, 31(11):1445–1452, 2010. 1

[31] Viktor Larsson, Kalle Åström, and Magnus Oskarsson. Efficient solvers for minimal problems by syzygy-based reduction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2383–2392, Honolulu, HI, USA, 2017. IEEE. 5

[32] Viktor Larsson, Magnus Oskarsson, Kalle Åström, Alge Wallis, Zuzana Kukelova, and Tomas Pajdla. Beyond Gröbner bases: Basis selection for minimal solvers. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3945–3954, Salt Lake City, UT, USA, 2018. IEEE. 5

[33] Viktor Larsson, Nicolas Zobernig, Kasim Taskin, and Marc Pollefeys. Calibration-free structure-from-motion with calibrated radial trifocal tensors. In *Proceedings of the 16th European Conference on Computer Vision (ECCV)*, pages 382–399, Glasgow, UK, 2020. Springer-Verlag. 1, 7

[34] Karel Lebeda, Jiri Matas, and Ondrej Chum. Fixing the locally optimized RANSAC. In *Proceedings of the British Machine Vision Conference (BMVC)*, pages 95.1–95.11, Surrey, UK, 2012. BMVA Press. 7, 8

[35] Reimar Lenz. Linsenfehlerkorrigierte eichung von halbleiterkameras mit standardobjektiven für hochgenaue 3D — messungen in echtzeit. In *Mustererkennung 1987*, pages 212–216, Braunschweig, Germany, 1987. Springer-Verlag. 2, 3, 4

[36] Bojian Liang and Nick Pears. Visual navigation using planar homographies. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 205–210, Washington, DC, USA, 2002. IEEE Robotics and Automation Society. 1

[37] David Liebowitz and Andrew Zisserman. Metric rectification for perspective images of planes. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 482–488, Santa Barbara, CA, USA, 1998. IEEE. 1

[38] Philipp Lindenberger, Paul-Edouard Sarlin, and Marc Pollefeys. LightGlue: Local feature matching at light speed. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 17581–17592, Paris, France, 2023. IEEE/CVF. 7

[39] Charles Loop and Zhengyou Zhang. Computing rectifying homographies for stereo vision. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 125–131, Fort Collins, CO, USA, 1999. IEEE. 1

[40] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision (IJCV)*, 60(2):91–110, 2004. 6

[41] Gaku Nakano. Inverse DLT method for one-sided radial distortion homography. In *Proceedings of the 27th International Conference on Pattern Recognition (ICPR)*, pages 448–462, Kolkata, India, 2024. Springer-Verlag. 2, 3, 4, 5, 6, 7

[42] Tudor Nicosevici, Nuno Gracias, Shahriar Negahdaripour, and Rafael Garcia. Efficient three-dimensional scene modeling and mosaicing. *Journal of Field Robotics*, 26(10):759–788, 2009. 1

[43] David Nistér. Preemptive RANSAC for live structure and motion estimation. In *Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV)*, pages 199–206, Nice, France, 2003. IEEE. 2

[44] A. G. Amitha Perera, Glen Brooksby, Anthony Hoogs, and Gianfranco Doretto. Moving object segmentation using scene understanding. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*, New York, NY, USA, 2006. IEEE. 1

[45] Trung T. Pham, Tat-Jun Chin, Jin Yu, and David Suter. The random cluster model for robust geometric fitting. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 36(8):1658–1671, 2014. 6

[46] Simon J. D. Prince, Ke Xu, and Adrian David Cheok. Augmented reality camera tracking with homographies. *IEEE Computer Graphics and Applications*, 22(6):39–45, 2002. 1

[47] Rahul Raguram, Ondřej Chum, Marc Pollefeys, Jiří Matas, and Jan-Michael Frahm. USAC: A universal framework for random sample consensus. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 35(8):2022–2038, 2012. 2

[48] Jürgen Richter-Gebert. *Perspectives on Projective Geometry*. Springer-Verlag, Berlin, Germany, 2011. 1, 2, 3

[49] A. Seidenberg. *Lectures In Projective Geometry*. Van Nostrand Reinhold Company, New York, NY, USA, 1962. 2, 3

[50] Yongduek Seo and Ki-Sang Hong. Weakly calibrated video-based augmented reality: Embedding and rendering through virtual camera. In *Proceedings IEEE and ACM International Symposium on Augmented Reality*, pages 129–136, Munich, Germany, 2000. IEEE. 1

[51] Gilles Simon, Andrew W. Fitzgibbon, and Andrew Zisserman. Markerless tracking using planar structures in the scene. In *Proceedings IEEE and ACM International Symposium on Augmented Reality*, pages 120–128, Munich, Germany, 2000. IEEE. 1

[52] Peter F. Sturm and Stephen J. Maybank. On plane-based camera calibration: A general algorithm, singularities, applications. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 432–437, Fort Collins, CO, USA, 1999. IEEE. 1, 2

[53] Richard Szeliski. Video mosaics for virtual environments. *IEEE Computer Graphics and Applications*, 16(2):22–30, 1996. 1

[54] Richard Szeliski. *Computer Vision*. Springar Nature, Cham, Switzerland, 2 edition, 2022. 2, 3

[55] P. H. S. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. *Computer Vision and Image Understanding*, 78(1):138–156, 2000. 2

[56] Bill Triggs. Autocalibration from planar scenes. In *Proceedings of the 5th European Conference on Computer Vision (ECCV)*, pages 89–105, Freiburg, Germany, 1998. Springer-Verlag. 1

[57] Bill Triggs, Philip F. McLauchlan, Richard I. Hartley, and Andrew W. Fitzgibbon. Bundle adjustment – a modern synthesis. In *Vision Algorithms: Theory and Practice*, pages 298–372, Corfu, Greece, 1999. Springer-Verlag. 2

[58] H. W. Turnbull. *Theory of Equations*. Oliver and Boyd, Edinburgh, UK, 2 edition, 1944. 4

[59] Toshio Ueshiba and Fumiaki Tomita. Plane-based calibration algorithm for multi-camera systems via factorization of homography matrices. In *Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV)*, pages 966–973, Nice, France, 2003. IEEE. 1

[60] Marcus Valtonen Örnhag. Radially distorted planar motion compatible homographies. In *Proceedings of the 9th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, pages 280–288, Valetta, Malta, 2020. SciTePress. 1

[61] Marcus Valtonen Örnhag and Mårten Wadenbäck. Enforcing the general planar motion model: Bundle adjustment for planar scenes. In *Revised Selected Papers from the 8th International Conference on Pattern Recognition Applications and Methods (ICPRAM)*, pages 119–135, Cham, Switzerland, 2020. Springer-Verlag.

[62] Marcus Valtonen Örnhag, Patrik Persson, Mårten Wadenbäck, Kalle Åström, and Anders Heyden. Efficient real-time radial distortion correction for UAVs. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, 2021. IEEE. 1, 3

[63] Oswald Veblen and John Wesley Young. *Projective Geometry (Vol. 1)*. Ginn and Company, Boston, MA, USA, 1910. 2, 3

[64] Mårten Wadenbäck, Kalle Åström, and Anders Heyden. Recovering planar motion from homographies obtained using a 2.5-point solver for a polynomial system. In *Proceedings of the 23rd IEEE International Conference on Image Processing (ICIP)*, pages 2966–2970, Phoenix, AZ, USA, 2016. IEEE. 1

[65] Julio Zaragoza, Tat-Jun Chin, Michael S. Brown, and David Suter. As-projective-as-possible image stitching with moving DLT. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2339–2346, Portland, OR, USA, 2013. IEEE. 1

[66] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 22(11):1330–1334, 2000. 1, 2

[67] Chen Zhu, Zihan Zhou, Zirang Xing, Yanbing Dong, Yi Ma, and Jingyi Yu. Robust plane-based calibration of multiple non-overlapping cameras. In *Proceedings of the Fourth International Conference on 3D Vision (3DV)*, pages 658–666, Stanford, CA, USA, 2016. IEEE. 1

[68] Jacek Zienkiewicz and Andrew Davison. Extrinsics autocalibration for dense planar visual odometry. *Journal of Field Robotics*, 32(5):803–825, 2015. 1