Computer Vision and Image Understanding 114 (2010) 220–233

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Efficient generic calibration method for general cameras with single centre of projection

Aubrey K. Dunne *, John Mallon, Paul F. Whelan

Vision Systems Group, Dublin City University, Dublin 9, Ireland

ARTICLE INFO

Article history:
Received 20 February 2008
Accepted 4 May 2009
Available online 15 July 2009

Keywords:
Catadioptric
Generic calibration
General model
Pose estimation

ABSTRACT

Generic camera calibration is a non-parametric calibration technique that is applicable to any type of vision sensor. However, the standard generic calibration method was developed such that both central and non-central cameras can be calibrated within the same framework. Consequently, existing parametric calibration techniques cannot be applied for the common case of cameras with a single centre of projection (e.g. pinhole, fisheye, hyperboloidal catadioptric). This paper proposes improvements to the standard generic calibration method for central cameras that reduce its complexity, and improve its accuracy and robustness. Improvements are achieved by taking advantage of the geometric constraints resulting from a single centre of projection in order to enable the application of established pinhole calibration techniques. Input data for the algorithm is acquired using active grids, the performance of which is characterised. A novel linear estimation stage is proposed that enables a well established pinhole calibration technique to be used to estimate the camera centre and initial grid poses. The proposed solution is shown to be more accurate than the linear estimation stage of the standard method. A linear alternative to the existing polynomial method for estimating the pose of additional grids used in the calibration is demonstrated and evaluated. Distortion correction experiments are conducted with real data for both an omnidirectional camera and a fisheye camera using the standard and proposed methods. Motion reconstruction experiments are also undertaken for the omnidirectional camera. Results show the accuracy and robustness of the proposed method to be improved over those of the standard method.

1. Introduction

There is currently a trend towards increased use of wide-angle dioptric and catadioptric cameras within the vision community due to the richer feature set and greater persistence of vision that these camera types provide. As a consequence of this trend, a number of models and calibration algorithms have recently been proposed for such cameras. The most basic models extend the pinhole camera model with one or two radial distortion terms [26,7,31,12]. These methods become less accurate for wide-angle and catadioptric lenses as the camera incorporates more distortion. Many of the common distortion models (polynomial, divisional, rational) can be augmented with an increasing number of parameters [13,30] to allow wider angle lenses to be calibrated. However, they are not suitable for fisheye or catadioptric lenses for which the field of view (FOV) exceeds 180°.

Several methods have been proposed that model wide-angle cameras as radially symmetric imagers [27,28], thus simplifying the unknown parameter set. In [27], distortion is modelled using a varying focal length instead of an image displacement approach, allowing cameras with FOV greater than 180° to be modelled. The complete class of single viewpoint catadioptric camera configurations was derived in [1], and this has been the basis for the development of parametric calibration models that are specific to a particular camera/lens configuration, most notably types of central catadioptric [16] and non-central catadioptric [17] cameras. The equivalence between catadioptric projections and mappings of the sphere was demonstrated in [8], resulting in a unifying model for catadioptric cameras. Nevertheless, only a few methods have been proposed that can model both dioptric (with FOV greater than 180°) and catadioptric cameras, i.e. a unifying model for all central cameras [27,2].

All the above calibration techniques assume a parametric camera model of some form, where the task is to estimate the (usually small) set of model parameters. In contrast, a non-parametric approach was proposed by Grossberg and Nayar [9]. This general camera model consists of a mapping in which each pixel is mapped to the direction of a half-ray in space, together with an anchor point. In principle, the ray direction for each pixel is completely independent of the ray directions of the surrounding pixels, thus allowing application to any type of central or non-central camera. The calibration technique described in [9] uses two images of a grid in different, known, positions. By determining the location

* Corresponding author. Fax: +353 1 7005508.
E-mail address: aubrey.dunne@eeng.dcu.ie (A.K. Dunne).

1077-3142/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.cviu.2009.05.005

seen by each pixel on each grid, the set of all camera ray directions can be determined. A generalisation to this calibration method, termed generic calibration, was proposed by Sturm and Ramalingam [24], wherein the world transformation between grid positions is not known a priori. Here, the calibration consists of determining the points seen by a pixel on each of three grids in unknown orientations. Effectively this becomes the estimation of the positions and orientations of each of the three grids, since knowledge of these allows the world ray-plane intersections to be determined.

The generic camera calibration process using planar calibration targets, as proposed in [24,25], can be summarised as follows:

(1) Take a minimum of three initial images of a calibration grid in different orientations. Images of a calibration grid in various additional orientations are required to completely cover the image.
(2) For each pixel, determine the location seen by that pixel on each grid.
(3) Linearly estimate the poses of the calibration grids, and the effective centre of projection if the camera is central, using this data and the known constraints.
(4) Refine the orientations of the initial grids and the ray directions in a bundle adjustment stage.
(5) Estimate the poses of the additional grids (in order to calibrate pixels that do not view any of the initial grids in step (1)) using geometric constraints followed by bundle adjustment.
(6) Store the ray directions (as Plücker matrices) in a look-up table.

This calibration method will hereafter be referred to as the standard generic method.

Different variants of the standard generic method for calibration using planar targets have been proposed for central, axial, and non-central cameras [24,21]. Nevertheless the generic calibration framework is the same for all camera types. That is, a collinearity constraint is applied to the intersection points of each camera ray with the grids involved in the linear estimation stage. The solution for the grid poses, and for the camera centre in the case of central cameras, is obtained by solving equations that are formed by enforcing this collinearity constraint. The central and non-central models vary only in the method for solving these equations. Standard central generic calibration is a specialisation of the general generic method, and so the link between the pinhole camera model and the geometric constraints of central cameras is not considered in the central generic method. Additionally, since the polynomial based pose estimation method of step (5) is the same for both central and non-central cameras, it does not account for known centrality. This paper presents a new generic calibration method for cameras with a single centre of projection, hereafter referred to as the proposed generic method, that achieves improved performance through the use of pinhole calibration techniques. It proposes alternatives to steps 2, 3 and 5 of the standard generic method that increase the accuracy of the calibration. The three key contributions in this paper are as follows.

Firstly, the issue of specifying accurate input data is addressed. In step (2), the location seen by each pixel on each calibration grid must be determined. When that location is not exactly at a grid feature point on a static calibration grid, it must be determined by homographic interpolation between the grid features, as is done in [24,25]. In theory, generic calibration can achieve pixel level calibration, and thus it seems appropriate to use pixel level data as input to the algorithm. Such data can be obtained by the use of spatio-temporally varying grids displayed on a flat screen monitor. We have termed these grids 'active grids', and while this method has been used before [9,22,5,4], no discussion has been proffered on their performance for calibration purposes. An explanation of active grids and an evaluation of their performance relative to standard localisation techniques is presented in Section 2. Secondly, a novel method for the linear estimation of the camera centre and the initial grid poses is proposed. Within standard generic calibration, the estimation of the camera centre and the initial grid poses in the central calibration variant is complicated, partly due to coupling of the variables to be extracted. In Section 3 active grids are shown to facilitate other, more intuitive and more accurate methods of determining the camera centre for central cameras. Thirdly, a linear pose estimation stage is proposed for use in generic calibration as an alternative to the polynomial pose estimation algorithm of Ramalingam et al. [21]. The linear pose estimation algorithm is discussed and evaluated in Section 4. Together, the above modifications serve to make the proposed generic method for central cameras both more robust and more accurate than the standard method. Simulations and experiments with real data are presented in Section 5 that demonstrate the improved performance. The effects of the modifications and the accuracy of the complete calibration are shown and discussed.

2. Active grids

Binary chessboard grids are typically used in camera calibration, since the corners of the chessboard grid squares can be easily extracted and accurately localised in images of the grids. This results in a dense set of grid to image correspondences. For the standard generic method, these correspondences must be used to determine the intersection points of camera rays with the grid (i.e. the location seen on the grid by each camera pixel). In most cases, the intersection points will not lie exactly on a grid corner. Therefore, homographic interpolation is employed in [24] to determine the intersection points based on the extracted image coordinates of the four closest grid corner points. However, this approach is unsuitable for high fidelity calibration, since any distortion present in the images of the calibration grids introduces a bias in the results. Fig. 1a shows a vector plot of the error residuals after homographic interpolation is applied to 500 random points on a 300 mm $\times$ 300 mm grid (simulated camera with radial distortion). The systematic bias in the plot increases with distance from the image centre, suggesting it is primarily due to the radial distortion. Interpolation bias can be reduced by decreasing the square sizes of the calibration grids, although this approach is limited by the limits of camera resolutions, or by applying collinearity constraints [21]. Homographic interpolation has the additional disadvantage of requiring local image continuity. The general camera model makes no continuity assumptions, and can thus model discontinuous cameras. By using calibration grids that require interpolation, the applicability of generic calibration is reduced.

The use of active grids overcomes this problem with homographic interpolation by providing a direct localisation of the point seen by *every* pixel viewing the active grid, thus enabling pixel-level calibration. An active grid is implemented using a flat-screen TFT monitor that displays a temporal sequence of spatially varying greyscale patterns. The location of any point on the active grid can be decoded from the intensity displayed at that point across the sequence of patterns. We have used patterns from the domain of structured light to encode location. This approach is similar in spirit to the approach used by Sagawa et al. [22] for distortion correction. The active grids implemented for the work in this paper consist of 22 different patterns consecutively displayed on the monitor. Fig. 1b shows the set of patterns that encode vertical location. A more detailed description of the implementation of active grids can be found in [6].

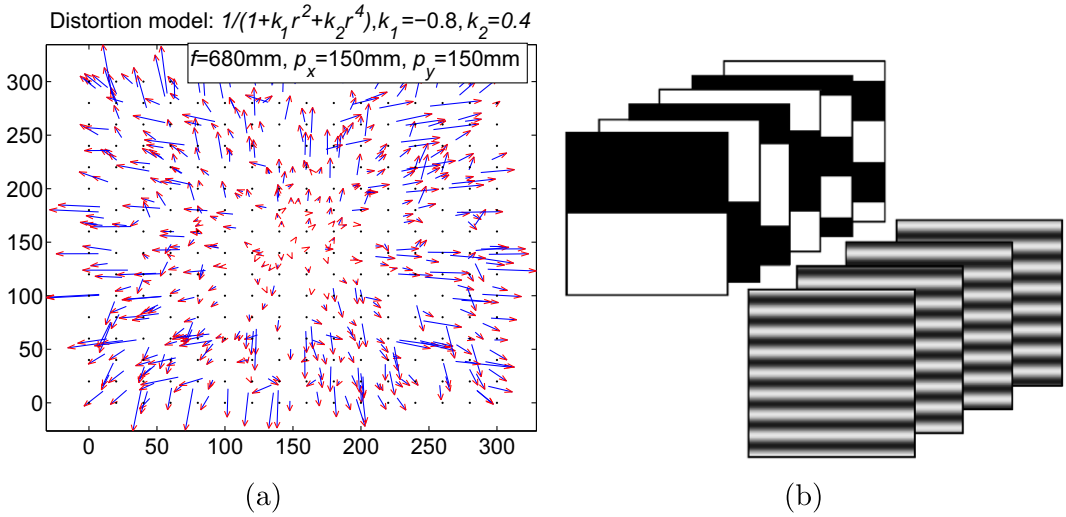

Fig. 1. (a) Vector plot of error residuals for homographic interpolation (20 mm grid pitch) showing bias. Vectors are scaled ×300; (b) binary and sinusoidal active grid patterns for encoding vertical location.

Active grids overcome the distortion bias associated with homographic interpolation and consequently are ideal for use in the generic calibration process. Their performance with respect to standard techniques for localising features on static calibration grids is next examined. Corner detection in chessboard patterns was recently shown to be invariant to both perspective bias and distortion bias, and so to outperform non-corner based patterns [15]. Consequently two chessboard corner localisation techniques, one derivative based and one saddle-point based, were used for the benchmarking process. If using a standard, static chessboard calibration grid, these methods are typical of the localisation techniques that would be used to subpixelly determine the grid corners. The comparison between the active grids method and these two standard methods is shown in Fig. 2. Details of the experimental setup are provided in [6]. The robustness of active grids to variations in camera-grid displacement, orientation, image blur and additive Gaussian noise is seen in Fig. 2 to be superior to that of the standard methods under almost all conditions. The robustness of active grids to image blur is an important benefit for the calibration process as it means that camera focus can be fixed during calibration. Note that these tests determine the relative performance of the methods for detecting the grid corner locations, and they do not address the errors resulting from localising points that are not grid corners (which would require bias inducing interpolation and thus reduce the performance of the static grid methods). It can be concluded from these results that active grids have localisation accuracy and robustness equal to or exceeding those of the standard methods for the detection of corners from chessboard calibration grids. Combined with the lack of interpolation bias, active grids are thus ideal for use in generic calibration, where pixel-level localisation is a key requirement of the calibration process.

3. Linear estimation

The purpose of the linear estimation stage in central generic calibration is to determine the grid poses and the position of the camera centre in the camera coordinate system attached to the base (usually first) grid. The camera centre is the single point through which all camera rays would pass if no reflection or refraction occurred, thus the accuracy of the entire calibration is directly dependent on the accuracy of the centre estimate. The linear estimation stage of the standard central generic method is based on a collinearity constraint: for each ray, the camera centre and the world coordinates of the intersection point of that ray with each grid are collinear. This can be expressed mathematically by stacking the global homogeneous coordinate for each intersection point in a 4×4 matrix. Collinearity is enforced by ensuring that all 3×3 subdeterminants of this matrix are zero. Determining the camera centre and plane positions and orientations from this starting point for the standard central generic method is not straightforward. Significant complexity arises due to coupling between many of the unknowns. See [24,25] for a description of the algorithm. The equations necessary to solve for the unknowns are presented in detail in [20]. Eqs. (20)–(39) in that paper, which are the same for both the minimal and non-minimal scenarios, indicate the level of involvement required to decouple the unknowns. As a further example of the standard method's complexity, the final equation to be solved is non-linear in the unknowns, and a solution is determined only by observing via simulation that several parameters are always equal to 0.

An alternative linear estimation stage is next proposed that is less complicated, and that is shown to be more accurate, than the method of standard generic calibration.

3.1. Synthetic pinhole calibration

The proposed linear estimation stage results from a novel interpretation of existing methods for the calibration of pinhole cameras. As known, pinhole calibration techniques are not suitable for wide FOV cameras due to the existence of severe non-linear image distortion that invalidates the pinhole projection model. For cameras with FOVs equal to and exceeding 180° the pinhole model itself is invalid, since there is no image plane location and positive focal length for which all 3-space points linearly projected through the camera centre can intersect. However, pinhole calibration methods are well established and it would be beneficial to utilise this established theory. The proposed linear estimation method enables such utilisation. The key idea is that an additional calibration grid, referred to as the base grid, is used as a synthetic image plane in the calibration process, thus forming a synthetic camera that is exactly pinhole. By placing the base grid in front of the general camera so as to intersect the camera rays on the object side of the camera optics, as shown in Fig. 3, a distortion free image is formed on the synthetic image plane. The synthetic image points are the points of intersection of the rays with the inserted synthetic

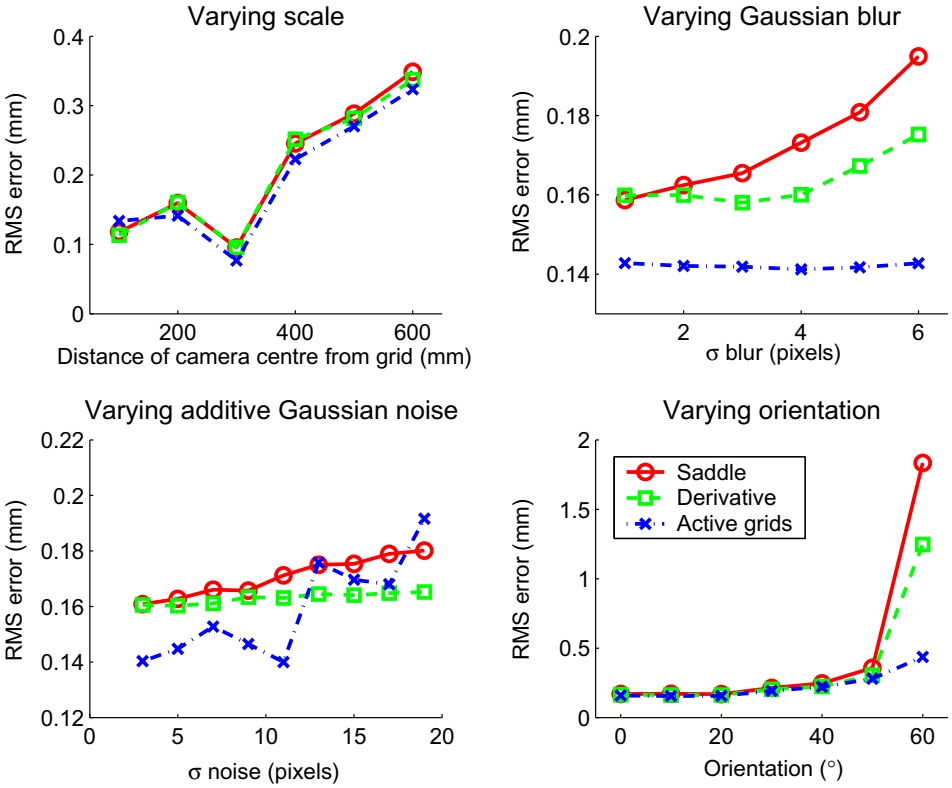

Fig. 2. Performance plots for saddle point localisation, derivative localisation and active grids localisation. Orientation is measured between the grid normal and the camera axis in the horizontal plane. Where not otherwise specified, the camera-grid distance is 200 mm, the orientation is 0°, and there is no blurring or additive noise.

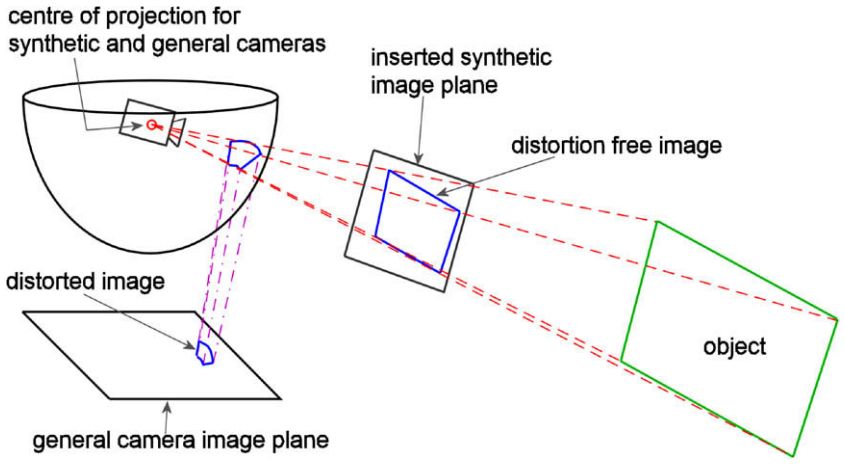

Fig. 3. Linear estimation of camera centre for proposed generic method. Synthetic image plane allows use of pinhole calibration techniques for determining centre.

image plane. If this plane is an active grid, the intersection locations can be determined directly as described in Section 2.

Consider the camera pixels that view the object in Fig. 3. The camera rays associated with these pixels are coincident at the centre of projection of the camera. Therefore the intersections of these rays with an additional grid, the synthetic image plane, will be a perspective projection of the object corners. By accurately determining the intersection points with the synthetic image plane of all the camera rays that intersect the object, a pinhole image of the object is formed on the synthetic image plane. Since the projection from the object through the synthetic image plane to the centre of projection preserves point collinearity, the synthetic image is free of all distortion. The synthetic camera can then be calibrated from at least two such synthetic images of a calibration grid in different positions. Any standard pinhole calibration method can be used to achieve this calibration. Since the centres of the synthetic and general cameras coincide, the desired estimate of the general camera centre is directly available from the synthetic camera calibration as $[p_x\ p_y\ f]^T$, where p_x and p_y are the principal point offsets of the synthetic camera, and f is the synthetic camera's focal length. The pose of grids two and three can also be extracted from the synthetic pinhole calibration using well known techniques [23]. Note that there is no constraint on the placement or the pose of the base grid acting as the synthetic image plane, once it is located

externally to the general camera. The only prerequisites for the proposed linear estimation are that the general camera is central and that the calibration targets are planar with known calibration patterns.

The non-linear calibration problem is thus linearised by moving the calibration from a point at which the optics are non-linear to a point at which they are linear. This new approach provides a key link between the established theory of pinhole calibration and the generic calibration of central cameras. A minimum of three grids is required for the proposed linear estimation stage – two for the pinhole calibration and one for the synthetic image plane – which is the same number as required for the standard generic method's linear estimation stage. The benefits of active grids as outlined in Section 2, in particular their ability to directly and accurately provide ray-grid intersections points, make them ideal for use as synthetic image planes in this method. Standard chessboards, in conjunction with homographic interpolation, could also be used to form the synthetic image plane. However, in that case the interpolation bias would introduce distortion into the synthetic camera and so it could no longer be precisely modelled with a pinhole model.

In [22] Sagawa proposes a distortion removal scheme in which an active grid is imaged, and the non-parametric mapping between the active grid locations and the image pixel locations is subsequently determined. This differs from the proposed calibration method in which a distortion-free image is directly formed for a synthetic pinhole camera with known image plane, completely removing the need for distortion correction.

A question arises as to which pinhole calibration technique should be used for the proposed linear estimation stage. To answer this, two well known pinhole calibration techniques, those of Sturm [26] and Wang and Liu [29], were incorporated into separate implementations of the proposed linear estimation stage. Both of these techniques are based on the same underlying constraints on the Image of the Absolute Conic (IAC), but they take different approaches to determining the solutions. The relative performance of these two implementations was evaluated, resulting in the conclusion that the method of Sturm is more accurate and robust than the method of Wang for this application (see [6] for more details). Therefore the pinhole calibration method of Sturm is used in the proposed linear estimation stage. This method is based on the IAC, ω, and the corresponding relationship between the IAC and the camera calibration matrix, K, given by $\omega \simeq K^{-T}K^{-1}$, where $\simeq$ denotes equality up to a non-zero multiple.

3.2. Determining camera centre and poses of initial grids

The proposed linear estimation stage for n calibration grids proceeds as follows. Let Q_{ij} be the intersection point of ray i with calibration grid j in a coordinate frame attached to grid j, and let the pose of calibration grid j be given by T_j. The perspectivity, H^j, due to central projection gives

$$Q_{i[1\ 2\ 4]} \simeq H^j Q_{ij[1\ 2\ 4]}, \quad \forall i \text{ and } j = 2, \ldots, n \tag{1}$$

Each H^j results in the following two linear equations in the elements of ω

$$\mathbf{h}_1^{jT}\omega\mathbf{h}_1^j - \mathbf{h}_2^{jT}\omega\mathbf{h}_2^j = 0 \tag{2}$$

$$\mathbf{h}_1^{jT}\omega\mathbf{h}_2^j = 0 \tag{3}$$

where $\mathbf{h}_i^j$ is the ith column of H^j. The synthetic camera's aspect ratio and skew are determined by the properties of the grid that acts as the synthetic camera's image plane. By using an active grid implemented on a TFT monitor that has square pixels and zero pixel skew, the aspect ratio and skew are 1 and 0, respectively. The unknown elements of ω are then $\{\omega_{11}, \omega_{13}, \omega_{23}, \omega_{33}\}$. The three degrees of freedom correspond to the values of p_x, p_y and f. Since the solution of ω is up to scale, there are three camera parameters to estimate, and in the minimal case there are four independent equations resulting from H^2 and H^3, so the system to be solved is actually overconstrained. The inhomogeneous camera centre, C, after solving linearly for ω using least squares, is given by

$$C = \begin{pmatrix} \frac{-\omega_{13}}{\omega_{11}} \\ \frac{-\omega_{23}}{\omega_{11}} \\ -\frac{\sqrt{\omega_{11}\omega_{33} - \omega_{13}^2 - \omega_{23}^2}}{\omega_{11}} \end{pmatrix} \tag{4}$$

The camera calibration matrix of the synthetic camera is

$$K = \begin{pmatrix} C_3 & 0 & C_1 \\ 0 & C_3 & C_2 \\ 0 & 0 & 1 \end{pmatrix} \tag{5}$$

where C_i is the ith element of C. Poses of the grids used in the linear estimation stage with respect to the base grid acting as the synthetic image plane can be extracted from the homographies H^j by factorisation. Letting $G^j = K^{-1}H^j$, and with $\mathbf{g}_i^j$ the ith column of G^j,

$$R_j = s\begin{pmatrix} \mathbf{g}_1^j & \mathbf{g}_2^j & \mathbf{g}_1^j \times \mathbf{g}_2^j \end{pmatrix} \tag{6}$$

$$\mathbf{t}_j = s\mathbf{g}_3^j + C \tag{7}$$

$$s = \frac{1}{mean\left(\|\mathbf{g}_1^j\| \ \|\mathbf{g}_2^j\|\right)} \tag{8}$$

where R_j is the rotation matrix describing the rotational component of the pose of grid j, $\mathbf{t}_j$ is the translational component of the pose of grid j, and s is a scale factor. The sign of s is chosen so that the planes are located on the same side of C as the synthetic image plane. An orthonormal rotation matrix can be obtained via the SVD as in [31].

The implementation of the proposed linear estimation stage requires the computation of homographies between large corresponding datasets. The size of the datasets is given by the number of pixels that see both grids, typically in the region of 50,000 point pairs (a random subset of the point pairs can be used to decrease homography computation time for very large point pair sets). Homographies are determined within a RANSAC framework that selects inliers as the point pairs that are in homographic correspondence with one another. The H^js are calculated using the standard DLT with normalisation [11], and the RANSAC parameters are selected based on the experimental results in Section 2. Outliers can exist in the decoded data due to image shot noise, non-linearities in the camera radiometric transfer function, and possible incorrect decoding of active grids due to sharp local discontinuities.

3.3. Bundle adjustment

In the standard generic calibration method, bundle adjustment is applied to the result of the linear estimation stage in order to improve the quality of the calibration result. The calibration result is the look-up table that maps camera pixels to ray directions in 3-space. For the general camera model, clearly it is the ray directions that should be adjusted in any bundle adjustment scheme, but there are several approaches to doing this. The rays are calculated as the join of the camera centre, C, and the 3-space ray-grid intersection points of ray i with plane j, P_{ij}. If the intersection points are known for more than one grid then the centroid of these points can be used. Also, $P_{ij} = T_j Q_{ij}$, where T_j is the pose of grid j. Thus the parameters that determine the calibration of each camera pixel are $[C, T_j]$. Consequently the ray directions can be updated directly, or updated by bundle adjusting one or both of $[C, T_j]$ for $j = 2, \ldots, n$.

The first possible scheme is to adjust the ray directions directly based on a ray-point distance metric. This is the scheme that is

applied in the standard generic calibration method [21]. In this case the ray directions, the camera centre C, and the T_js, $j = 2, \ldots, n$, are bundle adjusted to minimise the ray-point distances, with all rays forced to be concurrent at the current estimate of C. A second possible approach is to adjust the ray directions indirectly by optimising the T_js, and forcing the rays to pass through the adjusted C and the intersection points with the grids in the adjusted poses. For this approach only C and the T_js, $j = 2, \ldots, n$, are directly adjusted.

In either approach, the ray-point distance measure suggested by Ramalingam [21] can be used as the error metric. This metric measures the perpendicular distance between the ith ray and the P_{ij} intersection points, and in this paper it is determined efficiently, and without resorting to least-squares, using the dot product. Given the ray direction D_i, the closest point along D_i to P_{ij} is parameterised by λ_{ij} as

$$D_i \cdot (\lambda_{ij} D_i - (P_{ij} - C)) = 0 \tag{9}$$

where C is the inhomogeneous camera centre, resulting in the following solution for λ_{ij}

$$\lambda_{ij} = D_i \cdot (P_{ij} - C) \tag{10}$$

The ray-point distance, d_{ij}, is then given by

$$d_{ij} = \|\lambda_{ij} D_i - (P_{ij} - C)\| \tag{11}$$

$$= \|(D_i \cdot (P_{ij} - C)) D_i - (P_{ij} - C)\| \tag{12}$$

The bundle adjustment method that is applied in this paper is the second approach described above, where C and the T_js, $j = 2, \ldots, n$, are directly adjusted in order to minimise the geometric ray-point distance. After each bundle adjustment iteration, the ray directions are recalculated in a least squares sense [3] as the best fit rays to the new P_{ij}s, determined by the updated T_js, that pass through the updated C. This is a relaxation optimisation, in which the ray directions are indirectly updated at each iteration. It involves the minimisation of $3 + 6(n-1)$ parameters – 3 parameters for C, and 6 parameters for each T_j, $j = 2, \ldots, n$, using the Rodrigues representation. The alternative approach described above, in which the ray directions are directly updated in the bundle adjustment, requires two additional parameters to be minimised for each ray involved in the bundle adjustment (typically 1000s).

3.4. Simulated experiments

A comparison of the robustness to Gaussian noise of the linear estimation stages of the standard and proposed generic calibration methods is shown in Fig. 4. Errors in the estimation of the camera centre, and in the translation and rotation of the second and third grids involved in the linear estimation stage, are presented (averaged over 50 trials). The ray-point error is the perpendicular distance between each estimated ray and its known point of intersection with each calibration grid (see Section 3.3). These results are for a simulated camera with camera centre $[0\ 0\ 600]^T$ (in coordinate frame of base grid), and with focal length and distortion parameters chosen to simulate a wide angle camera with FOV of 100°. Results are shown for the standard generic method both with and without bundle adjustment of the grid transformations; bundle adjustment was not applied in the proposed linear estimation stage in these experiments.

The results clearly show that the proposed linear estimation stage without bundle adjustment outperforms that of the original generic method with bundle adjustment across all levels of noise tested. The results also indicate that bundle adjustment does not significantly improve the calibration result for the standard generic method, although the ray-point error is consistently reduced. Possibly this is due to a weak global minimum of the cost function, which is evidenced by the relatively larger improvement due to bundle adjustment at lower levels of noise compared to the improvement at higher noise levels. The translation and rotation error measures are indirect measures that are coupled in the transformations, and so their errors cannot be considered independently. Examination of the mean errors in the intersection point locations after applying the transformations estimated by the standard generic method, and by the standard generic method with bundle adjustment, reveal that the application of bundle adjustment does indeed reduce this error, despite the larger rotation errors for the standard method with bundle adjustment shown in Fig. 4.

4. Pose estimation

Pose estimation is required during generic calibration in order to increase the number of calibrated camera rays. Once the pose of an additional grid is estimated, the camera ray associated with each pixel that sees this additional grid can be included in the calibration. Exact solutions to the general pose estimation problem can be found for either three or four non-collinear point-image pairs by solving a fourth or higher degree polynomial [10]. However, closed form solutions to the general pose estimation problem for more than four points are not straightforward [14]. In the standard generic method [21] a geometric three point algorithm for estimating the pose is described that operates for both central and non-central cameras. Taking the central case, given calibrated rays with directions R_i and R_j, and the distance d_{ij} between their intersection points with the grid of unknown pose, the depths λ_i and λ_j of the intersection points along R_i and R_j can be computed by simultaneously solving $\|\lambda_i R_i - \lambda_j R_j\|^2 = d_{ij}^2$ for $i,j = (l,m,n)$, $i \neq j$. However, when included in a RANSAC framework, re-estimation of the pose using all inliers (typically the finally step in RANSAC) is not possible using this algorithm. The method is also very sensitive to additive noise (although a guided selection of sufficiently separated points can alleviate this problem). To overcome these drawbacks a linear least-squares solution to the pose estimation problem is described that is applicable for central generic calibration. Although the method does not minimise geometric error, it is linear, fast, always gives a solution, and can conveniently be incorporated within a RANSAC framework.

The method allows pinhole pose estimation to be applied to central generic cameras through the use of a synthetic image plane. By intersecting the previously calibrated camera rays that see points on the unknown grid with the synthetic plane, the distortion free pinhole projection of the unknown grid is formed on that plane. Consider this synthetic plane as the image plane of a synthetic pinhole camera, where the pinhole camera's centre is coincident with the camera centre of the general camera. With reference to Fig. 5, given a grid in the base position with world coordinate points $X_i \in P^3$, and a grid with an unknown pose T relative to the base grid containing unknown world points $X'_i \in P^3$, the goal is to determine the unknown pose T. Although general cameras are usually not pinhole, a solution is possible via the insertion of the synthetic image plane in a known orientation (the orientation selection is discussed later) between the camera centre and the grid with unknown pose, as shown in Fig. 5. Since the pose of the synthetic image plane is chosen, all the intrinsic and extrinsic parameters of the synthetic pinhole camera are known. They are the camera projection matrix P, camera calibration matrix K, camera rotation R_s, and camera centre C. Pose estimation can therefore proceed using the established pinhole camera method that is described in [23]. Note that the synthetic image plane is a mathematical construct only and is not physically realised. The general pose estimation problem for central cameras is therefore cast as a pinhole pose estimation problem for which an established solution is available.

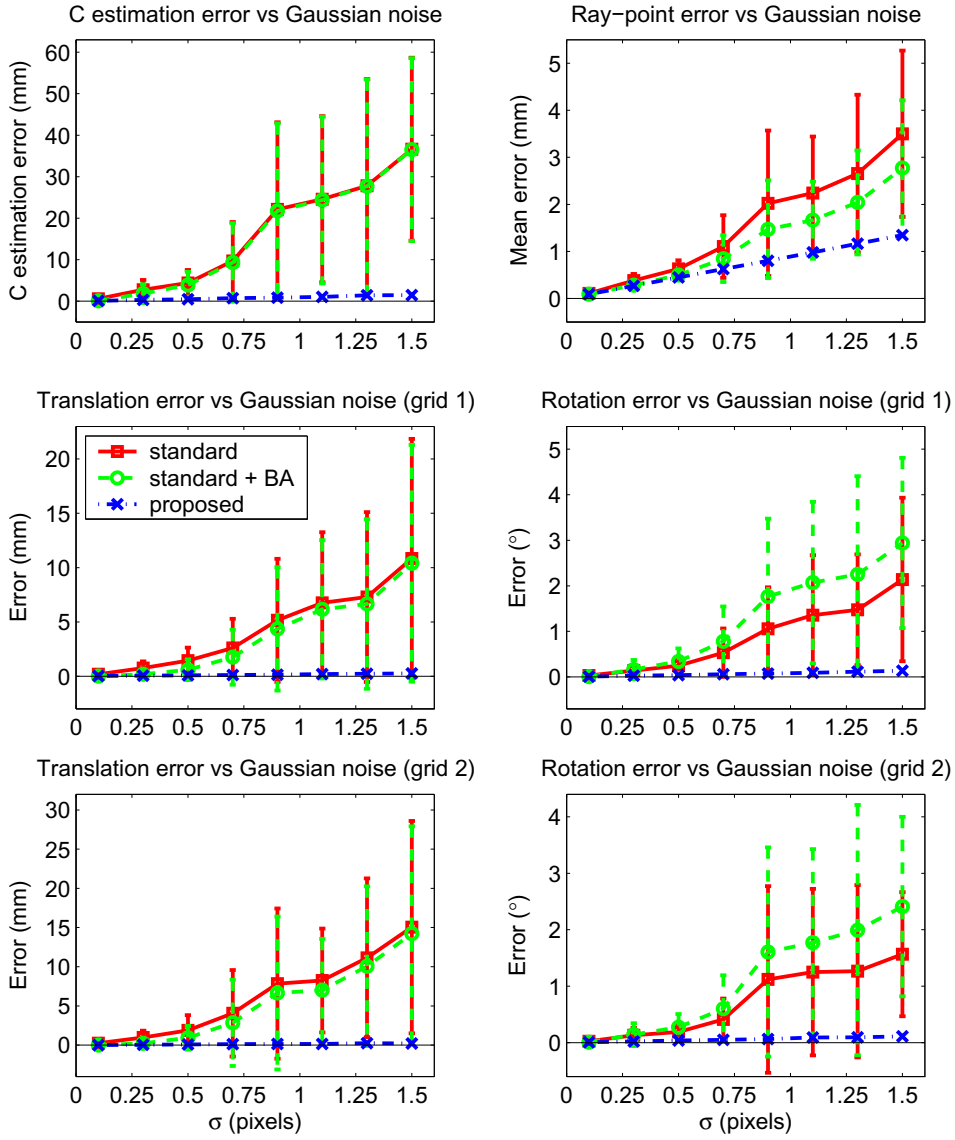

Fig. 4. Centre and transformation estimation performance plots versus Gaussian noise for standard generic method, standard generic method with bundle adjustment, and proposed generic method (BA = bundle adjustment). The rotation error is defined as the sum of the out-of-plane and in-plane rotation errors. Note that the SDs for the proposed generic method in these results are non-zero, but are significantly smaller in magnitude than the SDs of the standard generic method. The numerical data underlying these plots is contained in the document cviu_09.pdf available in the additional results section of the CIPA code archive – www.cipa.dcu.ie/code.html.

An outline derivation of the pose estimation equations can be obtained by considering the point transformations shown in Fig. 5 described by

$$x'_i \simeq PX'_i \tag{13}$$

$$x'_i \simeq HX_{i[1\ 2\ 4]} \tag{14}$$

$$X'_i \simeq TX_i \tag{15}$$

By combining these equations and substituting $P = KR_s[I - C]$, it follows that

$$(KR_s)^{-1}H \simeq (\mathbf{t}_1 \quad \mathbf{t}_2 \quad \mathbf{t}_3 - C) \tag{16}$$

where $\mathbf{t}_i$ is the ith column of T. Letting $G = (KR_s)^{-1}H$, and applying Eqs. (6)–(8), a solution for the rotation, $\widehat{R}$, and the translation, $\mathbf{t}$, of the grid with unknown pose can be determined. An orthonormal rotation R is obtained from $\widehat{R}$ via the SVD. Non-linear minimisation can subsequently be applied to the linearly estimated pose using the ray-point error metric described in Section 3.3.

One is free to choose the orientation of the synthetic image plane, after which the synthetic camera parameters are directly determined. Ideally the synthetic image plane should be as close as possible to perpendicular to the known rays involved in the pose estimation process. This orientation can be determined in a least-squares sense by minimising the sum of the angles between the calibrated rays and the normal of the synthetic image plane, in a similar way to [19]. The unit plane normal, $\mathbf{n}$, is found as the solution to

$$\text{argmin}_{\mathbf{n}} \sum_{i=1}^{m} \|[D_i]_\times \mathbf{n}\|_2 \quad \text{subject to } \|\mathbf{n}\| = 1 \tag{17}$$

where D_i is the unit vector representing ray i, and $[D_i]_\times$ is its corresponding skew-symmetric matrix.

The pose estimation schemes of the standard and proposed generic methods were evaluated against each other for both simulated and real data. Both methods are incorporated in RANSAC frameworks for all experiments. For the standard generic method pose

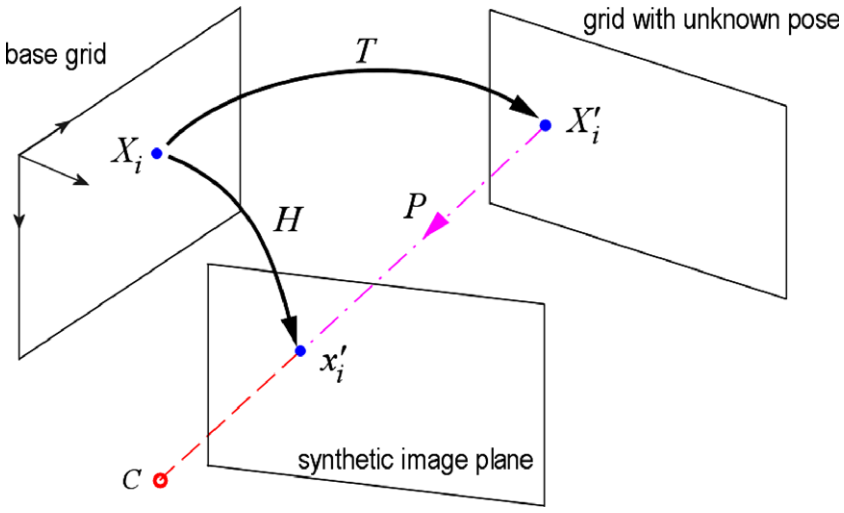

Fig. 5. Proposed linear pose estimation method using synthetic image plane.

estimation, the re-estimation stage in RANSAC is performed by non-linear minimisation of the error metric described in Section 3.3. In the proposed method, RANSAC re-estimation uses the linear pose estimation procedure described above. The robustness to Gaussian noise of the pose estimation methods was first evaluated for a synthetic wide angle camera with camera centre fixed at $[0\ 0\ 600]^T$. The translations and Euler rotations of the grid whose pose was to be estimated were randomly chosen from $[-150$ mm 150 mm$]$ and $[-30°\ \ 30°]$, respectively. The mean rotational and mean percentage translational errors over 50 trials are shown in Fig. 6. It is seen that the proposed generic method's pose estimation outperforms that of the standard generic method over all simulated levels of noise.

The two pose estimation methods were also evaluated against each other using real data, so that their robustness to outliers could be determined. The experiment consisted of a perspective camera (Kodak MegaPlus 1.4i) imaging an active grid in two random orientations. The camera was pre-calibrated using the plane-based method of Sturm [26], allowing the ray-pixel look-up table to be determined directly. A 3D laser scanner (depth resolution <0.1 mm) was used to determine the ground truth transformation between the two grids. The experimental results in Table 1 show the errors in the estimated relative pose between the grids for each pose estimation method. These results indicate the importance of using RANSAC for the linear pose estimation, as without RANSAC the linear pose estimates are seen to degrade significantly due to outliers in the data. The larger error magnitudes of the real experimental results relative to the simulated ones are due to the smaller grids and larger camera-grid distances used in the real experiments.

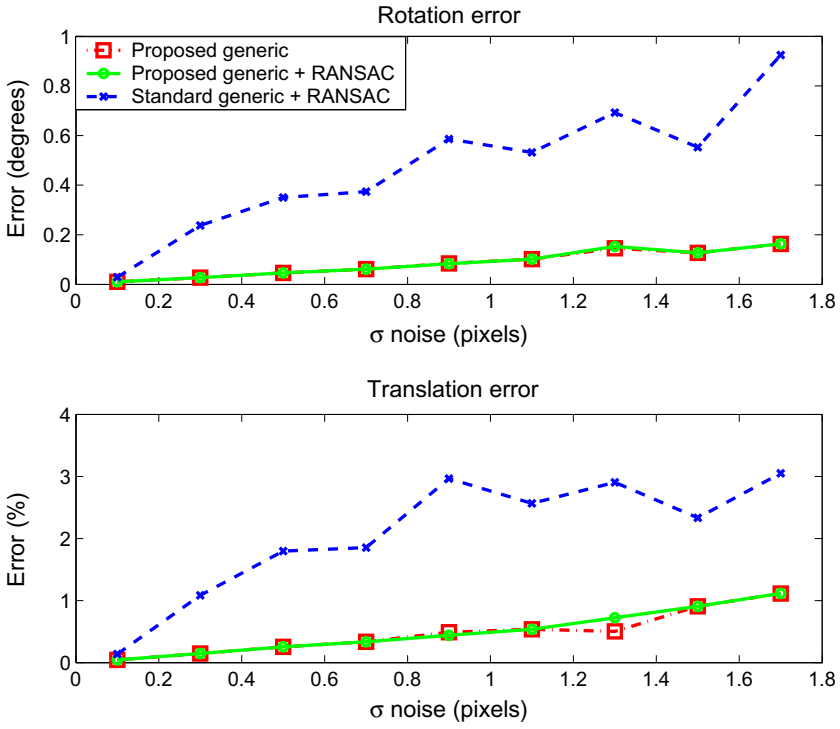

Fig. 6. Performance comparison for synthetic data of pose estimation stage of standard generic method with RANSAC, proposed generic method, and proposed generic method with RANSAC. The rotation error is defined as the sum of the out-of-plane and in-plane rotation errors.

Table 1
Pose estimation errors for real data using pose estimation stages of standard generic method with RANSAC, proposed generic method, and proposed generic method with RANSAC. The rotation error is defined as the sum of the out-of-plane and in-plane rotation errors.

	Standard + RANSAC	Proposed	Proposed + RANSAC
Rotation error (°)	1.6265	1.6617	0.8706
Translation error (mm)	9.2662	17.7704	7.6266

5. Experimental results

Both the standard and proposed generic methods are analysed for real data with respect to linear estimation calibration parameters, a ray-point error metric, distortion correction, and separate motion reconstruction tasks. An omnidirectional catadioptric camera and a camera with fisheye lens were used to capture the images for the experiments. The omnidirectional camera consists of a 360 OneVR hyperboloidal omnidirectional mirror[1] mounted on a Nikon D70 SLR digital camera. With the correct positioning and alignment this catadioptric configuration has a single centre of projection. However, the mirror could not be mounted directly onto the camera due to limitations on the minimum focusing distance of the camera lens, and thus an external bracket was used to fix the configuration. Correct alignment with this could not be guaranteed. The second camera used was a Nikon CoolPix 4500 digital camera attached to a Nikon FC-E8 fisheye converter lens, which has a 183° FOV. For each calibration method approximately 207° of the horizontal FOV and approximately 82° of the vertical FOV of the omnidirectional camera was calibrated; for the fisheye camera approximately 94% of the entire FOV was calibrated. Three grids were used in each calibration for the linear estimation stage, and a further three grids for the omnidirectional and seven grids for the fisheye cameras to extend the calibrated regions to include additional pixels. A minimal number of grids were used in the linear estimation stage of calibration in these experiments in order to examine the relative performance of the standard generic method and the proposed generic method in the most challenging case. For high fidelity calibration more than the minimum number of grids should be employed. Active grids were used for all grids during calibration, with the same images used as input to both calibration methods so that direct comparisons between the standard generic and proposed generic methods are not influenced by the type of input data. A RANSAC stage is applied to the locations decoded from the active grids in order to remove any incorrectly decoded location data. Normalisation is applied in the computation of all homographies.

The estimates of the camera centre, and the estimated out-of-plane and in-plane rotation angles and translation magnitudes for the poses of the second and third grids used in the linear estimation stages for each calibration method and for each camera are shown in Table 2. Both linearly estimated and bundle adjusted values are presented. The difference between the linearly estimated values of the camera centre for each method for the omnidirectional camera is 5.48 mm, compared to 2.88 mm for the fisheye camera. Similarly, linear estimates of the remaining calibration parameters for each calibration method are larger for the omnidirectional camera (within 15.8% of each other) than for the fisheye camera (within 6.77% of each other). Considering the bundle adjusted values for each camera, it is clear that they converge towards the same global values for each calibration method. Importantly, excepting one value, these global values are significantly closer to the linear estimates of the proposed method than to the standard method's linear estimates, indicating superior performance of the proposed method's linear estimation stage. In the case of the fisheye camera, the close agreement between the bundle adjusted values for each calibration method indicate that a global minimum in the solution space has been reached. In the case of the omnidirectional camera 469 iterations were performed during bundle adjustment. For the fisheye camera, the bundle adjustment determined a global minimum solution after 412 iterations when initialised with the standard method linear estimation values, but after only 174 iterations when initialised with the proposed method linear estimation values.

[1] Kaidan Inc., Feasterville, PA.

Table 2
Camera centre and grid transformation estimates for omnidirectional and fisheye camera calibrations, both before and after the application of bundle adjustment (BA). Centre and translations are measured in mm, rotations are measured in degrees.

Camera	Parameter	Standard method		Proposed method	
		Linear	+BA	Linear	+BA
Omnidirectional	C	168.05	168.00	167.90	168.07
		156.20	159.37	159.82	160.33
		−112.10	−115.43	−116.21	−116.90
	R_2	35.06	36.24	36.52	36.77
		4.42	4.82	4.87	4.94
	$\mathbf{t}_2$	135.20	138.63	139.66	140.76
	R_3	34.33	35.40	35.61	35.65
		8.25	9.47	9.56	9.60
	$\mathbf{t}_3$	188.78	191.11	192.13	192.45
Fisheye	C	104.71	106.17	107.00	106.20
		161.89	162.30	162.29	162.32
		−125.69	−124.71	−124.01	−124.69
	R_2	43.09	42.65	42.33	42.66
		5.17	4.90	4.84	4.89
	$\mathbf{t}_2$	121.80	120.86	120.22	120.88
	R_3	37.55	37.09	36.84	37.09
		6.03	5.90	5.83	5.90
	$\mathbf{t}_3$	288.07	284.46	282.39	284.41

The ray-point error metric, described in Section 3.3, is applied to each calibration dataset to give an indication of the relative errors in the calibrations (ground truths for the camera centres and the second and third grid positions are not known). Table 3 shows the mean and standard deviation of the ray-point errors for each method and for each camera, both before and after bundle adjustment (rays involved in linear estimation stage only). The ray-point errors for the omnidirectional camera calibration are smaller by a factor of 8 for the proposed generic method than for the standard generic method. There is less difference in the magnitude of the errors for the calibration methods when applied to the fisheye camera, but the mean and standard deviation of the errors before bundle adjustment for the proposed generic method are smaller than for the standard method. For both cameras, bundle adjustment reduces the ray-point errors of the standard generic method calibrations significantly. The relatively small ray-point errors for the omnidirectional camera after bundle adjustment indicate that misalignment of the omnidirectional mirror with the camera is not significant. The difference in performance between the standard and proposed calibration methods is less for the fisheye camera than for the omnidirectional camera. Additionally, the ray-point error is similar across both camera types for the linear estimates using the proposed generic calibration, but is different by a factor of 5 for the linear estimates of the standard generic calibration. The results in Tables 2 and 3 indicate both the importance of bundle adjustment for the standard generic calibration method, and the superior initialisation for bundle adjustment that the proposed generic calibration method provides. Note that bundle adjustment is incorporated in the calibrations with both the proposed generic method and the standard generic method for all remaining results presented.

The centrality of the camera configurations after completion of the linear estimation stage was examined to get further insight

Table 3
Ray-point errors (mm) for all rays involved in the linear estimation stage for each calibration method and for each camera (BA = bundle adjustment).

Method	Error type	Omnidirectional		Fisheye	
		Error	Error after BA	Error	Error after BA
Standard method	Mean	1.6440	0.1413	0.3235	0.0727
	SD	0.8186	0.0724	0.1579	0.0448
Proposed method	Mean	0.1924	0.1281	0.1397	0.0727
	SD	0.0906	0.0731	0.0709	0.0448

into the quality of the calibrations. This was done by determining the best-fit 3-space ray $\hat{i}$ to the world intersection points P_{ij} of ray i with grid j for $j = 1,\ 2,\ 3$, which are estimated in the calibration. A method for least squares ray fitting in 3-space has recently been presented in [3], but for the case of only three points an efficient solution can be calculated similarly to the 2D case by fixing a coordinate frame to the plane defined by the join of the P_{ij}s for each i, given by the null vector of $[P_{i1}\ P_{i2}\ P_{i3}]^T$. The rays $\hat{i}$ can then be intersected with a plane passing through the estimated camera centre C that is closest to perpendicular to the $\hat{i}$s (see Section 4 for a method of calculating this plane). The distribution of the intersection points on this plane indicates the extent of centrality of the camera calibration. Fig. 7 shows plots of the distributions of the $\hat{i}$ ray intersections for each calibration method for each camera. These plots are for the estimated camera centre and grid poses after application of the linear estimation stage only and before bundle adjustment is applied. It can be seen that the distribution is highly non-Gaussian and non-isotropic for the omnidirectional camera calibrated with the standard generic method. In contrast, the distribution is both more compact and closer to Gaussian for the same camera calibrated using the proposed method. The wide distribution for the standard generic method may partly be a result of inexact camera centrality, although the distribution for the proposed generic method indicates that it achieves a solution that is geometrically consistent with an approximately central configuration. For the fisheye camera, the results using the two calibration methods are very similar, although the distribution is marginally closer to Gaussian for the proposed generic method. Isotropic Gaussian distribution of the errors indicates that the camera centre estimate and grid poses are in geometric agreement, with the error resulting solely from the Gaussian noise in the intersection point locations. The fisheye camera centre estimated in the proposed linear estimation stage is within the convex hull of the intersection points. However this is not the case for the camera centre linearly estimated using the standard method, indicating greater inconsistency between the estimated camera centre and estimated grid poses for that method.

5.1. Distortion correction

Two distortion correction experiments were carried out in order to both qualitatively and quantitatively evaluate each of the calibration methods.

In the first experiment the calibration data was used to remove the inherent non-linear distortion from the calibrated regions of images of real scenes. For the omnidirectional camera, a portion of a cylindrical image was formed by intersecting the calibrated rays with a unit cylinder, the axis of which was coincident with the camera centre, and then unwrapping the cylinder to form a planar image. Fig. 8 shows the original images and the cylindrically unwarped images calculated for the same pixels using the calibration data from the standard generic and proposed generic methods.

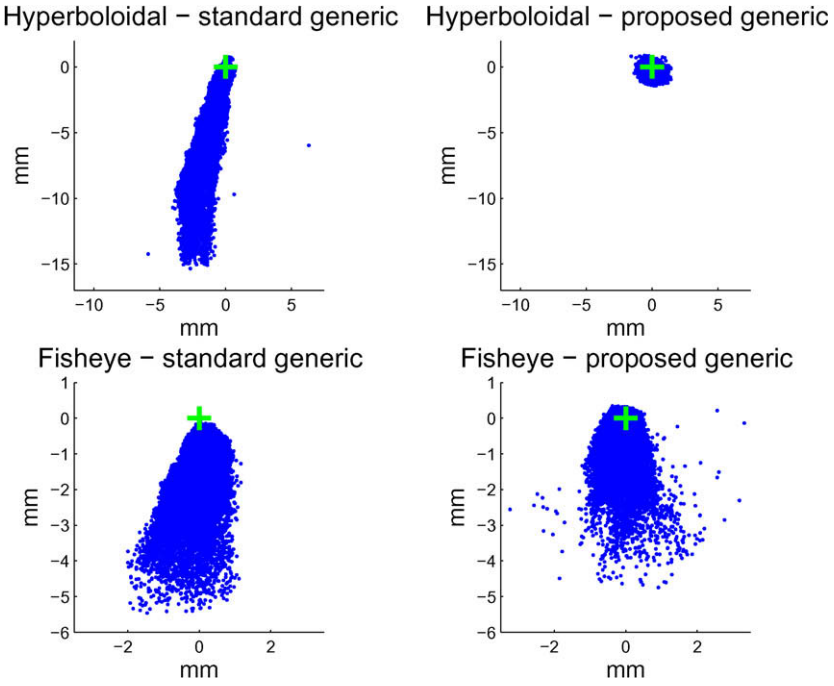

Fig. 7. Distribution of intersection points of best-fit rays for grids 1–3 with perpendicular plane passing through estimated camera centre. Camera centre shown with +. Only rays that intersect all three grids are considered. Note that the axes' scales are smaller for the fisheye plots.

As expected, real world straight lines that are parallel to the mirror axis (vertical) are mapped to straight lines in the images corrected using either method. However, some aberrations are visible in the images corrected using the standard generic method. In contrast, the corrected images formed using the proposed generic method have less aberration. Ellipses highlight the regions where aberrations are present in the distortion corrected images.

Distortion correction for the fisheye camera is best demonstrated by generating perspectively corrected images from the originals. This is readily achieved by intersecting the calibrated camera rays with a plane whose orientation is determined as in Section 4. The results after perspective correction using each calibration dataset, and for the same image pixels, are shown in Fig. 9. Similarly to the distortion corrected omnidirectional images, there are some visible aberrations in the images corresponding to the meeting points of mis-estimated grids. This difference is most noticeable in the image region highlighted with an ellipse in Fig. 9d. The corresponding image region in Fig. 9f has less aberration.

Quantitative evaluation of the calibrations was carried out by generating perspectively corrected images of planar chessboard grids. The plane onto which the corrected images were projected was selected as described in Section 4. Distortion residuals were measured for each image after applying a homography between the distortion corrected image grid corners and the known metric grid structure. Fig. 10 shows the distortion residuals for both the

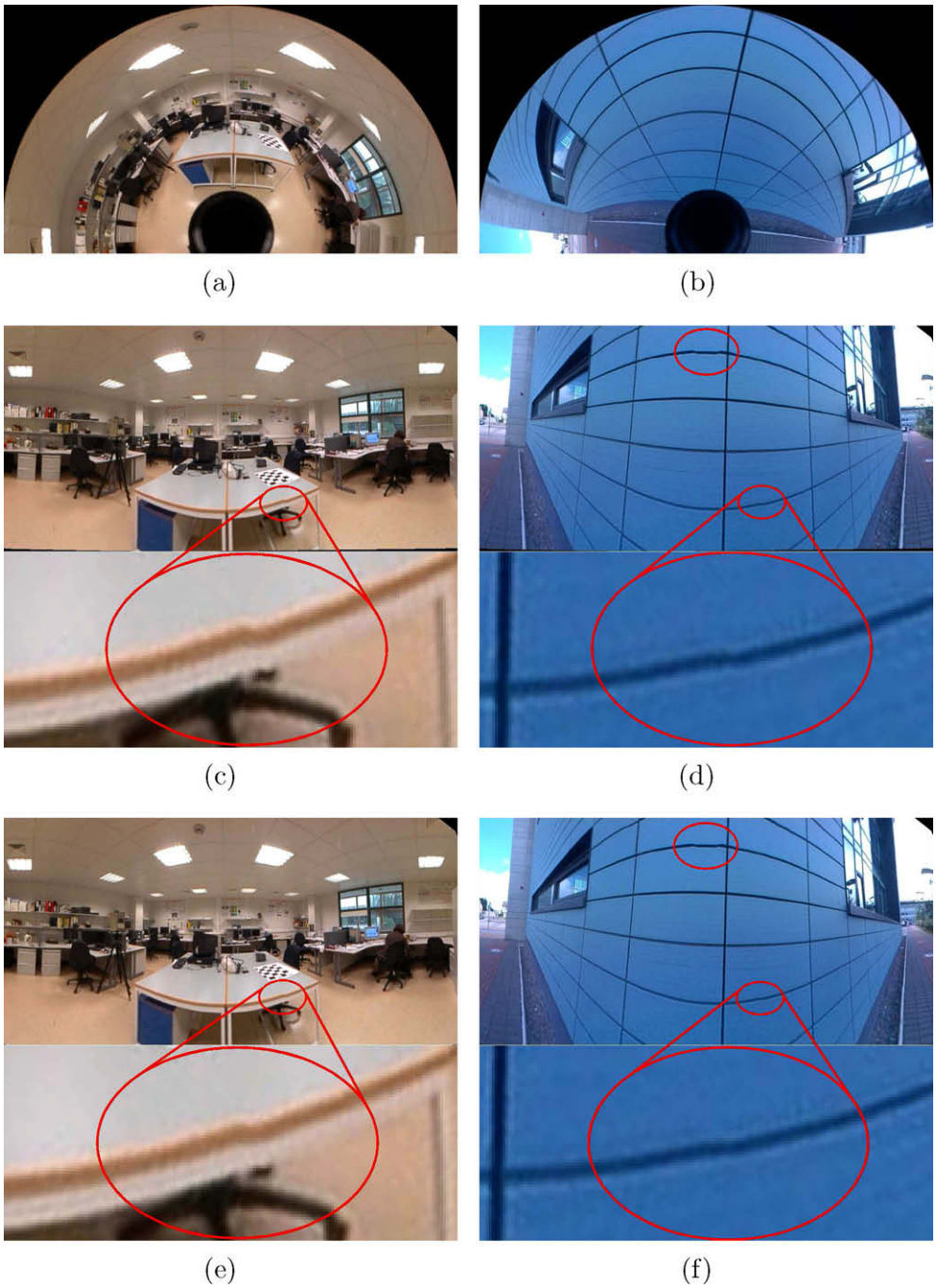

Fig. 8. Omnidirectional camera distortion correction results: (a and b) original omnidirectional images; (c and d) cylindrically unwarped images after standard generic calibration; (e and f) cylindrically unwarped images after proposed generic calibration.

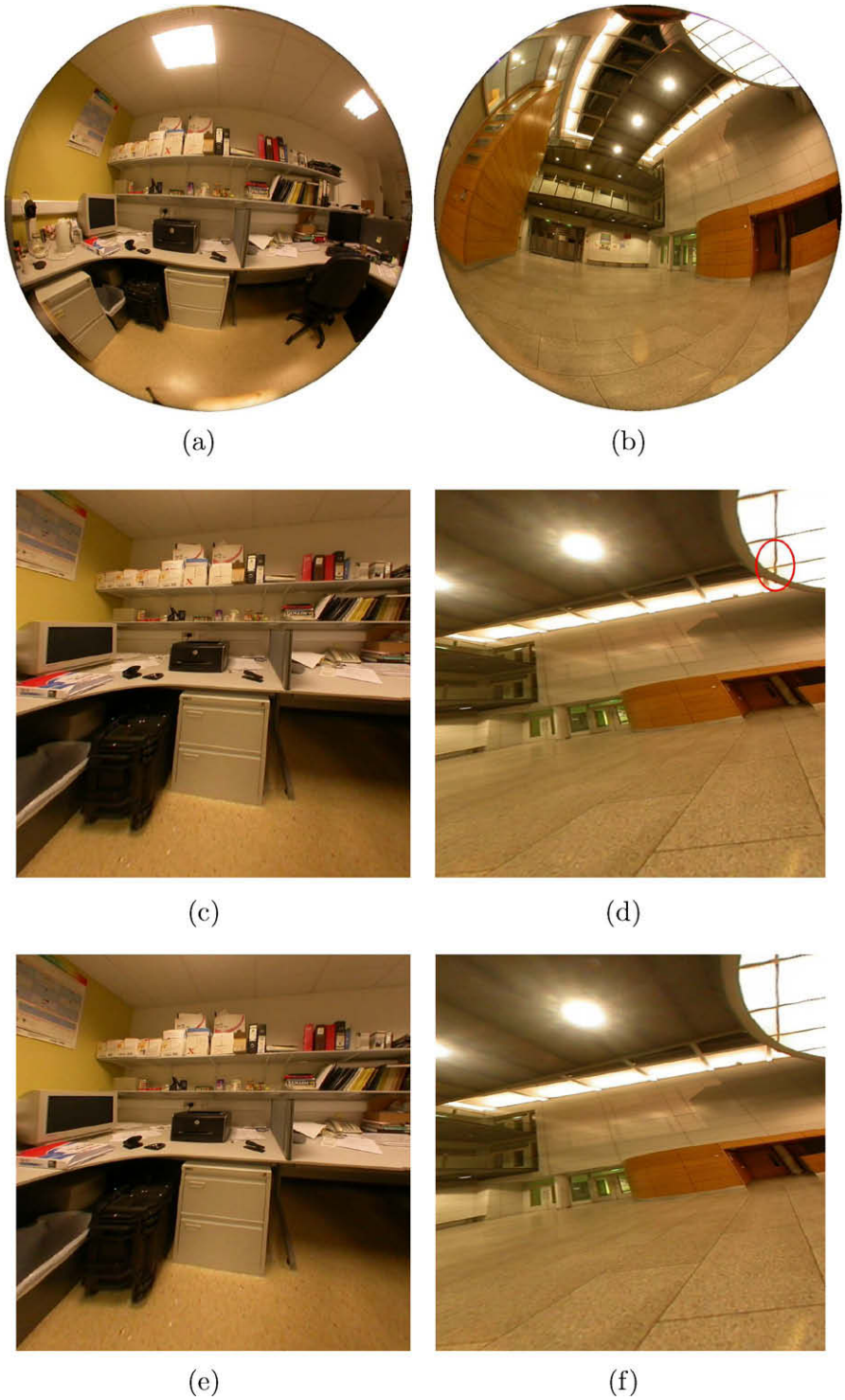

Fig. 9. Fisheye camera distortion correction results: (a and b) original fisheye images; (c and d) perspectively corrected images after standard generic calibration; (e and f) perspectively corrected images after proposed generic calibration.

standard and proposed methods for both cameras. In the case of the omnidirectional camera, no radial distortion bias is visible in either vector plot, but the plots do display large divergences along a roughly vertical line at the left of Fig. 10a and b. These coincide with areas where two grids with mis-estimated pose meet, and correspond to some of the aberrations seen in Fig. 8c–f. The divergences for the omnidirectional camera seen in Fig. 10 and the distortion correction residuals given in Table 4 are smaller for the vector plot using the proposed generic method than for the vector plot using the standard generic method, indicating a better calibration. For the fisheye camera the distortion residual plots are almost identical across the two calibration methods. The mean and standard deviation of the residuals for the fisheye camera using both methods are also very similar, which is as expected given the similarity in the bundle adjusted calibration values for each method (see Table 2), and since the distortion corrected is applied to the centre region of the image where the rays are determined by the linear estimation stage alone.

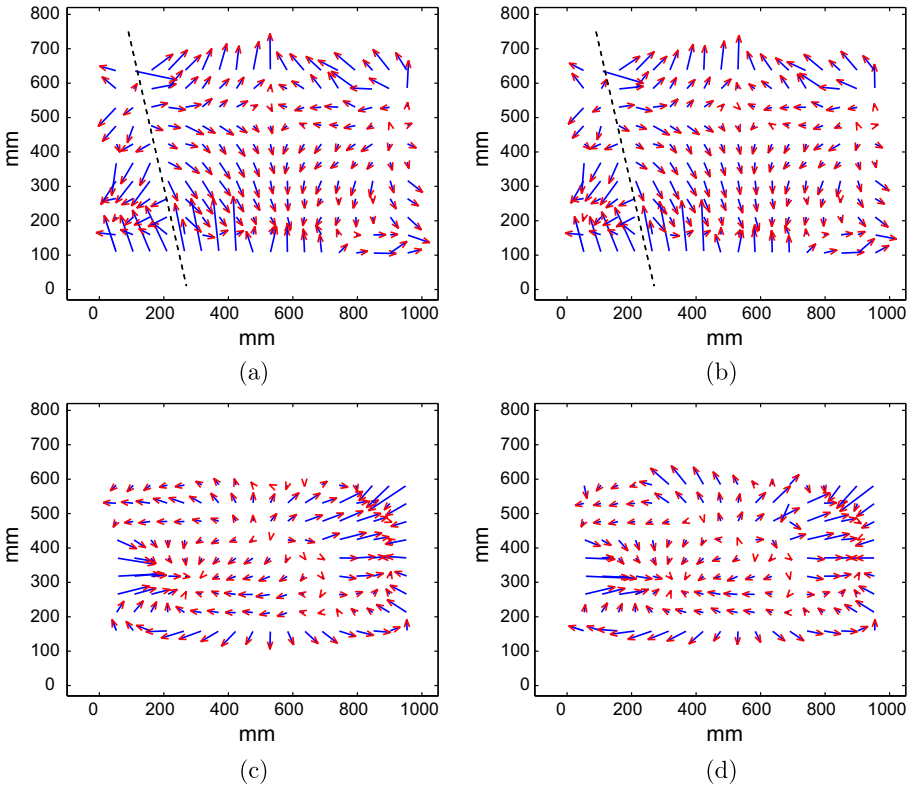

Fig. 10. Vector plots of residuals after perspective correction of a chessboard grid: omnidirectional camera using (a) standard generic method calibration data, and (b) proposed generic method calibration data; fisheye camera using (c) standard generic method calibration data, and (d) proposed generic method calibration data. Distortion corrected points are mapped to the same size metric grid for both cameras so as to enable direct comparison of the residuals. Vectors are scaled $\times 25$.

5.2. Motion reconstruction

Motion reconstruction experiments were conducted with the omnidirectional camera for the cases of pure translation and pure rotation. The similarity of the results for the fisheye camera calibration using the two calibration methods, and after applying bundle adjustment, mean that any motion reconstruction experiments would likely be too similar for the purpose of comparison.

The experimental setup consisted of a 3D calibration object (two orthogonal planar chessboard grids) rigidly mounted on a stage capable of controlled rotation and translation. For the translation experiment, the object was translated 100 mm in steps of 20 mm, and for the rotation experiment it was rotated by 110° in steps of 10°. Point matches were extracted across both image sequences, and were used to index the Plücker matrix lookup tables for each calibration method to get the corresponding ray direction information. The essential matrix, E, between each image pair was linearly estimated using the ray-based epipolar constraint

$$\mathbf{L}'E\mathbf{L} = 0 \tag{18}$$

where $\mathbf{L}$, $\mathbf{L}'$ are the first three components of the Plücker vectors derived from the Plücker matrices [19]. Rotations and translations are extracted from the essential matrices according to [18]. The motion reconstruction results are shown in Fig. 11. It can be seen, and was verified numerically, that the motion estimated with the proposed generic method is closer to linear in the case of translation, and closer to the ground truth value of 110° in the case of rotation, than for the standard generic method. For visualisation purposes the differences between the average translation vector and the estimated translation vectors are scaled $\times 10$ for each method.

Table 4
Residuals (mm) after distortion correction for omnidirectional and fisheye cameras using standard and proposed generic methods.

Method	Error type	Omnidirectional	Fisheye
Standard method	Mean	1.96	1.54
	SD	1.10	0.94
Proposed method	Mean	1.83	1.55
	SD	1.05	0.92

6. Conclusions

This paper proposes an improved method of generic camera calibration for cameras with a single centre of projection. The main contribution of the paper is a novel linear estimation stage based on a new interpretation of an existing technique that allows pinhole calibration techniques to be applied to the generic calibration of non-pinhole cameras. A performance evaluation of active grids for use in generic calibration, and a linear pose estimation stage for estimating the poses of additional grids in the calibration, are also presented. Individual components of the proposed method are separately evaluated using simulated data, with the results clearly showing that the proposed generic method outperforms the standard generic method in terms of accuracy and robustness to noise. The complete proposed generic method is also evaluated against the standard generic method using real data for both an omnidirectional camera and a fisheye camera, with the results for distortion correction and motion reconstruction tasks for the omnidirectional camera demonstrating the improved performance of the proposed generic method. For the real data, the proposed linear estimation stage achieves results that are shown to be significantly closer to the final camera solution, thereby improving efficiency by requiring less effort in bundle adjustment. It is concluded that the proposed generic method should be used instead of the standard generic method in order to attain the best results for the generic calibration of central cameras.

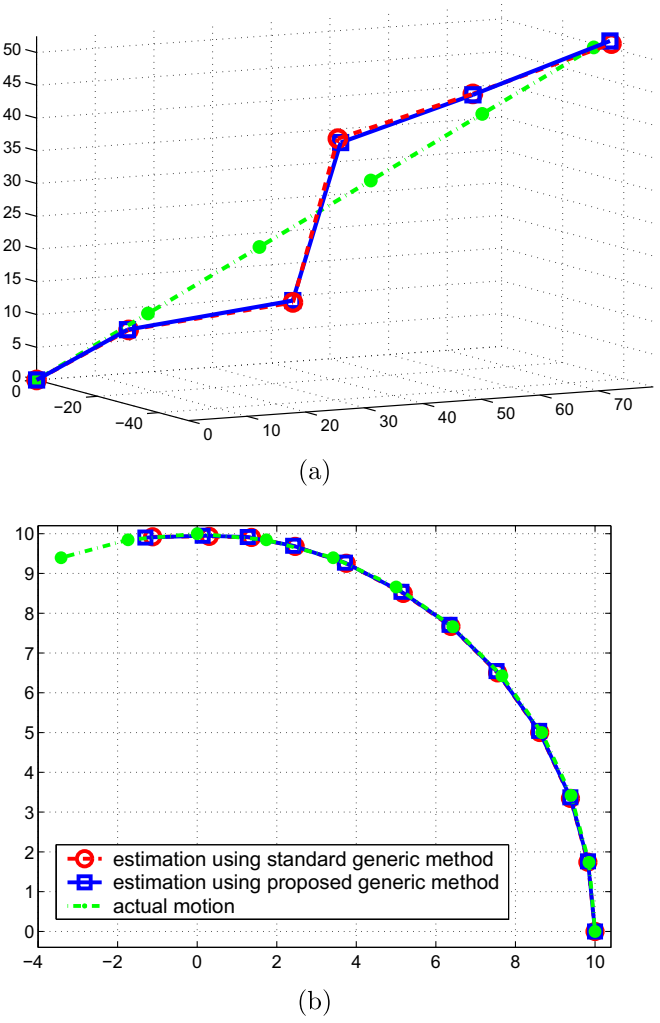

Fig. 11. (a) Translation and (b) rotation reconstruction using calibration data from standard generic method and proposed generic method. The numerical data underlying these plots is contained in the document cviu_09.pdf available in the additional results section of the CIPA code archive – www.cipa.dcu.ie/code.html.

Acknowledgment

This research is funded by the Irish Research Council for Science, Engineering and Technology funded by the National Development Plan.

References

[1] S. Baker, S. Nayar, A theory of catadioptric image formation, in: Proceedings of the 6th IEEE International Conference on Computer Vision, Bombay, India, 1998, pp. 35 – 42.
[2] J.P. Barreto, A unifying geometric representation for central projection systems, Computer Vision and Image Understanding 103 (3) (2006) 208–217.
[3] J.P. Barreto, J.M. Santos, P. Menezes, F. Fonseca, Ray-based calibration of rigid medical endoscopes, in: Proceedings of the 8th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras, Marseille, France', 2008.
[4] T. Bonfort, P. Sturm, P. Gargallo. General specular surface triangulation, in: Proceedings of the Asian Conference on Computer Vision, Hyderabad, India, vol. II, 2006, pp. 872–881.
[5] A.K. Dunne, J. Mallon, P.F. Whelan. A comparison of new generic camera calibration with the standard parametric approach, in: Proceedings of the IAPR Conference on Machine Vision Applications, Tokyo, Japan, vol. 1, 2007, pp. 114–117.
[6] A.K. Dunne, J. Mallon, P.F. Whelan. Efficient generic calibration method for general cameras with single centre of projection, in: Proceedings of the 11th IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, 2007.
[7] A.W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, USA, vol. 1, 2001, pp. 125–132.
[8] C. Geyer, K. Daniilidis, An unifying theory for central panoramic systems and pratical implications, In Proceedings of the 6th European Conference on Computer Vision, Dublin, Ireland, 2000, pp. 445–461.
[9] M.D. Grossberg, S.K. Nayar. A general imaging model and a method for finding its parameters, in: Proceedings of the 8th International Conference on Computer Vision, Vancouver, Canada, vol. 2, 2001, pp. 108–115.
[10] R.M. Haralick, C.N. Lee, K. Ottenberg, M. Nolle, Review and analysis of solutions of the three point perspective pose estimation problem, International Journal of Computer Vision 13 (3) (1994) 331–356.
[11] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, second ed., Cambridge, 2003.
[12] J. Heikkila, Geometric camera calibration using circular control points, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (10) (2000) 1066–1077.
[13] J. Kannala, S.S. Brandt, A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (8) (2006) 1335–1340.
[14] C. Lu, G.D. Hager, E. Mjolsness, Fast and globally convergent pose estimation from video images, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (6) (2000) 610–622.
[15] J. Mallon, P.F. Whelan, Which pattern? Biasing aspects of planar calibration patterns and detection methods, Pattern Recognition Letters 28 (8) (2007) 921–930.
[16] B. Micusik, T. Pajdla, Estimation of omnidirectional camera model from epipolar geometry, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Wisconsin, USA, vol. 1, 2003, pp. 485–490.
[17] B. Micusik, T. Pajdla, Structure from motion with wide circular field of view cameras, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (7) (2006) 1135–1149.
[18] D. Nister. An efficient solution to the five-point relative pose problem, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Wisconsin, USA, vol. II, 2003, pp. 195–202.
[19] S. Ramalingam, S.K. Lodha, P. Sturm, A generic structure-from-motion framework, Computer Vision and Image Understanding 103 (3) (2006) 218–228.
[20] S. Ramalingam, P. Sturm, Minimal solutions for generic imaging models, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Alaska, USA, June 2008, pp. 1–8.
[21] S. Ramalingam, P. Sturm, S. Lodha. Theory and experiments towards complete generic calibration. Rapport de Recherche 5562, INRIA, 2005.
[22] R. Sagawa, M. Takatsuji, T. Echigo, Y. Yagi. Calibration of lens distortion by structured-light scanning, in: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, Canada, 2005, pp. 1349–1354.
[23] P. Sturm. Algorithms for plane-based pose estimation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, South Carolina, USA, 2000, pp. 1010–1017.
[24] P. Sturm, S. Ramalingam. A generic calibration concept: theory and algorithms. Rapport de Recherche 5058, INRIA, 2003.
[25] P. Sturm, S. Ramalingam. A generic concept for camera calibration, in: Proceedings of the 8th European Conference on Computer Vision, Prague, Czech Republic, vol. 2, 2004, pp. 1–13.
[26] P.F. Sturm, S.J. Maybank. On plane-based camera calibration: a general algorithm, singularities, applications, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Colorado, USA, vol. 1, 1999, pp. 432–437.
[27] J. Tardif, P. Sturm, S. Roy, Self-calibration of a general radially symmetric distortion model, in: Proceedings of the 9th European Conference on Computer Vision, Graz, Austria, Lecture Notes in Computer Science, vol. 4, Springer, 2006, pp. 186–199. May.
[28] S. Thirthala, M. Pollefeys. Multi-view geometry of 1d radial cameras and its application to omnidirectional camera calibration, in: Proceedings of the 10th IEEE International Conference on Computer Vision, Beijing, China, 2005, pp. 1539–1546.
[29] J. Wang, Y. Liu. Characteristic line of planar homography matrix and its applications in camera calibration, in: Proceedings of the 18th International Conference on Pattern Recognition, Hong Kong, China, 2006, pp. 147–150.
[30] G.Q. Wei, S.D. Ma, Implicit and explicit camera calibration: theory and experiments, IEEE Transactions on Pattern Analysis and Machine Intelligence 16 (5) (1994) 469–480.
[31] Z. Zhang. A Flexible New Technique for Camera Calibration. Technical Report, Microsoft Research, 1998.