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
aubrey.dunne@eeng.dcu.ie

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 with the goal of generality, and it is therefore suboptimal for the common case of cameras with a single centre of projection (e.g. pinhole, fisheye, hyperboloidal catadioptric). This paper proposes novel 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. Input data for the algorithm is acquired using active grids, the performance of which is characterised. A new linear estimation stage to the generic algorithm is proposed incorporating classical pinhole calibration techniques, and it is shown to be significantly more accurate than the linear estimation stage of the standard method. A linear method for pose estimation is also proposed and evaluated against the existing polynomial method. Distortion correction and motion reconstruction experiments are conducted with real data for a hyperboloidal catadioptric sensor for both the standard and proposed methods. Results show the accuracy and robustness of the proposed method to be superior to 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 a 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 [1, 2, 3, 4]. These methods are 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 and increasing number of parameters [5, 6] to allow wider angle lenses to be calibrated. However, they are not suitable for fisheye or catadioptric lenses for which the field of view exceeds 180°.

Several methods have been proposed that model wide-angle cameras as radially symmetric imagers [7, 8], thus simplifying the unknown parameter set. In [7], distortion is modelled using a varying focal length instead of an image displacement approach, allowing cameras with fields of view greater than 180° to be modelled. The complete class of single viewpoint catadioptric camera configurations was derived in [9], 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 [10] and non-central catadioptric [11]. The equivalence between catadioptric projections and mappings of the sphere was demonstrated in [12], 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 [7][13].

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 [14]. 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 [14] uses two images of a grid in different, known, positions. By determining the location 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 [15], wherein the world transforma-

tion 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 calibration process as proposed in [15, 16] can be summarised as follows: Three images of a calibration grid in different orientations are acquired, and for each camera pixel the location it sees on each grid is determined. The ray anchor points (camera centre for central cameras) and grid orientations are then linearly estimated using this location data, and refined via bundle adjustment. Finally, the calibration area is extended by imaging the grid in new orientations that intersect with the previously calibrated region, and using geometric constraints and bundle adjustment to estimate their pose. All the ray directions are stored in a lookup table as Plücker matrices. This calibration method will hereafter be referred to as the standard generic method.

The standard generic method is applicable to any camera geometry, and thus the calibration process is very general in order to cope with both central and non-central cameras. Consequently, many of the inherent geometric constraints of central cameras are not taken advantage of when calibrating these cameras. This paper proposes a new generic calibration method for cameras with a single centre of projection (hereafter referred to as the proposed generic method) that is optimum given the constraints of central cameras. The three key contributions in this paper are as follows:

Firstly, the issue of specifying accurate input data is addressed. The standard generic method [16] uses homographic interpolation with chessboard grid patterns. 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 [14][17][18], no discussion has been preferred on their performance for calibration purposes. An explanation of active grids is given in §2, along with an evaluation of their performance relative to standard localisation techniques. Secondly, a novel method for the linear estimation of the camera centre is proposed. The estimation of the camera centre in the standard generic method is "rather complicated" according to Sturm himself [15] and is given without any geometrical interpretation. In §3 active grids are shown to allow other, more simple and more accurate methods of determining the camera centre for central cameras. Thirdly, an alternative pose estimation stage is proposed. The pose estimation stage proposed by Sturm is a 3-point technique that does not lend itself well to large scale single shot pose estimation. The new pose estimation algorithm is derived and evaluated in §4. Together,
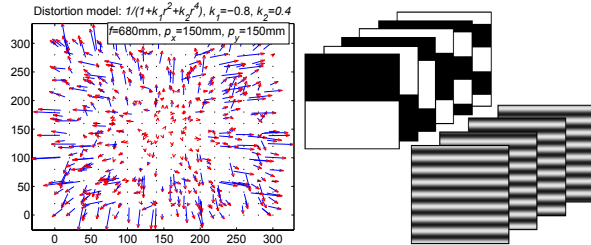

Figure 1. Vector plot of error residuals for homographic interpolation (20mm grid pitch) showing bias (left). Vectors are scaled ×20. Binary and sinusoidal active grid patterns for encoding vertical location (right).

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 §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 many 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 [15] 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 1 (left) shows a vector plot of the error residuals after homographic interpolation is applied to 500 random points on a $300mm \times 300mm$ 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 radial distortion.

Our 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 a flat-screen TFT monitor that is used to display a temporal sequence of spatially varying grayscale 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 in [17] for distortion correction. However, in that case the displayed patterns vary only temporally, resulting in difficulty resolving the boundaries of narrow stripes (the authors fall back on linear interpolation in these situations). In our method, each active grid requires 22 patterns to be consecutively displayed in order to fully encode the location data. Two patterns are used to determine thresholds, 12 patterns encode location on an 8×8 grid using Gray coded binary patterns, and 8 sinusoidal grayscale patterns encode location spatially within each square in this grid. The set of binary and sinusoidal patterns that encode vertical location are shown in Fig. 1 (right). Both the binary patterns and their inverses are displayed to make the decoding near white/black boundaries more robust. The phase of the sinusoidal patterns is shifted by 90° between consecutive patterns, and location is decoded from them as in [19].

Active grids overcome the distortion bias associated with homographic interpolation and consequently are ideal for use in the calibration process. The proposed generic method relies substantially on active grids for its accuracy, and therefore, since they have not previously been benchmarked against standard techniques for feature localisation, a performance evaluation of them is provided. 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 [20]. Consequently we used derivative corner localisation and saddle-point corner localisation techniques with a chessboard grid for the benchmarking process. The comparison between the active grids method and these two standard methods is shown in Fig 2. The experiments were conducted with real data by subpixally localising corners in the image of a chessboard grid, displayed on a TFT monitor, using the two standard methods. An active grid was then displayed on the monitor and decoded. The subpixel corner locations for the active grid method were determined by searching this decoded location data with the known metric grid dimensions. A second active grid was then placed in front of the camera, and the locations on this grid seen by the corner subpixels of each method were decoded directly. By mapping these locations to the known metric chessboard structure via homographies, the RMS error residuals for each method were determined. 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. Note that, where not otherwise specified, the camera-grid distance is 200mm, the orientation is 0°, and there is no blurring or additive noise. The excellent performance of the active grids in the comparison is partly due to their robustness to image sensor blooming in regions of high contrast: for active grids the highest resolution data is extracted from the sinusoidal pat-
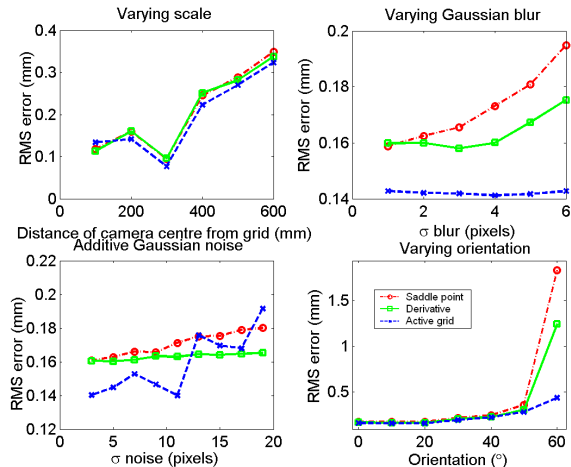

Figure 2. Performance plots for saddle point localisation, derivative localisation and active grids localisation. Orientation is measured between the grid normal and the camera axis in the horizontal plane.

terns, which contain only low contrast.

3. Linear Estimation

The purpose of the linear estimation stage in central generic calibration is to determine 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. The linear estimation stage of the standard 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 local homogeneous coordinate for each intersection point in a 4×4 matrix. Collinearity is enforced by ensuring the determinant of this matrix is zero. The algorithm for determining the camera centre and plane positions and orientations from this starting point for the standard generic method is "rather complicated" as stated by Sturm in [15]. See [15, 16] for a detailed explanation.

By taking a novel interpretation of existing methods for the calibration of pinhole cameras, an alternative, less complex, linear estimation stage is proposed. As known, pinhole calibration techniques are not suitable for wide field-of-view cameras due to the existence of severe non-linear image distortion that invalidates the linear projection model. The key idea of the proposed method is that an additional active grid is used as a synthetic image plane in the calibration, thus forming a synthetic pinhole camera. By placing the synthetic plane 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
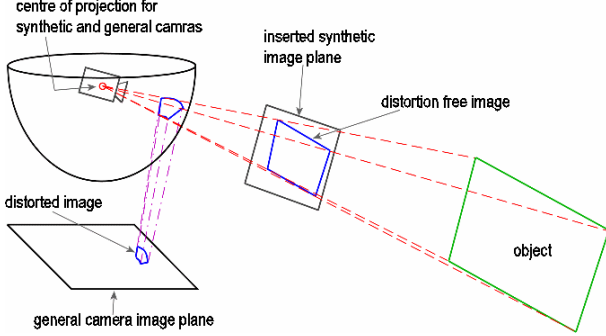
Figure 3. Linear estimation of camera centre for proposed generic calibration. Synthetic image plane allows use of pinhole calibration techniques for determining centre.
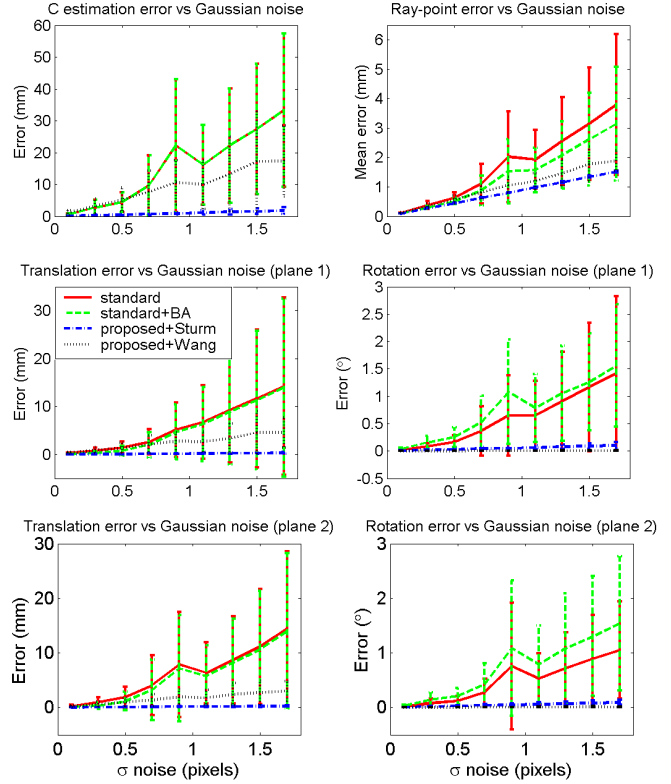

Figure 4. Centre estimation performance plots for standard generic method, standard generic method with bundle adjustment, and proposed generic method using both Sturm's and Wang's planar pinhole calibration techniques.

formed on the synthetic image plane. The synthetic camera can then be calibrated using any standard pinhole calibration method, with the desired estimate of the camera centre being $[p_x \ p_y \ f]^T$. The pose of grids two and three can also be extracted from the synthetic pinhole calibration using well known techniques [1][3]. In this way the non-linear calibration problem is converted to a linear calibration problem. This new approach provides a key link between the established theory of pinhole calibration and the generic calibration of central cameras, allowing the generic calibration of non-pinhole central cameras using pinhole calibration techniques. The number of grids required for the proposed generic method is three - two for the pinhole calibration plus one for the synthetic image plane - which is the same number as required for the standard generic method. Note that active grids are ideal for use as synthetic image planes in this method as they can directly provide the required ray-grid intersection points.

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 [1] and Wang [21], 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, but they take different approaches to determining the solutions. A comparison of the robustness to Gaussian noise of these two implementations of the proposed linear estimation stage, and of the standard generic method linear estimation stage, 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 calibration, are shown (averaged over 50 trials). The ray-point error is the perpendicular distance between each estimated ray and its known points of intersection with each calibration grid. These results are for a simulated camera with camera centre $[0 \ 0 \ 600]^T$ (in coordinate frame of first 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, as described in [16]. Bundle adjustment is not applied to the other two methods. The results clearly indicate that both of the implementations of the proposed linear estimation stage outperform that of the original generic method across all levels of noise tested. For the proposed generic method, the implementation using Sturm's calibration outperforms that using Wang's calibration in all cases, and therefore Sturm's pinhole technique is used in the proposed generic method. It is interesting to note that bundle adjustment does not appear to significantly improve the calibration result for the standard generic method. This is likely due to the error in the linear estimation of the camera centre (centre estimate is not bundle adjusted).

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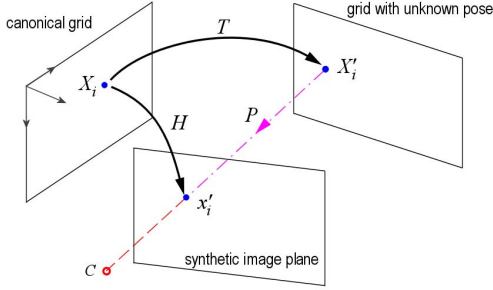
Figure 5. Proposed linear pose estimation method using synthetic image plane.

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 [22]. However, closed form solutions to the pose estimation problem for more than four points do not exist [23]. The most common approach to pose estimation in these cases is to minimise either the image space error or the object space error using standard nonlinear minimisation techniques. An iterative technique [23] has also been proposed.

In the standard generic method [15] a geometric three point algorithm for estimating the pose is described. 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 can be computed by simultaneously solving $|\lambda_i R_i - \lambda_j R_j|^2 = d_{ij}^2$ for $i, j = (l, m, n)$, $i \neq j$. Additional points are used to determine the correct pose from the eight possible solutions. A significant disadvantage of the algorithm is that, when included in a RANSAC framework, a linear re-estimation of the pose based on all inliers is not possible (typically the final step in RANSAC). The method is also very sensitive to additive noise (although a guided selection of sufficiently separated points can alleviate this problem), and is computationally expensive.

To overcome these difficulties a linear least-squares solution to the pose estimation problem is proposed. 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 is derived based on standard camera geometry and the pinhole camera model, similarly to [24]. With reference to Fig. 5, given a grid in the base position (canonical grid) with world coordinate points $X_i \in {}^3$, and a grid with an unknown pose $T \in {}^3$ relative to the canonical grid containing unknown world points $X'_i \in {}^3$, we wish to determine the unknown pose T. This is achieved via the insertion of a synthetic image plane placed in a known orientation (the orientation selection is discussed later) between the camera centre and the canonical grid, as shown in Fig. 5. Note that this synthetic image plane is a mathematical construct

only and is not physically realised. The key to the proposed approach is that the synthetic image plane allows the general pose estimation problem to be converted to a pinhole pose estimation problem. Points X_i can be projected onto the synthetic image plane by intersecting the previously calibrated rays with this plane. The projection is according to the pinhole model

$$x'_i = P X'_i \tag{1}$$

where $x_i \in {}^2$ are the imaged points on the synthetic image plane and P is the known 3×4 camera projection matrix associated with the synthetic image plane. Since the canonical grid is on the world $Z = 0$ plane, we also have

$$x'_i = H X_{i[1\ 2\ 4]]} \tag{2}$$

and

$$X'_i = T X_i \tag{3}$$

where H is a homography in ^{2}. Therefore

$$
\begin{aligned}
H X_{i[1\ 2\ 4]]} &= P T_{(:,[1\ 2\ 4])} X_{i[1\ 2\ 4]]} \\
&= K R [I_{(3\times3)} | -\tilde{C}] T_{(:,[1\ 2\ 4])} X_{i[1\ 2\ 4]} \tag{4}
\end{aligned}
$$

where $\tilde{C} \in {}^3$ is the inhomogeneous coordinate of the camera centre, K is the 3×3 camera calibration matrix for the synthetic image plane, and $R \in SO(3)$ is the world rotation of the synthetic image plane. Therefore

$$(KR)^{-1} H = \begin{pmatrix} T_{11} & T_{12} & T_{14} - \tilde{C}_1 \\ T_{21} & T_{22} & T_{24} - \tilde{C}_2 \\ T_{31} & T_{32} & T_{34} - \tilde{C}_3 \end{pmatrix} \tag{5}$$

Letting

$$G = (KR)^{-1} H \tag{6}$$

gives

$$sR' = \begin{pmatrix} G_1 & G_2 & G_1 \times G_2 \end{pmatrix} \tag{7}$$

where G_i is the ith column of G. An orthonormal R' is obtained via the SVD. The scale factor, s, can be obtained as

$$s = mean \frac{\begin{pmatrix} G_1 & G_2 \end{pmatrix}}{\begin{pmatrix} \hat{R}_1 & \hat{R}_2 \end{pmatrix}} \tag{8}$$

The translation is

$$\hat{t} = \frac{G_4 + \tilde{C}}{s} \tag{9}$$

The desired pose estimate is then

$$\hat{T} = \begin{pmatrix} \hat{R} & \hat{t} \\ 0_{(1\times3)} & 1 \end{pmatrix} \tag{10}$$

The orientation of the synthetic image plane should be chosen so as to be as perpendicular as possible to the known rays involved in the estimation process. This orientation is determined in a least-squares sense by minimising the sum
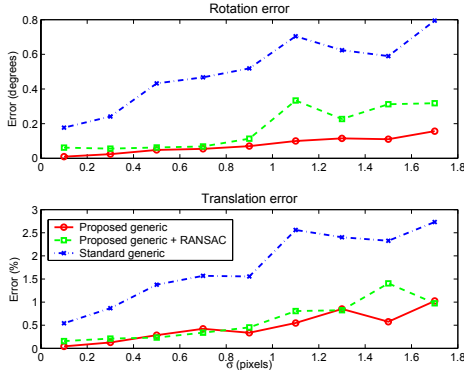
Figure 6. Performance comparison of pose estimation stage of proposed generic method, proposed generic method with RANSAC, and standard generic method.

Table 1. Camera centre and grid transformation estimates for catadioptric sensor calibration. Centre and translations are measured in mm, rotations are measured in degrees.

Method	Centre	R1	T1	R2	T2
Original method	169.29 152.06 -106.48	34.65 -141.74	130.58	35.27 -13.35	182.84
Proposed method	167.90 159.82 -116.21	36.52 -142.08	139.66	35.61 -13.70	192.13

Table 2. Ray-point errors (mm) for all rays involved in the linear estimation stage for each calibration method (BA = bundle adjustment).

Method	Error type	Error	Error after BA
Original method	Mean	3.2219	1.1634
	SD	1.4923	0.6583
Proposed method	Mean	0.1924	0.1314
	SD	0.0906	0.0727

of the angles between the calibrated rays and the synthetic image plane, as in [25].

The robustness to Gaussian noise of the standard generic method's pose estimation stage (embedded in RANSAC, followed by non-linear optimisation), and the pose estimation stage of the proposed generic method, both on it's own and embedded in a RANSAC framework, is evaluated for simulated data. The simulated camera centre is fixed at $[0\ 0\ 600]^T$, and the translations and Euler rotations of the grid whose pose is to be estimated are randomly chosen from $[-150mm\ 150mm]$ and $[-30^o\ 30^o]$ respectively. The mean rotational and mean percentage translational errors are shown in Fig. 6. It is seen that the proposed generic method's pose estimation consistently outperforms that of the standard generic method for all simulated levels of noise. Also, embedding the linear method in a RANSAC phase actually reduces the accuracy of the estimates. This is possibly due to the non-isotropic nature of the noise that results from projecting Gaussian noise from the image plane onto a non-parallel plane.

5. Experimental Results

Both the standard and proposed generic methods are analysed for real data with respect to a ray-point error metric, distortion correction, and separate motion reconstruction tasks. All images for these experiments were taken using a 360 OneVR hyperboloidal omnidirectional mirror [1] in conjunction with a Nikon D70 SLR digital camera. This catadioptric configuration has a single centre of projection. For each calibration method approximately 207° of the horizontal FOV and approximately 82° of the vertical FOV of the camera was calibrated, using three grids for the linear estimation of the camera centre, and a further three grids to extend the calibrated region to include additional pixels. Active grids were used for all grids during calibra-

tion, with the same images used as input to both 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 (RANSAC thresholds empirically chosen with reference to Fig. 2).

The first three grids used in the calibration were found to have 41467 common intersecting rays, of which 31398 were determined to be inliers. Table 1 shows, for each calibration method, the estimates of the camera centre, and the estimated YZ yaw-roll rotation angles and translation magnitudes for the second and third grids. Note the significant difference between the estimated values of the camera centre z coordinate for each method.

The ray-point error metric, described in §3, can be applied to each calibration dataset to give an indication of the relative errors in each calibration (actual positions of the camera centre and the second and third grids are not known). Table 2 shows the mean and standard deviation of the ray-point errors for each method, both before and after bundle adjustment. Bundle adjustment is applied to the proposed method here for comparative purposes only. The non-bundle-adjusted parameters are used in the remainder of the calibration with the proposed generic method, whereas the bundle adjusted results are used for calibration with the standard generic method (as per [26]). The ray-point error results clearly show that the configuration of the camera centre and the calibration grids is in greater geometric agreement for the proposed generic method than for the standard generic method. These results also agree with the simulated results in §3, specifically they show that the linear estimates of the proposed generic method are capable of
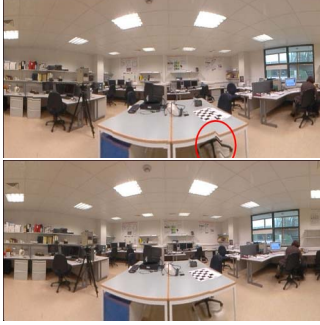
Figure 7. Cylindrically unwarped catadioptric image after standard generic calibration (top) and proposed generic calibration (bottom)
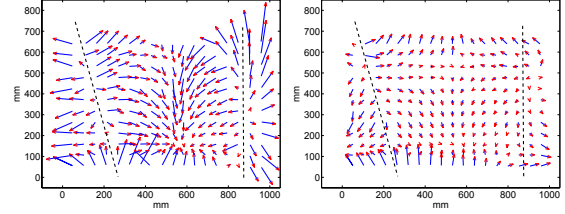

Figure 8. Vector plots of residuals after distortion correction of a chessboard grid for standard generic method (left) and proposed generic method (right). Vectors are scaled ×15

outperforming the bundle adjusted estimates of the standard generic 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 is used to remove the inherent non-linear distortion from the calibrated area of an omnidirectional image of a real scene. A portion of a cylindrical image is formed by intersecting the calibrated rays with a unit cylinder, the axis of which is coincident with the camera centre, and the cylinder is then unwrapped to form a planar image. Fig 7 shows the cylindrically unwarped images calculated using the calibration data from the standard generic and proposed generic methods. As expected, real world straight lines that are parallel to the mirror axis (vertical) are mapped to straight lines in both of the corrected images. However, some abberation is visible in the image corrected using the standard generic method calibration data (highlighted by ellipse). In contrast, the corrected image formed using the proposed generic method has significantly less aberration. Note that the field of view of the cylindrical unwarped image for the standard generic method is less than that of the proposed generic method due to the smaller estimate for the z coordinate of the camera centre using the standard generic method.

Quantitative evaluation of the calibrations was carried out by generating a perspectively corrected image of an 18×12 chessboard calibration grid, with square side length 53mm. The plane onto which the corrected images are projected is selected as described in §4. Distortion residuals are measured after applying a homography between the distortion corrected image of the grid and the known metric grid structure. Fig. 8 shows the distortion residuals for both the standard and proposed methods. No radial distortion bias is visible in either vector plot, but the plot for the standard generic method displays large divergences along roughly vertical lines at the left and right of the image. These

correspond to areas where two active grids with misestimated poses meet, and correspond to the aberrations seen in Fig. 7 (top). The divergences and residuals are smaller for the vector plot using the proposed generic method (mean RMS error = 2.23mm, STD = 1.06mm) than for the vector plot using the standard generic method (mean RMS error = 4.54mm, STD = 1.96mm) indicating a better calibration.

5.2. Motion Reconstruction

Motion reconstruction experiments were conducted for the cases of pure translation and pure rotation. 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 100mm in steps of 20mm, and for the rotation stage it was rotated by 90° in steps of 22.5^o. Point matches were manually extracted across both image sequences, and used to index the Plücker matrix lookup tables for each calibration method to get the corresponding ray direction information. The essential matrix, $\mathbf{E}$, between each image pair was linearly estimated using the ray-based epipolar constraint $\mathbf{L}'\mathbf{EL} = \mathbf{0}$, where $\mathbf{L}, \mathbf{L}'$ are the first 3 components of the Plücker vectors derived from the Plücker matrices [25]. Rotations and translations are extracted from the essential matrices according to [27]. The motion reconstruction results are shown in Fig.9. It can be seen that the motion estimated with the proposed generic method is closer to linear in the case of translation, and closer to 90° 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 ×4.

6. Summary and Conclusions

This paper proposes a novel method of generic camera calibration for cameras with a single centre of projection. The main contributions of the paper are a performance evaluation of active grids for use in calibration, an improved linear estimation stage based on a new interpretation of an existing technique that allows pinhole calibration techniques to be applied to the calibration of non-pinhole cameras, and
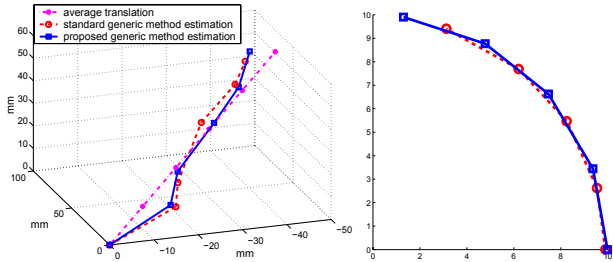
Figure 9. Translation (left) and rotation (right) reconstruction using calibration data from standard generic method and proposed generic method.

a new linear pose estimation stage. The individual components of the proposed method are separately evaluated using simulated data, with the results indicating that the proposed generic method outperforms the standard generic calibration method in terms of accuracy and robustness to noise. The proposed generic method is also evaluated against the standard generic method for real data using a hyperboloidal omnidirectional camera, with the results for distortion correction and motion reconstruction tasks showing the improved performance of the proposed generic method.

Acknowledgements

References

[1] P. Sturm and S. J. Maybank, "On plane-based camera calibration: A general algorithm, singularities, applications," in *CVPR*, vol. 1, 1999, pp. 432–437. 1, 4

[2] A. W. Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," in *CVPR*, vol. 1, 2001, pp. 125–132. 1

[3] Z. Zhang, "A flexible new technique for camera calibration," Microsoft research, Tech. Rep., 1998. 1, 4

[4] J. Heikkila, "Geometric camera calibration using circular control points," *PAMI*, vol. 22, no. 10, pp. 1066–1077, 2000. 1

[5] J. Kannala and S. S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fisheye lenses," *PAMI*, vol. 28, no. 8, pp. 1335–1340, 2006. 1

[6] G. Q. Wei and S. D. Ma, "Implicit and explicit camera calibration: theory and experiments," *PAMI*, vol. 16, no. 5, pp. 469–480, 1994. 1

[7] J. Tardif, P. Sturm, and S. Roy, "Self-calibration of a general radially symmetric distortion model," in *ECCV*, vol. 4, May 2006, pp. 186–199. 1

[8] S. Thirthala and M. Pollefeys, "Multi-view geometry of 1d radial cameras and its application to omnidirectional camera calibration," in *ICCV*, 2005, pp. 1539–1546. 1

[9] S. Baker and S. Nayar, "A theory of catadioptric image formation," in *ICCV*, 1998, pp. 35–42. 1

[10] B. Micusik and T. Pajdla, "Estimation of omnidirectional camera model from epipolar geometry," in *CVPR*, vol. 1, 2003, pp. 485–490. 1

[11] ——, "Structure from motion with wide circular field of view cameras," *PAMI*, vol. 28, no. 7, pp. 1135–1149, 2006. 1

[12] C. Geyer and K. Daniilidis, "An unifying theory for central panoramic systems and pratical implications," in *ECCV*, 2000, pp. 445–461. 1

[13] J. P. Barreto, "A unifying geometric representation for central projection systems," *CVIU*, vol. 103, no. 3, pp. 208–217, 2006. 1

[14] M. D. Grossberg and S. K. Nayar, "A general imaging model and a method for finding its parameters," in *ICCV*, vol. 2, 2001, pp. 108–115. 1, 2

[15] P. Sturm and S. Ramalingam, "A generic calibration concept: Theory and algorithms," Rapport de Recherche 5058, 2003. 1, 2, 3, 5

[16] ——, "A generic concept for camera calibration," in *ECCV*, vol. 2, 2004, pp. 1–13. 2, 3, 4

[17] R. Sagawa, M. Takatsuji, T. Echigo, and Y. Yagi, "Calibration of lens distortion by structured-light scanning," in *ICIRS*, 2005, pp. 1349–1354. 2, 3

[18] A. K. Dunne, J. Mallon, and P. F. Whelan, "A comparison of new generic camera calibration with the standard parametric approach," in *MVA*, vol. 1, 2007, pp. 114–117. 2

[19] J. Salvi, J. Pages, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognition*, vol. 37, no. 4, pp. 827–849, April 2004. 3

[20] J. Mallon and P. F. Whelan, "Which pattern? biasing aspects of planar calibration patterns and detection methods," *PRL*, vol. 28, no. 8, pp. 921–930, 2007. 3

[21] J. Wang and Y. Liu, "Characteristic line of planar homography matrix and its applications in camera calibration," in *ICPR*, 2006, pp. 147–150. 4

[22] R. M. Haralick, C. Lee, K. Ottenberg, and M. Nolle, "Analysis and solutions of the three point perspective pose estimation problem," Hamburg, Tech. Rep., 1991. 5

[23] C. Lu, G. D. Hager, and E. Mjolsness, "Fast and globally convergent pose estimation from video images," *PAMI*, vol. 22, no. 6, pp. 610–622, 2000. 5

[24] P. Sturm, "Algorithms for plane-based pose estimation," in *CVPR*, 2000, pp. 1010–1017. 5

[25] S. Ramalingam, S. K. Lodha, and P. Sturm, "A generic structure-from-motion framework," *CVIU*, vol. 103, no. 3, pp. 218–228, 2006. 6, 7

[26] S. Ramalingam, P. Sturm, and S. Lodha, "Theory and experiments towards complete generic calibration," Rapport de Recherche 5562, 2005. 6

[27] D. Nistér, "An efficient solution to the five-point relative pose problem," *PAMI*, vol. 26, no. 6, pp. 756–777, 2004. 7