Extrinsic Camera Calibration Without A Direct View Using Spherical Mirror

Amit Agrawal
Mitsubishi Electric Research Labs (MERL)
201 Broadway, Cambridge, MA, USA
agrawal at merl dot com

Abstract

We consider the problem of estimating the extrinsic parameters (pose) of a camera with respect to a reference 3D object without a direct view. Since the camera does not view the object directly, previous approaches have utilized reflections in a planar mirror to solve this problem. However, a planar mirror based approach requires a minimum of three reflections and has degenerate configurations where estimation fails. In this paper, we show that the pose can be obtained using a single reflection in a spherical mirror of known radius. This makes our approach simpler and easier in practice. In addition, unlike planar mirrors, the spherical mirror based approach does not have any degenerate configurations, leading to a robust algorithm.

While a planar mirror reflection results in a virtual perspective camera, a spherical mirror reflection results in a non-perspective axial camera. The axial nature of rays allows us to compute the axis (direction of sphere center) and few pose parameters in a linear fashion. We then derive an analytical solution to obtain the distance to the sphere center and remaining pose parameters and show that it corresponds to solving a 16^{th} degree equation. We present comparisons with state-of-art method using planar mirrors and show that our approach recovers more accurate pose in the presence of noise. Extensive simulations and results on real data validate our algorithm.

1. Introduction

The external calibration of a camera with a reference 3D object is a fundamental pose estimation procedure in computer vision. If 2D projections of three 3D points are visible, the extrinsic parameters (or pose), namely rotation and translation can be obtained from the three 2D-3D correspondences. This is known as the P3P problem [11, 21, 12] and is a basic building block in several applications such as 3D reconstruction, navigation, localization, etc. In this paper, we consider the case where the camera does not view the 3D points *directly*. This is a natural setting when considering multiple cameras with non-overlapping field of view (FOV) [18], calibrating a camera with a display for human-computer interaction [25] or when reconstructing specular/shiny objects requiring a calibrated camera and reference 3D points.

Previous approaches to solve this problem have used planar mirrors to observe the reflections of the reference object. The extrinsic parameters along with the location and orientation of planar mirrors can be obtained using the reflections. This is referred to as *mirror-based extrinsic calibration*.

As shown in Sturm and Bonfort [28], at least three reflections of the reference object are required to uniquely determine the extrinsic parameters using a planar mirror. The three reflections should be obtained using different orientations of the mirror. Even when using three views, there exist degenerate/singular configurations. The singular configuration occurs if all the mirror planes intersect in a single line. This happens if the mirror is rotated around a fixed axis, or if all mirror planes are parallel (intersection line is at infinity) [27]. Moreover, a particular algorithm can have specific degenerate cases depending on its implementation. For example, consider the state-of-art algorithm proposed in [31]. If the intersection of any two mirrors lie on the plane defined by the reference 3D points, it results in a degenerate configuration. In practice, a configuration close to a singular configuration will be unstable and will not be robust in the presence of noise. Previous approaches either assume that the mirror configuration is non-singular or use heuristics to detect the presence of such configurations. If a singular configuration is detected, more images need to be captured.

Thus, we believe that previous methods to solve this problem are lacking in several respects. In this paper, we overcome the limitations of the previous methods. We show that the mirror-based extrinsic calibration can be done using a *single* reflection from a spherical mirror with known radius. The location of the spherical mirror is not required to

Method	Mirror Shape	Mirrors	3D Points	Total 2D projections	Degenerate Mirror Configurations
Sturm and Bonfort [28]	Planar	3	4	12	Yes
Rodorigues *et al*. [27]	Planar	3	4	12	Yes
Kumar *et al*. [18]	Planar	5	4	20	Yes
Hesch *et al*. [14]	Planar	3	3	9	Yes
Takahashi *et al*. [31]	Planar	3	3	9	Yes
Ours	Spherical	1	8	8	No

Table 1. Ours is the first approach that uses a single reflection from a spherical mirror for mirror-based extrinsic calibration. Unlike previous approaches, our approach does not have any degenerate mirror configurations. In addition, we show in Section 6 that it outperforms the state-of-art method [31].

be known and is estimated simultaneously with the extrinsic parameters. Our approach does *not* require the boundary of the sphere to be visible in the image, which is commonly used to estimate the sphere center [16, 6, 7] and only uses the projection of reference 3D points. This makes our approach simple and practical to use. Moreover, no degenerate mirror configurations exist when using a spherical mirror, since we require only one view. In addition, we show that our approach outperforms the state-of-art method using planar mirrors [31] and recovers more accurate pose in the presence of noise. Our contributions are as follows

1.1. Contributions

- We demonstrate that mirror based extrinsic calibration can be done from a single reflection using a non-planar spherical mirror.
- We derive analytical solutions for estimating the spherical mirror location and pose of the reference object using a single reflection.

2. Related Work

Mirror based extrinsic calibration: Previous approaches either attach markers to planar mirrors to estimate the mirror poses [15, 20] or estimate the mirror poses along with the extrinsic calibration [28, 27, 18, 14, 31]. As discussed, a minimum of three reflections are required using planar mirror and degenerate configurations exists. Table 1 compares our approach with previous approaches.

Catadioptric Cameras: Both planar and spherical mirrors have been used with perspective camera as a catadioptric imaging system. Mirror reflections of a scene allow capture of multiple viewpoints using a single camera. This is referred to as planar catadioptric stereo [23, 10, 32, 9]. The epipolar geometry in this case is well-studied in literature and such systems have been used for 3D reconstruction. Recently, Reshetouski *et al*. [26] used three planar mirrors to obtain hundreds of views of an object for visual hull reconstruction.

Spherical Mirrors: Spherical mirror based catadioptric cameras have also been used for wide-angle 3D reconstruction [22, 5, 3, 16, 19], for obtaining wide-angle light fields [29] and navigation [17]. These approaches typically attach multiple spherical mirrors on a planar surface along with markers [5, 19] and use the markers to estimate the initial location of spheres or utilize the sphere boundary (contour) in the image. For example, Francken *et al*. [6] require several images of a moving sphere ($\approx$ 10) along with the sphere boundary to be visible in all images for extrinsic calibration. Similarly, Nitschkea *et al*. [25] used multiple eye reflections along with the cornea contour (limbus) information for display-camera calibration. In contrast, our approach does not require the sphere boundary to be visible and uses a single reflection.

Perhaps the closest work to ours is the recent work of Agrawal *et al*. [1], where multiple spherical mirrors based catadioptric system is calibrated using a planar checkerboard. The pose of the checkerboard is also recovered during calibration. Thus, calibration of such systems is similar to our problem of extrinsic mirror based calibration. The approach in [1] utilizes the fact that by using rays from two or more spherical mirrors, the pose can be obtained using a linear algorithm. For a single mirror, the linear system of [1] is degenerate and thus cannot be applied to our case.

Thus, if we use two spherical mirrors (in a single view) or take two views of a spherical mirror, we can apply the algorithm of [1] to perform extrinsic calibration. However, in this paper our goal is to achieve it using a single view of a single spherical mirror. For a spherical mirror, using two views results in a degenerate configuration when the sphere location is same in both views. In practice, if the sphere locations are close by, the pose estimate will not be robust in presence of noise. Thus, to completely avoid degenerate mirror configurations, we need a solution involving single view/mirror. In addition, [1] clearly states that they were unable to find an analytical solution for calibration using a single spherical mirror, which is our key contribution.

3. Extrinsic Camera Calibration

We consider a perspective camera and a reference 3D object which is not viewed directly by the camera. We work in the camera coordinate system with its center of projection

(COP) at origin. The camera is assumed to be internally calibrated off-line using standard approaches. Let $\mathbf{P}(i)_{i=1}^{K}$ denote K 3D points which are known in the reference object coordinate system. Our goal is to estimate the unknown pose (rotation R and translation $\mathbf{t}$) of these points in the camera coordinate system.

To perform the calibration, we place a spherical mirror of radius r at an unknown location $\mathbf{C}$ in-front of the camera and observe the reflection of the 3D points. Let $\mathbf{p}(i)_{i=1}^{K}$ be the 2D projection of the 3D points after reflection. Since the camera is calibrated, we know the camera rays $\mathbf{v}(i)_{i=1}^{K}$ corresponding to the projected 2D points. The goal is to then recover the desired pose $[R, \mathbf{t}]$ along with the unknown sphere location $\mathbf{C}$ using K 3D-2D correspondences.

Axial Geometry: It is well-known that a perspective camera looking into a spherical mirror corresponds to an axial imaging system [3]. Let $d = \parallel C \parallel$ be the distance of the sphere center from the COP. The axis A is defined as $\mathbf{A} = \mathbf{C}/d$ and corresponds to the normalized vector joining the COP and the sphere center.

4. Estimation of Axis

As described in [2, 1], the so called coplanarity constraint holds for axial systems, which allow estimation of axis using 8 3D-2D correspondences. In addition, the coplanarity constraints can also recover the rotation R and translation orthogonal to the axis (two degrees of translation). The translation along the axis cannot be recovered by the coplanarity constraints. For completeness, we re-derive the constraint below in this section and the derivation follows [2, 1]. In the next section, we derive an analytical solution for remaining parameters, namely, distance d to the sphere center and translation along the axis, which is a key contribution of this paper.

Let π be the plane of reflection given by the axis and a camera ray $\mathbf{v}(i)$. From Snell's law, the reflected ray should lie to π. Thus, the transformed 3D point $R\mathbf{P}(i) + \mathbf{t}$ should also lie on π. Therefore, the 3-vectors $\mathbf{A}$, $\mathbf{v}(i)$ and $R\mathbf{P}(i)+\mathbf{t}$ are co-planar, resulting in following co-planarity constraint

$$\mathbf{v}(i)^T(\mathbf{A} \times (R\mathbf{P}(i) + \mathbf{t})) = 0, \tag{1}$$

where $\times$ denotes cross-product. Let $[\mathbf{A}]_\times$ be the 3×3 skew-symmetric matrix obtained from the 3-vector $\mathbf{A}$. The above equation can be written as

$$\mathbf{v}(i)^T E\mathbf{P}(i) + \mathbf{v}(i)^T\mathbf{s} = 0, \tag{2}$$

where $E = [\mathbf{A}]_\times R$ and $\mathbf{s} = \mathbf{A} \times \mathbf{t}$. Since $\mathbf{A}^T\mathbf{s} = 0$, the component of translation $\mathbf{t}$ along the axis, $\mathbf{t}_A$, cannot be recovered using the co-planarity constraints. Notice that the form of the E matrix is similar to the essential matrix for relative pose estimation between two perspective views [13].

Using 8 correspondences, one can write a linear system as

$$\underbrace{\begin{bmatrix} (\mathbf{P}(1)^T \otimes \mathbf{v}(1)^T) & \mathbf{v}(1)^T \\ \vdots & \vdots \\ (\mathbf{P}(8)^T \otimes \mathbf{v}(8)^T) & \mathbf{v}(8)^T \end{bmatrix}}_{\mathbf{B}} \begin{bmatrix} E(:) \\ \mathbf{s} \end{bmatrix} = 0, \tag{3}$$

where $\otimes$ denotes Kronecker product and B is an 8×12 matrix.

Non-Planar Reference 3D points: If the reference 3D points $\mathbf{P}(i)$ do not lie on a plane, the rank of B is 8. As described in [2] (8-point algorithm), a four dimensional subspace solution can be obtained by SVD. The subspace coefficients are then obtained by using Nister's 5-point algorithm [24]. This results in solution(s) for the E matrix and the $\mathbf{s}$ vector. For each E matrix, four rotation matrices can be obtained [13].

Planar Reference 3D points: Suppose the reference 3D points $\mathbf{P}(i)$ lie on a plane. This is a common scenario when a checkerboard is used for calibration. Without loss of generality, we can align the checkerboard with the xy plane and set the z^{th} component of $\mathbf{P}(i)$, $\mathbf{P}^z(i)$, to zero. In that case, the columns 7, 8 and 9 of B goes to zero. As described in [1], let $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ be the columns of E. Removing the zero columns, the above linear system can be written as

$$B_r \begin{bmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \mathbf{s} \end{bmatrix} = 0, \tag{4}$$

where B_r is the reduced 8×9 matrix of rank 8. Let $B_r = U\Sigma V^T$ be the SVD of B_r. The standard SVD based solution is given by the last column of V. Thus, $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{s}$ can be computed using 8 correspondences. Since $\mathbf{A}^T E = 0$, $\mathbf{A}$ is orthogonal to $\mathbf{e}_1$ and $\mathbf{e}_2$. The axis A is then given by

$$\mathbf{A} = (\mathbf{e}_1 \times \mathbf{e}_2)/ \parallel (\mathbf{e}_1 \times \mathbf{e}_2) \parallel . \tag{5}$$

There are various ways to recover the rotation matrix R from only the first two columns of the E matrix. One way is described in [2], where the last column of E matrix, $\mathbf{e}_3$, is obtained using Demazure constraints [4]. As shown in [2], there are two solutions for $\mathbf{e}_3$, which only differ in sign. Thus, the solutions for E are given by $\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \end{bmatrix}$ and $\begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & -\mathbf{e}_3 \end{bmatrix}$. Since each E matrix corresponds to four rotation matrices, a total of eight rotation matrices are obtained. However, only four of these are proper rotation matrices. The other four corresponds to reflection matrices and have determinant -1 instead of 1. These can be easily discarded by checking the determinant value.

To summarize, for planar reference 3D points, we obtain a solution for axis and four solutions for rotation R and vector $\mathbf{s}$. In the next section, we describe an analytical solution to obtain d and $\mathbf{t}_A$ for a given solution of R and $\mathbf{s}$.

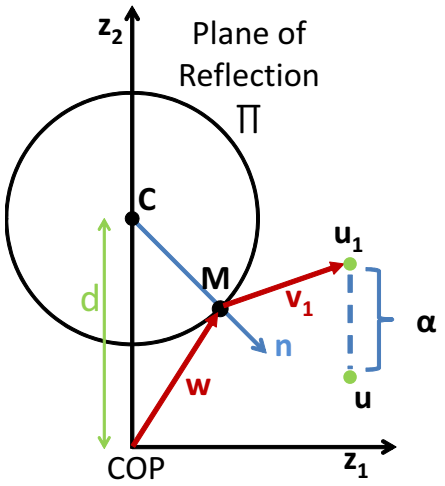

Figure 1. Estimation of unknown d and α on the plane of reflection π.

5. Analytical solution for d and $\mathbf{t}_A$

In this section, we describe an analytical solution to obtain the sphere distance d and translation along axis $\mathbf{t}_A$ assuming known axis, rotation R and vector $\mathbf{s}$. We again emphasise that in [1], authors state that finding an analytical solution for a single mirror is extremely difficult and they were unable to find such a solution. In [1], the full pose was recovered linearly using rays from two or more spheres. In contrast, since we assume a single view (single spherical mirror), the linear pose estimation algorithm of [1] cannot be applied.

We first transform the reference 3D points using known pose parameters. The translation orthogonal to the axis $\mathbf{A}$, $\mathbf{t}_{A^\perp}$ is obtained as $\mathbf{s} \times \mathbf{A}$. Let $\mathbf{Q}(i) = R\mathbf{P}(i) + \mathbf{t}_{A^\perp}$ be the transformed points. Let $\mathbf{t}_A = \alpha A$ be the unknown translation along axis, where α is a scalar.

When the axis is known, the analysis can be done on the plane of reflection π (Figure 1). Let $[\mathbf{z}_1, \mathbf{z}_2]$ denote an ortho-normal coordinate system on π, where $\mathbf{z}_2$ is along the axis. For a given camera ray $\mathbf{v}(i)$, let $\mathbf{z}_1 = \mathbf{z}_2 \times (\mathbf{z}_2 \times \mathbf{v}(i))$ be the orthogonal direction. The projection of $\mathbf{Q}(i)$ on π is given by $\mathbf{u} = [u_x, u_y]$, where $u_x = \mathbf{z}_1^T\mathbf{Q}(i)$ and $u_y = \mathbf{z}_2^T\mathbf{Q}(i)$. Since the translation along axis is unknown, the reflected camera ray should pass through $\mathbf{u}_1 = [u_x, u_y + \alpha]$. Similarly, the camera ray $\mathbf{v}(i)$ can be represented by the 2D vector $\mathbf{w} = [w_x, w_y]$ on π, where $w_x = \mathbf{z}_1^T\mathbf{v}(i)$ and $w_y = \mathbf{z}_2^T\mathbf{v}(i)$. $\mathbf{w}$ is normalized so that $\mathbf{w}^T\mathbf{w} = 1$.

The spherical mirror is represented as a circle on π with center at $\mathbf{C} = [0, d]$ and radius r. Let $\mathbf{M} = k\mathbf{w}$ be the common point on the mirror and the camera ray for some scalar k. Since $\mathbf{M}$ lies on the circle, $(\mathbf{M}-\mathbf{C})^T(\mathbf{M}-\mathbf{C}) = r^2$. Hence,

$$(kw_x)^2 + (kw_y - d)^2 = k^2 - 2kw_yd + d^2 = r^2. \quad (6)$$

The normal at $\mathbf{M}$ is given by $\mathbf{n} = \mathbf{M} - \mathbf{C}$. The reflected ray $\mathbf{v}_1$ is given by the Snell's law as $\mathbf{v}_1 = \mathbf{w} - 2\mathbf{n}(\mathbf{n}^T\mathbf{w})/r^2$. The reflected ray should pass through $\mathbf{u}_1$. Therefore,

$$0 = \mathbf{v}_1 \times (\mathbf{u}_1 - \mathbf{M}). \quad (7)$$

After substituting all the terms, the above equation simplifies to

$$K_1k^2 + K_2k + K_3 = 0, \quad (8)$$

where

$$\begin{aligned} K_1 &= 2w_x(d - \alpha - u_y) + 2u_xw_y, \\ K_2 &= 2d(\alpha w_xw_y - u_xw_y^2 - dw_xw_y \\ &\quad -u_x + u_yw_xw_y), \\ K_3 &= 2u_xw_yd^2 + r^2(\alpha w_x - u_xw_y + u_yw_x). \end{aligned}$$

By eliminating k between (6) and (8), we obtain a single equation in d and α, which is 6^{th} degree in d and quadratic in α. Since we have two unknowns (d and α), we need two equations from two correspondences. These can be written as

$$\begin{aligned} EQ_1: &\quad k_{11}\alpha^2 + k_{12}\alpha + k_{13} = 0, \\ EQ_2: &\quad k_{21}\alpha^2 + k_{22}\alpha + k_{23} = 0, \end{aligned}$$

where the coefficients k_{ij} depend on known values and the unknown parameter d. Eliminating α^2 between the equations, we get

$$\alpha = \frac{k_{11}k_{23} - k_{21}k_{13}}{k_{21}k_{12} - k_{11}k_{22}}. \quad (9)$$

Substituting α back in either EQ_1 or EQ_2 and simplifying, we obtain a 16^{th} degree equation[1] in the single unknown d. Solving this equation give 16 solutions for d. We remove imaginary solutions and enforce $d > r$ to get valid real solutions. Finally, the correct solution is obtained by checking with remaining $8 - 2 = 6$ correspondences. In particular, we choose the solution that minimizes the re-projection error over all eight correspondences. The translation along axis is then obtained as $\mathbf{t}_A = \alpha\mathbf{A}$.

Note that the axis computation requires 8 correspondences, while estimating d and $\mathbf{t}_A$ only requires two. In practice, one can choose any two correspondences out of 8 or test all possible ${}^8C_2 = 28$ choices.

5.1. Non-linear refinement

Once the initial estimate of pose and sphere location is obtained, the parameters can be refined by minimizing the image re-projection error for all K points. For a spherical mirror, the image projection of a 3D point can be obtained by solving a 4^{th} degree equation [3]. Let $\mathbf{x}$ denote the estimated pose and sphere location and let $\widehat{p}(i)$ be the the 2D

[1] Details are provided in the supplementary materials.

projection of $\mathbf{P}(i)$ using $\mathbf{x}$. We use `lsqnonlin` in Matlab to refine $\mathbf{x}$ by minimizing the root mean squared (RMS) re-projection error $J = \sqrt{\frac{1}{K}\sum_{i=1}^{K}(p(i) - \widehat{p}(i))^2}$. For our approach there are a total of 9 parameters (6 for pose and 3 for sphere center). For comparison with planar mirror based approach, we also implemented the corresponding non-linear refinement procedure, which involves $6 + 3 \times 3 = 15$ parameters (6 for pose and 3 for each mirror position). Our approach is summarized in Algorithm 1.

Input: Eight 2D-3D correspondences $\mathbf{v}(i) \leftrightarrow \mathbf{P}(i)$
Output: Sphere Location $\mathbf{C}$, Rotation R, Translation $\mathbf{t}$

Using eight correspondences, compute the axis $\mathbf{A}$ and solutions for rotation R and vector $\mathbf{s}$.
for *each solution of* R *and* $\mathbf{s}$ **do**
  Compute $\mathbf{t}_{A^\perp} = \mathbf{s} \times \mathbf{A}$.
  Choose two correspondences.
  Solve for d and $\mathbf{t}_A$. Remove invalid solutions.
  for *each valid solution of* d *and* $\mathbf{t}_A$ **do**
    Compute $\mathbf{C} = d\mathbf{A}$.
    Compute $\mathbf{t} = \mathbf{t}_A + \mathbf{t}_{A^\perp}$.
    Compute the RMS re-projection error using $\mathbf{C}$, R and $\mathbf{t}$ over all eight correspondences.
  end
end
Find solution that has minimum re-projection error over all eight correspondences.
Perform non-linear refinement to refine $\mathbf{C}$, R and $\mathbf{t}$.

Algorithm 1: Extrinsic calibration using spherical mirror.

6. Simulations

In this section, we present simulations on our approach along with comparisons with state-of-art planar mirror based approach [31]. We employ a realistic scenario using a camera with resolution of 1500×1500 pixels and focal length 2000 pixels (FOV of $41°$). A planar checkerboard with 9×6 squares, each of size 30 mm is used as a reference object. The ground truth rotation angles and translation of checkerboard in the camera coordinate system are $\theta = [2.2; -73.4; 5.3]$ degrees and $\mathbf{t} = [183.4; 134.6; 35]$ mm. We place a spherical mirror of radius 25.4 mm (1 inch) at the location $\mathbf{C} = [-11.5; -3.6; 55.0]$ in the camera-coordinate system and compute the ground truth projection of checkerboard corners using [3]. To evaluate performance with noise, we add different amount of Gaussian noise with standard deviation σ varying from 0 to 1 pixel in 2D projections. For each σ, we run 100 trials and average the error in the estimated rotation and translation. Translation error is computed as the norm of the translation error vector, normalized with the true translation magnitude. Rotation error is computed as the minimum angle of rotation between the estimated and the true rotation. For each σ and each trial, the input 8 points were randomly chosen.

Comparison with [31]: For comparison with [31], we observe the reflection of checkerboard corners using three planar mirrors. The location and orientation of planar mirrors used is similar to that in the demo example given at [30] (mirror distances are 351.8, 318.1 and 377.7 mm respectively.). For fair comparison, we made sure that the ground truth 2D projection of checkerboard corners using both spherical and planar mirrors occupy similar pixel area in the image. Figure 2 shows the 2D projections using spherical mirror (black) and using planar mirrors (red, green and blue). Notice that the spherical mirror projections are in-fact confined to a smaller pixel area.

Our algorithm uses 8 points while [31] requires only 3 points. Again for fair comparisons, we use the **same** 8 3D points for the planar mirror based approach in each trial. Notice that this results in $8 \times 3 = 24$ 2D projections for [31], whereas our approach uses only 8 2D projections. As shown in [31], using 3 points give 4 solutions for each P3P problem, leading to a potential of 64 solutions. However, we use all 8 correspondences to obtain a single least squares solution for each mirror[2]. In addition, we also use all 8 correspondences for computing the mirror orientations and checkerboard pose in the least square framework of [31]. For implementation, we used the matlab code provided at [30].

Simulation Results: Figure 2 compares the error in rotation and translation for the initial estimates (without non-linear refinement) obtained using [31] and by using our algorithm. Notice that the translation error is much smaller using our approach compared to [31] (11.9% versus 55.3% for $\sigma = 1$). However, our initial rotation estimate has a slightly larger average error ($4.3°$ versus $2.1°$ for $\sigma = 1$). Figure 3 shows the final error in rotation and translation after non-linear refinement. Notice that after non-linear refinement, the rotation error for our algorithm is similar to [31]. However, our approach provides a significantly smaller translation error compared to [31] (2.4% versus 6.0% for $\sigma = 1$). Thus, these simulations clearly demonstrate that our approach outperforms the state-of-art planar mirror based approach.

7. Analysis

Mirror Size: Intuitively, calibration using spherical mirror is possible using only a single view since a spherical mirror results in a non-central (non-perspective) camera. The degree of non-centrality depends on the radius of the mirror. Thus, it is expected that a larger mirror (with larger radius) will perform better than a smaller mirror. In practice, the deviations of camera rays from a perspective (central) model

[2] Better than using 3 points since we assume no outliers.

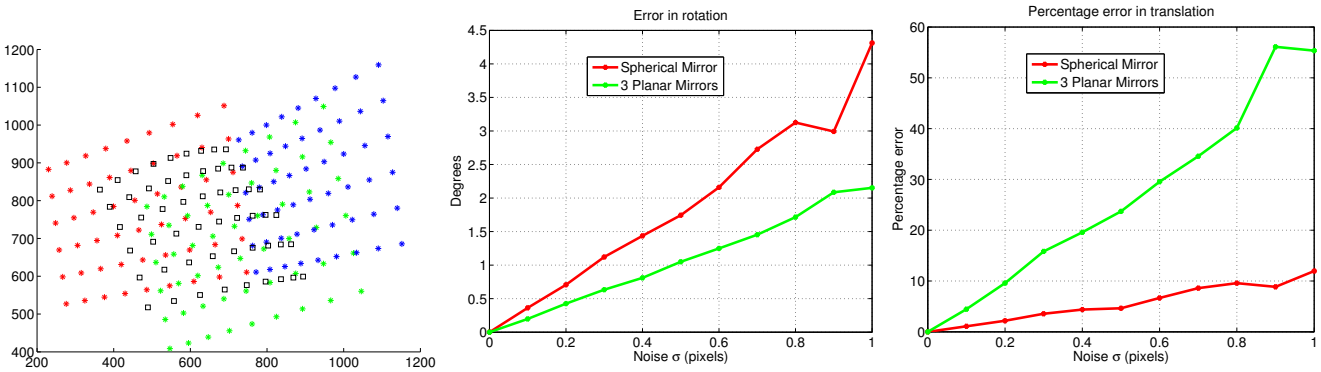

Figure 2. (Left) 2D projections of reference 3D points using planar mirrors (red, green, blue) and spherical mirror (black) used in simulation. (Middle and Right) Error in rotation and translation with noise for initial estimates.

Figure 3. (Left and Middle) Error in rotation and translation with noise for final estimates after non-linear refinement. (Right) Percentage error in estimated sphere location before and after non-linear refinement.

should be significantly larger than the deviations caused by imaging noise. We analyze the effect of mirror size on calibration accuracy using the same simulation scenario as in Section 6. Figures 4 and 5 show the errors in translation and rotation for different mirror radii but for the same location of spherical mirror. Notice that larger mirror size provides smaller error both for initial estimates and estimates after non-linear refinement.

Mirror radius error: Since our algorithm assumes that the mirror radius is known, we evaluate performance using an incorrect value of radius. Typical reflective balls are manufactured with diameter tolerance of $\pm 0.005" = \pm 0.127$ mm (available at mcmaster.com). We evaluate the performance for true radius $r = 25.4$ mm assuming 5% error ($= 1.27$ mm), which is ten times larger than the above tolerance. Figure 6 shows the error in estimated pose after non-linear refinement using incorrect radius of $24.13 = 25.4 - 1.27$ mm. Note that even with no noise ($\sigma = 0$), the pose error is non-zero due to incorrect radius. Compared to the correct radius case (Figure 3), the pose error is slightly larger (as expected). However, even when using incorrect radius, our estimates are better than those obtained using [31]. Thus, our approach can easily handle

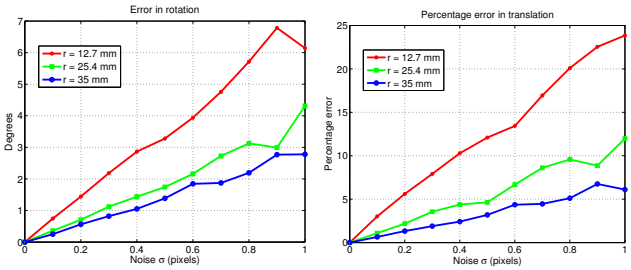

Figure 4. As the size of spherical mirror increases, pose estimation performance with noise improves. Plots show errors for initial estimates using mirror of different radii at the same location.

standard manufacturing tolerance of spherical balls.

8. Real Experiments

In this section, we show real experiments on extrinsic calibration using our approach and compare it with with [31]. We use a Canon Rebel XT camera with $18 - 55$ mm zoom lens. The lens was set to the maximum zoom setting. A checkerboard with square size 30 mm was placed outside the FOV of the camera. We captured one photo us-

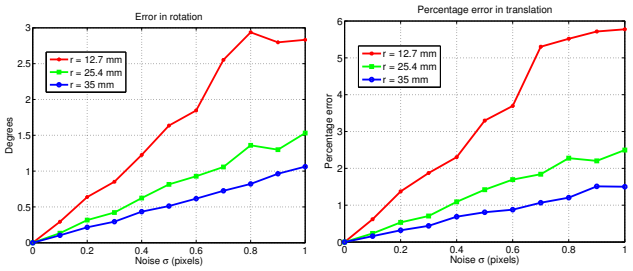

Figure 5. Pose estimation errors with varying mirror size after non-linear refinement of initial estimates. The errors in initial estimates are shown in Figure 4.

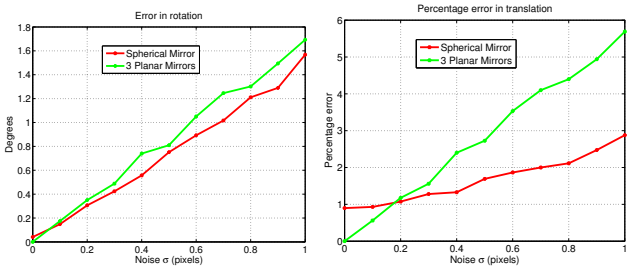

Figure 6. Pose estimation errors with 5% error in true mirror radius ($r = 25.4$ mm). The errors are slightly larger than shown in Figure 3 as expected. However, the estimates are better compared to [31].

ing a spherical ball[3] of diameter 3 inches and three photos using a planar mirror (Figure 7). Only the larger checkerboard in Figure 7 was used. The checkerboards were automatically detected using `LIBCBDETECT` software provided by Geiger *et al*. [8]. For spherical mirror, 48 corners were detected whereas for planar mirror, 30 corners were detected in each photo.

Table 2 compares the estimated pose and average re-projection error using our approach with [31]. Note that the 2D projections are confined to a significantly smaller pixel area for the spherical mirror. Despite this, the obtained pose is similar to that obtained using planar mirrors. Figure 8 shows the detected checkerboard points (red) along with the re-projected checkerboard points (green) using estimated the calibration parameters for our approach, which overlap nicely.

9. Discussions and Conclusions

To the best of our knowledge, we have presented the first approach for mirror-based extrinsic calibration using a single reflection from a non-planar spherical mirror. Our approach is automatic and does not require any user interaction. The spherical mirror provides a win-win situation: (a) only a single view is required, (b) there exist no degenerate mirror configurations, and (c) it provides better performance than the state-of-art planar mirror based approach in the presence of noise. We showed extensive analysis of

[3] available at mcmaster.com

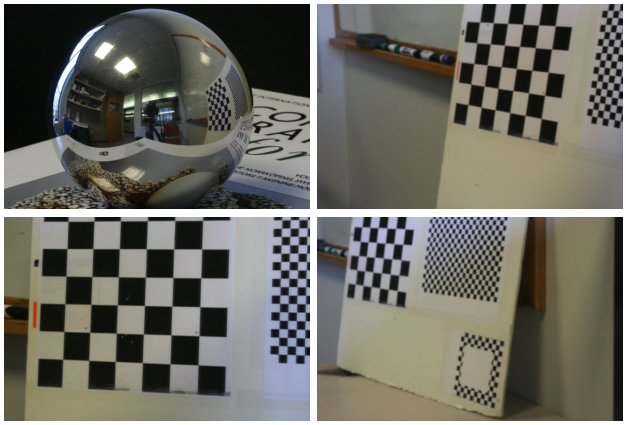

Figure 7. Single spherical mirror photo and three planar mirror photos used in the real experiment. The camera and the checkerboard was at the same location for all four photos.

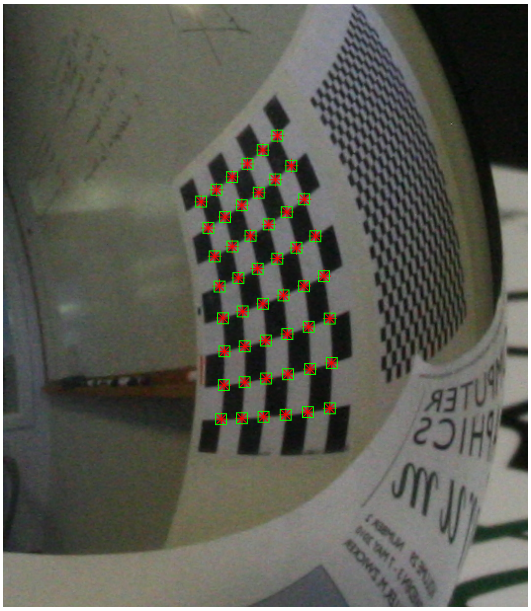

Figure 8. Detected corner points (red) and re-projected corner points (green) after pose estimation for our case. The RMS re-projection error is $J = 0.17$ pixels.

our approach with varying mirror size and assuming error in known mirror radius.

Since spherical mirrors are easy to manufacture and are low cost, we believe that our algorithm will be widely adopted. Our algorithm equivalently provides a method for calibrating spherical mirror based catadioptric systems using a single photo of a checkerboard. Techniques for display-camera calibration [25] often utilize reflections from an eye by modeling them as spherical mirror reflection. We believe that our approach will be useful in several vision applications such as specular surface reconstruction, robot navigation, catadioptric imaging, display-camera calibration and multi-camera calibration without overlapping field of view.

Method	$\theta_x, \theta_y, \theta_z$ (degree)	t_x, t_y, t_z (mm)	Re-projection error J (pixels)
Takahashi *et al*. [31]	$-79.30, 7.33, -80.48$	$316.45, 82.34, 435.46$	1.28
Ours	$-79.48, 6.69, -80.34$	$321.40, 78.59, 440.68$	0.17

Table 2. Comparison of checkerboard pose obtained using our method with [31] using real photos shown in Figure 7.

References

[1] A. Agrawal and S. Ramalingam. Single image calibration of multi-axial imaging systems. accepted in CVPR 2013. 2, 3, 4

[2] A. Agrawal, S. Ramalingam, Y. Taguchi, and V. Chari. A theory of multi-layer flat refractive geometry. In *CVPR*, pages 3346–3353, June 2012. 3

[3] A. Agrawal, Y. Taguchi, and S. Ramalingam. Analytical forward projection for axial non-central dioptric and catadioptric cameras. In *ECCV*, volume 6313, pages 129–143, Sept. 2010. 2, 3, 4, 5

[4] M. Demazure. Sur deux problemes de reconstruction. Technical Report 882, INRIA, 1988. 3

[5] Y. Ding, J. Yu, and P. Sturm. Multi-perspective stereo matching and volumetric reconstruction. In *ICCV*, pages 1827–1834, Sept. 2009. 2

[6] Y. Francken, C. Hermans, and P. Bekaert. Screen-camera calibration using a spherical mirror. In *Canadian Conf. Computer and Robot Vision*, pages 11–20, May 2007. 2

[7] Y. Francken, C. Hermans, and P. Bekaert. Screen-camera calibration using gray codes. In *Canadian Conf. Computer and Robot Vision*, pages 155–161, May 2009. 2

[8] A. Geiger, F. Moosmann, O. Car, and B. Schuster. Automatic camera and range sensor calibration using a single shot. In *Proc. IEEE Int'l Conf. Robotics and Automation*, pages 3936–3943, May 2012. 7

[9] J. Gluckman and S. Nayar. Catadioptric stereo using planar mirrors. *Int'l J. Computer Vision*, 44(1):65–79, Aug. 2001. 2

[10] A. Goshtasby and W. A. Gruver. Design of a single-lens stereo camera system. *Pattern Recognition*, 26:923–937, June 1993. 2

[11] J. A. Grunert. Das pothenotische problem in erweiterter gestalt nebst uber seine anwendungen in der geod isie. *Grunerts Archiv fiir Mathematik und Physik*, 1:238–248, 1841. 1

[12] R. M. Haralick, C.-N. Lee, K. Ottenberg, and M. Nölle. Review and analysis of solutions of the three point perspective pose estimation problem. *Int'l J. Computer Vision*, 13:331–356, 1994. 1

[13] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000. 3

[14] J. A. Hesch, A. I. Mourikis, and S. I. Roumeliotis. Extrinsic camera calibration using multiple reflections. In *ECCV*, pages 311–325, Sept. 2010. 2

[15] K. H. Jang, D. H. Lee, and S. K. Jung. A moving planar mirror based approach for cultural reconstruction: Research articles. *Comput. Animat. Virtual Worlds*, 15(3-4):415–423, July 2004. 2

[16] M. Kanbara, N. Ukita, M. Kidode, and N. Yokoya. 3D scene reconstruction from reflection images in a spherical mirror. In *ICPR*, volume 4, pages 874–879, Aug. 2006. 2

[17] Y. Kojima, R. Sagawa, T. Echigo, and Y. Yagi. Calibration and performance evaluation of omnidirectional sensor with compound spherical mirrors. In *OMNIVIS*, Oct. 2005. 2

[18] R. Kumar, A. Ilie, J.-M. Frahm, and M. Pollefeys. Simple calibration of non-overlapping cameras with a mirror. In *CVPR*, pages 1–7, June 2008. 1, 2

[19] D. Lanman, D. Crispell, M. Wachs, and G. Taubin. Spherical catadioptric arrays: Construction, multi-view geometry, and calibration. In *3DPVT*, pages 81–88, 2006. 2

[20] P. Lbraly, C. Lasmea, C. Deymier, O. Ait-Aider, E. Royer, and M. Dhome. Flexible extrinsic calibration of non-overlapping cameras using a planar mirror: Application to vision-based robotics. In *IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems*, pages 5640–5647, Oct. 2010. 2

[21] E. L. Merritt. Explicit three-point resection in space. *Photogrammetric Engineering*, 15(4):649–655, 1949. 1

[22] B. Micusik and T. Pajdla. Structure from motion with wide circular field of view cameras. *IEEE Trans. Pattern Anal. Machine Intell.*, 28(7):1135–1149, 2006. 2

[23] S. A. Nene and S. K. Nayar. Stereo with mirrors. In *ICCV*, pages 1087–1094, Jan. 1998. 2

[24] D. Nistér. An efficient solution to the five-point relative pose problem. *IEEE Trans. Pattern Anal. Machine Intell.*, 26(6):756–770, June 2004. 3

[25] C. Nitschkea, A. Nakazawaa, and H. Takemuraa. Display-camera calibration using eye reflections and geometry constraints. *Computer Vision and Image Understanding*, 115:835–853, June 2011. 1, 2, 7

[26] I. Reshetouski, A. Manakov, H.-P. Seidel, and I. Ihrke. Three-dimensional kaleidoscopic imaging. In *CVPR*, pages 353–360, June 2011. 2

[27] R. Rodrigues, J. P. Barreto, and U. Nunes. Camera pose estimation using images of planar mirror reflections. In *ECCV*, pages 382–395, Sept. 2010. 1, 2

[28] P. Sturm and T. Bonfort. How to compute the pose of an object without a direct view. In *ACCV*, pages 21–31, Jan. 2006. 1, 2

[29] Y. Taguchi, A. Agrawal, A. Veeraraghavan, S. Ramalingam, and R. Raskar. Axial-cones: Modeling spherical catadioptric cameras for wide-angle light field rendering. *ACM Trans. Graph.*, 29(6):172:1–172:8, Dec. 2010. 2

[30] K. Takahashi, S. Nobuhara, and T. Matsuyama. http://vision.kuee.kyoto-u.ac.jp/ nob/proj/mirror/. 5

[31] K. Takahashi, S. Nobuhara, and T. Matsuyama. A new mirror-based extrinsic camera calibration using an orthogonality constraint. In *CVPR*, June 2012. 1, 2, 5, 6, 7, 8

[32] X.-H. Ying, K. Peng, R. Ren, and H. Zha. Geometric properties of multiple reflections in catadioptric camera with two planar mirrors. In *CVPR*, pages 1126–1132, June 2010. 2