

# Single Image Calibration of Multi-Axial Imaging Systems

Amit Agrawal and Srikumar Ramalingam  
 Mitsubishi Electric Research Labs (MERL)  
 201 Broadway, Cambridge, MA 02139  
 [agrawal,ramalingam] at merl.com

## Abstract

Imaging systems consisting of a camera looking at multiple spherical mirrors (reflection) or multiple refractive spheres (refraction) have been used for wide-angle imaging applications. We describe such setups as multi-axial imaging systems, since a single sphere results in an axial system. Assuming an internally calibrated camera, calibration of such multi-axial systems involves estimating the sphere radii and locations in the camera coordinate system. However, previous calibration approaches require manual intervention or constrained setups. We present a fully automatic approach using a single photo of a 2D calibration grid. The pose of the calibration grid is assumed to be unknown and is also recovered. Our approach can handle unconstrained setups, where the mirrors/refractive balls can be arranged in any fashion, not necessarily on a grid.

The axial nature of rays allows us to compute the axis of each sphere separately. We then show that by choosing rays from two or more spheres, the unknown pose of the calibration grid can be obtained linearly and independently of sphere radii and locations. Knowing the pose, we derive analytical solutions for obtaining the sphere radius and location. This leads to an interesting result that 6-DOF pose estimation of a multi-axial camera can be done without the knowledge of full calibration. Simulations and real experiments demonstrate the applicability of our algorithm.

## 1. Introduction

Catadioptric imaging systems consist of a camera looking at single or multiple mirrors for wide-angle imaging. Except for a few special configurations as described in [3], such systems do not have an effective single view-point. The special configurations include a perspective camera placed at the focal point of a hyperbolic or an elliptical mirror, or an orthographic camera placed on the axis of a parabolic mirror. Common configurations such as camera viewing a spherical mirror or multiple mirrors always lead to a non-central imaging system [23]. On the other hand, a camera looking through single/multiple spherical refrac-



Figure 1. Camera looking at spherical mirrors (left) and refractive spheres (right). Can we calibrate such imaging systems automatically using a single photo of a checkerboard?

tive balls [24, 2] results in a non-central dioptric system. Such systems have been used in several applications such as vision-guided navigation [12], wide-angle 3D reconstruction [7, 11, 15, 14] and recognition [19], and for capturing wide-angle light fields [24]. In this paper, we describe such setups (see Figure 1) as *multi-axial* imaging systems, since a single spherical mirror/refractive ball results in an axial system [2, 24]. Without loss of generality, we use ‘sphere’ to refer to the spherical mirror as well as the refractive sphere.

To effectively use such systems in a practical application, one needs to have an easy calibration procedure. However, calibration of such systems typically require manual intervention or constrained setups. Taguchi *et al.* [24] manually marked the outline of each sphere in the captured image to get an initial estimate of sphere center and radius. This becomes cumbersome if several spheres are used. Techniques such as [8, 19] also used the image contour of sphere for localization. Some techniques require constrained setups, which are difficult in practice. For example, Kojima *et al.* [12] attach the imaging system on two turn-tables, and rotate it along two axis to observe the reflection of a light source. Nayar’s [15] sphereo system require sphere centers to lie on an plane orthogonal to the imaging plane, along with coplanar light sources such that their centroid lies on the optical axis for calibration. Other techniques require placing the mirrors on a planar grid and observe markers on the grid to obtain the grid pose. This along with mechanical design data is used to estimate the initial sphere locations [13, 7]. However, this restricts the placement of spheres and requires markers to be seen, losing some spa-

tial resolution for the scene reflected by the mirror.

We show that calibration of such multi-axial systems can be done using a *single* photo of a 2D calibration grid (e.g. checkerboard). The pose of the checkerboard is assumed to be unknown, and is also recovered. This makes our approach useful and easy in practice. Our approach is independent of the positioning of the spheres; they can be arranged in any fashion and not necessarily placed on a planar grid. It does not require estimation of the sphere contour in the captured photo. Our approach can handle spheres of different radii and is marker-less. We only assume that the camera is internally calibrated off-line using a standard approach such as [4].

The key to our approach is understanding of the underlying geometry of rays. While a single sphere results in a multi-perspective image and corresponds to an axial camera, multiple spheres correspond to a multi-axial camera where each 3D point is imaged multiple times via underlying axial cameras. The axial nature of rays allows us to compute the axis of each sphere separately. By choosing rays from two or more spheres, we demonstrate that the unknown pose of the checkerboard can be obtained *linearly*, and independently of sphere radii and locations. Thus, 6-DOF pose estimation can be done in a *semi-calibrated* setting, without the knowledge of sphere locations and radii. This leads to an interesting result, since pose estimation typically requires full calibration for both central and non-central cameras. Finally, we derive analytical solutions for estimating the radii and center of spheres with known checkerboard pose and axes.

### 1.1. Related Work

**Axial Systems:** A camera looking at a single mirror can be categorized as an (a) axial or (b) off-axis system, depending on its placement on the mirror axis. A general framework for calibrating axial cameras was proposed by Ramalingam *et al.* [18]. However, they require three images of a checkerboard in different orientations and their parameterization involves two rotations/translations. In contrast, our approach allows calibration from a single photo. Radial distortion based models assume known center of distortion [27] or model each distortion circle separately [26]. Instead, we use a global model parameterized via sphere radius and center. Micusik and Pajdla [14] proposed auto-calibration of single mirror based axial systems starting from a central approximation, using multiple images. A brute force approach to apply axial methods to multi-axial scenario is to calibrate each sphere separately. However, this does not utilize inter-sphere constraints and the fact that the same checkerboard is seen via all spheres. We show that using rays from two or more spheres allow computing the checkerboard pose linearly and *independently* of sphere parameters. This greatly simplifies the calibration, where the checkerboard pose is estimated first, followed by the esti-

mation of sphere parameters.

Other imaging systems such as flat refractive systems [1] have also been shown to be axial. We build upon the algorithm proposed in [1] to compute the axes and checkerboard pose.

**Off-Axis Systems:** Off-axis mirror calibration is harder than calibrating axial systems. Caglioti *et al.* [5] proposed a line based algorithm for calibrating off-axis catadioptric cameras. Nishino and Nayar [16] modeled a camera looking at an eye as an off-axis ellipsoidal catadioptric system. However, they require manual intervention and fit an ellipse to the apparent cornea contour (limbus) in the image for calibration. In contrast, our method can work even if the cross-section of the mirror is not visible. Generic camera calibration algorithms [22] have also been proposed to calibrate non-central cameras. In contrast, we use a parametric model. To the best of our knowledge, ours is the first algorithm that allows calibrating multiple refractive spheres.

## 2. Multi-Axial Cameras

Consider the multi-axial imaging systems, where a perspective camera observes the scene reflected via  $n$  spherical mirrors, or refracted via  $n$  refractive spheres. We work in the camera coordinate system, with its center of projection (COP) at origin. Let  $\mathbf{C}(i)_{i=1}^n$  be the center of the  $i^{th}$  sphere with radius  $r_i$ . Let  $\mathbf{P}(i)_{i=1}^K$  denote  $K$  3D points on the calibration plane which are known in its coordinate system. Without loss of generality, we assume the plane to be aligned with the  $xy$  plane so that  $\mathbf{P}^z(i) = 0 \forall i$ . Let  $[R, \mathbf{t}]$  be the unknown rigid transformation of the calibration plane in the camera coordinate system. We assume that the internal camera calibration has been done offline and hence we know the camera ray  $\mathbf{v}(i, j)$  for each 3D point  $\mathbf{P}(i)$  and each mirror  $j$ . Our goal is to compute the center and radius of each sphere, as well as the unknown pose of the calibration grid given  $nK$  3D-2D correspondences. As discussed, we use the term ‘sphere’ to describe both the spherical mirror and the refractive sphere. The estimation of axes and checkerboard pose is identical for both cases, and they differ only in the estimation of sphere radii and locations.

**Axial and Multi-Axial Geometry:** It is well-known that a single sphere corresponds to an axial imaging system [2]. Let  $d_i = \|\mathbf{C}_i\|$  be the distance of the  $i^{th}$  sphere center from the COP. The axis  $\mathbf{A}_i = \mathbf{C}_i/d_i$  is defined as the normalized vector joining the sphere center to the COP.

## 3. Axis Computation

Agrawal *et al.* [1] derived the *coplanarity constraint* for axial cameras considering flat refractive geometry. Since each sphere also corresponds to an axial camera, the coplanarity constraint can be used to compute the axis  $\mathbf{A}_i$  of each sphere. For completeness, we re-derive the constraint as below.

Consider the plane of reflection (refraction)  $\pi$  given by the camera ray  $\mathbf{v}(i, j)$  and the axis  $\mathbf{A}_j$  for the  $j^{\text{th}}$  sphere. From Snell's law, the reflected (refracted) ray should lie on  $\pi$ . Thus, the corresponding 3D point  $R\mathbf{P}(i) + \mathbf{t}$  should also lie on  $\pi$ , leading to

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

where  $\times$  denotes the cross-product and  $\mathbf{A}_j \times \mathbf{v}(i, j)$  is the normal to  $\pi$ . Let  $[\mathbf{A}_j]_{\times}$  be the  $3 \times 3$  skew-symmetric matrix obtained from  $\mathbf{A}_j$ . The above equation can be written as

$$0 = \mathbf{v}(i, j)^T E_j \mathbf{P}(i) + \mathbf{v}(i, j)^T \mathbf{s}_j,$$

where  $E_j = [\mathbf{A}_j]_{\times} R$  and  $\mathbf{s}_j = \mathbf{A}_j \times \mathbf{t}$ . Since  $\mathbf{s}_j^T \mathbf{A}_j = 0$ , the component of translation  $\mathbf{t}$  along the axis  $\mathbf{A}_j$ ,  $\mathbf{t}_A(j)$ , cannot be estimated using the coplanarity constraint.

Stacking the coplanarity equation for 8 correspondences, we get a linear system

$$\underbrace{\begin{bmatrix} (\mathbf{P}(1)^T \otimes \mathbf{v}(1, j)^T) & \mathbf{v}(1, j)^T \\ \vdots & \vdots \\ (\mathbf{P}(8)^T \otimes \mathbf{v}(8, j)^T) & \mathbf{v}(8, j)^T \end{bmatrix}}_{\mathbf{B}} \begin{bmatrix} E_j(\cdot) \\ \mathbf{s}_j \end{bmatrix} = 0, \quad (2)$$

where  $\otimes$  denotes Kronecker product and  $B$  is an  $11 \times 12$  matrix. Since  $\mathbf{P}^z(i) = 0$ , columns 7, 8 and 9 of  $B$  reduce to zero. Let  $B_r$  be the reduced  $8 \times 9$  matrix, whose rank is 8. Let  $\mathbf{e}_1^j, \mathbf{e}_2^j, \mathbf{e}_3^j$  be the columns of  $E_j$ . Then the above linear system is equivalent to

$$B_r \begin{bmatrix} \mathbf{e}_1^j \\ \mathbf{e}_2^j \\ \mathbf{s}_j \end{bmatrix} = 0. \quad (3)$$

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, we can compute the first two columns of  $E_j$  and  $\mathbf{s}_j$  using 8 correspondences. In [1], Demazure constraints [6] were used to compute the last column of  $E_j$ , which requires solving a  $4^{\text{th}}$  degree equation. However, it is not necessary and the axis can be computed directly from the first two columns of  $E_j$  as:  $\mathbf{A}_j = (\mathbf{e}_1^j \times \mathbf{e}_2^j) / \|(\mathbf{e}_1^j \times \mathbf{e}_2^j)\|$ . This is because  $\mathbf{A}_j^T E_j = 0$  and hence  $\mathbf{A}_j$  is orthogonal to both  $\mathbf{e}_1^j$  and  $\mathbf{e}_2^j$ .

Thus, by using 8 correspondences, the axis  $\mathbf{A}_j$  can be estimated for each sphere. Next, we show that the translation ambiguity along any individual axis can be removed by using rays from two or more spheres, since two or more axes will not be parallel in general.

## 4. Recovering the Checkerboard Pose

A brute-force approach for calibrating multi-axial system would be to calibrate each sphere *separately*, as an axial system. Note that the checkerboard pose in the camera coordinate system is identical with respect to all the spheres.

However, in presence of inevitable noise in images, the estimated pose will be different for each sphere. Such an approach does not utilize the fact that the pose is identical for all spheres, thereby not utilizing *inter-sphere* constraints.

In addition, calibrating each sphere separately leads to a more difficult and computationally expensive algorithm. Let us briefly discuss the calibration procedure for a single spherical mirror. The coplanarity constraint allows us to compute the axis  $\mathbf{A}_j$ , the matrix  $E_j$  and the vector  $\mathbf{s}_j$  for each mirror. As explained earlier, the translation along the axis,  $\mathbf{t}_A(j)$ , is not recovered. The rotation matrix  $R_j$  can be recovered from  $E_j$ , similar to obtaining rotation matrix from essential matrix. However, four rotation matrices are obtained from  $E_j$  due to twisted pair ambiguity [9]. Thus, there remains three calibration parameters for each mirror: (a) mirror radius  $r_j$ , (b) mirror distance  $d_j$  and (c)  $\mathbf{t}_A(j)$ . Solving for  $r_j, d_j$  and  $\mathbf{t}_A(j)$  requires solving a set of non-linear equations. These non-linear equations need to be solved *four* times, for each rotation matrix to find the correct solution for a given mirror. Thus, for  $n$  mirrors, if we calibrate each mirror separately, we need to solve the set of non-linear equations  $4n$  times (instead of  $n$  times), which is computationally inefficient.

Now we show that by using rays from two or more spheres, the translation ambiguity along any single axis can be removed. The pose can be obtained in (a) linear fashion, and (b) *independently* of sphere radii and distances. This has several advantages. Firstly, we can compute a common pose, instead of having to deal with  $n$  different poses. Secondly, the full pose can be recovered linearly without estimating sphere parameters, thus enabling pose estimation in a semi-calibrated setting. Finally, since the translation ambiguity is resolved, two parameters remain for each sphere (radius and distance) instead of three. This greatly simplifies the resulting non-linear equations. Solving for sphere radius  $r_j$  and distance  $d_j$  along with  $\mathbf{t}_A(j)$  proved to be difficult and we were not able to find an analytical solution to estimate  $r_j, d_j$  and  $\mathbf{t}_A(j)$  simultaneously.

### 4.1. Linear Estimation of Pose

We assume that the axes of spheres can be estimated as explained in Section 3. To solve for pose, we re-write the coplanarity constraint in terms of the unknown rotation  $R$  and translation  $\mathbf{t}$  as

$$0 = \mathbf{q}(i, j)^T R\mathbf{P}(i) + \mathbf{q}(i, j)^T \mathbf{t}, \quad (4)$$

where  $\mathbf{q}(i, j) = [\mathbf{A}_j]_{\times}^T \mathbf{v}(i, j)$ . Let  $R_{12}$  be the  $3 \times 2$  matrix corresponding to the first two columns of  $R$ . Stacking  $N$  equations, we get a linear system (similar to (3)) as,

$$K \begin{bmatrix} R_{12}(\cdot) \\ \mathbf{t} \end{bmatrix} = 0, \quad (5)$$

where  $K$  is an  $N \times 9$  matrix (the third column of  $R$  vanishes since  $\mathbf{P}^z(i) = 0$ ).  $K$  can be written as  $\begin{bmatrix} K_R & K_t \end{bmatrix}$ , where

$K_R$  is the  $N \times 6$  matrix corresponding to the rotation and  $K_t$  is the  $N \times 3$  matrix corresponding to the translation. We first compute the rotation and then compute the translation to recover the scale of translation also. Following [9], the solution for  $R_{12}$  can be obtained from

$$\underbrace{(K_t K_t^\dagger - I_N) K_R}_{G} R_{12}(\cdot) = 0, \quad (6)$$

where  $\dagger$  represents the pseudo-inverse and  $I_N$  is a  $N \times N$  identity matrix.

To recover the full translation, the rank of  $K_t$  should be 3. However, using rays from a single sphere, the rank of  $K_t$  is 2, which demonstrates the translation ambiguity along the axis. If we combine the coplanarity constraints from two mirrors  $j$  and  $k$  whose axes are not parallel,  $K_t$  has rank 3 and  $G$  has rank 5.  $R_{12}$  can then be recovered using standard SVD solution from (6). The last column of  $R$  is obtained as the cross product of the first two columns. In presence of noise, the closest rotation matrix is obtained by taking the SVD of  $R$  as  $R = USV^T$  and setting  $R = UV^T$ . After obtaining rotation,  $\mathbf{t}$  is given by  $\mathbf{t} = -K_t^\dagger K_R R_{12}(\cdot)$ . Thus, one can use all the rays from all the spheres to obtain a least square estimate of pose.

Intuitively, the translation ambiguity is resolved as follows.  $\mathbf{t}$  can be written as  $\mathbf{t} = \mathbf{s}_j + \alpha \mathbf{A}_j$ , where  $\alpha$  is an unknown scalar. For a single sphere, only  $\mathbf{s}_j$  and  $\mathbf{A}_j$  are recovered using coplanarity constraints and hence  $\alpha$  and  $\mathbf{t}$  are ambiguous. Using a second sphere  $k$ , we get another constraint  $\mathbf{t} = \mathbf{s}_k + \beta \mathbf{A}_k$ . Thus, we have

$$\mathbf{s}_k - \mathbf{s}_j = \alpha \mathbf{A}_j - \beta \mathbf{A}_k. \quad (7)$$

If  $\mathbf{A}_k$  and  $\mathbf{A}_j$  are not parallel,  $\alpha$  and  $\beta$  can be solved using the above 3-vector equation, recovering the full translation  $\mathbf{t}$  and resolving the ambiguity.

Let  $\mathbf{Q}(i) = \mathbf{R}\mathbf{P}(i) + \mathbf{t}$  be the transformed checkerboard points in the camera coordinate system. There is still a sign ambiguity remaining, since  $[-R, -\mathbf{t}]$  will also be a solution. This can be resolved using chirality constraints, by checking the sign of  $\mathbf{Q}^z(i)$ .

## 5. Solving for Spherical Mirror Parameters

Now we consider the mirror case and describe how to solve for the sphere radius and distance assuming known checkerboard pose and axes. When the axes are known, the analysis can be done on the plane of refraction  $\pi$  in 2D for each mirror. For simplicity, we drop the subscript  $j$  for the mirror in this section.

Let  $[\mathbf{z}_1, \mathbf{z}_2]$  denote an orthogonal coordinate system on  $\pi$ . We choose  $\mathbf{z}_2$  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  $\pi$  is given by  $\mathbf{u} = [u_x, u_y]$ , where  $u_x = \mathbf{z}_1^T \mathbf{Q}(i)$  and  $u_y = \mathbf{z}_2^T \mathbf{Q}(i)$ . Similarly, the camera ray  $\mathbf{v}(i)$  can be represented by the 2D vector  $\mathbf{w} = [w_x, w_y]$  on

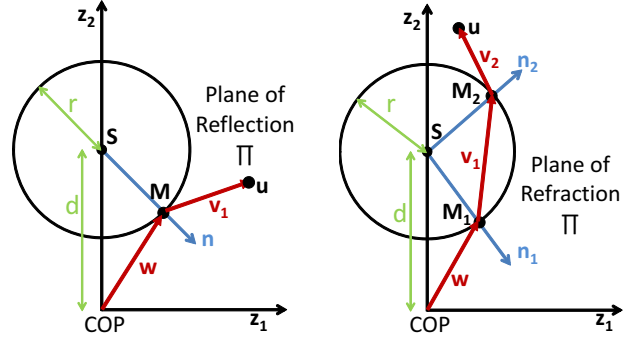


Figure 2. Estimation of sphere radius  $r$  and distance  $d$  for spherical mirror (left) and refractive ball (right) assuming known checkerboard pose and mirror axis.

$\pi$ , 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  $\mathbf{w}^T \mathbf{w} = 1$ . Let  $\beta = r^2$ .

Referring to Figure 2, the spherical mirror is represented as a circle on  $\pi$  with center at  $\mathbf{S} = [0, d]$  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,  $(\mathbf{M} - \mathbf{S})^T (\mathbf{M} - \mathbf{S}) = r^2 = \beta$ . Hence,

$$(kw_x)^2 + (kw_y - d)^2 = k^2 - 2kw_y d + d^2 = \beta. \quad (8)$$

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

$$0 = \mathbf{v}_1 \times (\mathbf{u} - \mathbf{M}). \quad (9)$$

After substituting all the terms, the above equation simplifies to

$$K_1 k^2 + K_2 k + K_3 = 0, \quad (10)$$

where  $K_1 = 2(dw_x + u_x w_y - u_y w_x)$ ,  $K_2 = -2d(u_x + u_x w_y^2 + dw_x w_y - u_y w_x w_y)$ , and  $K_3 = 2u_x w_y d^2 - \beta u_x w_y + \beta u_y w_x$ .

By eliminating  $k$  between (8) and (10), we obtain a single equation in  $d$  and  $\beta$ , which is 6<sup>th</sup> degree in  $d$  and quadratic in  $\beta$ . Since we have two unknowns ( $d$  and  $\beta$ ), we need two equations from two correspondences. Let  $EQ_1$  and  $EQ_2$  be the two such equations. By eliminating  $\beta$  between  $EQ_1$  and  $EQ_2$ , we obtain a 7<sup>th</sup> degree equation<sup>1</sup> in the single unknown  $d$ . Solving this equation gives 7 solutions for  $d$ . For each solution of  $d$ , the corresponding solution for  $\beta$  can be obtained from  $EQ_1$  (or  $EQ_2$ ). The correct solutions are obtained by removing imaginary solutions, enforcing  $d > 0$ ,  $\beta > 0$ ,  $d > \sqrt{\beta}$  and testing using additional correspondences.

**Known Radius:** In practice, a good initial estimate of mirror radius may be known. If so, we only need to solve a 6<sup>th</sup> degree equation in single unknown  $d$ .

<sup>1</sup>Details are provided in the supplementary materials.

## 5.1. Non-linear refinement

Non-linear refinement can be done by minimizing the image re-projection error for all checkerboard points. The projection of a 3D point can be obtained by solving a 4<sup>th</sup> degree equation for a spherical mirror as shown in [2]. Let  $\mathbf{x}$  denote all calibration parameters and  $\hat{p}(i, j)$  be the projection of  $\mathbf{P}(i)$  for  $i^{\text{th}}$  point and  $j^{\text{th}}$  mirror. Let  $p(i, j)$  be the corresponding detected checkerboard corner in the image. We use `lsqnonlin` in Matlab to refine  $\mathbf{x}$  by minimizing the total squared error  $J = \sum (p(i, j) - \hat{p}(i, j))^2$ .

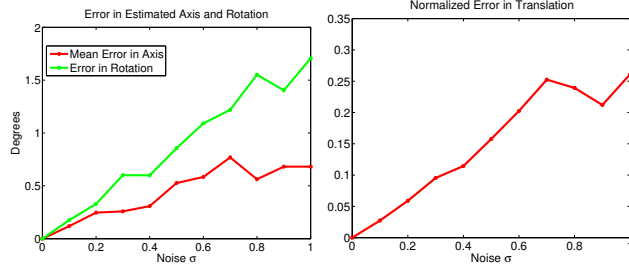


Figure 3. Simulation results for multiple spherical mirrors. Error in axis, translation and rotation after initial estimation.

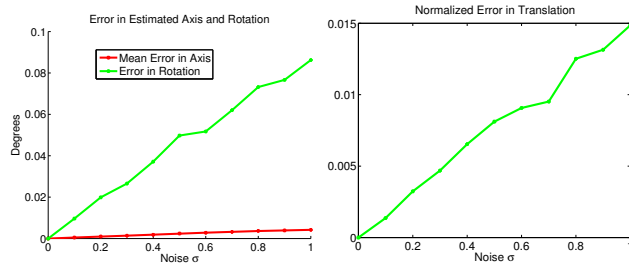


Figure 4. Error in axis, translation and rotation after non-linear refinement for simulation in Figure 3.

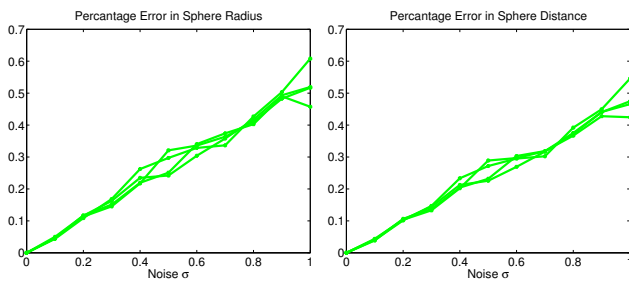


Figure 5. Simulation results for multiple spherical mirrors. Error in sphere locations and sphere radius after non-linear refinement for all four mirrors.

## 6. Simulations

We first show simulation results for axes and pose estimation. We use a realistic scenario with four spherical mirrors of radius 0.5 inch, viewed from a camera with focal length 55 mm and resolution  $2000 \times 2000$  pixels, at a distance of 120 mm. We assume a  $9 \times 6$  checkerboard grid

with each square of size 60 mm. We add different amount of Gaussian noise with standard deviation  $\sigma$  varying from 0 to 1 pixel. For each  $\sigma$ , we run 100 trials and average the error in the estimated rotation, translation and axes. 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 estimated and true rotation. Error in axis is computed as the angle between the estimated and true axes, and is averaged over all spheres. Figure 3 plots the error in axes, translation and rotation with increasing noise. The axes and rotation error are within few degrees. Thus, our approach can provide a good initial estimate for subsequent non-linear refinement.

Next, we analyze the complete estimation of all calibration parameters. After estimating the axes and pose, the mirror parameters are estimated, followed by non-linear refinement. Figure 4 plots the error in axes, translation and rotation after non-linear refinement. Figure 5 shows the error in sphere locations and radii for all four mirrors, as a percentage of the true values. Notice that the maximum error is less than 0.7%.

## 7. Multiple Refractive Spheres

Now we consider a camera looking through multiple refractive spheres (ball lenses) as shown in Figure 8. Similar to mirrors, the axis is defined as the vector joining the camera center to the center of the refractive sphere. The calibration parameters consist of centers, radii and refractive indices of the balls along with the unknown checkerboard pose. We assume that the refractive index  $\mu$  is known and is identical for all spheres. As described in Section 3 and Section 4.1, we can compute the checkerboard pose and axis of each refractive ball in exactly the same manner using the multi-axial property. Only the estimation of ball center and radius is different from the previous case due to refraction.

### 7.1. Recovering Refractive Sphere Parameters

As in before, the analysis can be done on the plane of refraction  $\pi$  (see Figure 2), since the entire light path remains on  $\pi$ . We follow Section 5. Let  $\mathbf{M}_1 = k\mathbf{w}$  be the common point on the refractive ball and the camera ray  $\mathbf{w}$ . Then we have

$$(kw_x)^2 + (kw_y - d)^2 = k^2 - 2kw_yd + d^2 = \beta, \quad (11)$$

similar to (8). The normal at  $\mathbf{M}_1$  is given by  $\mathbf{n}_1 = \mathbf{M}_1 - \mathbf{S}$ , where  $\mathbf{n}_1^T \mathbf{n}_1 = \beta$ . The first refracted ray  $\mathbf{v}_1$  is given by the Snell's law [2] as

$$\mathbf{v}_1 = \frac{1}{\mu} \mathbf{w} + b_1 \mathbf{n}_1, \quad (12)$$

where

$$b_1 = \frac{-\mathbf{w}^T \mathbf{n}_1 - \sqrt{(\mathbf{w}^T \mathbf{n}_1)^2 - (1 - \mu^2)\beta}}{\mu\beta}. \quad (13)$$

The second refraction point  $\mathbf{M}_2$  can be written as  $\mathbf{M}_2 = \mathbf{M}_1 + \lambda \mathbf{v}_1$  for some constant  $\lambda$ . Since  $\mathbf{M}_2$  lies on the sphere,

$$\begin{aligned} \beta &= (\mathbf{M}_2 - \mathbf{S})^T (\mathbf{M}_2 - \mathbf{S}) = (\mathbf{n}_1 + \lambda \mathbf{v}_1)^T (\mathbf{n}_1 + \lambda \mathbf{v}_1), \\ &= \beta + 2\lambda \mathbf{n}_1^T \mathbf{v}_1 + \lambda^2 \mathbf{v}_1^T \mathbf{v}_1. \end{aligned} \quad (14)$$

This gives  $\lambda = -2\mathbf{n}_1^T \mathbf{v}_1 / (\mathbf{v}_1^T \mathbf{v}_1)$ . The normal at  $\mathbf{M}_2$  is then given by  $\mathbf{n}_2 = \mathbf{M}_2 - \mathbf{S}$ . The outgoing ray  $\mathbf{v}_2$  can be written as  $\mathbf{v}_2 = \mu \mathbf{v}_1 + b_2 \mathbf{n}_2$ , for some  $b_2$ . Taking dot product with  $\mathbf{n}_2$  on both sides,  $b_2 = (\mathbf{v}_2^T \mathbf{n}_2 - \mu \mathbf{v}_1^T \mathbf{n}_2) / \beta$ . From the symmetry of the sphere, we have  $\mathbf{v}_2^T \mathbf{n}_2 = -\mathbf{w}^T \mathbf{n}_1$  and  $\mathbf{v}_1^T \mathbf{n}_2 = -\mathbf{v}_1^T \mathbf{n}_1$ . Thus,

$$b_2 = (-\mathbf{w}^T \mathbf{n}_1 + \mu \mathbf{v}_1^T \mathbf{n}_1) / \beta. \quad (15)$$

Finally, the outgoing ray  $\mathbf{v}_2$  should pass through  $\mathbf{u}$ . Thus,

$$\mathbf{v}_2 \times (\mathbf{u} - \mathbf{M}_2) = 0. \quad (16)$$

After substituting all the terms, the above equation simplifies to an equation which is 5<sup>th</sup> degree in  $d$ , 4<sup>th</sup> degree in  $k$  and quadratic in  $\beta$ . Eliminating  $k$  between this equation and (11), we obtain a single equation in  $d$  and  $\beta$ , which is 12<sup>th</sup> degree in  $d$  and fourth degree in  $\beta$ .

Thus, each correspondence results in a single equation in  $d$  and  $\beta$ . Similar to the mirror case, we need two correspondences for two equations. However, eliminating  $\beta$  turned out to be quite difficult due to high complexity of this equation. We therefore assume that the radius of refractive ball (and hence  $\beta$ ) is known a-priori<sup>2</sup> and solve for the single unknown  $d$  by solving the above derived 12<sup>th</sup> degree equation. Similar to the mirror case, we perform non-linear refinement after the above initial estimation. The projection of a 3D point can be obtained by solving a 10<sup>th</sup> degree equation for refractive sphere as shown in [2].

**Simulation:** We use similar settings as in Section 6 to simulate a camera looking through four refractive spheres of radius 12.7 mm and refractive index 1.5. The center of the spheres are placed at a mean distance of 80 mm from the camera. The checkerboard is placed at  $T = [3.4, 1.2, 250]$  mm from the camera with the rotation angles of  $[2.2, 3.4, 5.3]$  degrees. Neither the spherical array nor the checkerboard is fronto-parallel to the imaging plane. Figure 6 and Figure 7 shows the error in estimated axes, sphere distances and checkerboard pose using our algorithm after non-linear refinement, with Gaussian noise of increasing variance added to checkerboard corners. These plots show that accurate calibration can be obtained using our approach. For example, the maximum error in sphere distance is less than 0.7%.

## 8. Real Experiments

We performed real experiments using Canon Rebel XT camera having a resolution of  $3256 \times 2304$  pixels with 18 –

<sup>2</sup>In practice, radius may be known from manufacture to a good precision.

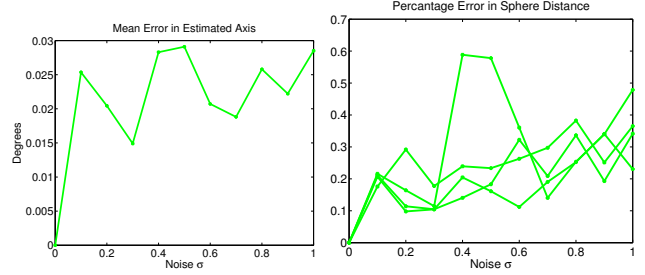


Figure 6. Simulation results for multiple refractive spheres assuming known sphere radii. Error in sphere axes and sphere locations.

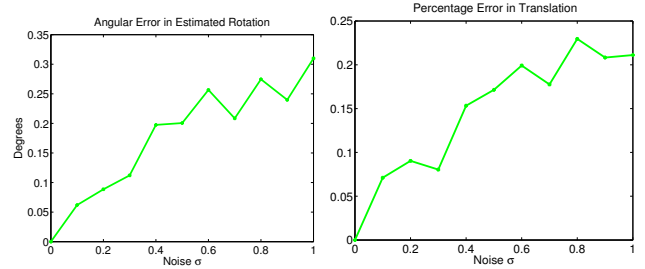


Figure 7. Simulation results for multiple refractive spheres assuming known sphere radii. Error in checkerboard rotation and translation.

55 mm Canon zoom lens. The lens was set to the maximum zoom setting. The camera was internally calibrated off-line using [4]. To avoid defocus issues, we use a small ( $f/10$ ) aperture.

**Refractive Spheres:** Figure 8 shows a camera looking through four acrylic spheres<sup>3</sup> of radius 0.5 inch and refractive index 1.49 at a checkerboard. Notice the checkerboard distortion due to refraction. Checkerboard corners were detected using [20]. To obtain the ground truth (GT) pose of the checkerboard, we took another photo (not shown) by directly viewing the checkerboard. Table 1 compares the ground truth pose with the estimated pose using our approach. Figure 8 also shows the detected corners (red) and re-projected corners as well as the sphere boundary (green) using the estimated calibration parameters. The final root mean square (RMS) reprojection error was 2.35 pixels.

**Spherical Mirrors:** In Figure 9, the camera looks at the scene reflected from four spherical mirrors<sup>4</sup> of radius 25.4 mm (1 inch) each. Notice the checkerboard distortion due to mirror reflection.

For the case of reflection, the camera does not view the checkerboard directly. To obtain the ground truth checkerboard pose, we need to estimate its extrinsic without a direct view [21, 10]. We use the recent method proposed by Takahashi *et al.* [25], which requires capturing three images of the checkerboard using a planar mirror in different orientations. The algorithm in [25] estimates the planar mirror poses as well as the checkerboard pose. Table 2 compares

<sup>3</sup>Available at plasticballs.com for 80 cents each.

<sup>4</sup>We use off-the-shelf stainless steel balls from McMaster-Carr.

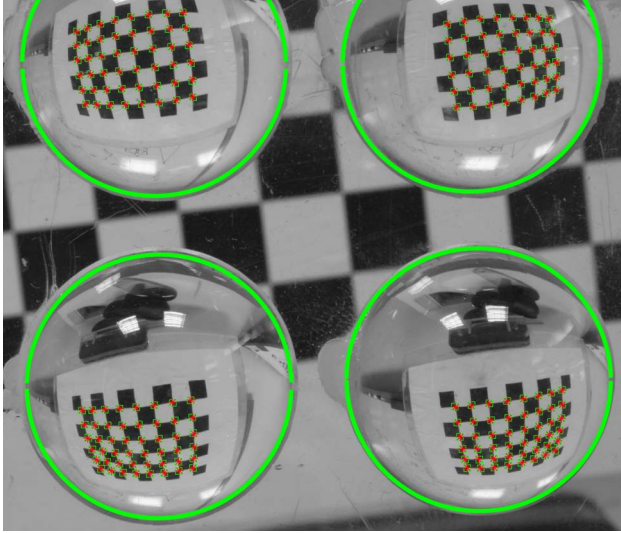


Figure 8. Real results on refractive spheres. Detected corners (red) and re-projected checkerboard points/sphere boundary (green) using estimated calibration parameters are shown.

|      | $\theta_x, \theta_y, \theta_z$ (degree) | $t_x, t_y, t_z$ (mm)   |
|------|---|------------------------|
| GT   | 176.22, -4.77, 83.33                    | -77.95, -89.49, 420.01 |
| Ours | 175.32, -3.00, 83.30                    | -79.24, -88.02, 422.57 |

Table 1. Comparison of checkerboard pose obtained using our method with ground truth (GT) pose for Figure 8. GT was computed by directly looking at the checkerboard. Rotation error equals  $1.98^\circ$  and the normalized translation error is 0.74%.

the ground truth pose (as estimated using [25]) with the estimated pose using our approach. The rotation error is  $1.2^\circ$  and the translation error is 5.28%. The estimated radii of mirrors were 25.54, 25.52, 25.45 and 25.65 mm, with a maximum error of 0.97%. Figure 8 also shows the detected corners (red) and the re-projected corners as well as the sphere boundary (green) using the estimated calibration parameters. Notice that the estimated contour matches well with the actual sphere boundary. The final RMS reprojection error after calibration was 0.5 pixels.

Figure 10 shows an example with two mirrors of different size (radius 12.7 and 25.4 mm respectively). Notice that the mirrors are not arranged on a grid. The estimated radii were 12.27 and 24.43 mm and the final RMS reprojection error was 0.48 pixels. Table 3 shows the estimated checkerboard pose along with the ground truth pose estimated using [25].

**Analysis:** The stability of pose estimation (Section 4.1) depends on axes being sufficiently non-parallel. If the angles between the underlying axes are small, the linear estimation of pose can become unstable. This could happen if  $d_i \gg r_i$ . For the above real experiments, the angle between the axes was only  $\approx 10^\circ$ , indicating that the pose estimation is quite stable.

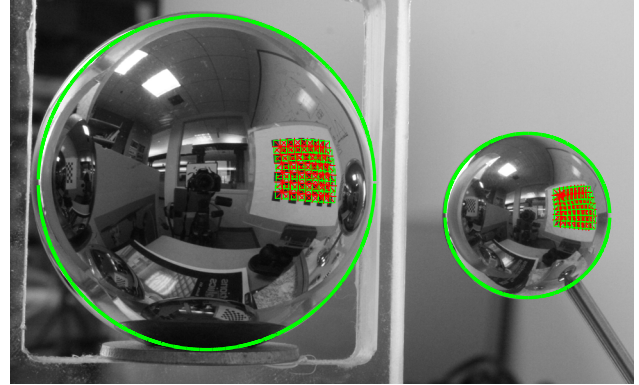


Figure 10. Our approach can handle mirrors of different size arranged in any fashion, not necessarily on a grid. Captured photo with detected (red) and re-projected checkerboard points (green) after calibration. Sphere boundary (green) using estimated parameters is also shown.

|      | $\theta_x, \theta_y, \theta_z$ (degree) | $t_x, t_y, t_z$ (mm)  |
|------|---|-----------------------|
| GT   | -144.53, 4.06, -94.12                   | 213.50, 61.29, 254.74 |
| Ours | -143.58, 4.45, -93.91                   | 226.32, 63.01, 242.43 |

Table 2. Comparison of checkerboard pose obtained using our method with ground truth (GT) pose for Figure 9. GT was obtained using the algorithm in [25]. Rotation error equals  $1.02^\circ$  and the normalized translation error is 5.28%.

|      | $\theta_x, \theta_y, \theta_z$ (degree) | $t_x, t_y, t_z$ (mm)  |
|------|---|-----------------------|
| GT   | -118.86, 1.38, -81.22                   | 196.34, 27.58, 307.81 |
| Ours | -118.87, 1.37, -81.08                   | 197.02, 27.28, 297.50 |

Table 3. Comparison of checkerboard pose obtained using our method with ground truth (GT) pose for Figure 10. GT was obtained using the algorithm in [25]. Rotation error equals  $0.14^\circ$  and the normalized translation error is 2.8%.

## 9. Discussions and Conclusions

We have presented a single image based calibration approach for *multi-axial* systems, consisting of a camera looking at multiple spherical mirrors or multiple refractive spheres. We believe that ours is the first algorithm that allows calibrating multiple refractive spheres based imaging system using a single photo.

We also showed an interesting theoretical result that full 6-DOF pose estimation of such multi-axial systems can be done in a semi-calibrated setting (without the knowledge of complete calibration), by choosing rays from two or more of the underlying axial systems. In general, three 2D-3D correspondences are required for pose estimation, for both central and non-central cameras [17] with complete calibration, since each correspondence provides two constraints. However, the coplanarity constraint is a single constraint for each correspondence. Therefore, pose estimation for multi-axial systems in semi-calibrated setting requires a minimum of six 2D-3D correspondences, instead of three. Nevertheless, this could be useful for localization, navigation and SLAM based approaches using wide-angle catadioptric sensors.

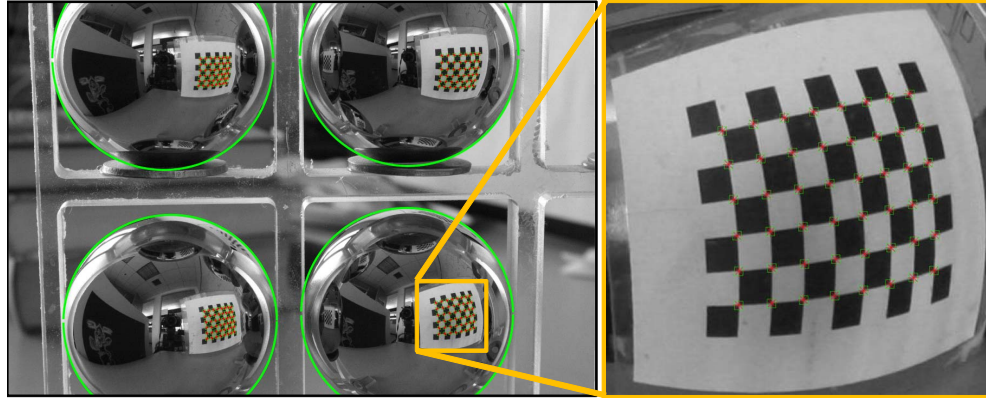


Figure 9. Real results on spherical mirrors. (Left) Setup. (Right) Captured photo with detected (red) and re-projected checkerboard points (green) after calibration. Sphere boundary (green) using estimated parameters is also shown.

## References

- [1] A. Agrawal, S. Ramalingam, Y. Taguchi, and V. Chari. A theory of multi-layer flat refractive geometry. In *CVPR*, pages 3346–3353, June 2012.
- [2] 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.
- [3] S. Baker and S. Nayar. A theory of single-viewpoint catadioptric image formation. *Int'l J. Computer Vision*, 35(2):175–196, Nov. 1999.
- [4] J.-Y. Bouguet. Camera calibration toolbox for matlab. <http://www.vision.caltech.edu/bouguetj>.
- [5] V. Caglioti, P. Taddei, G. Boracchi, S. Gasparini, and A. Giusti. Single-image calibration of off-axis catadioptric cameras using lines. In *OMNIVIS*, Oct. 2007.
- [6] M. Demazure. Sur deux problèmes de reconstruction. Technical Report 882, INRIA, 1988.
- [7] Y. Ding, J. Yu, and P. Sturm. Multi-perspective stereo matching and volumetric reconstruction. In *ICCV*, pages 1827–1834, Sept. 2009.
- [8] 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.
- [9] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [10] J. A. Hesch, A. I. Mourikis, and S. I. Roumeliotis. Extrinsic camera calibration using multiple reflections. In *ECCV*, pages 311–325, Sept. 2010.
- [11] 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, 2006.
- [12] Y. Kojima, R. Sagawa, T. Echigo, and Y. Yagi. Calibration and performance evaluation of omnidirectional sensor with compound spherical mirrors. In *OMNIVIS*, 2005.
- [13] D. Lanman, D. Crispell, M. Wachs, and G. Taubin. Spherical catadioptric arrays: Construction, multi-view geometry, and calibration. In *3DPVT*, pages 81–88, 2006.
- [14] 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.
- [15] S. Nayar. Sphereo: Determining depth using two specular spheres and a single camera. In *SPIE Conf. Optics, Illumination, and Image Sensing for Machine Vision III*, pages 245–254, Nov. 1988.
- [16] K. Nishino and S. Nayar. The World in an Eye. In *CVPR*, volume I, pages 444–451, June 2004.
- [17] D. Nistér and H. Stewénius. A minimal solution to the generalised 3-point pose problem. *J. Mathematical Imaging and Vision*, 27(1):67–79, 2007.
- [18] S. Ramalingam, P. Sturm, and S. K. Lodha. Theory and calibration algorithms for axial cameras. In *ACCV*, 2006.
- [19] R. Sagawa, N. Kurita, T. Echigo, and Y. Yagi. Compound catadioptric stereo sensor for omnidirectional object detection. In *IROS*, volume 3, pages 2612–2617, Sept. 2004.
- [20] D. Scaramuzza. Ocamcalib: Omnidirectional camera calibration toolbox for matlab. <https://sites.google.com/site/scarabotix/ocamcalib-toolbox>.
- [21] P. Sturm and T. Bonfort. How to compute the pose of an object without a direct view. In *ACCV*, pages 21–31, 2006.
- [22] P. Sturm and S. Ramalingam. A generic concept for camera calibration. In *ECCV*, 2004.
- [23] P. Sturm, S. Ramalingam, J.-P. Tardif, S. Gasparini, and J. Barreto. Camera models and fundamental concepts used in geometric computer vision. *Foundations and Trends in Computer Graphics and Vision*, 6(1-2):1–183, 2011.
- [24] 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.
- [25] K. Takahashi, S. Nobuhara, and T. Matsuyama. A new mirror-based extrinsic camera calibration using an orthogonality constraint. In *CVPR*, June 2012.
- [26] J.-P. Tardif, P. Sturm, M. Trudeau, and S. Roy. Calibration of cameras with radially symmetric distortion. *IEEE Trans. Pattern Anal. Machine Intell.*, 31(9):1552–1566, 2009.
- [27] S. Thirthala and M. Pollefeys. Multi-view geometry of 1D radial cameras and its application to omnidirectional camera calibration. In *ICCV*, volume 2, pages 1539–1546, 2005.