

CAMERA CALIBRATION WITH SPHERES: LINEAR APPROACHES

Hui Zhang, Guoqiang Zhang, Kwan-Yee Kenneth Wong

Department of Computer Science, The University of Hong Kong, Pokfulam Rd, Hong Kong
{hzhang, gqzhang, kykwong}@cs.hku.hk

ABSTRACT

This paper addresses the problem of camera calibration from spheres. By studying the relationship between the dual images of spheres and that of the absolute conic, a linear solution has been derived from a recently proposed non-linear semi-definite approach. However, experiments show that this approach is quite sensitive to noise. In order to overcome this problem, a second approach has been proposed, where the orthogonal calibration relationship is obtained by regarding any two spheres as a surface of revolution. This allows a camera to be fully calibrated from an image of three spheres. Besides, a *conic homography* is derived from the imaged spheres, and from its eigenvectors the orthogonal invariants can be computed directly. Experiments on synthetic and real data show the practicality of such an approach.

1. INTRODUCTION

Camera calibration, the determination of the camera intrinsic parameters, is traditionally achieved by imaging some special pattern with known metric structure [1][2]. However, these approaches involve the design and use of highly accurate tailor-made calibration patterns, which are often difficult and expensive to be manufactured. Apart from this, multi view vision systems are becoming more and more cost effective, and calibration of such a large number of cameras with a pattern is tedious and cumbersome. Besides, it is also very difficult to calibrate all the cameras simultaneously as the points on the calibration pattern may not be all visible simultaneously in all views. To overcome these difficulties, it is desirable to have some common simple object as the calibration target. In [3], Wong et. al. proposed to use a surface of revolution (SOR) to calibrate a camera. The silhouettes of SOR can be extracted more reliably compared with points and they facilitate more precise camera calibration. In addition, as long as the SOR is placed in the common view field of the cameras, its occluding contours are always visible from any position. Hence it can be used to accurately calibrate multiple cameras mounted at arbitrary locations simultaneously.

Sphere, a special type of SOR, is used in this paper as the calibration object. It has first been used in [4] to com-

pute the aspect ratio of the two image axes. Later, Daucher et. al. [5] introduced a multi-step nonlinear approach to estimate four camera parameters using spheres. More recently, Teramoto and Xu [6] related the absolute conic with three images of spheres and calibrated the camera by minimizing some geometric errors nonlinearly. Agrawal [7] derived similar constraints in the dual space, and the parameters were solved by minimizing some algebraic errors using semi-definite programming. Note that all the above approaches involve a nonlinear procedure which can result in a low calibration speed. In this paper, two novel linear approaches, namely the scalar and orthogonal approaches, have been proposed to overcome the above problems. The scalar approach is obtained by eliminating the imaged sphere centers in the calibration constraints in [7] so that the nonlinear semi-definite minimization is avoided. The orthogonal approach is inspired by [3], where any combination of two spheres can be regarded as a surface of revolution with its rotation axis uniquely defined by the joint line of the two sphere centers. The orthogonal constraints [8] can then be used to calibrate the camera. Besides, a *conic homography* can be derived from the imaged spheres and from its eigenvectors, the orthogonal invariants can be computed directly.

This paper is organized as follows. § 2 presents the fundamentals for camera calibration from the absolute conic. § 3 relates the dual image of a sphere to that of the absolute conic. § 4 introduces two novel linear approaches for camera calibration from spheres. § 5 shows the results of synthetic/real experiments. Conclusions are given in § 6.

2. CALIBRATION WITH THE ABSOLUTE CONIC

One advantage of direct calculation of the image of the absolute conic (IAC) is that it allows a linear calibration. The absolute conic was first introduced by Faugeras et al. [9]. It is a point conic on the plane at infinity that is invariant to rotation and translation as a set. Given the camera calibration matrix \mathbf{K} , it projects to the IAC as

$$\omega = \mathbf{K}^{-T} \mathbf{K}^{-1} = \begin{bmatrix} \omega_1 & \omega_3 & \omega_4 \\ \omega_3 & \omega_2 & \omega_5 \\ \omega_4 & \omega_5 & \omega_6 \end{bmatrix}. \quad (1)$$

The IAC, and its dual (DIAC) $\omega^* = \mathbf{K}\mathbf{K}^T$ [8] are imaginary point and line conics respectively, which have no real point lying on them. The camera matrix \mathbf{K} can be easily obtained from the IAC or DIAC by Cholesky decomposition [10].

The image of the absolute conics can be found by making use of the orthogonal constraints. The vanishing point \mathbf{v} of the normal direction to a plane and the vanishing line \mathbf{l} of the plane satisfy the pole-polar relationship w.r.t. the image of absolute conic ω [8], given by

$$\mathbf{l} = \omega \mathbf{v}. \quad (2)$$

Thus to fully calibrate a camera at least three such conjugate pairs are needed. To calibrate a zero skew/natural camera, at least two are needed.

3. THE APPARENT CONTOUR OF A SPHERE AND ITS DUAL

Without loss of generality, consider a camera $\mathbf{P} = \mathbf{K}[\mathbf{I}_3 | \mathbf{0}]$ viewing a sphere $\hat{\mathbf{Q}}$ centered at the Z -axis (see Fig.1(a)). The limb points $\mathbf{X} = [X_0 \cos \theta \ X_0 \sin \theta \ \gamma X_0 \ 1]^T$ of the sphere always form a planar circle \mathbf{C}_3 . The image points $\hat{\mathbf{x}}$ (see Fig.1(b)) of \mathbf{X} under \mathbf{P} can be defined as

$$\hat{\mathbf{x}} = \mathbf{K}[\mathbf{I} | \mathbf{0}] \begin{bmatrix} X_0 \cos \theta \\ X_0 \sin \theta \\ \gamma X_0 \\ 1 \end{bmatrix} = X_0 \mathbf{K} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} \cos \theta \\ \sin \theta \\ 1 \end{bmatrix}, \quad (3)$$

where $[X_0 \ 0 \ \gamma X_0 \ 1]$ is the generating point of the circle \mathbf{C}_3 . Since the point $\mathbf{X}_u = [\cos \theta \ \sin \theta \ 1]^T$ lies on the unit circle $\mathbf{C}_u = \text{diag}\{1, 1, -1\}$, the homography $\hat{\mathbf{H}} = \mathbf{K} \text{diag}\{1, 1, \gamma\}$ transforms \mathbf{C}_u to the image of \mathbf{C}_3 as $\hat{\mathbf{C}} = \hat{\mathbf{H}}^{-T} \mathbf{C}_u \hat{\mathbf{H}}^{-1}$. Now consider the sphere rotates to an arbitrary position about the camera center by a rotation \mathbf{R} . Let $\mathbf{H} = \mathbf{K}\mathbf{R} \text{diag}\{1, 1, \gamma\}$, the image of the sphere is thus given by

$$\mathbf{C} = \mathbf{H}^{-T} \mathbf{C}_u \mathbf{H}^{-1}. \quad (4)$$

In the dual space, the dual of \mathbf{C} is given by

$$\begin{aligned} \mathbf{C}^* &= \mathbf{K}\mathbf{R} \text{diag}\{1, 1, -\gamma^2\} \mathbf{R}^T \mathbf{K}^T \\ &= \mathbf{K}\mathbf{R}(\mathbf{I} + \text{diag}\{0, 0, -(\gamma^2 + 1)\}) \mathbf{R}^T \mathbf{K}^T \\ &= \mathbf{K}\mathbf{K}^T - (\gamma^2 + 1) \mathbf{K} \mathbf{r}_3 \mathbf{r}_3^T \mathbf{K}^T \\ &= \mathbf{K}\mathbf{K}^T - \mathbf{o}\mathbf{o}^T, \end{aligned} \quad (5)$$

where \mathbf{r}_3 is the third column of the rotation matrix \mathbf{R} and $\mathbf{o} = \sqrt{\gamma^2 + 1} \mathbf{K} \mathbf{r}_3$ is the image of the sphere center $\mathbf{R}[0 \ 0 \ Z \ 1]^T$ under \mathbf{P} . This result coincides with those derived in [7].

Note there exists a scalar for each sphere image, i.e.,

$$\begin{aligned} \beta_1 \mathbf{C}_1^* &= \omega^* - \mathbf{o}_1 \mathbf{o}_1^T \\ \beta_2 \mathbf{C}_2^* &= \omega^* - \mathbf{o}_2 \mathbf{o}_2^T \\ \beta_3 \mathbf{C}_3^* &= \omega^* - \mathbf{o}_3 \mathbf{o}_3^T. \end{aligned} \quad (6)$$

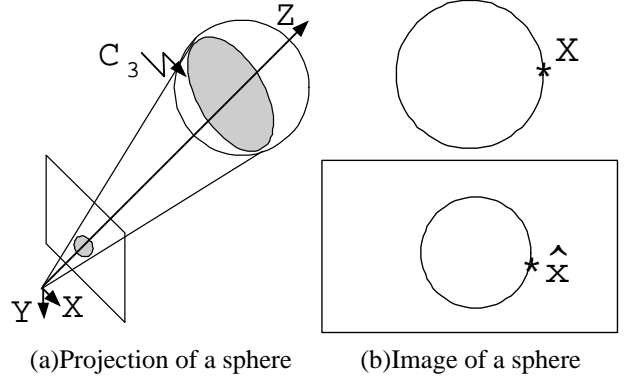


Fig. 1. A sphere is viewed by a camera $\mathbf{P} = \mathbf{K}[\mathbf{I}_3 | \mathbf{0}]$.

It can be easily derived that the ratios of the scalars β_j/β_i are the eigenvalues of $\mathbf{C}_j \mathbf{C}_i^*$ ($i, j = 1, 2, 3$ and $i \neq j$). The image of each sphere center \mathbf{o} can then be obtained by a linear combination of the first two eigenvectors of $\mathbf{C}_j \mathbf{C}_i^*$ [7]. Note although β_j/β_i and β_k/β_i can be chosen to be consistent with β_j/β_k ($i, j, k = 1, 2, 3$ and $i \neq j \neq k$), multiple choices exist in practise. The correct one can be retained by the heuristic that all the three resulting imaged sphere centers should lie within their corresponding apparent contours. Note there can even be no solution when noise is too large.

4. LINEAR CALIBRATION APPROACHES

4.1. Scalar Approach

The fundamental idea is to eliminate the images of the sphere centers in (6) so that apart from ω^* , only three unknowns remain, i.e., the scalars β_1, β_2 and β_3 . The equation (6) can be rewritten as

$$\omega^* - \beta_i \mathbf{C}_i^* = \mathbf{o}_i \mathbf{o}_i^T = \begin{bmatrix} o_{i1}^2 & o_{i1}o_{i2} & o_{i1}o_{i3} \\ o_{i1}o_{i2} & o_{i2}^2 & o_{i2}o_{i3} \\ o_{i1}o_{i3} & o_{i2}o_{i3} & o_{i3}^2 \end{bmatrix}, \quad (7)$$

where $\mathbf{o}_i = [o_{i1} \ o_{i2} \ o_{i3}]^T$ ($i = 1, 2, 3$) is the image of the i th sphere center. By observing the elements on the right hand side, three quadratic equations can be derived as

$$\begin{aligned} o_{i1}^2 o_{i2}^2 &= (\omega_{11}^* - \beta_i c_{i11}^*)(\omega_{22}^* - \beta_i c_{i22}^*) = (\omega_{12}^* - \beta_i c_{i12}^*)^2 \\ o_{i1}^2 o_{i3}^2 &= (\omega_{11}^* - \beta_i c_{i11}^*)(\omega_{33}^* - \beta_i c_{i33}^*) = (\omega_{13}^* - \beta_i c_{i13}^*)^2 \\ o_{i2}^2 o_{i3}^2 &= (\omega_{22}^* - \beta_i c_{i22}^*)(\omega_{33}^* - \beta_i c_{i33}^*) = (\omega_{23}^* - \beta_i c_{i23}^*)^2. \end{aligned} \quad (8)$$

From each combination of two conics, three linear equations can be obtained by eliminating the second order components of ω^* ,

$$\begin{aligned} a_2 \omega_{11}^* + a_1 \omega_{22}^* - 2b_{12} \omega_{12}^* - c_{12} &= 0 \\ a_3 \omega_{11}^* + a_1 \omega_{33}^* - 2b_{13} \omega_{13}^* - c_{13} &= 0 \\ a_3 \omega_{22}^* + a_2 \omega_{33}^* - 2b_{23} \omega_{23}^* - c_{23} &= 0, \end{aligned} \quad (9)$$

where $a_m = \beta_i C_{imm}^* - \beta_j C_{jmm}^*$, $b_{mn} = \beta_i C_{imn}^* - \beta_j C_{jmn}^*$, $c_{mn} = (\beta_i^2 C_{imn}^{*2} - \beta_j^2 C_{jmn}^{*2}) - (\beta_i^2 C_{imm}^* C_{inn}^* - \beta_j^2 C_{jmm}^* C_{jnn}^*)$ ($m, n=1, 2, 3$). Note in practise, however, errors in the estimated scalar β_i can cause instability in calibration. Another novel approach is then proposed to overcome this problem.

4.2. The Orthogonal Approach

The orthogonal relationship in (2) can be directly obtained from the image of two spheres.

Proposition I From images C_1 and C_2 of two circles, a homography $H_c = C_2 C_1^*$, termed as the conic homography, can be obtained so that its eigenvectors give a line of fixed points (axis) and a fixed point (vertex) not on the line.

Due to space restrictions, its proof for the general cases is omitted here, while the proof for the special case when the camera is viewing two spheres is given as follows.

Proof. For each pair of conics C_i^* and C_j^* ($i \neq j$), multiplying the line $\mathbf{l}_{ij} = \mathbf{o}_i \times \mathbf{o}_j$ which joins the images of the two sphere centers to both sides of (6) gives

$$\begin{aligned} \beta_i C_i^* (\mathbf{o}_i \times \mathbf{o}_j) &= (\omega^* - \mathbf{o}_i \mathbf{o}_i^T) (\mathbf{o}_i \times \mathbf{o}_j) \\ \beta_i C_i^* \mathbf{l}_{ij} &= \omega^* \mathbf{l}_{ij} = \mathbf{v}_{ij}. \end{aligned} \quad (10)$$

Here the axis \mathbf{l}_{ij} is the vanishing line of the plane passing through the camera center and the two sphere centers, the vertex \mathbf{v}_{ij} is the vanishing point of the plane normal direction. Consider also the image of C_j^* , it can be derived that

$$\beta_i C_i^* \mathbf{l}_{ij} - \beta_j C_j^* \mathbf{l}_{ij} = \omega^* \mathbf{l}_{ij} - \omega^* \mathbf{l}_{ij} = 0 \quad (11)$$

Hence \mathbf{l}_{ij} can be computed as the eigenvector of $C_j C_i^*$ with the eigenvalue β_j / β_i , and it can also be easily proved that \mathbf{v}_{ij} is the intersection of the two remaining eigenvectors.

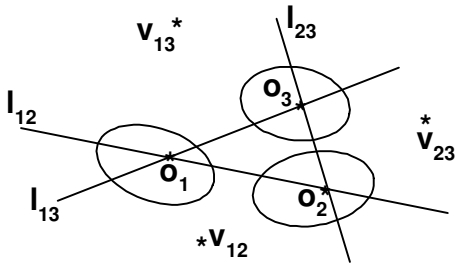


Fig. 2. Given three spheres, three SORs can be formed to give three pairs of axes and vertices.

From (2), two sphere images provide 2 linear constraints on the IAC so that from three sphere images, 6 constraints are obtainable for fully calibrating the camera (see Fig.2). Note alternatively the image of the sphere center can be regarded as the vanishing point of the plane formed by the

sphere limb points. However, it is easy to see that the resulting constraints are dependent on those above. Similarly, the constraints obtained from the scalar approaches can neither provide additional constraints. Hence when the number of spheres reduces to 2, a camera with more than 2 unknown parameters cannot be calibrated.

Note when the line joining two sphere centers passes through the camera center, the sphere limb points become concentric so that only 4 independent constraints can be obtained. Hence a camera with only four parameters, i.e., zero skew or unit aspect ratio, can be recovered. When the centers of the three spheres are collinear, only two constraints can be obtained and the camera cannot be calibrated. Additionally, increasing the number of the spheres can increase the precision of calibration only when any four sphere centers are not co-planar.

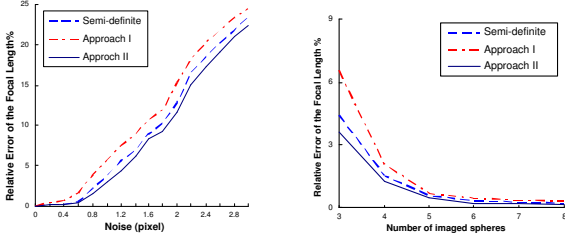
5. EXPERIMENTS AND RESULTS

5.1. Synthetic Data

The synthetic camera has fixed intrinsic parameters, with focal length $\alpha_x = 880$, $\alpha_y = 800$, skew $s = 0.1$ and principle point $(u_0, v_0) = (320, 240)$. The points on the apparent contour of each sphere are corrupted with Gaussian noises of 16 different levels from 0 to 3 pixels, and a conic is then fitted to the noisy points by some classical algorithms [11].

Given 3 sphere images, the first experiment is to calibrate the camera under different noise levels. For each level, 100 independent trials are performed using the scalar and orthogonal approaches, as well as Agrawal's semi-definite method. Fig.3(a) shows the average percentage errors of the focal length. The errors of the other parameters, which are not shown here, exhibit similar trend. It can be seen that the errors increase linearly with the noise level. Besides, the orthogonal approach has a better precision compared with the other two. From Fig.3(a), the scalar approach performs the worst since the estimated scalars are seriously affected by noises. Table 1.(a) shows the estimated parameters under the noise level of 1 pixel.

In the second experiment, the camera is calibrated with different numbers of spheres, from 3 to 10, under Gaussian noise of 1 pixel. Note the number of sphere combination does not increase linearly with the number of spheres N . It is in fact its combination of 3, i.e. $N C_3$. For each number of spheres used, 100 independent trials are performed using the scalar and orthogonal approaches, as well as Agrawal's semi-definite method. Fig.3(b) shows the average percentage errors of the focal length. It can be seen that the errors decrease to a level below 1% when the number of spheres attains four. Hence four spheres can be used to precisely calibrate the camera. Table 1.(b) shows the estimated parameters from four spheres under the noise level of 1 pixel.



(a) Under 16 noise levels (b) From different number of spheres

Fig. 3. Relative errors of the focal length.

Table 1. Estimated camera parameters from images of (a) Three spheres

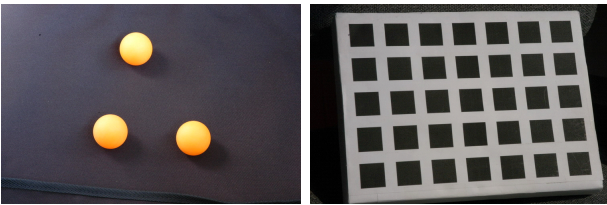
| Settings | α_x | α_y | s | (u_0, v_0) |
|---------------|------------|------------|-------|------------------|
| Semi-definite | 838.16 | 766.07 | 2.38 | (334.94, 238.52) |
| Scalar | 819.24 | 746.01 | -1.18 | (329.28, 261.05) |
| Orthogonal | 838.58 | 768.48 | 1.32 | (324.97, 242.24) |

(b) Four spheres

| Settings | α_x | α_y | s | (u_0, v_0) |
|---------------|------------|------------|------|------------------|
| Semi-definite | 906.55 | 823.70 | 1.24 | (314.47, 237.05) |
| Scalar | 869.93 | 790.77 | 1.41 | (323.01, 228.82) |
| Orthogonal | 896.36 | 812.58 | 1.24 | (318.83, 239.28) |

5.2. Real Scene

In the real scene experiment, an image of three pingpong balls (see Fig.4(a)) is taken with a Nikon100D CCD camera. The image resolution is 1505×1000 . The Canny edge detector [12] is first applied to find the points on the apparent contours of the spheres, to which conics are fitted with a least square approach [11]. The camera is calibrated with the proposed scalar and orthogonal approaches and the results are compared with those from Agrawal's semi-definite approach. The estimated parameters are listed in Table 2, where the result from the classical method of Zhang [2] is taken as the ground truth. Fig.4(b) shows the calibration pattern used with the Zhang's calibration method. From Table 2, it can be seen that the orthogonal approach performs the best while the scalar approach the worst.



(a) Image of three spheres (b) Image of planar calibration pattern

Fig. 4. Images used for calibration.

Table 2. Camera parameters estimated from the pingpong ball image with different approaches.

| Settings | α_x | α_y | s | (u_0, v_0) |
|---------------------|------------|------------|------|------------------|
| Zhang(ground truth) | 2721.5 | 2722.3 | 0.68 | (769.16, 504.38) |
| Semi-definite | 2673.1 | 2672 | 1.87 | (776.86, 486.98) |
| Scalar | 2460.7 | 2461.7 | 6.27 | (795.05, 488.54) |
| Orthogonal | 2758.1 | 2759.8 | 0.36 | (781.08, 491.83) |

6. CONCLUSIONS

This paper have proposed two calibration algorithms, namely the scalar and orthogonal approaches, by making use of the apparent contours of at least three spheres in a single image. Only linear procedures are involved, and the results are similar to or even better than those of the previous approaches. Their solutions can be used as a starting point for a maximum likelihood estimate which minimizes the geometric error between the measured edgels and the conic. Note the performance of calibration could be poor if the spheres are imaged near the image centers. However, the key limitation of previous approaches, namely the quality of the conic detection strongly affecting the accuracy of calibration results, can be alleviated using the orthogonal approach.

7. REFERENCES

- [1] Tsai R., "A versatile camera calibration technique for high accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. 3(4), pp. 323–344, 1987.
- [2] Zhang Z., "A flexible new technique for camera calibration," *IEEE Transactions on PAMI*, vol. 22(11), pp. 1330–1334, 2000.
- [3] Wong K.-Y. K., Mendonça P. R. S., and Cipolla R., "Camera calibration from surfaces of revolution," *IEEE Trans. on PAMI*, vol. 25(2), pp. 147–161, February 2003.
- [4] Penna M. A., "Camera calibration: A quick and easy way to determine the scale factor," *IEEE Trans. on PAMI*, vol. 13, no. 12, pp. 1240–1245, 1991.
- [5] Daucher D., Dhome M., and Lapreste J., "Camera calibration from spheres images," in *Proc. of 3rd ECCV*, 1994, pp. 449–454.
- [6] Teramoto H. and Xu G., "Camera calibration by a single image of balls: From conics to the absolute conic," in *Proc. of 5th ACCV*, 2002, pp. 499–506.
- [7] Agrawal M. and Davis L. S., "Camera calibration using spheres: A semi-definite programming approach," in *Proceedings of IEEE International Conf. on Computer Vision*, 2003.
- [8] Hartley R.I. and Zisserman A., *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, UK, 2000.
- [9] Faugeras O.D., Luong Q.-T., and Maybank S.J., "Camera self-calibration: Theory and experiments," pp. 321–334, 1992.
- [10] Gentle J.E., *Numerical Linear Algebra for Applications in Statistics*, Springer-Verlag, Berlin, 1998.
- [11] Fitzgibbon A. W., Pilu M., and Fisher R. B., "Direct least-squares fitting of ellipses," *IEEE Trans. on PAMI*, vol. 21, 1999.
- [12] Canny J., "A computational approach to edge detection," *IEEE Trans. on PAMI*, vol. 8(6), pp. 679–698, November 1986.