A Stratified Approach for Camera Calibration Using Spheres

Kwan-Yee K. Wong, *Member, IEEE,* Guoqiang Zhang, *Student-Member, IEEE*
and Zhihu Chen, *Non-Member, IEEE*

Abstract

This paper proposes a stratified approach for camera calibration using spheres. Previous works have exploited epipolar tangents to locate frontier points on spheres for estimating the epipolar geometry. It is shown in this paper that other than the frontier points, two additional point features can be obtained by considering the bitangent envelopes of a pair of spheres. A simple method for locating the images of such point features and the sphere centers is presented. An algorithm for recovering the fundamental matrix in a plane plus parallax representation using these recovered image points and the epipolar tangents from three spheres is developed. A new formulation of the absolute dual quadric as a cone tangent to a dual sphere with the plane at infinity being its vertex is derived. This allows the recovery of the absolute dual quadric, which is used to upgrade the weak calibration to a full calibration. Experimental results on both synthetic and real data are presented, which demonstrate the feasibility and the high precision achieved by our proposed algorithm.

Index Terms

Camera calibration, sphere, absolute dual quadric, epipolar tangent.

I. INTRODUCTION

Many computer vision algorithms exploit information from multiple cameras to resolve ambiguities arising from the loss of depth information in the imaging process. Examples include 3D model reconstruction [1]–[4], image-based rendering [5], [6], and traffic surveillance [7]. The successes of these algorithms often depend on the accuracy of the knowledge of both the internal (i.e, focal length, aspect ratio, and principal point) and external (i.e., rotation and translation) parameters of the cameras. The process of recovering the internal and external parameters of a camera is known as camera calibration, and it usually involves three major steps: i) taking images of some specially designed

The authors are with the Department of Computer Science, The University of Hong Kong, Pokfulam Road, Hong Kong. Email: {kykwong,gqzhang,zhchen}@cs.hku.hk

patterns with known metric geometry or with some special structures; ii) extracting the pattern features from the images; and iii) estimating the camera parameters using constraints derived from the extracted features. Calibrating multiple cameras is in general much more tedious and cumbersome than calibrating a single camera. Due to the distribution and differences in orientations of the cameras, it is often extremely difficult, if not impossible, to make the calibration pattern simultaneously visible in all views. One way to tackle this problem is to divide the cameras into smaller groups such that all cameras within the same group can see the calibration pattern. Each group of cameras is then calibrated separately, and finally all camera groups are registered into a common reference coordinate system.

Calibration methods proposed in the literature exhibit a trade-off between geometric accuracy and flexibility. Classical calibration techniques use a highly accurate tailor-made 3D calibration grid to achieve a high calibration accuracy [8], [9]. The coefficients of a projection matrix are first computed by *direct linear transformation* (DLT), which are then refined iteratively in an optimization process that minimizes the projection errors of the 3D reference points. In [10], Zhang proposed a calibration algorithm that uses images of a planar grid pattern taken under unknown orientations. Experimental results suggested that his approach is comparable to those classical techniques in terms of accuracy, while the use of a planar calibration pattern allows a considerable amount of flexibility. Other than the planar grid pattern, circle is another common planar pattern that has been exploited for camera calibration. In [11], Chen et al. developed an algorithm for camera calibration using two coplanar circles. In [12] and [13], the authors studied the constraints provided by concentric circles for camera calibration. Other configurations that have been studied include parallel circles [14], [15] and coaxial circles [16]. Note that the use of a planar calibration pattern lessens the above mentioned simultaneous visibility problem in calibrating multiple cameras. However, those algorithms using a planar pattern usually require the pattern being observed in multiple distinct orientations so as to provide enough constraints for performing the calibration. This again will increase the chance of occurrence of the simultaneous visibility problem.

In this paper, we present a novel technique for calibrating multiple cameras using spheres as a calibration object. The proposed method requires a minimum of three spheres (or equivalently a single sphere placed at three distinct locations) and places no assumptions on the camera internal parameters. It is computationally efficient as it involves no nonlinear optimization. Due to the complete symmetry exhibited by a sphere, its silhouette is always visible by a camera in any arbitrary orientation provided that the sphere is inside the field of view of the camera. Hence the use of spheres as a calibration object avoids the simultaneous visibility problem often encountered in calibrating multiple cameras. In fact, the proposed algorithm can work even under partial occlusions of the spheres. Preliminary results of this work have been published in [17].

The remainder of this paper is organized as follows. Section II gives a survey of works in the literature using spheres for recovering the camera internal and external parameters. Section III briefly reviews the properties of a sphere and its projection in an image. Section IV introduces the special point features induced by a pair of spheres, and describes a simple method for locating their images as well as the imaged sphere centers from the conic silhouettes of the spheres. By exploiting such recovered imaged points and the epipolar tangents, an algorithm for estimating the epipolar geometry from three or more sphere silhouettes is presented in Section V. The work of achieving a metric reconstruction is addressed in Section VI. Section VII presents experimental results on epipolar geometry recovery and multiple camera calibration, followed by conclusions in Section VIII.

II. RELATED WORK

In the literature of camera calibration, there exists a large number of algorithms which exploit different patterns, structures, or motions to derive constraints on the camera parameters. In this section, we will focus our discussion on a group of calibration methods that use spheres as a calibration object. The use of spheres in camera calibration dates back to 1991, when Penna [18] presented a simple method for recovering the aspect ratio of the two image axes from an image of a sphere. Under the assumption of a zero-skew camera, Daucher et al. [19] later proposed a multi-step nonlinear approach for estimating four camera internal parameters from sphere images. In [20], Teramoto and Xu derived a formula relating the conic silhouette of a sphere to the image of the absolute conic, and presented a nonlinear method for estimating the camera internal parameters by minimizing the reprojection error. Their method was further elaborated by Agrawal and Davis [21] who formulated the problem in the dual space and presented a method to recover the camera internal parameters using semi-definite programming. In [22], Ying and Zha used the same constraints as presented in [20] and [21], and developed two linear approaches based on the geometric interpretation and special structure in the algebra of the formulation respectively. In [23] and [24], Zhang et al. introduced the *conic homography* computed from the conic silhouettes of two spheres, and showed that the eigenvectors of such a homography can be used to derive a pole-polar constraint on the imaged absolute conic for camera calibration. Note that all the calibration techniques mentioned so far mainly focused on recovering the camera internal parameters. In [21] and [24], the authors also considered recovering the camera external parameters in multi-camera calibration. By assuming the radii of all spheres being equal, they first estimated the 3D positions of the sphere centers in each of the camera coordinate systems. They then obtained the relative rotation and translation between two cameras by registering the estimated sphere centers.

When treating spheres as special features (like 3D points, lines, and circles), important progresses have been made in the literature of weak camera calibration [25] (i.e., the recovery of fundamental

matrix for two views, or achieving a projective reconstruction from multiple views). In [26], Schulz-Mirbach and Weiss proposed that constraints on the epipolar geometry can be derived from conic correspondences induced by planar conics, but no experimental results were presented. By deriving a different form of constraints on the epipolar geometry, Kahl and Heyden [27] extended the work of Schulz-Mirbach and Weiss, and developed a method for recovering the fundamental matrix from conic correspondences induced by general quadrics. The minimum data required are four quadrics, with each providing two constraints from the two epipolar planes tangent to the quadric. Such constraints can be nicely derived from the analytic expression of the conic images of the quadric surfaces. Recently, Kaminski and Shashua proposed to recover the epipolar geometry from general planar and spatial curves [28]. Their contribution, however, is mainly theoretical.

This paper tackles the problem of recovering both the camera internal and external parameters from spheres. Note that two spheres together can be viewed as a surface of revolution, with the revolution axis passing through both sphere centers. The envelopes of the planes bitangent to the two spheres will form two cones, each with its vertex lying on the revolution axis (see Fig. 1). We will show that the images of these two vertices, and hence the image of the line passing through the sphere centers, can be computed directly from the conic silhouettes of the spheres. Given three spheres, there will be six such vertices, all lying on the plane passing through the sphere centers. By exploiting the image correspondences of these vertices and the sphere centers, and the point correspondences induced by the epipolar tangents [29], [30], the epipolar geometry can be recovered. The minimum number of spheres required is three, which is advantageous over the work in [27] which requires at least four spheres. By relating the absolute dual quadric to the spheres in the projective space, it will be shown that the weak calibration of the cameras can be easily upgraded to a full calibration by recovering the absolute dual quadric. Compared with our previous work [24], the method proposed in this paper does not require the assumption of the radii of all spheres being equal for recovering the camera external parameters. Besides, unlike the work in [24] which depends on the estimated internal parameters for reconstructing and registering the sphere centers, the proposed method estimates the internal parameters and the external parameters separately in two independent processes. This can improve the calibration accuracy by minimizing error accumulation. Experimental results demonstrate that our algorithm can be carried out easily and can achieve a high precision.

III. SPHERE AND ITS PROJECTION

Throughout this paper, we denote a scalar value using italic font (e.g., s), and a vector or matrix using bold font (e.g., $\mathbf{v}$, $\mathbf{M}$). The homogeneous coordinates of a point or line is denoted using a vector (e.g., $\mathbf{x}$), with its inhomogeneous counterpart indicated by a tilde (e.g., $\tilde{\mathbf{r}}$). The dual of a vector or matrix is denoted by an asterisk (e.g., $\mathbf{v}^*$, $\mathbf{M}^*$).
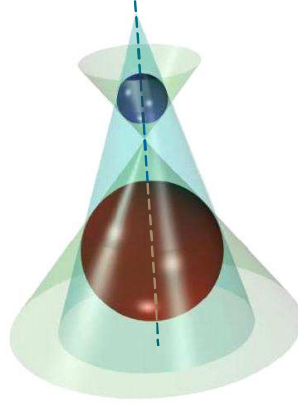
Fig. 1. Two spheres with the two bitangent cones. The vertices of the two cones lie on the revolution axis passing through the sphere centers.

In homogeneous coordinates, a quadric surface can be represented by a 4×4 symmetric matrix $\mathbf{Q}$. Any point $\mathbf{X}$ on the surface satisfies

$$\mathbf{X}^{\mathrm{T}}\mathbf{Q}\mathbf{X} = 0. \tag{1}$$

A dual quadric $\mathbf{Q}^*$ is an alternative expression of the quadric $\mathbf{Q}$ defined in terms of the planes $\mathbf{\Pi} = \mathbf{Q}\mathbf{X}$ tangent to the quadric surface. The corresponding equation of the dual quadric is

$$\mathbf{\Pi}^{\mathrm{T}}\mathbf{Q}^*\mathbf{\Pi} = 0, \tag{2}$$

where $\mathbf{Q}^*$ is the adjoint of $\mathbf{Q}$. If $\mathbf{Q}$ is invertible, $\mathbf{Q}^*$ is simply given by $\mathbf{Q}^{-1}$.

Sphere is a special kind of quadric, with the form

$$\mathbf{Q}_{\mathrm{s}} \sim \begin{bmatrix} \mathbf{I} & -\tilde{\mathbf{A}} \\ -\tilde{\mathbf{A}}^{\mathrm{T}} & \tilde{\mathbf{A}}^{\mathrm{T}}\tilde{\mathbf{A}} - r^2 \end{bmatrix}, \tag{3}$$

where the 3-vector $\tilde{\mathbf{A}}$ represents the sphere center in Cartesian coordinates and r is its radius. Note that the upper left 3×3 sub-matrix is an identity matrix. The sphere center $\mathbf{A}$ and the plane at infinity $\mathbf{\Pi}_\infty$ are in a pole-polar relationship w.r.t. $\mathbf{Q}_{\mathrm{s}}$, written as

$$\mathbf{\Pi}_\infty \sim \mathbf{Q}_{\mathrm{s}}\mathbf{A}. \tag{4}$$

Like the case of a circle which meets the line at infinity $\mathbf{L}_\infty$ at two circular points $\mathbf{I}$ and $\mathbf{J}$, a sphere intersects $\mathbf{\Pi}_\infty$ at the absolute conic $\mathbf{\Omega}$, as shown in Fig. 2.

The projection of a quadric $\mathbf{Q}$ under a camera with a projection matrix $\mathbf{P}$ is a conic $\mathbf{C}$ [9], formulated as
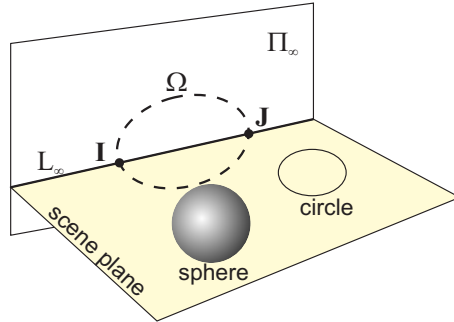
$$\mathbf{C} \sim (\mathbf{P}\mathbf{Q}^*\mathbf{P}^{\mathrm{T}})^*. \tag{5}$$

Fig. 2. A circle meets with the line at infinity $\mathbf{L}_\infty$ at two circular points $\mathbf{I}$ and $\mathbf{J}$, and a sphere intersects the plane at infinity $\mathbf{\Pi}_\infty$ at the absolute conic $\mathbf{\Omega}$.

Given a sphere with center $\tilde{\mathbf{A}}$ and radius r, its image under a camera with a projection matrix $\mathbf{P} = \mathbf{K}[\ \mathbf{I}\quad \mathbf{0}\ ]$ is therefore given by

$$\mathbf{C} \sim (\mathbf{K}\mathbf{K}^{\mathrm{T}} - (\mathbf{K}\tilde{\mathbf{A}}/r)(\mathbf{K}\tilde{\mathbf{A}}/r)^{\mathrm{T}})^*. \tag{6}$$

Hence the silhouette of a sphere in an image can be analytically expressed as a conic, or equivalently as a dual conic.

Given a set of fully calibrated cameras, a quadric can be reconstructed from its images [31]. Consider a non-degenerate quadric with nine degrees of freedom. From (5), it can be seen that one conic image places five constraints on the quadric. It appears that two conic images would provide enough constraints for the recovery of the quadric. Cross and Zisserman showed that, however, the ten constraints resulting from two conic images are linearly dependent, with two redundant ones [31]. Hence, at least three views are required to uniquely determine the quadric.

IV. POINT FEATURES INDUCED BY SPHERES

From (6), it is obvious that neither the imaged sphere center nor the image of any line passing through the sphere center can be determined from the image of a sphere. Nevertheless, we can address the problem from another point of view by considering two spheres together. Note that two spheres can be viewed as a surface of revolution, with the axis of revolution given by the line passing through the two sphere centers. The envelopes of the bitangent planes to the two spheres form two cones whose vertices lie on the axis passing through the sphere centers, as shown in Fig. 1.

The projections of these two cone vertices in the image can be computed directly from the conic silhouettes of the two spheres. Consider a camera viewing two spheres $\mathbf{Q}_1$ and $\mathbf{Q}_2$ with radii r_1 and r_2 respectively. Without loss of generality, let the centers of $\mathbf{Q}_1$ and $\mathbf{Q}_2$ be located at the world origin with coordinates $(\ 0\quad 0\quad 0\ )^{\mathrm{T}}$ and along the Z-axis with coordinates $(\ 0\quad 0\quad t\ )^{\mathrm{T}}$, respectively (see Fig. 3). Since the configuration of the two spheres is invariant under a rotation about the Z-axis, the

camera center can be set arbitrarily on the X-Z plane with coordinates $\begin{pmatrix} X_0 & 0 & Z_0 \end{pmatrix}^{\mathrm{T}}$. To compute the image of the bitangent cone with a vertex lying outside the line segment joining the two sphere centers, we consider a plane $\begin{pmatrix} A & B & C & D \end{pmatrix}^{\mathrm{T}}$ passing through the camera center and tangent to both spheres, with both spheres lying on the same side of the plane. This plane should satisfy

$$
\begin{cases}
AX_0 + CZ_0 + D = 0 \\
D/\sqrt{A^2 + B^2 + C^2} = r_1 \\
(Ct + D)/\sqrt{A^2 + B^2 + C^2} = r_2
\end{cases}
\tag{7}
$$

The first equation above expresses the fact that the camera center is lying on the plane being considered. The second equation enforces that the plane is at a distance r_1 from the center of $\mathbf{Q}_1$ and is hence tangent to $\mathbf{Q}_1$. Similarly the last equation gives the constraint that the plane is tangent to $\mathbf{Q}_2$. Note the distances from the plane to the two sphere centers are both taken as positive to ensure that the spheres lie on the same side of the plane. Solving (7) gives

$$
\begin{cases}
A = (-Z_0 r_2 + Z_0 r_1 - t r_1)/(t X_0) \\
B = \pm\sqrt{1 - C^2 - A^2} \\
C = (r_2 - r_1)/t \\
D = r_1
\end{cases}
\tag{8}
$$

It follows from (8) that there will be two solutions which correspond to the two planes as depicted in Fig. 3.

It can be easily verified that the point $\mathbf{V}_{\mathrm{out}}$ with coordinates $\begin{pmatrix} 0 & 0 & \frac{t r_1}{r_1 - r_2} \end{pmatrix}^{\mathrm{T}}$ lies on both bitangent planes. Hence, the intersection line of the two bitangent planes is determined by the camera center and $\mathbf{V}_{\mathrm{out}}$. Note that $\mathbf{V}_{\mathrm{out}}$ is a point on the Z-axis, and it is in fact the vertex of the cone bitangent to the two spheres. Its coordinates are determined only by the two spheres, and it can thus provide a point correspondence between different views. The two bitangent planes project as two lines in the image which form the image of the cone, and are externally bitangent to the conic silhouettes of the two spheres. The intersection point of these two bitangent lines therefore gives the image of $\mathbf{V}_{\mathrm{out}}$.

The analysis is similar for the other cone with its vertex lying within the line segment connecting the two sphere centers (see Fig. 4). Due to the fact that the two spheres should be located on different sides of the bitangent plane, the set of equations now becomes

$$
\begin{cases}
AX_0 + CZ_0 + D = 0 \\
D/\sqrt{A^2 + B^2 + C^2} = r_1 \\
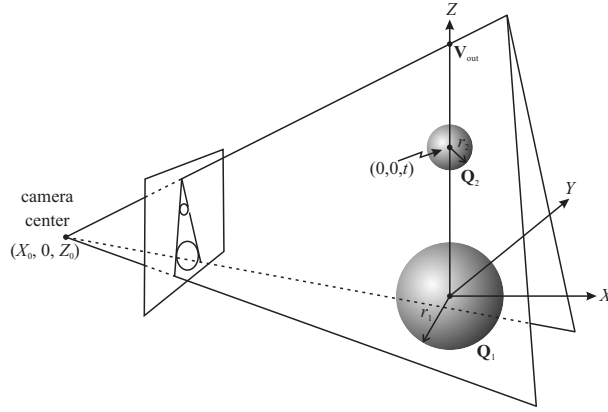(Ct + D)/\sqrt{A^2 + B^2 + C^2} = -r_2
\end{cases}
\tag{9}
$$

Fig. 3.   Bitangent planes to the two spheres with the intersection line passing through the camera center and $\mathbf{V}_{\text{out}}$. The two spheres are on the same side of the bitangent planes.

Solving (9) gives

$$
\begin{cases}
A = (Z_0 r_2 + Z_0 r_1 - t r_1)/(t X_0) \\
B = \pm\sqrt{1 - C^2 - A^2} \\
C = -(r_2 + r_1)/t \\
D = r_1
\end{cases}
\tag{10}
$$

Similarly, the point $\mathbf{V}_{\text{in}} = \begin{pmatrix} 0 & 0 & \frac{t r_1}{r_1 + r_2} \end{pmatrix}^{\text{T}}$ lying on the intersection line of the two bitangent planes is determined only by the two spheres. It can provide one more point correspondence between different views. Its image can be obtained as the intersection of the internal bitangents to the two conic silhouettes in the image (see Fig. 4). For the case when the two conic silhouettes partially overlap with each other, the internal bitangents become two complex vectors, of which the intersection is still the image of $\mathbf{V}_{\text{in}}$. The image of the revolution axis can be readily obtained as the line passing through the images of $\mathbf{V}_{\text{out}}$ and $\mathbf{V}_{\text{in}}$.

The above analysis shows that a pair of spheres can provide two point features lying on the axis passing through the two sphere centers. With three spheres, there will be six such point features and the three sphere centers can be obtained as the intersections of the axes defined by the point features. Hence, nine point features can be deduced from an image of three spheres. It should be noted that all these nine points are lying on a common plane. This implies that these points only provide the geometry of the plane, which are not sufficient for determining the epipolar geometry (see [9]).

By extending the above conclusion, N spheres can provide $C_2^N \times 2 + N = N^2$ point features when $N > 2$. For the case of four or more spheres in general positions, these points should be distributed in a 3D world frame rather than on a common plane. Hence, the epipolar geometry of two views can be recovered easily from such point features using classical algorithms [32]. Weak calibration of the cameras can be achieved by applying the factorization algorithm [9], [33]. Factorization technique
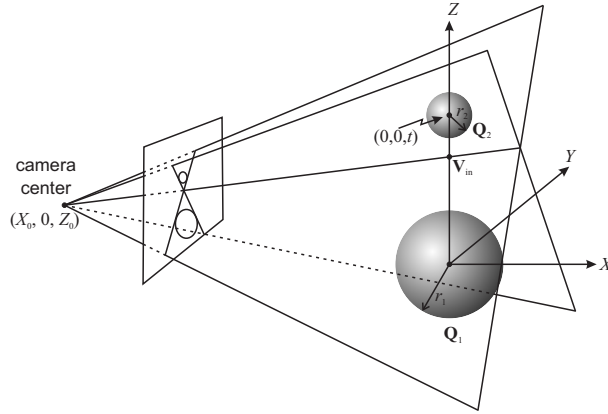
Fig. 4. Bitangent planes to the two spheres with the intersection line passing through the camera center and $\mathbf{V}_{\text{in}}$. The two spheres are on different sides of the planes.

is advantageous over other methods in that it treats each view equally and recover all the projection matrices simultaneously.

V. RECOVERING EPIPOLAR GEOMETRY FROM THREE SPHERES

In this section, the problem of recovering epipolar geometry using only three spheres is tackled. Both the point features and the epipolar tangent constraints are exploited in estimating the fundamental matrices between view pairs.

A. Plane Plus Parallax Geometry

Consider a pair of cameras with projection matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ respectively. Let $\mathbf{e}_i$ be the epipole formed on the image of the camera $\mathbf{P}_i$. The fundamental matrix can be written in a plane plus parallax representation, given by [9]

$$\mathbf{F} \sim \mathbf{H}^{-\text{T}}[\mathbf{e}_1]_\times \sim [\mathbf{e}_2]_\times \mathbf{H}, \tag{11}$$

where $\mathbf{H}$ is a homography between the two views induced by a world plane not passing through the two camera centers. $\mathbf{H}$ can be determined from a minimum of four point correspondences over the two views. By exploiting the homography, the epipole $\mathbf{e}_i$ can be recovered from two more points not lying on the plane used for inducing the homography $\mathbf{H}$ [9].

In [34], Cross et al. demonstrated that if the silhouette of an object in the first view is transferred to the second view via a homography induced by a world plane, the outer bitangents to the two silhouettes in the second view should be epipolar lines, and their intersection gives the epipole. This situation is illustrated in Fig. 5. Backprojecting the silhouettes $\mathbf{S}_1$ and $\mathbf{S}_2$ of the object from both views gives two silhouettes $\mathbf{S}_1'$ and $\mathbf{S}_2'$ in the world plane $\mathbf{\Pi}$. The geometry that the epipolar plane $\mathbf{O}_1\mathbf{O}_2\mathbf{X}$ being tangent to the object at the frontier point $\mathbf{X}$ results in a line $\mathbf{L}$, which is the intersection

of the plane $\mathbf{\Pi}$ with this epipolar plane, being bitangent to both $\mathbf{S}'_1$ and $\mathbf{S}'_2$. It follows that the outer bitangent to the images of $\mathbf{S}'_1$ and $\mathbf{S}'_2$ in the second view is an epipolar line. Hence, the epipole $\mathbf{e}_i$ can be determined by the two epipolar bitangents, and the fundamental matrix can thus be recovered by (11).
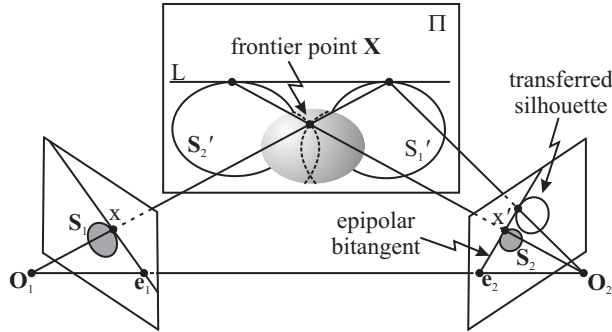

Fig. 5. By transferring the silhouette in the first view to the second view via a plane induced homography, the outer bitangents to the two silhouettes in the second view will be epipolar lines and their intersection gives the epipole.

B. Recovering Epipoles

In our work of recovering two view geometry from three spheres, a homography $\mathbf{H}$ between the two views is first computed from the point correspondences induced by the spheres (see Section IV). Note that the plane used for inducing the homography is the plane passing through the three sphere centers. The degenerate case that the plane passes through the camera centers can be detected easily and may be avoided by carefully positioning the spheres. As indicated in [30], [35], generally, the silhouettes of an object in two views can provide two outer epipolar tangent points, which correspond to two frontier points on the surface of the object (see Fig. 5). This implies that three spheres can provide six frontier points for the computation of the fundamental matrix. Under perspective projection, the frontier points would not all lie on the plane used for inducing the homography simultaneously, and hence they can provide the viewing geometry off the plane. The epipole $\mathbf{e}_i$ can thus be recovered from them using the obtained homography. In our work, since the silhouette of a sphere is analytically expressed as a conic, the transformation of the silhouettes and further the computation of bitangents and epipoles can be performed easily. Note that the conic silhouette and the transferred conic silhouette may be very close to each other (see Fig. 6(c)). This situation would, however, not affect the computation of the epipolar bitangents since the conics representing the silhouettes can be estimated with high precision.
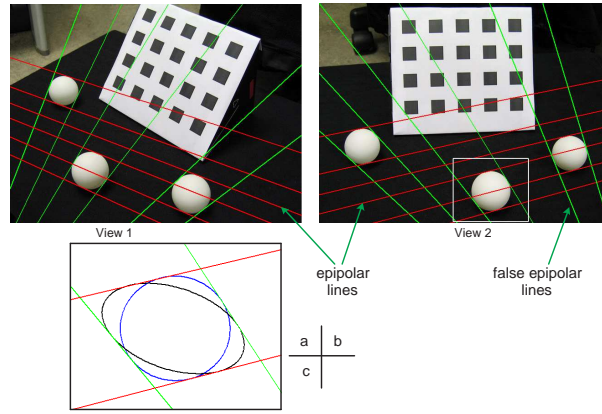
Fig. 6. (a) and (b) are two views of a world scene with three spheres and a planar pattern. Each silhouette and the corresponding transferred silhouette provide four outer bitangents. Two of them are true epipolar lines, and the other two are epipolar lines corresponding to a virtual camera. (c) shows an enlarged view of the silhouette and the transferred one in the second view.

C. Multiple Solutions

Given the conic silhouette of a sphere in one view and its conic silhouette transferred from another view via a homography $\mathbf{H}$ induced by the plane $\mathbf{\Pi}$ containing the three sphere centers, four bitangents to the silhouettes can be located analytically (see Fig. 6(c)). These four bitangents correspond to two sets of epipolar lines, and give rise to two solutions for the epipole. One of the epipoles is the true solution, and the other one is the image of a virtual camera. The existence of the epipole for the virtual camera can be explained by the symmetry in the configuration of the three spheres w.r.t. the plane $\mathbf{\Pi}$. Consider two cameras $\mathbf{P}_1$ and $\mathbf{P}_2$ located at $\mathbf{O}_1$ and $\mathbf{O}_2$ respectively (see Fig. 7). Transferring the conic silhouettes from the image of $\mathbf{P}_1$ to the image of $\mathbf{P}_2$ via $\mathbf{H}$ is equivalent to finding the intersections of the plane $\mathbf{\Pi}$ with the viewing cones formed from the conic silhouettes in $\mathbf{P}_1$, and projecting such intersections into the image of $\mathbf{P}_2$. Now consider a virtual camera $\mathbf{P}_1'$ located at $\mathbf{O}_1'$ which is a mirrored position of $\mathbf{O}_1$ w.r.t. the plane $\mathbf{\Pi}$, and having a mirrored and opposite viewing direction (see Fig. 7). Under this configuration, the virtual image of the three spheres in $\mathbf{P}_1'$ will be identical to the image of $\mathbf{P}_1$, and the viewing cones formed from the conic silhouettes in $\mathbf{P}_1'$ will intersect the plane $\mathbf{\Pi}$ at exactly the same locations as those formed from the conic silhouettes in $\mathbf{P}_1$. This implies that the homography $\mathbf{H}$ will transfer both the conic silhouettes in $\mathbf{P}_1$ and $\mathbf{P}_1'$ to the same conic silhouettes in $\mathbf{P}_2$. Consequently, there will be two sets of epipolar bitangents, one for $\mathbf{P}_1$ and another for $\mathbf{P}_1'$. The intersections of the epipolar bitangents will then give two epipoles, of which one is the image of $\mathbf{O}_1$ and the other is the image of $\mathbf{O}_1'$.

The true epipole can be easily distinguished from the two solutions. A direct method is to select the correct one manually by integrating the world scene information from images, like in Fig. 6(a) and
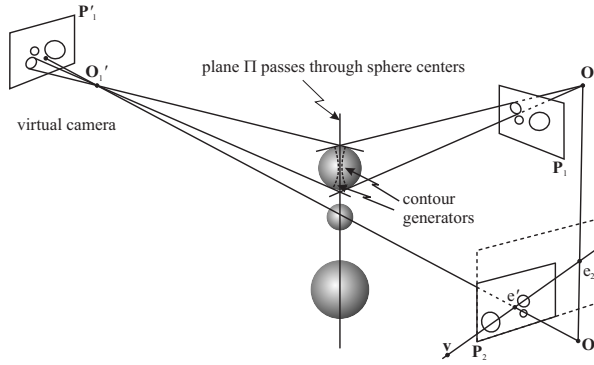
Fig. 7. The camera center $\mathbf{O}_1'$ of the virtual camera and the camera center $\mathbf{O}_1$ of the camera $\mathbf{P}_1$ are bilateral symmetric w.r.t. the plane $\mathbf{\Pi}$ passing through the three sphere centers. Both real cameras are on the same side of the plane $\mathbf{\Pi}$, and the virtual camera is on the other side. $\mathbf{v}$ is the vanishing point of the normal direction of the plane $\mathbf{\Pi}$.

Fig. 6(b). If the radii of the three spheres are equal, the true epipole can be selected automatically. As shown in Fig. 7, the line connecting $\mathbf{O}_1'$ and $\mathbf{O}_1$ is orthogonal to the plane $\mathbf{\Pi}$. Its image $\mathbf{e}_2\mathbf{e}'$ in the image of $\mathbf{P}_2$ should pass through the vanishing point $\mathbf{v}$ of the normal direction of $\mathbf{\Pi}$. For the case where both $\mathbf{O}_1$ and $\mathbf{O}_2$ are on the same side of $\mathbf{\Pi}$, the three points $\mathbf{v}$, $\mathbf{e}'$ and $\mathbf{e}_2$ keep the relationship that $\mathbf{e}'$ is within the line segment $\mathbf{v}\mathbf{e}_2$. This relationship can be exploited to select the true epipole. It is known that $\mathbf{v}$ and the vanishing line $\mathbf{l}$ of $\mathbf{\Pi}$ are in a pole-polar relationship w.r.t. the imaged absolute conic $\boldsymbol{\omega}$, given by

$$\mathbf{l} \sim \boldsymbol{\omega}\mathbf{v} \sim \mathbf{K}^{-\mathrm{T}}\mathbf{K}^{-1}\mathbf{v},$$

where $\mathbf{K}$ is the calibration matrix [9]. When the radii of the spheres are equal, the image of $\mathbf{V}_{\mathrm{out}}$ will lie on $\mathbf{l}$. Three spheres can provide three such points which are sufficient to determine $\mathbf{l}$. As for $\mathbf{K}$, it can be easily recovered from three spheres by applying the algorithm presented in [23]. $\mathbf{v}$ can thus be computed from the obtained $\mathbf{l}$ and $\mathbf{K}$. The real epipole $\mathbf{e}_2$ is chosen to be the one further away from $\mathbf{v}$. On the contrary, if the two cameras are on different sides of the plane $\mathbf{\Pi}$, the one closer to $\mathbf{v}$ is chosen instead. If the radii of the three spheres happen to be unequal, the vanishing point $\mathbf{v}$ can still be recovered by involving a third view. Like the line $\mathbf{e}_2\mathbf{e}'$ in Fig. 7, the third camera provides another line passing through $\mathbf{v}$. This line together with the line $\mathbf{e}_2\mathbf{e}'$ determines $\mathbf{v}$ uniquely.

D. Summary

To conclude, three spheres can provide nine point features (planar geometry) and six frontier points (viewing geometry off the plane determined by the nine points) for computing the fundamental matrix between two views. The fundamental matrix can be easily recovered in the form of a plane plus parallax expression with the point features providing the homography and the frontier points

determining the epipole. Extending the situation to N spheres where $N > 2$, there would be N^2 point features and $2N$ frontier points for two views. If the spheres are placed generally instead of having all sphere centers lying on a common plane, the point features alone can determine the epipolar geometry using classical algorithms as discussed in Section IV. Alternatively, if our algorithm is applied instead, the situation of having two solutions for the epipole would not happen as long as not all sphere centers are on the same plane. This is because the additional spheres break the symmetry in the configuration.

VI. METRIC RECONSTRUCTION

A. Recovery of the Absolute Dual Quadric

The work of upgrading a weak calibration to a full calibration can be achieved by identifying the absolute dual quadric $\mathbf{Q}_\infty^*$ or the absolute conic $\mathbf{\Omega}$ in the projective space [9]. As introduced in Section III, a sphere intersects the plane at infinity $\mathbf{\Pi}_\infty$ at the absolute conic $\mathbf{\Omega}$ and this situation is preserved under arbitrary projective transformation. Hence, to recover $\mathbf{\Omega}$ is equivalent to determining the plane at infinity $\mathbf{\Pi}_\infty$ in the projective space. Since $\mathbf{\Pi}_\infty$ and the sphere center are in a pole-polar relationship w.r.t. the sphere, the task of locating the plane at infinity $\mathbf{\Pi}_\infty$ is equivalent to reconstructing the spheres (which are quadrics in the projective space) and determining the sphere centers. Both approaches can be performed by solving linear equations (see Section III for quadric reconstruction).

Generally, it is not simple to compute the intersection of a sphere with the plane at infinity $\mathbf{\Pi}_\infty$. Hence, the absolute dual quadric $\mathbf{Q}_\infty^*$ is often computed instead of the absolute conic $\mathbf{\Omega}$. By taking $\mathbf{\Pi}_\infty$ as a vertex and the dual sphere $\mathbf{Q}^*$ as a point quadric, a cone with vertex $\mathbf{\Pi}_\infty$ and tangent to $\mathbf{Q}^*$ is given by

$$\mathbf{Q}_{co} \sim (\mathbf{\Pi}_\infty^T \mathbf{Q}^* \mathbf{\Pi}_\infty) \mathbf{Q}^* - (\mathbf{Q}^* \mathbf{\Pi}_\infty)(\mathbf{Q}^* \mathbf{\Pi}_\infty)^T. \tag{12}$$

It can be shown that $\mathbf{Q}_{co}$ is the absolute dual quadric $\mathbf{Q}_\infty^*$ in the projective space (see proof in Appendix). Hartley and Zisserman [9] showed that if $\mathbf{Q}_\infty^*$ is decomposed as

$$\mathbf{Q}_\infty^* = \mathbf{H}_r \begin{bmatrix} \mathbf{I}_{3\times3} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \mathbf{H}_r^T, \tag{13}$$

then $\mathbf{H}_r^{-1}$ is a 3D (point) transformation that takes the projective frame to a Euclidean frame. It follows that the projection matrices are rectified by post-multiplication with $\mathbf{H}_r$, written as $\mathbf{P}_M = \mathbf{P}\mathbf{H}_r$. Hence, the reconstruction ambiguity is removed and the cameras can be fully calibrated.

To achieve a high precision of metric reconstruction, the obtained rectifying transformation $\mathbf{H}_r$ is refined further. It is known that the calibration matrix $\mathbf{K}$ usually has very small skew value, and the principal point is close to the image center. This situation is exploited in our work to refine $\mathbf{H}_r$. We

assume that the camera has zero skew and different cameras have the same principal point. The cost function to be minimized can be expressed as

$$\xi(\mathbf{H}_r) = \lambda \sum_i s_i + \sum_i \sum_{j \neq i} \| \begin{pmatrix} u_{0i} \\ v_{0i} \end{pmatrix} - \begin{pmatrix} u_{0j} \\ v_{0j} \end{pmatrix} \|, \tag{14}$$

where λ is a weight factor. In each iteration, the skew s_i and the principal point $\begin{pmatrix} u_{0i} & v_{0i} \end{pmatrix}^{\mathrm{T}}$ are obtained by decomposing $\mathbf{P}_M^i$ in the form of $\mathbf{P} = \mathbf{K}[\ \mathbf{R} \quad \mathbf{t}\ ]$. The algorithm for computing the rectifying transformation $\mathbf{H}_r$ is briefly summarized below:

1) Reconstruct the spheres in the projective frame.

2) Compute the sphere centers by triangulation.

3) Recover the plane at infinity $\mathbf{\Pi}_\infty$ from the reconstructed spheres and their centers.

4) Compute the rectifying transformation $\mathbf{H}_r$ using (12) and (13).

5) Refine $\mathbf{H}_r$ by minimizing (14).

B. Recovery of the Calibration Matrix

When there are less than three views, the spheres cannot be reconstructed and therefore the approach described in the last subsection for upgrading the weak calibration cannot be applied. One possible but not desirable solution is to obtain the camera internal parameters separately using other existing methods (e.g., [10]), and use them to upgrade the weak calibration. Alternatively, the camera calibration matrix $\mathbf{K}$ can be recovered from a single image of three or more spheres.
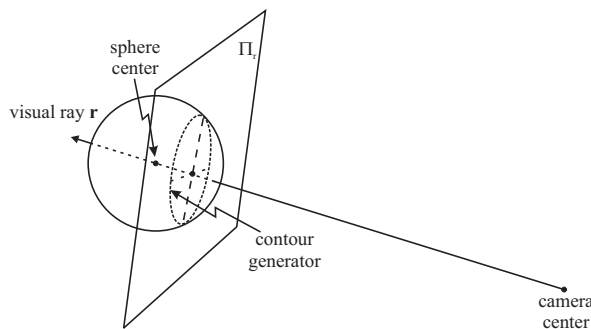

Fig. 8. The contour generator of the sphere is a circle lying on the plane $\mathbf{\Pi}_r$. The visual ray $\mathbf{r}$ going through the sphere center from the camera center also passes through the circle center.

Consider the projection of a sphere under a camera again, as shown in Fig. 8. One observes that the visual ray $\mathbf{r}$ going through the sphere center is orthogonal to the plane $\mathbf{\Pi}_r$ where the contour generator lies. It follows that the image of the sphere center and the vanishing line of the plane $\mathbf{\Pi}_r$ are in a pole-polar relationship w.r.t. the imaged absolute conic $\boldsymbol{\omega}$. Hence one imaged sphere can place two constraints on $\boldsymbol{\omega}$. In the presence of three or more spheres, the imaged sphere centers

can be extracted. The vanishing line corresponding to each sphere can be computed as the polar to the imaged sphere center w.r.t. the conic silhouette of the sphere. Since $\mathbf{K}$ has only five degrees of freedom, a minimum of three spheres is sufficient to calibrate the internal parameters of the camera.

Alternatively, by viewing any two spheres as a surface of revolution, different forms of constraints on $\boldsymbol{\omega}$ can be derived (see [23], [36]). It is shown that the imaged axis of the surface of revolution and the vanishing point of the normal direction of the plane defined by the axis and the camera center are in a pole-polar relationship w.r.t. $\boldsymbol{\omega}$. Denote the two conic silhouettes as $\mathbf{C}_i$ and $\mathbf{C}_j$ respectively. The imaged axis is shown to be one of the eigenvectors of the matrix $\mathbf{C}_i\mathbf{C}_j^*$ and the cross product of the other two eigenvectors of $\mathbf{C}_i\mathbf{C}_j^*$ gives the vanishing point [24]. It can be proved that the set of constraints derived in this way are essentially equivalent to those introduced in the last paragraph.

To get a high precision of $\mathbf{K}$, the spheres are required to be placed further away from the image center while capturing the images. This is because as the conic image of a sphere is close to the image center, the vanishing line of the plane $\mathbf{\Pi}$ in Fig. 8 will be far away from the image, which, as a result, will be estimated badly. On the other hand, the spheres cannot be moved too far away from the image center, as the image may have remarkable radial distortion effect near the edges of the image. In situations where strong radial distortions are observed in the images, standard algorithms for correcting lens distortions (e.g., [37]) may be applied prior to the calibration of the internal and external parameters. Applying the stratified camera calibration approach (i.e., weak calibration followed by metric reconstruction) as discussed in the last subsection, on the other hand, can avoid these issues easily.

VII. EXPERIMENTAL RESULTS

The performance of the proposed calibration method was evaluated using both synthetic and real data. Errors in the location of the induced point features, symmetric transfer errors of the estimated fundamental matrices, as well as 3D reconstruction errors are presented. Besides, its accuracy was also compared against a commonly used point-based approach proposed by Zhang [10]. Experimental results demonstrate that good quality 3D model can be reconstructed using the calibration results. This demonstrates the usefulness and effectiveness of the proposed algorithm for calibrating multiple views.

A. Synthetic Data Experiments

In the first synthetic data experiment, three spheres with unequal radii were generated to test the accuracy of the induced point features estimated from the imaged spheres, and that of the fundamental matrix estimated in the presence of noise. Fig. 9 shows two views of the configuration of the spheres. To simulate the effect of image noise, the silhouette of a sphere was first obtained analytically using

(5), and 150 points were sampled along the silhouette. Random uniform noise was then added to these sample points by perturbing them in the directions normal to their local tangents. Finally, a conic was fitted to these noisy points using a linear least squares method. In this experiment, five different noise levels (0.5, 0.7, 1.0, 1.2 and 1.5 pixels) were used and 20 independent trials were performed for each noise level. The procedure for recovering epipolar geometry from three spheres is briefly summarized below:

1) Obtain the silhouette of each sphere analytically using (5).

2) Sample the silhouette of each sphere and add noise to the sample points in directions normal to the local tangent directions.

3) Fit conic to the noisy sample points of each sphere.

4) Extract point features from the noisy conics and compute a homography $\mathbf{H}$ over two views.

5) Transfer the conics in one view to the other view using $\mathbf{H}$ and determine the epipole $\mathbf{e}$ using bitangents.

6) Refine epipole $\mathbf{e}$ using frontier points.
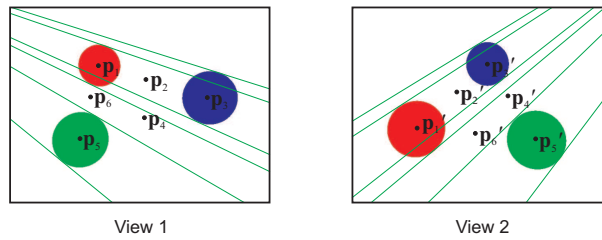

Fig. 9.   Two views of three synthetic spheres. The lines in the image are the ground truth epipolar tangents.

In Step 4 of the above algorithm, only six of the nine induced point features were exploited in the computation of $\mathbf{H}$. This is because the projections of $\mathbf{V}_{\text{out}}$ are often far away from the image centers, and including these points in the computation affects the precision of $\mathbf{H}$ greatly. To achieve a high precision in the estimated epipolar geometry, the epipole $\mathbf{e}$ obtained in Step 5 is refined further. The cost function to be minimized is the symmetric transfer error given by the geometric distances between the epipolar tangent points and the corresponding epipolar lines, as shown in Fig. 10, written as

$$\xi(\mathbf{e}) = \frac{1}{2n} \sum_{i=1}^{n} \left(d(\mathbf{p}_i', \mathbf{F}(\mathbf{e})\mathbf{p}_i) + d(\mathbf{p}_i, \mathbf{F}^{\mathrm{T}}(\mathbf{e})\mathbf{p}_i') \right), \tag{15}$$

where n is the number of frontier points and the function $d(.,.)$ is the point-to-line Euclidean distance expressed in pixels, and $\mathbf{p}_i$ and $\mathbf{p}_i'$ are the epipolar tangent points.

Table I shows the root mean squared (rms) distances between every estimated feature point $\mathbf{p}_i$ in the first view (see Fig. 9) and its ground truth location. It can be observed that the rms distance
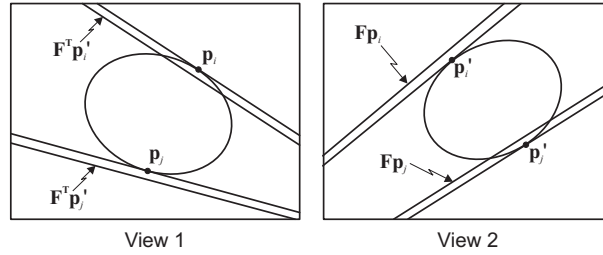
Fig. 10.   Two views of a sphere with an associated fundamental matrix $\mathbf{F}$. $\mathbf{p}_i$ and $\mathbf{p}'_i$ are the epipolar tangent points, and $\mathbf{F}\mathbf{p}_i$ and $\mathbf{F}^{\mathrm{T}}\mathbf{p}'_i$ are the corresponding epipolar lines.

TABLE I

THE RMS DISTANCES (IN PIXELS) BETWEEN EVERY ESTIMATED POINT FEATURE AND ITS GROUND TRUTH IN THE FIRST
VIEW UNDER DIFFERENT NOISE LEVELS.

noise	$\mathbf{p}_1$	$\mathbf{p}_2$	$\mathbf{p}_3$	$\mathbf{p}_4$	$\mathbf{p}_5$	$\mathbf{p}_6$
0.5 pixels	0.066	0.094	0.072	0.087	0.075	0.063
0.7 pixels	0.081	0.120	0.090	0.109	0.077	0.073
1.0 pixels	0.127	0.178	0.149	0.156	0.105	0.106
1.2 pixels	0.140	0.199	0.159	0.179	0.146	0.139
1.5 pixels	0.189	0.173	0.188	0.218	0.217	0.196

increases with the noise level. The point $\mathbf{V}_{\mathrm{out}}$ is usually located far from the image center and its estimation is usually very unstable. Hence $\mathbf{V}_{\mathrm{out}}$ is not included in the above analysis. Point features $\mathbf{p}'_i$ in the second view exhibited similar statistic characteristics, and the corresponding results are omitted here. The estimated fundamental matrix over the two views is evaluated by computing the symmetric transfer errors of the six ground truth point features. The errors are then averaged over the 20 trials under each noise level. Table II displays the list of the averaged symmetric transfer errors under different noise levels. As expected, the error increases with the noise level.

In the second synthetic data experiment, the performance of the proposed algorithm was compared against a commonly used point-based algorithm proposed by Zhang [10]. The calibration and reconstruction accuracies of the proposed method using four and seven spheres, respectively, were compared

TABLE II

THE AVERAGED SYMMETRIC TRANSFER ERRORS (IN PIXELS) FOR THE SIX GROUND TRUTH POINT FEATURES UNDER
DIFFERENT NOISE LEVELS.

noise (pixels)	0.5	0.7	1.0	1.2	1.5
avg transfer err	0.0543	0.0657	0.1076	0.1161	0.1409

against those of Zhang's method using 49 points and 100 points, respectively. The synthetic spheres used were placed in general positions to ensure that the point features extracted from them would not all lie on the same plane. The procedure for achieving metric reconstruction from four/seven spheres is briefly summarized below:

1) Use the procedure in the first synthetic data experiment to compute the fundamental matrices.

2) Obtain initial values for the projective depths of the point features from the estimated fundamental matrices.

3) Recover the projection matrices using Factorization algorithm.

4) Recover the rectifying transformation $\mathbf{H}_r$ in the projective space, with the weight factor $\lambda = 1$ in (14).

5) Rectify the obtained projection matrices using $\mathbf{H}_r$.

A 7×7 point pattern and a 10×10 point pattern were used in Zhang's method [10]. These point patterns were first projected onto each view, and random uniform noises were then added to each of the points. In this experiment, nine different noise levels (0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0 pixels) were used and 50 independent trials were performed for each noise level. Fig. 11 shows the images used in this experiment. The estimated camera parameters were evaluated using the ground truth values, including the coordinates of the principal point (u_0, v_0) and the scale factors α and β along the image axes. The relative errors for α and β, and absolute errors for u_0 and v_0 were measured. The average errors for each of the parameters are shown in Fig. 12. Reconstruction results were also evaluated in this experiment. 125 points inside a regular cube were projected onto each view and random uniform noises were added to each of these image points. These noisy image points and the estimated projection matrices were then used to reconstruct the 3D points. Note that the reconstruction results were up to a similarity transformation. To compute the reconstruction errors, a similarity transformation was computed for each of the reconstructions and reconstruction errors were measured against the ground truth after the alignment. Fig. 13 shows the 3D relative errors, 2D reprojection errors, aspect ratios and angle errors between adjacent edges of the reconstruction results under different noise levels.
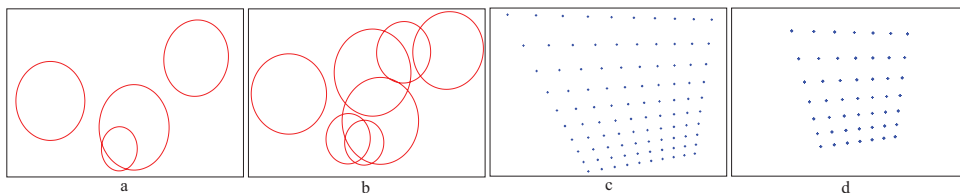

Fig. 11.   Images used in the second synthetic data experiment. From left to right: (a) 4 spheres, (b) 7 spheres, (c) 100 points, (d) 49 points.
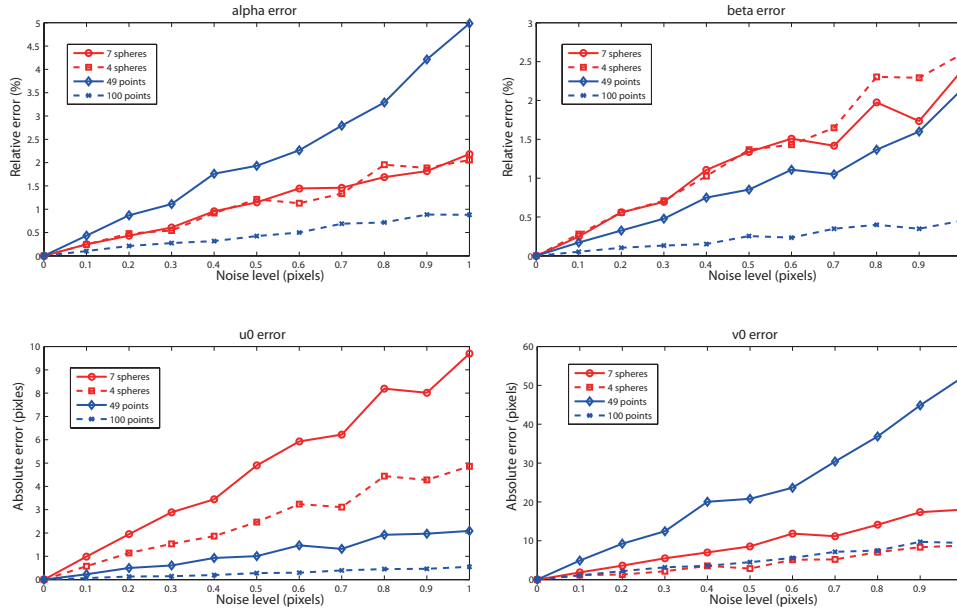
Fig. 12. Comparison of the calibration results using the proposed method and using Zhang's method [10].

It can be seen from Fig. 12 and Fig. 13 that the proposed method can achieve calibration and reconstruction results comparable to that of Zhang's method.The aspect ratios and angle errors between adjacent edges are close to 1 and 0 degree, respectively. The 3D errors and the reprojection errors in the 4 spheres case were about 1.1% and 0.7 pixel, respectively, under a noise level of 1 pixel. Note that only ten feature points (four sphere centers and six internal bitangent points) were used in the 4 spheres case, but yet it could achieve reconstruction result comparable to that obtained with Zhang's method using 100 points (3D error 0.8% and reprojection error 0.7 pixel), and much better result than that obtained with Zhang's method using 49 points (3D error 4.1% and reprojection error 1.9 pixels). As shown in Fig. 12 and Fig. 13, both the calibration and reconstruction errors virtually increase linearly with the noise level. It is also noted that using more spheres does not improve the calibration and reconstruction results. One possible explanation might be that some spheres were seriously occluded by the others (see Fig. 11(b)), leading to inaccurate feature points computation.

B. Real Data Experiments

In the first real data experiment, three ping pong balls were used to recover the epipolar geometry and a planar grid pattern was included in the images for evaluating the performance. The camera used in the experiment was a Canon A95 digital camera, and the image resolution was 640×480. Two views of the spheres and the grid pattern are shown in Fig. 6. After the refinement of the epipole, the averaged symmetric transfer error of the frontier points of the spheres is 0.022 pixels. The grid corners on the planar grid pattern were extracted by first applying Canny edge detection, followed
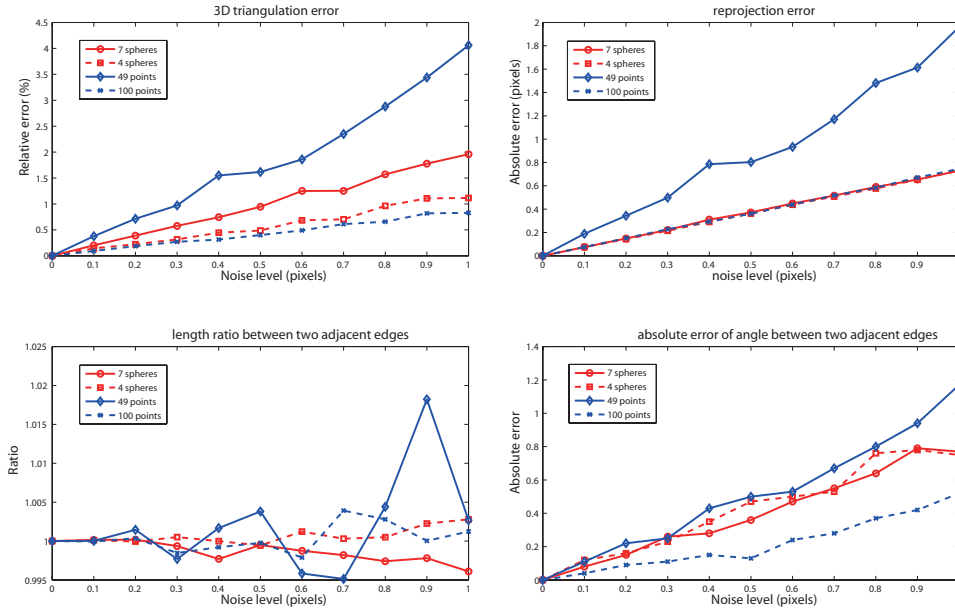
Fig. 13.   Comparison of the reconstruction results using the proposed method and Zhang's method [10].

by fitting line segments to the grid edges and finally intersecting the lines segments. The averaged symmetric transfer error of the grid corner points resulting from the estimated fundamental matrix is 0.16 pixels.

In the second real data experiment, only point features induced by four spheres were exploited in the computation of the fundamental matrix to test the accuracy of the estimated point features. The four spheres used were placed in general positions to ensure that the point features extracted from the spheres would not all lie on the same plane. A Canon A95 digital camera was used and the image resolution was 640×480. Two views of the configuration are shown in Fig. 14. Like in the three spheres experiment, six of the sixteen point features (i.e., the projections of $\mathbf{V}_{\text{out}}$) were not included in the computation of the fundamental matrix. The averaged symmetric transfer error of the frontier points resulting from the estimated fundamental matrix is 0.04 pixels, and that of the planar grid pattern corner points is 0.32 pixels. These demonstrate the high precision of the estimated fundamental matrix.

In the third real data experiment, seven spheres and a box with a planar grid pattern attached to it were used. Similar to the previous experiment, only point features (excluding $\mathbf{V}_{\text{out}}$) extracted from spheres were exploited in estimating the fundamental matrices and performing a full calibration of the cameras, whereas the planar grid pattern was used to evaluate the calibration performance. Four images were taken at different positions with a Nikon100D digital camera, which are shown in Fig. 15. The image resolution was 800×531. The rms reprojection error of the point features extracted from spheres is 0.028 pixels, and that of the planar grid pattern corner points is 0.123 pixels.
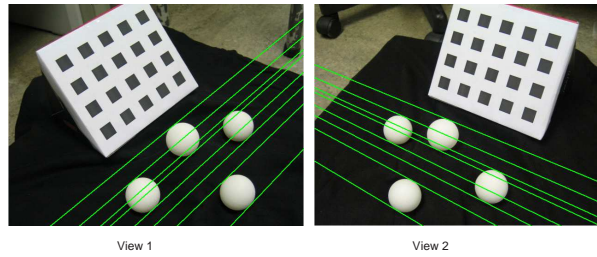
View 1   View 2

Fig. 14. Two views of four spheres and a planar grid pattern. The lines in the image are the computed outer epipolar tangents.

The length ratios between adjacent edges of the reconstructed planar grid pattern is 0.993, and the four angles are 89.4, 90.7, 89.3 and 90.6 degrees, respectively. The reconstructed box model is shown in Fig. 16. The model was reconstructed by manually selecting the imaged box corners, performing triangulation, and rendering textures on it.
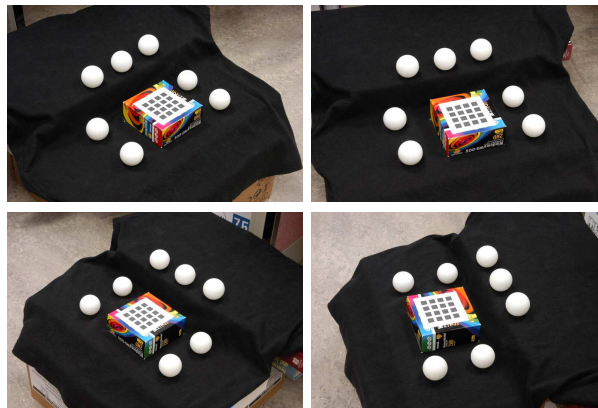

Fig. 15. Four views of seven ping pong balls and a box with a planar grid pattern attached to it.

VIII. CONCLUSIONS

We have proposed a simple and effective algorithm for calibrating multiple cameras using spheres. We found that besides the outer epipolar tangent constraints, point features could be extracted from the silhouettes of the spheres. We have shown how the point features and the outer epipolar tangent constraints from three spheres can be combined nicely in the recovery of the fundamental matrix. By relating dual spheres and the plane at infinity to the absolute dual quadric, a novel method of recovering the absolute dual quadric in the projective space has been developed for upgrading the projective frame to a metric frame. The experimental results demonstrate the applicability and effectiveness of our algorithm.

Fig. 16. Four views of the reconstructed box.

When some spheres are being totally occluded in some images in performing the camera calibration, and there are no more than two imaged spheres in one of those views, our algorithm would fail. This is because such a view cannot provide enough constraints for computing the fundamental matrix relating any other view. Conics in different views are currently being matched manually. It is possible to automate the matching process by using spheres with different colors. Alternatively, images of a single sphere being placed at three or more distinct locations can be captured simultaneously by all cameras. This will avoid the occlusion problem and make the matching problem trivial.

APPENDIX

The dual absolute quadric $\mathbf{Q}_\infty^*$ can be nicely expressed in terms of a dual sphere $\mathbf{Q}^*$ and the plane at infinity $\mathbf{\Pi}_\infty$, written as

$$\mathbf{Q}_\infty^* = (\mathbf{\Pi}_\infty^{\mathrm{T}} \mathbf{Q}^* \mathbf{\Pi}_\infty) \mathbf{Q}^* - (\mathbf{Q}^* \mathbf{\Pi}_\infty)(\mathbf{Q}^* \mathbf{\Pi}_\infty)^{\mathrm{T}}. \tag{16}$$

The proof follows by replacing $\mathbf{Q}^*$ and $\mathbf{\Pi}_\infty$ in (16) with their expressions in the metric frame. Consider a sphere with center $\tilde{\mathbf{A}}$ and radius r, its dual can be written as

$$\mathbf{Q}^* = \begin{bmatrix} \mathbf{I} - \tilde{\mathbf{A}}\tilde{\mathbf{A}}^{\mathrm{T}}/r^2 & -\tilde{\mathbf{A}}/r^2 \\ -\tilde{\mathbf{A}}^{\mathrm{T}}/r^2 & -1/r^2 \end{bmatrix}. \tag{17}$$

In the metric frame, the plane at infinity $\mathbf{\Pi}_\infty$ is given by $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}^\mathrm{T}$. Consequently, the right hand side of (16) can be rewritten as

$$-\mathbf{Q}^*/r^2 - \begin{pmatrix} -\tilde{\mathbf{A}}/r^2 \\ -1/r^2 \end{pmatrix} \begin{pmatrix} -\tilde{\mathbf{A}}^\mathrm{T}/r^2 & -1/r^2 \end{pmatrix}$$

$$= -\mathbf{Q}^*/r^2 - \begin{pmatrix} \tilde{\mathbf{A}}\tilde{\mathbf{A}}^\mathrm{T}/r^4 & \tilde{\mathbf{A}}/r^4 \\ \tilde{\mathbf{A}}^\mathrm{T}/r^4 & 1/r^4 \end{pmatrix}$$

$$\sim \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^\mathrm{T} & 0 \end{pmatrix}$$

$$= \mathbf{Q}^*_\infty.$$

By the principle of duality, $\mathbf{\Pi}_\infty$ can be regarded as a vertex and the dual sphere $\mathbf{Q}^*$ as a point quadric in (16), and $\mathbf{Q}^*_\infty$ can be geometrically interpreted as a cone with vertex $\mathbf{\Pi}_\infty$ and tangent to $\mathbf{Q}^*$.

REFERENCES

[1] W. N. Martin and J. K. Aggarwal, "Volumetric descriptions of objects from multiple views," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 5, no. 2, pp. 150–158, March 1983.

[2] C. H. Chien and J. K. Aggarwal, "Volume/surface octrees for the representation of three-dimensional objects," *Computer Vision, Graphics and Image Processing*, vol. 36, no. 1, pp. 100–113, October 1986.

[3] E. Boyer and M. O. Berger, "3D surface reconstruction using occluding contours," *Int. Journal of Computer Vision*, vol. 22, no. 3, pp. 219–233, March 1997.

[4] C. Liang and K.-Y. K. Wong, "Robust recovery of shapes with unknown topology from the dual space," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 12, pp. 2205–2216, December 2007.

[5] S. Laveau and O. D. Faugeras, "3-D scene representation as a collection of images and fundamental matrices," in *Proc. Int. Conf. on Pattern Recognition*, vol. A, Jerusalem, Israel, October 1994, pp. 689–691.

[6] S. Avidan and A. Shashua, "Novel view synthesis in tensor space," in *Proc. Conf. Computer Vision and Pattern Recognition*, San Jaun, PR, June 1997, pp. 1034–1040.

[7] J. Douret and R. Benosman, "A multi-cameras 3D volumetric method for outdoor scenes: A road traffic monitoring application." in *Proc. Int. Conf. on Pattern Recognition*, vol. III, Cambridge, UK, August 2004, pp. 334–337.

[8] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses," *IEEE Trans. Robotics and Automation*, vol. 3, no. 4, pp. 323–344, August 1987.

[9] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge, UK: Cambridge University Press, 2000.

[10] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, November 2000.

[11] Q. Chen, H. Y. Wu, and T. Wada, "Camera calibration with two arbitrary coplanar circles," in *Proc. 8th European Conf. on Computer Vision*, vol. III, Prague, Czech Republic, May 2004, pp. 521–532.

[12] J. S. Kim, P. Gurdjos, and I. S. Kweon, "Geometric and algebraic constraints of projected concentric circles and their applications to camera calibration," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 4, pp. 637–642, April 2005.

[13] G. Jiang and L. Quan, "Detection of concentric circles for camera calibration," in *Proc. 10th Int. Conf. on Computer Vision*, vol. I, Beijing, China, October 2005, pp. 333–340.

[14] Y. H. Wu, H. J. Zhu, Z. Y. Hu, and F. C. Wu, "Camera calibration from the quasi-affine invariance of two parallel circles," in *Proc. 8th European Conf. on Computer Vision*, vol. I, Prague, Czech Republic, May 2004, pp. 190–202.

[15] P. Gurdjos, P. F. Sturm, and Y. H. Wu, "Euclidean structure from $n >= 2$ parallel circles: Theory and algorithms," in *Proc. 9th European Conf. on Computer Vision*, vol. I. Springer, May 2006, pp. 238–252.

[16] C. Colombo, D. Comanducci, and A. del Bimbo, "Camera calibration with two arbitrary coaxial circles," in *Proc. 9th European Conf. on Computer Vision*, vol. I, Graz, Austria, May 2006, pp. 265–276.

[17] G. Zhang and K.-Y. K. Wong, "Motion estimation from spheres," in *Proc. Conf. Computer Vision and Pattern Recognition*, vol. 1, New York, USA, June 2006, pp. 1238–1243.

[18] M. A. Penna, "Camera calibration: A quick and easy way to determine the scale factor," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 12, pp. 1240–1245, December 1991.

[19] N. Daucher, M. Dhome, and J. T. La Preste, "Camera calibration from spheres images," in *Proc. 3rd European Conf. on Computer Vision*, vol. A, Stockholm, Sweden, May 1994, pp. 449–454.

[20] H. Teramoto and G. Xu, "Camera calibration by a single image of balls: from conics to the absolute conic," in *Proc. Asian Conf. on Computer Vision*, Melbourne, Australia, January 2002, pp. 499–506.

[21] M. Agrawal and L. S. Davis, "Camera calibration using spheres: a semi-definite programming approach," in *Proc. 9th Int. Conf. on Computer Vision*, Nice, France, October 2003, pp. 782–789.

[22] X. Ying and H. Zha, "Linear approaches to camera calibration from sphere images or active intrinsic calibration using vanishing points," in *Proc. 10th Int. Conf. on Computer Vision*, vol. I, Beijing, China, October 2005, pp. 596–603.

[23] H. Zhang, G. Zhang, and K.-Y. K. Wong, "Camera calibration with spheres: Linear approaches," in *Proc. Int. Conf. on Image Processing*, vol. II, Genova, September 2005, pp. 1150–1153.

[24] H. Zhang, K.-Y. K. Wong, and G. Zhang, "Camera calibration from images of spheres," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 3, pp. 499–503, March 2007.

[25] O. Faugeras, Q. T. Luong, and T. Papadopoulo, *The Geometry of Multiple Images*. MIT Press, 2001.

[26] H. Schulz-Mirbach and I. Weiss, "Projective reconstruction from curve correspondences in uncalibrated views," Tech. Rep. TR-402-94-014, 1994. [Online]. Available: citeseer.ist.psu.edu/schulz-mirbach94projective.html

[27] F. Kahl and A. Heyden, "Using conic correspondence in two images to estimate the epipolar geometry," in *Proc. 6th Int. Conf. on Computer Vision*, Bombay, India, January 1998, pp. 761–766.

[28] J. Y. Kaminski and A. Shashua, "Multiple view geometry of general algebraic curves," *Int. Journal of Computer Vision*, vol. 56, no. 3, pp. 195–219, February 2004.

[29] P. R. S. Mendonça, K.-Y. K. Wong, and R. Cipolla, "Epipolar geometry from profiles under circular motion," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 604–616, June 2001.

[30] K.-Y. K. Wong and R. Cipolla, "Structure and motion from silhouettes," in *Proc. 8th Int. Conf. on Computer Vision*, vol. II, Vancouver, BC, Canada, July 2001, pp. 217–222.

[31] G. Cross and A. Zisserman, "Quadric reconstruction from dual-space geometry," in *Proc. 6th Int. Conf. on Computer Vision*, Bombay, India, January 1998, pp. 25–31.

[32] Z. Zhang, "Determining the epipolar geometry and its uncertainty: A review," *Int. Journal of Computer Vision*, vol. 27, no. 2, pp. 161–195, March 1998.

[33] P. F. Sturm and B. Triggs, "A factorization based algorithm for multi-image projective structure and motion," in *Proc. 4th European Conf. on Computer Vision*, vol. II, Cambridge, UK, April 1996, pp. 709–720.

[34] G. Cross, A. W. Fitzgibbon, and A. Zisserman, "Parallax geometry of smooth surfaces in multiple views," in *Proc. 7th Int. Conf. on Computer Vision*, Corfu, Greece, September 1999, pp. 323–329.

[35] S. N. Sinha, M. Pollefeys, and L. McMillan, "Camera network calibration from dynamic silhouettes," in *Proc. Conf. Computer Vision and Pattern Recognition*, vol. I, Washington, DC, June 2004, pp. 195–202.

[36] K.-Y. K. Wong, P. R. S. Mendonça, and R. Cipolla, "Camera calibration from surfaces of revolution," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 147–161, February 2003.

[37] F. Devernay and O. D. Faugeras, "Straight lines have to be straight," *Machine Vision and Applications*, vol. 13, pp. 14–24, 2001.