The HKU Scholars Hub   The University of Hong Kong   香港大學學術庫

Title	**Motion estimation from spheres**
Author(s)	**Zhang, G; Wong, KYK**
Citation	**Proceedings Of The Ieee Computer Society Conference On Computer Vision And Pattern Recognition, 2006, v. 1, p. 1238-1243**
Issued Date	**2006**
URL	**http://hdl.handle.net/10722/45552**
Rights	**©2006 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.**

Motion Estimation from Spheres

Guoqiang Zhang and Kwan-Yee K. Wong
Department of Computer Science, University of Hong Kong,
Hong Kong SAR, China
{gqzhang, kykwong}@cs.hku.hk

Abstract

This paper addresses the problem of recovering epipolar geometry from spheres. Previous works have exploited epipolar tangencies induced by frontier points on the spheres for motion recovery. It will be shown in this paper that besides epipolar tangencies, N^2 point features can be extracted from the apparent contours of the N spheres when $N > 2$. An algorithm for recovering the fundamental matrices from such point features and the epipolar tangencies from 3 or more spheres is developed, with the point features providing a homography over the view pairs and the epipolar tangencies determining the epipoles. In general, there will be two solutions to the locations of the epipoles. One of the solutions corresponds to the true camera configuration, while the other corresponds to a mirrored configuration. Several methods are proposed to select the right solution. Experiments on using 3 and 4 spheres demonstrate that our algorithm can be carried out easily and can achieve a high precision.

1. Introduction

In computer vision and computer graphics, many studies have been conducted on motion estimation. For the case of point correspondences, the geometric information has been well studied, like the computation of fundamental matrix between two views [13], trifocal tensor of three views [3] and point-based factorization algorithm of N views [9]. Recently, other features, like lines, conics and silhouettes of arbitrary objects, are being studied and exploited in motion estimation. Unlike point features, the use of line correspondences in motion estimation involves at least three images to derive constraints on the viewing geometry [10]. For the case of silhouettes, the epipolar tangents to the silhouettes are exploited for recovering epipolar geometry. In [2], Cross et al. developed an algorithm for motion estimation from point features and epipolar tangencies . In [11], Wong and Cipolla first dealt with the circular motion problem using the two outer epipolar tangencies. Silhouettes from general viewpoints are then registered with the circular motion sequence. In [8], Sinha et al. calibrated a camera network (internal parameters and external parameters) using epipolar tangents from dynamic silhouettes.

The literature is somewhat sparse when it comes to the problem of recovering camera geometry from conic correspondences. In [7], Schulz-Mirbach and Weiss proposed that the constraints on epipolar geometry can be derived from conic correspondences induced by planar conics, but no results had been presented. In [5], Kahl and Heyden extended their work to general quadrics, like spheres and ellipsoids, by deriving a different form of constraints on the epipolar geometry. The minimum data required are 4 quadrics with each one providing two constraints, which is essentially the same as the case of using silhouettes of objects, and the constraints from the two outer epipolar tangents can be nicely derived from the analytic expression of the images of quadric surfaces (conics). Recently, Kaminsky and Shashua proposed to recover epipolar geometry from general planar curves and spatial curves [6], of which the contribution is mainly on theorem.

Another motivation for the work stems from the need of calibrating a camera network, which is a necessary and important step in many computer vision studies. Calibrating such a large number of cameras using a general 3D reference object (like a planar pattern) is tedious and cumbersome because such a pattern will not be visible in all views simultaneously. Taking the sphere as the reference object can, however, easily overcome this problem. The work of recovering camera internal parameters from spheres has been addressed in [1, 12]. In [1], in addition to recovering camera internal parameters, Agrawal et al. also mentioned the possibility of motion estimation. However, only the sphere centers are exploited in their work and also the radii of all the spheres are required to be equal.

In this paper, we use spheres to recover the epipolar geometry. Note that 2 spheres together can be viewed as a surface of revolution. The envelops of planes bitangent to these 2 spheres form 2 cones with the 2 vertices lying on the axis of the surface of revolution (see Figure 1). We will show

that the images of these 2 vertices can be computed from the apparent contours of the spheres in the image. 3 spheres can provide 6 vertices, with all these points lying on the plane passing through the 3 sphere centers. Together with the epipolar tangent constraints (viewing geometry not on the plane where the vertices lie), the epipolar geometry can be recovered from the point correspondences of such vertices and the sphere centers. The advantage of the proposed method is straightforward. Besides epipolar tangent constraints, point features are shown to be extracted from the imaged spheres. The minimum number of spheres required is 3, which is advantageous over the work in [5]. Also the recovered points have high precision because of the good estimation of conics (sphere images). Experiments demonstrate that our algorithm can be carried out easily and can achieve a high precision.

The reminder of this paper is organized as follows. Section 2 briefly reviews the projection properties of the sphere surfaces. Section 3 describes how to extract the point features from the apparent contours of the spheres. Section 4 presents the algorithm for recovering epipolar geometry from 3 or more spheres. Section 5 gives the experimental results on using 3 and 4 spheres. A short conclusion is given in Section 6.

2. Projection of A Sphere

A sphere is a special kind of quadrics, of which the projection properties have been described in details in [4]. 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

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

A dual quadric $\mathbf{Q}^*$ is an alternative expression of the quadric $\mathbf{Q}$ in terms of the planes tangent to the quadric surface. The corresponding equation for a tangent plane π is $\pi^{\mathrm{T}}\mathbf{Q}^*\pi = 0$, where $\mathbf{Q}^*$= adjoint of $\mathbf{Q}$, or $\mathbf{Q}^{-1}$ if $\mathbf{Q}$ is invertible.

It is indicated in [4] that the projection of a quadric $\mathbf{Q}$ under the camera matrix $\mathbf{P}$ is a conic $\mathbf{C}$. The projection is formulated as

$$\mathbf{C}^* = \mathbf{P}\mathbf{Q}^*\mathbf{P}^{\mathrm{T}}, \quad (1)$$

where $\mathbf{C}^*$ is the dual conic. Considering a sphere with center $\mathbf{a}$ and radius r, under the camera matrix $\mathbf{P} = \mathbf{K}[\ \mathbf{I}\quad \mathbf{0}\ ]$, its image, the dual conic $\mathbf{C}^*$ is given by

$$\mathbf{C}^* = \mathbf{K}\mathbf{K}^{\mathrm{T}} - (\mathbf{K}\mathbf{a}/r)(\mathbf{K}\mathbf{a}/r)^{\mathrm{T}}. \quad (2)$$

$\mathbf{C}$ can be obtained as the inverse of $\mathbf{C}^*$. Hence the apparent contour of a sphere in the image can be analytically expressed as a conic, or equivalently as a dual conic.
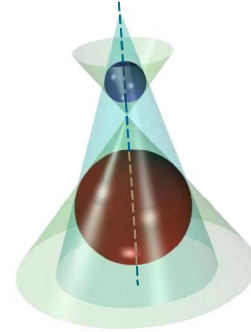

Figure 1. Two spheres with the two bitangent cones. The two vertices lie on the axis passing through the sphere centers.

3. Point Features from Imaged Spheres

From equation (2), it can be concluded that neither the imaged sphere center nor the image of any axis passing through the sphere center can be extracted from the image of a single sphere. Nevertheless, we can address the problem from another point of view by considering 2 spheres together. Note that 2 spheres can be viewed as a surface of revolution, with the axis of revolution passing through the 2 sphere centers. The envelopes of bitangent planes to the 2 sphere surfaces are 2 cones whose vertices lie on the axis as shown in Figure 1.

The projections of these 2 cone vertices in the image can be extracted from the apparent contours of the 2 spheres which are 2 conics. Consider the cone with its vertex lying outside the line segment connecting the 2 sphere centers. Without loss of generality, we set the center of Sphere 1 with radius r_1 to be coincided with the world origin (see Figure 2). The center of sphere 2 with radius r_2 is assumed to lie on the Z-axis with coordinates $(\ 0\quad 0\quad t\ )^{\mathrm{T}}$. Since the two spheres are invariant while rotating about the Z-axis, the camera center can be set lying on the X-Z plane with coordinates $(\ X_0\quad 0\quad Z_0\ )^{\mathrm{T}}$. The plane $(\ A\quad B\quad C\quad D\ )^{\mathrm{T}}$ passing through the camera center and tangent to both spheres (with both spheres lying on the same side of the plane) should satisfy

$$\begin{cases} AX_0 + CZ_0 + D = 0 \\ D/\sqrt{A^2 + B^2 + C^2} = r_1 \\ (Ct + D)/\sqrt{A^2 + B^2 + C^2} = r_2 \end{cases}. \quad (3)$$

The solution is derived to be

$$\begin{cases} D = r_1 \\ C = (r_2 - r_1)/t \\ A = (-Z_0 r_2 + Z_0 r_1 - t r_1)/(t X_0) \\ B = \pm\sqrt{1 - C^2 - A^2} \end{cases}. \quad (4)$$

Equation (4) reveals that there are two solutions which correspond to the two planes as indicated in Figure 2.
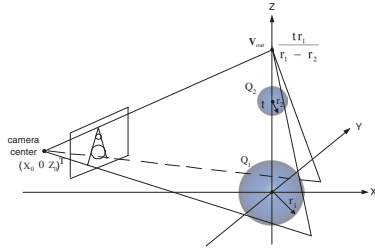
Figure 2. Bitangent planes to the 2 spheres with the intersection line passing through the camera center and $\mathbf{V}_{out}$. The 2 spheres are on the same side of the tangent planes.

Figure 3. Bitangent planes to the 2 spheres with the intersection line passing through the camera center and $\mathbf{V}_{in}$. The 2 spheres are on different sides of the planes.

It can be easily verified that the point $\mathbf{V}_{out}$ with co-ordinates $\begin{pmatrix} 0 & 0 & \frac{tr_1}{r_1-r_2} \end{pmatrix}^T$ lies on both tangent planes. Hence, the intersection line of the two planes is determined by the camera center and $\mathbf{V}_{out}$. Note that $\mathbf{V}_{out}$ is a point on the Z-axis and is determined only by the 2 spheres. It can thus provide a point correspondence between different views. It can be seen easily that these 2 tangent planes project as 2 lines in the image, which are tangent to the apparent contours of both spheres. The intersection point of these 2 lines gives the image of $\mathbf{V}_{out}$.

The analysis is similar for the other cone with its vertex lying within the line segment connecting the 2 sphere centers, (see Figure 3). Due to the fact that the 2 spheres should be located on different sides of the tangent plane, the set of equations now becomes

$$
\begin{cases}
AX_0 + CZ_0 + D = 0 \\
D/\sqrt{A^2 + B^2 + C^2} = r_1 \\
(Ct + D)/\sqrt{A^2 + B^2 + C^2} = -r_2
\end{cases}. \qquad (5)
$$

By simple deduction, the solution is

$$
\begin{cases}
D = r_1 \\
C = -(r_2 + r_1)/t \\
A = (Z_0 r_2 + Z_0 r_1 - t r_1)/(t X_0) \\
B = \pm\sqrt{1 - C^2 - A^2}
\end{cases}. \qquad (6)
$$

Similarly, the point $\mathbf{V}_{in} = \begin{pmatrix} 0 & 0 & \frac{tr_1}{r_1+r_2} \end{pmatrix}^T$ on the intersection line of the 2 planes is determined only by the 2 spheres. It can provide one more point correspondence between different views. Its image can be obtained as the intersection of the inner bitangents to the 2 conics in the image, as shown in Figure 3. For the case when the 2 conics partially overlap with each other, the bitangents become 2 complex vectors, of which the intersection is still the image of $\mathbf{V}_{in}$. The image of the axis can be readily obtained as the line passing through the images of $\mathbf{V}_{out}$ and $\mathbf{V}_{in}$.

From the above analysis, we know that 2 spheres can provide 2 point correspondences and the axis passing through the 2 sphere centers. It is easy to see that for 3 spheres, there are totally 9 point correspondences, 6 from
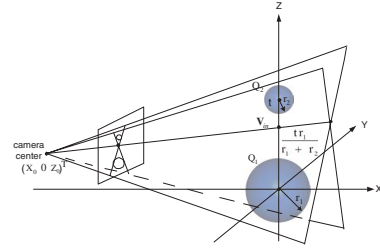
the vertices of the cones and 3 from the sphere centers (obtained as the intersections of the axis images). It should be noted that all these 9 points are lying on a single plane. This infers that these points only provide the geometry of the plane, which are not sufficient for determining the epipolar geometry (see [4]). To extend the conclusion, N spheres can provide N^2 point correspondences ($N > 2$). For the case of 4 or more spheres in general positions, the points should be distributed in a 3D world scene rather than on the same plane. Hence, the camera geometry can be extracted easily from the point features using classic algorithms [13, 3].

4. Recovery of Epipolar Geometry

4.1. Planar Parallax Geometry

Consider a pair of camera matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ with epipole $\mathbf{e}_i$, formed on the image of the camera $\mathbf{P}_i$. The fundamental matrix can be written in a plane plus parallax representation, given by [4]

$$
\mathbf{F} = \mathbf{H}^{-T}[\mathbf{e}_1]_\times = [\mathbf{e}_2]_\times \mathbf{H}, \qquad (7)
$$

where $\mathbf{H}$ is a homography between the two views induced by a world plane not passing through the two camera centers, and $[\mathbf{e}_2]_\times$ is the 3×3 skew matrix satisfying $[\mathbf{e}_2]_\times \mathbf{x} = \mathbf{e}_2 \times \mathbf{x}$. $\mathbf{H}$ 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 2 more points not on the plane for inducing the homography [4].

In [2], Cross et al. demonstated that if the apparent contour of the object in the first view is transferred to the second view via a homography induced by a world plane, the outer bitangents to the apparent contour and the transferred contour in the second view should be epipolar lines, and their intersection point gives the epipole. This situation is illustrated in Figure 4. Backprojecting the apparent contours $\mathbf{S}_1$ and $\mathbf{S}_2$ of the object from both views gives 2 contours $\mathbf{S}_1'$ and $\mathbf{S}_2'$ in the world plane π. 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 plane π with this epipolar plane, being bitangent to these two con-
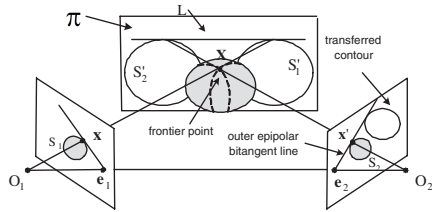
Figure 4. Epipolar bitangency. The 2 contours $\mathbf{S}'_1$ and $\mathbf{S}'_2$ on the plane π are the intersections of the plane π with the 2 cones formed by backprojecting the apparent contours $\mathbf{S}_1$ and $\mathbf{S}_2$ of the object in both views, respectively. The transferred contour in View 2 is obtained by projecting $\mathbf{S}'_1$ into View 2. $\mathbf{x}$ and $\mathbf{x}'$ are the corresponding outer epipolar tangent points.

tours. It follows that the outer bitangent to the images of the 2 contours $\mathbf{S}'_1$, $\mathbf{S}'_2$ in View 2 is an epipolar line. Hence, the epipole $\mathbf{e}_i$ can be determined by two outer epipolar bitangent lines, and the fundamental matrix can thus be recovered by (7).

4.2. Recovering Epipoles

In our work of recovering two view geometry from 3 spheres, a homography $\mathbf{H}$ between the two views is first computed from the point correspondences induced by the spheres (see Section 3). Note that the world plane is the one passing through the 3 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 [11, 8], generally, the silhouettes of an object between two views can provide 2 outer epipolar tangent points, which correspond to 2 frontier points on the surface of the object (see Figure 4). This implies that 3 spheres can provide 6 frontier points, for the computation of the fundamental matrix. Under perspective projection, the frontier points would not all lie on the plane for inducing the homography simultaneously, or it is equivalent to say that they 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 apparent contour of a sphere is analytically expressed as a conic, the transformation of the apparent contours and further the computation of epipoles can be performed easily.

Note that the 3 spheres are bilateral symmetric w.r.t. the plane passing through the 3 sphere centers. This geometry induces 2 epipoles in both View 1 and View 2, one for the real solution and the other one for the projection of a virtual camera center, as shown in Figure 5. Note that the apparent contour and the transferred contour in Figure 5(c) are very close to each other. This situation would, however, not affect the computation of bitangent epipolar lines. This is because the conics representing the contours can be estimated with high precision. Now let's consider the appear-
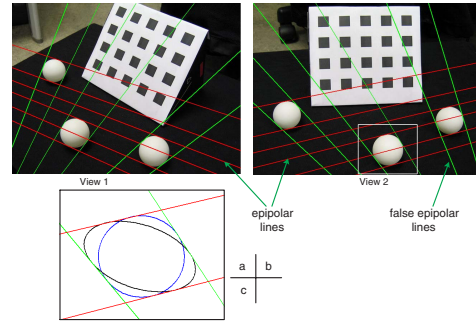

Figure 5. (a) and (b) are two views of a world scene with 3 spheres and a planar pattern. Each apparent contour and the corresponding transferred contour provide 4 outer bitangent lines. 2 of them are the expected epipolar lines, and the other 2 the epipolar lines w.r.t. a virtual camera. (c) shows the apparent contour and the transferred contour in the marked rectangular in View 2.
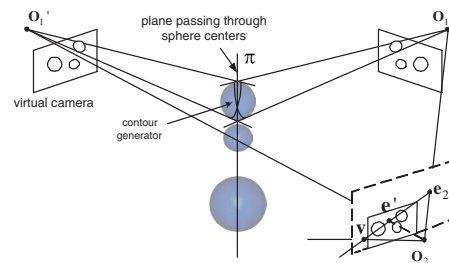

Figure 6. The virtual camera and Camera 1 are bilateral symmetric w.r.t. the plane π passing through the 3 sphere centers. Both real cameras are on one side of plane π, and the virtual camera is on the other side. $\mathbf{v}$ is the vanishing point of the norm direction of plane π.

ance of the virtual camera by taking View 2 as an example. As shown in Figure 6, the virtual camera is a mirror version of Camera 1 about the plane π. Hence backprojecting the apparent contours of a sphere from Camera 1 and the virtual camera gives the 2 identical contours on the plane π. Projecting these 2 identical contours onto View 2 results in 4 outer bitangent epipolar lines, 2 for the virtual camera and 2 for Camera 1. Correspondingly, the intersection points are the projections of the virtual camera center $\mathbf{O}'_1$ and the center of Camera 1 $\mathbf{O}_1$. The analysis for View 1 is similar.

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 Figure 5 (a), (b). If the radii of the 3 spheres are equal, the true epipole can be automatically selected instead. As shown in Figure 6, the line connecting $\mathbf{O}'_1$ and $\mathbf{O}_1$ is orthogonal to plane π. Its image line $\mathbf{e}_2\mathbf{e}'$ in View 2 should pass through vanishing point $\mathbf{v}$. For the case where both $\mathbf{O}_1$ and $\mathbf{O}_2$ are on the same side of π, the 3 points $\mathbf{v}$, $\mathbf{e}'$ and $\mathbf{e}_2$ keep the relationship that $\mathbf{e}'$ is within the line seg-

ment $\mathbf{v}\mathbf{e}_2$. This relationship can be exploited to get the true epipole. It is known that $\mathbf{v}$ and the vanishing line $\mathbf{l}$ of π have a pole-polar relationship w.r.t. the *absolute conic* $\boldsymbol{\omega}$, given by

$$\mathbf{l} = \boldsymbol{\omega}\mathbf{v} = \mathbf{K}^{\text{-T}}\mathbf{K}^{-1}\mathbf{v},$$

where $\mathbf{K}$ is the calibration matrix [4]. When the radii of the spheres are equal, it is not difficult to conclude that the image of $\mathbf{V}_{out}$ in Figure 4 is lying on $\mathbf{l}$. 3 spheres can provide 3 such points, which are sufficient to determine $\mathbf{l}$. As for $\mathbf{K}$, it can be easily recovered from 3 spheres by applying the algorithm presented in [12]. $\mathbf{v}$ can thus be computed from the obtained $\mathbf{l}$ and $\mathbf{K}$. The real epipole $\mathbf{e}_2$ is chosen to be further away from $\mathbf{v}$. On the contrary, if the two cameras are on different sides of the plane π, the one closer to $\mathbf{v}$ is chosen instead. If the radii of the 3 spheres happen to be unequal, vanishing point $\mathbf{v}$ can still be recovered by involving a third view. Like the line $\mathbf{e}_2\mathbf{e}'$ in Figure 6, the third camera provides another line passing through $\mathbf{v}$. This line together with line $\mathbf{e}_2\mathbf{e}'$ determine $\mathbf{v}$ uniquely.

To conclude, 3 spheres can provide 9 point features (planar geometry) and 6 frontier points (viewing geometry off the plane determined by the 9 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 (N > 3) spheres, there should be N^2 point features and 2N frontier points for two views. The point features alone can determine the epipolar geometry using classic algorithms as discussed in Section 3. 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 additional spheres break the symmetry in the configuration. Compared with the algorithm in [5] which requires a minimum number of 4 quadrics in the computation of fundamental matrix, our method needs only 3 spheres, and further it can be easily performed.

5. Experimental Results

To test the usefulness and effectiveness of our algorithm, experiments on using 3 and 4 spheres, respectively, were conducted. For the 3 spheres case, both simulated and real data experiments were performed. The camera used in the real data experiment was a Canon A95 digital camera.

5.1. Three Spheres Experiment

To understand the implementation of the algorithm, the procedure is briefly summarized below:
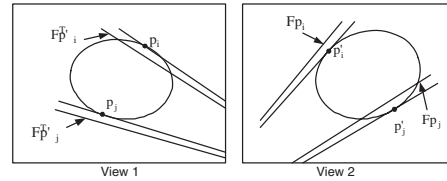
1. Estimate the conics in the two images.

Figure 7. Two views of a sphere with an associated fundamental matrix $\mathbf{F}$. $\mathbf{p}$ and $\mathbf{p}'$ are the outer epipolar tangent points. $\mathbf{Fp}$ and $\mathbf{F}^{\text{T}}\mathbf{p}'$ are the corresponding epipolar lines.

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

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

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

In Step 2 of the algorithm, only 6 of 9 point features were exploited in the computation of $\mathbf{H}$. This is because the projections of $\mathbf{V}_{out}$ are 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 3 is refined further. The cost function to be minimized is the residual error given by the geometric distances between the outer epipolar tangent points and the transformed epipolar lines as shown in Figure 7, written as

$$\boldsymbol{\xi}(\mathbf{e}) = \frac{1}{2n}\sum_{i=1}^{n} d(\mathbf{p}_i', \mathbf{F}(\mathbf{e})\mathbf{p}_i) + d(\mathbf{p}_i, \mathbf{F}^{\text{T}}(\mathbf{e})\mathbf{p}_i'), \quad (8)$$

where n is the number of froniter points and function d is the point-to-line Euclidean distance expressed in pixels.

In the synthetic data experiment, 3 spheres with unequal radii were generated and rendered with different colors. The image resolution was 640×480. Two views of the scene were taken in the recovery of the epipolar geometry as shown in Figure 8. The apparent contours of the spheres were extracted using Canny's edge detector and conics were fitted to these contours by a robust least square ellipse fitting algorithm. After refinement of the epipole $\mathbf{e}$, the residual error $\boldsymbol{\xi}$ is not zero but 0.0216 pixels. This may be due to the fact that there exist quantization noise in the rendered image, and the conics failed to represent the apparent contours without error. The ground truth for the epipole $\mathbf{e}_i$ is shown in the first row of Table 1, and in the second row are the recovered parameters by our algorithm. The third row displays the difference of the parameters in pixels. The errors in the table are quite small compared with the ground truth. The maximum distance between the recovered feature points over the two images (not including $\mathbf{V}_{out}$) and the ground truth is 0.14 pixels. The point $\mathbf{V}_{out}$ is unstable, and the error would become bigger as it moves away from the image center.
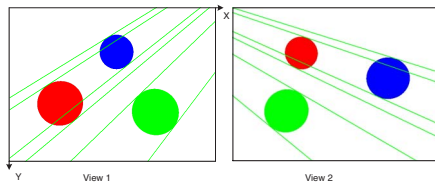
Figure 8. Two views of the 3 synthetic spheres. The lines in the image are the recovered outer tangent epipolar lines.

Data	$\mathbf{e}_{1x}$	$\mathbf{e}_{1y}$	$\mathbf{e}_{2x}$	$\mathbf{e}_{2y}$
Ground truth	1074.1	-375.28	-551.61	-174.37
Recovered	1083.2	-382.95	-542.54	-169.95
Difference(pixels)	9.2	7.7	9.1	4.4

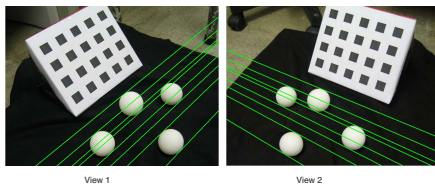Table 1. Epipoles for synthetic data experiment.

Figure 9. Two views of the 4 spheres and a planar pattern. The lines in the image are the computed outer tangent epipolar lines.

In the real data experiment, 3 ping pong balls were used, and a planar pattern was taken for evaluating the performance. The image revolution was 640×480. The two views of the spheres and the pattern are shown in Figure 5(a),(b). After refinement of the epipole, the residual error $\boldsymbol{\xi}$ is 0.022 pixels. By using the obtained fundamental matrix, the computed residual error for the points on the planar pattern is 0.16 pixels.

5.2. Four Spheres Experiment

To test the accuracy of the recovered point features, the points from 4 spheres were only exploited in the computation of the fundamental matrix. In the experiment, 4 spheres were placed in general positions to ensure that the point features extracted from the spheres would not lie on the same plane. The two views are shown in Figure 9. Like the 3 spheres experiment, 6 of the 16 point features were not included in the computation of the fundamental matrix. The residual error of the frontier points is 0.04 pixels. The points from the planar pattern were then exploited to evaluate the performance, and the residual error $\boldsymbol{\xi}$ is 0.32 pixels, which demonstrates the high precision of the obtained fundamental matrix.

6. Conclusion

We have proposed a simple and effective algorithm for recovering the epipolar geometry from spheres. We found that besides the outer epipolar tangent constraints, point features over the views could be extracted from the apparent contours of the spheres. We have shown how the point features and the outer epipolar tangent constraints from 3 spheres can be combined nicely in the recovery of the fundamental matrix. The method does not require any initial guess of parameters nor involve any high dimensional non-linear minimization. The procedure of recovering epipolar geometry from 4 or more spheres is discussed. The experiments on using both 3 and 4 spheres demonstrate the robustness and the effectiveness of our algorithm.

References

[1] M. Agrawal and L. Davis. Camera calibration using spheres: a semi-definite programming approach. In *Proc. 9th Int. Conf. on Computer Vision*, pages 782–789, 2003.

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

[3] R. Hartley. Lines and points in three views and the trifocal tensor. *Int. Journal of Computer Vision*, 22(2):125–140, 1997.

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

[5] F. Kahl and A. Heyden. Using conic correspondence in two images to estimate the epipolar geometry. In *Proc. 6th Int. Conf. on Computer Vision*, pages 761–766, 1998.

[6] J. Y. Kaminski and A. Shashua. Multiple view geometry of general algebraic curves. *Int. J. Comput. Vision*, 56(3):195–219, 2004.

[7] H. Schulz-Mirbach and I. Weiss. Projective reconstruction from curve correspondences in uncalibrated views. Technical Report TR-402-94-014, 1994.

[8] S. Sinha, M. Pollefeys, and L. McMillan. Camera network calibration from dynamic silhouettes. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages I: 195–202, 2004.

[9] P. Sturm and B. Triggs. A factorization based algorithm for multi-image projective structure and motion. In *Proc. 4th European Conf. on Computer Vision*, pages II:709–720, 1996.

[10] J. Weng, T. Huang, and N. Ahuja. Motion and structure from line correspondences: Closed-form solution, uniqueness, and optimization. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(3):318–336, March 1992.

[11] K.-Y. K. Wong and R. Cipolla. Structure and motion from silhouettes. In *Proc. 8th Int. Conf. on Computer Vision*, volume II, pages 217–222, Vancouver, BC, Canada, July 2001.

[12] H. Zhang, G. Zhang, and K.-Y. K. Wong. Camera calibration with spheres: Linear approaches. In *Proc. International Conference on Image Processing*, Genova, September 2005.

[13] Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. *Int. Journal of Computer Vision*, 27(2):161–195, March 1998.