<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>Polygonal light source estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Schnieders, D; Wong, KYK; Dai, Z</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>The 9th Asian Conference on Computer Vision (AACV 2009), Xi’an, China, 23-27 September 2009. In Lecture Notes in Computer Science, 2009, v. 5996, p. 96-107</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2010</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/144830">http://hdl.handle.net/10722/144830</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>The original publication is available at <a href="http://www.springerlink.com">www.springerlink.com</a></td>
</tr>
</tbody>
</table>
Abstract. This paper studies the problem of light estimation using a specular sphere. Most existing work on light estimation assumes distant point light sources, while this work considers an area light source which is estimated in 3D space by reconstructing its edges. An empirical analysis on existing methods for line estimation from a single view is carried out, and it is shown that line estimation for a single view of a sphere is an ill-conditioned configuration.

By considering a second identical sphere, a closed form solution for single view polygonal light estimation is proposed. In addition, this paper also proposes an iterative approach based on two unknown views of just a single sphere. Experimental results on both synthetic and real data are presented.

1 Introduction

The calibration of light sources plays an important role in both computer graphics and computer vision. For instance, combining computer-generated models with the real world, as in augmented reality, requires known light positions for realistic rendering. Many computer vision techniques make the common assumption of distant point light sources, which conveniently reduces the complexity in modeling the image formation and allows a simple light source position estimation. As an example, consider the classic shape from shading (SfS) technique that recovers the 3D shape of an object by relating the intensity values to the normal vectors and light source direction. Motivated by the possibility that SfS could be extended to deal with area light sources, this paper studies the problem of recovering polygonal area light sources from images of a specular sphere.

There exists a relatively large amount of research dealing with distant point illuminant estimation, and many early results were published in the context of SfS [123]. A survey of those and related methods can be found in [4]. Other related work on light estimation includes a method developed by Yang and Yuille [5] which estimates light directions from a Lambertian sphere by locating the occluding boundary of light sources. Their method was extended by Zhang and Yang [6] who introduced the concept of critical points which have their normal vectors perpendicular to the light direction. Later Wang and Samaras [7] further extended this method and estimated light from a single view of an object with
known geometry by a robust minimization of a global error. Wong et. al. [8] introduced a method for recovering both the light directions and camera poses from a single sphere of unknown radius. In addition to the estimation of light directions, Takai et. al. [9] proposed a method for estimation of near point light sources and ambient light from a single image. This is achieved by employing a pair of reference spheres as light probes.

Because of the relative high complexity involved in estimating area light sources, the literature in this field is rather sparse. In [10], Debevec estimated global illumination in the context of augmented reality without assuming specific types of light source, but did not estimate parameters such as distance and size of the illuminants.

In this paper an area light source is recovered from an image of the specular highlight it produces on a sphere. Unlike Zhou and Kambhamettu’s method [11] which uses an iterative approach for estimating area light sources from specularities observed on two spheres, this paper provides a closed form solution by treating an area light source as a polygon in 3D space composed of a set of lines. Such lines are independently determined as the intersections of the reflection rays. We call this type of light source a polygonal light source.

There exists previous work on estimating a line from a single view. Lanman et. al. [12] formulated the problem and solved it in theory, but practical results remained inaccurate. In [13], the authors solved their inaccurate line estimation by carefully estimating all parameter of their system. In this paper it will be shown that even with ground truth calibration, single view line estimation from a single view of a sphere cannot be accurately solved. A closed form solution of two spheres is therefore proposed. This paper also develops an iterative approach based on two unknown views of just a single sphere. The rotation relating the two unknown views can be estimated by assuming a rectangular light source.

A related paper by Gasparini and Sturm [14] also utilizes 3D lines in non-central images. In their interesting work, a system that deals with the structure from motion problem for general camera models is proposed.

The rest of the paper is organized as follows. Section 2 considers the problem of line estimation from a single view. In the first part of this section, a theoretical description for the estimation is given, while the second part provides empirical results which show that single view line estimation is an ill-conditioned problem. Knowing that a single view of a single sphere is insufficient, Section 3 formulates the problem for (a) a single view of two spheres and (b) two views of a single sphere. Experimental results on both synthetic and real data are shown in Section 4 followed by conclusions in Section 5.

2 Line Estimation from a Single View

Consider a pinhole camera viewing a line $L$. Together with the camera center the line defines a plane. Any line lying on this plane will project to the same line on the image, which makes line reconstruction from a single view ambiguous.

Instead of considering the image $l$ of the line, let us now consider an image $l_s$ of the line $L$ formed by its reflection on a sphere. While $l$ is a 2D line, $l_s$ will in
general be a curve. The viewing lines defined by the back-projection of the points on $l_s$ will intersect the sphere and reflect according to the law of reflection. The resulting rays, which we call reflection rays will intersect $\mathcal{L}$. In contrast to the case of a pinhole camera viewing a line, the reflection rays will not meet at a single point, and this enables line reconstruction from a single view.

**Proposition 1.** The reflection rays constructed from an image of the reflection of a line $\mathcal{L}$ on a sphere will intersect two lines, namely the line $\mathcal{L}$ and a line $\mathcal{A}$ passing through the sphere center and the camera center.

**Proof.** Let us denote the back-projection of a point $x \in l_s$ as the viewing line $\mathcal{V}$ and its reflection on the sphere with center $S$ as the reflection line $\mathcal{R}$. The viewing line will leave the camera center $C$, pass through the point $x$, and intersect the sphere at a point $P$. Let $V$ and $R$ be the unit vectors in the directions of the viewing line and the reflection line respectively. The law of reflection states that the incident angle must be equal to the reflection angle, and the reflection direction is therefore given by $R = (2N \cdot V)N - V$, where $N$ is the unit normal vector at point $P$. The reflection line $\mathcal{R}$ passes through $P$ in the direction $R$ and will, by construction, intersect the line $\mathcal{L}$ at some point $L$. All the reflection rays constructed in such a way will intersect the line $\mathcal{L}$.

To show the intersection with the other line, note that the lines $\mathcal{V}$, $\mathcal{R}$ and $\mathcal{N}$ are coplanar, where $\mathcal{N}$ is defined as the line from $S$ in direction $N$. As the camera center $C$ is on $\mathcal{V}$ and the sphere center $S$ is on $\mathcal{N}$, it follows that the line $\mathcal{A}$ from the camera center $C$ to the sphere center $S$ also lies on the same plane as $\mathcal{V}$, $\mathcal{R}$ and $\mathcal{N}$, and making an angle of $\gamma = 180 - (\alpha + \beta)$ with $\mathcal{R}$, where $\alpha = \angle(N, R)$ and $\beta = \angle(A, N)$ (see Fig. 1). This applies to all reflection rays and it follows that any reflection ray $\mathcal{R}$ will intersect $\mathcal{A}$ and $\mathcal{L}$.

In 1874, Schubert published his work *Kalkül der Abzählenden Geometrie* in which he showed that the number of lines intersecting four arbitrary lines will be zero, one, two or infinite [15]. Unless the four lines lie on a doubly ruled surface

![Fig. 1. The image of a line $\mathcal{L}$ on a sphere with center $S$ is determined by reflected viewing rays $\mathcal{R}$ which will intersect two lines, the line $\mathcal{L}$ and a line $\mathcal{A}$ passing through $S$ and camera center $C$](image-url)
(single sheet hyperboloid, hyperbolic paraboloid or plane), they will not produce infinite intersecting lines [14,16]. Therefore in the general case the reflection rays in our system will intersect at most two lines, and under a degenerate case the reflection rays lie on a doubly ruled surface and produce infinite intersecting lines. In practice the degenerate case can easily be detected, because the reflection rays will produce a Pl¨ucker hyperplane matrix with a nullspace larger than two dimensions.

**Corollary 1.** Reconstruction of a line from its reflection on a sphere becomes possible by solving for the two lines intersecting its reflection rays and selecting the one which does not pass through the camera center.

### 2.1 Line Intersection in Pl¨ucker Space

In order to formulate line intersections algebraically, we adopt the 6-vector Pl¨ucker line coordinates representation for directed lines in $P^3$ [17]. Two points $P = (p_x, p_y, p_z, 1)$ and $Q = (q_x, q_y, q_z, 1)$ define a line $L$ as

$$L = (l_0, l_1, l_2, l_3, l_4, l_5) = (p_x q_y - q_x p_y, p_x q_z - q_x p_z, p_x - q_x, p_y q_z - q_y p_z, p_z - q_z, q_y - p_y).$$  \hfill (1)$$

With this notation, lines in $P^3$ are mapped to homogeneous points $L = (l_0, l_1, l_2, l_3, l_4, l_5)$ or hyperplanes $\overline{L} = (l_4, l_5, l_3, l_2, l_0, l_1)$ in 5 dimensional Pl¨ucker coordinate space.

A major advantage of this representation is the simplicity of the incidence operation. Given two lines $A$ and $B$, the incidence operation is the inner product between the homogeneous Pl¨ucker representation of line $A$ and the hyperplane Pl¨ucker representation of line $B$

$$A \cdot B = a_0 b_4 + a_1 b_5 + a_2 b_3 + a_3 b_2 + a_4 b_0 + a_5 b_1.$$  \hfill (2)$$

Since the inner product will be zero for intersecting lines, solving for $n$ lines $I_1, I_2, ..., I_n$ that intersect $m$ given lines $L_1, L_2, ..., L_m$ is equivalent to finding the $n$-dimensional nullspace of a matrix formed by the Pl¨ucker hyperplane representations of the given lines:

$$Mx = \begin{bmatrix} \overline{L}_1 \\ \overline{L}_2 \\ \vdots \\ \overline{L}_m \end{bmatrix}, \quad x = 0.$$  \hfill (3)$$

Finding the set of lines $x$ that map $M$ to a null vector, implies that for each row $i$ the inner product $\overline{L}_i \cdot x$ equals zero. Given the $m$ reflection lines from the previous section and the task of finding the $n = 2$ intersecting lines, we can simply solve for those lines by finding the nullspace of the matrix $M$ with singular value decomposition.
\[
M = U \Sigma V^T = \begin{bmatrix}
  u_{11} & \cdots & u_{1m} \\
  \vdots & \ddots & \vdots \\
  u_{m1} & \cdots & u_{mm} \\
\end{bmatrix}
\begin{bmatrix}
  \sigma_1 \\
  \vdots \\
  0 \\
  \vdots \\
  0 \\
\end{bmatrix}
\begin{bmatrix}
  v_{11} & \cdots & v_{16} \\
  \vdots & \ddots & \vdots \\
  v_{61} & \cdots & v_{66} \\
\end{bmatrix}^T.
\]
(4)

For \( n = 2 \), \( M \) is a rank four matrix and will span a two dimensional subspace that can be parameterized by the two points \( a = (v_{15}, \ldots, v_{65}) \) and \( b = (v_{16}, \ldots, v_{66}) \), which correspond to the two smallest singular values \( \sigma_5 \) and \( \sigma_6 \). Fortunately not all points on the 5-dimensional line \( \mathcal{L}(t) = at + b \) are 3-dimensional lines, but just those lines \( \mathcal{A} \) that satisfy
\[
\mathcal{A} \cdot \overline{\mathcal{A}} = 0.
\]
(5)
Teller and Hohmeyer were the first to formulate and solve this problem by intersecting the line \( \mathcal{L}(t) \) with all points that satisfy (5). This produces the quadratic equation
\[
(a \cdot \overline{a})t^2 + 2(a \cdot \overline{b}) + (b \cdot \overline{b}) = 0,
\]
(6)
for which the two real roots correspond to the two intersecting lines. As a result of proposition 1, the nullspace of matrix \( M \) will in general be two dimensional. In practice nearly coplanar reflection lines will result in a nullspace with higher dimensions. Due to numerical instabilities, incorrect solutions will likely be selected. An empirical analysis has been performed which considers this problem and is described in the next section.

2.2 Empirical Analysis

In order to study the feasibility of the theoretical formulation above, we analyze synthetically generated images empirically. A specular sphere reflecting a line was rendered using an OpenGL Shading Language program. The reflected line light source was detected by thresholding and subsequent Bezier spline curve fitting, which allowed sub-pixel-accurate sample points for computing the reflection lines. As we are dealing with synthetic data, all camera and sphere parameters are readily available and the reflection lines can be determined. One practical concern about the theoretical formulation in Section 2 is the numerical instability in the case of nearly coplanar reflection lines. Coplanar reflection lines are undesirable because all lines on that plane will intersect the reflection lines and as a result the nullspace selection will be unstable. We use the average angle that lines make with a best fitting plane, as a planarity measurement.

In the first experiment we rendered a synthetic sphere of radius \( S_r = 1 \) and center \( S = (0, 0, 0)^T \) with a synthetic camera. Without loss of generality, let the camera be located at \( C = (0, 0, -5)^T \) pointing in the negative z-direction and a line be positioned parallel to the x-axis piercing the y-axis at \( L = (0, -5, 0) \). Let \( A \) be the vector from \( S \) to \( C \) and \( B \) the shortest vector from \( S \) to the line light source. The distance \( |A| \) between the camera center and the sphere center as
Fig. 2. This plot shows the plane fitting error against the viewing angle. A linear relationship between this angle and the planarity measurement can be identified, with better results for larger angles.

well as the distance $|B|$ between the line and the sphere center is kept constant at 5 units, while the camera is being rotated around the x-axis in 10 degree intervals. Let us define the angle $\angle(A, B)$ as the viewing angle. Fig. 2 plots the plane fitting error against this viewing angle. A linear relationship between the viewing angle and the planarity measurement can be identified, with better results (smaller error) for larger viewing angles.

Fig. 3 plots the plane fitting error while translating the light source along y-axis (left) and translating the camera center along z-axis (right) with constant viewing angle. These plots show a strong relationship between light and camera distances to the plane fitting error. Shorter distances between sphere center and light produce least coplanar reflection rays compared to larger distances. The opposite is true for distances between sphere center and camera center.

The synthetic experiments above show that the coplanarity of the reflection rays depends on the two distances $|A|$ and $|B|$ as well as the viewing angle. An increase in the plane fitting error can be observed with large viewing angles, small light distances and relatively large camera distances. This is undesirable

Fig. 3. Plots for plane fitting error while translating the light source along the y-axis (left) and translating the camera center along the z-axis (right) with constant viewing angle
for line reconstruction as it prevents robust single view line estimation from spherical reflections under general conditions. Apart from this, the experiments in this section show that the reflection lines are often close to coplanar, with a maximum plane fitting error of about 3 degrees, which is insufficient for accurate line estimation in practical situations.

3 Estimation of a Polygonal Light Source

In this section, we apply the theoretical formulation developed above for the problem of polygonal light source estimation by treating the light source as a polygon composed of a set of lines. Each line in the set will be independently reconstructed and a polygon will be calculated from the reconstructed lines. It has been shown in the previous section that in practice a line cannot be uniquely reconstructed from a single view of a single sphere, and setups with an additional sphere or an additional view are considered.

In the following, two procedures that estimate polygonal light sources given just the intrinsic camera parameters $K$ are considered. Section 3.2 introduces a closed form solution given a single view of two spheres of the same radius, while Section 3.3 introduces an iterative approach based on two unknown views of a single sphere. Both of the methods require a known translation between the camera center $C$ to the sphere center $S$. For this reason a solution for estimating the sphere center from its silhouette is described in the following section.

3.1 Where Is the Sphere?

The sphere silhouette, being a conic, can be represented by a 3x3 symmetric matrix $C_{sil}$, given by

$$C_{sil} = (PQ_s^*P^T)^* = (KK^T - (KS/S_r)(KS/S_r)^T)^*,$$

(7)

where $Q_s^*$ denotes the dual to the quadric $Q_s$, which represents the sphere with center $S$ and radius $S_r$. Here the pinhole camera is given as $P = K[I 0]$.

In order to recover the sphere center $C$ from $C_{sil}$, the effect of $K$ is first removed by normalizing the image using $K^{-1}$. The conic $C_{sil}$ will be transformed to a conic $\hat{C}_{sil} = K^TCK$ in the normalized image. This conic $\hat{C}_{sil}$ can be diagonalized into

$$\hat{C}_{sil} = MDM^T = M \begin{bmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix} M^T,$$

(8)

where $M$ is an orthogonal matrix whose columns are the eigenvectors of $\hat{C}_{sil}$, and $D$ is a diagonal matrix consisting of the corresponding eigenvalues. The matrix $M^T$ defines a rotation that will transform $\hat{C}_{sil}$ to the circle $D$ with radius
\[ r = \sqrt{-\frac{b}{a}} \text{ centered at the origin. This transformation corresponds to rotating} \]
\[ \text{the camera about its center until its principle axis passes through the sphere} \]
\[ \text{center. The distance } d \text{ between the camera center and sphere center is given as} \]
\[ d = S_r \frac{\sqrt{1+r^2}}{r}. \]  
\[ (9) \]

Finally, the sphere center can be recovered as
\[ C = M [0 0 d]^T \]
\[ = dm_3, \]
\[ (10) \]
where \( m_3 \) is the third column of \( M \).

3.2 Two Spheres and a Single View

This section gives a solution for polygonal light estimation by introducing a second, identical sphere into the scene. Two identical spheres for light estimation have been utilized before [19]. In contrast to previous work, this section provides a closed form solution. Note that the iterative method provided by Zhou et. al. [19] gives no guarantee for convergence.

Firstly the relative locations of both identical spheres are estimated by the method given in section 3.1. The reflecting rays \( R_1, ..., R_m \) for the first sphere and the reflecting rays \( R'_1, ..., R'_n \) for the second sphere will form the equation
\[ Mx = \begin{bmatrix} R_1 \\ \vdots \\ R_m \\ R'_1 \\ \vdots \\ R'_n \end{bmatrix} x = 0. \]
\[ (11) \]

The reflection lines for a single sphere will be relatively coplanar. However, by including the reflection lines from a second sphere, an intersecting line can be determined, which corresponds to an edge of the light source. Repeating this process for all sides of the polygonal light source results in \( n \) lines in space. Correspondence for the \( n \) sides of the light source can be achieved easily because the order of the edges will not change in the specular reflection.

3.3 Two Views of a Single Sphere

In the following method two images of a single sphere are taken from two distinct viewpoints. Projection matrices for the two cameras can be written as
\[ P_1 = K_1[I \ C_1] \]
\[ P_2 = K_2[E \ C_2], \]
\[ (12) \]
where \( K_1 \) and \( K_2 \) are the two camera calibration matrices (assumed to be known) and \( C_1 \) and \( C_2 \) are determined from the method described in Section 3.1. The sphere center is the world origin and the first camera is chosen as a reference view. The unknown 3x3 rotation matrix \( E \) is independent from the image \( C_{sil} \) of the sphere but can be determined from the specular reflection. This follows from the fact that the location of the highlight on the sphere surface depends on the cameras location as well as the light location.

Given the correct rotation matrix \( E \), the reflection lines for both views corresponding to an edge of the light source can be determined. Note that the resulting reflection lines for a single view will be relatively coplanar, but by including the reflection lines from the other view, an intersecting line can be determined which gives an edge of the light source. This process is repeated for all edges of the light source.

To solve for the unknown rotation, an optimization over the 3-dimensional rotation space is performed using the cost

\[
E_{\text{cost}} = w_1 \alpha_r + w_2 d_r + k,
\]

where \( w_1 \) and \( w_2 \) are weight coefficients determined experimentally, \( \alpha_r \) is the average angle between the reflection lines and \( d_r \) the average distance between them.

To add additional constraint on the rotation \( E \), a rectangular light source is assumed and \( k \) is a measurement of how rectangular the given lines \( \{W, X, Y, Z\} \) are, and is defined as

\[
k = \angle(W, Y) + \angle(X, Z) + |\angle(W, X) - 90| + |\angle(Y, Z) - 90| + |\angle(Y, Z) - 90| + |\angle(Z, W) - 90|.
\]

Instead of performing an optimization directly on the parameterized search space, an initial global minimum is found by subdividing the search space. The optimization is subsequently initialized with the global minimum of the subdivision. This procedure avoids an early termination of the optimization in a local minimum.

### 4 Experimental Results

The closed form solution of Section 3.2 and the iterative method of Section 3.3 for recovering a polygonal light source have been implemented. Experiments on both synthetic and real data were carried out, and the results are presented in the following sections.

#### 4.1 Synthetic Experiments

A synthetic experiment with a rectangular light source and (a) a single view of two spheres and (b) two views of a single sphere has been performed. An
OpenGL Shading Language program has been written to model the specular reflection of the polygonal light sources on the spheres. Both the silhouette of the spheres and the edges of the specular highlight were extracted automatically by thresholding and subsequent fitting of Bezier spline curves.

All ground truth data is available for the edges of the light source and estimation errors can be calculated as the angle between the estimated edges and ground truth edges. Their average was 3.77 degree for experiment (a) and 1.63 degree for experiment (b). The distance between those lines was used as a second error measurement and their average was 0.079S_\text{r} for experiment (a) and 0.058S_\text{r} for experiment (b), where S_\text{r} is the radius of the sphere.

4.2 Experiments on Real Data

For the first experiment on real data, two identical plastic white cue snooker balls were imaged from a single viewpoint. The spheres were put below a standard rectangular fluorescent office lamp. The light source has a dimension of 270mm x 1170mm while the snooker balls’ diameter is 57mm. The intrinsic parameters of the camera were obtained using Zhang’s camera calibration method [20]. Cubic Bezier-spline snake was applied to extract the contours of the sphere in the images, and conics were then fitted to these contours using a direct least squares method [21]. Edges from the specular reflection of the rectangular light source were picked and matched manually. One of the spheres is shown in a crop of the image in Fig. 4(a), and Fig. 4(b) illustrates a synthetically generated view of the sphere reflecting the estimated light source. We compared the size of

![Fig. 4](image)

Fig. 4. (a) A standard rectangular fluorescent office lamp is illuminating a plastic white cue snooker ball. (b) A synthetically generated view of the sphere reflecting the estimated light source for the two sphere single view case. (c) A blue snooker ball is enlightened by a smaller rectangular desk light source. (d) The synthetically generated view with estimated light source for the two view single sphere case.
the estimation result with the specification of the light source and found an approximate error of $13mm \times 52mm$ for the $270mm \times 1170mm$ light source.

Two views of a single sphere were taken in a second experiment. This time a blue snooker ball was enlightened by a smaller rectangular desk light source with a dimension of $68mm \times 33mm$. A crop of the image is shown in Fig. 4(c), and Fig. 4(d) illustrates a synthetically generated view of the sphere reflecting the estimated light source. In this case the estimation result had an error of $5mm \times 3mm$ for the $68mm \times 33mm$ light source.

5 Conclusion

This paper recovers a polygonal light source from the image of a specular sphere. Its main contributions are

1. an empirical analysis, which shows that line estimation from a single view of a single sphere is not possible in practice;
2. a closed form solution for recovering a polygonal light source from a single view of two spheres; and
3. an iterative approach for rectangular light source estimation based on two views of a single sphere.

Experiments on both synthetic and real images show promising results. In future research, we would like to study the possibility of extending SfS to handle the more complex lighting conditions of a polygonal light source.

References