Geometric Point Light Source Calibration

Jens Ackermann, Simon Fuhrmann, and Michael Goesele

TU Darmstadt, Germany

Abstract
We present a light position calibration technique based on a general arrangement of at least two reflective spheres in a single image. Contrary to other techniques we do not directly intersect rays for triangulation but instead solve for the optimal light position by evaluating the image-space error of the light highlights reflected from the spheres. This approach has been very successful in the field of Structure-from-Motion estimation. It has not been applied to light source calibration because determining the reflection point on the sphere to project the highlight back in the image is a challenging problem. We show a solution and define a novel, non-linear error function to recover the position of a point light source. We also introduce a light position estimation that is based on observing the light source directly in multiple images which does not use any reflections. Finally, we evaluate both proposed techniques and the classical ray intersection method in several scenarios with real data.

Categories and Subject Descriptors (according to ACM CCS): I.4.1 [Computer Graphics]: Digitization and Image Capture—Reflectance

1. Introduction

Estimation of light positions is an elementary building block for various computer vision and computer graphics topics. Augmented reality applications rely on known illumination to relight and seamlessly blend synthetic objects into a real scene. Similarly, known (point) light source positions are the basis of many shape or reflectance recovery techniques such as photometric stereo.

Although computer vision moved in recent years towards more uncontrolled setups, most reconstruction techniques are still designed to operate in a controlled indoor environment. This is especially true if quality is the prevailing focus such as in scientific applications, the movie and game industry. We focus on this setting where accuracy is of special importance and the scene can be controlled to a certain degree. In those scenarios it is usually also inappropriate to work with the approximation of a directional light source since the light cannot be placed sufficiently far away.

We propose a new method to recover the position, and thus also the direction, of a point light source. For our setup, we require a single image and two or more mirror spheres that are placed in the scene. In order to calibrate the light source, we assume that the radius and the position of the spheres are known, where the latter can easily be derived

from the images and the sphere radius alone. Our contributions are

- a novel calibration method that is based on minimizing the image-space error of the light highlights reflected from the spheres,
- a highly accurate calibration method that directly triangulates the light source in a multi-view setting, and
- a thorough comparison and evaluation of both methods and the tradiational ray intersection approach on real images with carefully acquired ground truth measurements.

Contrary to other works in this area we also study the impact of the spatial arrangement and the number of spheres on robustness and accuracy of the solution. Besides the new calibration techniques introduced here, we give the reader a good understanding of the available methods and their performance with respect to one another. We also discuss the weaknesses and strengths of the individual approaches regarding accuracy and implementation effort.

1.1. Related Work

There is a large body of literature on light source estimation. Some works exploit cast shadows [PWSP11], sample the complete incoming light-field [SSI99, KY04] or estimate the light source from stationary images [WMTG05]. Approaches that minimize an intensity error compare actual im-

ages of a scene with known geometry and reflectance against renderings with the current light estimate [HNI05, WC01, XW08]. However, obtaining the exact scene geometry is often difficult and error prone. A straightforward way to circumvent this problem is to use simple shapes such as cubes or spheres because their geometry is known.

Many techniques [DY13, WS02, WSL08, ZK02] assume an infinitely distant illumination and can only recover the direction of the light. Some of the ideas in these works can be readily used to estimate positions of near point light sources if applied to multiple spheres. Masselus et al. [MDA02] demonstrate that once light directions are known with respect to several scene points the corresponding rays can be intersected to yield a light position. In particular, they use four diffuse spheres and invert a linear shading model to obtain the directions. They do not perform a quantitative evaluation on real images. Powell et al. [PSG01] show that obtaining the respective light directions is especially easy for reflective spheres at known positions. They use two spheres in a special setup with a fixed baseline of 11cm and assume that reflection points in 3D correspond accurately to detected image highlights. The framework of Zhou et al. [ZK04] is based on images of specular spheres placed at different locations to triangulate an area light source. They do not evaluate the impact the number of images has on their results. Nayar [Nay89] uses mirroring spheres for 3D reconstruction and shows a strong relation to multi-view stereo. He evaluates his reconstruction framework in the context of light source triangulation.

All these methods assume that highlights on the spheres are detected accurately in the image and then shoot rays from the camera towards the spheres. The reflected rays are then intersected in 3D. The triangulation problem is well known in the context of image-based scene reconstruction [HS97] and it is preferable to minimize the reprojection error instead of computing the closest point to all rays in 3D space.

Aoto et al. [ATS*12] are the only ones to consider this error for triangulation of near light sources. Their setup consists of a hollow glass sphere with known position and radius. Due to the inherent difficulty of computing the 3D position of the light source reflection on the surface of the sphere, the authors exploit a characteristic of epipolar geometry to triangulate the light source using the two light highlights on the front and the back side of the sphere. This limits their approach to a small baseline defined by the diameter of the sphere and consequently yields unstable results for distant light sources. In contrast, our approach enables us to use an arbitrary baseline which cannot be achieved with a single glass sphere.

1.2. Overview

In Section 2 we will first present the formulation for the *forward calibration* which casts rays from the camera towards

the spheres and then finds the light source by intersecting the reflection rays in 3D. We then introduce our new light source calibration method based on minimizing the image-space error of the light source projection. We also investigate a straightforward but novel calibration technique where the light source is directly visible in at least two images of the scene, and the light position can be obtained through triangulation. In Section 3, we describe our lab setup including how we place the mirror spheres and how we obtain ground truth measurements. In Section 4 we evaluate the methods introduced in Section 3 with respect to the ground truth and with the help of a multi-view evaluation technique. We present our findings and wrap up in Section 5.

2. Light Source Calibration

In this section we will first present the most common way of obtaining the light position L by shooting rays towards the highlights on the sphere. The reflections of these rays are then intersected in 3D. This is what we call the *forward calibration* as rays are shot *forward* from the camera. We will then introduce our new *backward calibration* which evaluates the error in image space by tracing rays from the light source to the spheres and back to the camera. A third method which directly triangulates the light position with high accuracy is presented.

For all explanations we assume that the sphere position S and radius r are known in the camera's coordinate system. We will discuss ways to obtain the sphere position from the image and r alone in Section 3.1.

2.1. Forward Calibration

The most commonly represented method to perform light calibration is by finding the closest point in 3D to a series of rays. Masselus et al. [MDA02] obtain these rays for diffuse spheres by inverting a linear shading equation. For mirror spheres, the typical approach is to shoot rays u through the observed highlight pixels [PSG01, Nay89]. It is then straightforward to solve a quadratic polynomial to obtain the intersection R with a known sphere. Reflecting at the intersection normal $N = (R - S)/r$ gives the ray $v = u - 2(N^t u)N$ originating at R.

Once the rays v toward the light source are known, the light position L is given as the position that minimizes the squared distance to all rays. The orthogonal projection of $\hat{L} = L - R$ onto the ray v yields a decomposition $\hat{L} = \hat{L}_\parallel + \hat{L}_\perp$ with $\hat{L}_\parallel = (v^t \hat{L}) \cdot v$. The orthogonal distance $\|\hat{L}_\perp\|$ can then be expressed with matrices $A = (\mathrm{id} - v^t v)$ and $b = AR$ as

$$d = \|\hat{L} - \hat{L}_\parallel\| = \|(\mathrm{id} - vv^t)\hat{L}\| = \|A \cdot L - b\|. \quad (1)$$

We minimize the squared distance to all rays simultaneously:

$$\min \sum_{i=1}^{n} d_i^2 = \min \|(A_1^t, \ldots, A_n^t)^t \cdot L - (b_1^t, \ldots, b_n^t)^t\|^2 \quad (2)$$

2.2. Backward Calibration

In the case of the backward calibration our idea is to optimize the light source position by minimizing the projection errors of the reflections R. Let $R_i(L)$ be the reflection of the light L in sphere i, and H_i the detected highlight. The task is to minimize

$$\arg\min_L \sum_i \|\pi(KR_i(L)) - H_i\|^2 \qquad (3)$$

where K is the calibration matrix of the camera and π a projection operator. This case is more difficult because we do not know which ray to intersect with the sphere. To our knowledge, it has not been studied for light calibration with a general constellation of mirror spheres. Again, we assume the sphere position S and radius r to be known in the camera coordinate system. The challenge is to compute the highlight position R in 3D for a given light source position L. Figure 1 illustrates this situation.
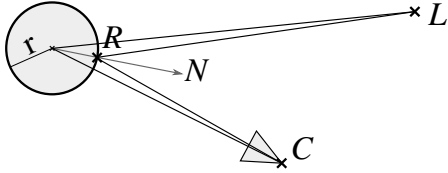

Figure 1: *Reflection geometry. The difficulty in backward calibration is to determine the point of reflection R that is generated from light source L and reflected into camera C.*

We first translate the camera coordinate system into the known sphere center S which yields a camera position $\widetilde{C} = -S$. For a light source position $\widetilde{L} = L - S$ there exists exactly one point $\widetilde{R}$ on the surface of the sphere that reflects towards the camera. Note that $\widetilde{R}$ does not in general bisect the angle between $\widetilde{L}$ and $\widetilde{C}$ but rather the angle between $\widetilde{L} - \widetilde{R}$ and $\widetilde{C} - \widetilde{R}$. Aoto et al. [ATS*12] do not compute the reflection point $\widetilde{R}$ for a general arrangement and only remark the difficulty of that problem. We show how this can be solved and review the steps taken by Eberly [Ebe] leading to a quartic equation.

If $\widetilde{L}$ and $\widetilde{C}$ are not parallel, we can use them as basis vectors and decompose the unknown point as $\widetilde{R} = x\widetilde{C} + y\widetilde{L}$. A first constraint is then given by the radius r as

$$r^2 = \widetilde{R}^t\widetilde{R} = x^2\widetilde{C}^t\widetilde{C} + 2xy\widetilde{C}^t\widetilde{L} + y^2\widetilde{L}^t\widetilde{L}. \qquad (4)$$

We obtain a second constraint by reflecting $\widetilde{C}$ across the line described by $\widetilde{R}$:

$$\widetilde{C}' = 2\frac{\widetilde{C}^t\widetilde{R}}{\widetilde{R}^t\widetilde{R}}\widetilde{R} - \widetilde{C} = 2\frac{x\widetilde{C}^t\widetilde{C} + y\widetilde{C}^t\widetilde{L}}{r^2}\widetilde{R} - \widetilde{C} =: 2\alpha\widetilde{R} - \widetilde{C}. \qquad (5)$$

The reflected point $\widetilde{C}'$ lies on the line from $\widetilde{R}$ to $\widetilde{L}$. Thus,

$\widetilde{C}' - \widetilde{R}$ is parallel to $\widetilde{L} - \widetilde{R}$:

$$0 = (\widetilde{L} - \widetilde{R}) \times (\widetilde{C}' - \widetilde{R}) = (\widetilde{L} - \widetilde{R}) \times \left((2\alpha - 1)\widetilde{R} - \widetilde{C}\right) \qquad (6)$$

$$= (2\alpha - 1)x\widetilde{L} \times \widetilde{C} - \widetilde{L} \times \widetilde{C} + y\widetilde{L} \times \widetilde{C} \qquad (7)$$

$$= (2\alpha x - x - 1 + y)\widetilde{L} \times \widetilde{C} \qquad (8)$$

Since $\widetilde{L}$ and $\widetilde{C}$ were assumed not to be parallel ($\widetilde{L} \times \widetilde{C} \neq 0$) it follows that

$$0 = 2\alpha x - x - 1 + y \qquad (9)$$

$$= 2r^{-2}(x\widetilde{C}^t\widetilde{C} + y\widetilde{C}^t\widetilde{L})x - x - 1 + y \qquad (10)$$

$$= 2r^{-2}\widetilde{C}^t\widetilde{C}x^2 + 2r^{-2}\widetilde{C}^t\widetilde{L}xy - x + y - 1 \qquad (11)$$

Equation (4) and Equation (11) are two polynomials in the coordinates of $\widetilde{R}$. Introducing $c := r^{-2}\widetilde{C}^t\widetilde{C}$, $b := r^{-2}\widetilde{C}^t\widetilde{L}$, $a := r^{-2}\widetilde{L}^t\widetilde{L}$, and separating y in Equation (11) yields

$$y = \frac{1 - 2cx^2 + x}{2bx + 1}. \qquad (12)$$

We insert this result into Equation (4) and reorder:

$$0 = 4c(ac - b^2)x^4 - 4(ac - b^2)x^3 \qquad (13)$$
$$+ (a + 2b - 4ac + c)x^2 + 2(a - b)x$$
$$+ a - 1$$

We know that this fourth order polynomial equation has at least one real solution because the reflection exists in all non-degenerate cases. We obtain it with a standard technique (see [BS08]) which instead computes the roots of

$$x^2 + \frac{\beta + A}{2}x + \left(z + \frac{\beta z - \delta}{A}\right) \qquad (14)$$

with $\beta = -1/c$, $\delta = \frac{a-b}{2c(ac-b^2)}$, $\gamma = \frac{a+2b-4ac+c}{4c(ac-b^2)}$, $A = \pm\sqrt{8z + \beta^2 - 4\gamma}$, $e = \frac{a-1}{4c(ac-b^2)}$, and z any real solution of the cubic equation

$$8z^3 - 4\gamma z^2 + (2\beta\delta - 8e)z + e(4\gamma - \beta^2) - \delta^2 = 0. \qquad (15)$$

We pick the positive solution x of Equation (14) which corresponds to $\widetilde{R}$ lying *between* $\widetilde{L}$ and $\widetilde{C}$. With y from Equation (12) the reflection point is given as $\widetilde{R} = x\widetilde{C} + y\widetilde{L}$.

Finally, we translate back into the camera coordinate system and obtain $R = \widetilde{R} + S$. The projection of this point into the image contributes to the overall error according to Equation (3). We then solve the resulting non-linear least squares problem using the Ceres [Cer] optimization library. In our tests, we did not observe the optimization getting stuck in local minima when restarting with different initial conditions.

2.3. Direct Light Position Triangulation

Another way of obtaining the light source position is to include the light directly in the images of the scene. This is often not applicable if the light source is far away from the scene. If feasible, however, this method yields impressive

results as we will show in our evaluation in Section 4. A related approach has been proposed by Frahm et al. [FKGK05] in the context of augmented reality with light source estimation. In contrast to their approach, we do not use light tracking but robust camera calibration with bundle adjustment.

In order to find the 3D positions of the spheres, the 2D coordinates of the sphere centers $p_{i,j}$ need to be known in every image I_i. There are several ways to obtain these coordinates. A manual approach is to fit an ellipse to the mirror spheres as we explain in Section 3.1. This yields the sphere center in 2D as well as the sphere position 3D. A second approach is to take a photo of the scene with a camera ring flash (*Canon MR-14EX TTL*) as proposed by Lensch et al. [LKG*03]. The flash will create a highlight on every sphere in the scene. Each highlight is centered around the ray from the camera through the sphere center. The highlights can be detected in the images and Structure-from-Motion techniques are applied to recover the position of the spheres as follows:

Given the sphere centers $p_{i,j}$ in the images I_i we can use the 5-point algorithm [Nis04] for relative pose estimation on the first two cameras. Note that this approach requires at least five spheres in the scene or other means of calibrating the cameras, such as markers. All remaining cameras can be added using the 3-point absolute pose algorithm [NS04]. Finally, if the light source is visible in at least two images, it can directly be triangulated. Standard bundle adjustment is applied to substantially improve the positions of the spheres and the light position.

3. Capture Setup and Preprocessing

In this section we describe our capture setup which includes the scene with the spheres, our metric floor mat which is the basis for our ground truth measurements, and the camera and camera calibration we use. Figure 2 shows our setup.

Mirror Spheres: In this setup we distributed the mirror spheres at arbitrary but known position on the floor mat. We use eight mirror spheres but only require a minimum of two spheres to calibrate the light source. Using more spheres naturally increases the robustness of the approaches. We evaluate in Section 4 to which extend fewer spheres degrade the accuracy of the results. Three of the eight spheres are placed at an elevated position on three stands that are 5cm, 10cm and 15cm above ground. This avoids degenerate (planar) 3D point constellations in the Structure-from-Motion scene reconstruction described in Section 4.3. The quality of the spheres is quite relevant. We experimented with spheres of varying grade and even slight geometric inaccuracies on the surface can lead to highlights that are offset by several pixels and markedly influence the stability of the results. We use quality bearing balls with a diameter of 6cm.

Metric Canvas: In order to obtain ground truth positions for both the spheres and the light source, a calibration target
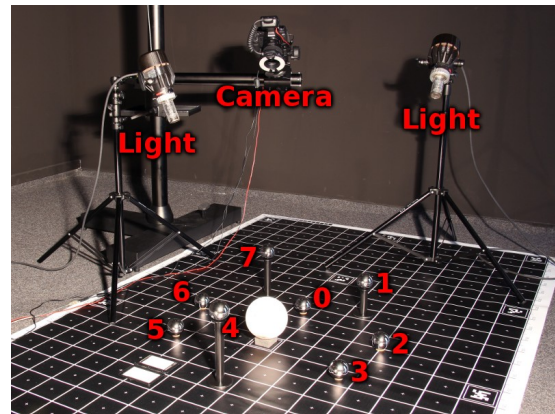

Figure 2: *The capture setup. The image shows the camera with ring flash attached, the light sources (we use only one at a time) and the spheres with corresponding numbers. The floor mat is the basis for our ground truth measurements.*

with metric information has been printed on a large canvas. We used this canvas as floor mat and carefully placed the spheres at known positions. The ground truth light positions have been measured using a plummet from the center of the light bulb to the floor mat. We expect that the accuracy of our measurements is in the order of millimeters for both the spheres and the light. This seems sufficient as the errors of the light estimation are orders of magnitudes larger.

Camera: We captured all photos using a *Canon 5D Mark II* camera with a *Canon EF 35mm F1.4L* prime lens. The intrinsic parameters of camera and lens have been calibrated prior to the evaluation using OpenCV [Ope]. The calibration determines the exact focal length (we kept the focus point fixed for all photos), the principal point, and the radial distortion parameters. It is performed on the detected corners of a checker board and reduces the reprojection error from several pixels down to subpixel accuracy.

Light: We used a *K5600 Joker-Bug 800* HMI lamp which produces a high light output by exciting a pressurized mercury vapor in the bulb. This lamp is particularly well suited for our task because it provides a good point light source.

3.1. Preprocessing

In a preprocessing step we first determine the distance d of each sphere from the camera center and the projection p of the sphere center onto the image plane. The sphere will project as an ellipse with parameters directly computable from the known camera intrinsics [HZ06] and the radius of the sphere. We manually adjust p and d until the rendered ellipse matches the image of the sphere. This procedure could be automated by first segmenting the sphere, fitting an ellipse, and then recovering p and d as proposed by Wong

Evaluation of Forward and Backward Calibration to Ground Truth [cm]				
Calibration and Dataset	Standard Deviation	RMS Distance	Min Error	Max Error
L_1 (fwd)	(2.8, 1.0, 2.9)	7.1	1.5	13.0
L_1 (bwd)	(1.2, 1.2, 2.0)	6.0	3.1	8.2
L_2 (fwd)	(1.1, 1.2, 1.5)	3.4	1.3	4.6
L_2 (bwd)	(1.1, 1.0, 1.3)	2.8	0.9	4.1

Table 1: *Evaluation results for forward and backward calibration on two data sets. Light positions were estimated in all camera frames. We show the standard deviation of the light position and the RMS distance to the ground truth position, as well as the minimum and maximum error.*

Figure 3: Left: *Positions of spheres (circles), light sources (squares) in the xy-plane and the camera (triangle). The grid lines are spaced 50cm apart. Note that at some positions $(A+C$ and $E+I)$ only the z-coordinate changes (the bigger square always corresponds to the light mentioned first) and that three spheres are placed at elevated positions.* Right: *Color coding of the z-component with height in cm.*

Figure 4: *Visualization of the directional error in the forward calibration (left) and the backward calibration (right). The horizontal center line corresponds to the ray from the origin passing through L_1. The color encodes the error from low (blue) to high error (red) and shows that the error function is less sensitive along the direction of the ray.*

et al. [WSL08]. However, manual parameter fitting seems appropriate for two reasons: Firstly, segmenting the mirror spheres is a hard problem due to low contrast between the spheres and the background. Secondly, manual parameter selection leads to higher accuracy in the order of at most a pixel for p and a few millimeters for d.

We also run an automatic highlight detection that reliably selects the point of the light reflection on each sphere with subpixel accuracy. A simple but reliable procedure is to first apply a non-maximum suppression on the intensity image with a large radius. Then, for each maximum, we use the average pixel position of all pixel in a small radius around the maximum with an intensity value of at least $t < 1$ times the maximum intensity. We use HDR images and $t = 0.5$.

4. Evaluation

In this section we first evaluate the techniques we introduced in Section 2, namely the *forward calibration* and the *backward calibration*. We do this for both varying camera positions and different light source locations. Afterwards, we analyze how the number of spheres influences the calibration. This aspect is typically disregarded in other works which assume a fixed number. Finally, we evaluate the direct light source triangulation.

4.1. Dependency on Reflection Geometry

For a fixed set of spheres, the reflection geometry (see Figure 1) depends only on the relative positions of the camera and light source. We investigate the robustness of the forward and backward calibration with respect to varying constellations of those. We first captured two data sets with 8 images from varying view ports each. The ground truth light positions L_1, L_2 for the two data sets are as follows:

$$L_1 = (102.6, 0.0, 114.5) \qquad L_2 = (55, -35, 74.5)$$

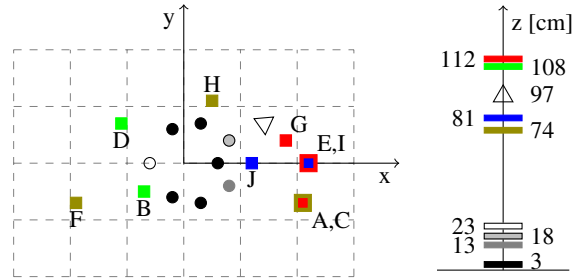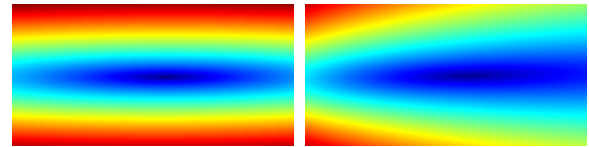After calibration the light position is given in the local coordinate system of the camera. To study the variance and to compare against the measured ground truth we have to transform the light position in a global coordinate system. To do this, we determine a rigid, least-squares optimal transformation [Ume91] from the estimated 3D sphere positions to the ground truth sphere positions derived from our metric canvas. Of course, this transformation will also include a small alignment error.

Table 1 shows the evaluation results for the forward and backward calibration. In particular, we computed the standard deviation of estimated light positions for all eight view ports. We also computed the root means square (RMS) distance to the ground truth light position. Both methods, the forward and the backward calibration perform similarly. We also notice that the error in dataset L_1 is larger than for L_2. The distance between light and spheres is about 1.5m for L_1 and 1.0m for L_2, so the larger error is plausible.

After varying the view port, we captured an additional dataset and moved the light source to 10 different positions while keeping the camera fixed. Figure 3 gives an overview over the sampled light positions and mirror sphere centers in the *xy*-plane together with a color-coding of the *z*-component.

Position	A	B	C	D	E	F	G	H	I	J
Distance from Origin [cm]	133.4	116.0	157.2	125.9	156.8	125.4	144.8	95.5	136.3	100.4
Forward Error [cm]	4.2	6.4	6.8	7.7	8.4	10.6	7.9	4.2	5.7	3.6
Backward Error [cm]	2.8	6.3	4.6	8.7	7.4	8.9	7.3	4.8	3.4	2.9

Table 2: *Distance between ground truth and our estimates for varying light source positions according to Figure 3.*

Number of Spheres	Standard Deviation [cm]	RMS Distance to Ground Truth [cm]
2 (fwd)	(5.0, 1.5, 5.8)	11.0
2 (bwd)	(4.9, 1.6, 5.9)	11.0
3 (fwd)	(2.2, 0.7, 2.6)	8.8
3 (bwd)	(2.0, 0.7, 2.5)	8.3
4 (fwd)	(1.3, 0.4, 1.5)	8.6
4 (bwd)	(1.5, 0.5, 1.9)	8.0
5 (fwd)	(0.9, 0.3, 1.1)	8.5
5 (bwd)	(1.2, 0.4, 1.5)	7.8
6 (fwd)	(0.7, 0.2, 0.7)	8.5
6 (bwd)	(0.9, 0.3, 1.2)	7.7
7 (fwd)	(0.4, 0.1, 0.5)	8.4
7 (bwd)	(0.6, 0.2, 0.8)	7.5

Table 3: *Different number of spheres $n = 2, \ldots, 7$ used for calibrating light position E. Statistics in each row are computed over all $\binom{8}{n}$ possible combinations of spheres and evaluated for both techniques.*

Figure 5: *The forward and backward error for any combination of two spheres with light position E.*

Table 2 lists the distances of our estimates to the ground truth position. It can be observed that the backward calibration yields slightly better results for most light positions, however, the difference between the approaches is small. The results also suggest a correlation between error and the distance of the light, which is to be expected in any triangulation-based system. Another observation is that the error orthogonal to the light direction is typically very small whereas the error in the direction of the light is large. In order to analyze this directional uncertainty, we sampled the error function for both the forward and the backward calibration on a 2D slice along the direction of the light. See Figure 4 for a visualization of the error function and Section 5 for further discussion of this behavior.

4.2. Dependency on the Number of Spheres

The impact of the number of spheres on the robustness is rarely considered in light source estimation. The goal is rather to place as few calibration objects in the scene as possible. While our technique requires only a minimum of two spher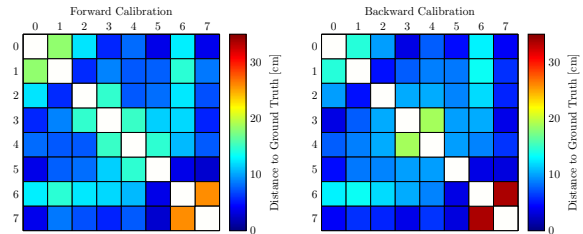es, the results so far have been computed with eight spheres. We believe that in many image-based reconstruction setups increased robustness is well worth the effort.

We run the proposed algorithm for all possible combinations of n out of 8 spheres. To reduce the number of possible combinations, we perform this evaluation only on light position E from Table 2. This position promises a challenging configuration because it has a relatively large error even for $n = 8$ and the light position has the largest distance to the spheres. The standard deviation for fixed n in the second column of Table 3 gives an indication of the stability with respect to different sphere configurations and baselines. As expected, we see a strong decrease with growing number of spheres for both techniques.

The RMS distance of all $\binom{8}{n}$ results to the measured ground truth is summarized in the third column of Table 3 for $n = 2, \ldots, 7$. We observe that the error for both calibration methods decreases with an increasing number of spheres, see also Figure 6 which illustrates this for all light positions.

We take a more detailed look at the distribution of errors for all combinations but restrict our analysis to the case $n = 2$ which resembles the setting used in the majority of related approaches [Nay89, PSG01, TMNM09]. Figure 5 plots the distance to the ground truth for all possible combinations of two spheres for the forward and backward calibration at light position E. For 88% of the pairs the error is lower than 15cm and only three combinations lead to larger deviations.

Figure 7 shows the positional error for all pairs that contain sphere number 0 for all light positions. Again, most of the errors are below 15cm. We also observe that none of the combinations outperforms the others for all light positions. Thus, we cannot detect a preferred arrangement for the spheres in the scene.
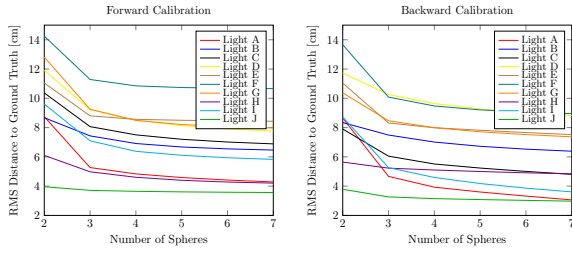
Figure 6: *The forward and backward error for different number of spheres for all light positions. Each RMS error is computed over all combinations of n spheres out of 8.*

Figure 7: *Error for all light positions and all pairs of spheres that contain sphere number 0.*

Evaluation of direct light triangulation [cm]		
Ground Truth	Estimate	Error
$(102.6, 0.0, 114.5)$	$(102.4, -0.2, 113.7)$	0.9
$(55.0, -35.0, 74.5)$	$(55.0, -35.1, 74.3)$	0.2

Table 4: *Evaluation of light source positions obtained through direct triangulation followed by bundle adjustment. The result is within millimeters from the measured ground truth position.*

4.3. Pose Estimation and Direct Triangulation

Many image-based reconstruction tasks require to observe the target object from multiple camera positions, such as multi-view photometric stereo [BAG12]. To estimate the camera pose either tracking markers have to be placed in the scene or features on the object need to be detected. If light estimation with mirror spheres is performed in such a context, the spheres can directly be used for pose estimation with the help of a ring flash, see Section 2.3, and additionally to obtain a highly accurate light position from at least two other images that contain the light source.

For each of the two datasets L_1 and L_2 we took additional images that contained both the spheres and the direct light source. Because the light source is extremely bright we used the *B+W Gray Filter 72mm 110 E 1000x* which reduces the incoming light intensity by about 10 F-stops. This yields an extremely well localized point light in the image. The light can be automatically detected with sub-pixel accuracy using the same technique described in Section 3.1.

The results for triangulating the light position are given in Table 4. As can be seen, for Dataset 1, the positions are highly accurate with errors of less than a centimeter although the distances between the cameras and the light source were about 4.5m. For the Dataset 2, the distances between light source and the cameras were about 2.5m and the positional error is in the order of the uncertainty of the ground truth measurements.
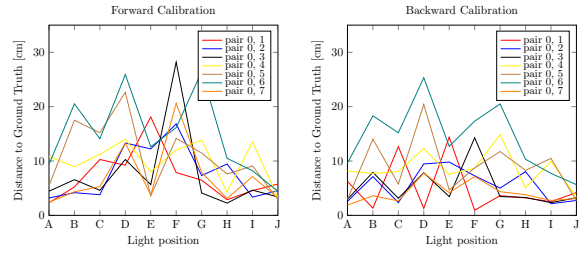
5. Discussion

In this paper we presented a novel light calibration approach based on minimizing reprojection errors of the light reflections on the surface of a reflective sphere. This approach has, to the best of our knowledge, neither been implemented nor evaluated due to the difficulty of determining the 3D position of the light reflection. We also presented a method that directly triangulates the visible light source in the images. Although this is a straightforward approach, we have not encountered this technique in the literature. Further, we presented a thorough comparison of all approaches with the goal to give an overview of the performance of several light calibration techniques. We will now give our assessment and interpretation of the results.

Undoubtedly the direct triangulation approach yields the best results in practice. This is mainly due to the fact that, depending on the baseline of the cameras, this method does not suffer from the directional uncertainty of all the other methods (compare Figure 4). However, due to the more complicated multi-view setup and the constraint of observing the spheres and light source at the same time, the use cases of this approach are limited. In particular multi-view photometric stereo and image-based acquisition approaches can benefit from this technique. They rely on accurate predictions of the irradiance on a target object which falls off quadratically with distance from the light source. A realistic example with the light source 1m away and an erroneous estimate of 1.1m already leads to a factor of $1.1^{-2} \approx 0.8$ in the predicted irradiance. This underlines the importance of accurate light positioning and puts the obtained results into perspective.

The results for the forward and backward calibration are accurate in the orders of centimeters but the two techniques yield very comparable results. The inaccuracies are due to the fact that these methods have large directional uncertainty mainly depending on the baseline and distance of the light source. It is on the one hand disappointing that our backward calibration approach does not result in considerably better localization of the light source. On the other hand this is good news for practical applications: It does not seem to be necessary to go through the hassle of implementing this
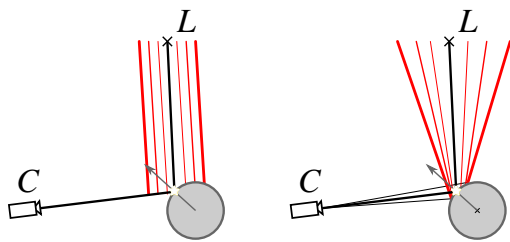
Figure 8: *Illustration of the error function for forward calibration (left) and backward calibration (right). The red lines represent ISO-surfaces of the error function.*

approach which is mathematically more involved and leads to a non-linear optimization problem.

Both forward and backward calibration have large directional uncertainty but they manifest in a different way. Where the forward calibration error expands cylindrically around each light direction, the backward calibration error function models the perspective aspect of the camera which allows for larger errors further away from the spheres. This can clearly be observed in Figure 4 and is illustrated in Figure 8. Although this is a useful property and is commonly referred to as the Gold Standard method [HZ06] in bundle adjustment literature, this aspect seems to be of limited relevance in light source calibration. We belief the controlled setup where spheres are usually close together and the light source has similar distance to all spheres allows the forward calibration to behave sufficiently well. It will be interesting to see if the gap in performance between forward and backward calibration becomes more evident with more general sphere constellations.

References

[ATS*12] AOTO T., TAKETOMI T., SATO T., MUKAIGAWA Y., YOKOYA N.: Position estimation of near point light sources using a clear hollow sphere. In *ICPR* (2012). 2, 3

[BAG12] BELJAN M., ACKERMANN J., GOESELE M.: Consensus multi-view photometric stereo. In *DAGM-OAGM* (2012). 7

[BS08] BRONŠTEJN I., SEMENDJAJEW K.: *Taschenbuch der Mathematik*. Harri Deutsch Verlag, 2008. 3

[Cer] Ceres: A Nonlinear Least Squares Minimizer. code.google.com/p/ceres-solver. accessed 2013-06-21. 3

[DY13] DOSSELMANN R., YANG X. D.: Improved method of finding the illuminant direction of a sphere. *SPIE JEI* (2013). 2

[Ebe] EBERLY D.: Computing a point of reflection on a sphere. www.geometrictools.com. accessed June 15th 2013. 3

[FKGK05] FRAHM J.-M., KOESER K., GREST D., KOCH R.: Markerless augmented reality with light source estimation for direct illumination. In *CVMP* (2005). 4

[HNI05] HARA K., NISHINO K., IKEUCHI K.: Light source position and reflectance estimation from a single view without the distant illumination assumption. *TPAMI* (2005). 2

[HS97] HARTLEY R. I., STURM P. F.: Triangulation. *Computer Vision and Image Understanding 68*, 2 (1997). 2

[HZ06] HARTLEY R., ZISSERMAN A.: *Multiple view geometry in computer vision*. Cambridge University Press, 2006. 4, 8

[KY04] KANBARA M., YOKOYA N.: Real-time estimation of light source environment for photorealistic augmented reality. In *ICPR* (2004). 1

[LKG*03] LENSCH H. P. A., KAUTZ J., GOESELE M., HEIDRICH W., SEIDEL H.-P.: Image-based reconstruction of spatial appearance and geometric detail. *ACM ToG* (2003). 4

[MDA02] MASSELUS V., DUTRÉ P., ANRYS F.: The free form light stage. In *EG Workshop on Rendering* (2002). 2

[Nay89] NAYAR S. K.: Sphereo: Determining depth using two specular spheres and a single camera. In *Robotics Conferences* (1989). 2, 6

[Nis04] NISTÉR D.: An efficient solution to the five-point relative pose problem. *TPAMI* (2004). 4

[NS04] NISTÉR D., STEWENIUS H.: A minimal solution to the generalised 3-point pose problem. In *J. of Math. Imaging and Vision* (2004). 4

[Ope] OpenCV. docs.opencv.org/doc/tutorials/calib3d/camera_calibration/camera_calibration.html. accessed 2013-06-21. 4

[PSG01] POWELL M. W., SARKAR S., GOLDGOF D. B.: A simple strategy for calibrating the geometry of light sources. *TPAMI* (2001). 2, 6

[PWSP11] PANAGOPOULOS A., WANG C., SAMARAS D., PARAGIOS N.: Illumination estimation and cast shadow detection through a higher-order graphical model. In *CVPR* (2011). 1

[SSI99] SATO I., SATO Y., IKEUCHI K.: Acquiring a radiance distribution to superimpose virtual objects onto a real scene. *Trans. on Visualization and Computer Graphics* (1999). 1

[TMNM09] TAKAI T., MAKI A., NIINUMA K., MATSUYAMA T.: Difference sphere: An approach to near light source estimation. *Computer Vision and Image Understanding* (2009). 6

[Ume91] UMEYAMA S.: Least-squares estimation of transformation parameters between two point patterns. *TPAMI* (1991). 5

[WC01] WEBER M., CIPOLLA R.: A practical method for estimation of point light-sources. In *BMVC* (2001). 2

[WMTG05] WINNEMÖLLER H., MOHAN A., TUMBLIN J., GOOCH B.: Light waving: Estimating light positions from photographs alone. *CGF* (2005). 1

[WS02] WANG Y., SAMARAS D.: Estimation of multiple illuminants from a single image of arbitrary known geometry. In *ECCV* (2002). 2

[WSL08] WONG K.-Y. K., SCHNIEDERS D., LI S.: Recovering light directions and camera poses from a single sphere. In *ECCV* (2008). 2, 5

[XW08] XU S., WALLACE A.: Recovering surface reflectance and multiple light locations and intensities from image data. *Pattern Recognition Letters* (2008). 2

[ZK02] ZHOU W., KAMBHAMETTU C.: Estimation of illuminant direction and intensity of multiple light sources. In *ECCV* (2002). 2

[ZK04] ZHOU W., KAMBHAMETTU C.: A unified framework for scene illuminant estimation. In *BMVC* (2004). 2