

## Semi-calibrated Near Field Photometric Stereo

Fotios Logothetis<sup>1</sup>, Roberto Mecca<sup>\*1,2</sup>, Roberto Cipolla<sup>1</sup>

{fl1302, rm822, rc10001}@cam.ac.uk

<sup>1</sup>Department of Engineering, University of Cambridge, United Kingdom

<sup>2</sup>Department of Mathematics, University of Bologna, Italy

### Abstract

3D reconstruction from shading information through Photometric Stereo is considered a very challenging problem in Computer Vision. Although this technique can potentially provide highly detailed shape recovery, its accuracy is critically dependent on a numerous set of factors among them the reliability of the light sources in emitting a constant amount of light.

In this work, we propose a novel variational approach to solve the so called semi-calibrated near field Photometric Stereo problem, where the positions but not the brightness of the light sources are known. Additionally, we take into account realistic modeling features such as perspective viewing geometry and heterogeneous scene composition, containing both diffuse and specular objects. Furthermore, we also relax the point light source assumption that usually constraints the near field formulation by explicitly calculating the light attenuation maps. Synthetic experiments are performed for quantitative evaluation for a wide range of cases whilst real experiments provide comparisons, qualitatively outperforming the state of the art.

### 1. Introduction

Extracting 3D clues from shading information in multiple images has been known in the Computer Vision community since Woodham [49] introduced the problem of Photometric Stereo (PS). The goal is to reconstruct the 3D shape of a static object knowing its shading under different illumination conditions. First attempts employed several assumptions in order to deal with a solvable problem; those included diffuse reflectance, orthographic viewing geometry as well as uniform directional lighting. Up to now, most of the literature regarding PS still relies on the same “fully orthographic” set of assumptions.

**Motivation** PS approaches can be divided into two main categories. The first one solves calibrated PS by requir-

<sup>\*</sup>Roberto Mecca is a Marie Curie Fellow of the Istituto Nazionale di Alta Matematica, Italy.

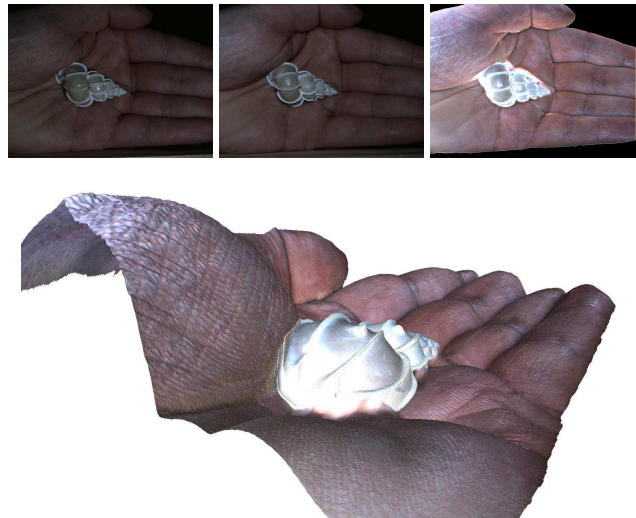


Figure 1. The first row shows two samples of near field images acquired under different lighting and different exposure (both unknown) and the computed albedo. Below it is the shape reconstructed using the proposed semi-calibrated PS approach.

ing an a priori knowledge of the light source’s geometry, position and brightness. The second category solves the uncalibrated PS problem where the information about the illumination of the scene is unknown. Recently, Cho *et al.* [9] introduced the idea of semi-calibrated PS where the brightness but not the positions of the light sources was considered unknown. Indeed, accurate knowledge of the brightness requires careful measurements with either specialized equipment (LUX meters) or reference objects with known geometry. Moreover the brightness commonly tends to vary from acquisition to acquisition, and this variation is mainly caused by three important reasons. First of all, the lighting status of a light source: most LEDs used as light sources have a non-negligible ramp-up time when turned on corresponding to constantly increasing brightness; this is followed by a gradual decrease of brightness because of heating-up that affects their electrical properties. This is especially relevant for high speed acquisitions performed with high power emitters. Under these circumstances, electrical

instabilities causing LED flickering are not uncommon [34]. The second reason is the deterioration of the light sources, unavoidably happening across time. This imposes a light calibration every time images have to be acquired over a relatively long lapse of time. Finally, it may be desirable to shoot images with auto-exposure settings in order to minimize the number of underexposed and saturated pixels. If this is the case, the effective brightness (*i.e.* brightness time exposure) of the respective light source cannot be known a priori.

**Contribution** In this work, we deal with the problem of semi-calibrated near field PS so that the geometry of the light sources is reasonably assumed to be point-wise and their positions are assumed known.

Additionally, we consider the shading model in favor of perspective viewing geometry and both diffuse and specular reflectance presented in [27].

Finally we propose a robust step-wise variational numerical solver where the lighting factors, shape and specular coefficient are approximated sequentially. This allows to relax the assumption that point light sources spread light uniformly in all directions by computing the spatially varying maps of the light intensities.

The reliability of our method for providing highly accurate 3D reconstructions and approximating lighting parameters is demonstrated quantitatively over a large set of synthetic and real data. Regarding real world experiments, our method is compared against other standard methods for estimating the brightness of light sources (*i.e.* using a light light-meter or reference objects).

## 2. Related Works

Starting from the seminal work of Woodham [49] where the shape from PS problem relied on fully orthographic assumptions, several approaches have been presented to make this technique more reliable. To deal with actual physical effects, perspective viewing geometry has been introduced by Horn [19] and then adopted by Tankus and Kiryati [45] in the PS context. Papadimitri and Favaro [35] proposed a new approach to the perspective viewing deriving a new but substantially equivalent parameterization of the pinhole camera modeling.

**Calibrated near field** In order to get a reliable reconstruction when the field of view is wide, apart from having an accurate pinhole camera parameterization, it becomes fundamental to model properly how light spreads from the source. This brings into play the concept of non-uniform light propagation and, most importantly, non negligible light dissipation.

The way the light spreads clearly depends on the shape of the source. Several methods consider flat light sources [11]

mostly aiming at specific applications. However, the majority of the literature in this field considers as the most relevant parameterization the point light source initially proposed by Iwahori *et al.* [20] and Clark [10]. More recently Xie *et al.* [52] suggested incremental steps from fully orthographic assumptions to perspective near field context to provide accurate reconstruction according to specific mesh deformations.

For the near field scenario, when light sources are close to the object, shadows appear frequently. There are several works dealing with shadows, aiming to extract geometrical information [44] or to avoid biased shading information [18] especially when only a few light sources are considered. For example, Chandraker *et al.* [8] proposed a graph-cut method to estimate the lighting visible pixels. Barsky and Petrou [3] took into account both highlights and shadows using only four light sources. However, different approaches for 3D shape recovery such as multi-view stereo techniques including [47, 53] have employed direct visibility computation steps. This has the potential of higher reliability than the heuristics presented above.

**Image ratios** Images acquired for PS reconstruction have constant overlapping information due to the monocular procedure; this overlapping information allows for simplification of the problem by considering ratios of images. After being introduced by Davis and Soderblom [13], this approach has been often used [23, 24, 51, 26, 18, 48, 42, 30, 28]. Mecca *et al.* [29] proposed a novel approach embedding the nonlinear parameterization of a point light source into a differential problem of hyperbolic PDEs. Lately, this approach has been further extended for colored surfaces [39] including diffuse and specular reflectance [27] and for acquisition in presence of ambient light [25]. Finally, Chandraker *et al.* [7] showed that the image ratio formulation can provide level set information for any surface with isotropic BRDF.

**Photometric calibration** Light calibration, intended as an estimation of light positions/directions and intensities, often requires specific equipment [14, 15] or a dedicated procedure [50] that can differ depending on the scenario. For example, PS for outdoor scenes generally uses intrinsic light calibration according to geo-position information [43, 2, 21, 1] or triangulating with reflective spheres [31]. Alternatively, in a dark environment, light calibration can be performed considering reference objects where the shape and reflectance are known. Goldman *et al.* [17] used spheres of the same materials of the object under observation, dealing with spatially varying BRDFs. Instead, the variational formulation proposed by Quéau *et al.* [38] needed additional information on the depth of the boundary (Dirichlet boundary condition). Shi *et al.* [40] disambiguate the generalized bas-relief (GBR) ambiguity [4] and

provided radiometric calibration analyzing color/intensity profiles in the RGB and irradiance-time domains. As most of the previous works, Papadimitri and Favaro [36] still assumed diffuse reflectance only for the near field uncalibrated PS. Very few attempts have been made to solve the uncalibrated PS problem for heterogeneous surfaces. For example, Georgiades [16] proposed an approach capable of computing the depth up to a concave/convex ambiguity. Shi *et al.* [41] parametrised general BRDF with bi-polynomial retaining the low frequency part to the detriment of the high component.

Finally, Cho *et al.* [9] tackled directly the semi-calibrated problem proving that it has a closed-form solution for the fully orthographic scenario that assumes purely diffuse reflection. To make the computation faster, they proposed an alternating minimization scheme that extends the classic PS problem with the joint estimation of constant light intensities. With the aim to relax the semi-calibrated modeling for both diffuse and specular surfaces for the near field scenario, we propose a new approach able to deal with spatially varying light intensities enhanced by the computation of shadow maps.

### 3. Problem formulation

In order to deal with near field PS, multiple physical effects have to be modeled simultaneously. Most of them require non-linear parameterization and this yields irradiance equations which are rather complicated to manage. This is especially evident when taking into consideration the system of equations that contains the information from multiple images.

#### 3.1. Irradiance equation

We assume to have  $N_{img}$  images (*i.e.* light sources) having  $N_{px}$  number of pixels. We exploit the monocular aspect of the PS problem by considering image ratios. To do that, we use the following single lobe irradiance equation for the  $k^{th}$  light source proposed in [27] unifying the Lambertian shading model [22] with the Blinn-Phong one [37, 5]:

$$i_k(\mathbf{x}) = \rho(\mathbf{x})\phi_k a_k(\mathbf{x}, z) (\bar{\mathbf{N}}(\mathbf{x}) \cdot \bar{\mathbf{W}}_k(\mathbf{x}, z))^{\frac{1}{c(\mathbf{x})}} \quad (1)$$

where lowercase letters indicate scalars, bold letters and uppercase bold letters are vectors of  $\mathbb{R}^2$  and  $\mathbb{R}^3$  respectively. The bar over a vector means that it is normalized (*i.e.*  $\bar{\mathbf{N}} = \frac{\mathbf{N}}{|\mathbf{N}|}$ ).

The vector  $(\mathbf{x}, z)$  is the prospective point of the surface  $\Sigma \subset \mathbb{R}^3$  (so that  $\mathbf{x}$  is the image point and  $z$  is the depth),  $\phi_k$  is the brightness of the light source,  $\rho(\mathbf{x})$  is the albedo and  $0 < c(\mathbf{x}) \leq 1$  is the shininess coefficient. The weighted lighting vector  $\mathbf{W}_k(\mathbf{x}, z)$  averages the light direction  $\mathbf{L}_k(\mathbf{x}, z)$  with the viewer direction  $\mathbf{V} = (\mathbf{x}, f)$  as

follows

$$\mathbf{W}_k(\mathbf{x}, z) = \bar{\mathbf{L}}_k(\mathbf{x}, z) + \min \left\{ 1, \frac{|1 - c(\mathbf{x})|}{\varepsilon} \right\} \bar{\mathbf{V}}(\mathbf{x}) \quad (2)$$

with  $f$  as focal length. The transition phase from the diffuse ( $c(\mathbf{x}) = 1$ ) to the specular reflection is described by  $\varepsilon$ . The direction of light is parametrized as the difference between the surface point  $\chi \in \Sigma$  and the position of the point light source  $\mathbf{P}_k$  both depending on the image point  $\mathbf{x}$ , that is

$$\mathbf{L}_k(\mathbf{x}, z) = \chi(\mathbf{x}) - \mathbf{P}_k(\mathbf{x}). \quad (3)$$

The perspective normal vector  $\mathbf{N}(\mathbf{x})$  is parametrised according to the notation provided in [35], that is

$$\mathbf{N}(\mathbf{x}) = \frac{1}{f}(f\nabla z(\mathbf{x}), -f - z(\mathbf{x}) - \mathbf{x} \cdot \nabla z(\mathbf{x})). \quad (4)$$

Finally, instead of parametrizing analytically the radial and distance dissipation of a point light source as in [27], we relax the concept of point light source by considering  $a_k(\mathbf{x}, z)$  and  $\phi_k$  as an unknown attenuation map and light source brightness respectively. In the next sections, we discuss how the irradiance equation can be manipulated in order to calculate the unknowns:  $z$ ,  $\rho$ ,  $c$  and  $\phi_k$ ,  $a_k \forall k$ .

#### 3.2. Modeling with image ratios

Differently from the two-lobes based irradiance equations [33, 46, 12, 32], the single lobe formulation (1) allows for a significant simplification of the mathematical modeling through the use of image ratios. Indeed, dividing equations for images  $h$  and  $k$ , and raising both side of the equation to the power of  $c(\mathbf{x})$  gives:

$$\left( \frac{i_h(\mathbf{x}, z)\phi_k a_k(\mathbf{x}, z)}{i_k(\mathbf{x}, z)\phi_h a_h(\mathbf{x}, z)} \right)^{c(\mathbf{x})} = \frac{\mathbf{N}(\mathbf{x}) \cdot \bar{\mathbf{W}}_h(\mathbf{x}, z)}{\mathbf{N}(\mathbf{x}) \cdot \bar{\mathbf{W}}_k(\mathbf{x}, z)}. \quad (5)$$

By substituting the parametrisation of the normal from equation (4), we get the following albedo independent PDE:

$$\mathbf{b}_{hk}(\mathbf{x}, z) \cdot \nabla z(\mathbf{x}) = s_{hk}(\mathbf{x}, z). \quad (6)$$

By denoting the vector components with superscript indexes and removing the dependencies for readability, the vector function  $\mathbf{b}_{hk}$  and the scalar function  $s_{hk}$  can be written as follows (see [27] for the details regarding the derivation)

$$\mathbf{b}_{hk} = \left[ (\phi_k a_k i_h)^c (f\bar{w}_k^1 - x^1\bar{w}_k^3) - (\phi_h a_h i_k)^c (f\bar{w}_h^1 - x^1\bar{w}_h^3), \right. \\ \left. (\phi_k a_k i_h)^c (f\bar{w}_k^2 - x^2\bar{w}_k^3) - (\phi_h a_h i_k)^c (f\bar{w}_h^2 - x^2\bar{w}_h^3) \right]$$

and

$$s_{hk} = (f + z) \left( (\phi_k a_k i_h)^c \bar{w}_k^3 - (\phi_h a_h i_k)^c \bar{w}_h^3 \right). \quad (7)$$

In addition, the ratio equation (5) can be used to update the shininess parameter  $c$ . Re-arranging it, we get:

$$c = \frac{\log(\mathbf{N} \cdot \bar{\mathbf{W}}_h) - \log(\mathbf{N} \cdot \bar{\mathbf{W}}_k)}{\log(i_h) + \log(\phi_k a_k) - \log(i_k) - \log(\phi_h a_h)}. \quad (8)$$

### 3.3. Albedo and Light source brightness

We formulate the problem for computing the albedo, the light source brightness and the light attenuation maps such that they can be directly computed depending on the geometry of the scene  $z$  and the reflectance coefficient  $c(\mathbf{x})$  of the irradiance equation (1).

First of all, normals can be computed by numerically differentiating the depth. The weighted lighting vector  $\mathbf{W}_k(\mathbf{x}, z)$  is also trivially computed using the current estimate of the depth and the light source positions.

Rearranging the irradiance equation (1), we get the following relation for the albedo:

$$\rho(\mathbf{x}) = \frac{i_k(\mathbf{x})}{\phi_k a_k(\mathbf{x}, z) D_k(\mathbf{x}, z)} \quad (9)$$

where  $D_k(\mathbf{x}, z) = (\bar{\mathbf{N}}(\mathbf{x}) \cdot \bar{\mathbf{W}}_k(\mathbf{x}, z))^{\frac{1}{c(\mathbf{x})}}$ .

We note that for each pixel  $\mathbf{x}$ , there are  $N_{img}$  equations (9). All of them can be stacked into a matrix giving the following (over-constrained) linear system:

$$\begin{bmatrix} \phi_1 a_1(\mathbf{x}, z) D_1(\mathbf{x}, z) \\ \vdots \\ \phi_{N_{img}} a_{N_{img}}(\mathbf{x}, z) D_{N_{img}}(\mathbf{x}, z) \end{bmatrix} \rho(\mathbf{x}) = \begin{bmatrix} i_1(\mathbf{x}) \\ \vdots \\ i_{N_{img}}(\mathbf{x}) \end{bmatrix}. \quad (10)$$

For the light source brightness, the equation (9) can be equivalently re-arranged as:

$$\phi_k = \rho(\mathbf{x}) \frac{i_k(\mathbf{x})}{a_k(\mathbf{x}, z) D_k(\mathbf{x}, z)}. \quad (11)$$

We note that for each light source  $k$  there are  $N_{px}$  equations (11), one per each pixel ordered as  $(\mathbf{x}_1, \dots, \mathbf{x}_{N_{px}})$ . Stacking these into a linear system, we get:

$$\begin{bmatrix} \rho(\mathbf{x}_1) a_k(\mathbf{x}_1, z) D_k(\mathbf{x}_1, z) \\ \vdots \\ \rho(\mathbf{x}_{N_{px}}) a_k(\mathbf{x}_{N_{px}}, z) D_k(\mathbf{x}_{N_{px}}, z) \end{bmatrix} \phi_k = \begin{bmatrix} i_k(\mathbf{x}_1) \\ \vdots \\ i_k(\mathbf{x}_{N_{px}}) \end{bmatrix}. \quad (12)$$

### 3.4. Attenuation map

Most near field PS methods [20, 29, 10, 52] assume that light attenuates radially due to the point shape of the light source. However, in practice this assumption might not be very realistic, providing deformation for the shape recovered. Ideally, we would like to jointly calculate the light attenuation at each pixel for each image considering  $a_k(\mathbf{x}, z)$  as a pixel-wise unknown. This renders the PS problem helplessly under-constrained with many more unknowns than equations. To overcome this limitation, we employ the assumption that the light attenuation  $a_k(\mathbf{x}, z)$  is locally constant. Note that this is a much weaker assumption than assuming that the lighting is locally directional. Indeed,

$\mathbf{L}_k(\mathbf{x}, z)$  is expected to have a non-negligible variation even between nearby pixels.

Thus, we seek to solve for  $a_k(\mathbf{x}, z)$  by assuming that it has a constant value over a small patch of  $N_{patch}$  pixels surrounding each pixel  $\mathbf{x}$ . For every patch, each pixel in each images provides one equations, thus the total number of equations is  $N_{img} N_{patch}$ . The unknowns to be calculated are:  $N_{patch}$  albedos  $\rho(\mathbf{x})$ ,  $N_{patch}$  depths  $z(\mathbf{x})$ ,  $N_{patch}$  shininess coefficients  $c(\mathbf{x})$  and  $N_{img}$  attenuation  $a_k(\mathbf{x}, z)$ .

Thus, for the number of equations to exceed the number of unknowns, we require that:

$$N_{img} N_{patch} > 3N_{patch} + N_{img} \iff N_{patch} > \frac{N_{img}}{N_{img} - 3}. \quad (13)$$

The inequality (13) is easily satisfied for  $N_{patch} = 9$  (i.e. a 3x3 pixel patch) and  $5 \lesssim N_{img} \lesssim 50$  (typical number of images used in PS problems). Under this assumption establishing that the search for  $a_k(\mathbf{x}, z)$  is not under-constrained, we re-arrange equation (1) to get:

$$a_k(\mathbf{x}, z) = \frac{i_k(\mathbf{x})}{\rho(\mathbf{x}) \phi_k D_k(\mathbf{x}, z)}. \quad (14)$$

Reusing the strategy followed in section 3.3, we note that for each light source  $k$  and each patch, there are  $N_{patch}$  equations (14), one per each pixel ordered as  $(\mathbf{x}_1, \dots, \mathbf{x}_{N_{patch}})$ . Stacking these into a linear system, we get:

$$\begin{bmatrix} \rho(\mathbf{x}_1) \phi_k D_k(\mathbf{x}_1, z) \\ \vdots \\ \rho(\mathbf{x}_{N_{patch}}) \phi_k D_k(\mathbf{x}_{N_{patch}}, z) \end{bmatrix} a_k(\mathbf{x}, z) = \begin{bmatrix} i_k(\mathbf{x}_1) \\ \vdots \\ i_k(\mathbf{x}_{N_{patch}}) \end{bmatrix}. \quad (15)$$

### 3.5. Scale ambiguity

There is one issue that arises from the above-mentioned discussion, which is a global two parameter ambiguity between the scale of  $\phi_k$ , the scale of  $a_k(\mathbf{x}, z)$  and the scale of  $\rho(\mathbf{x})$ . Indeed, one can replace  $\phi_k$  by  $m_1 \phi_k$ ,  $a_k(\mathbf{x}, z)$  by  $m_2 a_k(\mathbf{x}, z)$  and  $\rho(\mathbf{x})$  by  $\frac{\rho(\mathbf{x})}{m_1 m_2}$ , for any scalars  $m_1, m_2$ , and the irradiance equation (1) is unchanged.

To overcome these ambiguities, we set  $\|\Phi\| = 1$ , with  $\Phi = [\phi_1, \dots, \phi_{N_{img}}]^T$ . We also set  $\text{mean}[\rho(\mathbf{x})] = 1$ . Although this may seem counter intuitive with respect to the usual definition of  $\rho(\mathbf{x})$  to be always between zero and one, the irradiance equation is consistent with  $\rho(\mathbf{x})$  of any scale.

## 4. Computational Approach

In section 3 we presented how the irradiance equation (1) can be manipulated so as to calculate the various unknowns. Here we present a robust alternating optimization strategy to jointly estimate depth, light source brightness, (scaled) albedo, light attenuation maps and reflectance coefficients.



### 4.1. Depth Computation

We note that there are  $\binom{N_{img}}{2}$  versions of equation (6), one for each pair of images. By following a similar method with [27], we add a zero-order Tikhonov regularizer that constraints the mean  $z$  to be close to  $z_0$ . This is essential for ensuring that the differential problem based on (6) has a unique solution. Hence, we seek to minimise:

$$\min_z \sum_{h,k} \|\mathbf{b}_{h,k} \cdot \nabla z - s_{h,k}\|_{L^1} + \lambda \|z - z_0\|_{L^2} \quad (16)$$

where the  $L^1$  norm is chosen for robustness to noise and outliers.

Equation (16) is minimized by using the ADMM scheme proposed in [27]. Furthermore, since  $\mathbf{b}$  and  $s$  implicitly depend on  $z$ , in the iterative process they are calculated by using the current estimates of the depth values.

In section 4.4 we describe the steps of the alternating optimization.

### 4.2. Shadow Map estimation

Most PS methods estimate normals locally and so it is common to use heuristics to detect cast shadows. Since we follow an iterative, global geometry refinement algorithm, it is reasonable to calculate cast shadows directly through ray-tracing using the previous estimate of the geometry. Although this procedure is very expensive in terms of computational time, it ensures that the surface is globally consistent with the light sources making the computation of the lighting factors more reliable.

### 4.3. Pixel based selection strategy

As described in sections 3.3 and 3.4, the computation of  $\rho(\mathbf{x})$  and all  $\phi_k$  and  $a_k(\mathbf{x}, z)$  requires the solving the linear equations (10), (12) and (15) respectively. With the aim to gain robustness to outliers and noise, we solve these system considering  $L_1$  relaxation [6]. In addition, a dynamic selection of reliable pixels is done by ignoring those pixels where the normal is almost perpendicular to weighted lighting vector and thus very little light is expected to be reflected of the surface. Hence, pixels where  $\bar{\mathbf{N}}(\mathbf{x}) \cdot \bar{\mathbf{W}}_k(\mathbf{x}, z) < 0.1$  are considered unreliable and the respective rows at equations (10), (12) and (15) are removed. In addition, pixels found to be in shadow (see section 4.2) or pixels with very low image values ( $i_k(\mathbf{x}) < 0.05$ ) or saturated ( $i_k(\mathbf{x}) > 0.99$ ) are also ignored<sup>1</sup>.

### 4.4. Alternating optimization

Up to this point, we have discussed how different sets of variables can be used to easily calculate the rest. Of course, in the beginning of the problem very little information about

<sup>1</sup> We assume normalized image values with 0 corresponding to black, 1 corresponding to white.

---

#### Algorithm 1: Semi-Calibrated PS

---

**Input** : Images, light source positions, rough mean distance  $z_0$

**Output**: Depth map  $z$ , light source brightnesses  $\Phi$ , albedo  $\rho$ , attenuation maps and shininess parameter  $c$

**Initialization**:

$z^0 = z_{mean}$ ,  $c^0(\mathbf{x}) = 1$ ,  $\rho^0(\mathbf{x}) = 1$ ,  $\phi_k = \frac{1}{\sqrt{N_{img}}} \forall k$ ;

calculate fields  $\mathbf{L}_k^0(\mathbf{x}, z^0) \quad \forall k$ ;

set  $a_k^0(\mathbf{x}, z) = \frac{1}{\|\mathbf{L}_k^0(\mathbf{x}, z^0)\|^2} \quad \forall k$ ;

**while**  $|z^{t+1} - z^t| > 10^{-4} \times |z^t|$  **do**

**if**  $t > 1$  compute shadow maps. **endif**;

    calculate fields  $\mathbf{W}_k^t(\mathbf{x}, z^t) \quad \forall k$ ;

    compute  $\Phi^{t+1}$  solving (12);

    for every pair  $i_h$  and  $i_k$ , calculate fields

$\mathbf{b}_{hk}(z^t, a_h^t, a_k^t, \phi_h^{t+1}, \phi_k^{t+1}, \mathbf{W}_h^t, \mathbf{W}_k^t)$ ,

$s_{hk}(z^t, a_h^t, a_k^t, \phi_h^{t+1}, \phi_k^{t+1}, \mathbf{W}_h^t, \mathbf{W}_k^t)$ ;

    and then approximate  $z^{t+1}$  solving (16);

$\rho^{t+1}(\mathbf{x})$  solving (10);

$a_k^{t+1}(\mathbf{x}, z^{t+1})$  solving (15)  $\forall k$ ;

$c^{t+1}(z^{t+1}, a_h^{t+1}, a_k^{t+1}, \mathbf{W}_h^t, \mathbf{W}_k^t)$  solving (8);

**end**

---

the geometry and the photometric properties of the scene is known and all the unknowns have to be jointly estimated. In addition, as each step depends on the results of the previous steps, the order of these steps is very crucial so as to converge to an acceptable solution at a reasonable rate.

First of all, we perform the  $\Phi$  update step according to (12). A single scalar per image is calculated using the whole image data, thus this step is very robust to outliers and errors. The dependence of the system (12) on  $z$  is overcome by initializing  $z^0 = z_{mean}$ , the mean distance between the camera and the object. The albedo is initialized to  $\rho^0(\mathbf{x}) \equiv 1$ , the attenuation using the point source model  $a_k^0(\mathbf{x}) = \frac{1}{\|\mathbf{L}_k^0(\mathbf{x}, z^0)\|^2}$  and  $\phi_k = \frac{1}{\sqrt{N_{img}}}$ ,  $\forall k$ . Finally, we initialize  $c(\mathbf{x}) = 1$  which corresponds to diffuse reflectance.

The next step updates the  $z$  values using (16). The variational solver is very robust to noise and other forms of sparse corruptions but cannot deal with the systematic error coming from using wrong  $\Phi$ . Indeed, this is exactly the core purpose of the paper; inaccurate estimates for  $\Phi$  can lead to substantial deformation of the recovered shape. This is demonstrated experimentally in section 5 with the comparison with [27]. Furthermore, the main equation for this step (6) is albedo independent hence it is reasonable to precede this computation to the albedo's computation.

Next step is the attenuation map calculation. This is a local operation and much more sensitive to inaccuracies. For each pixel, we choose  $3 \times 3 = 9 = N_{patch}$  sized patches contain-

ing all of its neighbors to estimate the attenuation. As the patches are overlapping, continuity is implicitly enforced.

Finally we perform the  $c(\mathbf{x})$  update at the end. This is the most numerically unstable step as it involves a ratio of differences of logarithms (8) thus it is essential to have as good estimates of the rest of the parameters as possible.

The whole alternating optimisation procedure is summarized in Algorithm 1.

## 5. Experiments

We evaluated our algorithm with synthetic and real data covering a range of different situations. Firstly, in order to make our approach comparable with the state of the art on semi-calibrated PS [9] (by using our own implementation of their alternating minimisation method), we adapted it to the easier, fully orthographic and diffuse scenario. This was easily achieved by fixing  $c = 1$ ,  $a_k = 1$  and  $f = \infty$  ( $10^6$  in practice). Secondly, in order to test the reliability of our approach, we considered synthetic data generated with the Cook and Torrance reflection model [12]. Furthermore, we compared it to [27] (using code and a data set available online<sup>2</sup>) which is the most similar in terms of irradiance model. Although [27] does not focus on solving the semi-calibrated PS, the comparison with this work provides clues regarding shape deformation occurring when reasonable but incorrect brightness is used.

Our algorithm was implemented in MATLAB and run on a server machine with an AMD Opteron CPU. The computation time for the 2 MPixel x 9 images datasets was about one hour with peak memory usage of around 30GB. Around 90% of the computation maps was spent on shadow and attenuation maps.

### 5.1. Synthetic Data

We generated three synthetic data sets with eight images using the “Armadillo” model from The Stanford 3D Scanning Repository<sup>3</sup> where the “Lena” image was used for the albedo, see Figure 2. The first dataset was made under the classic PS assumptions having directional lighting, orthographic projection and diffuse reflection, in order to provide a fair comparison to [9]. The second dataset was a near-field, diffuse reflection scenario where the object had a size of 2cm and is placed 4cm away from a virtual pin-hole having focal length of 12mm. The light sources were symmetrically distributed around the camera at two circles of radius 3 and 5 cm respectively. To apply [9] to the near-field scenario, we used as lighting directions the  $\mathbf{L}_k$  at the center of the object. As a result, the errors grew towards the boundaries of the image (Figure 3). Finally, our third

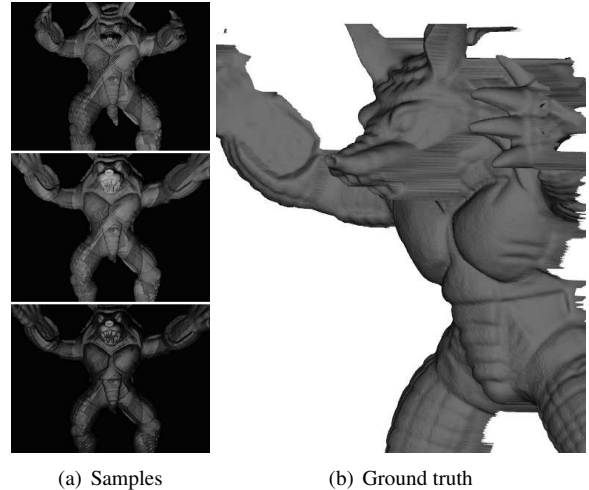


Figure 2. Synthetic data. Left column is sample images for the 3 cases examined having “Lena” as albedo. From top to bottom: far field diffuse, near field diffuse, near field Cook and Torrance.

dataset had the same arrangement as the near field scenario but the object was rendered by using the Cook and Torrance reflection model. This created specular highlights inconsistent with our irradiance equation (1) thus the robustness of our method was tested with more physical derived synthetic data. To stress all compared algorithms to their limits, we chose the brightness of the light sources  $\Phi$  with significant variation. In fact, the brightest light source was 5 times brighter than the dimmest (see Figure 4).

The quantitative evaluation of the algorithms was achieved through the mean error between the obtained normals vs the ground truth showed by the bottom line of Figure 3. The evaluation of the predicted light source intensities is shown in Figure 4. It was achieved by finding the angle between  $\Phi$  and the ground truth  $\Phi_{gt}$ . As  $\Phi$  is always a unit vector (section 3.5), we get  $E_\Phi = \arccos(\Phi \cdot \Phi_{gt})$ .

The proposed approach significantly outperforms the competitors in both near field cases and gets a slightly higher error than [9] in the classic PS scenario. That is because our  $z$  estimation step propagates errors around the occlusion boundaries at the face of the Armadillo in contrast to [9] which recovers normals directly.

### 5.2. Real Data

We evaluated our algorithm on several real datasets shown in Figure 1 and Figure 5. To make the tests as challenging as possible, we combined multiple objects with various reflections including a plastic baseball player figurine, a marble statue, a 3D printed plastic version of the Armadillo and a shell.

Our capture setup consists of a FL3-U3-20E4C camera of Point Grey mounting a 12mm lens. Our light sources are 9 OSRAM Platinum DRAGON (white) LEDs placed

<sup>2</sup>[researchgate.net/publication/310310648\\_A\\_Single\\_Lobe\\_Photometric\\_Stereo\\_Approach\\_for\\_Heterogeneous\\_Material](https://researchgate.net/publication/310310648_A_Single_Lobe_Photometric_Stereo_Approach_for_Heterogeneous_Material)

<sup>3</sup><http://graphics.stanford.edu/data/3Dscanrep>

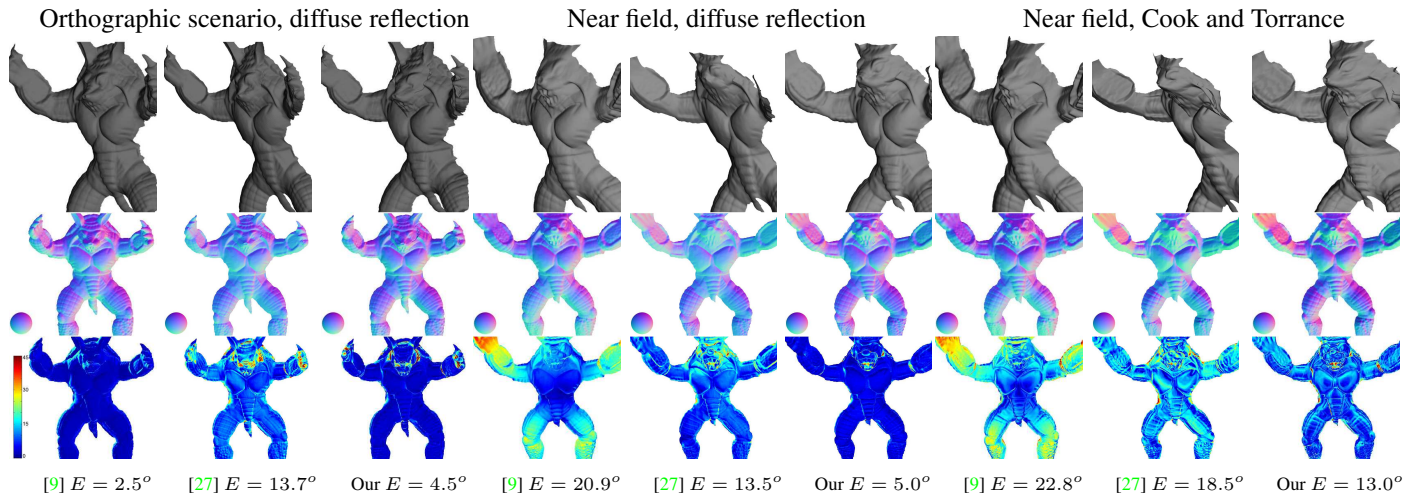


Figure 3. From top line to bottom: reconstructions, normals and angular error for normals.

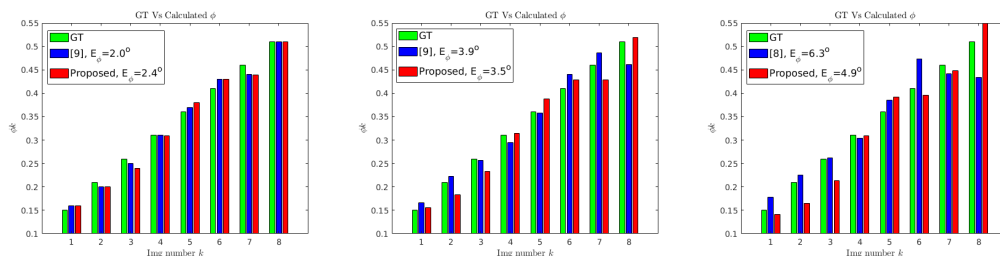


Figure 4. Comparison of proposed method vs [9] for calculating  $\Phi$  for all three synthetic datasets (Figure 2).

in a circular and symmetric disposition over three concentric rings of radius 3.5, 4.5 and 5.5cm respectively. The mean depth is estimated using a ruler. For scenes at roughly 10-20cm away from the camera, the 1mm limit of precision translates to around 0.5-1% uncertainty. To have a reliable ground truth estimate of the light source brightness, we followed very slow acquisition sequences where each LED was given around 100ms to reach the end of its ramp up time. Between capturing from different LEDs we had a delay of a few seconds to avoid overheating.

To simulate significant variation of light source brightness in order to stretch our algorithm to its full potential, we shot images with varying exposure.

Qualitative comparison with competitors is shown in Figures 6 and 7. The performance of the proposed approach can be qualitatively appreciated for the computation of the light source brightness  $\Phi$  showed in Figure 9. For this purpose we evaluated the actual brightness considering two approaches that provided equivalent results. The first approach involved placing the sensor of a LUX meter in the middle of the field of view at a known distance. The second and more reliable approach was image based. For each LED, we took a picture of a planar surface perfectly aligned to the image plane painted with a water based barium sul-

fate coating formulated to yield paint with high diffuse reflection. Then, the  $\Phi$  estimation was done by ignoring the  $z$  and  $c$  update steps on our algorithm (known geometry & reflectance on the reference object).

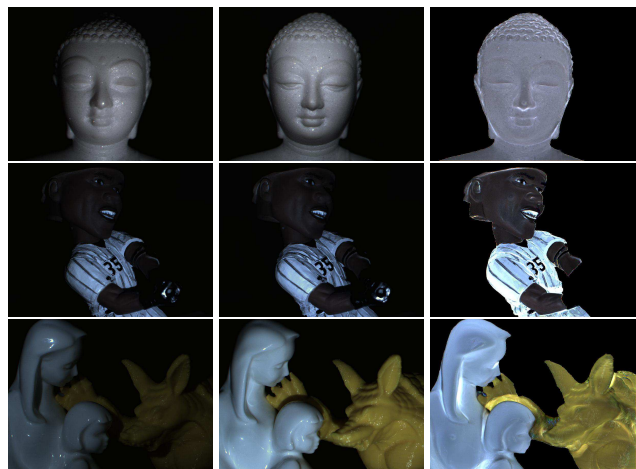


Figure 5. Darkest (Left) and brightest (middle) samples of our real datasets and the computed albedos (Right). Note that the coloring of the albedos is done by noting each pixel's position on the Bayer pattern of the raw data.



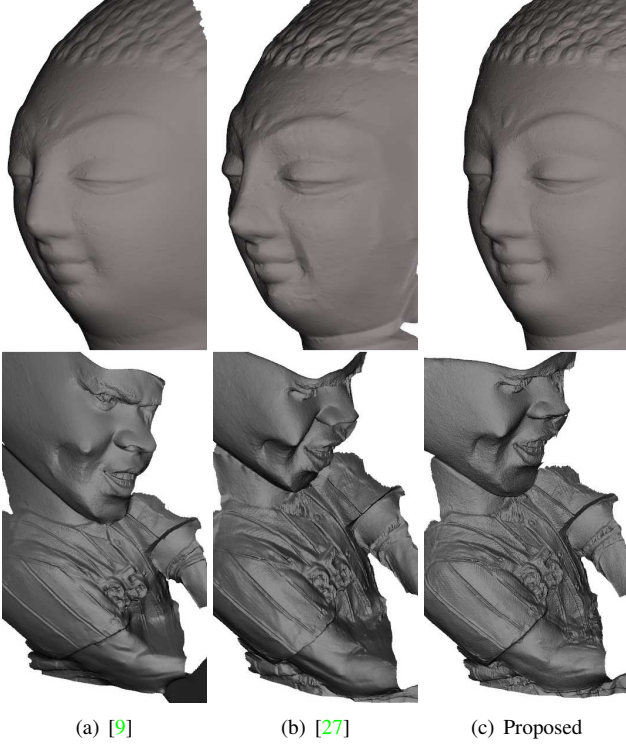


Figure 6. Reconstructions: head of a Buddha statue, plastic base-ball player.

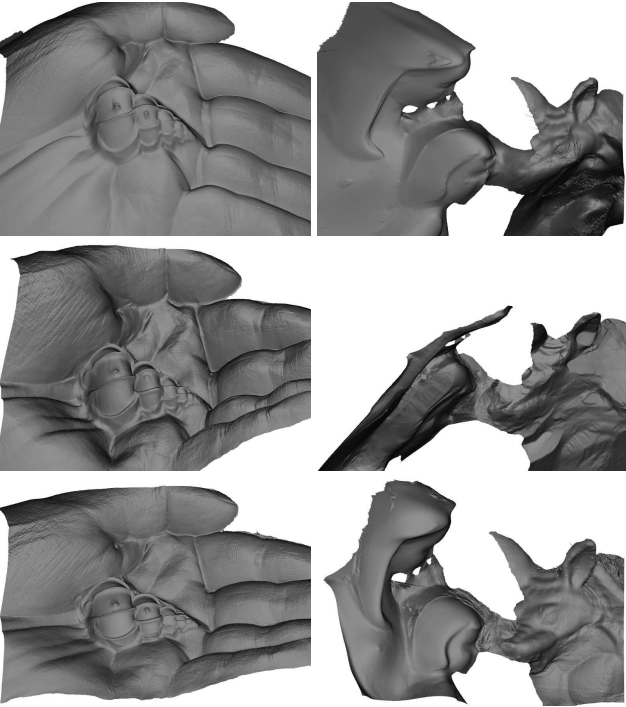


Figure 7. Reconstructions: Top [9], middle [27], bottom Proposed.



Figure 8. Close-up views of the textured reconstructions obtained using the proposed method.

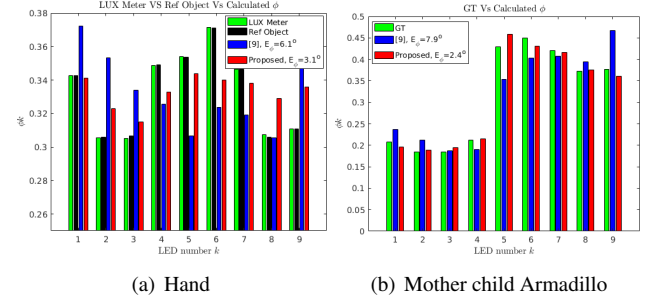


Figure 9. Calculated  $\Phi$  vs ground truth. (left) hand data set, use both LUX meter and reference object as GT (they agree). (right). Mother-child- Armadillo data set shot with variable exposures. The GT  $\Phi$  is the LUX meter's reading times the exposure.

## 6. Conclusion and Further Work

In this work we approached the semi-calibrated near field PS by presenting a robust variational solver approximating the spatially varying light attenuation as well as recovering the shape, the shininess parameter and the light source's intrinsic brightness. Beside assuming perspective viewing geometry, our approach considers the irradiance equation presented in [27] that deals with both diffuse and specular reflectance having a heuristic single lobe formulation that yields a twofold benefit. First of all, simplifications due to image ratio can be exploited in the variational solver. Secondly, objects made of heterogeneous materials can be reconstructed. We proved our method over both synthetic and real world data showing its actual functionality in comparison with state of the art approaches and devices capable of providing light measurements.

The main limitation of our approach is its inability to deal with surfaces with a vastly different reflectance from the Blinn-Phong specular model or the Lambertian (diffuse) model. In addition, the light attenuation map recovery has a high computational cost requiring around half an hour of processing on top end server hardware.

In the same context of semi-calibrated near field PS, an important future work would be to consider non-negligible ambient light. This makes the problem much harder to solve since the photometric parallax is reduced by the presence of an additional offset. Finally, we foresee the applicability of our light attenuation method estimation in the case of a significantly attenuating medium such as water.



## References

- [1] A. Abrams, C. Hawley, and R. Pless. Heliometric stereo: Shape from sun position. In *European Conference on Computer Vision (ECCV)*, pages 357–370, 2012. 2
- [2] J. Ackermann, F. Langguth, S. Fuhrmann, and M. Goesele. Photometric stereo for outdoor webcams. In *Computer Vision and Pattern Recognition (CVPR)*, pages 262–269, 2012. 2
- [3] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 25(10):1239–1252, 2003. 2
- [4] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *International Journal of Computer Vision (IJCV)*, 35(1):33–44, 1999. 2
- [5] J. F. Blinn. Models of light reflection for computer synthesized pictures. In *Conference on Computer Graphics and Interactive Techniques, (SIGGRAPH)*, pages 192–198, 1977. 3
- [6] E. J. Candès, M. B. Wakin, and S. P. Boyd. Enhancing sparsity by reweighted  $l_1$  minimization. *Journal of Fourier Analysis and Applications*, 14(5):877–905, 2008. 5
- [7] M. Chandraker, J. Bai, and R. Ramamoorthi. On Differential Photometric Reconstruction for Unknown, Isotropic BRDFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 35(12):2941–2955, 2013. 2
- [8] M. K. Chandraker, S. Agarwal, and D. J. Kriegman. Shadowcuts: Photometric stereo with shadows. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007. 2
- [9] D. Cho, Y. Matsushita, Y.-W. Tai, and I. Kweon. Photometric stereo under non-uniform light intensities and exposures. In *European Conference on Computer Vision (ECCV)*, pages 170–186, 2016. 1, 3, 6, 7, 8
- [10] J. J. Clark. Active photometric stereo. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 29–34, 1992. 2, 4
- [11] J. J. Clark. Photometric stereo using LCD displays. *Image and Vision Computing*, 28(4):704–714, 2010. 2
- [12] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *ACM Trans. Graph.*, 1(1):7–24, 1982. 3, 6
- [13] P. A. Davis and L. A. Soderblom. Modeling crater topography and albedo from monoscopic viking orbiter images: 1. methodology. *Journal of Geophysical Research: Solid Earth*, 89(B11):9449–9457, 1984. 2
- [14] P. E. Debevec, T. Hawkins, C. Tchou, H. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *Conference on Computer Graphics and Interactive Techniques, (SIGGRAPH)*, pages 145–156, 2000. 2
- [15] P. E. Debevec, A. Wenger, C. Tchou, A. Gardner, J. Waese, and T. Hawkins. A lighting reproduction approach to live-action compositing. *ACM Trans. Graph.*, 21(3):547–556, 2002. 2
- [16] A. S. Georgiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. In *International Conference on Computer Vision (ICCV)*, pages 816–825, 2003. 3
- [17] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially-varying brdfs from photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 32(6):1060–1071, 2010. 2
- [18] C. Hernández, G. Vogiatzis, and R. Cipolla. Overcoming shadows in 3-source photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 33(2):419–426, 2011. 2
- [19] B. K. P. Horn. Obtaining shape from shading information. *The Psychology of Computer Vision*, Winston, P. H. (Ed.), pages 115–155, 1975. 2
- [20] Y. Iwahori, H. Sugie, and N. Ishii. Reconstructing shape from shading images under point light source illumination. In *International Conference on Pattern Recognition (ICPR)*, pages 83–87, 1990. 2, 4
- [21] J. Jung, J. Y. Lee, and I. S. Kweon. One-day outdoor photometric stereo via skylight estimation. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4521–4529, 2015. 2
- [22] J. Lambert. *Photometria sive De mensura et gradibus luminis, colorum et umbrae*. Sumptibus viduae Eberhardi Klett, typis Christophori Petri Detleffsen, 1760. 3
- [23] K. M. Lee and C.-C. Kuo. Shape from photometric ratio and stereo. *Journal of Visual Communication and Image Representation*, 7(2):155 – 162, 1996. 2
- [24] S. Lee and M. Brady. Integrating stereo and photometric stereo to monitor the development of glaucoma. *Image and Vision Computing*, 9(1):39–44, 1991. 2
- [25] F. Logothetis, R. Mecca, Y. Quéau, and R. Cipolla. Near-field photometric stereo in ambient light. In *British Machine Vision Conference (BMVC)*, 2016. 2
- [26] R. Mecca and Y. Quéau. Unifying Diffuse and Specular Reflections for the Photometric Stereo Problem. In *Workshop on Applications of Computer Vision (WACV)*, 2016. 2
- [27] R. Mecca, Y. Quéau, F. Logothetis, and R. Cipolla. A single lobe photometric stereo approach for heterogeneous material. *SIAM Journal on Imaging Sciences*, 2016. 2, 3, 5, 6, 7, 8
- [28] R. Mecca, E. Rodolà, and D. Cremers. Realistic photometric stereo using partial differential irradiance equation ratios. *Computers & Graphics*, 51:8–16, 2015. 2
- [29] R. Mecca, A. Wetzler, A. Bruckstein, and R. Kimmel. Near Field Photometric Stereo with Point Light Sources. *SIAM Journal on Imaging Sciences*, 7(4):2732–2770, 2014. 2, 4
- [30] R. Mecca, A. Wetzler, R. Kimmel, and A. M. Bruckstein. Direct shape recovery from photometric stereo with shadows. In *International Conference on 3D Vision (3DV) 2013*, pages 382–389, 2013. 2
- [31] S. K. Nayar. Sphereo: Determining depth using two specular spheres and a single camera, 1989. 2
- [32] S. K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Transactions on Robotics and Automation*, 6(4):418–431, 1990. 3

- [33] S. K. Nayar, K. Ikeuchi, and T. Kanade. Surface reflection: Physical and geometrical perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 13(7):611–634, 1991. 3
- [34] OSRAM - Opto Semiconductors. *Platinum DRAGON LED*, 2009. 2
- [35] T. Papadhimetri and P. Favaro. A New Perspective on Uncalibrated Photometric Stereo. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013. 2, 3
- [36] T. Papadhimetri and P. Favaro. Uncalibrated Near-Light Photometric Stereo. In *British Machine Vision Conference (BMVC)*, 2014. 3
- [37] B. T. Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, 1975. 3
- [38] Y. Quéau, F. Lauze, and J.-D. Durou. A L1-TV Algorithm for Robust Perspective Photometric Stereo with Spatially-Varying Lightings. In *Scale Space and Variational Methods in Computer Vision (SSVM)*, 2015. 2
- [39] Y. Quéau, R. Mecca, and J.-D. Durou. Unbiased photometric stereo for colored surfaces: A variational approach. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016. 2
- [40] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan. Self-calibrating photometric stereo. In *Conference on Computer Vision and Pattern Recognition, (CVPR)*, pages 1118–1125, 2010. 2
- [41] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi. Bi-polynomial modeling of low-frequency reflectances. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 36(6):1078–1091, 2014. 3
- [42] W. Smith and F. Fang. Height from photometric ratio with model-based light source selection. *Computer Vision and Image Understanding*, 145:128–138, 4 2016. 2
- [43] K. Sunkavalli, F. Romeiro, W. Matusik, T. Zickler, and H. Pfister. What do color changes reveal about an outdoor scene? In *Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, June 2008. 2
- [44] K. Sunkavalli, T. E. Zickler, and H. Pfister. Visibility subspaces: Uncalibrated photometric stereo with shadows. In *European Conference on Computer Vision (ECCV)*, pages 251–264, 2010. 2
- [45] A. Tankus and N. Kiryati. Photometric stereo under perspective projection. In *International Conference on Computer Vision (ICCV)*, pages 611–616, 2005. 2
- [46] K. E. Torrance and E. M. Sparrow. Radiometry. chapter Theory for Off-specular Reflection from Roughened Surfaces, pages 32–41. Jones and Bartlett Publishers, Inc., USA, 1992. 3
- [47] G. Vogiatzis, P. H. S. Torr, and R. Cipolla. Multi-view stereo via volumetric graph-cuts. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 391–398, 2005. 2
- [48] L. B. Wolff and E. Angelopoulou. 3-d stereo using photometric ratios. In *European Conference on Computer Vision (ECCV)*, pages 247–258, 1994. 2
- [49] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):134–144, 1980. 1, 2
- [50] C. Wu, S. G. Narasimhan, and B. Jaramaz. A Multi-Image Shape-from-Shading Framework for Near-Lighting Perspective Endoscopes. *International Journal of Computer Vision (IJCV)*, 86(2-3):211–228, 2010. 2
- [51] T.-P. Wu, K.-L. Tang, C.-K. Tang, and T.-T. Wong. Dense photometric stereo: A markov random field approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 28(11):1830–1846, 2006. 2
- [52] W. Xie, C. Dai, and C. C. L. Wang. Photometric stereo with near point lighting: A solution by mesh deformation. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4585–4593, June 2015. 2, 4
- [53] G. Zeng, S. Paris, L. Quan, and F. Sillion. Progressive surface reconstruction from images using a local prior. In *International Conference on Computer Vision (ICCV)*, pages 1230–1237, 2005. 2