A Closed-Form Solution to Uncalibrated Photometric Stereo via Diffuse Maxima

Paolo Favaro
Heriot-Watt University
Edinburgh, UK
p.favaro@hw.ac.uk

Thoma Papadhimitri
Heriot-Watt University
Edinburgh, UK
tp71@hw.ac.uk

Abstract

In this paper we propose a novel solution to uncalibrated photometric stereo. Our approach is to eliminate the so-called generalized bas relief (GBR) ambiguity by exploiting points where the Lambertian reflection is maximal. We demonstrate several noteworthy properties of these maxima: 1) Closed-form solution: A single diffuse maximum constrains the GBR ambiguity to a semi-circle in 3D space; 2) Efficiency: As few as two diffuse maxima in different images identify a unique solution; 3) GBR-invariance: The estimation error of the GBR parameters is completely independent of the true parameters. Furthermore, our algorithm is remarkably robust: It can obtain an accurate estimate of the GBR parameters even with extremely high levels of outliers in the detected maxima (up to 80% *of the observations). The method is validated on real data and achieves state-of-the-art results.*

1. Introduction

The Lambertian model remains one of the most widely adopted models for diffuse reflectance in computer vision. Despite such a widespread use, this model has not been exploited to its full potential. In this paper, we show that locations of maximal diffuse brightness carry very useful geometrical information about shape and light. We demonstrate their potential in the case of *uncalibrated photometric stereo*, where no prior knowledge about the illumination, geometry, and reflectance is available. Recall that in uncalibrated photometric stereo, in absence of light calibration, the normal field of the object can be obtained from the Lambertian reflectance model only up to a 9-parameter linear ambiguity [12]. This ambiguity can be further reduced to 3 parameters, the so-called *generalized bas-relief ambiguity* (GBR) [4], via the integrability constraint, which enforces that a valid surface can be reconstructed from the estimated normal field.

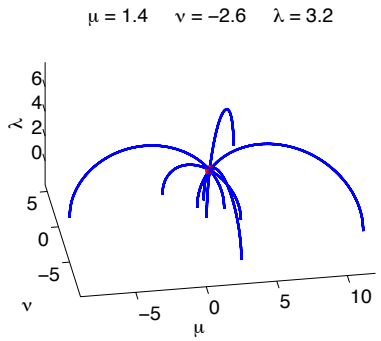

Figure 1. GBR ambiguity solved via LDR constraints. The integrability constraint alone allows the GBR parameters (μ, ν, λ) to take any value in the 3D space. A single LDR maximum constrains the GBR parameters to a semi-circle. Other LDR maxima define other semi-circles all intersecting at the true GBR parameters.

Despite its apparent simplicity, finding a way to fix the GBR parameters with assumptions as realistic as possible is still an open research problem. We make the assumption that one can detect and exploit maxima of the *Lambertian diffuse reflectance* (LDR) component (as opposed to the specular reflectance component [7, 8, 16, 19] or the temporal maxima [15]). We show that each of these maxima restricts the space of GBR ambiguities to a semi-circle in 3-D space, and that as few as two maxima from differently illuminated images yield a unique solution for the GBR parameters (see Fig. 1). The semi-circles can be efficiently computed in closed-form (see sec. 5.1).

A second fundamental challenge that was not addressed so far is that the estimation error of the GBR parameters should not depend on the true parameters, a condition that we call *GBR-invariance*. In other words, an algorithm should yield the same estimated parameters, regardless of which solution satisfying the integrability constraint one starts from. Current state-of-the-art methods suffer from this shortcoming and obtain a stable estimate by running their algorithms several times with different initial solutions

and by averaging all the estimates. In contrast, our algorithm needs to be run only once as we demonstrate in sec. 5.

2. Prior Work

Photometric stereo [24] is a method for estimating shape and reflectance of an object, using three [24] or more [27] images of a static object taken from a fixed viewpoint and under different lighting conditions. When the illumination directions and intensities are known, the problem can be solved via a linear system.[1] When no prior knowledge about the illumination, geometry, and reflectance is available, the problem is called *uncalibrated photometric stereo* (UPS). UPS has been addressed with a variety of techniques that make explicit or implicit use of the GBR ambiguity. For instance, in [2], the GBR ambiguity is solved by minimizing the entropy of the albedo distribution following the argument that incorrect GBR parameters result in spreading albedo values. This approach relies on the assumption that albedos are based on a few intensity values. A similar philosophy is to group the normal-albedo distribution based on color appearance [21]. The use of color, when available, allows clustering based on the chromaticity. Another example is to introduce additional constraints to determine the GBR ambiguity by exploiting specularities of glossy surfaces [7, 8]. This method relies on the ability to correctly detect specularities and on the assumption that objects in the scene have a non negligible specular component. Lagger and Fua [16] notice that both specular and diffuse maxima provide useful information about the illumination direction. However, they discard diffuse maxima and focus only on specular ones. In other investigations, the GBR ambiguity has been eliminated by exploiting inter-reflections [6], by considering the Torrance and Sparrow reflectance model [10] or by considering specular and glossy surfaces [23]. Other approaches instead exploit shadows [18, 14], dimensionality reduction [20], the bilateral symmetry of isotropic BRDFs [1] or consider changing the viewpoint and exploit the Helmoltz reciprocity principle [29]. Hertzmann et al. [13] use a reference object in the scene to generate intensity-normal look-up tables, and can handle fairly general reflectance functions (more precisely, all materials of a single object must be a linear combination of a given basis of materials). A recent work of Chandraker et al. [5] exploits a particular image acquisition setup and spatial and temporal image derivatives to obtain an uncalibrated photometric invariant that can deal with a general isotropic BRDF. Other methods generalize the lighting environment and implicitly extend the Lambertian model so that different normals see a different portion of the illumination hemisphere [3].

[1] This result holds under the Lambertian image formation model, orthographic projection, by considering distant point light sources and when shadows and interreflections can be ignored.

Contributions. We introduce LDR maxima to solve the GBR ambiguity in UPS. Our main contributions are:

- A closed-form solution for a single LDR maximum;
- A robust estimator to find the GBR parameters that tolerates up to 80% incorrect LDR maxima;
- A proof that the solution is GBR-invariant.

Finally, we experimentally demonstrate that our method handles a wide range of real-world scenes, is the most computationally efficient UPS algorithm, and achieves state-of-the-art results.

3. Photometric Stereo

Photometric stereo is the task of recovering the surface of an object with normal field $\mathbf{n}$, given K input images taken from a fixed viewpoint, and illuminated by known distant point light sources with direction $\mathbf{l}$. We denote the unit-normal to the surface at the pixel index p with $\mathbf{n}_p \in S^2$ for $p = 1, \dots, P$, where P is the total number of pixels in an image. Similarly, we denote with $\mathbf{l}_k \in S^2$ the unit-normal corresponding to the k-th illumination direction. Now, let us call $\mathbf{L}_k = e_k \mathbf{l}_k$ the *light vector*, where e_k denotes the light intensity. Also, let $\mathbf{N}_p = \rho_p \mathbf{n}_p$ represent the *generic normal* vector and ρ_p be the albedo coefficient. Then, the intensity measured at a pixel index p with illumination k for the Lambertian case (and under the assumption of a linear response of the camera and orthographic projection) is:

$$I_{p,k} = \rho_p \langle \mathbf{n}_p, \mathbf{l}_k \rangle e_k = \mathbf{N}_p^T \mathbf{L}_k \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes inner product and $(\cdot)^T$ the transpose of a vector (or of a matrix). By sorting the pixels in lexicographical order we can rearrange all the generic normal vectors in a matrix $\mathbf{N} = [\mathbf{N}_1 \ \dots \ \mathbf{N}_P] \in \mathbb{R}^{3 \times P}$, which we call *normal matrix*; similarly we can rearrange the light vectors into a matrix $\mathbf{L} = [\mathbf{L}_1 \ \dots \ \mathbf{L}_K] \in \mathbb{R}^{3 \times K}$, which we call *light matrix*. Then, we can write eq. (1) in the following compact matrix form:

$$\mathbf{I} = \mathbf{N}^T \mathbf{L} \tag{2}$$

where $\{\mathbf{I}\}_{p,k} = I_{p,k}$. If light directions and intensities are given, solving photometric stereo is equivalent to solving the linear system (2) in the unknown normal matrix $\mathbf{N}$. Finally, extracting the albedo ρ_p from the generic normal vectors $\mathbf{N}_p$ can be done via $\rho_p = \|\mathbf{N}_p\|_2$.

4. Uncalibrated Photometric Stereo

If the light matrix is unknown, $\mathbf{N}$ and $\mathbf{L}$ can be obtained up to a linear ambiguity $\mathbf{G} \in GL(3)$ [12], where:

$$\mathbf{I} = \mathbf{N}^T \overbrace{\mathbf{G}^{-1}\mathbf{G}}^{identity} \mathbf{L}. \tag{3}$$

It can be shown [4] that if we impose the *integrability constraint*, then the ambiguities take the form

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix} \tag{4}$$

where the 3 parameters $\lambda \neq 0$, $\mu, \nu \in \mathbb{R}$ represent the group of GBR transformations. In this formulation it is apparent that solving UPS amounts to fixing these 3 parameters. For simplicity, we also fix the sign of λ to be positive. Define $\hat{\mathbf{N}} \doteq \mathbf{G}^{-T}\mathbf{N}$, the *pseudo-normals*, and $\hat{\mathbf{L}} \doteq \mathbf{G}\mathbf{L}$, the *pseudo-lights*. An initial pair $(\hat{\mathbf{N}}, \hat{\mathbf{L}})$ can be computed with the algorithm of [28]. Our task is to find $\hat{\lambda}$, $\hat{\mu}$, and $\hat{\nu}$ such that $\hat{\mathbf{G}}^T\hat{\mathbf{N}} = \mathbf{N}$ and $\hat{\mathbf{G}}^{-1}\hat{\mathbf{L}} = \mathbf{L}$ where

$$\hat{\mathbf{G}}^T = \begin{bmatrix} 1 & 0 & \hat{\mu} \\ 0 & 1 & \hat{\nu} \\ 0 & 0 & \hat{\lambda} \end{bmatrix} \quad \text{and} \quad \hat{\mathbf{G}}^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{\hat{\mu}}{\hat{\lambda}} & -\frac{\hat{\nu}}{\hat{\lambda}} & \frac{1}{\hat{\lambda}} \end{bmatrix}. \tag{5}$$

5. Lambertian Diffuse Reflectance Maxima

The model in eq. (1) is an approximation of the true image formation process. In practice, real scenes have non Lambertian objects, lights may lie close to the objects and have extended surfaces, and other effects such as specularities, shadows, inter-reflections, and noise need to be taken into account. We separate the input images into a component due to the Lambertian reflectance and one due to the specular reflectance by applying the recent *low-rank + sparse outliers* model [17, 26] (see sec. 6).

We assume that the scene contains curved objects (with positive Gaussian curvature) such that there exists a nonempty set $\mathcal{S}_k$ of spatial maxima (where "spatial" refers to the image domain) of the inner product $\langle \mathbf{n}_p, \mathbf{l}_k \rangle$, for a given k-th illumination. We call such points *Lambertian diffuse reflectance* maxima. Notice that we have a maximum only when $\mathbf{n}_p$ and $\mathbf{l}_k$ point in the same direction. Since both vectors are unit-normal, they must also coincide.

5.1. A Closed-Form Solution

Suppose that the set $\mathcal{S}_k$ is given. For each k, an LDR maximum at $p \in \mathcal{S}_k$ constrains the GBR ambiguity to the following set

$$(\tilde{\mu}, \tilde{\nu}, \tilde{\lambda}) \in \arg\max_{\hat{\mu},\hat{\nu},\hat{\lambda}} \left\langle \frac{\hat{\mathbf{G}}^T\hat{\mathbf{N}}_p}{\|\hat{\mathbf{G}}^T\hat{\mathbf{N}}_p\|}, \frac{\hat{\mathbf{G}}^{-1}\hat{\mathbf{L}}_k}{\|\hat{\mathbf{G}}^{-1}\hat{\mathbf{L}}_k\|} \right\rangle; \tag{6}$$

by using $I_{p,k} = \langle \hat{\mathbf{G}}^T\hat{\mathbf{N}}_p, \hat{\mathbf{G}}^{-1}\hat{\mathbf{L}}_k \rangle$, one obtains

$$(\tilde{\mu}, \tilde{\nu}, \tilde{\lambda}) \in \arg\min_{\hat{\mu},\hat{\nu},\hat{\lambda}} \|\hat{\mathbf{G}}^T\hat{\mathbf{N}}_p\| \|\hat{\mathbf{G}}^{-1}\hat{\mathbf{L}}_k\|. \tag{7}$$

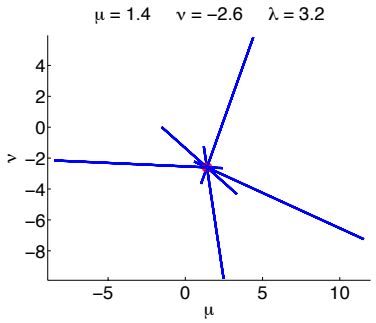

Figure 2. GBR ambiguity solved via LDR constraints. Each segment is the set of solutions in the parameters $\tilde{\mu}$ and $\tilde{\nu}$ for one LDR maximum. Segments must intersect at the correct GBR solution.

Proposition 5.1 *The set of maxima in $(\tilde{\mu}, \tilde{\nu})$ of eq. (8) is the segment between the 2-D points $[\mu_0\ \nu_0]^T$ and $[\mu_1\ \nu_1]^T$, where*

$$\begin{aligned} \mu_0 &= \frac{-\hat{l}_{2,k}^2\hat{n}_{1,p} + \hat{l}_{1,k}\hat{l}_{2,k}\hat{n}_{2,p} + \hat{l}_{1,k}\hat{l}_{3,k}\hat{n}_{3,p}}{\hat{n}_{3,p}(\hat{l}_{1,k}^2 + \hat{l}_{2,k}^2)} \\ \nu_0 &= \frac{\hat{l}_{1,k}\hat{l}_{2,k}\hat{n}_{1,p} - \hat{l}_{1,k}^2\hat{n}_{2,p} + \hat{l}_{2,k}\hat{l}_{3,k}\hat{n}_{3,p}}{\hat{n}_{3,p}(\hat{l}_{1,k}^2 + \hat{l}_{2,k}^2)} \\ \mu_1 &= -\frac{\hat{n}_{1,p}}{\hat{n}_{3,p}} \qquad \nu_1 = -\frac{\hat{n}_{2,p}}{\hat{n}_{3,p}} \end{aligned} \tag{8}$$

and we defined $\hat{\mathbf{L}}_k = [\hat{l}_{1,k}\ \hat{l}_{2,k}\ \hat{l}_{3,k}]^T$ and $\hat{\mathbf{N}}_p = [\hat{n}_{1,p}\ \hat{n}_{2,p}\ \hat{n}_{3,p}]^T$. Given a 2-D point $[\tilde{\mu}\ \tilde{\nu}]^T = [\mu_0\ \nu_0]^T + (1-\alpha)([\mu_1\ \nu_1]^T - [\mu_0\ \nu_0]^T)$ onto the segment, where $\alpha \in [0,1]$, the third GBR parameter is uniquely determined via

$$\tilde{\lambda} = \sqrt{\alpha(1-\alpha)}\frac{|\hat{\mathbf{N}}_p^T\hat{\mathbf{L}}_k|}{|\hat{n}_{3,p}|\sqrt{\hat{l}_{1,k}^2 + \hat{l}_{2,k}^2}} \tag{9}$$

and the trajectory of the 3-D point $[\tilde{\mu}\ \tilde{\nu}\ \tilde{\lambda}]^T$ in the parameter $\alpha \in [0,1]$ forms a semi-circle of radius $\frac{I_{p,k}}{2|\hat{n}_{3,p}|\sqrt{\hat{l}_{1,k}^2+\hat{l}_{2,k}^2}}$.

Proof. See [9]. ∎

The next result tells us how many LDR maxima we need to solve the GBR ambiguity.

Lemma 5.2 *If two LDR maxima correspond to two pseudo-lights, whose first two components are not proportional to each other, then the GBR ambiguity can be solved uniquely.*

Proof. See [9]. ∎

To illustrate the results, we show in Fig. 1 and Fig. 2 how LDR maxima constrain the GBR parameters segments in the (μ, ν) plane and to semi-circles in the (μ, ν, λ) space. All segments or semi-circles intersect at the same point, which corresponds to the correct solution.

Next, we prove a remarkable result: The estimation error of the GBR parameters obtained with LDR maxima

does not depend on the true GBR parameters or the initial pseudo-normals $\hat{\mathbf{N}}$ and pseudo-lights $\hat{\mathbf{L}}$. In practice, this means that our method yields the same solution every time it is run, unlike competing methods (see section 6).

Proposition 5.3 *The constraints on the GBR parameters introduced by an LDR maximum are GBR-invariant.*

Proof. See [9]. ∎

5.2. Existence of LDR Maxima

One may wonder whether LDR maxima are commonly found in images. Indeed, degenerate surfaces that do not generate LDR maxima do exist (*e.g.*, a slope). However, we experimentally find an average of 251 LDR maxima per dataset (see section 6). These statistics are very favorable as only 2 maxima are sufficient to solve UPS.

5.3. Detection and Localization of LDR Maxima

In this section we describe a procedure to detect LDR maxima. The main idea is that we can do so by detecting local *intensity maxima*. In general, however, intensity maxima do not correspond to LDR maxima. In fact, detection of the maxima of $I_{p,k}$ requires

$$\nabla I_{p,k} = \nabla \rho_p \mathbf{N}_p^T \mathbf{l}_k e_k + \rho_p \nabla \mathbf{N}_p^T \mathbf{l}_k e_k = 0, \tag{10}$$

while the LDR maxima are determined by

$$\nabla \mathbf{N}_p^T \mathbf{l}_k = 0. \tag{11}$$

Thus, even small variations of the albedo could result in a displacement of the detected maximum. However, we show that an LDR maximum can be found within a small disc around the detected intensity maximum, whose radius Δ depends on the albedo gradient magnitude and the curvature of the surface.

We approximate the surface of the object around an intensity maximum with a hemisphere. Let

$$\mathbf{P} \doteq [p_1 \; p_2 \; \sqrt{R^2 - p_1^2 - p_2^2}]^T \in \mathbb{R}^3, \tag{12}$$

with $p_1^2 + p_2^2 \leq R^2$, be a point on the hemisphere with radius $R \in \mathbb{R}_+$. If a pixel index p corresponds to the pixel coordinates $[p_1 \; p_2]^T$, then the normals to this surface region can be directly defined as $\mathbf{N}_p = \mathbf{P}/\|\mathbf{P}\|$. Notice that in this case the correct LDR maximum lies at the pixel coordinates $R[l_{1,k} \; l_{2,k}]^T$. If $[p_1 \; p_2]^T$ denotes the detected intensity maximum, we are interested in finding an estimate of the disc radius Δ via the distance

$$\Delta = \left\| R \begin{bmatrix} l_{1,k} \\ l_{2,k} \end{bmatrix} - \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \right\|. \tag{13}$$

Suppose that the surface is Lambertian and we capture an image with illumination direction $\mathbf{l}_k$, where $l_{3,k} > 0$, then the measured intensity is $I_{p,k} = \rho_p \langle \mathbf{P}/\|\mathbf{P}\|, \mathbf{l}_k \rangle e_k$. Detection of the maxima of $I_{p,k}$ requires

$$\nabla I_{p,k} = \nabla \rho_p \frac{\mathbf{P}^T \mathbf{l}_k}{\|\mathbf{P}\|} e_k + \rho_p \frac{\nabla \mathbf{P}^T}{\|\mathbf{P}\|} \left(\mathbf{I_d} - \frac{\mathbf{P}\mathbf{P}^T}{\|\mathbf{P}\|} \right) \mathbf{l}_k e_k = 0 \tag{14}$$

where $\mathbf{I_d}$ denotes the identity matrix. Notice that since

$$\nabla \mathbf{P}^T = \begin{bmatrix} 1 & 0 & -\frac{p_1}{\sqrt{R^2 - p_1^2 - p_2^2}} \\ 0 & 1 & -\frac{p_2}{\sqrt{R^2 - p_1^2 - p_2^2}} \end{bmatrix} \tag{15}$$

we have $\nabla \mathbf{P}^T \mathbf{P} = \mathbf{0}$, which, combined with eq. (15), yields the constraint

$$\nabla \rho_p \mathbf{P}^T \mathbf{l}_k + \rho_p \nabla \mathbf{P}^T \mathbf{l}_k = 0. \tag{16}$$

Assume that the ratio between the albedo gradient and the albedo magnitude is bound by a constant $B > 0$, *i.e.*, $\left\| \frac{\nabla \rho_p}{\rho_p} \right\| \leq B$. Then, we can write

$$\left\| \nabla \mathbf{P}^T \mathbf{l}_k \right\| = \left\| \frac{\nabla \rho_p}{\rho_p} \mathbf{P}^T \mathbf{l}_k \right\| \leq \left\| R \frac{\nabla \rho_p}{\rho_p} \right\| \leq BR, \tag{17}$$

where we used the identity $\|\mathbf{P}\| = R$. Finally, we obtain

$$\left\| \begin{bmatrix} l_{1,k} \\ l_{2,k} \end{bmatrix} - \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \frac{l_{3,k}}{\sqrt{R^2 - p_1^2 - p_2^2}} \right\| \leq BR. \tag{18}$$

For a reliable detection and localization we consider LDR maxima that are not too far off the detected maxima. This means that locally, we can approximate $\mathbf{P}$ with the hemisphere tangent plane at $R\mathbf{L}_k$. This means that $\sqrt{R^2 - p_1^2 - p_2^2} \simeq R l_{3,k}$ and we obtain

$$\Delta = \left\| R \begin{bmatrix} l_{1,k} \\ l_{2,k} \end{bmatrix} - \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \right\| \leq BR^2. \tag{19}$$

Therefore, the correct localization of LDR maxima will be a tradeoff between the curvature of the local surface and the albedo gradient: Surfaces with high curvature can tolerate high albedo gradients and, vice versa, surfaces with small curvature can tolerate little albedo variation. We experimentally find that $\Delta = 1$ is a good tradeoff.

Furthermore, as done in [16], we improve the selection of LDR maxima by discarding maxima that appear at the same spatial location under different illuminations (2 are sufficient to decide) as they are generated by the albedo texture. Third, we discard maxima with low pixel intensities, *i.e.*, maxima from the k-th input image with intensity less than $(\max_p I_{p,k} - \min_p I_{p,k})/2$. We experimentally found that such maxima tend to be less reliable than those with high intensity.

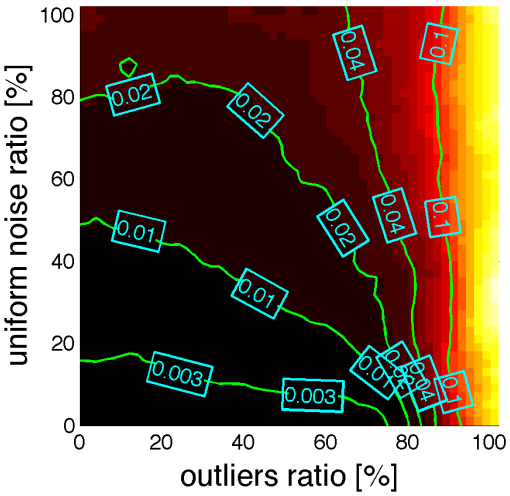

Figure 3. Color coding of the estimation error against outliers and noise ratios. Error levels are shown with solid green lines and the corresponding values are indicated in the cyan box.

5.4. Dealing with Outliers

In practice, the sets $\{\mathcal{S}_k\}_{k=1,\ldots,K}$ contain many outliers, due, for instance, to specularities. We take outliers into account by using a robust estimation procedure.

To simplify the notation, let us define $\mathbf{y} \doteq [\mu\ \nu\ \lambda]^T$, *i.e.*, the vector of GBR parameters. Let S be the sum of the cardinalities of the sets $\{\mathcal{S}_k\}_{k=1,\ldots,K}$. Also, let $\{\mathbf{y}_i\}_{i=1,\ldots,S(S-1)/2}$ be the set of all the intersections between any two semi-circles defined by two detected maxima lying in one of the sets $\{\mathcal{S}_k\}_{k=1,\ldots,K}$. We consider the intersections $\{\mathbf{y}_i\}_{i=1,\ldots,S(S-1)/2}$ as samples of a random variable $\mathbf{Y}$ defining the GBR parameters. We can obtain an approximation to the distribution $\mathbf{f_Y}$ of the random variable $\mathbf{Y}$ via the empirical distribution

$$\mathbf{f_Y}(\mathbf{y}) \simeq \frac{2}{S(S-1)} \sum_{i=1}^{\frac{S(S-1)}{2}} \delta\left(\mathbf{y} - \mathbf{y}_i\right). \qquad (20)$$

Then, the GBR parameters can be chosen as the median $\bar{\mathbf{y}}$ of the empirical distribution, *i.e.*,

$$(\tilde{\mu}, \tilde{\nu}, \tilde{\lambda}) \equiv \bar{\mathbf{y}} = \arg\min_{\hat{\mathbf{y}}} \mathbf{E}_{\mathbf{f_Y}}\left[|\hat{\mathbf{y}} - \mathbf{y}|\right] \qquad (21)$$

where $\mathbf{E}_{\mathbf{f_Y}}$ denotes the expectation with respect to the distribution $\mathbf{f_Y}$.

5.5. Robustness

Because the proposed algorithm can compute the GBR parameters very efficiently, we can evaluate how well it can tolerate several levels of outliers and errors in the detected LDR maxima. We show that the algorithm can tolerate very high levels of outliers, while it is more sensitive to persistent misalignments between normals and lights.

We synthetically generate a set of $K = 12$ light directions. Then, we generate $P = 500$ maxima by matching each normal to one of the light directions. With reference to Fig. 3, to simulate outliers in the detected LDR maxima, we corrupt `outliers ratio` $\times P$ normals. To simulate errors in the detected LDR maxima, we add to all the normals uniform noise in the range `uniform noise ratio` $\times[-0.1\ +0.1]$ (the persistent misalignment). The performance is computed as the relative error $\epsilon = \frac{\|\mathbf{y}-\bar{\mathbf{y}}\|_2}{\|\mathbf{y}\|_2}$, where $\mathbf{y}$ is the correct vector of GBR parameters and $\bar{\mathbf{y}}$ the estimated one. In Fig. 3 we show the results for several outlier and noise ratios. To ease visualization we smooth the error map and then quantize it. The outlier ratio is shown in the abscissa and ranges from 0% (left) to 100% (right). Similarly, the uniform noise ratio is shown in the ordinate axis and ranges from 0% (bottom) to 100% (top). Notice that the darkest region (bottom-left) denotes no more than 0.3% relative error and it allows up to 75% outliers even when the remaining 25% of the maxima have around 10% uniform additive noise.

6. Experiments

To validate our method we use 9 real object datasets: **Redfish** and **Octopus** of **5** images each (courtesy of Neil Alldrin[2]); **Cat**, **Buddha**, **Owl**, **Horse** and **Rock** of **12** images each (courtesy of Dan Goldman and Steven Seitz[3]); **Puppet** of **15** images taken in our lab with a Canon 5D Mark II; **Face** of **64** images (Yale Face Database [11]). All the experiments were run with Matlab on a Mac (2.66 GHz CPU, Intel Core Duo) platform.

Image Acquisition of the Puppet. To take the photos of the Puppet, we used a Canon 5D Mark II camera and a 30-watt incandescent light source attached on a tripod. We placed the light source in 15 different positions (far enough from the object in order to satisfy the distant point light source assumption) and took 3 photos for each position. The 3 photos were then averaged to reduce noise. Ambient image subtraction was also performed. Because the light source has constant intensity, we scaled each image with a random scalar between 0.5 and 1.5 (this was done for the face dataset as well). An object mask was also created by simple thresholding.

Image Pre-Processing. Eq. (2) dictates that the intensity matrix $\mathbf{I}$ should have rank 3. Because real world data are affected by noise, shadows, specularities and other non-Lambertian effects, the rank will be different, typically bigger, than 3. Hence, we perform a pre-processing step based

[2] http://neilalldrin.com/research

[3] http://www.cs.washington.edu/education/courses/csep576/05wi/projects/project3/project3.htm

Figure 4. **Removal of non Lambertian effects.** Left: one of the input images. Middle: detected non Lambertian components (shadows and specularities) via the low-rank matrix recovery pre-processing step. Right: Low-rank and noise-free data to be used in our algorithm.

on a recent convex optimization approach [17, 26] that recovers the low-rank intensity matrix without having any information about what entries are corrupted:

$$\min_{\mathbf{A},\mathbf{E}} ||\mathbf{A}||_* + \gamma||\mathbf{E}||_1 \quad s.t \quad \mathbf{I} = \mathbf{A} + \mathbf{E} \qquad (22)$$

where $\mathbf{A}$ is the noise-free low-rank data and $\mathbf{E}$ is a sparse matrix with the outliers. $||\cdot||_*$ and $||\cdot||_1$ denote the nuclear and l_1 norm respectively. $\gamma > 0$ is a weighting parameter that defines the amount of outliers in the data. This parameter depends on the input images and, as suggested in [25], we fix it to $\gamma = \frac{\kappa}{\sqrt{P}}$ where κ is a constant and P is the total number of pixels of a single input image. We fixed $\kappa = 1.7$ for all datasets having at least 12 images and $\kappa = 3$ for the others. Fig. 4 shows how the pre-processing algorithm removes shadows and specularities in the case of the Face dataset.

Implementation Details. We impose the integrability constraint as described in [28] and the surface pseudo-normals $\hat{\mathbf{N}}$ and pseudo-lights $\hat{\mathbf{L}}$ are obtained up to the GBR ambiguity. The initial group of all the LDR maxima is detected by using the built-in Matlab function "imregionalmax" after applying a small Gaussian smoothing to all the input images. This function gives as output a set of single pixels for each image corresponding to the intensity local maxima. As described in sec. 5.3 we select as local maxima all pixels in a disc neighborhood of radius 1 pixel. Then, we discard all local maxima that appear (at least twice) at the same spatial location for different illumination conditions and those that have low intensity values.

Results. Because no ground truth data is available for these datasets, we consider as ground truth for the normal maps the ones obtained from the calibrated photometric stereo method. We compare our experimental results with the state-of-the-art in uncalibrated photometric stereo [21, 2]. For the accuracy evaluation we consider the mean angular error of the estimated normals with respect to the calibrated case. In Table 1 we show the performance comparison of all methods in the case of datasets with 12 input images (top) and in the case of datasets with 5 and 15 input images (bottom). Notice that $\sigma < 10^{-12}$ in all the reconstructions as predicted by our analysis (see Proposition 4.1). Since our method provides a GBR-invariant solution, we always get the same estimate no matter what pseudo-normals and pseudo-lights we start from. This is a major advantage over prior work.

Table 1. Performance comparison with the entropy minimization (EM) method [2] and the Self Calibrating Photometric Stereo (SCPS) method [21]. We show the mean (ϵ) and standard deviation (σ) of the angular error of the estimated normal maps.

12 images	A/E	Rock	Buddha	Horse	Cat	Owl
EM	ϵ	22.16	15.05	20.65	15.39	18.48
	σ	1.88	2.19	3.85	3.78	5.58
SCPS	ϵ	24.88	13.58	21.01	6.15	10.47
	σ	7.42	4.93	9.57	2.83	4.75
Our method	ϵ	**11.61**	**4.98**	**4.80**	**5.37**	**6.63**
	σ	**0**	**0**	**0**	**0**	**0**
Our method with GT	ϵ	**2.67**	**3.11**	**2.27**	**2.72**	**5.76**
	σ	**0**	**0**	**0**	**0**	**0**

5 images	A/E	Octopus	Red Fish	**15 images**	Puppet
EM	ϵ	9.03	8.63	EM	26.36
	σ	0.76	1.14		2.39
SCPS	ϵ	13.23	7.60	SCPS	NA
	σ	9.85	4.32		NA
Our method	ϵ	**6.64**	**5.60**	Our method	**12.15**
	σ	**0**	**0**		**0**
Our method with GT	ϵ	**2.87**	**4.47**	Our method with GT	**5.63**
	σ	**0**	**0**		**0**

Fig. 5 shows the comparison of the obtained normal and depth maps for the methods [21, 2] in the case of the Buddha dataset. The depth map is obtained by integrating the normal map with a Poisson solver [22]. Figure 6 shows all the depth maps obtained by our method across all the datasets (bottom row) and the depth maps obtained from the calibrated photometric stereo method (middle row). Finally, the average time [4] to run our method, implemented in non-optimized Matlab code, on all the datasets is 13.5 seconds. Notice that the average time for [2] is 62 seconds, while for [21] is 10 minutes. Moreover, since our standard deviation is almost zero, we only need to run the algorithm once, while the other methods need to be run several times. To appreciate the fine details of the reconstructed surfaces, in Fig 7 we show the results obtained in the case of the **Octopus** dataset. We also tested our method on face reconstruction by using data from Yale's face database and the reconstructed depth map and surfaces are shown in Fig. 8.

[4] Notice that the running time of our algorithm depends mostly on the number of detected local maxima and much less on the size of the input images.

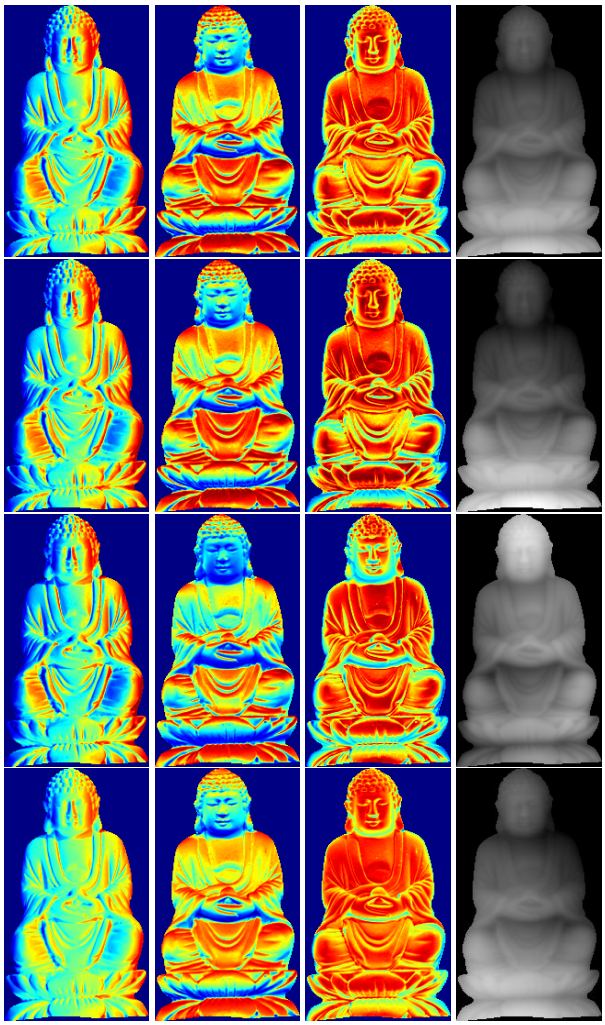

Figure 5. **Comparison of reconstructed normal and depth maps.** From left to right, each column corresponds to the first (x axis), second (y axis), and third (z axis) components of the normal map $\mathbf{n}$ and the depth map, respectively. Top row: Results obtained from the calibrated photometric stereo method. Second row: Results obtained from our method. Third row: Results obtained from the entropy minimization method. Bottom row: Results obtained from the SCPS method.

Discussion. In the last row of each table, we also show the performance when the selected LDR maxima are correct (based on the ground truth obtained from the calibrated photometric stereo). Although we achieve the lowest error rates compared to prior methods, the performance based on the ground truth shows that the algorithm would not achieve 0 error even if the detected maxima were correct. As the theory is exact, this limitation can only be due to non Lambertian components still present in the model. Indeed, the pre-processing step [17, 26] may not fully remove specularities and shadows.

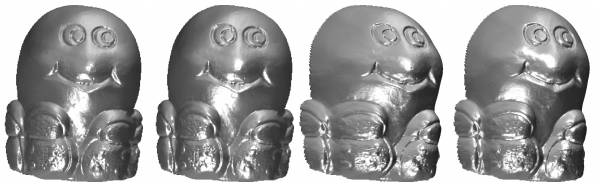

Figure 7. **Octopus reconstruction.** First and third images: Rendered surfaces obtained with the calibrated photometric stereo method. Second and fourth images: Rendered surfaces obtained with our method.

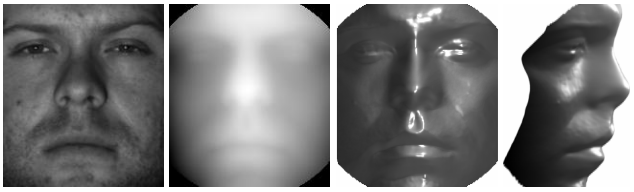

Figure 8. **Results on Yale's Face database.** Left: one of the input images. Second from the left: reconstructed depth map. Third and fourth from the left: Two views of the reconstructed surface.

7. Conclusion

In this paper we presented a simple and fast method for solving the Generalized Bas Relief (GBR) ambiguity of the uncalibrated photometric stereo problem. Our method makes no stringent assumptions about the distribution of lights, only that they are sufficiently different, and can handle a wide variety of albedos and surfaces. We introduce a novel Lambertian constraint for the distribution of lights and normals, which can be computed in closed-form. We use this new constraint to design a very robust algorithm that achieves the best results to date and provides the same estimate regardless of the GBR ambiguity. Key to our approach is a reliable detection of Lambertian diffuse reflectance maxima in the input images and a robust formulation of the GBR estimation process.

8. Acknowledgements

We wish to thank Neil Alldrin for making his code available and Boxin Shi for providing results on our datasets.

References

[1] N. Alldrin and D. Kriegman. Toward reconstructing surfaces with arbitrary isotropic reflectance: A stratified photometric stereo approach. *International Journal of Computer Vision*, pages 1–8, 2007. 2

[2] N. Alldrin, S. Mallick, and D. Kriegman. Resolving the generalized bas-relief ambiguity by entropy minimization. *Computer Vision and Pattern Recognition*, pages 1 –7, 17-22 2007. 2, 6

[3] R. Basri and D. Jacobs. Photometric stereo with general, unknown lighting. *IEEE Conference on Computer Vision and Pattern Recognition*, pages II:374–381, 2001. 2

[4] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *Int'l Journal of Computer Vision*, pages 35(1):33–44, 1999. 1, 3

Figure 6. **Reconstructed depth maps.** Top row: One of the input images. Middle row: Graylevel-coded depth map reconstructed with calibrated photometric stereo. Bottom row: Graylevel-coded depth map obtained with our method.

[5] M. Chandraker, J. Bai, and R. Ramamoorthi. A theory of photometric reconstruction for unknown isotropic reflectances. *IEEE Conference on Computer Vision and Pattern Recognition*, 2011. 2

[6] M. K. Chandraker, F. Kahl, and D. J. Kriegman. Reflections on the generalized bas-relief ambiguity. In *Conference on Computer Vision and Pattern Recognition*, pages 788–795, 2005. 2

[7] O. Drbohlav and M. Chantler. Can two specular pixels calibrate photometric stereo? *Proc. of Int'l Conf. on Computer Vision*, pages 1850–1857, 2005. 1, 2

[8] O. Drbohlav and R. Sara. Specularities reduce ambiguity of uncalibrated photometric stereo. *European Conference on Computer Vision*, pages 46–62, 2002. 1, 2

[9] P. Favaro and T. Papadhimitri. A closed-form solution to uncalibrated photometric stereo via diffuse maxima. *Technical Report - Heriot-Watt University*, 2012. 3, 4

[10] A. Georghiades. Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo. *Proc. of Int'l Conf. on Computer Vision*, pages 816–823, 2003. 2

[11] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23:643–660, 2001. 5

[12] H. Hayakawa. Photometric stereo under a light-source with arbitrary motion. *Journal of the Optical Society of America*, 11(11):3079–3089, 1994. 1, 2

[13] A. Hertzmann and S. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying brdfs. *Pattern Analysis and Machine Intelligence*, 27(8):1254–1264, 2005. 2

[14] H. P. K. Sunkavalli, T. Zickler. Visibility subspaces: uncalibrated photometric stereo with shadows. *European Conference on Computer Vision, Part II*, pages 251–264, 2010. 2

[15] S. J. Koppal and S. G. Narasimhan. Clustering appearance for scene analysis. *IEEE Conference on Computer Vision and Pattern Recognition*, 2:1323 – 1330, June 2006. 1

[16] P. Lagger and P. Fua. Retrieving multiple light sources in the presence of specular reflections and texture. *Comput. Vis. Image Underst.*, 111:207–218, August 2008. 1, 2, 4

[17] Z. Lin, M. Chen, and Y. Ma. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices. *UIUC Technical Report UILU-ENG-09-2215*, 2009. 3, 6, 7

[18] T. Okabe, I. Sato, and Y. Sato. Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting conditions. *Proc. of Int'l Conf. on Computer Vision*, pages 1693–1700, 2009. 2

[19] M. Oren and S. K. Nayar. A theory of specular surface geometry. *Int. J. Comput. Vision*, 24:105–124, September 1997. 1

[20] I. Sato, T. Okabe, Q. Yu, and Y. Sato. Shape reconstruction based on similarity in radiance changes under varying illumination. *Proc. of Int'l Conf. on Computer Vision*, pages 1–8, 14-21 2007. 2

[21] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan. Self-calibrating photometric stereo. *In IEEE Conference on Computer Vision and Pattern Recognition*, pages 1118–1125, 2010. 2, 6

[22] T. Simchony, R. Chellappa, and M. Shao. Direct analytical methods for solving poisson equations in computer vision problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:435–446, 1990. 6

[23] P. Tan, S. Mallick, L. Quan, D. Kriegman, and T. Zickler. Isotropy, reciprocity and the generalized bas-relief ambiguity. In *Computer Vision and Pattern Recognition Conference*, pages 1 –8, 17-22 2007. 2

[24] R. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980. 2

[25] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices by convex optimization. In *proceedings of Neural Information Processing Systems (NIPS)*, December 2009. 6

[26] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Asian Conference on Computer Vision*, pages 703–717, 2011. 3, 6, 7

[27] T. Wu and C. Tang. Dense photometric stereo using a mirror sphere and graph cut. *Computer Vision and Pattern Recognition*, pages 140–147, 2005. 2

[28] A. Yuille and D. Snow. Shape and albedo from multiple images using integrability. *Computer Vision and Pattern Recognition*, pages 158–164, 1997. 3, 6

[29] T. Zickler, P. Belhumeur, and D. Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *Int'l Journal of Computer Vision*, 49(2-3):215–227, 2002. 2