Unifying Diffuse and Specular Reflections for the Photometric Stereo Problem

Roberto Mecca*
Engineering Department, University of Cambridge (United Kingdom)
roberto.mecca@eng.cam.ac.uk

Yvain Quéau
Université de Toulouse (France)†
yvain.queau@enseeiht.fr

Abstract

After thirty years of researching, the photometric stereo technique for 3D shape recovery still does not provide reliable results if it is not constrained into very well-controlled scenarios. In fact, dealing with realistic materials and lightings yields a non-linear bidirectional reflectance distribution function which is primarily difficult to parametrize and then arduous to solve. With the aim to let the photometric stereo approach face more realistic assumptions, in this work we firstly introduce a unified irradiance equation describing both diffuse and specular reflection components in a general lighting setting. After that, we define a new equation we call unifying due to its basic features modeling the photometric stereo problem for heterogeneous materials. It is provided by making the ratio of irradiance equations holding both diffuse and specular reflections as well as non-linear light propagation features simultaneously. Performing a wide range of experiments, we show that this new approach overcomes state-of-the-art since it leads to a system of unifying equations which can be solved in a very robust manner using an efficient variational approach.

1. Introduction

Extracting 3D shape information from the amount of light reflected by a static object is a task as difficult as the spreading of light from the source and the material of the object are modeled realistically.

Initial studies aiming at solving the Photometric Stereo (PS) problem [37] extended the Shape from Shading (SfS) problem [10] by adding supplementary information from additional images, making the PS problem easier to solve than SfS. On the other hand, several important simplifications were done at that time, among which: orthographic viewing geometry, uniform light directions and intensities, diffuse reflectance, absence of shadows and discontinuities.

Although those assumptions allowed to prove well-posedness of the PS problem, they constrained this technique to be employed only for very controlled setups.

Contribution In this work, we propose a method able to perform PS in much less controlled setups, by simultaneously handling several nonlinear physical effects in order to make it as close as possible to real-world scenarios. We move beyond the simplifications above, by proposing:

- a single lobe irradiance equation for both diffuse and specular reflections as well as general lightings and perspective viewing geometry;
- a unifying formulation for the PS problem based on partial differential irradiance equation ratios;
- a L^1 norm-based variational method computing directly the depth and the shininess parameter, being robust to shadows and allowing depth discontinuities.

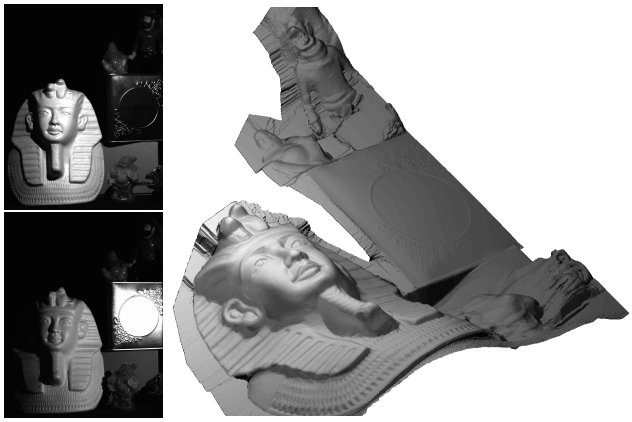


Figure 1. Overview of our approach. On the left, we show 2 out of the 10 images used in this experiment, for a medium-scale scene where shapes having diffuse or specular reflections are present. On the right, we show the reconstructed scene.

*Marie Curie fellow of the Istituto Nazionale di Alta Matematica.

†Experimental setups were provided by Toulouse Tech Transfer, and this collaboration was funded by CNRS GdR 2286 (MIA).

To the best of our knowledge, this approach overcomes state-of-the-art since it represents the first purely data-driven approach dealing with heterogeneous surfaces. An overview of the capability of our approach is presented in Figures 1 and 2.

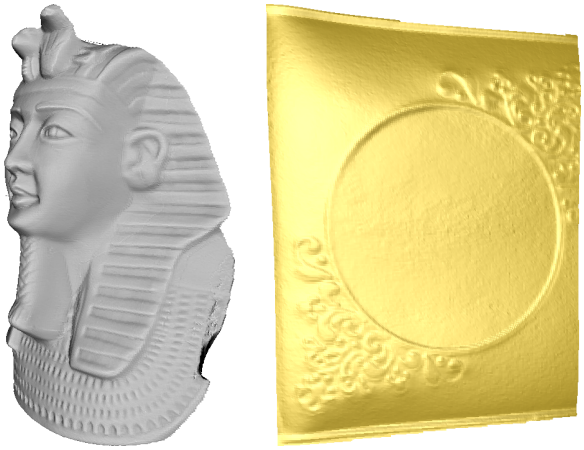


Figure 2. Results of our algorithm on two real-world objects with different reflectance properties, presented in Figure 1. Robustness to non-directional lightings, perspective effects, shadows and non-Lambertian reflectance is enforced.

1.1. Related works

After spending more than thirty years of research, the scientific community did not get to a definitive solution for the PS problem in a very general framework. However, several attempts have been made in order to relax the original assumptions (consisting in orthographic viewing geometry, uniform light propagation and Lambertian reflection, [37]) dealing with realistic irradiance equations.

Camera modeling Perspective viewing geometry for photometric 3D-reconstruction has been introduced by Bruckstein [3] and later developed by Prados and Faugeras [29]. Recently, Papadhimitri and Favaro [26] presented a new perspective parametrization to solve the PS problem without knowledge of the light directions, i.e. Uncalibrated PS (UPS).

Such important feature allowed PS to be exploited for important applications as shape recovery from endoscopic images [25, 6]. In this particular case the anisotropic spread of light cannot be considered negligible due to the proximity of the light source to the inspected surface.

Lightings Starting from initial works dealing with uniform light direction assumption, lighting has been relaxed into considering nearby point light sources spreading light radially. Initially this idea has been proposed by Iwahori *et al*. [14] and Clark [5] later. More recently, Migita *et al*. [21] presented an optimization method for shape recovery while moving a point light source. Papadhimitri and Favaro [27] used the near field lighting for the UPS problem.

However, all those works modeling anisotropic light propagation by adopting point light source parametrization, assumed diffuse reflection which limits the applicability of the PS technique.

Reflectance Shape recovery from specular shading still remains a challenging goal since very common materials provide specular highlights that prevent reasonable reconstructions from the PS technique.

Regarding shading models for specular highlights, several dedicated irradiance equations have been presented so far. Firstly, Torrance and Sparrow [34] presented a physical model based on radiometry principles. Later, Phong [28] showed an empirical model which basically extended the cosine law making it depend also from the viewer direction. The Blinn-Phong shading model [1] extended further the previous one by eliminating some limitation in the analytical formulation maintaining a reliable effectiveness [24].

Outliers Regarding robustness to outliers, most recent works proposed sparsity-enhancing estimators. For instance, the images can be a priori processed according to a low-rank constraint, as suggested by Wu *et al*. [38]. Such an approach was recently improved by Wang *et al*. [35], who proposed to deal with the non-convex nature of the problem through a proximal strategy. Ikehata *et al*. [13] introduced robust estimators derived from the l^1 norm to recover the normals. Eventually, points where the normal cannot be defined, because of the presence of edges or depth discontinuities, can be handled by robust normal field integration methods such as those presented in [7, 30].

Image Ratios New PS models based on non-linear PDEs have attracted increasing interest in the last few years [4, 18, 20, 31]. Most of these approaches considered *image ratios* in order to yield photometric invariant equations [9, 15], and modeled the irradiance equations via PDEs. Mecca *et al*. [18, 20] used specific irradiance equations for diffuse surfaces, proving uniqueness of solution by characteristic strip expansion. Chandraker *et al*. [4] considered more general irradiance equations with unknown light sources, and computed the photometric invariants describing the surface through its isocontours.

Image ratios results more suitable for the PS problem due to the fact that photometric invariant equations are independent on several factors, among which the albedo [36].

This paper is organized by presenting the single lobe irradiance equation in Section 2. The mathematical derivation of the new unifying equation for the PS problem is provided in Section 3. The variational approach solving the PS problem with the new differential formulation is shown in Section 4, yielding the results shown in Section 5.

2. Image formation model

With the aim to formulate a mathematical model as general as possible, we have to deal with two important features of the image formation model, which are the modeling of the pinhole camera and that of the reflectance.

2.1. Camera model

We start by considering the projection of the tridimensional surface Σ into the image plane introduced by Papadhimitri and Favaro in [26]. This allows to easily extend the perspective viewing projection to the orthographic one just by making the focal length tend to infinity. The 3D real world coordinates $(\xi, \eta, \zeta) \in \Sigma$ are projected into the image plane $\overline{\Omega} = \Omega \cup \partial\Omega$ according to the following parametrization

$$(\xi(x,y), \eta(x,y), \zeta(x,y)) = \left(x\frac{f+z(x,y)}{f}, y\frac{f+z(x,y)}{f}, z(x,y) \right) \tag{1}$$

where z is the *depth*. This yields the direction of the outgoing normal to the surface given by

$$\mathbf{n}(x,y) = \left(\nabla z(x,y), -\frac{f+z(x,y)}{f} - \frac{(x,y)\cdot\nabla z(x,y)}{f} \right) \tag{2}$$

where the derivatives of z are considered in the image coordinates, that is $\nabla z(x,y) = \left(\frac{\partial z}{\partial x}(x,y), \frac{\partial z}{\partial y}(x,y)\right)$. Let us denote the normalized vector as follows

$$\overline{\mathbf{n}}(x,y) = \frac{\mathbf{n}(x,y)}{|\mathbf{n}(x,y)|} \tag{3}$$

extending the same notation to the rest of the paper, considering $\bar{\cdot}$ as unit vector.

2.2. Single lobe reflectance model

Coherently with having images as input data including diffuse and specular reflections, we consider single lobe irradiance equation for both reflections. Since diffuse and specular components separation may not always be a reliable procedure to accurately recover shapes done by heterogeneous materials [11, 16, 33], we consider a single irradiance equation that simultaneously parametrizes diffuse and specular reflections, by extending the Blinn-Phong shading model [1] as follows

$$I_i(x,y) = \rho(x,y) a_i(x,y,z) \left(\overline{\mathbf{n}}(x,y,z) \cdot \overline{\mathbf{h}}_i(\mathbf{l}_i, \mathbf{v}) \right)^{\frac{1}{c(x,y)}}. \tag{4}$$

Here ρ and a_i are positive scalar functions representing, respectively, the *albedo* and the attenuation of light, while $c(x,y) > 0$ describes more general reflectance properties. Eventually, the vector function $\mathbf{h}_i = (h_i^1, h_i^2, h_i^3)$ combines information about the view direction $\mathbf{v} = \mathbf{v}(x,y,z)$ and the i^{th} light source $\mathbf{l}_i = \mathbf{l}_i(x,y,z)$. With the aim to unify diffuse and specular components parametrized by a unique irradiance equation we choose $\mathbf{h}_i$ to be as follows

$$\mathbf{h}_i(x,y,z) = \bar{\mathbf{l}}_i(x,y,z) + \min\left\{1, \frac{|1-c(x,y)|}{\varepsilon}\right\} \overline{\mathbf{v}}(x,y,z), \tag{5}$$

where ϵ is a fixed parameter describing the transition between matte and glossy material, allowing the continuous transition state between diffuse and specular component. In the next part we provide an intuitive and effective interpretation of the vector $\mathbf{h}_i$ by showing the dependency on the shininess parameter $c(x,y)$. Let us remark that this new reflectance model generalizes many previous models used in PS, as described hereafter.

2.3. Dealing with well-known settings

Besides being independent from the camera projection model, thanks to the parametrization (2), (4) can be used to deal with surfaces made by heterogeneous materials depending on $c(x,y)$:

Diffuse reflectance by having $c = 1$ (i.e. $\mathbf{h}_i = \mathbf{l}_i$) , we obtain the purely Lambertian shading model;

Specular reflectance setting $0 < c \leq 1-\epsilon$ (i.e. $\mathbf{h}_i = \bar{\mathbf{l}}_i + \overline{\mathbf{v}}$) leads to purely specular Blinn-Phong type reflectance, where the size of the specular lobe depends on the value of c;

Intermediate state is well parametrized taking $1-\epsilon < c < 1$ ($\epsilon = 0.01$ for all our experiments). This provides a transition phase between diffuse and specular components.

Beyond the previous well-known settings, challenging **spatially-varying reflectances** can be modelled by using the piecewise constant space dependency of c over the image domain. This allows us to take into account the very difficult problem of dealing with heterogeneous materials. In fact, the presented algorithm is capable to approximate both z and c.

Moreover, by manipulating the definitions of a_i and $\mathbf{l}_i$, this new model yields several well-known light configurations.

Uniform lighting When $a_i \equiv \phi_i$ and $\mathbf{l}_i$ is independent from the coordinates (x,y,z), it corresponds to the directional lighting model, ϕ_i being a scalar parameter representing the intrinsic intensity of the source;

Point light source When $\mathbf{l}_i(x,y,z) = (\xi_i, \eta_i, \zeta_i) - (\xi(x,y), \eta(x,y), \zeta(x,y))$, one obtains the point light source model, for a source at position (ξ_i, η_i, ζ_i);

Light attenuation In order to deal with more realistic physical effects, we can also consider inverse-of-squared distance light attenuation

$$a_{di}(x,y,z) = \frac{1}{|\mathbf{l}_i(x,y,z)|^2} \tag{6}$$

and an anisotropic angular factor that, without loss of generality, we choose as

$$a_{ai}(x,y,z) = (\bar{\mathbf{l}}_i(x,y,z) \cdot \overline{\mathbf{p}}_i)^{\mu} \tag{7}$$

taking inspiration from [20], where $\overline{\mathbf{p}}_i$ is the principal lighting direction of the source, i.e. its orientation. This attenuation term holds for most of commercial LEDs [23] and it is usually calibrated by manufacturers, under the form of luminous intensity diagrams. Isotropic sources, which are valid approximations for small angles, are obtained by setting $\mu = 0$, while stronger anisotropy effects can be simulated by increasing the value of μ [19]. Considering both attenuation factors, we get

$$a_i(x,y,z) = \frac{\phi_i a_{ai}(x,y,z)}{|\mathbf{l}_i(x,y,z)|^2}. \tag{8}$$

Despite the apparent variety of the configurations handled by (4), we will show in the following that the 3D-reconstruction problem can be reformulated by the same simple quasilinear PDE, handling the general viewing geometry, reflectance and lighting models previously described.

3. The unifying equation

As already shown in [20], the ratios of irradiance equations as (4) seen as partial differential irradiance equations simplifies the PS problem since it becomes independent by photometric invariant (albedo) and irrational nonlinearities (normalization of the normal vector). We now extend this theory to the more realistic case studied in this paper.

3.1. Partial differential irradiance equation ratios

With the aim to extend such methodology to the more general and realistic irradiance equations (4), we divide those coming from the i^{th} and j^{th} light source, leading to the following equation

$$\underbrace{\frac{(I_i)^c}{(a_i)^c\,\overline{\mathbf{h}}_i \cdot \mathbf{n}}}_{\text{Equation for } I_i} = \overbrace{\frac{\rho^c}{|\mathbf{n}|}}^{\text{Equation for } I_j} = \frac{(I_j)^c}{(a_j)^c\,\overline{\mathbf{h}}_j \cdot \mathbf{n}} \tag{9}$$

with the same dependencies as above, that we neglect from now on whenever the reader does not necessary needs them. Denoting the vector field

$$\begin{aligned}\mathbf{b}_{ij} = \Big(&(a_j I_i)^c \Big(\overline{h}_j^1 - \frac{x}{f}\overline{h}_j^3\Big) - (a_i I_j)^c \Big(\overline{h}_i^1 - \frac{x}{f}\overline{h}_i^3\Big), \\ &(a_j I_i)^c \Big(\overline{h}_j^2 - \frac{y}{f}\overline{h}_j^3\Big) - (a_i I_j)^c \Big(\overline{h}_i^2 - \frac{y}{f}\overline{h}_i^3\Big)\Big)\end{aligned} \tag{10}$$

and the scalar function

$$s_{ij} = \frac{f+z}{f}\Big((a_j I_i)^c\, \overline{h}_j^3 - (a_i I_j)^c\, \overline{h}_i^3\Big) \tag{11}$$

by considering the first and the last part of the chain of equalities (9), we obtain the following quasilinear PDE

$$\mathbf{b}_{ij}(x,y,z) \cdot \nabla z(x,y) = s_{ij}(x,y,z) \tag{12}$$

that we will call as **unifying equation** for the PS problem. It elegantly describes the interaction between two irradiance scenarios depicted in I_i and I_j, whatever the models for the camera projection, the surface reflectance and the type of lighting.

By stacking the $\binom{N}{2}$ vector fields $\mathbf{b}_{ij}$, such that $(i,j) \in \{1 \dots N\}^2$ and $i < j$, into a matrix field $\mathbf{B} : \overline{\Omega} \to \mathbb{R}^{2\times\binom{N}{2}}$, and the $\binom{N}{2}$ scalar fields s_{ij} into a vector field $\mathbf{s} : \overline{\Omega} \to \mathbb{R}^{\binom{N}{2}}$, our new mathematical formulation of the PS problem with N images reads as the following system of unifying equations

$$\mathbf{B}^{\top} \nabla z = \mathbf{s}. \tag{13}$$

3.2. Advantages over previous work

Merging PS data We remark that most of the previous works proposing robust approaches to the PS problem as [13, 22, 31, 35, 38] derived the mathematical formulation by considering each irradiance equation independently from the others. Since our model is derived by coupling irradiance equations considering their ratios, the unifying equation exploits the single view acquisition simplifying the complexity of the problem due to heterogeneous materials.

Robustness An important advantage directly follows from this coupling. It is based on the number of unifying equations when solving PS with $N > 2$. In our formulation, the number of equations to be solved does not increase linearly with the number of images since N images yield $\binom{N}{2}$ unifying equations. Clearly these equations tend to be redundant, providing a natural framework for robust recovery of the shape. Hence, robust estimators enforcing sparsity of the residual [13] are well adapted.

Missing data If, for some pixel (x,y), no information is available in any of the images $I^i(x,y), i \in \{1 \dots N\}$, due for instance to a null albedo (no light reflected at all), the corresponding set of unifying equations (13) is not informative, since it reads as $\mathbf{0} = \mathbf{0}$. Thus, there may be points on the surface that are left undetermined. Denoting by Ω_{missing} the set of such points (which is easily computed a priori by thresholding the values of $\sum_{i=1}^{N} I^i$), we can deal with this issue by modifying the definition of $\mathbf{B}$ and $\mathbf{s}$ according to

$$\mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad \text{over } \Omega_{\text{missing}} \tag{14}$$

$$\mathbf{s} = \mathbf{0} \qquad \text{over } \Omega_{\text{missing}} \tag{15}$$

which will enforce $\nabla z = 0$ over this subset Ω_{missing}. This can be seen as a built-in hole filling of the non-informative

areas to ensure some smoothness and prevent artifacts. This can be very useful when dealing with very specular objects, since in such case the information is concentrated in small areas: performing such self-filling will smoothly ensure continuity between the areas where shape reconstruction is possible. Hence, it becomes possible to reconstruct specular surfaces with few images, while several hundreds are used in state-of-the art approaches [12].

4. Variational resolution

In order to enhance robustness to noise (inherent to the acquisition process) and outliers (such as shadows, or non-differentiable elements of the surface), we consider L^1 optimization, following the recent sparsity-enhancing approaches described in [13, 31]. Adapting such framework to the context of quasilinear system of PDEs (13), we consider the following optimization problem

$$\min_z \left\| \mathbf{B}(z)^\top \nabla z - \mathbf{s}(z) \right\|_{L^1(\overline{\Omega})}, \tag{16}$$

where $L^1(\overline{\Omega})$ is the traditional space of functions whose absolute value is Lebesgue integrable.

We emphasize that this variational problem is conceptually very different from the state-of-the-art sparsity-enhancing technique [13], where l^1 optimization is considered *locally*, in order to approximate the normal to the surface at each pixel. In fact, our formulation considers the global minimum over the image domain, having as unknown the depth, as for instance in [31]. Proceeding so, integrability of normals is not an issue, since piecewise-smooth surfaces are recovered directly, without relying on a posteriori use of dedicated solvers, which are either fast but not robust to discontinuities [32], or robust to discontinuities but slow [7, 30].

4.1. Orthographic camera and directional lightings

In the specific case of orthographic camera and directional lightings, neither $\mathbf{B}$ nor $\mathbf{s}$ depend on z. Hence, the functional to minimize is not coercive since it depends only on ∇z, and not on z. Yet, the knowledge of z in just one point suffices to make it coercive. Alternatively, any least-squares prior z_0 on the solution can be introduced, turning the initial problem (16) into its zero-order Tikhonov regularized version

$$\min_z \left\| \mathbf{B}^\top \nabla z - \mathbf{s} \right\|_{L^1(\overline{\Omega})} + \lambda \left\| z - z_0 \right\|^2_{L^2(\overline{\Omega})} \tag{17}$$

with $\lambda > 0$ very small in order not to bias the results (we systematically used $\lambda = 10^{-9}$). In our experiments, z_0 is a uniform function, which basically fixes the mean value of z, and hence the mean distance from the object to the camera. Dealing with the (functional) L^1 norm requires a totally different machinery from [13], since it involves the partial derivatives of z. In this view, it is convenient to rewrite (17) under its ADMM form [2]:

$$\begin{cases} \min\limits_{z,\mathbf{g}} \left\| \mathbf{g} \right\|_{L^1(\overline{\Omega})} + \lambda \left\| z - z_0 \right\|^2_{L^2(\overline{\Omega})} \\ \text{s.t.} \quad \mathbf{g} = \mathbf{B}^\top \nabla z - \mathbf{s} \end{cases}. \tag{18}$$

Introducing the auxiliary variables $\mathbf{g}$ and $\mathbf{u}$, this new problem can be solved using the alternating direction scheme:

$$\mathbf{g}^{k+1} = \operatorname*{argmin}_{\mathbf{g}} \|\mathbf{g}\|_{L^1(\overline{\Omega})} + \alpha \left\| \mathbf{g} - \left(\mathbf{B}^\top \nabla z^k - \mathbf{s} - \mathbf{u}^k \right) \right\|^2_{L^2(\overline{\Omega})} \tag{19}$$

$$z^{k+1} = \operatorname*{argmin}_{z} \frac{\lambda}{\alpha} \|z - z_0\|_{L^2(\overline{\Omega})} + \left\| \mathbf{B}^\top \nabla z - \left(\mathbf{g}^{k+1} + \mathbf{s} + \mathbf{u}^k \right) \right\|^2_{L^2(\overline{\Omega})} \tag{20}$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{g}^{k+1} - \mathbf{B}^\top \nabla z^{k+1} + \mathbf{s} \tag{21}$$

starting from $(z^0, \mathbf{g}^0, \mathbf{u}^0) = (z_0, \mathbf{B}^\top \nabla z_0 - \mathbf{s}, \mathbf{0})$. This scheme can be proven to converge from almost any descent parameter α, whose choice only affects the convergence rate. Eq. (19) can be solved pointwise by shrinkage, and (20) by Gauss-Seidel iterations. Let us remark that no boundary condition is required using this approach [8], while state-of-the-art differential methods for PS require the depth to be known on the boundary [4].

4.2. Perspective viewing and lighting geometry

In the case of perspective camera and/or point light sources, the fields $\mathbf{B} = \mathbf{B}(z)$ and $\mathbf{s} = \mathbf{s}(z)$ depend explicitly on the unknown z. To handle this issue, we consider the semi-implicit scheme

$$z^{k+1} = \operatorname*{argmin}_{z} \left\| \mathbf{B}(z^k)^\top \nabla z - \mathbf{s}(z^k) \right\|_{L^1(\overline{\Omega})} + \lambda \left\| z - z^k \right\|^2_{L^2(\overline{\Omega})} \tag{22}$$

starting from an initial solution $z^0 = z_0$. Convergence towards a local minimum is guaranteed, and we experimentally observed that it was reached in only a few iterations.

4.3. Approximating the shininess parameter c

Since the unifying irradiance equation allows a simultaneous parametrization of diffuse and specular reflectance, the approximation of the shininess coefficient c is an important achievement for the material understanding of the depicted scene. In fact, the shininess parameter can be determined within the process, by enforcing sparsity residuals on the ratio equations (9). A typical iterative scheme writes as:

$$\begin{cases} z^{k+1} = \operatorname*{argmin}\limits_{z} \left\| \mathbf{B}(z^k, c^k)^\top \nabla z - \mathbf{s}(z^k, c^k) \right\|_{L^1(\overline{\Omega})} \\ c^{k+1} = \operatorname*{argmin}\limits_{c} \left\| \left(\frac{I_i a_j(z^k)}{I_j a_i(z^k)} \right)^c - \frac{\overline{\mathbf{h}}_i(z^k, c^k) \cdot \mathbf{n}(z^k)}{\overline{\mathbf{h}}_j(z^k, c^k) \cdot \mathbf{n}(z^k)} \right\|_{L^1(\overline{\Omega})} \end{cases} \tag{23}$$

where the z update can be obtained as described earlier. Taking the logarithm of both members of the second resid-

ual, the c update is given by

$$c^{k+1} = \text{median}\left\{\log\left((I_i a_j)/(I_j a_i)\right) / \log\left((\overline{\mathbf{h}}_i \cdot \mathbf{n})/(\mathbf{h}_j \cdot \mathbf{n})\right)\right\}. \tag{24}$$

As challenging proof of concept, we demonstrate the working principle of our method on synthetic data from the MERL dataset [17] as well as real image scenarios.

5. Experiments

With the aim of showing a fair comparison of our method with the state of the art algorithms, we firstly consider the simplified version of our model leading to the orthographic scenario explained in Section 4.1. Afterwards, more realistic results considering the perspective viewing and lighting geometry will be shown.

5.1. Robustness to realistic outliers

To quantitatively evaluate the robustness of our method, we performed tests on synthetic data, considering first the same setup as in [13]. That restricts our model by considering orthographic projection, directional lightings and known reflectance ($c = 1$) with outliers consisting in self-shadows and small additive specular spots generated according to the bichromatic Blinn-Phong model (Figure 3).

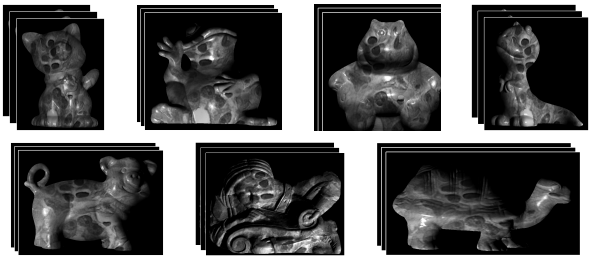


Figure 3. Synthetic images used for quantitative evaluation, generated according to the dichromatic Blinn-Phong reflectance model, with self-shadowing effects. Taking $c = 1$, both shadows and highlights are outliers to the model (4), and hence useful to evaluate the robustness of our method to real-world outliers.

To create the data, we used the seven real-world sets of normals provided in the Harvard's dataset[1], that we integrated into a depth map using our solver (setting $\mathbf{B} = \mathbf{I}$), before recomputing the normals using finite differences. The shading images were calculated under 8 different directional lightings and overlapped with the `hestain.png` built-in image from Matlab, considered as albedo.

Quantitative comparisons of our results with traditional least-squares [37] and state-of-the-art l^1 normal recovery [13] can be found in Table 1. Specular highlights are in this case considered as outliers: reflectance models composed of two lobes can hence be handled using the proposed one-lobe model, which captures the dominant lobe, the other one being trated as outlier.

	Cat					Frog					Hippo					Lizard				
	$\mathbf{n}$		z		t	$\mathbf{n}$		z		t	$\mathbf{n}$		z		t	$\mathbf{n}$		z		t
	ME	std	ME	std		ME	std	ME	std		ME	std	ME	std		ME	std	ME	std	
[37]	7.7	**4.6**	19	15	**30**	6.6	**4.0**	14	12	**37**	8.3	**5.5**	23	15	**31**	7.6	**4.8**	13	9.2	**32**
[13]	6.7	5.3	17	13	305	5.7	4.3	12	11	390	7.2	6.0	21	14	545	7.0	5.8	12	8.6	311
Ours	**6.0**	5.7	**13**	**10**	590	**4.6**	5.1	**9.1**	**8.8**	724	**6.9**	6.5	**20**	**11**	554	**6.9**	5.9	**12**	**8.5**	284

	Pig					Scholar					Turtle				
	$\mathbf{n}$		z		t	$\mathbf{n}$		z		t	$\mathbf{n}$		z		t
	ME	std	ME	std		ME	std	ME	std		ME	std	ME	std	
[37]	7.1	**4.5**	13	12	**39**	6.8	**4.2**	21	11	**97**	7.4	**5.0**	9.3	8.8	**33**
[13]	6.3	5.0	12	11	417	5.7	4.5	18	10	1039	6.6	5.5	8.5	8.7	352
Ours	**5.8**	5.6	**10**	**9.2**	769	**5.1**	4.8	**13**	**7.4**	1865	**6.5**	6.2	**7.6**	**8.0**	526

Table 1. Robustness to outliers, for least-squares [37] and l^1 estimation of the normals [13], and the proposed L^1 estimation of the depth, regarding the synthetic data presented in Figure 3. We show the error on the normals orientations $\mathbf{n}$ in degrees, the absolute error on the depth z in pixels, and the CPU time t in seconds (CPU times were evaluated considering Matlab codes executed on a i7 at 3.4 GHz). Considering the depth, rather than the normals, improves the results because it includes an implicit smoothing.

5.2. Handling arbitrary reflectances

Here, we question the ability of the proposed BRDF model (4) to handle real-world reflectances (Figures 4 and 5). To this purpose, we recovered the shape of a piece of sphere rendered according to the BRDFs of the MERL database [17]. First, we applied the proposed variational scheme using $N = 56$ images while imposing $c = 1$: some partly retro-reflective materials (`pickled-oak-256`) are handled since this effect can be considered as outlier. Yet, since many materials in this database are mostly specular, it is not realistic to assume that specularities are outliers to a diffuse model, hence the slopes are over-estimated in most cases (`alum-bronze`). On the other hand, when simultaneously estimating the shape and the reflectance as discussed in Section 4.3, most materials of the database are reasonably recovered. Notable exceptions include very dark materials (`black-obsidian`) or highly retro-reflective materials (`polyethylene`).

We also used this dataset to perform experiments on the required number of input images. Figure 4 shows that reliable 3D-reconstruction results of highly specular materials are obtained from as few as twenty images. This has to be compared with the hundreds of images used in state-of-the-art [12]: reducing so much the required number of inputs is made possible by increasing quadratically the number of equations due to image ratios, and by using robust variational recovery.

5.3. Qualitative evaluation on real-world datasets

We eventually performed tests on a real-world dataset representing the current limits of PS having several objects of different materials, inducing depth discontinuities and shadows.

[1] http://vision.seas.harvard.edu/qsfs/Data.html

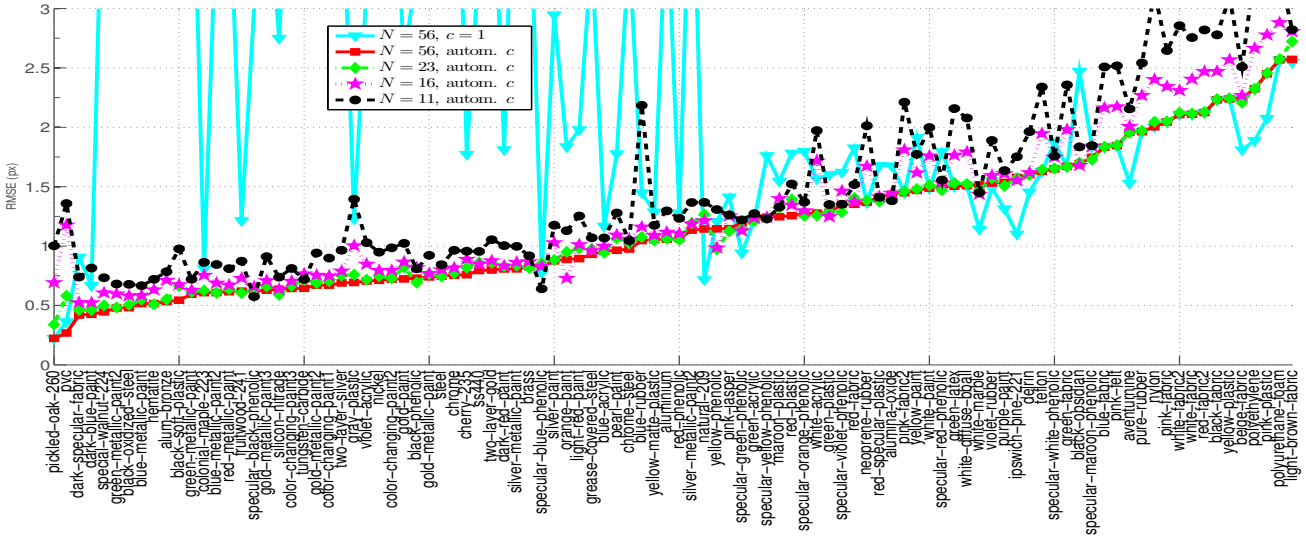


Figure 4. RMSE on the depth, as a function of the BRDF. Most of specular materials are correctly handled with a dozen of images and automatic reflectance estimation.

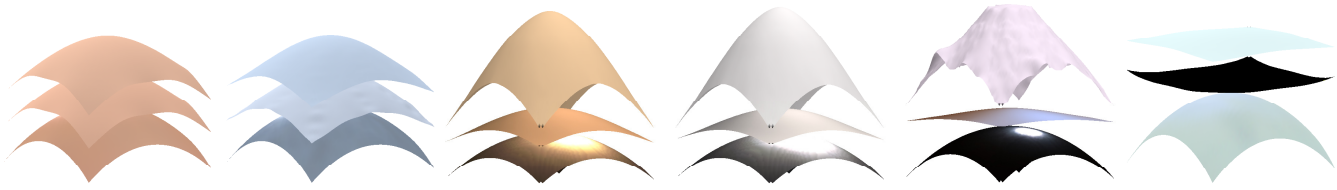


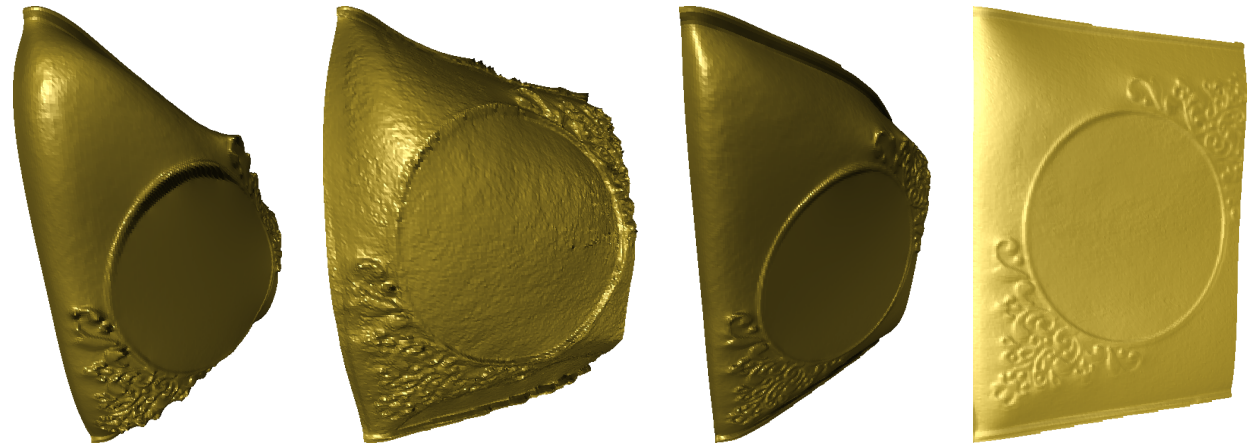
Figure 5. 3D-Reconstruction results for several materials from the MERL database, sorted according to their ranking in Figure 4. Bottom row shows the ground-truth surface and reflectance, middle row shows the shape recovered while estimating automatically c using (23). Top row shows the results with $c = 1$.

The dataset we used is presented in Figure 1. It consists of a scene with several objects placed approximately at 50cm of a calibrated pinhole camera ($f \approx 32$mm), lit successively by 10 calibrated LEDs located approximately at 30cm from the scene. Besides the effects of perspective camera and non-directional lightings, the originality of this scene is the presence of numerous objects, while PS is usually applied to one single object. The difficulties induced are the presence of objects with various reflectances (especially the tea box, which is metallic), the stronger shadowing effects, the presence of depth discontinuities and the missing data (the legs of the Haddock character are black). To deal with the piecewise constantly varying reflectance, we manually segmented the objects. Then, we applied the alternating scheme (23) to each object individually, providing us with an initial 3D-reconstruction (cf. figure 5) as well as with an initial guess for the c values. Then, as shown in Figure 1, we reconstructed the full scene. Depth discontinuities are successfully recovered together with $c \approx 1$ to the bust (plaster), $c \approx 0.05$ to the tea box (metal), $c \approx 0.4$ to the Haddock character (plastic), $c \approx 0.35$ to both Buddha dolls, and $c \approx 1$ to the rest of the scene.

While Figure 2 shows the results obtained with the full scene pictured by 1M pixels images, Figure 1 presents the reconstructions of single objects. It is worth noticing the quality of the reconstruction of the specular tea box and the recovery of the discontinuities for the full scene. Eventually, these results are compared in Figures 6 and 7 to several existing methods for Lambertian PS. Namely the classical PS approach proposed by Woodham [37], the robust approach from [13] and a recent work on PS with pointwise sources [27]. Figure 6 demonstrates the advantage of *global* L^1-based recovery over *local* sparsity-enhancing techniques [13], and the importance of modelling appropriately the lighting and the viewing. Besides these important considerations, Figure 7 shows that our approach even provides satisfactory 3D-reconstructions of specular objects.

[37] [13] [27] Ours

Figure 6. Results of several PS algorithms on the bust dataset (plaster). With the classical approach [37] that assumes uniform lighting and orthographic viewing, artifacts due to shadows are visible. By *locally* enforcing robustness as proposed in [13], these artifacts are still visible. The *global* formulation of our approach provides better results, while also correcting the distortion due to pointwise lighting and perspective viewing, as also assumed by [27].



[37] [13] [27] Ours

Figure 7. Results of several PS algorithms on the teabox dataset (metal). Our approach provides a more reliable 3D-reconstruction of this purely specular object, compared to other methods based on the Lambertian assumption.

6. Conclusion and perspectives

In this work we introduced a unifying irradiance equation describing both diffuse and specular reflection components in the most general lighting setting. We derived a new formulation based on coupling those irradiance equations by considering their ratios. The resulting mathematical model consists in a set of quasi-linear PDEs having several advantages with respect to the usual tools aimed at solving the PS problem. In fact, instead of considering specular highlights as outliers, we exploit geometric information of the shape by modeling specular reflectance with a Blinn-Phong extension of the cosine law for diffuse reflection. Furthermore, besides being independent from the albedo, the number of unifying equations does not increase linearly with respect to the available images. Indeed, instead of N (irradiance) equations, robustness is guaranteed by having a quadratic set of $\binom{N}{2}$ (unifying) equations. The numerical tool we used for solving such system of quasi-linear PDEs is based on a variational approach performing L^1 minimization. On the other hand, researching on speeding up our algorithm is foreseen in order to make the computation as close as possible to real-time 3D shape providers.

We remark that this work makes very challenging goals closer to be achieved. For instance, our approach could be extended so as to estimate parameters related to the material of the surface, lightings (uncalibrated PS), *etc*.

References

[1] J. F. Blinn. Models of light reflection for computer synthesized pictures. In *SIGGRAPH*, 1977. 2, 3

[2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, 3(1):1–122, 2011. 5

[3] A. M. Bruckstein. On shape from shading. *CVGIP*, 44(2):139–154, 1988. 2

[4] M. Chandraker, J. Bai, and R. Ramamoorthi. On differential photometric reconstruction for unknown, isotropic brdfs. *PAMI*, 35(12):2941–2955, 2013. 2, 5

[5] J. Clark. Active photometric stereo. In *CVPR*, 1992. 2

[6] T. Collins and A. Bartoli. 3d reconstruction in laparoscopy with close-range photometric stereo. In *MICCAI*. 2012. 2

[7] J. D. Durou, J.-F. Aujol, and F. Courteille. Integrating the normal field of a surface in the presence of discontinuities. In *EMMCVPR*, 2009. 2, 5

[8] J. D. Durou and F. Courteille. Integration of a Normal Field without Boundary Condition. In *ICCV Workshops*, 2007. 5

[9] C. Hernández, G. Vogiatzis, and R. Cipolla. Shadows in three-source photometric stereo. In *ECCV*, 2008. 2

[10] B. K. P. Horn. Obtaining shape from shading information. *The Psychology of Computer Vision*, pages 115–155, 1975. 1

[11] K. Hyeongwoo, J. Hailin, S. Hadap, and K. Inso. Specular reflection separation using dark channel prior. In *CVPR*, 2013. 3

[12] S. Ikehata and K. Aizawa. Photometric stereo using constrained bivariate regression for general isotropic surfaces. In *CVPR*, 2014. 5, 6

[13] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In *CVPR*, 2012. 2, 4, 5, 6, 7, 8

[14] Y. Iwahori, H. Sugie, and N. Ishii. Reconstructing shape from shading images under point light source illumination. In *ICPR*, 1990. 2

[15] S. Lee and M. Brady. Integrating stereo and photometric stereo to monitor the development of glaucoma. *IVC*, 9(1):39 – 44, 1991. 2

[16] S. Mallick, T. Zickler, D. Kriegman, and P. Belhumeur. Beyond lambert: reconstructing specular surfaces using color. In *CVPR*, 2005. 3

[17] W. Matusik, H. Pfister, M. Brand, and L. McMillan. A data-driven reflectance model. *ACM Transactions on Graphics*, 22(3):759–769, 2003. 6

[18] R. Mecca and M. Falcone. Uniqueness and approximation of a photometric shape-from-shading model. *SIAM J. Imag. Sci.*, 6(1):616–659, 2013. 2

[19] R. Mecca, E. Rodolà, and D. Cremers. Realistic photometric stereo using partial differential irradiance equation ratios. *Computers & Graphics*, 51:8–16, 2015. 4

[20] R. Mecca, A. Wetzler, A. Bruckstein, and R. Kimmel. Near field photometric stereo with point light sources. *SIAM J. Imag. Sci.*, 7(4):2732–2770, 2014. 2, 4

[21] T. Migita, S. Ogino, and T. Shakunaga. Direct bundle estimation for recovery of shape, reflectance property and light position. In *ECCV*. 2008. 2

[22] D. Miyazaki and K. Ikeuchi. Photometric stereo under unknown light sources using robust SVD with missing data. In *ICIP*, 2010. 4

[23] I. Moreno, M. Avendaño-Alejo, and R. I. Tzonchev. Designing light-emitting diode arrays for uniform near-field irradiance. *Applied Optics*, 45(10):2265–2272, 2006. 4

[24] A. Ngan, F. Durand, and W. Matusik. Experimental analysis of brdf models. In *EGSR*, 2005. 2

[25] T. Okatani and K. Deguchi. Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center. *CVIU*, 66(2):119–131, 1997. 2

[26] T. Papadhimitri and P. Favaro. A new perspective on uncalibrated photometric stereo. In *CVPR*, 2013. 2, 3

[27] T. Papadhimitri and P. Favaro. Uncalibrated near-light photometric stereo. In *BMVC*, 2014. 2, 7, 8

[28] B. T. Phong. Illumination for computer generated pictures. *Commun. ACM*, 18(6):311–317, 1975. 2

[29] E. Prados and O. D. Faugeras. Shape from shading: A well-posed problem? In *CVPR*, 2005. 2

[30] Y. Quéau and J.-D. Durou. Edge-preserving integration of a normal field: Weighted least squares, TV and L1 approaches. In *SSVM*, 2015. 2, 5

[31] Y. Quéau, F. Lauze, and J.-D. Durou. A L1-TV algorithm for robust perspective photometric stereo with spatially-varying lightings. In *SSVM*, 2015. 2, 4, 5

[32] T. Simchony, R. Chellappa, and M. Shao. Direct analytical methods for solving poisson equations in computer vision problems. *PAMI*, 12(5):435–446, 1990. 5

[33] R. Tan and K. Ikeuchi. Separating reflection components of textured surfaces using a single image. *PAMI*, 27(2):178–193, 2005. 3

[34] K. E. Torrance and E. M. Sparrow. Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.*, 57(9):1105–1112, 1967. 2

[35] Y. Wang, C. M. Lee, L. Cheong, and K. Toh. Practical matrix completion and corruption recovery using proximal alternating robust subspace minimization. *IJCV*, 111(3):315–344, 2015. 2, 4

[36] L. B. Wolff and E. Angelopoulou. 3-d stereo using photometric ratios. In *ECCV*, 1994. 2

[37] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):134–144, 1980. 1, 2, 6, 7, 8

[38] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *ACCV*, 2010. 2, 4