A Single-Lobe Photometric Stereo Approach for Heterogeneous Material[*]

Roberto Mecca[†], Yvain Quéau[‡], Fotios Logothetis[§], and Roberto Cipolla[§]

Abstract. Shape from shading with multiple light sources is an active research area, and a diverse range of approaches have been proposed in recent decades. However, devising a robust reconstruction technique still remains a challenging goal, as the image acquisition process is highly nonlinear. Recent Photometric Stereo variants rely on simplifying assumptions in order to make the problem solvable: light propagation is still commonly assumed to be uniform, and the Bidirectional Reflectance Distribution Function is assumed to be diffuse, with limited interest for specular materials. In this work, we introduce a well-posed formulation based on partial differential equations (PDEs) for a unified reflectance function that can model both diffuse and specular reflections. We base our derivation on ratio of images, which makes the model independent from photometric invariants and yields a well-posed differential problem based on a system of quasi-linear PDEs with discontinuous coefficients. In addition, we directly solve a differential problem for the unknown depth, thus avoiding the intermediate step of approximating the normal field. A variational approach is presented ensuring robustness to noise and outliers (such as black shadows), and this is confirmed with a wide range of experiments on both synthetic and real data, where we compare favorably to the state of the art.

1. Introduction. The problem of reconstructing a static object using multiple images taken from a fixed point of view but with varying illumination is well known in the computer vision community as the Photometric Stereo (PS) problem. Initial PS methods extended the shape from shading problem (SfS) [8, 21, 22, 23, 24, 32] (which attempts to recover three-dimensional (3D) structure from a single image) by adding supplementary information contained in additional images. In contrast with SfS, PS can be unambiguously solved with three images even when the albedo is unknown [62], and even with two images when the albedo is known [52]. However, most classical PS approaches have relied on important simplifications, such as

(a) orthographic viewing geometry;
(b) uniform directional illumination;

(c) diffuse reflectance;

(d) absence of shadows and discontinuities;

(e) monochromatic lighting and reflectance.

Although those assumptions allow one to prove the well-posedness of the PS problem, they constrained this 3D shape recovery technique to be employable only under a very controlled setup.

Contribution. With the aim to extend and derive a well-posed mathematical formulation for the model presented in [37], in this work we propose a method capable of performing the PS technique 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. Taking inspiration from [40], we relax all the above mentioned assumptions from (a) to (e) by proposing the following:

- an irradiance equation capable of handling both diffuse and specular reflectance, as well as pointwise and directional light sources, and perspective viewing geometry;
- a *unifying equation* for the PS problem based on partial differential irradiance equation ratios;
- an l^1-norm-based variational method for solving the set of fundamental equations that ensures robustness to shadows and depth discontinuities;
- direct depth recovery, without integrating the normals, thanks to a differential parameterization of the surface.

To the best of our knowledge, this approach is a step forward with respect to the state of the art since it relaxes the aforementioned assumptions to make PS more realistic.

1.1. Related works.
After more than thirty years of research, the scientific community has not yet found a definitive solution for the general PS problem. However, several attempts have been made in order to relax the assumptions (a)–(e) in order to deal with realistic irradiance equations.

Camera modeling. Perspective viewing geometry for photometric 3D reconstruction has been introduced by Horn [21], then used by Bruckstein [8], and later developed in equivalent manners by Tankus and Kiryati [55], Prados and Faugeras [49], and Durou and his coworkers [6, 15]. Recently, Papadhimitri and Favaro [46] presented a new perspective parameterization to solve the PS problem without knowledge of the light directions, i.e., uncalibrated PS (UPS).

This significant feature allowed PS to be exploited for such important applications as shape recovery from endoscopic images [13, 45]. In this particular case the nonuniform 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 and directional lighting assumptions, light propagation models have been extended to consider nearby point light sources spreading light radially. Initially this idea was proposed by Iwahori, Sugie, and Ishii [29], and later by Clark [12]. More recently, nearby point light sources have also been considered in alternating optimization procedures ("bundle adjustments") for structure-from-motion [41] and UPS problems [47].

However, all those works modeling nonuniform light propagation by adopting point light source parameterization assume diffuse reflection, which limits the applicability of the PS technique.

Reflectance. Most of the research done till now for the PS technique has assumed purely diffuse reflectance as the Bidirectional Reflectance Distribution Function (BRDF). Unlike (a) and (b), shape recovery from specular shading still remains a challenging goal since most of the common materials provide specular highlights that prevent reasonable reconstructions by the PS technique.

Regarding shading models for specular highlights, several dedicated irradiance equations have been presented so far. First, Torrance and Sparrow [56] presented a physical model based on radiometry principles. Later, Phong [48] showed an empirical model which basically extended the cosine law, making it depend also on the viewer direction. The Blinn–Phong shading model [4] extended further the previous one by eliminating some limitation in the analytical formulation. Then Cook and Torrance [14] provided a well-known specular model based on a strongly nonlinear physical theory. Other important models for specular BRDFs can be found in [33], and interesting comparisons among some of them have been performed in [44].

Instead of simplifying the PS problem by removing specularity, other works dealt with the images as they are, having both specular and diffuse components. Nayar, Ikeuchi, and Kanade [43] assumed hybrid surfaces by summing diffuse and specular components using the Beckmann and Spizzichino [3] reflection model, allowing them to locally reconstruct the shape of the object. Ikeuchi [28] faced the PS problem with specular reflectance by introducing a smoothness prior on the surface, which was shown to be realistic for several industrial applications, though this approach is limited to a typology of surfaces effectively holding this smoothness constraint. Wolff [60] proposed a reflectance model for combined diffuse and specular reflections from dielectric materials. Since the modeling involved purely physical parameters in terms of the Fresnel coefficient and the Dirac delta function, the results achieved can be considered mostly theoretical rather than actually applicable for shape reconstruction. Goldman et al. [17] described the irradiance equation as a weighted sum of fundamental diffuse and specular materials, and they solved it by iteratively adding weights inside a minimization process in order to include specular reflection. Yang and Ahuja [65] introduced a PS method for estimating diffuse reflection and surface normals from color images. Using a dichromatic reflection model, they considered surface chromaticity as a matching invariant for PS, assuming that at least one of the color images was diffuse. Chung and Jia [11] adopted a two-step optimization approach to estimate the normals and all reflectance parameters. In the first step, the visual information from shadows was used to estimate the normals at some surface points. These recovered normals played a key role in estimating the global reflectance parameters according to the Ward BRDF [59]. In the second step, with the reduced number of unknowns, a robust iterative optimization process was applied estimating the surface normals and the diffuse albedo for all surface points.

Although very sophisticated approaches modeling specular highlights have been presented, several advances have also been made in creating more robust algorithms while considering highlights as missing parts that additionally include shadows, saturations, self-occlusions, etc.

Outliers. Regarding robustness to outliers, most recent works proposed sparsity-enhancing estimators. For instance, the SVD-based approach on which the UPS relies [18] can be enhanced to deal with shadows, which were a priori detected and treated as missing data by Miyazaki and Ikeuchi [42]. An alternative is to preprocess the images according to a low-rank

constraint, as suggested by Wu et al. [63]. Such an approach, relying on convex relaxation of an l^0-norm-based minimization problem, was recently improved by Wang et al. [58], who proposed dealing with the nonconvex nature of the problem using a proximal strategy. A similar approach, but more specifically dedicated to perspective UPS, was considered in [50], extending the approach by Ikehata et al. [27], who 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 [16].

Clearly, such reinforcement of the PS technique is mandatory for real-world applications. Yet robustness can also be improved by considering extra information contained in RGB images, rather than relying solely on luminance.

Color. In most PS methods, the images are assumed to have gray levels. Color images can be used to separate the diffuse part from the specular one, as in the methods presented above, but usually only the one-channel diffuse part is used in the reconstruction. Clearly, not all information is exploited when proceeding in this manner. On the other hand, methods dealing with all color channels assume either that the material is monochromatic and the lightings are colored [7] or that the inverse is the case [2].

All these works brought the PS problem further on one or the other specific point, but they all require estimating both the shape and the albedo. Another strategy involves using image ratios that make it possible to eliminate the albedo from the unknowns.

Image ratios. Considering *image ratios* yields photometric invariants [19, 34, 61]. Such invariants have been successfully used in image matching under varying illumination [30], in robust approaches to PS [19, 64], or in general PS settings with unknown [10] or pointwise [40, 38, 51] lightings. Image ratios also play a key role in linearizing the differential approaches to PS based on nonlinear PDEs.

Differential approaches to PS. Recently, new PS models based on nonlinear PDEs have attracted increasing interest [9, 10, 36, 39, 40, 50, 54]. These approaches model the irradiance equations via nonlinear PDEs. This system of nonlinear PDEs can be solved using calculus of variations [50] or proximal methods [20]. However, such methods remain computationally very intensive: it is simpler to eliminate the nonlinearity, through image ratios, as done in [36, 39, 40, 51, 54]. Chandraker, Bai, and Ramamoorthi [10] considered more general irradiance equations with unknown light sources and computed the photometric invariants describing the surface through its isocontours [1]. However, in all these works, the shape reconstruction process required additional initial or boundary information, which are not required in the linear least-squares variational approach proposed recently in [51, 54]. Mecca and coauthors [36, 39, 40] used specific irradiance equations for diffuse surfaces only, proving uniqueness of the solution by characteristic strip expansion and assuming known light information (i.e., direction or position). However, the approach adopting such an expansion performs unfavorably under noise. This is due to the fact that for each step of the expansion, only a single pair of images is taken into account despite a larger amount of images being available. In addition, depth information is only propagated through a single direction, in contrast to robust variational solvers that simultaneously estimate the depth at all pixels [20, 50, 51, 54].

In the same framework, a mathematical study of a purely specular reflectance using the Blinn–Phong shading model has been proposed by Tozza et al. [57] employing an additional

shading model practically separated from the diffuse one. In fact, although the approach deals with both diffuse and specular reflectance, the actual differential models are applied separately after a preliminary segmentation of the shape domain [25]. Furthermore, simplifying assumptions related to the orthographic viewing geometry and directional lighting are considered in order to deal with linear PDEs. Relaxing those hypotheses, Mecca, Rodolà, and Cremers [38] proposed two differential models dealing separately with diffuse and specular reflectance while introducing the parameterization of more realistic physical effects such as perspective viewing and light attenuation. Even if they did not provide any mathematical proof regarding the well-posedness of the differential problems, the experimental part is very interesting. In fact, they showed that the instability of the characteristic strip expansion due to noise increases significantly while reconstructing specular shapes.

In order to deal with these drastic limitations, in this work we unify both diffuse and specular reflectances by modifying the Blinn–Phong model into a single-lobe approach. Furthermore, in order to make the solution of the differential formulation robust, we propose a variational solver and compare it with the method based on the characteristic strip expansion to confirm the theorized gain in robustness.

This paper is organized as follows: The unified irradiance equation is presented in section 2. The mathematical derivation of the new equation for the PS problem is provided in section 3. Two numerical solvers are presented in section 4. The first shows the well-posedness of a fast-marching-based approach where the depth is recovered. The second introduces a variational solver over a larger amount of unknowns, including one of the reflectance parameters. Experimental results are presented in section 5.

2. Image formation model. With the aim being to formulate a mathematical model as general as possible, we have to deal with two fundamental features of the image formation process regarding the modeling of the pinhole camera and 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 [46]. This allows one to easily extend the perspective viewing projection to the orthographic one just by making the focal length tend to infinity. The projection takes a 3D point, described by the real-world coordinates $(\xi, \eta, \zeta) \in \Sigma$, into the image plane $\overline{\Omega} = \Omega \cup \partial\Omega$ according to

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

where z is the *depth*. According to this parameterization, the direction of the outgoing normal to the surface is given by

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

where the derivatives of z are considered in the image coordinates and are denoted in the following: $\nabla z(x,y) = \left(\frac{\partial z}{\partial x}(x,y), \frac{\partial z}{\partial y}(x,y) \right)$. Let us denote the normalized vector by

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

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

2.2. Reflectance model. In order to formulate the PS problem in a general manner, we consider the irradiance equation given by the following inner product for the ith light source:

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

where ρ, a_i, and c are positive scalar functions representing, respectively, the *albedo*, the attenuation of light, and the shininess. The albedo ρ and the shininess c are independent of the source number, since they are only related to the surface reflectance. In this work, we make no hypothesis on the albedo and the shininess. The vector function $\mathbf{h}_i = (h_i^1, h_i^2, h_i^3)$ combines information about the viewer direction $\mathbf{v} = \mathbf{v}(x,y,z)$ and the ith light source $\mathbf{l}_i = \mathbf{l}_i(x,y,z)$. With our aim being to unify diffuse and specular components parameterized by a single-lobe irradiance equation, we choose $\mathbf{h}_i$ to be as follows:

$$(5) \qquad \mathbf{h}_i = \overline{\mathbf{l}}_i + \min\left\{ 1, \frac{|1-c|}{\varepsilon} \right\} \overline{\mathbf{v}}.$$

Here ϵ is a parameter describing the material of the surface, allowing combination of diffuse and specular reflectances by describing the continuous transition state between diffuse and specular component.

Besides being independent from the camera projection model, thanks to the parameterization (2), this new general model can be used to deal with surfaces made by heterogeneous materials depending on c and ϵ:

- by choosing $c \equiv 1$, we obtain the purely diffuse (Lambertian) reflection model;
- setting $0 < c \leq 1 - \epsilon$ leads to purely specular reflectance (Blinn–Phong type), where the size of the specular lobes depends on the value of c;
- the choice $1 - \epsilon < c < 1$ provides a transition state between diffuse and specular component;
- spatially varying reflectances can be modeled by using the space dependency of c.

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

Moreover, by manipulating the definitions of a_i and $\mathbf{l}_i$, this new model allows one to deal with well-known light configurations.

Uniform lighting. When $a_i(x,y,z) \equiv \phi_i \in \mathbb{R}$ 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 light attenuations due to distance and angular factors (anisotropy) [40], defined, respectively, by

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

and

$$(7) \qquad a_{ai}(x, y, z) = g(\bar{\mathbf{l}}_i(x,y,z) \cdot \bar{\mathbf{p}}_i)$$

for some anisotropy function g depending on the source type, and where $\bar{\mathbf{p}}_i$ is the principal lighting direction of the source, i.e., its orientation. This function g is usually calibrated by manufacturers under the form of luminous intensity diagrams. A good approximation holding for a lot of commercial LEDs is the imperfect Lambertian primary source model, which reads

$$(8) \qquad a_{ai}(x, y, z) = (\bar{\mathbf{l}}_i(x,y,z) \cdot \bar{\mathbf{p}}_i)^{\mu}.$$

Lambertian ("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. Putting it all together, this yields the following general attenuation factor:

$$(9) \qquad a_i(x, y, z) = \phi_i \frac{a_{ai}(x, y, z)}{|\mathbf{l}_i(x, y, z)|^2}.$$

Despite being empirical, the reflectance model (4) generalizes many previous models used in PS. In the literature on PS, all these specific situations require a different machinery. However, we will show in the following that the 3D reconstruction problem can be reformulated by the same simple quasi-linear PDE, handling the general viewing geometry, reflectance, and lighting models previously described. Thus, we provide a unified approach to the PS problems in many real-world scenarios.

The purpose of using a unique irradiance equation for diffuse and specular components is coherent with having a unique image as input data including both reflections. Since diffuse and specular component separation is not a reliable procedure to recover a shape made up of heterogeneous materials, we consider a unified irradiance equation with the aim of extracting geometric information from both diffuse and, most importantly, specular highlights of the observed object.

3. The unifying equation. As already shown in [40] the ratio of irradiance equations as (4) seen as partial differential irradiance equations simplifies the PS problem since it becomes independent of the albedo and the 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 ratio. With the aim being to extend such a methodology to the more general and realistic irradiance equations (4), we divide those coming from the ith and jth light sources, leading to (10) below. From now on, we neglect the dependencies to improve the readability of the equations.

$$(10) \qquad \underbrace{\frac{(I_i)^c}{(a_i)^c \, \bar{\mathbf{h}}_i \cdot \mathbf{n}}}_{\text{Equation for } I_i} = \frac{\rho^c}{|\mathbf{n}|} = \overbrace{\frac{(I_j)^c}{(a_j)^c \, \bar{\mathbf{h}}_j \cdot \mathbf{n}}}^{\text{Equation for } I_j}.$$

By considering the first and the last part of the chain of equalities (10), after some algebra we obtain the following quasi-linear PDE:

$$(11) \qquad \mathbf{b}_{ij}(x,y,z) \cdot \nabla z(x,y) = s_{ij}(x,y,z),$$

where

$$
\mathbf{b}_{ij} = \Big((a_j I_i)^c \Big(\overline{h}_j^1 - \frac{x}{f} \overline{h}_j^3 \Big) - (a_i I_j)^c \Big(\overline{h}_i^1 - \frac{x}{f} \overline{h}_i^3 \Big),
$$
$$(12) \qquad (a_j I_i)^c \Big(\overline{h}_j^2 - \frac{y}{f} \overline{h}_j^3 \Big) - (a_i I_j)^c \Big(\overline{h}_i^2 - \frac{y}{f} \overline{h}_i^3 \Big) \Big)$$

and

$$(13) \qquad s_{ij} = \frac{f+z}{f} \Big((a_j I_i)^c \, \overline{h}_j^3 - (a_i I_j)^c \, \overline{h}_i^3 \Big),$$

a vector field and a scalar field, respectively.

We will call (11) the *unifying equation* for the PS problem, since it elegantly describes the interaction between two irradiance scenarios depicted in I_i and I_j, regardless of what the models are for the camera projection, the surface reflectance, and the type of lighting.

As a preliminary and important result, we initially consider the unifying equation (10) enhanced by the Dirichlet boundary condition as follows:

$$(14) \qquad \begin{cases} \mathbf{b}_{ij}(x,y,z) \cdot \nabla z(x,y) = s_{ij}(x,y,z), & (x,y) \in \Omega, \\ z(x,y) = g(x,y), & (x,y) \in \partial\Omega, \end{cases}$$

showing that this problem is well-posed in the next section.

3.2. Well-posedness by using the unifying equation. In this part, we focus on proving the well-posedness of the quasi-linear differential problem (14), which is based on a single pair of images. We employ the characteristic strip expansion method by drawing inspiration from [36, 39]. Let us remark that the fundamental improvement of our work with respect to [40] does not prevent showing the existence and uniqueness of a weak, Lipschitz continuous solution. Let us start by proving the following result.

Lemma 1. *The vector field* $\mathbf{b}_{ij}(x,y,z)$ *does not vanish in* Ω, *that is,*

$$(15) \qquad |\mathbf{b}_{ij}(x,y,z)| \neq 0 \quad \forall (x,y) \in \Omega, \quad \forall z \in C^1(\overline{\Omega}).$$

Proof. We prove it by contradiction in the most general case by assuming that there exists a point $(\widetilde{x}, \widetilde{y}) \in \Omega$ and a function $z \in C^1(\overline{\Omega})$ such that

$$(16) \qquad (a_j I_i)^c \Big(\overline{h}_j^1 - \frac{\widetilde{x}}{f} \overline{h}_j^3 \Big) - (a_i I_j)^c \Big(\overline{h}_i^1 - \frac{\widetilde{x}}{f} \overline{h}_i^3 \Big) = 0,$$

$$(17) \qquad (a_j I_i)^c \Big(\overline{h}_j^2 - \frac{\widetilde{y}}{f} \overline{h}_j^3 \Big) - (a_i I_j)^c \Big(\overline{h}_i^2 - \frac{\widetilde{y}}{f} \overline{h}_i^3 \Big) = 0,$$

where all the functions are evaluated at point $(\widetilde{x}, \widetilde{y})$, i.e., $I_{i,j} = I_{i,j}(\widetilde{x}, \widetilde{y})$, $\overline{h}_{i,j}^{1,2,3} = \overline{h}_{i,j}^{1,2,3}(\widetilde{x}, \widetilde{y}, \widetilde{z})$ and $z = z(\widetilde{x}, \widetilde{y})$.

In order to simplify the notation, we remove the dependency on $(\widetilde{x}, \widetilde{y})$ and consider fixed values for the involved quantities.

By considering the image irradiance equation (4) for (16) and (17), we get

$$(18) \qquad (a_i a_j \rho)^c \overline{\mathbf{n}} \cdot \overline{\mathbf{h}}_i \left(\overline{h}_j^1 - \frac{x}{f} \overline{h}_j^3 \right) - (a_i a_j \rho)^c \overline{\mathbf{n}} \cdot \overline{\mathbf{h}}_j \left(\overline{h}_i^1 - \frac{x}{f} \overline{h}_i^3 \right) = 0,$$

$$(19) \qquad (a_i a_j \rho)^c \overline{\mathbf{n}} \cdot \overline{\mathbf{h}}_i \left(\overline{h}_j^2 - \frac{y}{f} \overline{h}_j^3 \right) - (a_i a_j \rho)^c \overline{\mathbf{n}} \cdot \overline{\mathbf{h}}_j \left(\overline{h}_i^2 - \frac{y}{f} \overline{h}_i^3 \right) = 0,$$

which can be simplified as follows:

$$(20) \qquad \mathbf{n} \cdot \left(\mathbf{h}_i \left(h_j^1 - \frac{x}{f} h_j^3 \right) - \mathbf{h}_j \left(h_i^1 - \frac{x}{f} h_i^3 \right) \right) = 0,$$

$$(21) \qquad \mathbf{n} \cdot \left(\mathbf{h}_i \left(h_j^2 - \frac{y}{f} h_j^3 \right) - \mathbf{h}_j \left(h_i^2 - \frac{y}{f} h_i^3 \right) \right) = 0.$$

If we solve the system made by (20) and (21) with respect to z_x and z_y (in $\mathbf{n}$ as in (2)), we get the following solution:

$$(22) \qquad z_x^* = \frac{(h_i^2 h_j^3 - h_i^3 h_j^2)(f + z)}{f h_i^2 h_j^1 - f h_i^1 h_j^2 + h_i^3 h_j^2 x - h_i^2 h_j^3 x - h_i^3 h_j^1 y + h_i^1 h_j^3 y},$$

$$(23) \qquad z_y^* = -\frac{(h_i^1 h_j^3 - h_i^3 h_j^1)(f + z)}{f h_i^2 h_j^1 - f h_i^1 h_j^2 + h_i^3 h_j^2 x - h_i^2 h_j^3 x - h_i^3 h_j^1 y + h_i^1 h_j^3 y}.$$

Once we compute the image values substituting z_x^* and z_y^* into (4), we get $I_i = 0$ and $I_j = 0$, which clearly contradicts the assumption of having nonnull image data to reconstruct the shape. ∎

In this theoretical study, we will assume that there is no such missing data. We will later propose a variational approach to handling this issue.
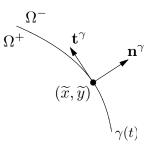
Now, we continue the proof of the last fact, which completes the proof for the well-posedness. Let us introduce this result by remarking that if $\gamma \in \Omega$ is a regular curve where ∇z is discontinuous (jump discontinuity, since we are assuming z as Lipschitz continuous), then the images are discontinuous on γ; see Figure 1. Then we assume that γ is given as well. Furthermore, we have that γ is the curve of discontinuity also for $\mathbf{b}$ and s since they depend explicitly on the images.

Theorem 1. *Let $\gamma(t)$ be a regular curve of discontinuity for the image functions. Let $(\widetilde{x}, \widetilde{y}) \in \gamma(t)$, and let $\mathbf{n}^\gamma(\widetilde{x}, \widetilde{y})$ be the outgoing normal with respect to the set Ω^+ (located on the right of $\gamma(t)$); then the following inequality holds:*

$$(24) \qquad \left[\lim_{\substack{(x,y) \to (\widetilde{x}, \widetilde{y}) \\ (x,y) \in \Omega^+}} \mathbf{b}_{ij}(x, y, z(x, y)) \cdot \mathbf{n}^\gamma(\widetilde{x}, \widetilde{y}) \right] \left[\lim_{\substack{(x,y) \to (\widetilde{x}, \widetilde{y}) \\ (x,y) \in \Omega^-}} \mathbf{b}_{ij}(x, y, z(x, y)) \cdot \mathbf{n}^\gamma(\widetilde{x}, \widetilde{y}) \right] \geq 0.$$

Proof. We start by considering the following quantities for both images I_i and I_j:

$$(25) \qquad \lim_{\substack{(x,y) \to (\widetilde{x}, \widetilde{y}) \\ (x,y) \in \Omega^+}} I_{i,j}(x, y) = I_{i,j}^+, \qquad \lim_{\substack{(x,y) \to (\widetilde{x}, \widetilde{y}) \\ (x,y) \in \Omega^-}} I_{i,j}(x, y) = I_{i,j}^-,$$

Figure 1. *The two subsets Ω^+ and Ω^- separated by the curve $\gamma(t)$ contain the well-defined classical solution. The vectors $\mathbf{n}^\gamma$ and $\mathbf{t}^\gamma$ are, respectively, the normal and the tangent vectors of $\gamma(t)$.*

which define the vector field $\mathbf{b}$ as follows:

$$\lim_{\substack{(x,y)\to(\widetilde{x},\widetilde{y})\\(x,y)\in\Omega^+}} \mathbf{b}_{ij}(x,y,z) = \left((a_j I_i^+)^c\left(\overline{h}_j^1 - \frac{\widetilde{x}}{f}\overline{h}_j^3\right) - (a_i I_j^+)^c\left(\overline{h}_i^1 - \frac{\widetilde{x}}{f}\overline{h}_i^3\right),\right.$$
$$\left.(a_j I_i^+)^c\left(\overline{h}_j^2 - \frac{\widetilde{y}}{f}\overline{h}_j^3\right) - (a_i I_j^+)^c\left(\overline{h}_i^2 - \frac{\widetilde{y}}{f}\overline{h}_i^3\right)\right) = (b_1^+, b_2^+),$$

$$\lim_{\substack{(x,y)\to(\widetilde{x},\widetilde{y})\\(x,y)\in\Omega^-}} \mathbf{b}_{ij}(x,y,z) = \left((a_j I_i^-)^c\left(\overline{h}_j^1 - \frac{\widetilde{x}}{f}\overline{h}_j^3\right) - (a_i I_j^-)^c\left(\overline{h}_i^1 - \frac{\widetilde{x}}{f}\overline{h}_i^3\right),\right.$$
$$\left.(a_j I_i^-)^c\left(\overline{h}_j^2 - \frac{\widetilde{y}}{f}\overline{h}_j^3\right) - (a_i I_j^-)^c\left(\overline{h}_i^2 - \frac{\widetilde{y}}{f}\overline{h}_i^3\right)\right) = (b_1^-, b_2^-).$$

Now, after calling $\mathbf{n}^\gamma$ at point $(\widetilde{x},\widetilde{y})$ as (n_1^γ, n_2^γ) and the components of the vector field $\mathbf{b}_{ij}(x,y,z)$ as (b_1^+, b_2^+) and (b_1^-, b_2^-) in each limit case, we write (24) as follows:

$$(26) \qquad b_1^+ b_1^- (n_1^\gamma)^2 + b_2^+ b_2^- (n_2^\gamma)^2 + (b_1^+ b_2^- + b_2^+ b_1^-)n_1^\gamma n_2^\gamma \geq 0.$$

Then, after substituting (4) into the previous inequality, we have

$$\left[(a_i a_j \rho)^c \overline{\mathbf{n}}^+ \cdot \overline{\mathbf{h}}_i\left(\overline{h}_j^1 - \frac{x}{f}\overline{h}_j^3\right) - (a_i a_j \rho)^c \overline{\mathbf{n}}^+ \cdot \overline{\mathbf{h}}_j\left(\overline{h}_i^1 - \frac{x}{f}\overline{h}_i^3\right)\right]$$
$$\left[(a_i a_j \rho)^c \overline{\mathbf{n}}^- \cdot \overline{\mathbf{h}}_i\left(\overline{h}_j^1 - \frac{x}{f}\overline{h}_j^3\right) - (a_i a_j \rho)^c \overline{\mathbf{n}}^- \cdot \overline{\mathbf{h}}_j\left(\overline{h}_i^1 - \frac{x}{f}\overline{h}_i^3\right)\right](n_1^\gamma)^2$$
$$+\left[(a_i a_j \rho)^c \overline{\mathbf{n}}^+ \cdot \overline{\mathbf{h}}_i\left(\overline{h}_j^2 - \frac{y}{f}\overline{h}_j^3\right) - (a_i a_j \rho)^c \overline{\mathbf{n}}^+ \cdot \overline{\mathbf{h}}_j\left(\overline{h}_i^2 - \frac{y}{f}\overline{h}_i^3\right)\right]$$
$$\left[(a_i a_j \rho)^c \overline{\mathbf{n}}^- \cdot \overline{\mathbf{h}}_i\left(\overline{h}_j^2 - \frac{y}{f}\overline{h}_j^3\right) - (a_i a_j \rho)^c \overline{\mathbf{n}}^- \cdot \overline{\mathbf{h}}_j\left(\overline{h}_i^2 - \frac{y}{f}\overline{h}_i^3\right)\right](n_2^\gamma)^2$$
$$+\left(\left[(a_i a_j \rho)^c \overline{\mathbf{n}}^+ \cdot \overline{\mathbf{h}}_i\left(\overline{h}_j^1 - \frac{x}{f}\overline{h}_j^3\right) - (a_i a_j \rho)^c \overline{\mathbf{n}}^+ \cdot \overline{\mathbf{h}}_j\left(\overline{h}_i^1 - \frac{x}{f}\overline{h}_i^3\right)\right]\right.$$
$$\left[(a_i a_j \rho)^c \overline{\mathbf{n}}^- \cdot \overline{\mathbf{h}}_i\left(\overline{h}_j^2 - \frac{y}{f}\overline{h}_j^3\right) - (a_i a_j \rho)^c \overline{\mathbf{n}}^- \cdot \overline{\mathbf{h}}_j\left(\overline{h}_i^2 - \frac{y}{f}\overline{h}_i^3\right)\right]$$
$$+\left[(a_i a_j \rho)^c \overline{\mathbf{n}}^+ \cdot \overline{\mathbf{h}}_i\left(\overline{h}_j^2 - \frac{y}{f}\overline{h}_j^3\right) - (a_i a_j \rho)^c \overline{\mathbf{n}}^+ \cdot \overline{\mathbf{h}}_j\left(\overline{h}_i^2 - \frac{y}{f}\overline{h}_i^3\right)\right]$$

$$\left[(a_i a_j \rho)^c \overline{\mathbf{n}}^- \cdot \overline{\mathbf{h}}_i \left(\overline{h}_j^1 - \frac{x}{f} \overline{h}_j^3 \right) - (a_i a_j \rho)^c \overline{\mathbf{n}}^- \cdot \overline{\mathbf{h}}_j \left(\overline{h}_i^1 - \frac{x}{f} \overline{h}_i^3 \right) \right] \right) n_1^\gamma n_2^\gamma \geq 0,$$

which simplifies to the following inequality:

(27)

$$\left[\mathbf{n}^+ \cdot \mathbf{h}_i \left(h_j^1 - \frac{x}{f} h_j^3 \right) - \mathbf{n}^+ \cdot \mathbf{h}_j \left(h_i^1 - \frac{x}{f} h_i^3 \right) \right] \left[\mathbf{n}^- \cdot \mathbf{h}_i \left(h_j^1 - \frac{x}{f} h_j^3 \right) - \mathbf{n}^- \cdot \mathbf{h}_j \left(h_i^1 - \frac{x}{f} h_i^3 \right) \right] (n_1^\gamma)^2$$

$$+ \left[\mathbf{n}^+ \cdot \mathbf{h}_i \left(h_j^2 - \frac{y}{f} h_j^3 \right) - \mathbf{n}^+ \cdot \mathbf{h}_j \left(h_i^2 - \frac{y}{f} h_i^3 \right) \right] \left[\mathbf{n}^- \cdot \mathbf{h}_i \left(h_j^2 - \frac{y}{f} h_j^3 \right) - \mathbf{n}^- \cdot \mathbf{h}_j \left(h_i^2 - \frac{y}{f} h_i^3 \right) \right] (n_2^\gamma)^2$$

$$+ \left(\left[\mathbf{n}^+ \cdot \mathbf{h}_i \left(h_j^1 - \frac{x}{f} h_j^3 \right) - \mathbf{n}^+ \cdot \mathbf{h}_j \left(h_i^1 - \frac{x}{f} h_i^3 \right) \right] \left[\mathbf{n}^- \cdot \mathbf{h}_i \left(h_j^2 - \frac{y}{f} h_j^3 \right) - \mathbf{n}^- \cdot \mathbf{h}_j \left(h_i^2 - \frac{y}{f} h_i^3 \right) \right] \right.$$

$$+ \left. \left[\mathbf{n}^+ \cdot \mathbf{h}_i \left(h_j^2 - \frac{y}{f} h_j^3 \right) - \mathbf{n}^+ \cdot \mathbf{h}_j \left(h_i^2 - \frac{y}{f} h_i^3 \right) \right] \left[\mathbf{n}^- \cdot \mathbf{h}_i \left(h_j^1 - \frac{x}{f} h_j^3 \right) - \mathbf{n}^- \cdot \mathbf{h}_j \left(h_i^1 - \frac{x}{f} h_i^3 \right) \right] \right) n_1^\gamma n_2^\gamma \geq 0.$$

In order to evaluate the difference between the two image values, we define the following quantities:

$$(28) \qquad i_{i,j}^+ = \mathbf{n}^+ \cdot \mathbf{h}_{i,j} = \left(\nabla^+ z, -\frac{f+z}{f} - \frac{(x,y) \cdot \nabla^+ z}{f} \right) \cdot \mathbf{h}_{i,j},$$

$$(29) \qquad i_{i,j}^- = \mathbf{n}^- \cdot \mathbf{h}_{i,j} = \left(\nabla^- z, -\frac{f+z}{f} - \frac{(x,y) \cdot \nabla^- z}{f} \right) \cdot \mathbf{h}_{i,j},$$

where $\nabla^+ z = (z_x^+, z_y^+)$ and $\nabla^- z = (z_x^-, z_y^-)$ are the limits of the gradient of z to $(\widetilde{x}, \widetilde{y})$ from Ω^+ and Ω^-, respectively. Now, we preliminarily substitute the first part of (28) and (29) into (27), yielding the inequality

(30)

$$\left[i_i^+ \left(h_j^1 - \frac{x}{f} h_j^3 \right) - i_j^+ \left(h_i^1 - \frac{x}{f} h_i^3 \right) \right] \left[i_i^- \left(h_j^1 - \frac{x}{f} h_j^3 \right) - i_j^- \left(h_i^1 - \frac{x}{f} h_i^3 \right) \right] (n_1^\gamma)^2$$

$$+ \left[i_i^+ \left(h_j^2 - \frac{y}{f} h_j^3 \right) - i_j^+ \left(h_i^2 - \frac{y}{f} h_i^3 \right) \right] \left[i_i^- \left(h_j^2 - \frac{y}{f} h_j^3 \right) - i_j^- \left(h_i^2 - \frac{y}{f} h_i^3 \right) \right] (n_2^\gamma)^2$$

$$+ \left(\left[i_i^+ \left(h_j^1 - \frac{x}{f} h_j^3 \right) - i_j^+ \left(h_i^1 - \frac{x}{f} h_i^3 \right) \right] \left[i_i^- \left(h_j^2 - \frac{y}{f} h_j^3 \right) - i_j^- \left(h_i^2 - \frac{y}{f} h_i^3 \right) \right] \right.$$

$$+ \left. \left[i_i^+ \left(h_j^2 - \frac{y}{f} h_j^3 \right) - i_j^+ \left(h_i^2 - \frac{y}{f} h_i^3 \right) \right] \left[i_i^- \left(h_j^1 - \frac{x}{f} h_j^3 \right) - i_j^- \left(h_i^1 - \frac{x}{f} h_i^3 \right) \right] \right) n_1^\gamma n_2^\gamma \geq 0.$$

Finally, once we write one image value i^- (for example) with respect to i^+ as

$$i_{i,j}^- = i_{i,j}^+ + \mathbf{n}^- \cdot \mathbf{h}_{i,j} - \mathbf{n}^+ \cdot \mathbf{h}_{i,j} = i_{i,j}^+ + \left(\nabla^- z, -\frac{f+z}{f} - \frac{(x,y) \cdot \nabla^- z}{f} \right) \cdot \mathbf{h}_{i,j}$$

$$- \left(\nabla^+ z, -\frac{f+z}{f} - \frac{(x,y) \cdot \nabla^+ z}{f} \right) \cdot \mathbf{h}_{i,j} = i_{i,j}^+ + \left(\boldsymbol{\delta}, -\frac{(\widetilde{x}, \widetilde{y}) \cdot \boldsymbol{\delta}}{f} \right) \cdot \mathbf{h}_{i,j},$$

where $\boldsymbol{\delta} = (\delta_1, \delta_2) = (z_x^- - z_x^+, z_y^- - z_y^+)$, we get the following inequality:

$$(31) \qquad \left[f\big(i_i^+(h_j^1 n_1^\gamma + h_j^2 n_2^\gamma) - i_j^+(h_i^1 n_1^\gamma + h_i^2 n_2^\gamma)\big) + (h_j^3 i_i^+ - h_i^3 i_j^+)(n_1^\gamma \widetilde{x} + n_2^\gamma \widetilde{y}) \right]$$

$$(32) \qquad \left[f\big(i_i^+(h_j^1 n_1^\gamma + h_j^2 n_2^\gamma) - i_j^+(h_i^1 n_1^\gamma + h_i^2 n_2^\gamma)\big) + (h_j^3 i_i^+ - h_i^3 i_j^+)(n_1^\gamma \widetilde{x} + n_2^\gamma \widetilde{y}) \right.$$

$$(33) \qquad \left. + (\delta_2 n_1^\gamma - \delta_1 n_2^\gamma)\big(f^2(h_j^1 h_i^2 - h_i^1 h_j^2) + f(h_i^2 h_j^3 \widetilde{x} - h_j^2 h_i^3 \widetilde{x} + h_j^1 h_i^3 \widetilde{y} - h_i^1 h_j^3 \widetilde{y})\big) \right] \geq 0.$$

By noticing that

$$(34) \quad \delta_2 n_1^\gamma - \delta_1 n_2^\gamma = n_1^\gamma\big(z_y^- - z_y^+\big) - n_2^\gamma\big(z_x^- - z_x^+\big) = -z_x^- n_2^\gamma + z_y^- n_1^\gamma - \big(-z_x^+ n_2^\gamma + z_y^+ n_1^\gamma\big)$$
$$= \lim_{\substack{(x,y)\to(\widetilde{x},\widetilde{y}) \\ (x,y)\in\Omega^+}} \nabla_{(-n_2^\gamma, n_1^\gamma)} z(x,y) - \lim_{\substack{(x,y)\to(\widetilde{x},\widetilde{y}) \\ (x,y)\in\Omega^-}} \nabla_{(-n_2^\gamma, n_1^\gamma)} z(x,y) = \nabla_{\mathbf{t}^\gamma}^+ z - \nabla_{\mathbf{t}^\gamma}^- z = 0,$$

due to the orthogonality of $(-n_2^\gamma, n_1^\gamma) = \mathbf{t}^\gamma$ with respect to the normal vector (n_1^γ, n_2^γ) of the discontinuity curve $\gamma(t)$, we get the inequality verified independently from c and $\mathbf{h}$. ∎

From a theoretical point of view, Lemma 1 and Theorem 1 guarantee the existence of a unique weak (Lipschitz) solution for (14), extending the procedure adopted in [36, 39, 40]. Such a procedure is very constructive since it is based on the characteristic strip expansion, allowing a fast approximation of the surface starting from the known Dirichlet boundary condition across the whole image domain.

We emphasize that this is a remarkable result, due to the fact that this entire formulation, being independent of such fundamental parameters as c and $\mathbf{h}$, could potentially be generalized for any isotropic BRDFs. That is, since the previous theorem does not depend on the particular choice of c and $\mathbf{h}$, we wonder what would be the most suitable choice of those parameters in order to deal with an irradiance equation that is as close as possible to a physical BRDF. In other words,

$$(35) \qquad [c^*(x,y,z), \mathbf{h}^*(x,y,z)] = \operatorname*{argmin}_{c,\mathbf{h}} \|I(c,\mathbf{h}) - I_{BRDF}\|,$$

where $I(c, \mathbf{h})$ is as presented in (4) and I_{BRDF} is intended as a more physically reliable BRDF, such as, for example, the complete Blinn–Phong model or the Cook–Torrance one. Both of these irradiance equations will be used in the experimental part (section 5.3) for comparison purposes and denoted, respectively, I_{BP} and I_{CT}.

We leave this part as an interesting future work.

4. Numerics on the unifying equation. In this section we present two conceptually different numerical solvers for the unifying equation. The first, based on the fast-marching scheme, is an extension of the differential treatment introduced in [39] and adopted in [40, 57, 38]. However, since such an approximation method uses a single pair of images for each pixel, it remains a theoretical framework which lacks robustness when it comes to real-world data. Indeed, this procedure turns out to be too sensitive to noise. In particular, especially when

mostly specular images are taken into account, artifacts appear by the direction of propagation for the characteristic strip expansion [38]. Hence, we introduce a second solver that relies on a variational approach aimed at increasing the robustness of the shape recovery. This plays a basic role since specular highlights provide a more difficult scenario in which the low signal-to-noise ratio attained at the darker image regions renders the reconstruction more susceptible to pixel noise.

In the following preliminary result, we show that each pair of vector fields coming in turn from different pairs of images is a linearly independent vector in the image plane.

Theorem 2. *Let $\mathbf{b}_{ij}(x, y, z)$ be the vector field of* (12) *where $(i, j) \in \Gamma$. Then $\forall (i_1, j_1), (i_2, j_2) \in \Gamma \ \forall (x, y) \in \bar{\Omega}$ and $\forall z \in Lip(\bar{\Omega})$ we have*

$$\mathbf{b}_{i_1 j_1}(x, y, z) \cdot \mathbf{b}_{i_2 j_2}(x, y, z) \neq \pm |\mathbf{b}_{i_1 j_1}(x, y, z)| |\mathbf{b}_{i_2 j_2}(x, y, z)|. \tag{36}$$

Proof. In order to avoid too many parameters, let us fix the indices (i_1, j_1) and (i_2, j_2) as $(1, 2)$ and $(1, 3)$, respectively. We prove by contradiction that $\mathbf{b}_{(1,2)}$ and $\mathbf{b}_{(1,3)}$ are never parallel, assuming that there exist a point $(\widetilde{x}, \widetilde{y}) \in \Omega$ and a Lipschitz function $\widetilde{z}$ such that

$$\mathbf{b}_{(1,2)}(\widetilde{x}, \widetilde{y}, \widetilde{z}) \cdot \mathbf{b}_{(1,3)}(\widetilde{x}, \widetilde{y}, \widetilde{z}) = \pm |\mathbf{b}_{(1,2)}(\widetilde{x}, \widetilde{y}, \widetilde{z})| |\mathbf{b}_{(1,3)}(\widetilde{x}, \widetilde{y}, \widetilde{z})|. \tag{37}$$

To improve readability we omit the dependence on $(\widetilde{x}, \widetilde{y}, \widetilde{z})$. Now, by squaring both sides, we have

$$\left(b_{(1,2)}^1 b_{(1,3)}^1 + b_{(1,2)}^2 b_{(1,3)}^2 \right)^2 = \left(\left(b_{(1,2)}^1 \right)^2 + \left(b_{(1,2)}^2 \right)^2 \right) \left(\left(b_{(1,3)}^1 \right)^2 + \left(b_{(1,3)}^2 \right)^2 \right). \tag{38}$$

Using the terms from (12), replacing the respective irradiance equation in (4) for I_1, I_2, and I_3, and eliminating the nonnull quantities like the albedo and the light attenuations, this can be rewritten as follows:

$$\begin{aligned}
&\bar{\mathbf{n}} \cdot \bar{\mathbf{h}}_3 (f \bar{h}_1^1 \bar{h}_2^2 - f \bar{h}_1^2 \bar{h}_2^1 + \bar{h}_1^3 \bar{h}_2^2 x - \bar{h}_1^2 \bar{h}_2^3 x + \bar{h}_1^2 \bar{h}_2^3 y - \bar{h}_1^3 \bar{h}_2^1 y) \\
&+ \bar{\mathbf{n}} \cdot \bar{\mathbf{h}}_2 (f \bar{h}_1^2 \bar{h}_3^1 - f \bar{h}_1^1 \bar{h}_3^2 - \bar{h}_1^3 \bar{h}_3^2 x + \bar{h}_1^2 \bar{h}_3^3 x - \bar{h}_1^1 \bar{h}_3^3 y + \bar{h}_1^3 \bar{h}_3^1 y) \\
&+ \bar{\mathbf{n}} \cdot \bar{\mathbf{h}}_1 (f \bar{h}_2^1 \bar{h}_3^2 - f \bar{h}_2^2 \bar{h}_3^1 + \bar{h}_2^3 \bar{h}_3^2 x - \bar{h}_2^2 \bar{h}_3^3 x + \bar{h}_2^1 \bar{h}_3^3 y - \bar{h}_2^3 \bar{h}_3^1 y) = 0.
\end{aligned}$$

After noticing that all the terms in the previous equation are divided by the same nonnull quantities as $|\mathbf{n}|$ and $|\mathbf{h}_i|$ for $i = 1, 2, 3$, we can consider the nonnormalized vectors (i.e., we can remove the bar on top of each factor). Then, substituting (2) for the normal like $\mathbf{n}$, we get

$$h_1^1 h_2^3 h_3^2 - h_1^3 h_2^1 h_3^2 - h_1^1 h_2^2 h_3^3 + h_1^2 h_2^1 h_3^3 + h_1^3 h_2^2 h_3^1 - h_1^2 h_2^3 h_3^1 = 0, \tag{39}$$

arriving at the final contradiction, since it means that the light sources had to be placed at collinear points, which is a condition commonly excluded in the PS field. ■

This result says that at each point the vector fields $\mathbf{b}_{i_1 j_1}$ and $\mathbf{b}_{i_2 j_2}$ span the image plane in every direction for every surface z. Furthermore, this result allows us to adopt the same strategy presented in [39], driving the characteristic strip expansion to reconstruct the surface. As an experimental consequence, we consider this result important also for the variational

approach. In fact, the hypothesis of noncollinear light sources must still be valid for the functional to minimize.

Let us define the discretized image domain as $\bar{\Omega}_d = \Omega_d \cup \partial\Omega_d$ as the grid of points (x_h, y_k), where $x_h = h\Delta_x$ and $y_k = k\Delta_y$ with $h = -n, \dots, n$ and $k = -m, \dots, m$.

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

$$(40) \qquad \mathbf{B}^\top \nabla z = \mathbf{s}.$$

4.1. The fast-marching approach. In order to use the characteristic strip approach, we linearly combine the available $\binom{N}{2}$ unifying equations in (40) with the aim of driving the characteristic expansion along the most suitable path across the image domain, as introduced in [39]. The numerical strategy will depend on the direction we want by choosing image pairs that have vector fields which allow different flow directions.

4.1.1. Upwind numerical scheme. We derive the upwind numerical discretization of (11) after introducing its vanishing viscosity formulation as follows:

$$(41) \qquad \mathbf{b}_{ij}(x, y, z) \cdot \nabla z(x, y) = \varepsilon_x \frac{\partial^2 z}{\partial x^2} + \varepsilon_y \frac{\partial^2 z}{\partial y^2} + s_{ij}(x, y, z),$$

where $\varepsilon_x = |b_{ij}^1| \frac{\Delta_x}{2}$ and $\varepsilon_y = |b_{ij}^2| \frac{\Delta_y}{2}$. That yields the following discretized equation:

$$(42) \quad b_{h,k}^1 \frac{Z_{h+1,k} - Z_{h-1,k}}{2\Delta_x} + b_{h,k}^2 \frac{Z_{h,k+1} - Z_{h,k-1}}{2\Delta_y}$$
$$= |b_{h,k}^1| \frac{\Delta_x}{2} \frac{Z_{h+1,k} - 2Z_{h,k} + Z_{h-1,k}}{\Delta_x^2} + |b_{h,k}^2| \frac{\Delta_y}{2} \frac{Z_{h,k+1} - 2Z_{h,k} + Z_{h,k-1}}{\Delta_y^2} + s_{h,k},$$

where we denote $\mathbf{b}_{ij}(x_h, y_k, z(x_h, y_k))$ as $\mathbf{b}_{h,k} = (b_{h,k}^1, b_{h,k}^2)$ and $s_{ij}(x_h, y_k)$ as $s_{h,k}$ by omitting the dependency on the pair of images (i, j).

After some algebra, we get the following fixed point scheme:

$$(43) \quad Z_{h,k}^{(t+1)} = \frac{|b_{h,k}^1(Z_{h,k}^{(t)})| Z_{i - \mathrm{sgn}\,(b_{h,k}^1(Z_{h,k}^{(t)})), k}^{(t)} + |b_{h,k}^2(Z_{h,k}^{(t)})| Z_{h, k - \mathrm{sgn}\,(b_{h,k}^2(Z_{h,k}^{(t)}))}^{(t)} + \Delta s_{h,k}(Z_{h,k}^{(t)})}{|b_{h,k}^1(Z_{h,k}^{(t)})| + |b_{h,k}^2(Z_{h,k}^{(t)})|},$$

where $\Delta_x = \Delta_y = \Delta$. The well-posedness of (43) is ensured by Lemma 1 since it guarantees that the sum $|b_{h,k}^1(Z_{h,k}^{(t)})| + |b_{h,k}^2(Z_{h,k}^{(t)})|$ does not vanish.

We must emphasize one severe limitation of the characteristic strip expansion method: it is not suitable for approximating several parameters for (42). In the next part we show that the proposed variational method can approximate simultaneously several parameters as $z(x, y)$ and $c(x, y)$. In the experimental section a quantitative comparison shows the superiority of the following variational method while dealing with noisy images and shadows.

4.2. The variational approach. In order to enhance robustness to noise and outliers such as shadows or nondifferentiable elements of the surface, we treat such effects as *sparse* phenomena and rely on l^1 optimization [27, 50]. Contrarily to the work by Ikehata et al. [27], we consider *global* l^1 optimization over the whole image domain Ω and all image ratios (i, j). Thus, we solve the quasi-linear system of PDEs (40) in the L^1 sense by considering the following optimization problem:

$$(44) \qquad \min_z \left\| \mathbf{B}(z)^\top \nabla z - \mathbf{s}(z) \right\|_{L^1(\overline{\Omega})},$$

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

Some advantages of this formulation over state-of-the-art approaches are listed below.

Merging PS data. We remark that most of the previous works proposing robust approaches to the PS problem, e.g., [27, 42, 50, 58, 63], derived the mathematical formulation by considering each irradiance equation independently of the others. Since our model is derived by coupling irradiance equations considering their ratio, the unifying equation represents a more suitable mathematical tool for the PS problem. In fact, it exploits the single view acquisition, simplifying the complexity of the problem due to heterogeneous materials.

Robustness. An important advantage directly follows from this coupling, one that is based on the number of unifying equations when solving PS with $N > 2$. In fact, most of the approaches attempt to solve a linear system having N equations provided by the irradiance equation for each image. Instead, 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 are well adapted.

Colored scenes and materials. We remark that color images are no longer a problem: since each color channel provides one unifying equation, they can be stacked in the same way as above. This is another remarkable result since we do not need to assume that either the source or the material is monochromatic. This assumption is usually introduced because it is impossible to separate the colored light intensity from the colored albedo. Since our formulation is independent of the albedo, only the directions and intensities of the light sources have to be explicitly considered. This means that, for real-world applications, calibration of the light must be performed with respect to each color channel.

Missing data. If, for some pixel (x, y), no information is available in any of the images $I_i(x, y), i \in \{1, \ldots, N\}$, due, for instance, to a null albedo (no light reflected at all), the corresponding set of unifying equations (40) 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

$$(45) \qquad \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \qquad\qquad \text{over } \Omega_{\text{missing}},$$

$$(46) \qquad \mathbf{s} = \mathbf{0} \qquad\qquad\qquad\qquad \text{over } \Omega_{\text{missing}},$$

which will enforce $\nabla z = 0$ over this subset Ω_{missing}. This can be seen as a built-in hole filling of the noninformative areas to ensure some smoothness and prevent artifacts. This can be very useful when dealing with very specular objects, since in such cases 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 [26].

We emphasize that this variational problem is conceptually very different from the state-of-the-art sparsity-enhancing technique [27], 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. Proceeding thusly, it is less likely to produce artifacts in, for instance, areas with very low intensity in all images, since they will be implicitly smoothed by the global minimization procedure. Furthermore, it allows one to directly recover piecewise-smooth surfaces without relying on a posteriori solvers, which are either fast but not robust to discontinuities [53], or robust to discontinuities but slow [16]. Our approach using global l^1 minimization is closer to what has been studied in [50], but with the major difference being that the initial PDE (40) is quasi-linear. Similar quasi-linear PDEs were also used in [51, 54], but using only least-squares optimization; this led to highly efficient numerics, but they were not as robust as the proposed l^1 optimization.

4.3. Orthographic camera and directional lightings. In the specific case of orthographic cameras and directional lightings, neither $\mathbf{B}$ nor $\mathbf{s}$ depends 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 (44) into its zero-order Tikhonov regularized version,

$$(47) \qquad \min_z \left\| \mathbf{B}^\top \nabla z - \mathbf{s} \right\|_{L^1(\overline{\Omega})} + \lambda \left\| z - z_0 \right\|_{L^2(\overline{\Omega})}^2 ,$$

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 [27], since it involves the partial derivatives of z. In this view, it is convenient to rewrite (47) under its ADMM form [5]:

$$(48) \qquad \begin{cases} \min_{z, \mathbf{g}} \|\mathbf{g}\|_{L^1(\overline{\Omega})} + \lambda \|z - z_0\|_{L^2(\overline{\Omega})}^2 \\ \text{s.t.} \quad \mathbf{g} = \mathbf{B}^\top \nabla z - \mathbf{s}. \end{cases}$$

Introducing the scaled dual variable $\mathbf{u}$, this new problem can be solved using the alternating direction scheme:

$$(49) \qquad \mathbf{g}^{k+1} = \underset{\mathbf{g}}{\operatorname{argmin}} \, \|\mathbf{g}\|_{L^1(\overline{\Omega})} + \alpha \left\| \mathbf{g} - \left(\mathbf{B}^\top \nabla z^k - \mathbf{s} - \mathbf{u}^k \right) \right\|_{L^2(\overline{\Omega})}^2 ,$$

$$(50) \qquad z^{k+1} = \underset{z}{\operatorname{argmin}} \frac{\lambda}{\alpha} \|z - z_0\|_{L^2(\overline{\Omega})} + \left\| \mathbf{B}^\top \nabla z - \left(\mathbf{g}^{k+1} + \mathbf{s} + \mathbf{u}^k \right) \right\|_{L^2(\overline{\Omega})}^2 ,$$

$$(51) \qquad \mathbf{u}^{k+1} = \mathbf{u}^k + \mathbf{g}^{k+1} - \mathbf{B}^\top \nabla z^{k+1} + \mathbf{s}$$

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. Equation (49) can be solved pointwise by shrinkage, and (50) by Gauss–Seidel iterations. By discretizing directly (50) and then writing the optimality condition, we avoid introducing any boundary condition, while state-of-the-art differential methods for PS require the depth to be known on the boundary [10].

4.4. 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

$$(52) \qquad z^{k+1} = \operatorname*{argmin}_z \left\| \mathbf{B}(z^k)^\top \nabla z - \mathbf{s}(z^k) \right\|_{L^1(\overline{\Omega})} + \lambda \left\| z - z^k \right\|_{L^2(\overline{\Omega})}^2$$

starting from an initial solution $z^0 = z_0$. Proof of convergence towards a local minimum is left for future work, but we experimentally observed that it was reached in only a few iterations.

4.5. Approximating the shininess parameter c. Since the unifying irradiance equation allows a simultaneous parameterization 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 (10). A typical iterative scheme can be written as

$$(53) \qquad \begin{cases} z^{k+1} = \operatorname*{argmin}_z \left\| \mathbf{B}(z^k, c^k)^\top \nabla z - \mathbf{s}(z^k, c^k) \right\|_{L^1(\overline{\Omega})}, \\ c^{k+1} = \operatorname*{argmin}_c \left\| \left(\frac{I_i a_j(z^k)}{I_j a_i(z^k)} \right)^c - \frac{\overline{\mathbf{h}}_i(z^k, c^k) \cdot \mathbf{n}(z^k)}{\overline{\mathbf{h}}_j(z^k, c^k) \cdot \mathbf{n}(z^k)} \right\|_{L^1(\overline{\Omega})}, \end{cases}$$

where the z update can be obtained as described earlier.

Regarding the c update (53), it comes down to solving in a L^1 sense:

$$(54) \qquad \left(\frac{I_i a_j(z^k)}{I_j a_i(z^k)} \right)^c = \frac{\overline{\mathbf{h}}_i(z^k, c^k) \cdot \mathbf{n}(z^k)}{\overline{\mathbf{h}}_j(z^k, c^k) \cdot \mathbf{n}(z^k)}.$$

Taking the logarithm of both members, this is rewritten as

$$(55) \qquad c \log\left((I_i a_j)/(I_j a_i)\right) = \log\left((\overline{\mathbf{h}}_i \cdot \mathbf{n})/(\mathbf{h}_j \cdot \mathbf{n})\right).$$

Solving this equation in the L^1 sense, the c update is given by

$$(56) \qquad c^{k+1} = \operatorname{median}\left\{ \log\left((\overline{\mathbf{h}}_i \cdot \mathbf{n})/(\mathbf{h}_j \cdot \mathbf{n})\right) / \log\left((I_i a_j)/(I_j a_i)\right) \right\}.$$

We note that (56) allows for estimating one c value per pixel, and thus it is possible to deal with scenes containing objects of various reflectances without the need for performing any segmentation. However, if a scene is known to contain a single object of uniform reflectance, c can be set to the mean of c to get a more accurate estimate.
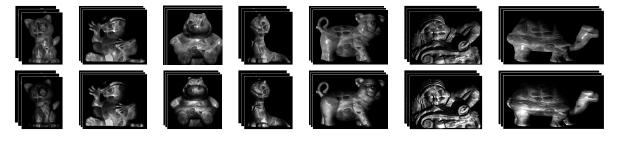
As challenging proof of concept, we demonstrate the working principle of our method on synthetic data from the MERL dataset [35] as well as real image scenario.

5. Experiments. We now demonstrate experimentally, both on synthetic and real-world datasets, the benefits of using the proposed approaches while showing the superiority of the variational method over the characteristics solver and comparing them with other approaches.

5.1. Quantitative evaluation on synthetic datasets. To quantitatively evaluate the robustness of our method, we first performed tests on synthetic data. To provide fair comparison with state-of-the-art methods, we considered orthographic projection ($f = \infty$), directional lightings, and two well-established reflectance models, namely, the purely diffuse Lambertian model ($c = 1$) and the purely specular Blinn–Phong model ($c = 0.4$). In both cases, in order to evaluate the robustness to outliers, we simulated self-shadowing effects, and in the case of Lambertian reflectance we also added specular spots (according to the bichromatic Blinn–Phong model).

To create the data, we used the seven real-world sets of normals provided in the Harvard dataset,[1] which we integrated into a depth map using our solver (setting $\mathbf{B}$ as in (45)). Then the normals were recomputed by finite differences and a synthetic albedo map was computed by overlapping the `hestain.png` built-in image from MATLAB. Finally, the images were generated under 8 different directional lightings (Figure 2).

Figure 2. *Synthetic images used for quantitative evaluation. Top: images simulated according to Lambert's law, with self-shadowing effects and additive specular spots. Both of these effects are outliers to Lambert's law, and hence are useful to evaluate the robustness of our method to real-world outliers. Bottom: specular images computed using $c = 0.4$, so as to produce wide lobes.*

We first evaluated the performances of the proposed variational recovery of the depth with respect to the presence of outliers (self-shadows and specular spots) and compared it to the traditional least-squares [62] and state-of-the-art l^1 [27] estimation of the normals. Then we evaluated its ability to recover objects whose reflectance is purely specular, and compared it with the results obtained using state-of-the-art robust Lambertian PS [27] and general reflectance PS [26].

In both cases, we evaluated the mean error (ME) and the standard deviation (std) on the orientation of the normals (computing in our case the normals by the same finite differences as above) and on the depth (for [62] and [27], and the depth was computed with the same l^1 solver we used to generate the ground truth). All codes were written in MATLAB and executed on an i7 at 3.4 GHz. We considered as stopping criteria reaching a relative residual inferior to 10^{-4}. The results shown in Tables 1 and 2 indisputably prove both the robustness of our approach and its ability to reconstruct specular objects with few images.

---

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

Table 1

*Robustness to outliers, for least-squares estimation of the normals [62], l^1 estimation of the normals [27], and the proposed l^1 estimation of the depth, on synthetic data with orthographic projection, Lambertian reflectance with self-shadows and specular spots. We show the error on the normals orientations **n** in degrees and the absolute error on the depth z in pixels. Our method systematically outperforms state-of-the-art methods regarding robustness to outliers: considering the depth, rather than the normals, seems to improve the results because the inherent smoothing it brings will tend to limit the appearance of artifacts, for instance, in low-albedo areas. It is remarkable that, though the normals are approximated using finite differences, after running our method, they are more accurate than those recovered using dedicated methods.*

	Cat				Frog			
	n		z		**n**		z	
	ME	std	ME	std	ME	std	ME	std
[62]	7.7	**4.6**	19	15	6.6	**4.0**	14	12
[27]	6.7	5.3	17	13	5.7	4.3	12	11
Ours	**6.0**	5.7	**13**	**10**	**4.6**	5.1	**9.1**	**8.8**

	Hippo				Lizard			
	n		z		**n**		z	
	ME	std	ME	std	ME	std	ME	std
[62]	8.3	**5.5**	23	15	7.6	**4.8**	13	9.2
[27]	7.2	6.0	21	14	7.0	5.8	12	8.6
Ours	**6.9**	6.5	**20**	**11**	**6.9**	5.9	**12**	**8.5**

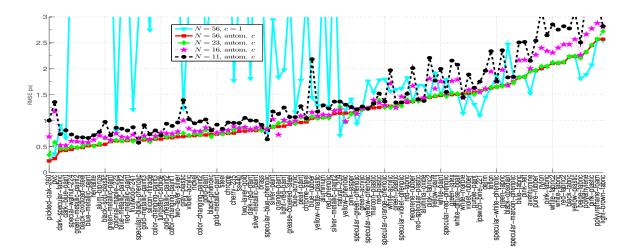|	Pig				Scholar				Turtle			
	n		z		**n**		z		**n**		z	
	ME	std	ME	std	ME	std	ME	std	ME	std	ME	std
[62]	7.1	**4.5**	13	12	6.8	**4.2**	21	11	7.4	**5.0**	9.3	8.8
[27]	6.3	5.0	12	11	5.7	4.5	18	10	6.6	5.5	8.5	8.7
Ours	**5.8**	5.6	**10**	**9.2**	**5.1**	4.8	**13**	**7.4**	**6.5**	6.2	**7.6**	**8.0**

5.2. Handling arbitrary reflectances. In this section, we question the ability of the proposed BRDF model (4) to handle real-world reflectances (Figures 3 and 4). To this purpose, we recovered the shape of a piece of sphere rendered according to the BRDFs of the MERL database [35]. First, we applied the proposed variational scheme using $N = 56$ images while imposing $c = 1$; some partly retroreflective materials (`pickled-oak-256`) are handled since this effect can be considered an outlier. However, 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 overestimated in most cases (`alum-bronze`). On the other hand, when simultaneously estimating the shape and the reflectance, as discussed in section 4.5, most materials of the database are reasonably recovered. Notable exceptions include very dark materials (`black-obsidian`) or highly retroreflective materials (`polyethylene`).

We also used this dataset to perform experiments on the required number of input images. Figure 3 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 approaches [26]: 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.

Table 2

Specular surface reconstruction. We compare the results of robust Lambertian estimation of the normals [27], general isotropic reflectance estimation of the normals [26], and the proposed method (c = 0.4) on the same data as in Table 1, but with purely specular images. Since considering purely specular reflectance is significantly different from assuming outliers to the Lambertian model, it is no surprise that [27] performs poorly. Given the low number of images we take into account, even the general isotropic reflectance method from [26] fails to recover consistent normals in areas where little information is available, though such artifacts tend to be smoothed after integration. On the other hand, since our algorithm uses an explicit BRDF model, and because the proposed PDE/variational approach tends to smooth artifacts, our method provides accurate 3D reconstruction even with few data.

	Cat $\mathbf{n}$		Cat z		Frog $\mathbf{n}$		Frog z	
	ME	std	ME	std	ME	std	ME	std
[27]	8.5	5.3	17	13	7.2	4.7	13	11
[26]	11	6.1	12	9.6	9.7	5.3	11	7.6
Ours	**2.4**	**1.1**	**7.9**	**4.1**	**2.3**	**1.0**	**5.9**	**4.1**

	Hippo $\mathbf{n}$		Hippo z		Lizard $\mathbf{n}$		Lizard z	
	ME	std	ME	std	ME	std	ME	std
[27]	8.9	6.0	19	14	8.3	5.3	10	7.8
[26]	11	6.6	15	10	11	6.3	9.1	6.3
Ours	**2.8**	**1.2**	**7.9**	**5.3**	**3.2**	**1.1**	**15**	**7.7**

	Pig $\mathbf{n}$		Pig z		Scholar $\mathbf{n}$		Scholar z		Turtle $\mathbf{n}$		Turtle z	
	ME	std	ME	std	ME	std	ME	std	ME	std	ME	std
[27]	9.0	5.2	14	11	8.1	4.9	20	12	8.5	5.8	10.9	8.6
[26]	11	6.6	13	8.4	11.2	6.4	21	13	11	6.3	13	8.8
Ours	**2.6**	**1.1**	**5.0**	**3.4**	**3.0**	**1.0**	**11.8**	**8.6**	**3.1**	**1.2**	**6.8**	**4.3**

Figure 3. *RMSE on the depth as a function of the BRDF. Most specular materials are correctly handled with a dozen images and automatic reflectance estimation.*

pickled-oak-260   pvc   alum-bronze
1   #2   #10

two-layer-silver   black-obsidian   polyethylene
#24   #83   #97

Figure 4. *3D reconstruction results for several materials from the MERL database, sorted according to their ranking in Figure* 3. *Bottom row shows the ground-truth surface and reflectance; middle row shows the shape recovered while estimating automatically c using* (53)*; top row shows the results with c = 1.*

5.3. Synthetic experiments comparing with double-lobe BRDFs. We generated more challenging synthetic datasets by relaxing the assumptions of the orthographic viewing geometry and the directional lighting. For that, we considered a virtual camera of 12mm focal length placed around 70mm away from an object of size around 30mm. In addition, we employed more realistic shading models, namely, full Blinn–Phong (BP) and Cook–Torrance (CT). More precisely, the assumed irradiance equation for the full Blinn–Phong model is given by

$$(57) \qquad I_{BP} = a(\rho_s(N \cdot H)^m + \rho_d(N \cdot L)),$$

where ρ_s is specular albedo, ρ_d is diffuse albedo ($\rho_s + \rho_d = 1$), a is the light attenuation defined in (9), and $m = 10$ is the specularity coefficient. We note that (57) is quite different from our irradiance equation (4), as it has separate lobes for the diffuse and the specular reflection.

We also considered the Cook–Torrance reflectance model given by

$$(58) \qquad I_{CT} = a\left(\rho_s \frac{FDG}{\pi(N \cdot L)(N \cdot V)} + \rho_d(N \cdot L)\right), \text{ where}$$

$$(59) \qquad F = F_0 + (1 - F_0)(1 - (V \cdot H))^5,$$

$$(60) \qquad D = \frac{1}{\pi R^2 (N \cdot H)^4} e^{\left(\frac{(N \cdot H)^2 - 1}{R^2 (N \cdot H)^2}\right)},$$

$$(61) \qquad G = \min\left(1, \frac{2(N \cdot H)(N \cdot V)}{(V \cdot H)}, \frac{2(N \cdot H)(N \cdot L)}{(L \cdot H)}\right),$$
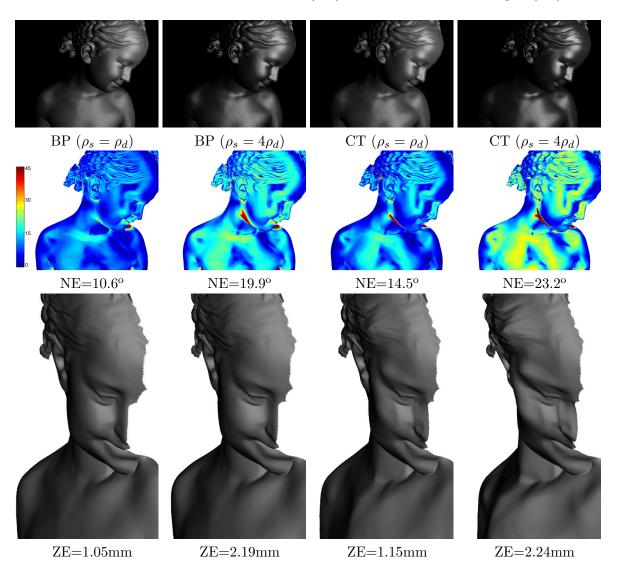
setting $F_0 = 0.75$, $R = 0.09$ in the experiments.

For each reflection model we performed two experiments, one with equal specular and diffuse albedos ($\rho_s = \rho_d = 0.5$) and one with mostly specular ($\rho_s = 0.75$, $\rho_d = 0.25$). This was

done in order to test our single-lobe BRDF on the more realistic scenario of additive specular and diffuse reflectances.

Each experiment contained 6 images of resolution 800×600 pixels with light sources placed on a ring around the camera at 50mm. The images were rendered with OpenGL code adapted from http://renderman.pixar.com.

The results are shown in Figure 5. For quantifying the reconstruction quality, we use two metrics: the mean error in normal orientation (NE) and the mean error in depth (ZE).

BP ($\rho_s = \rho_d$)	BP ($\rho_s = 4\rho_d$)	CT ($\rho_s = \rho_d$)	CT ($\rho_s = 4\rho_d$)
NE=10.6°	NE=19.9°	NE=14.5°	NE=23.2°
ZE=1.05mm	ZE=2.19mm	ZE=1.15mm	ZE=2.24mm

Figure 5. *Synthetic data rendered with various reflection models. Row one: 1 sample image for each dataset. Row two: normal error maps. Row three: reconstructions.*

5.4. Robustness to noise and shadows. Finally, we tested the robustness of our algorithm to noise and shadows. We start with the Cook–Torrance rendered sequence of Figure 5 with

$\rho_s = \rho_d$ (i.e., third column) and added up to 15% Gaussian noise as shown in the first three columns (from left to right) of Figure 6. In the last three columns of Figure 6, images with missing data (i.e., "shadows") are simulated by setting a subset of image pixels black. We chose a repeated small rectangles pattern of width 10 pixels and height 5 to 15 pixels. Each one of the 6 images of the sequence has the patterns slightly displaced in order to avoid having regions that are always in shadow.

The last three rows of Figure 6 emphasize the robustness of our L_1 variational solver, showing a comparison with other state-of-the-art methods [27, 38, 47]. Note that in order to get depth estimates for [27, 47], our variational solver was used on top of the obtained normals. Finally, we compare the variational approach with the upwind characteristic strip expansion of (43). Due to the lack of evident experimental difference between the implementation proposed in [38] and (43), we compare the characteristic strip expansion as similarly conceived from [38]. In order to compare it fairly with the proposed variational method, we use an estimate of the shininess coefficient computed from the variational approach in order to approximate the only depth using the characteristic strip expansion.

5.5. Qualitative evaluation on real-world datasets.

Medium-scale scene. We performed tests on several real-world datasets featuring objects of different materials, depth discontinuities, and shadows. Again we compare with other state-of-the-art methods as explained in section 5.4. As [38] performs very badly in the presence of outliers (see last row of Figure 6), we do not use it here. Samples from these datasets are presented in Figure 7.
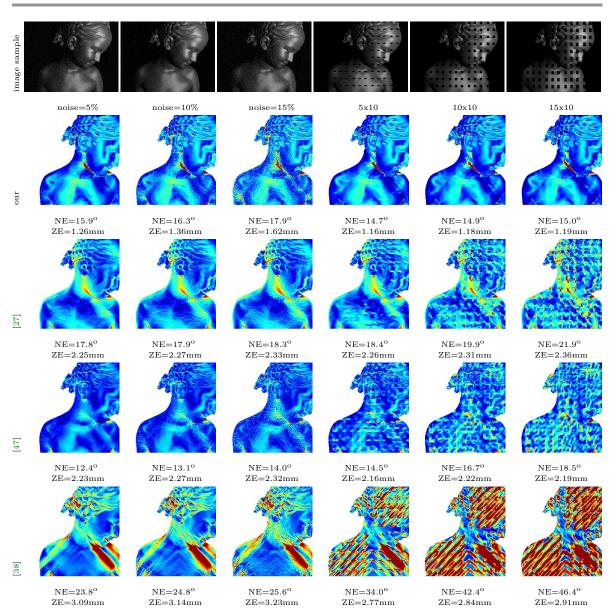
The fist dataset contained an Egyptian statue and a metallic tea box placed approximately 50cm away from a calibrated pinhole camera ($f \approx 32$mm). We used 8 LEDs with known positions, situated between the camera and the objects (approximately at 30cm from the scene). To get the best reconstruction quality we manually segmented the two objects of interest and forced the optimization to select a uniform shininess parameter c (see explanation after (56)).

The second dataset contained a Buddha statue placed 12cm away from a camera with $f = 12$mm. We used 8 LEDs with known positions on a ring of radius 5cm around the camera. We let the optimization select different c values for each pixel so as to be able to deal with small, highly specular spots that exist on the surface of this statue.

The final dataset featured a scene a various objects taken with the same setup as the first dataset. We note that this is a particularly challenging dataset as the specular objects in the middle of the scene (tea box and ball) are very badly illuminated, and the objects cast shadows on each other (we present all 8 images of the dataset here as a clear illustration of the challenge of this data). We note that we only performed manual background subtraction on these images and let our algorithm do the rest in order to push it to its limits. Nonetheless, we are able to get a reasonable reconstruction clearly outperforming the competitors (see Figure 10).

Additionally, these results are compared in Figures 8 and 9 to several existing methods for Lambertian PS, namely, the classical PS approach proposed by Woodham [62], the robust approach from [27] dealing with both diffuse and specular reflectance, and a recent work on PS with pointwise sources [47]. Figure 8 demonstrates the advantage of *global* L^1-based recovery

Figure 6. *The first three columns show a comparison with increasing noise. Instead, the last three columns compare the robustness of the methods having data with increasing missing parts (black patches on the images).*

over *local* sparsity-enhancing techniques [27], and the importance of modeling appropriately the lighting and the viewing. Besides these important considerations, Figure 9 shows that our approach even provides satisfactory 3D reconstructions of specular objects.

Micro-scale scene. Finally, we tested the ability of our approach to handle the images recorded using a high focal camera, using a device similar to the portable one introduced by Johnson et al. [31]. The LEDs illuminating the scene being inside the device, the light rays reach the surface after bouncing off the walls of the device, rather than directly. Hence, it

Figure 7. *Top row: two samples images from our first two datasets (Egyptian & Box and Buddha). Middle and bottom row: all 8 images of our third dataset consisting of a scene of multiple objects with varying reflectance properties.*
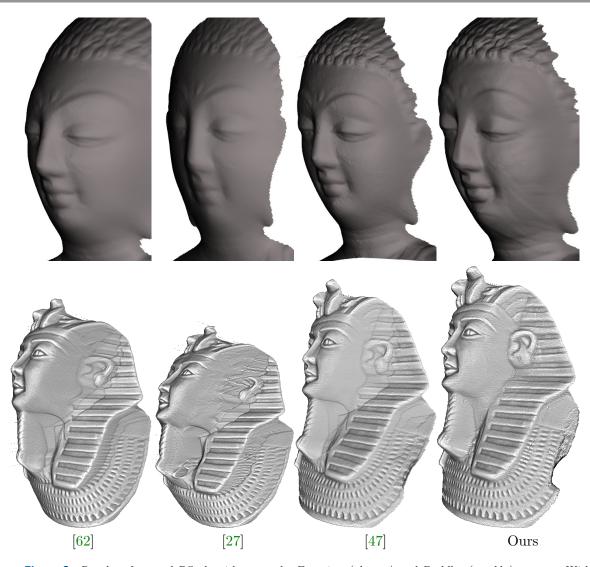
is not possible to consider any explicit lighting model such as the pointwise one. To deal with this issue, the lightings were calibrated using a 3D-printed planar rig of patterns with known geometry and albedo, from which the light directions and intensities can be extracted at the center of each pattern. Eventually, these are interpolated, so as to obtain dense sets $\mathbf{l}_i : \bar{\Omega} \to \mathbb{R}^3$ and $\phi_i : \bar{\Omega} \to \mathbb{R}$, allowing us to use our algorithm without any change, simply setting the attenuation terms to $a_i = \phi_i$. This process was performed with respect to each color channel.

We tested this approach on a very challenging dataset consisting of 5 RGB pictures, of size 687×964, of a highly specular coin (Figure 11). The camera calibration procedure provided the mean camera distance $z_0 \equiv 69$mm in the same time as the focal value $f \approx 150$mm. We set $\alpha = 0.2$ in this experiment and let the algorithm run until convergence. Convergence happened after only 3 iterations (approx. 10 seconds): indeed, there is neither shadow nor discontinuity in this scene, and hence the l^1 optimization converges almost instantly.

While the approach presented in [31] introduced a chemical gel between the camera and the object, in order to make the reflectance of the object as diffuse as possible, we were able to obtain a very reasonable 3D reconstruction using the images "as they are."
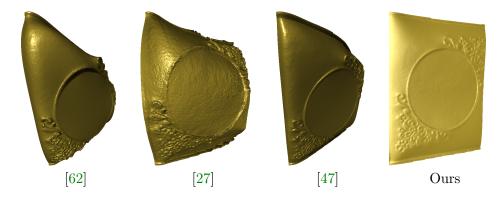
6. Conclusion and perspectives. In this work we introduced a unifying irradiance equation describing both diffuse and specular reflection components in the most general lighting setting. We derived a new formulation based on coupling those irradiance equations by considering their ratios. The resulting mathematical model consists of a set of quasi-linear PDEs

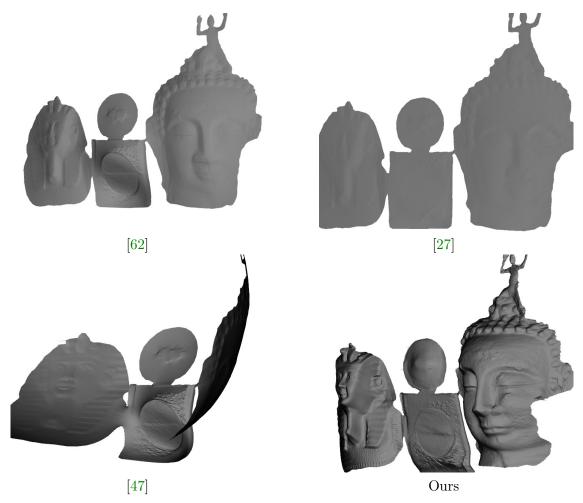| [62] | [27] | [47] | Ours |

Figure 8. *Results of several PS algorithms on the Egyptian (plaster) and Buddha (marble) statues. With the classical approach* [62] *that assumes uniform lighting and orthographic viewing, artifacts due to shadows are visible. By* locally *enforcing robustness as proposed in* [27], *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* [47].
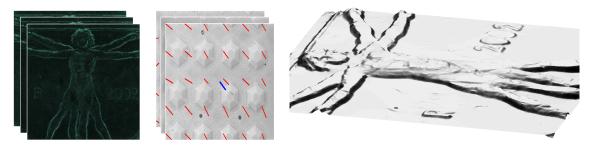
having several advantages with respect to the usual tools aimed at solving the PS problem. In fact, instead of considering specular highlights as outliers, we exploit geometric information of the shape by modeling specular reflectance with a Blinn–Phong extension of the cosine law for diffuse reflection. 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 a system of quasi-linear PDEs is based on a variational approach performing l^1 minimization. On the other hand,

[62]   [27]   [47]   Ours

Figure 9. *Results of several PS algorithms on the tea box (metal) dataset. Our approach provides a more reliable 3D reconstruction of this purely specular object, compared to other methods based on the Lambertian assumption.*

[62]   [27]

[47]   Ours

Figure 10. *Results of several PS algorithms on multiple objects scene. The overall lighting of this scene is very challenging (as objects cast shadows on each other; see Figure 7), and we also have to deal with non-uniform reflectance properties completely automatically. Thus, the quality of reconstruction for the Egyptian and the tea box are worse compared to Figures 8 and 9.*

Figure 11. *Reconstruction of a specular coin using arbitrary spatially varying lightings. From left to right: three of the five RGB images, corresponding calibrated lightings associated to the red channel, and surface reconstructed using our algorithm, with c = 0.2. Our general formulation allows us to reconstruct surfaces even without any lighting model, thanks to the space dependency of the fields* **B** *and* **s**.

research 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 brings very challenging goals closer to fruition. For instance, such a big number of equations can be used for estimating parameters related to the material of the surface, lightings (uncalibrated PS), etc.

REFERENCES

[1] N. G. ALLDRIN AND D. J. KRIEGMAN, *Toward reconstructing surfaces with arbitrary isotropic reflectance: A stratified photometric stereo approach*, in Proceedings of the 2007 IEEE International Conference on Computer Vision (ICCV), 2007.

[2] S. BARSKY AND M. PETROU, *The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows*, IEEE Trans. Pattern Anal. Mach. Intell., 25 (2003), pp. 1239–1252.

[3] P. BECKMANN AND A. SPIZZICHINO, *The Scattering of Electromagnetic Waves from Rough Surfaces*, 1987.

[4] J. F. BLINN, *Models of light reflection for computer synthesized pictures*, in Proceedings of the 4th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'77), ACM, New York, 1977, pp. 192–198.

[5] 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. Learning, 3 (2011), pp. 1–122.

[6] M. BREUSS, E. CRISTIANI, J.-D. DUROU, M. FALCONE, AND O. VOGEL, *Perspective shape from shading: Ambiguity analysis and numerical approximations*, SIAM J. Imaging Sci., 5 (2012), pp. 311–342, https://doi.org/10.1137/100815104.

[7] G. J. BROSTOW, C. HERNÁNDEZ, G. VOGIATZIS, B. STENGER, AND R. CIPOLLA, *Video normals from colored lights*, IEEE Trans. Pattern Anal. Mach. Intell., 33 (2011), pp. 2104–2114.

[8] A. M. BRUCKSTEIN, *On shape from shading*, Comput. Vision Graphic Image Process., 44 (1988), pp. 139–154.

[9] J. CHABROWSKI AND Z. KEWEI, *On variational approach to photometric stereo*, Ann. Inst. H. Poincaré Anal. Non Linéaire, 10 (1993), pp. 363–375.

[10] M. CHANDRAKER, J. BAI, AND R. RAMAMOORTHI, *On differential photometric reconstruction for unknown, isotropic BRDFs*, IEEE Trans. Pattern Anal. Mach. Intell., 35 (2013), pp. 2941–2955.

[11] H. J. CHUNG AND J. JIA, *Efficient photometric stereo on glossy surfaces with wide specular lobes*, in Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Anchorage, AK, 2008.

[12] J. J. CLARK, *Active photometric stereo*, in Proceedings of the 1992 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Champaign, États-Unis, 1992, pp. 29–34.

[13] T. COLLINS AND A. BARTOLI, 3*D reconstruction in laparoscopy with close-range photometric stereo*, in Medical Image Computing and Computer-Assisted Intervention (MICCAI), Lecture Notes in Comput. Sci. 7511, Springer-Verlag, Berlin, 2012, pp. 634–642.

[14] R. L. COOK AND K. E. TORRANCE, *A reflectance model for computer graphics*, ACM Trans. Graph., 1 (1982), pp. 7–24.

[15] F. COURTEILLE, A. CROUZIL, J. D. DUROU, AND P. GURDJOS, *Towards shape from shading under realistic photographic conditions*, in Proceedings of the 17th International Conference on Pattern Recognition, 2004, pp. 277–280.

[16] J.-D. DUROU, J.-F. AUJOL, AND F. COURTEILLE, *Integrating the normal field of a surface in the presence of discontinuities*, in Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR), Bonn, Germany, Lecture Notes in Comput. Sci. 5681, Springer-Verlag, Berlin, 2009, pp. 261–273.

[17] D. B. GOLDMAN, B. CURLESS, A. HERTZMANN, AND S. M. SEITZ, *Shape and spatially-varying BRDFs from photometric stereo*, IEEE Trans. Pattern Anal. Mach. Intell., 32 (2010), pp. 1060–1071.

[18] H. HAYAKAWA, *Photometric stereo under a light source with arbitrary motion*, J. Opt. Soc. Amer. A, 11 (1994), pp. 3079–3089.

[19] C. HERNANDEZ, G. VOGIATZIS, AND R. CIPOLLA, *Overcoming shadows in 3-source photometric stereo*, IEEE Trans. Pattern Anal. Mach. Intell., 33 (2011), pp. 419–426.

[20] L. HOELTEN, Y. QUÉAU, M. BREUSS, AND G. RADOW, *Optimised photometric stereo via non-convex variational minimisation*, in British Machine Vision Conference (BMVC), York, 2016.

[21] B. K. P. HORN, *Obtaining shape from shading information*, in The Psychology of Computer Vision, P. H. Winston, ed., McGraw-Hill, New York, 1975, pp. 115–155.

[22] B. K. P. HORN, *Height and gradient from shading*, Internat. J. Comput. Vision, 5 (1990), pp. 37–75.

[23] B. K. P. HORN AND M. J. BROOKS, *The variational approach to shape from shading*, Comput. Vision Graphic Process., 33 (1986), pp. 174–208.

[24] B. K. P. HORN, R. S. SZELISKI, AND A. L. YUILLE, *Impossible shaded images*, IEEE Trans. Pattern Anal. Mach. Intell., 15 (1993), pp. 166–169.

[25] K. HYEONGWOO, J. HAILIN, S. HADAP, AND I.-S. KWEON, *Specular reflection separation using dark channel prior*, in Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Portland, OR, 2013, pp. 1460–1467.

[26] S. IKEHATA AND K. AIZAWA, *Photometric stereo using constrained bivariate regression for general isotropic surfaces*, in Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, 2014, pp. 2187–2194.

[27] S. IKEHATA, D. WIPF, Y. MATSUSHITA, AND K. AIZAWA, *Robust photometric stereo using sparse regression*, in Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Providence, RI, 2012, pp. 318–325.

[28] K. IKEUCHI, *Determining surface orientations of specular surfaces by using the photometric stereo method*, IEEE Trans. Pattern Anal. Mach. Intell., 3 (1981), pp. 661–669.

[29] Y. IWAHORI, H. SUGIE, AND N. ISHII, *Reconstructing shape from shading images under point light source illumination*, in Proceedings of the 1990 IEEE International Conference on Pattern Recognition (ICPR), Vol. 1, Atlantic City, NJ, 1990, pp. 83–87.

[30] D. W. JACOBS, P. N. BELHUMEUR, AND R. BASRI, *Comparing images under variable illumination*, in Proceedings of the 1998 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Santa Barbara, CA, 1998, pp. 610–617.

[31] M. K. JOHNSON, F. COLE, A. RAJ, AND E. H. ADELSON, *Microgeometry capture using an elastomeric sensor*, ACM Trans. Graphics, 30 (2011), 46.

[32] R. KIMMEL AND A. M. BRUCKSTEIN, *Tracking level sets by level sets: A method for solving the shape from shading problem*, J. Computer Vision Image Understanding, 62 (1995), pp. 47–58.

[33] M. KURT AND D. EDWARDS, *A survey of BRDF models for computer graphics*, SIGGRAPH Comput. Graphics, 43 (2009), pp. 4:1–4:7.

[34] S. LEE AND M. BRADY, *Integrating stereo and photometric stereo to monitor the development of glaucoma*, Image Vision Comput., 9 (1991), pp. 39–44.

[35] W. MATUSIK, H. PFISTER, M. BRAND, AND L. MCMILLAN, *A data-driven reflectance model*, ACM Trans. Graphics, 22 (2003), pp. 759–769.

[36] R. MECCA AND M. FALCONE, *Uniqueness and approximation of a photometric shape-from-shading model*, SIAM J. Imaging Sci., 6 (2013), pp. 616–659, https://doi.org/10.1137/110857258.

[37] R. MECCA AND Y. QUÉAU, *Unifying diffuse and specular reflections for the photometric stereo problem*, in Proceedings of the 2016 IEEE Workshop on Applications of Computer Vision (WACV), Lake Placid, NY, 2016.

[38] R. MECCA, E. RODOLÀ, AND D. CREMERS, *Realistic photometric stereo using partial differential irradiance equation ratios*, Comput. Graphics, 51 (2015), pp. 8–16.

[39] R. MECCA, A. TANKUS, A. WETZLER, AND A. M. BRUCKSTEIN, *A direct differential approach to photometric stereo with perspective viewing*, SIAM J. Imaging Sci., 7 (2014), pp. 579–612, https://doi.org/10.1137/120902458.

[40] R. MECCA, A. WETZLER, A. M. BRUCKSTEIN, AND R. KIMMEL, *Near field photometric stereo with point light sources*, SIAM J. Imaging Sci., 7 (2014), pp. 2732–2770, https://doi.org/10.1137/140968100.

[41] T. MIGITA, S. OGINO, AND T. SHAKUNAGA, *Direct bundle estimation for recovery of shape, reflectance property and light position*, in Proceedings of the 2008 European Conference on Computer Vision (ECCV), Marseille, France, Lecture Notes in Comput. Sci. 5304, Springer-Verlag, Berlin, 2008, pp. 412–425.

[42] D. MIYAZAKI AND K. IKEUCHI, *Photometric stereo under unknown light sources using robust SVD with missing data*, in Proceedings of the 2010 International Conference on Image Processing, ICIP 2010, Hong Kong, China, 2010, pp. 4057–4060.

[43] S. NAYAR, K. IKEUCHI, AND T. KANADE, *Determining shape and reflectance of hybrid surfaces by photometric sampling*, IEEE Trans. Robotics Automat., 6 (1990), pp. 418–431.

[44] A. NGAN, F. DURAND, AND W. MATUSIK, *Experimental analysis of BRDF models*, in Proceedings of the 16th Eurographics Conference on Rendering Techniques, EGSR '05, Aire-la-Ville, Switzerland, 2005, pp. 117–126.

[45] 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*, Computer Vision Image Understanding, 66 (1997), pp. 119–131.

[46] T. PAPADHIMITRI AND P. FAVARO, *A new perspective on uncalibrated photometric stereo*, in Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Portland, OR, 2013, pp. 1474–1481.

[47] T. PAPADHIMITRI AND P. FAVARO, *Uncalibrated near-light photometric stereo*, in British Machine Vision Conference (BMVC), Nottingham, UK, 2014.

[48] B. T. PHONG, *Illumination for computer generated pictures*, Commun. ACM, 18 (1975), pp. 311–317.

[49] E. PRADOS AND O. FAUGERAS, *Shape from shading: A well-posed problem?*, in Proceedings of the 2005 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 2, San Diego, CA, 2005, pp. 870–877.

[50] Y. QUÉAU, F. LAUZE, AND J.-D. DUROU, *A L^1-TV algorithm for robust perspective photometric stereo with spatially-varying lightings*, in Scale Space and Variational Methods in Computer Vision (SSVM), Lège Cap-Ferret, France, Lecture Notes in Comput. Sci. 9087, Springer-Verlag, Berlin, 2015, pp. 498–510.

[51] Y. QUÉAU, R. MECCA, AND J.-D. DUROU, *Unbiased photometric stereo for colored surfaces: A variational approach*, in Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016.

[52] Y. QUÉAU, R. MECCA, J. D. DUROU, AND X. DESCOMBES, *Photometric stereo with only two images: A theoretical study and numerical resolution*, submitted.

[53] T. SIMCHONY, R. CHELLAPPA, AND M. SHAO, *Direct analytical methods for solving Poisson equations in computer vision problems*, IEEE Trans. Pattern Anal. Mach. Intell., 12 (1990), pp. 435–446.

[54] W. SMITH AND F. FANG, *Height from photometric ratio with model-based light source selection*, Computer Vision Image Understanding, 145 (2016), pp. 128–138.

[55] A. TANKUS AND N. KIRYATI, *Photometric stereo under perspective projection*, in Proceedings of the 2005 IEEE International Conference on Computer Vision (ICCV), Vol. 1, Beijing, China, 2005, pp. 611–616.

[56] K. E. TORRANCE AND E. M. SPARROW, *Theory for off-specular reflection from roughened surfaces*, J. Opt. Soc. Amer., 57 (1967), pp. 1105–1112.

[57] S. TOZZA, R. MECCA, M. DUOCASTELLA, AND A. D. BUE, *Direct differential photometric stereo shape recovery of diffuse and specular surfaces*, J. Math. Imaging Vision, 56 (2016), pp. 57–76.

[58] Y. WANG, C. M. LEE, L. CHEONG, AND K. TOH, *Practical matrix completion and corruption recovery using proximal alternating robust subspace minimization*, Internat. J. Comput. Vision, 111 (2015), pp. 315–344.

[59] G. J. WARD, *Measuring and modeling anisotropic reflection*, in Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '92), ACM, New York, 1992, pp. 265–272.

[60] L. WOLFF, *On the relative brightness of specular and diffuse reflection*, in Proceedings of the 1994 IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, 1994, pp. 369–376.

[61] L. B. WOLFF AND E. ANGELOPOULOU, *3-D stereo using photometric ratios*, in European Conference on Computer Vision (ECCV), Stockholm, Sweden, Lecture Notes in Comput. Sci. 801, Springer-Verlag, New York, 1994, pp. 247–258.

[62] R. J. WOODHAM, *Photometric method for determining surface orientation from multiple images*, Optical Engrg., 19 (1980), pp. 134–144.

[63] L. WU, A. GANESH, B. SHI, Y. MATSUSHITA, Y. WANG, AND Y. MA, *Robust photometric stereo via low-rank matrix completion and recovery*, in Proceedings of the 2010 Asian Conference on Computer Vision (ACCV), Queenstown, New Zealand, Lecture Notes in Comput. Sci. 6494, Springer-Verlag, New York, 2010, pp. 703–717.

[64] T.-P. WU, K. L. TANG, C.-K. TANG, AND T. T. WONG, *Dense photometric stereo: A Markov random field approach*, IEEE Trans. Pattern Anal. Mach. Intell., 28 (2006), pp. 1830–1846.

[65] Q. YANG AND N. AHUJA, *Surface reflectance and normal estimation from photometric stereo*, Computer Vision Image Understanding, 116 (2012), pp. 793–802.