



SMI 2015

Realistic photometric stereo using partial differential irradiance equation ratios

R. Mecca^{a,*}, E. Rodolà^{b,1}, D. Cremers^b^a Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5, 40126 Bologna, Italy^b Computer Vision Group, Technische Universität München, Department of Computer Science, Boltzmannstrasse 3, 85748 Garching, Germany

ARTICLE INFO

Article history:

Received 10 March 2015

Received in revised form

15 May 2015

Accepted 16 May 2015

Available online 29 May 2015

Keywords:

Photometric stereo

Partial differential equations

Image ratios

Diffuse component

Specular component

ABSTRACT

Shape from shading with multiple light sources is an active research area and a diverse range of approaches have been proposed in the last decades. However, devising a robust reconstruction technique still remains a challenging goal due to several highly non-linear physical factors being involved in the image acquisition process. Recent attempts at tackling the problem via photometric stereo rely on simplified hypotheses in order to make the problem solvable. Light propagation is still commonly assumed to be uniformly oriented, and the BRDF assumed to be diffuse, with limited interest for materials giving specular reflection. Taking into account realistic point light sources, in this paper we introduce a well-posed formulation based on partial differential equations for both diffuse and specular shading models. We base our derivation on the popular approach of image ratios, which makes the model independent from photometric invariants. The practical effectiveness of our method is confirmed with a wide range of experiments on both synthetic and real data, where we compare favorably to the state of the art.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Nowadays, the importance of having three-dimensional objects on a computer directly imported from the real world is increasing due to the fact that many applicative fields need to observe, manipulate or reproduce reality. However, among the several existing techniques for 3D scanning, only few can be actually considered as emerging technologies in the market. Structured light [1] and multi-view stereo [2] are gradually moving to smart devices, making them able to provide a rough depiction of a depth field rather than an accurate 3D reconstruction. Reconstruction accuracy depends on a number of factors which may prevent a precise measurement of the observed scene, and for these reasons additional constraints need to be imposed on the acquisition environment.

Photometric stereo (PS) allows highly accurate reconstructions under the requirement to work in a controlled setup, making it applicable in limited scenarios. However, starting from the seminal work of Woodham [3], the use of complex shading models has not evolved much in modern PS approaches [4,5]. Indeed, the simplifying approach that models the captured images as an inner product between uniform light direction and outgoing surface normal is still commonly

found in the literature [6]. Shape reconstruction from shading information is a difficult problem, due to the complexity of the underlying physical process describing how a light beam bounces on the surface. Thus, it becomes fundamental to take into account the parametrization of all elements that influence the image formation. Although the behavior of light itself (e.g., propagation and attenuation) does require to be carefully modeled, the bidirectional reflectance direction function (BRDF) represents the true bridge between the real world and the depicted one. Most of the literature dealing with PS assume diffuse reflection (i.e., uniform in all directions), reducing the mathematical model to a linear problem where the normal field can be easily computed [7] and finally integrated [8]. Realistically, this approach contains too many assumptions which fail as soon as the method is applied in many real-world applications. There are at least two reasons why the reconstruction of *specular* surfaces still remains a challenging task in the PS field. First, the BRDF for specular reflections is highly non-linear, which means that mathematically it is difficult to have a straightforward solvability. Second, the specular reflection has a very susceptible signal (see Fig. 1, upper part), easily corruptible by environment and acquisition noise.

Instead of considering shape reconstruction in the wild [9,10], in this paper we take into account a controlled lighting setup in which point light sources are used to illuminate the observed object. This kind of setup is commonly used in the PS field [11,12] by virtue of its applicability in several interesting applications [13–15]. We introduce a new differential model which allows to extract

* Corresponding author.

E-mail addresses: roberto.mecca@unibo.it (R. Mecca), rodola@in.tum.de (E. Rodolà), cremers@in.tum.de (D. Cremers).¹ Supported by an Alexander von Humboldt fellowship.



Fig. 1. Full reconstruction of the “bimba” model from 6 partial views under 5 point light sources. The images used for the reconstruction (top, one light source only) consist solely of the *specular* component of the reflected light, yet our method provides a near-perfect reconstruction of the model. The artifact on top of the head is due to the surface region not being visible in any of the partial views. See Section 4.1 for more details on the reconstruction process.

shape information even for specular surfaces (Fig. 1). The approach is based on the Blinn–Phong shading model [16], which has been shown to be a realistic BRDF [17]. The model we propose improves upon a recent proposal [18], which only takes into account *diffuse* surfaces. Similarly to this previous work, we formulate the problem using non-linear PDEs as our main mathematical tools.

1.1. Related work

As mentioned in the previous section, dealing with general BRDFs is a challenging problem due to the non-linearities arising in the image formation process. Previous studies consider an object of the same material as a reference [19], or attempt to get rid of the specularly by eliminating highlights in a pre-processing step [20,21]. Dichromatic reflectance models have also been considered as a basis for diffuse and specular reflection separation, in particular [22] makes use of a dark channel prior and [23] iteratively compares the intensity logarithmic differentiation of the input image. The dichromatic model can be employed to perform shape reconstruction of the diffuse component only [24,25]. Yang and Abuja [26] use both the diffuse and specular components for the reconstruction by assuming the illumination chromaticity to be known, and at least one of the input images to be free of specularly.

Several works take into account general irradiance equations, leading to difficult problems that are then solved after substantial simplifications. For example, Ikehata et al. [4,5] use the purely diffuse irradiance equation for general surfaces, and consider the specular component as a sparse error. In order to provide an

accurate solution, they introduce a regression procedure which requires tens of images. The use of such a big amount of data partially justifies the assumption of ignoring specular reflection (due to its sparsity). On the other hand, expensive algorithms based on energy minimization tend to be slow and cannot provide real-time shape recovery. Furthermore, light propagation is considered to be uniform, not allowing realistic features such as radial propagation and attenuation of light to be taken into account.

When the linearization of real physical effects comes into play, it induces deformations on the recovered surface. Motivated by this, we formulate a new PS differential model for specular surfaces that takes into account point light sources. For this purpose we make use of non-linear PDEs, a mathematical tool that has attracted increasing interest in the last few years [27,28,11]. These approaches consider *image ratios* in order to yield photometric invariants, and model the irradiance equations via PDEs. Mecca et al. [27,28,18] use specific irradiance equations for diffuse surfaces, and the uniqueness of solution is proved by characteristic strip expansion and assuming known light information (*i.e.*, direction or position). Chandraker et al. [11,29] consider more general irradiance equations with unknown light sources, and compute the photometric invariants describing the surface through its isocontours [30]. However, the shape reconstruction process requires additional initial or boundary information.

All the works mentioned so far take into account *uniform* light propagation, which restricts the problem to a very specific and controlled environment. Parametrization of realistic point light sources is not new [15,31], yet very few approaches apply this idea to shape reconstruction. Wu et al. [13] use two point light sources

in an endoscopic scenario. Rather than assuming two separate irradiance equations, they couple them in such a way to lead to a shape-from-shading problem, where the unique irradiance equation is assumed to be their sum.

With the aim to make the modeling of our PS technique as realistic as possible, we consider the reconstruction of diffuse and specular surfaces as separate cases. We extend the parametrization of point light sources described in [18], and modify it so as to perfectly fit a well-posed differential problem based on image ratios of partial differential irradiance equations, leading to quasi-linear PDEs. Besides tackling a more realistic scenario, our approach is computationally efficient as it can be implemented via fast marching methods. To the best of our knowledge, this is the first successful attempt at tackling this more general and realistic scenario.

1.2. Contribution

In this paper we introduce a new photometric stereo technique for 3D shape recovery. Among other advantages, our model gets rid of many unrealistic restrictions imposed by recent parametrizations [18], which we extend and improve upon in several ways. In particular, despite the non-linearities arising from the image formation process, light parametrization and attenuation, our model allows to successfully recover diffuse and (most notably) specular surfaces.

The key contributions of this paper can be summarized as follows:

- Our approach takes into account *any* kind of light attenuation, including but not limited to radial and distance-based dissipation effects.
- Our parametrization allows for *general* light positions (Fig. 2). By contrast, previous approaches either consider uniform lighting, or require the light sources to be attached to the optical plane of the camera [18,13,11].
- We define a well-posed differential model for *specular* surfaces by making use of image ratios of the Blinn–Phong shading model [16]. We extend the differential model to pinhole cameras, where the viewer direction changes radially.
- We provide a parametrization for realistic light attenuation and propagation of *point* light sources illuminating specular surfaces.
- Our method can be implemented efficiently via fast marching, and allows accurate recovery from few images in a matter of seconds using 1.2 Megapixel data.

2. Mathematical model

With the aim to formulate a suitable mathematical model for realistic scenarios, we start by following the parametrization defined in [18], and introduce several fundamental extensions and generalizations.

Let us briefly recall the parametrization of the surface which takes into account the perspective deformation based on the pinhole camera model (Fig. 2). Given a surface Σ , we consider the depth $z(x, y)$ defined in the image domain $\bar{\Omega} = \Omega \cup \partial\Omega$ to \mathbb{R} , such that

$$\mathbf{u}(x, y) = (\xi(x, y), \eta(x, y), \zeta(x, y)) = \left(-x \frac{z(x, y)}{f}, -y \frac{z(x, y)}{f}, z(x, y) \right). \quad (1)$$

We indicate with $f > 0$ the focal length of the camera, where $\zeta < -f < 0$, and the triplet $(\xi(x, y), \eta(x, y), \zeta(x, y)) = (\xi, \eta, \zeta)$ represents the real world coordinates as a function of image coordinates.

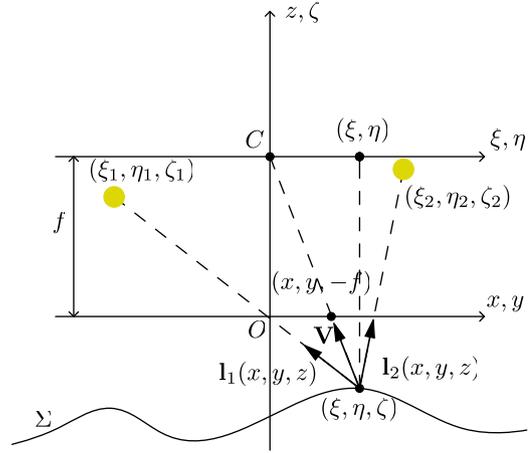


Fig. 2. A schematic section of the derivation of the perspective view geometry. In the perspective world (i.e. the image coordinate system given by $Oxyz$) the light reflected at image point (x, y) comes from the real point (ξ, η, ζ) of the surface.

The outgoing normal vector to Σ can be computed as follows:

$$\mathbf{n}(x, y) = \frac{z}{f^2} (f \nabla z(x, y), z(x, y) + (x, y) \cdot \nabla z(x, y)), \quad (2)$$

whose normalized version follows the notation:

$$\bar{\mathbf{n}}(x, y) = \frac{\mathbf{n}(x, y)}{|\mathbf{n}(x, y)|}. \quad (3)$$

Here and in the following sections, we will use $\bar{\mathbf{v}}$ to indicate that a given vector \mathbf{v} is normalized.

2.1. Unconstrained light source with attenuation

Differently from [18], we parametrize point light sources placed at general locations and thus remove the constraint that fixes them onto the optical plane. For general placements $\mathbf{P}_j = (\xi_j, \eta_j, \zeta_j)$ we have the following light directions:

$$\mathbf{l}_j(x, y, z) = \mathbf{P}_j - \mathbf{u}(x, y) = \left(\xi_j + x \frac{z(x, y)}{f}, \eta_j + y \frac{z(x, y)}{f}, \zeta_j - z(x, y) \right). \quad (4)$$

This parametrization allows us to easily include other features such as light attenuation. For example, the classical light attenuation due to energy reduction inversely proportional to (squared) distance between light source and object can be easily computed as $|\mathbf{l}_j(x, y, z)|^{-2}$.

Additionally, radial propagation of light with respect to the light source needs to be taken into account since light emitted by LEDs does *not* have uniform intensity, depending on the angle of exposure to the light source itself. Typical diagrams of radial dissipation of light have the shape shown in Fig. 3, and can be effectively computed as follows:

$$\cos(\theta)^\mu(x, y, z) = (\bar{\mathbf{l}}_j(x, y, z) \cdot (0, 0, 1))^\mu = \frac{(\zeta_j - z)^\mu}{|\mathbf{l}_j(x, y, z)|^\mu}, \quad (5)$$

where θ is the angle of exposure, $\mu > 0$ is an attenuation coefficient depending on the light source, and $\mathbf{l}_j(\mathbf{x}, \mathbf{y}, z)$ is the j -th light source placed at (ξ_j, η_j, ζ_j) .

Both attenuation effects can be combined into the following attenuation factor:

$$a_j(x, y, z) = \frac{(\zeta_j - z)^\mu}{|\mathbf{l}_j(x, y, z)|^{\mu+2}}. \quad (6)$$

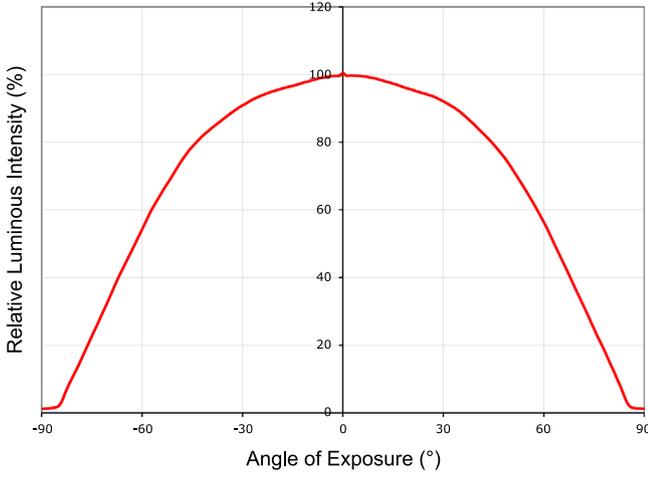


Fig. 3. The intensity of light for a directional light source. This graph is reproduced from the data sheet of the CREE XLamp MP-L EasyWhite LEDs, which were used for the real tests.

Remark. The true potential of this approach can be appreciated by considering that the model allows *any* kind of non-linear light attenuations. As we will show in the next sections, even when considering a general form for $a_j(x, y, z)$ we still obtain a well-posed differential problem.

2.2. Lambertian model

We start by briefly recalling the irradiance equation for the diffusive reflection of light. Consider the cosine law given by the following inner product:

$$D_j(x, y) = \rho_d(x, y) a_j(x, y, z) \bar{\mathbf{l}}_j(x, y, z) \cdot \bar{\mathbf{n}}(x, y), \quad (7)$$

where $D_j : \bar{\Omega} \rightarrow [0, 1]$ is the diffusive image function taken under the j -th light source, and $\rho_d(x, y)$ is the unknown diffusive albedo.

We derive our differential formulation by considering the image ratios given by D_j/D_i . This makes the problem easier to solve, since the non-linearities of the first partial derivatives of z disappear due to the simplification of $|\mathbf{n}_j(x, y, z)|$. In fact, after some algebra the ratio D_j/D_i leads to the following quasi-linear PDE:

$$\begin{aligned} & \left(D_i \frac{|\mathbf{l}_i|}{a_i} (f \xi_j - x \zeta_j) - D_j \frac{|\mathbf{l}_j|}{a_j} (f \xi_i - x \zeta_i) \right) z_x \\ & + \left(D_i \frac{|\mathbf{l}_i|}{a_i} (f \eta_j - y \zeta_j) - D_j \frac{|\mathbf{l}_j|}{a_j} (f \eta_i - y \zeta_i) \right) z_y \\ & = D_j \frac{|\mathbf{l}_j|}{a_j} (z \zeta_i - z^2) - D_i \frac{|\mathbf{l}_i|}{a_i} (z \zeta_j - z^2). \end{aligned} \quad (8)$$

These equations can be succinctly expressed as follows:

$$\begin{cases} \mathbf{b}_d(x, y, z) \cdot \nabla z(x, y) = k_d(x, y, z), & (x, y) \in \Omega \\ z(x, y) = g(x, y) & (x, y) \in \partial\Omega \end{cases} \quad (9)$$

where $g(x, y)$ is the Dirichlet boundary condition.

2.3. Specular model

We consider the well-known irradiance equation for specular surfaces, given by the Blinn–Phong shading model by the following inner product:

$$S_j(x, y) = \rho_s(x, y) a_j(x, y, z) (\bar{\mathbf{n}}(x, y) \cdot \bar{\mathbf{h}}_j(x, y, z))^c, \quad (10)$$

where $S_j : \bar{\Omega} \rightarrow [0, 1]$ is the specular image function, $c > 0$ is a specularity coefficient, and $\rho_s(x, y)$ is the unknown specular albedo.

Function $\mathbf{h}_j(x, y, z)$ is defined as

$$\mathbf{h}_j(x, y, z) = \bar{\mathbf{v}}(x, y, z) + \bar{\mathbf{l}}_j(x, y, z), \quad (11)$$

where $\bar{\mathbf{v}}(x, y, z)$ is the direction of the viewer. We emphasize that $\bar{\mathbf{v}}$ also depends explicitly on the image points (x, y) and on the depth z itself. Following this important observation, we parametrize this dependency by deriving a new Blinn–Phong shading formulation which is suitable for the differential model we present. In our setup, the viewer direction can be expressed as

$$\mathbf{v}(x, y, z) = (x, y, -f). \quad (12)$$

We remark the ductility of this differential approach: in fact, instead of considering an orthogonal viewer direction $(0, 0, 1)$ as it is commonly done, we model the pinhole camera accordingly.

Similarly to the previous case, in order to deal with a solvable differential formulation we consider the specular image ratio S_j/S_i instead of the non-linear system constructed by taking equations as in (10). After some more algebra, this leads to the following quasi-linear PDE:

$$\begin{aligned} & \left(a_j^{1/c} S_i^{1/c} (f \bar{h}_j + x \bar{h}_j^3) - a_i^{1/c} S_j^{1/c} (f \bar{h}_i + x \bar{h}_i^3) \right) z_x \\ & + \left(a_j^{1/c} S_i^{1/c} (f \bar{h}_j + y \bar{h}_j^3) - a_i^{1/c} S_j^{1/c} (f \bar{h}_i + y \bar{h}_i^3) \right) z_y \\ & = a_i^{1/c} S_j^{1/c} z \bar{h}_i^3 - a_j^{1/c} S_i^{1/c} z \bar{h}_j^3 \end{aligned} \quad (13)$$

and in turn

$$\begin{cases} \mathbf{b}_s(x, y, z) \cdot \nabla z(x, y) = k_s(x, y, z) & \text{on } \Omega \\ z(x, y) = g(x, y) & \text{on } \partial\Omega \end{cases}. \quad (14)$$

Observe that the structure of the differential problems for both the specular (9) and diffusive component (14) is actually the same, and both formulations can be shown to be mathematically well-posed admitting a unique weak Lipschitz solution.

3. Numerical scheme

In the previous section we introduced two differential formulations for the diffuse and specular reflections, respectively given in Eqs. (9) or (14), which can be solved using an identical numerical methodology for quasi-linear PDEs. Such formulations are based on pairs of images, but can be easily extended to handle $N > 2$ images by exploiting the quasi-linearity of the PDEs. Using multiple images allows us to safely ignore the boundary conditions, since depth information is not traveling from the boundary over the image domain. We adapted the numerical scheme [18] for our purposes so as to handle the unconstrained light parametrization and the specular shading model.

By keeping the diffuse and specular models separate, we consider linear combinations of those PDEs once we fix an initial point for the reconstruction (usually the central point), where the depth or initial guess thereof must be known. In our experiments we measured such distance in order to understand the artifacts of the shape recovery not depending on it. The updating formula is given by the following upwind numerical scheme:

$$z_{ij}^{(r+1)} = \frac{|b_{ij}^1(z_{ij}^{(r)})| z_{i+\text{sgn}(b_{ij}^1(z_{ij}^{(r)}))j}^{(r)} + |b_{ij}^2(z_{ij}^{(r)})| z_{ij+\text{sgn}(b_{ij}^2(z_{ij}^{(r)}))}^{(r)} - \Delta k_{ij}(z_{ij}^{(r)})}{|b_{ij}^1(z_{ij}^{(r)})| + |b_{ij}^2(z_{ij}^{(r)})|} \quad (15)$$

where Δ is the size of the discretized image domain. The vector field $\mathbf{b}(x_i, y_j) = (b_{ij}^1, b_{ij}^2)$ and $k(x_i, y_j) = k_{ij}$ are suitable linear combinations of the $\binom{N}{2}$ quasi-linear PDEs. Using a fast-marching procedure, we couple pairs of (9) or (14) in order to compute the directional derivative according to eight principal directions (2 horizontals, 2 verticals and 4 diagonals) which span the two-

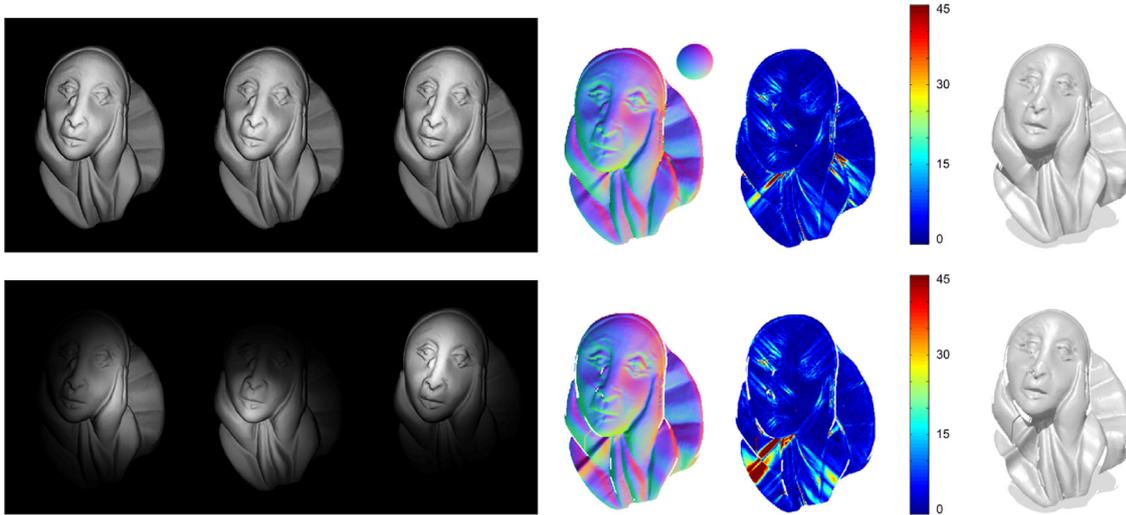


Fig. 4. Reconstruction of the Lambertian “pierrot” model under freely positioned light sources with realistic radial dissipation. In each row we show the three images used for the reconstruction (left), followed by the reconstructed normals, the angular error (in degrees) with respect to the ground-truth normals, and the final reconstructed surface. The two rows correspond to light sources with almost null ($\mu = 1.1$) and strong ($\mu = 30$) attenuation coefficients respectively. Notice how the reconstruction remains fairly accurate even when the signal is very weak; see, for instance, the visible part of the folded collar in the second row.

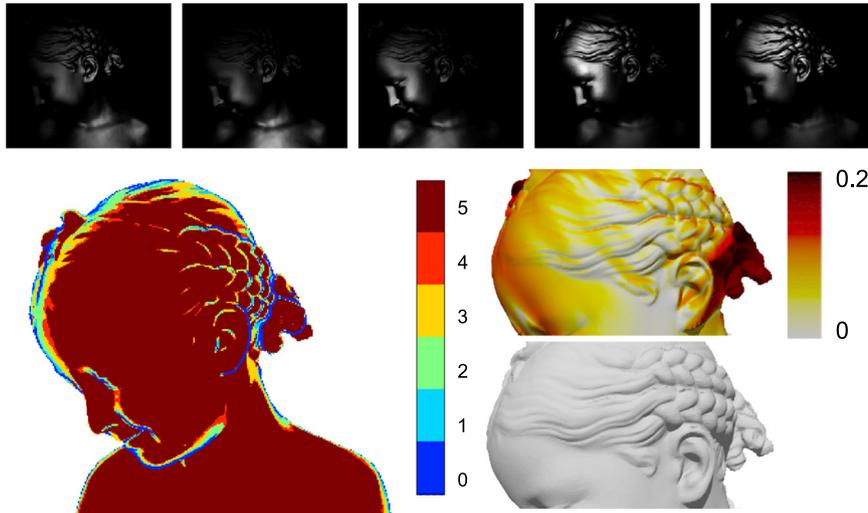


Fig. 5. Reconstruction under realistic shadows and missing data. In this example, the weak signal is due to both the model being a specular surface, and to the strong radial dissipation of the light sources ($\mu = 40$). The shadow map (bottom left) counts the number of lights that “see” each point, as calculated from the 5 images used for the reconstruction (top row). On the right we show a closeup of the ground-truth model (bottom), and the reconstructed surface with a per-vertex error map plotted over it (top). Notice that the maximum error is equal to 2 mm^2 .

dimensional image domain [28]. We choose the specific equation by taking the pair of images having the highest gray scale value at pixel (x_i, y_j) .

4. Experimental results

We performed a wide range of experiments on synthetic and real data. In order to provide a quantitative evaluation of our method, we generated synthetic data from the “bimba”, “pierrot”, and “bumpy sphere” 3D models from the AIM@Shape Repository. Specifically, for each model we took several snapshots from different points of view, with a virtual perspective pinhole camera placed at $\sim 45 \text{ cm}$ from the object. Each object was rescaled to have a diameter of $\sim 30 \text{ cm}$; camera resolution was fixed at 640×480 , with no simulated lens distortion, and the camera plane placed at $z = 0$ as in Fig. 2. Other variables such as number and position of lights, specularity, noise, and light attenuation are investigated with more specific experiments in the following sections.

In these experiments we initialize the reconstruction process (Section 3) with the ground-truth depth value of the central pixel; this is required only for comparison purposes, as by doing so we fix the scale of the reconstructed surface to be the same as the original object.

4.1. Sensitivity analysis

Aim of this section is to study the sensitivity of our model with respect to different parametrizations, and under the action of various sources of nuisance. We provide two measures of error in order to give a quantitative assessment of our results, namely, the *angular error* between the reconstructed and ground-truth normals (degrees), and the point-to-point *mean squared error (MSE)* between the recovered 3D surface and the ground-truth object (mm^2).

Light attenuation: As described in Section 2.1, our model takes into account general light positions as well as the combined attenuation due to radial dissipation of light and distance from the source. In Fig. 4 we show the reconstruction of a diffuse surface

from three images generated by manually displaced point lights. Notice that the depth discontinuity between the right side of the face and the folded collar (~ 15 cm deeper) is handled well, and the collar can still be recovered with low error. Average MSE values for the two reconstructions are 0.97 mm^2 (first row) and 2.33 mm^2 (second row). We emphasize that although for simplicity in this experiment we use lights with the same attenuation properties, our model can accommodate arbitrary dissipation models for each individual light source.

Missing data and shadows: When dealing with complex objects, shadows and missing data may occur frequently also depending on the measurement conditions and the physical properties of the object itself. For this experiment we chose a challenging setting: a specular surface, with complex details and strong light attenuation. The results are shown in Fig. 5. Notice how, due to the intertwining strands of hair and the signal decay, many object pixels are either covered by few images or carry feeble information. Nevertheless, the reconstructed surface remains crisp, with no holes, and the high-frequency details are preserved.

Noise: In a separate set of experiments we investigated the sensitivity of our method to sensor noise, which can indeed occur in practical scenarios especially under low light conditions. To this end, we injected additive Gaussian noise into the synthetic images of both diffuse and specular surfaces (5 images per reconstruction), which we analyzed separately. Fig. 6 reports the results of this experiment. It is interesting to note that, while the presence of noise induces a clear decrease in accuracy, the main features of the (diffuse) object still remain intact even under strong light dissipation. The effect of noise is more visible with specular surfaces, as illustrated in Fig. 7. In this more difficult scenario, the low signal-to-noise ratio attained at the darker image regions renders the reconstruction more susceptible to pixel noise.

In this situation, integrating the information captured from different vantage points can be an effective solution when applicable. In Fig. 1 we show an example of full reconstruction of the “bimba” model from six partial views around the object. For each viewpoint, we captured 5 images with different light sources. The partial views were individually reconstructed using our method for specular surfaces (*i.e.*, solving Eq. (14)), then rigidly aligned using the ground-truth motions, and finally merged together using Poisson integration [32] on the resulting oriented point cloud. In practical scenarios, where the ground-truth poses are not available, the rigid motions relating the partial views can still be easily obtained by robust methods [33,34].

4.2. Comparisons

In this section we compare our method with other approaches from the literature. As already discussed in Section 1.1, our model is the first to deal with specular data, realistic light propagation, and arbitrary light sources all in a unified framework. For this reason, a direct comparison with other methods relying on simpler models might seem unfair. However, this experiment allows us to position our approach and clarify its benefits with respect to the existing literature. For the purpose, we chose the traditional approach based on the computation of the normal field [3] and surface reconstruction [8] as a baseline, and the more recent state-of-the-art method of Ikehata et al. [5] which also works with specular surfaces, but does not consider light attenuation effects. We keep the comparison as fair as possible by (1) radially arranging the light positions on the camera plane at 2 cm from the center of projection, (2) by completely removing sensor noise, and (3) by disabling light attenuation in the specular case.

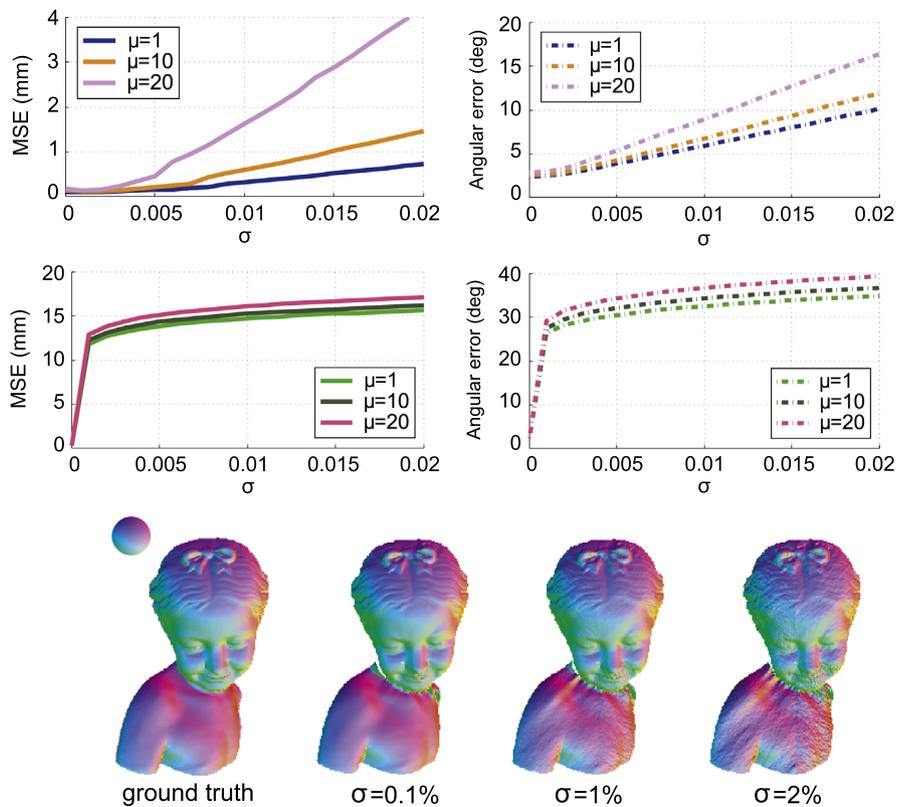


Fig. 6. First row: reconstruction error of diffuse surfaces, as a function of noise (σ) and light attenuation (μ). The curves are averaged over a collection of 8 objects; random pixel noise with standard deviation σ is added to the captured images, which are normalized to have values in $[0, 1]$. Second row: the same experiment performed over specular surfaces. Third row: reconstructed normals of a diffuse surface at different noise levels (here $\mu = 30$).

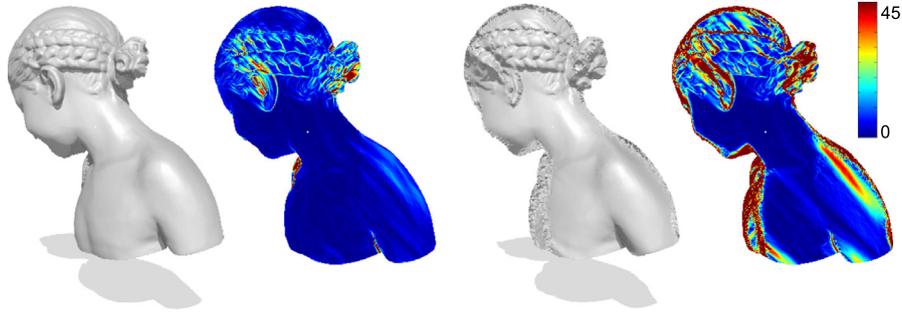


Fig. 7. Reconstruction of a specular surface in the noise-free case (left) and with 0.1% additive pixel noise (right). The colored images show the angular error (in degrees) of the reconstructed normals in the two cases. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

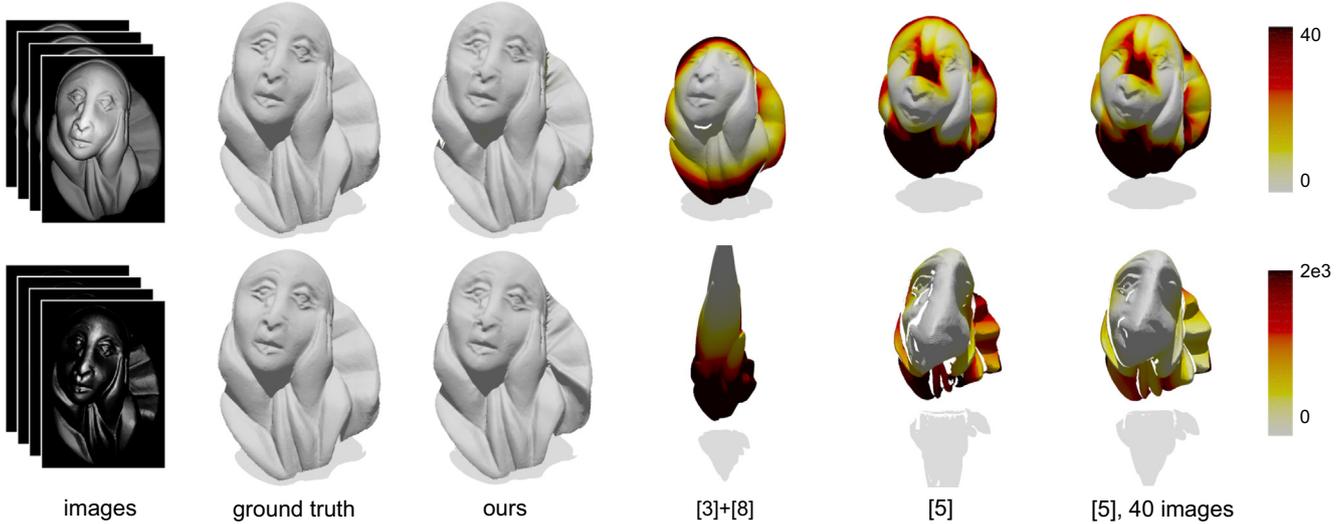


Fig. 8. Comparisons between our method and the approaches described in [3,8,5]. First row: diffuse surface with mild light attenuation ($\mu = 2$). Second row: specular surface, no attenuation. All methods take as input the same 4 images, except for the last column. MSE maps are plotted on top of each reconstruction (notice the different error scales among the two rows).

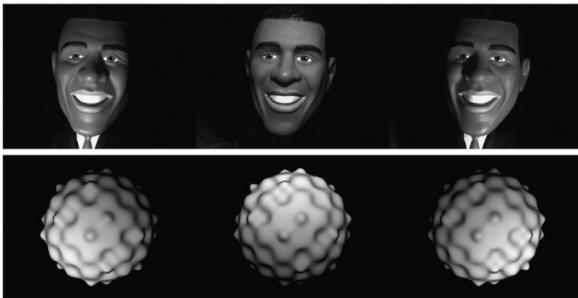


Fig. 9. Simulation of the image acquisition process. Model parameters are estimated roughly by adjusting them to get similar appearance as the real images (shown on top). In this diffuse example we use $f = 16$ mm and $\mu = 20$ (bottom).

The results of this experiment are shown in Fig. 8. All methods, including ours, have the error maps textured on the reconstructed surfaces. The baseline approach [3,8] does not deal well with attenuated light and it cannot handle specular reflections at all, hence the reconstruction process diverges completely in the latter case. The second method [5], on the other hand, attains much better results in both experiments. The effect of radial dissipation (first row) is less evident than in the former case; however, the surface still diverges from the correct geometry as light gradually diffuses from the middle. The specular object (second row) is partially recovered in correspondence to the brighter areas, although with large error; the reconstruction slightly improves when 40 images are given as input (last column). These results are probably due to

the strong signal decay that characterizes the source images, and to the inability of the method to deal with purely specular data.

4.3. Real data

In the last experiment we evaluate the practical applicability of our model for reconstruction of real world objects. Our setup consists of a pinhole Basler BIP2-1300c IP camera with 1.2 Mega-pixel, mounted on a rigid support; three ultra-bright white LEDs are radially displaced on the same plane of the camera with radius equal to 7.5 cm, and controlled by an Arduino Nano board. For this experiment we used a 2 Euro coin and a toy figure of ~ 5 cm diameter, both placed at 25 cm from the camera; this distance value was then used as the initial depth for the reconstruction process in order to fix the scale of the object.

The resulting reconstructions are depicted in Fig. 10. We would like to point out that in these experiments we did *not* perform any photometric calibration of the scene, nor did we perform an intrinsic/extrinsic calibration of the camera parameters and of the optical distortion. While a high-accuracy calibration would arguably increase the quality of the results [35], we preferred not to do so in an attempt to keep the reconstruction pipeline as simple and practical as possible, and at the same time evaluate its resilience to inaccurate measurements.

Note, however, that our model does require estimates for the focal length f (Eq. (4)), attenuation factor μ (Eq. (6)), and specular coefficient c (Eq. (10)). We obtained these values by a simple simulation of the image acquisition process: Given the

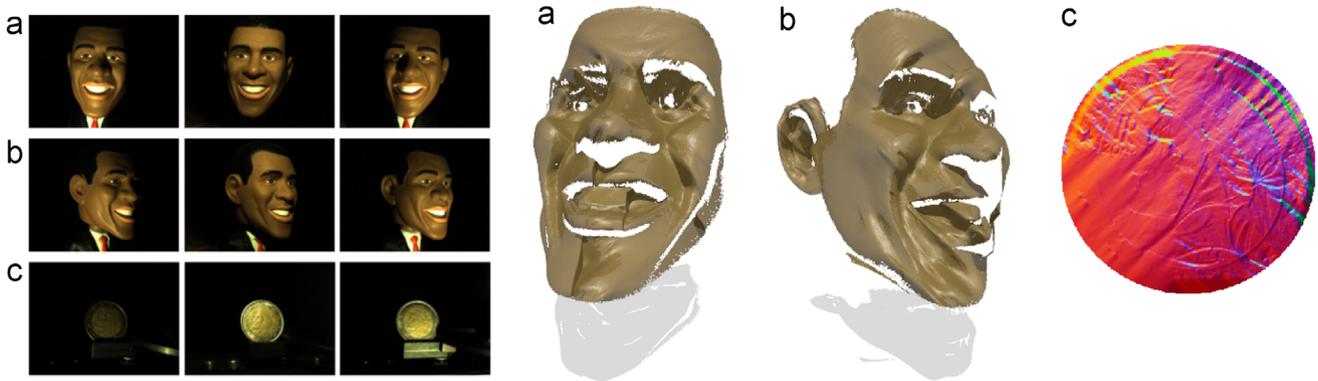


Fig. 10. Reconstruction experiments with real data. The toy figure is captured from two different points of view, which are then reconstructed independently; the Euro coin is shown with the normals to better capture the details. Despite the inaccurate calibration, the recovered geometries capture the main features of the true objects with little deformation. Missing parts are mainly due to the presence of either completely black or white regions in the depicted object. Also notice how the reconstructions are invariant to the underlying unknown albedo.

Table 1

Runtime at increasing number of images. All times are given in seconds, and averaged over a collection of 8 objects. Image resolution is 640×480 .

Number of Images	3	5	10	50
Diffuse	1.09	3.11	4.01	6.41
Specular	1.16	3.73	5.82	11.87

approximate object size, distance from the camera, CCD width (~ 4.79 mm from the technical specification) and light positions, we reproduced the projective model depicted in Fig. 2, and took a synthetic snapshot of the “bumpy sphere” under this virtual setup. This allowed us to get rough estimates for the required parameters, which were then used without further optimization. In Fig. 9 we show an example of this estimation procedure.

4.4. Runtimes

The numerical schemes (Section 3) were implemented in C++, and executed in Matlab using an Intel Core i7 3.40 GHz with 32 GB RAM. Being based on fast marching, the code can be easily parallelized; however, we did not perform any such optimization in order to keep the runtimes easily interpretable. Table 1 reports the execution times of our pipeline on image sets of increasing size. In all cases, the process reached convergence in less than 10 iterations.

Note how the specular case reports larger execution times on average. This is due to the different visibility conditions with respect to the diffuse setting, which causes the marching process to seek for better candidates over the whole collection of images whenever a dark pixel is encountered.

5. Conclusions

In this paper we tackled the problem of 3D shape recovery from multiple light sources. The proposed model overcomes the state of the art in terms of modeling physical factors; it is able to deal with photometric invariants, specular and diffusion shading models including point light sources with arbitrary attenuation. We introduced a physically motivated model based on PDEs, and demonstrated its capability to deal with challenging and realistic cases on a wide range of experiments. The method allows to obtain faithful reconstructions efficiently and in noisy conditions, confirming its practical applicability and promoting further exciting directions of research.

References

- [1] Koninckx T, van Gool L. Real-time range acquisition by adaptive structured light. *IEEE Trans Pattern Anal Mach Intell* 2006;28(3):432–45.
- [2] Kolev K, Tanskanen P, Speciale P, Pollefeys M. Turning mobile phones into 3d scanners. In: 2014 IEEE conference on computer vision and pattern recognition (CVPR); 2014. p. 3946–53.
- [3] Woodham RJ. Photometric method for determining surface orientation from multiple images. *Opt Eng* 1980;19(1):134–44.
- [4] Ikehata S, Aizawa K. Photometric stereo using constrained bivariate regression for general isotropic surfaces. In: 2014 IEEE conference on computer vision and pattern recognition (CVPR); 2014. p. 2187–94.
- [5] Ikehata S, Wipf D, Matsushita Y, Aizawa K. Robust photometric stereo using sparse regression. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2012. p. 318–25.
- [6] Horn BKP, Brooks MJ. Shape from shading. Cambridge, Massachusetts: The MIT Press; 1989 (ISBN: 0-262-08159-8).
- [7] Tankus A, Kiryati N. Photometric stereo under perspective projection. In: Tenth IEEE international conference on computer vision, ICCV 2005, vol. 1; 2005. p. 611–6.
- [8] Frankot R, Chellappa R. A method for enforcing integrability in shape from shading algorithms. *IEEE Trans Pattern Anal Mach Intell* 1988;10(4):439–51.
- [9] Barron JT, Malik J. Shape, albedo, and illumination from a single image of an unknown object. In: Proceedings of the 2012 IEEE conference on computer vision and pattern recognition (CVPR), CVPR '12. Washington, DC, USA: IEEE Computer Society; 2012. p. 334–41.
- [10] Abrams A, Hawley C, Pless R. Heliometric stereo: shape from sun position. In: Fitzgibbon A, Lazebnik S, Perona P, Sato Y, Schmid C, editors, computer vision, ECCV 2012, lecture notes in computer science. Berlin, Heidelberg: Springer; 2012. p. 357–70.
- [11] Chandraker M, Bai J, Ramamoorthi R. On differential photometric reconstruction for unknown, isotropic brdfs. *IEEE Trans Pattern Anal Mach Intell* 2013;35(12):2941–55.
- [12] Hansen MF, Atkinson GA, Smith LN, Smith ML. 3d face reconstructions from photometric stereo using near infrared and visible light. *Comput Vis Image Understand* 2010;114(8):942–51.
- [13] Wu C, Narasimhan SG, Jaramaz B. A multi-image shape-from-shading framework for near-lighting perspective endoscopes. *Int J Comput Vis* 2009;211–28.
- [14] Okatani T, Deguchi K. Shape reconstruction from an endoscope image by shape from shading technique for a point light source at the projection center. *Comput Vis Image Understand* 1997;66(2):119–31.
- [15] Deguchi K, Okatani T. Shape reconstruction from an endoscope image shape-from-shading technique for a point light source at the projection center. In: Workshop on MMBIA. IEEE Computer Society; 1996. p. 290–8.
- [16] Blinn JF. Models of light reflection for computer synthesized pictures. In: Proceedings of the 4th annual conference on computer graphics and interactive techniques, SIGGRAPH'77. New York, NY, USA: ACM; 1977. p. 192–8.
- [17] Ngan A, Durand F, Matusik W. Experimental analysis of brdf models. In: Proceedings of the sixteenth eurographics conference on rendering techniques, EGSR'05; 2005. p. 117–26.
- [18] Mecca R, Wetzler A, Bruckstein A, Kimmel R. Near field photometric stereo with point light sources. *SIAM J Imaging Sci* 2014;7(4):2732–70.
- [19] Hertzmann A, Seitz S. Example-based photometric stereo: shape reconstruction with general, varying brdfs. *IEEE Trans Pattern Anal Mach Intell* 2005;27(8):1254–64.
- [20] Barsky S, Petrou M. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans Pattern Anal Mach Intell* 2003;25(10):1239–52.
- [21] Yang Q, Wang S, Ahuja N. Real-time specular highlight removal using bilateral filtering. In: Daniilidis K, Maragos P, Paragios N, editors, computer vision, ECCV

- 2010, lecture notes in computer science, vol. 6314. Berlin, Heidelberg: Springer; 2010. p. 87–100.
- [22] Hyeonwoo K, Hailin J, Hadap S, Inso K. Specular reflection separation using dark channel prior. In: 2013 IEEE conference on computer vision and pattern recognition (CVPR); 2013. p. 1460–7.
- [23] Tan R, Ikeuchi K. Separating reflection components of textured surfaces using a single image. *IEEE Trans Pattern Anal Mach Intell* 2005;27(2):178–93.
- [24] Zickler T, Mallick SP, Kriegman DJ, Belhumeur PN. Color subspaces as photometric invariants. *Int J Comput Vis* 2008;79(1):13–30.
- [25] Shen H-L, Zheng Z-H. Real-time highlight removal using intensity ratio. *Appl Opt* 2013;52(19):4483–93.
- [26] Yang Q, Ahuja N. Surface reflectance and normal estimation from photometric stereo. *Comput Vis Image Understand* 2012;116(7):793–802.
- [27] Mecca R, Falcone M. Uniqueness and approximation of a photometric shape-from-shading model. *SIAM J Imaging Sci* 2013;6(1):616–59.
- [28] Mecca R, Tankus A, Wetzler A, Bruckstein A. A direct differential approach to photometric stereo with perspective viewing. *SIAM J Imaging Sci* 2014;7(2):579–612.
- [29] Chandraker M, Bai J, Ramamoorthi R. A theory of differential photometric stereo for unknown BRDFs. In: IEEE conference on computer vision and pattern recognition; 2011. p. 2505–12.
- [30] Alldrin NG, Kriegman DJ. Toward reconstructing surfaces with arbitrary isotropic reflectance: a stratified photometric stereo approach. In: International conference on computer vision; 2007. p. 1–8.
- [31] Iwahori Y, Sugie H, Ishii N. Reconstructing shape from shading images under point light source illumination. In: Proceedings of the 10th international conference on pattern recognition, vol. 1; 1990. p. 83–7.
- [32] Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction. In: Proceedings of the fourth eurographics symposium on geometry processing, SGP'06. Aire-la-Ville, Switzerland: Eurographics Association; 2006. p. 61–70.
- [33] Albarelli A, Rodolà E, Torsello A. Fast and accurate surface alignment through an isometry-enforcing game. *Pattern Recognit* 2015;48(7):2209–26.
- [34] Torsello A, Rodolà E, Albarelli A. Multiview registration via graph diffusion of dual quaternions. In: 2011 IEEE conference on computer vision and pattern recognition (CVPR); 2011. p. 2441–8.
- [35] Bergamasco F, Albarelli A, Rodolà E, Torsello A. Can a fully unconstrained imaging model be applied effectively to central cameras? In: 2013 IEEE conference on computer vision and pattern recognition (CVPR); 2013. p. 1391–8.