LED-based Photometric Stereo: Modeling, Calibration and Numerical Solution

Yvain Quéau, Bastien Durix, Tao Wu, Daniel Cremers, François Lauze, Jean-Denis Durou

▶ **To cite this version:**

Yvain Quéau, Bastien Durix, Tao Wu, Daniel Cremers, François Lauze, et al.. LED-based Photometric Stereo: Modeling, Calibration and Numerical Solution. 2017. hal-01399607v2

HAL Id: hal-01399607

https://hal.science/hal-01399607v2

Preprint submitted on 4 Jul 2017 (v2), last revised 4 Sep 2017 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Noname manuscript No.
(will be inserted by the editor)

LED-based Photometric Stereo: Modeling, Calibration and Numerical Solution

Yvain Quéau · Bastien Durix · Tao Wu · Daniel Cremers · François Lauze · Jean-Denis Durou

the date of receipt and acceptance should be inserted later

Abstract We conduct a thorough study of photometric stereo under nearby point light source illumination, from modeling to numerical solution, through calibration. In the classical formulation of photometric stereo, the luminous fluxes are assumed to be directional, which is very difficult to achieve in practice. Rather, we use light-emitting diodes (LEDs) to illuminate the scene to reconstruct. Such point light sources are very convenient to use, yet they yield a more complex photometric stereo model which is arduous to solve. We first derive in a physically sound manner this model, and show how to calibrate its parameters. Then, we discuss two state-of-the-art numerical solutions. The first one alternatingly estimates the albedo and the normals, and then integrates the normals into a depth map. It is shown empirically to be independent from the initialization, but convergence of this sequential approach is not established. The second one directly recovers the depth, by formulating photometric stereo as a system of PDEs which are partially linearized using image ratios. Although the sequential approach is avoided, initialization matters a lot and convergence is not established either. Therefore, we introduce a provably convergent alternating reweighted least-squares scheme for solving the original system of PDEs, without resorting to image ratios for linearization. Finally, we extend this study to the case of RGB images.

Keywords 3D-reconstruction · Photometric stereo · Point light sources · Variational methods · Alternating reweighted least-squares.

Yvain Quéau
Department of Computer Science
Technical University of Munich, Germany
E-mail: yvain.queau@tum.de

Bastien Durix
IRIT, UMR CNRS 5505
Université de Toulouse, France
E-mail: bastien.durix@enseeiht.fr

Tao Wu
Department of Computer Science
Technical University of Munich, Germany
E-mail: tao.wu@tum.de

Daniel Cremers
Department of Computer Science
Technical University of Munich, Germany
E-mail: cremers@tum.de

François Lauze
Department of Computer Science
University of Copenhagen, Denmark
E-mail: francois@di.ku.dk

Jean-Denis Durou
IRIT, UMR CNRS 5505
Université de Toulouse, France
E-mail: durou@irit.fr

1 Introduction

3D-reconstruction is one of the most important goals of computer vision. Among the many techniques which can be used to accomplish this task, shape-from-shading [27] and photometric stereo [64] are *photometric techniques*, as they use the relationship between the gray levels of the image, the shape of the scene, supposedly opaque, its reflectance and the luminous flux that illuminates it.

We first introduce some notations that will be used throughout this paper. We describe a point $\mathbf{x}$ on the scene surface by its coordinates $[x, y, z]^\top$ in a frame originating from the optical center C of the camera, such that the plane Cxy is parallel to the image plane and the Cz axis coincides with the optical axis and faces the scene (cf. Fig. 1). The coordinates $[u, v]^\top$ of a

point $\mathbf{p}$ in the image (pixel) are relative to a frame Ouv whose origin is the principal point O, and whose axes Ou and Ov are parallel to Cx and Cy, respectively. If f refers to the focal length, the conjugation relationship between $\mathbf{x}$ and $\mathbf{p}$ is written, in perspective projection:

$$\begin{cases} x = \dfrac{z}{f}\, u, \\ y = \dfrac{z}{f}\, v. \end{cases} \tag{1.1}$$

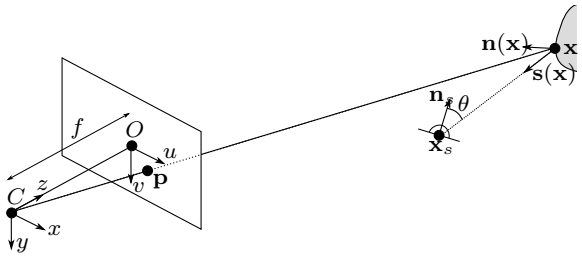


Fig. 1 Schematic representation of the geometric setup. A point $\mathbf{x} = [x, y, z]^\top \in \mathbb{R}^3$ on the scene surface and a pixel $\mathbf{p} = [u, v]^\top \in \mathbb{R}^2$ in the image plane are conjugated according to Eq. (1.1). Eq. (2.19) states that, when the scene is illuminated by a LED located in $\mathbf{x}_s \in \mathbb{R}^3$, the gray level $I(\mathbf{p})$ of the pixel $\mathbf{p}$ conjugated to $\mathbf{x}$ is a function of the angle between the lighting vector $\mathbf{s}(\mathbf{x})$ and the normal $\mathbf{n}(\mathbf{x})$ to the surface in $\mathbf{x}$ (Lambertian reflectance), of the angle θ between the principal direction $\mathbf{n}_s$ of the LED and $\mathbf{s}(\mathbf{x})$ (anisotropy), and of the distance $\|\mathbf{x} - \mathbf{x}_s\|$ between the surface point and the light source location (inverse-of-square falloff).

3D-reconstruction consists in estimating, in each pixel $\mathbf{p}$ of a part Ω of the image domain, its conjugate point $\mathbf{x}$ in 3D-space. Eq. (1.1) shows that it suffices to find the depth z to determine $\mathbf{x} = [x, y, z]^\top$ from $\mathbf{p} = [u, v]^\top$. The only unknown of the problem is thus the *depth map* z, which is defined as follows:

$$\begin{aligned} z : \quad \Omega \subset \mathbb{R}^2 \quad &\to \mathbb{R}^+ \\ \mathbf{p} = [u, v]^\top &\mapsto z(\mathbf{p}). \end{aligned} \tag{1.2}$$

We are interested in this article in 3D-reconstruction of *Lambertian surfaces* by photometric stereo. The *reflectance* in a point of such a surface is completely characterized by a coefficient ρ, called *albedo*, which is 0 if the point is black and 1 if it is white. Photometric stereo is nothing else than an extension of shape-from-shading: instead of a single image, the former uses $m \geqslant 3$ shots I^i, $i \in [1, m]$, taken from the same angle, but under varying lighting. Considering multiple images allows to circumvent the difficulties of shape-from-shading: photometric stereo techniques are able to *unambiguously* estimate the shape as well as the albedo i.e., without any resort to additional prior.

A parallel and uniform illumination can be characterized by a vector $\mathbf{s} \in \mathbb{R}^3$ oriented towards the light source, whose norm is equal to the *luminous flux density*. We call $\mathbf{s}$ the *lighting vector*. For a Lambertian surface, the classical modeling of photometric stereo is written, in each pixel $\mathbf{p} \in \Omega$, as the following system[1]:

$$I^i(\mathbf{p}) = \rho(\mathbf{x})\ \mathbf{s}^i \cdot \mathbf{n}(\mathbf{x}), \qquad i \in [1, m], \tag{1.3}$$

where $I^i(\mathbf{p})$ denotes the gray level of $\mathbf{p}$ under a parallel and uniform illumination characterized by the lighting vector $\mathbf{s}^i$, $\rho(\mathbf{x})$ denotes the albedo in the point $\mathbf{x}$ conjugate to $\mathbf{p}$, and $\mathbf{n}(\mathbf{x})$ denotes the unit-length outgoing normal to the surface in this point. Since there is a bijective correspondence between the points $\mathbf{x}$ and the pixels $\mathbf{p}$, we write for convenience $\rho(\mathbf{p})$ and $\mathbf{n}(\mathbf{p})$, in lieu of $\rho(\mathbf{x})$ and $\mathbf{n}(\mathbf{x})$. Introducing the notation $\mathbf{m}(\mathbf{p}) = \rho(\mathbf{p})\,\mathbf{n}(\mathbf{p})$, (1.3) can be rewritten in matrix form:

$$\mathbf{I}(\mathbf{p}) = \mathbf{S}\,\mathbf{m}(\mathbf{p}), \tag{1.4}$$

where vector $\mathbf{I}(\mathbf{p}) \in \mathbb{R}^m$ and matrix $\mathbf{S} \in \mathbb{R}^{m \times 3}$ are defined as follows:

$$\mathbf{I}(\mathbf{p}) = \begin{bmatrix} I^1(\mathbf{p}) \\ \vdots \\ I^m(\mathbf{p}) \end{bmatrix} \qquad \text{and} \qquad \mathbf{S} = \begin{bmatrix} \mathbf{s}^{1\top} \\ \vdots \\ \mathbf{s}^{m\top} \end{bmatrix}. \tag{1.5}$$

As soon as $m \geqslant 3$ *non-coplanar* lighting vectors are used, matrix $\mathbf{S}$ has rank 3. The (unique) least-squares solution of System (1.4) is then given by

$$\mathbf{m}(\mathbf{p}) = \mathbf{S}^\dagger\,\mathbf{I}(\mathbf{p}), \tag{1.6}$$

where $\mathbf{S}^\dagger$ is the *pseudo-inverse* of $\mathbf{S}$. From this solution, we easily deduce the albedo and the normal:

$$\rho(\mathbf{p}) = \|\mathbf{m}(\mathbf{p})\| \qquad \text{and} \qquad \mathbf{n}(\mathbf{p}) = \frac{\mathbf{m}(\mathbf{p})}{\|\mathbf{m}(\mathbf{p})\|}. \tag{1.7}$$

The normal field estimated in such a way must eventually be *integrated* so as to obtain the depth map, knowing that the boundary conditions, the shape of domain Ω as well as depth discontinuities significantly complicate this task [54].

To ensure lighting directionality, as required by Model (1.3), it is necessary to achieve a complex optical setup [44]. It is much easier to use light-emitting diodes (LEDs) as light sources, but with this type of light sources, we should expect significant changes in the modeling, and therefore in the numerical solution. The aim of our work is to conduct a comprehensive and detailed study of photometric stereo under point light source illumination such as LEDs.

[1] Equality (1.3) is in fact a proportionality relationship: see the expression (2.10) of $I(\mathbf{p})$.

Related works – Modeling the luminous flux emitted by a LED is a well-studied problem, see for instance [45]. One model which is frequently considered in computer vision is that of nearby point light sources. This model involves an inverse-of-square law for describing the attenuation of lighting intensity with respect to distance, which has long been identified as a key feature for solving shape-from-shading [31] and photometric stereo [12]. Attenuation with respect to the deviation from the principal direction of the source (anisotropy) can also be considered [7].

If the surface to reconstruct lies in the vicinity of a plane, it is possible to capture a map of these attenuation coefficients using a white planar reference object. Conventional photometric stereo [64] can then be applied to the images compensated by the attenuation maps [3,39,61]. Otherwise, it is necessary to include these terms in the photometric stereo model, which yields a nonlinear inverse problem to be solved.

This is easier to achieve if the parameters of the illumination model have been calibrated beforehand. Lots of methods exist for estimating a source location [1,4,11,17,21,53,59,62]. Such methods typically triangulate this location by resorting to specular spheres, see for instance the recent work [36]. Calibrating anisotropy is a more challenging problem, which was tackled recently in [47,67] by using images of a planar surface. Some photometric stereo methods also circumvent calibration by (partly or completely) automatically inferring lighting during 3D-reconstruction [35,36,37,43,50,57].

Still, even in the calibrated case, designing numerical schemes for solving photometric stereo under nearby point light sources remains difficult. When only two images are considered, the photometric stereo model can be simplified using image ratios. This yields a quasi-linear PDE [41,42] which can be solved by provably convergent front propagation techniques, provided that a boundary condition is known. To improve robustness, this strategy has been adapted to the multi-images case in [37,38,40,55], using variational methods. However, convergence guarantees are lost. Instead of considering such a differential approach, another class of methods [2,8,13,28,33,46,50,69] rather modify the classical photometric stereo framework [64], by alternatingly estimating the normals and the albedo, integrating the normals into a depth map, and updating the lighting based on the current depth. Yet, no convergence guarantee does exist. A method based on mesh deformation has also been proposed in [68], but convergence is not established either.

Contributions – In contrast to existing works which focus either on modeling, calibrating or solving photometric stereo with near point light sources such as LEDs, the objective of this article is to propose a comprehensive study of all these aspects of the problem. Building upon our previous conference papers [55,57,58], we introduce the following innovations:

- We present in Section 2 an accurate model for photometric stereo under point light source illumination. As in recent works [37,38,41,42,40,46,47,67], this model takes into account the nonlinearities due to distance and to the anisotropy of the LEDs. Yet, it also clarifies the notions of albedo and of source intensity, which are shown to be relative to a reference albedo and to the radiometric parameters of the camera. This section also introduces a practical calibration procedure for the location, the orientation and the relative intensity of a LED.

- Section 3 reviews and improves two state-of-the-art numerical solutions in several manners. We first modify the alternating method [2,8,13,28,33,46,50,69], by introducing an estimation of the shape scale, in order to recover the absolute depth without any prior. We then study the PDE-based approach which employs image ratios for eliminating the nonlinearities [37,38,40,55], and empirically show that local minima can be avoided by employing an augmented Lagrangian strategy. Nevertheless, neither of these state-of-the-art methods is provably convergent.

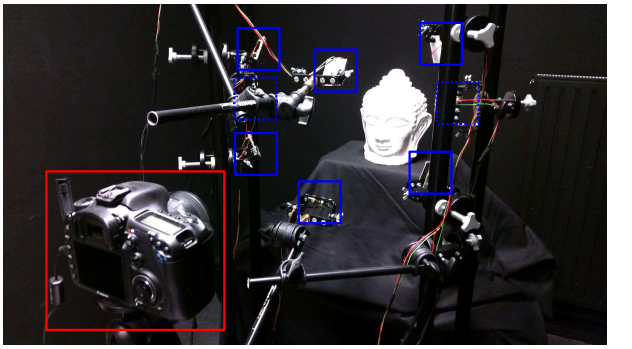
- Therefore, we introduce in Section 4 a new, provably convergent method, inspired by the one recently proposed in [57]. It is based on a tailored alternating reweighted least-squares scheme for approximately solving the non-linearized system of PDEs. Following [58], we further show that this method is easily extended in order to address shadows and specularities.

- In Section 5, we build upon the analysis conducted in [55] in order to tackle the case of RGB-valued images, before concluding and suggesting several future research directions in Section 6.

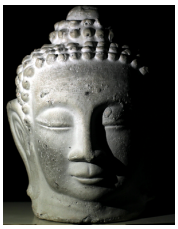
2 Photometric Stereo under Point Light Source Illumination

Conventional photometric stereo [64] assumes that the primary luminous fluxes are parallel and uniform, which is difficult to guarantee. It is much easier to illuminate a scene with LEDs.

Keeping this in mind, we have developed a photometric stereo-based setup for 3D-reconstruction of faces, which includes $m = 8$ LEDs[2] located at about 30 cm of the scene (see Fig. 2-a). The face is photographed by a Canon EOS 7D camera with focal length $f = 35\ mm$. Triggering the shutter in burst mode, while synchronically lighting the LEDs, provides us with $m = 8$ images such as those of Figs. 2-b, 2-c and 2-d. In this section, we model the luminous flux emitted by a LED, and show how to estimate the parameters involved. We finally show how to model photometric stereo under point light source illumination.



(a)



(b)

(c)

(d)

Fig. 2 (a) Our photometric stereo-based experimental setup for 3D-reconstruction of faces using a Canon EOS 7D camera (highlighted in red) and $m = 8$ LEDs (highlighted in blue). The walls are painted in black in order to avoid the reflections between the scene and the environment. (b-c-d) Three out of the $m = 8$ images obtained by this setup.

2.1 Modeling the Luminous Flux Emitted by a LED

For the LEDs we use, the characteristic illuminating volume is of the order of one cubic millimeter. Therefore, in comparison with the scale of a face, each LED can be seen as a point source located at a point $\mathbf{x}_s \in \mathbb{R}^3$. At any point $\mathbf{x} \in \mathbb{R}^3$, the lighting vector $\mathbf{s}(\mathbf{x})$ is necessarily radial i.e., collinear with the unit-length vector $\mathbf{u}_r = \frac{\mathbf{x}-\mathbf{x}_s}{\|\mathbf{x}-\mathbf{x}_s\|}$. Using spherical coordinates (r, θ, ϕ) of $\mathbf{x}$ in a frame having $\mathbf{x}_s$ as origin, it is written

$$\mathbf{s}(\mathbf{x}) = -\frac{\Phi(\theta,\phi)}{r^2}\,\mathbf{u}_r, \tag{2.1}$$

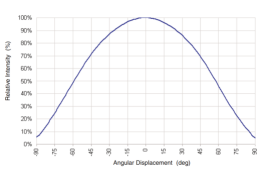
where $\Phi(\theta,\phi) \geqslant 0$ denotes the *intensity* of the source[3], and the $1/r^2$ attenuation is a consequence of the conservation of luminous energy in a non-absorbing medium. Vector $\mathbf{s}(\mathbf{x})$ is purposely oriented in the opposite direction from that of the light, in order to simplify the writing of the Lambertian model.

Model (2.1) is very general. We could project the intensity $\Phi(\theta,\phi)$ on the *spherical harmonics* basis, which allowed Basri et al. to model the luminous flux in the case of uncalibrated photometric stereo [6]. We could also sample $\Phi(\theta,\phi)$ in the vicinity of a plane, using a plane with known reflectance [3,39,61].

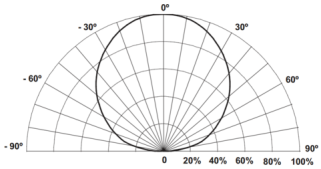
Using the specific characteristics of LEDs may lead to a more accurate model. Indeed, most of the LEDs emit a luminuous flux which is invariant by rotation around a *principal direction* indicated by a unit-length vector $\mathbf{n}_s$ [45]. If θ is defined relatively to $\mathbf{n}_s$, this means that $\Phi(\theta,\phi)$ is independent from ϕ. The lighting vector in $\mathbf{x}$ induced by a LED located in $\mathbf{x}_s$ is thus written

$$\mathbf{s}(\mathbf{x}) = \frac{\Phi(\theta)}{\|\mathbf{x}_s-\mathbf{x}\|^2}\,\frac{\mathbf{x}_s-\mathbf{x}}{\|\mathbf{x}_s-\mathbf{x}\|}. \tag{2.2}$$

The dependency on θ of the intensity Φ characterizes the anisotropy of the LED. The function $\Phi(\theta)$ is generally decreasing over $[0, \pi/2]$ (cf. Fig. 3).



(a)



(b)

Fig. 3 Intensity patterns of the LEDs used (source: http://www.lumileds.com/uploads/28/DS64-pdf). (a) Anisotropy function $\Phi(\theta)/\Phi_0$ as a function of θ. (b) Polar representation. These diagrams show us that $\theta_{1/2} = \pi/3$, which corresponds to $\mu = 1$ according to Eq. (2.4) (Lambertian source).

An anisotropy model satisfying this constraint is that of "imperfect Lambertian source":

$$\Phi(\theta) = \Phi_0\,\cos^{\mu}\theta, \tag{2.3}$$

which contains two parameters $\Phi_0 = \Phi(0)$ and $\mu \geqslant 0$, and models both isotropic sources ($\mu = 0$) and Lambertian sources ($\mu = 1$). Model (2.3) is empirical, and more

[2] We use white LUXEON Rebel LEDs: http://www.luxeonstar.com/luxeon-rebel-leds.

[3] The intensity is expressed in lumen per steradian ($lm.sr^{-1}$), or candela (cd).

elaborate models are sometimes considered [45], yet it has already been used in photometric stereo [37,38,40,41,46,47,57,67], including the case where all the LEDs are arranged on a plane parallel to the image plane i.e., when $\mathbf{n}_s = [0,0,1]^\top$ [42]. This model has proven itself and, moreover, LEDs manufacturers provide the angle $\theta_{1/2}$ such that $\Phi(\theta_{1/2}) = \Phi_0/2$, from which we deduce, using (2.3), the value of μ:

$$\mu = -\frac{\log(2)}{\log(\cos\theta_{1/2})}. \tag{2.4}$$

As shown in Fig. 3, the angle $\theta_{1/2}$ is $\pi/3$ for the LEDs we use. From Eq. (2.4), we deduce that $\mu = 1$, which means that these LEDs are Lambertian. Plugging the expression (2.3) of $\Phi(\theta)$ into (2.2), we obtain

$$\mathbf{s}(\mathbf{x}) = \Phi_0 \cos^\mu\theta \frac{\mathbf{x}_s - \mathbf{x}}{\|\mathbf{x}_s - \mathbf{x}\|^3}, \tag{2.5}$$

where we explicitly keep μ to address the most general case. Model (2.5) thus includes seven parameters: three for the coordinates of $\mathbf{x}_s$, two for the unit vector $\mathbf{n}_s$, plus Φ_0 and μ. Note that $\mathbf{n}_s$ appears in this model through the angle θ.

In its uncalibrated version, photometric stereo allows the 3D-reconstruction of a scene without knowing the lighting. Uncalibrated photometric stereo has been widely studied, including the case of nearby point light sources [28,35,43,50,69], but if this is possible, we should rather calibrate the lighting[4].

2.2 Calibrating the Luminous Flux Emitted by a LED

Most calibration methods for a point light source [1,4,11,17,21,53,59,62] do not take into account the attenuation of the luminous flux density as a function of the distance to the source, nor the possible anisotropy of the source, which may lead to relatively imprecise results. To our knowledge, there are few calibration procedures taking into account these phenomena. In [67], Xie et al. use a single pattern, which is partially bright and partially Lambertian, to calibrate a LED. We intend to improve this procedure using two patterns, one bright and the other Lambertian. The brilliant one will be used to determine the location of the LEDs by triangulation, and the Lambertian one to determine some other parameters by minimizing the reprojection error, as recently proposed by Pintus et al. in [52].

[4] It is also necessary to calibrate the camera, since the 3D-frame is attached to it. We assume that this has been made beforehand.

Specular Spherical Calibration Pattern – The location $\mathbf{x}_s$ of a LED can be determined by triangulation. In [53], Powell et al. advocate the use of a spherical mirror. To estimate the locations of the $m = 8$ LEDs for our setup, we use a billiard ball. Under perspective projection, the edge of the silhouette of a sphere is an ellipse, which we detect using a dedicated algorithm [51]. It is then easy to determine the 3D-coordinates of any point on the surface, as well as its normal, since the radius of the billiard ball is known. For each pose of the billiard ball, detecting the reflection of the LED allows us to determine, by reflecting the line of sight on the spherical mirror, a line in 3D-space passing through $\mathbf{x}_s$. In theory, two poses of the billiard ball are enough to estimate $\mathbf{x}_s$, even if two lines in 3D-space do not necessarily intersect, but the use of ten poses improves the robustness of the estimation. A similar procedure has recently been used in [36].

Lambertian Model – To estimate the principal direction $\mathbf{n}_s$ and the intensity Φ_0 in Model (2.5), we use a Lambertian calibration pattern. A surface is Lambertian if the apparent clarity of any point $\mathbf{x}$ located on it is independent from the viewing angle. The *luminance* $L(\mathbf{x})$, which is equal to the luminous flux emitted per unit of solid angle and per unit of apparent surface, is independent from the direction of emission. However, the luminance is not characteristic of the surface, as it depends on the *illuminance* $E(\mathbf{x})$ (denoted E from French "éclairement"), that is to say on the luminous flux per unit area received by the surface in $\mathbf{x}$. The relationship between luminance and illuminance[5] is written, for a Lambertian surface:

$$L(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\pi} E(\mathbf{x}), \tag{2.6}$$

where the albedo $\rho(\mathbf{x}) \in \mathbb{R}^+$ is defined as the proportion of luminous energy which is reemitted i.e., $\rho(\mathbf{x}) = 1$ if $\mathbf{x}$ is white, and $\rho(\mathbf{x}) = 0$ if it is black.

The parameter $\rho(\mathbf{x})$ is enough to characterize the reflectance (or BRDF) of a Lambertian surface. In addition, the illuminance at a point $\mathbf{x}$ of a (not necessarily Lambertian) surface with normal $\mathbf{n}(\mathbf{x})$, lit by the lighting vector $\mathbf{s}(\mathbf{x})$, is written

$$E(\mathbf{x}) = \mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}). \tag{2.7}$$

Focusing the camera on a point $\mathbf{x}$ of the scene surface, the illuminance $\epsilon(\mathbf{p})$ of the image plane, at pixel $\mathbf{p}$ conjugate to $\mathbf{x}$, is related to the luminance $L(\mathbf{x})$ by the following "almost linear" relationship [26]:

$$\epsilon(\mathbf{p}) = \beta \cos^4\alpha(\mathbf{p}) L(\mathbf{x}), \tag{2.8}$$

[5] A luminance is expressed in $lm.m^{-2}.sr^{-1}$ (or $cd.m^{-2}$), an illuminance in $lm.m^{-2}$, or lux (lx).

where β is a proportionality coefficient characterizing the clarity of the image, which depends on several factors such as the lens aperture, the magnification, etc. Regarding the factor $\cos^4\alpha(\mathbf{p})$, where $\alpha(\mathbf{p})$ is the angle between the line of sight and the optical axis, it is responsible for darkening at the periphery of the image. This darkening phenomenon should not be confused with vignetting, since it occurs even with ideal lenses [16].

With current photosensitive receptors, the *gray level* $J(\mathbf{p})$ at pixel $\mathbf{p}$ is almost proportional[6] to its illuminance $\epsilon(\mathbf{p})$, except of course in case of saturation. Denoting γ this coefficient of quasi-proportionality, and combining equalities (2.6), (2.7) and (2.8), we get the following expression of the gray level in a pixel $\mathbf{p}$ conjugate to a point $\mathbf{x}$ located on a Lambertian surface:

$$J(\mathbf{p}) = \gamma\,\beta\,\cos^4\alpha(\mathbf{p})\,\frac{\rho(\mathbf{x})}{\pi}\,\mathbf{s}(\mathbf{x})\cdot\mathbf{n}(\mathbf{x}). \tag{2.9}$$

We have already mentioned that there is a bijective correspondence between a point $\mathbf{x}$ and its conjugate pixel $\mathbf{p}$, which allows us to denote $\rho(\mathbf{p})$ and $\mathbf{n}(\mathbf{p})$ instead of $\rho(\mathbf{x})$ and $\mathbf{n}(\mathbf{x})$. As the factor $\cos^4\alpha(\mathbf{p})$ is easy to calculate in each pixel $\mathbf{p}$ of the photosensitive receptor, since $\cos\alpha(\mathbf{p}) = \frac{f}{\sqrt{\|\mathbf{p}\|^2+f^2}}$, we can very easily compensate for this source of darkening and will manipulate now the "corrected gray level":

$$I(\mathbf{p}) = \frac{J(\mathbf{p})}{\cos^4\alpha(\mathbf{p})} = \gamma\,\beta\,\frac{\rho(\mathbf{p})}{\pi}\,\mathbf{s}(\mathbf{x})\cdot\mathbf{n}(\mathbf{p}). \tag{2.10}$$

Lambertian Planar Calibration Pattern – To estimate the parameters $\mathbf{n}_s$ and Φ_0 in Model (2.6) i.e., to achieve *photometric calibration*, we use a second calibration pattern consisting of a checkerboard printed on a white paper sheet, which is itself stuck on a plane (cf. Fig. 4), with the hope that the unavoidable outliers to the Lambertian model will not influence the accuracy of the estimates too much.

The use of a convex calibration pattern (plane, in this case) has a significant advantage: the lighting vector $\mathbf{s}(\mathbf{x})$ at any point $\mathbf{x}$ of the surface is purely *primary* i.e., it is only due to the light source, without "bouncing" on other parts of the surface of the target, provided that the walls and surrounding objects are covered in black (see Fig. 2-a). Thanks to this observation, we can replace the lighting vector $\mathbf{s}(\mathbf{x})$ in Eq. (2.10) by the expression (2.5) which models the luminous flux emitted by a LED. From (2.5) and (2.10), we deduce the gray level $I(\mathbf{p})$ of the image of a point $\mathbf{x}$ located on this calibration pattern, illuminated by a LED:

$$I(\mathbf{p}) = \gamma\,\beta\,\frac{\rho(\mathbf{p})}{\pi}\,\Phi_0\,\cos^\mu\theta\,\frac{(\mathbf{x}_s-\mathbf{x})\cdot\mathbf{n}(\mathbf{p})}{\|\mathbf{x}_s-\mathbf{x}\|^3}. \tag{2.11}$$

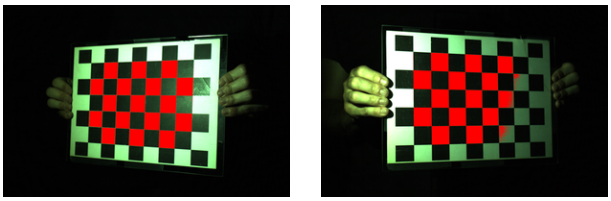


Fig. 4 Two out of the q poses of the planar Lambertian calibration pattern used for the photometric calibration of the LEDs. The parts of the white cells highlighted in red are used for estimating the principal directions and the intensities of the LEDs.

If $q \geqslant 3$ poses of the checkerboard are used, numerous algorithms exist for unambiguously estimating the coordinates of the points $\mathbf{x}^j$ of the pattern, for different poses $j \in [1,q]$. These algorithms also allow the estimation of the q normals $\mathbf{n}^j$ (we omit the dependency in $\mathbf{p}$ of $\mathbf{n}^j$, since the pattern is planar), and the intrinsic parameters of the camera[7]. As for the albedo, if the use of white paper does not guarantee that $\rho(\mathbf{p}) \equiv 1$, it still seems reasonable to assume $\rho(\mathbf{p}) \equiv \rho_0$ i.e., to assume a uniform albedo in the white cells. We can group all the multiplicative coefficients of the right hand side of (2.11) into one coefficient

$$\Psi = \gamma\,\beta\,\frac{\rho_0}{\pi}\,\Phi_0. \tag{2.12}$$

With this definition, and knowing that θ is the angle between vectors $\mathbf{n}_s$ and $\mathbf{x}-\mathbf{x}_s$, Eq. (2.11) can be rewritten, in a pixel $\mathbf{p}$ of the set Ω^j containing the white pixels of the checkerboard in the j^{th} pose (these pixels are highlighted in red in the images of Fig. 4):

$$I^j(\mathbf{p}) = \Psi\left[\mathbf{n}_s\cdot\frac{\mathbf{x}^j-\mathbf{x}_s}{\|\mathbf{x}^j-\mathbf{x}_s\|}\right]^\mu\frac{(\mathbf{x}_s-\mathbf{x}^j)\cdot\mathbf{n}^j}{\|\mathbf{x}_s-\mathbf{x}^j\|^3}. \tag{2.13}$$

Since $\mathbf{x}_s$ is already estimated, and the value of μ is (indirectly) provided by the manufacturer, the only unknowns in Eq. (2.13) are Ψ and $\mathbf{n}_s$. Two cases may occur:

- If the LED to calibrate is isotropic i.e., if $\mu = 0$, then it is useless to estimate $\mathbf{n}_s$, and Ψ can be estimated in a least-squares sense, by solving

$$\min_{\Psi}\sum_{j=1}^{q}\sum_{\mathbf{p}\in\Omega^j}\left[I^j(\mathbf{p}) - \Psi\,\frac{(\mathbf{x}_s-\mathbf{x}^j)\cdot\mathbf{n}^j}{\|\mathbf{x}_s-\mathbf{x}^j\|^3}\right]^2, \tag{2.14}$$

[6] Using RAW images is necessary to ensure such proportionality.

[7] To perform these operations, which allow us to perform the geometric calibration of the camera, we use the *Computer Vision* toolbox from Matlab.

whose solution is given by

$$\Psi = \frac{\displaystyle\sum_{j=1}^{q} \sum_{\mathbf{p}\in\Omega^j} I^j(\mathbf{p}) \, \frac{(\mathbf{x}_s - \mathbf{x}^j)\cdot\mathbf{n}^j}{\|\mathbf{x}_s - \mathbf{x}^j\|^3}}{\displaystyle\sum_{j=1}^{q} \sum_{\mathbf{p}\in\Omega^j} \left[\frac{(\mathbf{x}_s - \mathbf{x}^j)\cdot\mathbf{n}^j}{\|\mathbf{x}_s - \mathbf{x}^j\|^3}\right]^2}. \tag{2.15}$$

- In the other case (if $\mu > 0$), Eq. (2.13) can be rewritten

$$\underbrace{\Psi^{1/\mu}\,\mathbf{n}_s}_{\mathbf{m}_s} \cdot (\mathbf{x}^j - \mathbf{x}_s) = \left[I^j(\mathbf{p})\,\frac{\|\mathbf{x}_s - \mathbf{x}^j\|^{3+\mu}}{(\mathbf{x}_s - \mathbf{x}^j)\cdot\mathbf{n}^j}\right]^{1/\mu}. \tag{2.16}$$

The least-squares estimation of vector $\mathbf{m}_s$ defined in (2.16) is thus written

$$\min_{\mathbf{m}_s} \sum_{j=1}^{q} \sum_{\mathbf{p}\in\Omega^j} \left[\mathbf{m}_s \cdot (\mathbf{x}^j - \mathbf{x}_s) - \left[I^j(\mathbf{p})\,\frac{\|\mathbf{x}_s - \mathbf{x}^j\|^{3+\mu}}{(\mathbf{x}_s - \mathbf{x}^j)\cdot\mathbf{n}^j}\right]^{1/\mu}\right]^2. \tag{2.17}$$

This linear least-squares problem can be solved, for instance, using the pseudo-inverse. From this estimate, we easily deduce those of parameters $\mathbf{n}_s$ and Ψ:

$$\mathbf{n}_s = \frac{\mathbf{m}_s}{\|\mathbf{m}_s\|} \qquad \text{and} \qquad \Psi = \|\mathbf{m}_s\|^\mu. \tag{2.18}$$

In both cases, it is impossible to deduce from the estimate of Ψ that of Φ_0, because in the definition (2.12) of Ψ, the product $\gamma\,\beta\,\frac{\rho_0}{\pi}$ is unknown. However, since this product is the same for all LEDs (deactivating all automatic settings of the camera makes β and γ constant), this is not a problem for photometric stereo.

Fig. 5 shows two views of a schematic representation of the experimental setup of Fig. 2, where LEDs parameters were estimated using our calibration procedure.

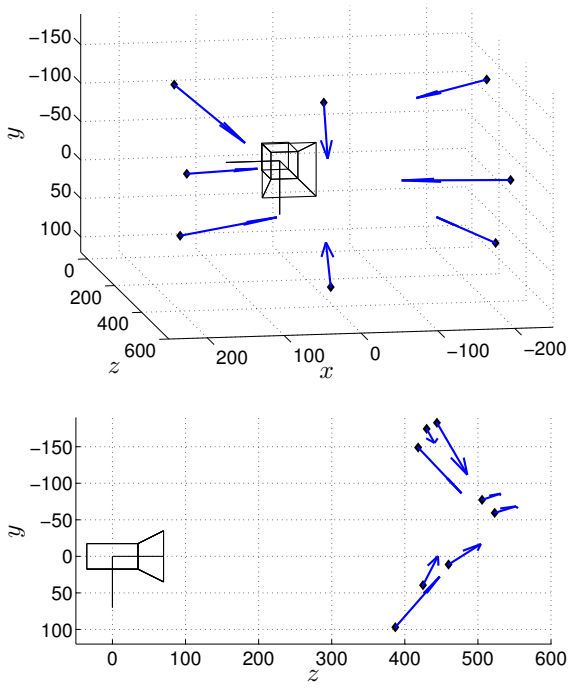


Fig. 5 Two views of a schematic representation of the experimental setup of Fig. 2-a, where the arrows symbolize the LEDs (unit mm). The camera center is located in $(0,0,0)$. A black marker characterizes the location $\mathbf{x}_s$ of each LED, the orientation of a blue arrow its principal direction $\mathbf{n}_s$, and the length of this arrow its intensity Ψ (up to a common factor).

2.3 Modeling Photometric Stereo with Point Light Sources

If the luminous flux emitted by a LED is described by Model (2.5), then we obtain from (2.11) and (2.12) the following equation for the gray level at pixel $\mathbf{p}$:

$$I(\mathbf{p}) = \Psi\,\frac{\rho(\mathbf{p})}{\rho_0}\left[\mathbf{n}_s \cdot \frac{\mathbf{x}-\mathbf{x}_s}{\|\mathbf{x}-\mathbf{x}_s\|}\right]^\mu \frac{(\mathbf{x}_s-\mathbf{x})\cdot\mathbf{n}(\mathbf{p})}{\|\mathbf{x}_s-\mathbf{x}\|^3}. \tag{2.19}$$

Thus, we must solve, in each pixel $\mathbf{p}\in\Omega$, the following system of equations, for $i\in[1,m]$:

$$I^i(\mathbf{p}) = \Psi^i\,\frac{\rho(\mathbf{p})}{\rho_0}\left[\mathbf{n}_s^i \cdot \frac{\mathbf{x}-\mathbf{x}_s^i}{\|\mathbf{x}-\mathbf{x}_s^i\|}\right]^{\mu^i} \frac{(\mathbf{x}_s^i-\mathbf{x})\cdot\mathbf{n}(\mathbf{p})}{\|\mathbf{x}_s^i-\mathbf{x}\|^3}, \tag{2.20}$$

where the values of parameters $\mathbf{x}_s^i$, $\mathbf{n}_s^i$ and Ψ^i, $i\in[1,m]$, are known, thanks to the calibration procedure described in Section 2.2. Let us introduce a new definition of the albedo relative to the albedo ρ_0 of the Lambertian calibration pattern:

$$\overline{\rho}(\mathbf{p}) = \frac{\rho(\mathbf{p})}{\rho_0}. \tag{2.21}$$

Since the parameters μ^i, $i\in[1,m]$, are provided by the manufacturer, the only unknowns in System (2.20) are the depth z of the 3D-point $\mathbf{x}$ conjugate to $\mathbf{p}$, its (relative) albedo $\overline{\rho}(\mathbf{p})$ and its normal $\mathbf{n}(\mathbf{p})$. To solve this system, the introduction of the auxiliary variable $\overline{\mathbf{m}}(\mathbf{p}) = \overline{\rho}(\mathbf{p})\,\mathbf{n}(\mathbf{p})$ may seem relevant, since this vector is not constrained to have unit-length, but we will see that this trick loses part of its interest. Defining the following m vectors, $i\in[1,m]$:

$$\mathbf{t}^i(\mathbf{x}) = \Psi^i\left[\mathbf{n}_s^i \cdot \frac{\mathbf{x}-\mathbf{x}_s^i}{\|\mathbf{x}-\mathbf{x}_s^i\|}\right]^{\mu^i} \frac{\mathbf{x}_s^i-\mathbf{x}}{\|\mathbf{x}_s^i-\mathbf{x}\|^3}, \tag{2.22}$$

then System (2.20) is rewritten:

$$I^i(\mathbf{p}) = \mathbf{t}^i(\mathbf{x})\cdot\overline{\mathbf{m}}(\mathbf{p}), \qquad i\in[1,m], \tag{2.23}$$

or, in matrix form:

$$\mathbf{I}(\mathbf{p}) = \mathbf{T}(\mathbf{x})\,\overline{\mathbf{m}}(\mathbf{p}), \tag{2.24}$$

where $\mathbf{I}(\mathbf{p}) \in \mathbb{R}^m$ has been defined in (1.5) and $\mathbf{T}(\mathbf{x}) \in \mathbb{R}^{m\times 3}$ is defined as follows:

$$\mathbf{T}(\mathbf{x}) = \begin{bmatrix} \mathbf{t}^1(\mathbf{x})^\top \\ \vdots \\ \mathbf{t}^m(\mathbf{x})^\top \end{bmatrix}. \tag{2.25}$$

Eq. (2.24) is similar to (1.4). Knowing the matrix field $\mathbf{T}(\mathbf{x})$ would allow us to estimate its field of pseudo-inverses in order to solve (2.24), just as calculating the pseudo-inverse of $\mathbf{S}$ allows us to solve (1.4). However, the matrix field $\mathbf{T}(\mathbf{x})$ depends on $\mathbf{x}$, and thus on the *unknown depth*. We will see that this simple difference induces major changes when it comes to the numerical solution.

The dependency on $\mathbf{x}$ of matrix field $\mathbf{T}(\mathbf{x})$, which significantly complicates the solution of Problem (2.24), has yet one advantage: if some gray levels $I^i(\mathbf{p})$ do not fit the Lambertian model, it is possible to eliminate them, because the number of lines in each matrix $\mathbf{T}(\mathbf{x})$ may vary from one pixel to another. As advocated in [10], we systematically eliminate in Section 3 the highest gray level, which may come from a specular highlight, as well as the two lowest ones, which may correspond to shadows. More elaborate methods for ensuring robustness will be discussed in Section 4.

3 Existing Variational Approaches for Solving Photometric Stereo under Point Light Source Illumination

In this section, we study two variational approaches from the literature for solving photometric stereo under point light source illumination.

The first one inverts the nonlinear image formation model by recasting it as a sequence of simpler subproblems [2,8,13,28,33,46,50,69]. It consists in estimating the normals and the albedo, assuming that the depth map is fixed, then integrating the normals into a new depth map, and to iterate. We show in Section 3.1 how to improve this standard method in order to estimate *absolute* depth, without resorting to any prior.

The second one first linearizes the image formation model by resorting to image ratios, then directly estimates the depth by solving the resulting system of PDEs in an approximate manner [37,38,40,55]. We show in Section 3.2 that state-of-the-art solutions, which resort to fixed point iterations, may be trapped in local minima. This shortcoming can be avoided by rather using an augmented Lagrangian algorithm.

Despite such improvements, we will see that these state-of-the-art methods remain unsatisfactory, because their convergence is not established.

3.1 Scheme Inspired by the Classical Numerical Solution of Photometric Stereo

For solving Problem (2.24), it seems quite natural to adapt the solution (1.6) of the linear model (1.4). To linearize (2.24), we have to assume that matrix $\mathbf{T}(\mathbf{x})$ is known. If we proceed iteratively, this can be made possible by replacing, at iteration $(k+1)$, $\mathbf{T}(\mathbf{x})$ by $\mathbf{T}(\mathbf{x}^{(k)})$. This very simple idea has led to several numerical solutions [2,8,13,28,33,46,50,69], which all require some kind of a priori knowledge on the depth. On the contrary, the scheme we propose here requires none, which constitutes a significant improvement. If $\mathbf{x}^{(0)}$ denotes an initial 3D-shape, this new scheme consits in iterating:

1. Solve Problem (2.24) in the least-squares sense in each $\mathbf{p} \in \Omega$, replacing $\mathbf{T}(\mathbf{x})$ by $\mathbf{T}(\mathbf{x}^{(k)})$, which provides a new estimation of $\overline{\mathbf{m}}(\mathbf{p})$:
$$\overline{\mathbf{m}}^{(k+1)}(\mathbf{p}) = \mathbf{T}(\mathbf{x}^{(k)})^\dagger \mathbf{I}(\mathbf{p}). \tag{3.1}$$
2. Deduce a new estimation of the normal $\mathbf{n}(\mathbf{p})$:
$$\mathbf{n}^{(k+1)}(\mathbf{p}) = \frac{\overline{\mathbf{m}}^{(k+1)}(\mathbf{p})}{\|\overline{\mathbf{m}}^{(k+1)}(\mathbf{p})\|}. \tag{3.2}$$
3. Integrate the new normal field $\mathbf{n}^{(k+1)}$, into an updated 3D-shape $\mathbf{x}^{(k+1)}$, up to a scale factor.
4. Estimate this scale factor by nonlinear optimization.
5. Increment k as long as $k \leq k_{\max}$.

For this scheme to be completely specified, we need to set the initial 3D-shape $\mathbf{x}^{(0)}$. We use as initial guess a fronto-parallel plane at distance z_0 from the camera, z_0 being a rough estimate of the mean distance from the camera to the surface.

Integration of Normals – Stages 1 and 2 of the scheme above are trivial and can be achieved pixelwise, but Stages 3 and 4 are trickier. From the equalities in (1.1), it is easy to deduce that the (non-unit-length) vector

$$\overline{\mathbf{n}}(\mathbf{p}) = \begin{bmatrix} f\,\partial_u z(\mathbf{p}) \\ f\,\partial_v z(\mathbf{p}) \\ -z(\mathbf{p}) - \mathbf{p}\cdot\nabla z(\mathbf{p}) \end{bmatrix} \tag{3.3}$$

is normal to the surface (the vector $\nabla z(\mathbf{p}) = [\partial_u z(\mathbf{p}), \partial_v z(\mathbf{p})]^\top$ is the gradient of z in $\mathbf{p}$). Expression (3.3) shows that integrating the (unit-length) normal field $\mathbf{n}$ allows to estimate the depth z only up to a scale factor $\kappa \in \mathbb{R}$, since:

$$\mathbf{n}(\mathbf{p}) \propto \begin{bmatrix} f\,\partial_u z(\mathbf{p}) \\ f\,\partial_v z(\mathbf{p}) \\ -z(\mathbf{p}) - \mathbf{p}\cdot\nabla z(\mathbf{p}) \end{bmatrix} \propto \begin{bmatrix} f\,\partial_u (\kappa\, z)(\mathbf{p}) \\ f\,\partial_v (\kappa\, z)(\mathbf{p}) \\ -(\kappa\, z)(\mathbf{p}) - \mathbf{p}\cdot\nabla (\kappa\, z)(\mathbf{p}) \end{bmatrix}. \tag{3.4}$$

The collinearity of $\overline{\mathbf{n}}(\mathbf{p})$ and $\mathbf{n}(\mathbf{p}) = [n_1(\mathbf{p}), n_2(\mathbf{p}), n_3(\mathbf{p})]^\top$ leads to the system

$$\begin{cases} n_3(\mathbf{p})\, f\, \partial_u z(\mathbf{p}) + n_1(\mathbf{p})\, [z(\mathbf{p}) + \mathbf{p} \cdot \nabla z(\mathbf{p})] = 0, \\ n_3(\mathbf{p})\, f\, \partial_v z(\mathbf{p}) + n_2(\mathbf{p})\, [z(\mathbf{p}) + \mathbf{p} \cdot \nabla z(\mathbf{p})] = 0, \end{cases} \tag{3.5}$$

which is homogeneous in $z(\mathbf{p})$. Introducing the change of variable $\tilde{z} = \log(z)$, which is valid since $z > 0$, (3.5) is rewritten

$$\begin{cases} [f\, n_3(\mathbf{p}) + u\, n_1(\mathbf{p})]\, \partial_u \tilde{z}(\mathbf{p}) + v\, n_1(\mathbf{p}) \partial_v \tilde{z}(\mathbf{p}) = -n_1(\mathbf{p}), \\ u\, n_2(\mathbf{p}) \partial_u \tilde{z}(\mathbf{p}) + [f\, n_3(\mathbf{p}) + v\, n_2(\mathbf{p})]\, \partial_v \tilde{z}(\mathbf{p}) = -n_2(\mathbf{p}). \end{cases} \tag{3.6}$$

The determinant of this system is equal to

$$f\, n_3(\mathbf{p})\, [u\, n_1(\mathbf{p}) + v\, n_2(\mathbf{p}) + f\, n_3(\mathbf{p})] = f\, n_3(\mathbf{p})\, [\overline{\mathbf{p}} \cdot \mathbf{n}(\mathbf{p})]\,, \tag{3.7}$$

if we denote

$$\overline{\mathbf{p}} = [u, v, f]^\top. \tag{3.8}$$

It is then easy to deduce the solution of (3.6):

$$\nabla \tilde{z}(\mathbf{p}) = -\frac{1}{\overline{\mathbf{p}} \cdot \mathbf{n}(\mathbf{p})} \begin{bmatrix} n_1(\mathbf{p}) \\ n_2(\mathbf{p}) \end{bmatrix}. \tag{3.9}$$

Let us now come back to Stages 3 and 4 of the proposed scheme. The new normal field is $\mathbf{n}^{(k+1)}(\mathbf{p})$, from which we can deduce the gradient $\nabla \tilde{z}^{(k+1)}(\mathbf{p})$ thanks to Eq. (3.9). By integrating this gradient between a point $\mathbf{p}_0$, chosen arbitrarily inside Ω, and any point $\mathbf{p} \in \Omega$, and knowing that $z = \exp\{\tilde{z}\}$, we obtain:

$$z^{(k+1)}(\mathbf{p}) = z^{(k+1)}(\mathbf{p}_0)\, \exp\left\{ \int_{\mathbf{p}_0}^{\mathbf{p}} \nabla \tilde{z}^{(k+1)}(\mathbf{q}) \cdot \mathrm{d}\mathbf{q} \right\}. \tag{3.10}$$

This integral can be calculated along one single path inside Ω going from $\mathbf{p}_0$ to $\mathbf{p}$, but since the gradient field $\nabla \tilde{z}^{(k+1)}(\mathbf{p})$ is never rigorously integrable in practice, this calculus usually depends on the choice of the path [66]. The most common parry to this well-known problem consists in resorting to a variational approach, see for instance [54] for some discussion.

Expression (3.10) confirms that the depth can only be calculated, from $\mathbf{n}^{(k+1)}(\mathbf{p})$, up to a scale factor equal to $z^{(k+1)}(\mathbf{p}_0)$. Let us determine this scale factor by minimization of the reprojection error of Model (2.24) over the entire domain Ω. Knowing that, from (1.1) and (3.8), we get $\mathbf{x} = \frac{z}{f}\, \overline{\mathbf{p}}$, this comes down to solving the following nonlinear least-squares problem:

$$z^{(k+1)}(\mathbf{p}_0) = \underset{w \in \mathbb{R}^+}{\arg\min}\ \mathcal{E}_{\text{alt}}(w) := \sum_{\mathbf{p} \in \Omega} \Big\| \mathbf{I}(\mathbf{p}) - \mathbf{T}\Big(\frac{w}{f} \exp\left\{ \int_{\mathbf{p}_0}^{\mathbf{p}} \nabla \tilde{z}^{(k+1)}(\mathbf{q}) \cdot \mathrm{d}\mathbf{q} \right\} \overline{\mathbf{p}} \Big)\, \overline{\mathbf{m}}^{(k+1)}(\mathbf{p}) \Big\|^2, \tag{3.11}$$

which allows us to eventually write the shape update (Stages 3 and 4):

$$\mathbf{x}^{(k+1)} = \frac{z^{(k+1)}(\mathbf{p}_0)}{f}\, \exp\left\{ \int_{\mathbf{p}_0}^{\mathbf{p}} \nabla \tilde{z}^{(k+1)}(\mathbf{q}) \cdot \mathrm{d}\mathbf{q} \right\} \overline{\mathbf{p}}. \tag{3.12}$$

Experimental Validation – Despite the lack of theoretical guarantee, convergence of this scheme is empirically observed, provided that the initial 3D-shape $\mathbf{x}^{(0)}$ is not too distant from the scene surface. For the curves in Fig. 6, several fronto-parallel planes with equation $z \equiv z_0$ were tested as initial guess. The mean distance from the camera to the scene being approximately 700 mm, it is not surprising that the fastest convergence is observed for this value of z_0. Besides, this graph also shows that that the initial scale can be under-estimated quite a lot, but over-estimating it severly slows down the process.

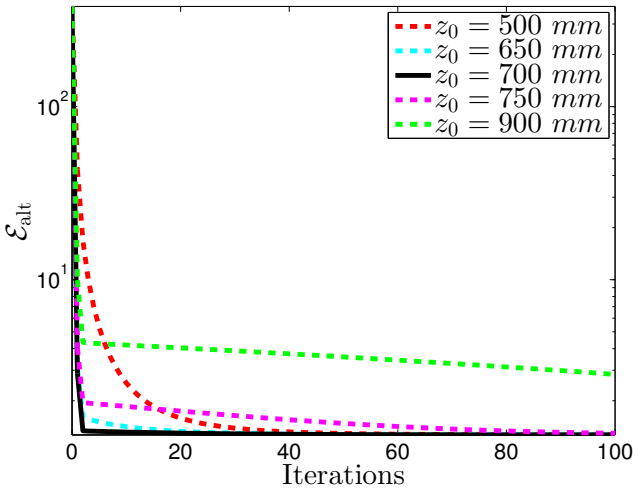


Fig. 6 Evolution of the energy $\mathcal{E}_{\text{alt}}$ of the alternating approach, defined in (3.11), in function of the iterations, when the initial shape is a fronto-parallel plane with equation $z \equiv z_0$. The used data are the $m = 8$ images of the plaster statuette of Fig. 2. The proposed scheme consists in alternating normal estimation, normal integration and scale estimation. It converges towards the same solution (at different speeds), for the five tested values of z_0.

Fig. 7 allows to compare the shape obtained by photometric stereo, from sub-images of size 920×1178 in full resolution (bounding box of the object), which contain 773794 points, with the ground truth obtained by laser scanning, which contains 1753010 points. The points density is thus almost the same on the front of the statuette, since we did not reconstruct its back. However, our result is achieved in less than ten seconds (five iterations of a Matlab code on a recent I7 processor), instead of several hours for the ground truth, while we also estimate the albedo.

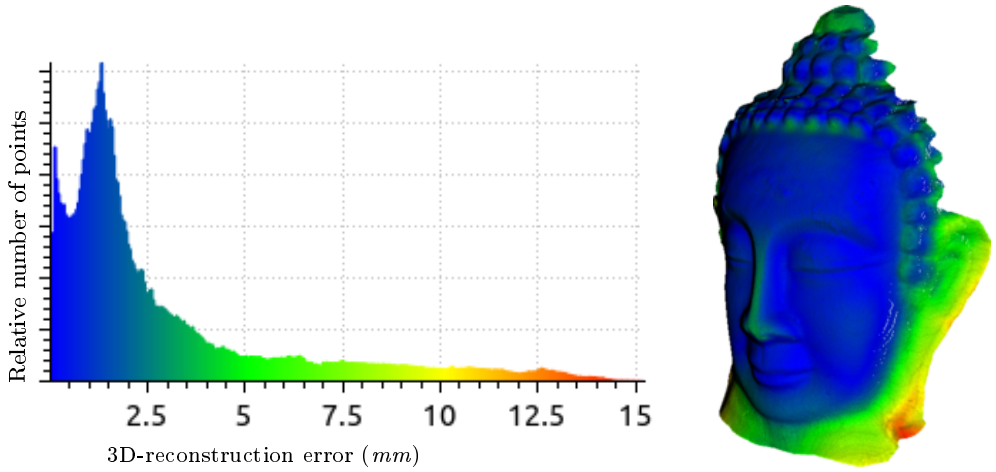


Fig. 8 (a) Histogram of point-to-point distances between the alternating 3D-reconstruction and the ground truth. The median value is 1.3 *mm*. (b) Spatial distribution of these distances. The histogram peaks is not located in zero. This indicates a bias, probably due to normal integration, which can be avoided by resorting to a differential approach based on PDEs.

Fig. 7 (a) 3D-reconstruction and (b) albedo obtained with the proposed alternating scheme. (c) Ground truth shape obtained by laser scanning. Photometric stereo not only provides a 3D-shape qualitatively similar to the laser scan, but it also provides the albedo.

Fig. 8-a shows the histogram of point-to-point distances between our result (Fig. 7-a) and the ground truth (Fig. 7-c). The median value is 1.3 *mm*. The spatial distribution of these distances, which is illustrated in Fig. 8-b, shows that the largest distances are observed on the highest slopes of the surface. This clearly comes from the fact that, even for a diffuse material such as plaster, the Lambertian model is not valid under skimming lighting.

More realistic reflectance models, such as the one proposed by Oren and Nayar in [48], would perhaps improve accuracy of the 3D-reconstruction in these points, but the problem of steep areas also comes from normal integration. In the next section, we describe a different formulation of photometric stereo which permits to avoid integration, by solving a system of PDEs in z.

3.2 Direct Estimation of the Depth Map using Image Ratios

The scheme proposed in Section 3.1 suffers from several defects. We just mentioned that it requires to integrate the gradient $\nabla\tilde{z}^{(k+1)}(\mathbf{p})$ at each iteration. This is not achieved by the naive formulation (3.12), but using more sophisticated methods which allow to overcome the problem of non-integrability [14]. Still, bias due to inaccurate normals estimation should not have to be corrected during integration. Instead, it seems more justified to directly estimate the depth map, without resorting to intermediate normals estimation. This can be achieved by recasting photometric stereo as a system of quasilinear PDEs, using image ratios.

Differential Reformulation of Problem (2.24) – Recall that the coordinates of the 3D-point $\mathbf{x}$ conjugate to a pixel $\mathbf{p}$ are completely characterized by the depth $z(\mathbf{p})$:

$$\mathbf{x} = \frac{z(\mathbf{p})}{f}\begin{bmatrix}\mathbf{p}\\ f\end{bmatrix}. \tag{3.13}$$

The vectors $\mathbf{t}^i(\mathbf{x})$ defined in (2.22) thus depend on the unknown depth values $z(\mathbf{p})$. Using once again the change of variable $\tilde{z} = \log z$, we consider from now on each $\mathbf{t}^i$, $i \in [1, m]$, as a vector field depending on the unknown map $\tilde{z}$:

$$\begin{aligned}\mathbf{t}^i(\tilde{z}) : \Omega &\to \mathbb{R}^3\\ \mathbf{p} &\mapsto \mathbf{t}^i(\tilde{z})(\mathbf{p}) = \Psi^i\left[\mathbf{n}_s^i \cdot \frac{\mathbf{v}^i(\tilde{z})(\mathbf{p})}{\|\mathbf{v}^i(\tilde{z})(\mathbf{p})\|}\right]^{\mu^i}\frac{\mathbf{v}^i(\tilde{z})(\mathbf{p})}{\|\mathbf{v}^i(\tilde{z})(\mathbf{p})\|^3},\end{aligned} \tag{3.14}$$

where we recall that the source intensity Ψ^i and its orientation $\mathbf{n}_s^i$ can be calibrated beforehand (see Section 2), and where each field $\mathbf{t}^i(\tilde{z})$ depends in a nonlinear way on the unknown (log-) depth map $\tilde{z}$, through the vector field $\mathbf{v}^i(\tilde{z}) : \Omega \to \mathbb{R}^3$, which goes from each surface point $\mathbf{x} \in \mathbb{R}^3$ (cf. Eq. (3.13)) towards the (also calibrated) i-th light source location $\mathbf{x}_s^i$:

$$\begin{aligned} &\mathbf{v}^i(\tilde{z}) : \Omega \to \mathbb{R}^3 \\ &\qquad \mathbf{p} \mapsto \mathbf{v}^i(\tilde{z})(\mathbf{p}) = \mathbf{x}_s^i - \tfrac{\exp(\tilde{z}(\mathbf{p}))}{f} \begin{bmatrix} \mathbf{p} \\ f \end{bmatrix}. \end{aligned} \tag{3.15}$$

Knowing that the (non-unit-length) vector $\overline{\mathbf{n}}(\mathbf{p})$ defined in (3.3), divided by $z(\mathbf{p})$, is normal to the surface, we can rewrite System (2.20), in each pixel $\mathbf{p} \in \Omega$:

$$I^i(\mathbf{p}) = \frac{\overline{\rho}(\mathbf{p})}{d(\tilde{z})(\mathbf{p})} \mathbf{t}^i(\tilde{z})(\mathbf{p}) \cdot \begin{bmatrix} f \nabla \tilde{z}(\mathbf{p}) \\ -1 - \mathbf{p} \cdot \nabla \tilde{z}(\mathbf{p}) \end{bmatrix}, \ i \in [1, m], \tag{3.16}$$

with

$$d(\tilde{z})(\mathbf{p}) = \sqrt{f^2 \left\| \nabla \tilde{z}(\mathbf{p}) \right\|^2 + (-1 - \mathbf{p} \cdot \nabla \tilde{z}(\mathbf{p}))^2}. \tag{3.17}$$

System (3.16) can be written in the even more compact following manner:

$$I^i(\mathbf{p}) = \frac{\overline{\rho}(\mathbf{p})}{d(\tilde{z})(\mathbf{p})} \left[\mathbf{Q}(\mathbf{p}) \, \mathbf{t}^i(\tilde{z})(\mathbf{p}) \right] \cdot \begin{bmatrix} \nabla \tilde{z}(\mathbf{p}) \\ -1 \end{bmatrix}, \ i \in [1, m], \tag{3.18}$$

by further introducing the following definition, $\forall \mathbf{p} = [u, v]^\top \in \Omega$:

$$\mathbf{Q}(\mathbf{p}) = \begin{bmatrix} f & 0 & -u \\ 0 & f & -v \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.19}$$

Partial Linearization of (3.18) *using Image Ratios* – In comparison with Eqs. (2.20), the PDEs (3.18) explicitly depend on the unknown map $\tilde{z}$, and thus remove the need for alternating normals estimation and integration. However, these equations contain two difficulties: they are nonlinear and cannot be solved locally. We can eliminate the nonlinearity due to the coefficient of normalization $d(\tilde{z})(\mathbf{p})$. Indeed, neither the relative albedo $\overline{\rho}(\mathbf{p})$, nor this coefficient, depend on the index i of the LED. We deduce from any pair $(i, j) \in [1, m]^2$, $i < j$, of equations forming (3.18), the following equalities:

$$\begin{aligned} \frac{\overline{\rho}(\mathbf{p})}{d(\tilde{z})(\mathbf{p})} &= \frac{I^i(\mathbf{p})}{\mathbf{a}^i(\tilde{z})(\mathbf{p}) \cdot \nabla \tilde{z}(\mathbf{p}) - b^i(\tilde{z})(\mathbf{p})} \\ &= \frac{I^j(\mathbf{p})}{\mathbf{a}^j(\tilde{z})(\mathbf{p}) \cdot \nabla \tilde{z}(\mathbf{p}) - b^j(\tilde{z})(\mathbf{p})}, \end{aligned} \tag{3.20}$$

with the following definitions of $\mathbf{a}^i(\tilde{z})(\mathbf{p})$ and $b^i(\tilde{z})(\mathbf{p})$ (and similar definitions of $\mathbf{a}^j(\tilde{z})(\mathbf{p})$ and $b^j(\tilde{z})(\mathbf{p})$), denoting $\mathbf{t}^i(\tilde{z})(\mathbf{p}) = [t_1^i(\tilde{z})(\mathbf{p}), t_2^i(\tilde{z})(\mathbf{p}), t_3^i(\tilde{z})(\mathbf{p})]^\top$:

$$\mathbf{a}^i(\tilde{z})(\mathbf{p}) = f \begin{bmatrix} t_1^i(\tilde{z})(\mathbf{p}) \\ t_2^i(\tilde{z})(\mathbf{p}) \end{bmatrix} - t_3^i(\tilde{z})(\mathbf{p}) \, \mathbf{p}, \tag{3.21}$$

$$b^i(\tilde{z})(\mathbf{p}) = t_3^i(\tilde{z})(\mathbf{p}). \tag{3.22}$$

From the equalities (3.20), we obtain the following PDE in z:

$$\begin{aligned} &\underbrace{\left[I^i(\mathbf{p}) \, \mathbf{a}^j(\tilde{z})(\mathbf{p}) - I^j(\mathbf{p}) \, \mathbf{a}^i(\tilde{z})(\mathbf{p}) \right]}_{\mathbf{a}^{i,j}(\tilde{z})(\mathbf{p})} \cdot \nabla \tilde{z}(\mathbf{p}) \\ &\qquad = \underbrace{\left[I^i(\mathbf{p}) \, b^j(\tilde{z})(\mathbf{p}) - I^j(\mathbf{p}) \, b^i(\tilde{z})(\mathbf{p}) \right]}_{b^{i,j}(\tilde{z})(\mathbf{p})}. \end{aligned} \tag{3.23}$$

The fields $\mathbf{a}^{i,j}(\tilde{z})$ and $b^{i,j}(\tilde{z})$ defined in (3.23) depend on $\tilde{z}$ but not on $\nabla \tilde{z}$: Eq. (3.23) is thus a *quasi-linear* PDE. It could be solved by the characteristic strips expansion method [41, 42] if we were dealing with $m = 2$ images only, but using a larger number of images is necessary in order to design a robust 3D-reconstruction method. Since we are provided with $m > 2$ images, we follow [19, 37, 38, 40, 55, 60] and write $\binom{m}{2}$ PDEs such as (3.23) formed by the $\binom{m}{2}$ pairs $(i, j) \in [1, m]^2$, $i < j$. Forming the matrix field $\mathbf{A}(\tilde{z}) : \Omega \to \mathbb{R}^{\binom{m}{2} \times 2}$ by concatenation of the row vectors $\mathbf{a}^{i,j}(\tilde{z})(\mathbf{p})^\top$, and the vector field $\mathbf{b}(\tilde{z}) : \Omega \to \mathbb{R}^{\binom{m}{2}}$ by concatenation of the scalar values $b^{i,j}(\tilde{z})(\mathbf{p})$, the system of PDEs to solve is written:

$$\mathbf{A}(\tilde{z}) \, \nabla \tilde{z} = \mathbf{b}(\tilde{z}) \quad \text{over } \Omega. \tag{3.24}$$

This new differential formulation of photometric stereo seems simpler than the original differential formulation (3.18), since the main source of nonlinearity, due to the denominator $d(\tilde{z})$, has been eliminated. However, it presents two difficulties. First, the PDEs (3.24) are generally incompatible and hence do not admit an exact solution. It is thus necessary to estimate an approximate one, by resorting to a variational approach. Second, they do not allow to estimate the scale of the scene. Indeed, when all the depth values simultaneously tend to infinity, then both members of (3.24) tend to zero (because the coordinates of $\mathbf{t}^i$ do so, cf. (3.14)). Thus, a large, distant shape will always "better" fit these PDEs than a small, nearby one (cf. Figs. 11 and 10). A "locally optimal" solution close to a very good initial estimate should thus be sought. Assuming that each of the $\binom{m}{2}$ equalities in System (3.24) is satisfied up to an additive, zero-mean, Gaussian noise[8], one should estimate such

[8] In fact, any noise assumption should be formulated on the images, and not on Model (3.24), which was obtained by

a solution by solving the following variational problem:

$$\min_{\tilde{z}:\Omega\to\mathbb{R}} \mathcal{E}_{\text{rat}}(\tilde{z}) := \|\mathbf{A}(\tilde{z})\,\nabla\tilde{z} - \mathbf{b}(\tilde{z})\|^2_{L^2(\Omega)}. \tag{3.25}$$

Fixed Point Iterations for Solving (3.25) – Overall, we are looking for a solution of (3.24) which must "stay close" to some reasonable initial estimate, in order to avoid the trivial solution. It has been proposed in [37, 38, 40, 55] to iteratively estimate this solution, by uncoupling the (linear) estimation of $\tilde{z}$ from the (nonlinear) estimation of $\mathbf{A}(\tilde{z})$ and $\mathbf{b}(\tilde{z})$. This can be achieved by rewriting (3.25) as the following constrained optimization problem:

$$\begin{aligned} &\min_{\tilde{z}:\Omega\to\mathbb{R}} \|\mathbf{A}\,\nabla\tilde{z} - \mathbf{b}\|^2_{L^2(\Omega)} \\ &\text{s.t.} \quad \begin{bmatrix}\mathbf{A}\\ \mathbf{b}\end{bmatrix} = \begin{bmatrix}\mathbf{A}(\tilde{z})\\ \mathbf{b}(\tilde{z})\end{bmatrix}, \end{aligned} \tag{3.26}$$

and resorting to a fixed point iterative scheme:

$$\tilde{z}^{(k+1)} = \arg\min_{\tilde{z}:\Omega\to\mathbb{R}} \|\mathbf{A}^{(k)}\,\nabla\tilde{z} - \mathbf{b}^{(k)}\|^2_{L^2(\Omega)}, \tag{3.27}$$

$$\begin{bmatrix}\mathbf{A}^{(k+1)}\\ \mathbf{b}^{(k+1)}\end{bmatrix} = \begin{bmatrix}\mathbf{A}(\tilde{z}^{(k+1)})\\ \mathbf{b}(\tilde{z}^{(k+1)})\end{bmatrix}. \tag{3.28}$$

The linear least-squares variational problem (3.27) can be solved by discretizing the gradient operator ∇ by finite differences (we use forward, first-order finite differences, with a Neumann boundary condition). The discrete optimization problem obtained this way admits as necessary optimality condition the normal equations, which form a linear system whose matrix is symmetric, positive semi-definite. This matrix is usually too larged to be inverted directly. Therefore, iterative methods such as the conjugate gradient algorithm should be preferred.

ADMM Iterations for Solving (3.25) – Fig. 9 shows that the previous fixed point scheme does not always decrease the energy after each iteration. We suspect that this is due to the choice of "freezing" the nonlinearities of the variational problem (3.25). Instead, $\tilde{z}$ can be estimated not only from the linearized parts, but also from the nonlinear ones, in the spirit of the ADMM algorithm, a standard procedure which dates back to the 70's [15, 18], but has been revisited recently [9].

considering ratios of gray levels: if the noise on gray levels is Gaussian, then that on ratios is Cauchy-distributed [24]. Hence, the least-squares solution (3.25) is the best linear unbiased estimator, but it is not the optimal solution, which would require non-convex estimators [14]. We leave this, however, as future work.

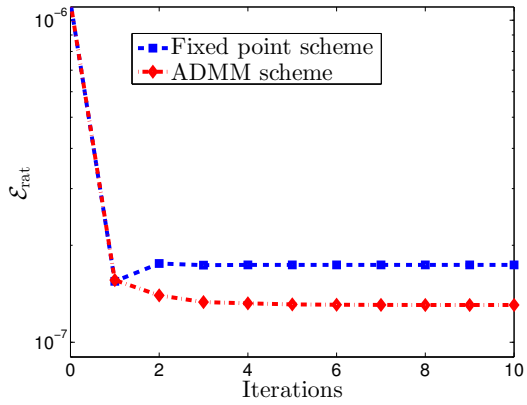


Fig. 9 Evolution of the energy $\mathcal{E}_{\text{rat}}$ of the ratios-based approach, defined in (3.25), in function of the iterations, for the data of Fig. 2 (the initial shape is a fronto-parallel plane with equation $z \equiv 700\ mm$). With the fixed point scheme, the energy is not always decreased after each iteration, contrarily to the ADMM scheme we are going to introduce.

In this view, we introduce an auxiliary variable $\overline{z}$ and reformulate Problem (3.25) as follows:

$$\begin{aligned} &\min_{\overline{z},\tilde{z}} \|\mathbf{A}(\overline{z})\,\nabla\tilde{z} - \mathbf{b}(\overline{z})\|^2_{L^2(\Omega)} \\ &\text{s.t.}\ \ \tilde{z} = \overline{z}. \end{aligned} \tag{3.29}$$

In order to solve the constrained optimization problem (3.29), let us introduce a dual variable w and a descent step ν. A local solution of (3.29) is then obtained at convergence of the following iterative scheme:

1. Update $\tilde{z}$ by using the linear part, "while keeping $\tilde{z}$ close to $\overline{z}^{(k)}$":

$$\tilde{z}^{(k+1)} = \arg\min_{\tilde{z}} \left\|\mathbf{A}(\overline{z}^{(k)})\nabla\tilde{z} - \mathbf{b}(\overline{z}^{(k)})\right\|^2_{L^2(\Omega)} + \frac{1}{2\,\nu}\left\|\tilde{z} - \overline{z}^{(k)} + w^{(k)}\right\|^2_{L^2(\Omega)}. \tag{3.30}$$

2. Update $\overline{z}$ by using the nonlinear part, "while keeping $\overline{z}$ close to $\tilde{z}^{(k+1)}$":

$$\overline{z}^{(k+1)} = \arg\min_{\overline{z}} \left\|\mathbf{A}(\overline{z})\,\nabla\tilde{z}^{(k+1)} - \mathbf{b}(\overline{z})\right\|^2_{L^2(\Omega)} + \frac{1}{2\,\nu}\left\|\tilde{z}^{(k+1)} - \overline{z} + w^{(k)}\right\|^2_{L^2(\Omega)}. \tag{3.31}$$

3. Update the dual variable w:

$$w^{(k+1)} = w^{(k)} + \tilde{z}^{(k+1)} - \overline{z}^{(k+1)}. \tag{3.32}$$

The first stage (3.30) is a linear least-squares problem which can be solved in the same way as for the previous scheme, using the normal equations of its discrete formulation. Note that the presence of the regularization term now guarantees the positive definiteness of the matrix of the system, and therefore convergence for the conjugate gradient algorithm.

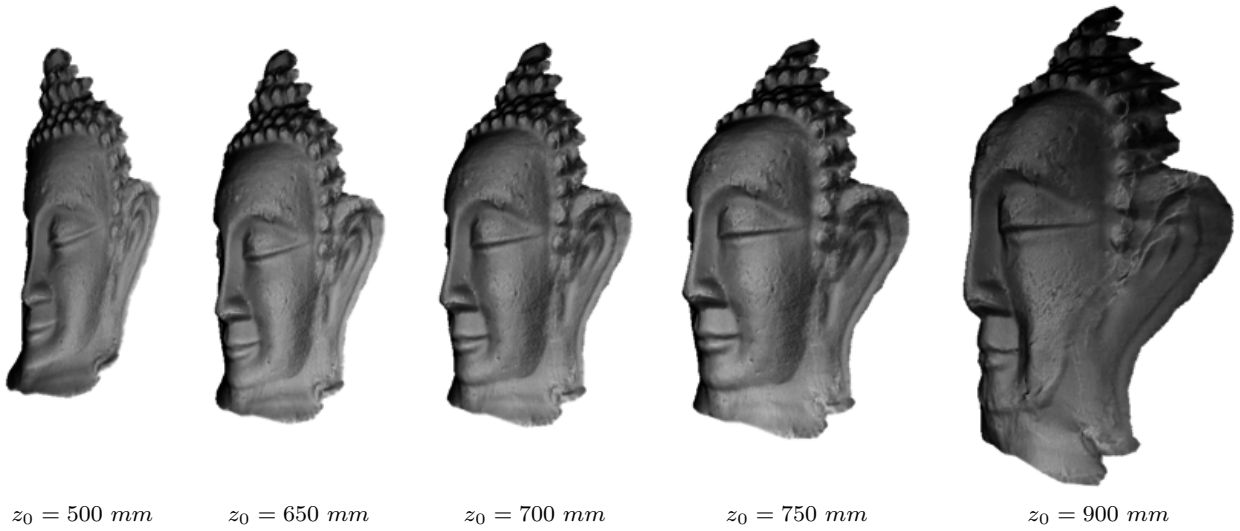


Fig. 10 3D-reconstructions after 10 iterations of the ADMM scheme, taking as initial guess different fronto-parallel planes $z \equiv z_0$. The median of the distances to ground truth is, from left to right: 3.05 mm, 2.88 mm, 1.68 mm, 2.08 mm and 5.86 mm. When the initial guess is too close to the camera, the 3D-reconstruction is flattened, while the scale is overestimated when starting too far away from the camera (although this yields a lower energy, see Fig. 11).

Thanks to the auxiliary variable $\overline{z}$, which decouples $\nabla\tilde{z}$ and $\tilde{z}$ in Problem (3.29), the second stage (3.31) is a *local* nonlinear least-squares problem: in fact, $\nabla\overline{z}$ is not involved in this problem, which can be solved pixelwise. Problem (3.31) thus reduces to a nonlinear least-squares estimation problem of one real variable, which can be solved by a standard method such as the Levenberg-Marquardt algorithm.

Because of the nonlinearity of Problem (3.31), it is unfortunately impossible to guarantee convergence for this ADMM scheme, which depends on the initialization and on the descent parameter ν [9]. A reasonable initialization strategy consists in using the solution provided by the scheme of Section 3.1. As for the descent step ν, we iteratively calculate its optimal value according to the *Penalty Varying Parameter* procedure described in [9]. Finally, the iteration stops when the relative variation of the criterion of Problem (3.31) falls under a threshold equal to 10^{-4}. Fig. 9 shows that with such choices, Problem (3.25) is solved more efficiently than with the fixed point scheme.

Fig. 11 shows that whatever the initial guess, this ADMM scheme always decreases the energy after each iteration. Yet, the final solution strongly depends on the initial guess, as confirmed by Fig. 10.

Fig. 12 shows the 3D-reconstruction obtained by refining the results of Section 3.1 using the ADMM scheme. At first sight, the shape depicted in Fig. 12-a seems hardly different from that of Fig. 7-a, but the comparison of histograms in Figs. 8-a and 12-b indicates that bias has been significantly reduced. This shows the superiority of direct depth estimation over alternating normal estimation and integration. Nevertheless, the lack of convergence guarantees and the need for appropriate initialization remain limiting bottlenecks. The method discussed in the next section overcomes both these issues.

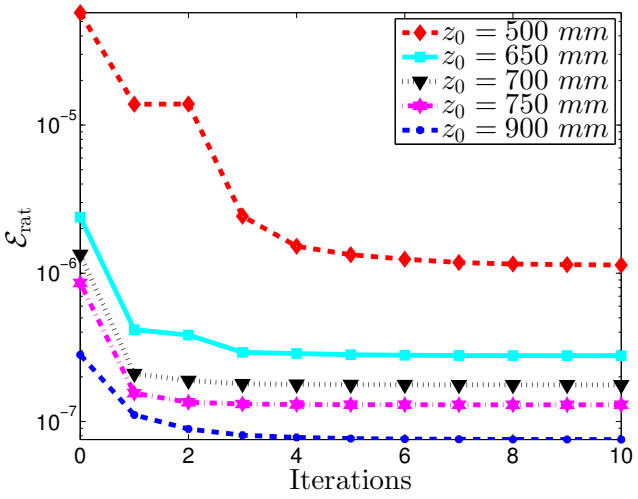
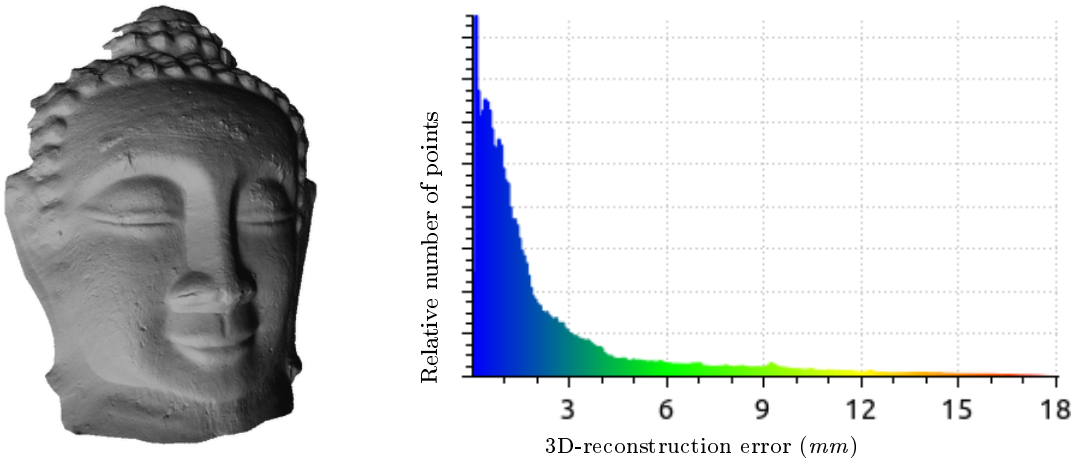


Fig. 11 Evolution of the energy $\mathcal{E}_{\text{rat}}$ defined in (3.25), in function of the iterations, for the data of Fig. 2. Using as initialization $\tilde{z} \equiv \log(z_0)$, the ADMM scheme always converges towards a local minimum, yet this minimum strongly depends on the value of z_0. Besides, a lower final energy does not necessarily means a better 3D-reconstruction, as shown in Fig. 10. A careful initial guess is thus of primary importance.



(a) (b)

Fig. 12 (a) 3D-reconstruction obtained with the ADMM scheme, using the result from Fig. 7-a as initial guess. (b) Histogram of point-to-point distances between this shape and the ground truth (cf. Fig. 7-c). The median value is 1.2 *mm*. The bias observed in the histogram of Fig. 8-a has been significantly reduced.

4 A Provably Convergent Scheme for Photometric Stereo under Point Light Source Illumination

When it comes to solving photometric stereo under point light source illumination, there are two main sources of difficulties: the quasilinear dependency of the lighting vectors on the depth map, and the presence of the nonlinear coefficient ensuring that the normal vectors have unit-length.

The alternating strategy from Section 3.1 solves the former issue by freezing the lighting vectors at each iteration, and the latter by simultaneously estimating the normal vector and the albedo. The objective function tackled in this approach, which is based on the reprojection error, seems to be the most relevant. Indeed, the final result is empirically found to be independent from the initialization, although convergence is not established.

On the other hand, the differential strategy from Section 3.2 explicitly tackles the nonlinear dependency of lighting on the depth, and eliminates the other nonlinearity using image ratios. Directly estimating depth reduces bias, but the objective function derived from image ratios admits local minima and convergence is still not established.

Therefore, an ideal numerical solution should: i) build upon a differential approach, in order to reduce bias, ii) avoid linearization using ratios, in order not to get trapped in local minima, and iii) be provably convergent. The variational approach presented in this section, initially presented in [57], satisfies these three criteria.

4.1 Proposed Discrete Variational Framework

Each PDE (3.18) is nonlinear, because $d(\tilde{z})$ depends in a nonlinear way on $\nabla\tilde{z}$ (see Eq. (3.17)). In Section 3.2, we have shown that this nonlinear coefficient can be eliminated using image ratios, yet this leads to local minima. On the other hand, explicitly considering this nonlinear coefficient within a variational framework is not an easy task [25]. Instead, we take inspiration from the way conventional photometric stereo [64] is linearized and integrate the nonlinearity inside the albedo variable, as we proposed recently in [57,58]. To this end, instead of estimating in each pixel $\mathbf{p}$ the unknown value $\overline{\rho}(\mathbf{p})$, we rather estimate the following one:

$$\tilde{\rho}(\mathbf{p}) = \frac{\overline{\rho}(\mathbf{p})}{d(\tilde{z})(\mathbf{p})}. \tag{4.1}$$

The system of nonlinear PDEs (3.18) is then rewritten as

$$I^i(\mathbf{p}) = \tilde{\rho}(\mathbf{p}) \ \left[\mathbf{Q}(\mathbf{p})\,\mathbf{t}^i(\tilde{z})(\mathbf{p})\right] \cdot \begin{bmatrix} \nabla\tilde{z}(\mathbf{p}) \\ -1 \end{bmatrix}, \qquad i \in [1,m], \tag{4.2}$$

which is a system of *quasilinear* PDEs in $(\tilde{\rho},\tilde{z})$, because $\mathbf{t}^i(\tilde{z})$ only depends on $\tilde{z}$, and not on $\nabla\tilde{z}$. Once $\tilde{\rho}$ and $\tilde{z}$ are estimated, it is straightforward to recover the "real" albedo $\overline{\rho}$ using (4.1).

Let us now denote $j = 1 \dots n$ the indices of the pixels inside Ω, I^i_j the gray level of pixel j in image I^i, $\tilde{\boldsymbol{\rho}} \in \mathbb{R}^n$ and $\tilde{\mathbf{z}} \in \mathbb{R}^n$ the vectors stacking the values $\tilde{\rho}_j$ and $\tilde{z}_j$, $\mathbf{t}^i_j(\tilde{z}_j) \in \mathbb{R}^3$ the vector $\mathbf{t}^i(\tilde{z})$ at pixel j, which smoothly (though nonlinearly) depends on $\tilde{z}_j$, and $\mathbf{Q}_j$ the matrix defined in Eq. (3.19) at pixel j. Then, the

discrete counterpart of Eq. (4.2) is written as the following system of nonlinear equations in $(\tilde{\boldsymbol{\rho}}, \tilde{\mathbf{z}})$:

$$I_j^i = \tilde{\rho}_j \ \left[\mathbf{Q}_j \, \mathbf{t}_j^i(\tilde{z}_j)\right] \cdot \begin{bmatrix} (\nabla \tilde{\mathbf{z}})_j \\ -1 \end{bmatrix}, \qquad i \in [1,m],\ j \in [1,n], \tag{4.3}$$

where $(\nabla \tilde{\mathbf{z}})_j \in \mathbb{R}^2$ represents a finite differences approximation of the gradient of $\tilde{z}$ at pixel j[9].

Our goal is to jointly estimate the albedo values $\tilde{\boldsymbol{\rho}} \in \mathbb{R}^n$ and the depth values $\tilde{\boldsymbol{z}} \in \mathbb{R}^n$ from the set of nonlinear equations (4.3), as solution of the following discrete optimization problem:

$$\min_{\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}}} \mathcal{E}(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}}) := \sum_{j=1}^{n} \sum_{i=1}^{m} \phi\left(r_j^i(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}})\right), \tag{4.4}$$

where the residual $r_j^i(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}})$ depends *locally* on $\tilde{\boldsymbol{\rho}}$, but *globally* on $\tilde{\boldsymbol{z}}$:

$$r_j^i(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}}) = \tilde{\rho}_j \left\{\zeta_j^i(\tilde{\boldsymbol{z}})\right\}_+ - I_j^i, \tag{4.5}$$

where:

$$\zeta_j^i(\tilde{\boldsymbol{z}}) = \mathbf{t}_j^i(\tilde{z}_j)^\top \mathbf{Q}_j^\top \begin{bmatrix} (\nabla \tilde{\boldsymbol{z}})_j \\ -1 \end{bmatrix}. \tag{4.6}$$

For fair comparison with the algorithms in Section 3, one can use $\{x\}_+ = x$ and $\phi(x) = \phi_{\text{LS}}(x) = x^2$. To improve robustness, self-shadows can be explicitly handled by using $\{x\}_+ = \max\{x, 0\}$, and the estimator ϕ can be chosen as any function which is non-negative, even, twice continuously differentiable, and monotonically increasing over $\mathbb{R}^+$ such that, for all $x \in \mathbb{R}$:

$$\frac{\phi'(x)}{x} \geq \phi''(x). \tag{4.7}$$

A typical example is Cauchy's robust M-estimator[10]:

$$\phi_{\text{Cauchy}}(x) = \lambda^2 \log(1 + x^2/\lambda^2), \tag{4.8}$$

where the parameter λ is user-defined (we always use the value $\lambda = 0.1$). We refer the reader to [58] for an in-depth discussion on the choice of robust estimators for photometric stereo, as well as for empirical evidences that this robust variational framework outperforms other state-of-the-art approaches such as low-rank factorization techniques [65], per-pixel robust estimation of the normal vector [30], or L^1 norm-based differential ratios [40].

[9] In our experiments, we use the same discretization as in Section 3.2, for fair comparison.

[10] An advantage of our formulation is to be generic i.e., independent from the choice of the function ϕ and of the operator $\{\cdot\}_+$.

4.2 Numerical Solution of (4.4)

Our goal is to find a local minimizer $(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)$ for (4.4), which must satisfy the following first-order conditions:

$$\frac{\partial \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) = \sum_{j=1}^{n} \sum_{i=1}^{m} \phi'(r_j^i(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)) \frac{\partial r_j^i}{\partial \tilde{\rho}_j}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) = \mathbf{0}, \tag{4.9}$$

$$\frac{\partial \mathcal{E}}{\partial \tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) = \sum_{j=1}^{n} \sum_{i=1}^{m} \phi'(r_j^i(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)) \frac{\partial r_j^i}{\partial \tilde{z}_j}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) = \mathbf{0}, \tag{4.10}$$

with:

$$\frac{\partial r_j^i}{\partial \tilde{\rho}_j}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) = \{\zeta_j^i(\tilde{\boldsymbol{z}}^*)\}_+, \tag{4.11}$$

$$\frac{\partial r_j^i}{\partial \tilde{z}_j}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) = \tilde{\rho}_j^* \chi(\zeta_j^i(\tilde{\boldsymbol{z}}^*))(\zeta_j^i)'(\tilde{\boldsymbol{z}}^*), \tag{4.12}$$

where χ is the (sub-)derivative of $\{\cdot\}_+$, which is a constant function equal to 1 if $\{x\}_+ = x$, and the Heaviside function if $\{x\}_+ = \max\{x, 0\}$.

For this purpose, we derive an alternating reweighted least-squares (ARLS) scheme. Suggested by its name, the ARLS scheme alternates Newton-like steps over $\tilde{\boldsymbol{\rho}}$ and $\tilde{\boldsymbol{z}}$, which can be interpreted as iteratively reweighted least-squares iterations.

Algorithm 1 (alternating reweighted least-squares)

1: *Initialize* $\tilde{\boldsymbol{\rho}}^{(0)}, \tilde{\boldsymbol{z}}^{(0)} \in \mathbb{R}^n$. *Set* $k := 0$.
2: ***loop***
3: *Generate* $\frac{\partial \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\boldsymbol{z}}^{(k)})$ *according to* (4.9), *and a positive semi-definite regularization* $H_\rho(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\boldsymbol{z}}^{(k)})$ *of the Hessian* $\frac{\partial^2 \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}}^2}(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\boldsymbol{z}}^{(k)})$ *according to* (4.21).
4: *Compute the Newton step* $\tilde{\boldsymbol{\rho}}^{(k+1)} = \tilde{\boldsymbol{\rho}}^{(k)} + \delta\tilde{\boldsymbol{\rho}}^{(k)}$, *where* $\delta\tilde{\boldsymbol{\rho}}^{(k)}$ *is the minimal-norm solution of the linear system*

$$H_\rho(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\boldsymbol{z}}^{(k)}) \delta\tilde{\boldsymbol{\rho}}^{(k)} = -\frac{\partial \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\boldsymbol{z}}^{(k)}), \tag{4.13}$$

which admits the closed-form solution (4.19).
5: *Generate* $\frac{\partial \mathcal{E}}{\partial \tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)})$ *according to* (4.10), *and a positive semi-definite regularization* $H_z(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)})$ *of the Hessian* $\frac{\partial^2 \mathcal{E}}{\partial \tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)})$, *according to* (4.25).
6: *Compute the Newton step* $\tilde{\boldsymbol{z}}^{(k+1)} = \tilde{\boldsymbol{z}}^{(k)} + \delta\tilde{\boldsymbol{z}}^{(k)}$, *where* $\delta\tilde{\boldsymbol{z}}^{(k)}$ *is the minimal-norm solution of the linear system*

$$H_z(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)}) \delta\tilde{\boldsymbol{z}}^{(k)} = -\frac{\partial \mathcal{E}}{\partial \tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)}), \tag{4.14}$$

which can be achieved by conjugate gradient iterations.
7: *If the stopping criterion is not satisfied, then set* $k := k + 1$.

The proposed algorithm only requires an initial shape estimate $\tilde{\boldsymbol{z}}^{(0)}$, an initial albedo estimate $\tilde{\boldsymbol{\rho}}^{(0)}$ and a stopping criterion. In our experiments, we use constant vectors as initializations for $\tilde{\boldsymbol{z}}$ and $\tilde{\boldsymbol{\rho}}$ i.e., the surface is initially approximated by a plane with uniform albedo. We experimentally found out that the albedo initialization has no real impact on the process, while that of the depth mostly changes the speed of convergence. Iterations are stopped when the relative difference between two successive values of the energy $\mathcal{E}$ falls below a threshold set to 10^{-3}.

We now describe in more care the updates (4.13) and (4.14). Similar to the famous iteratively reweighted least-squares (IRLS) algorithm, ARLS solves the original (possibly non-convex) problem (4.4) iteratively, by recasting it as a series of simpler quadratic programs. Given a current estimate of the solution, IRLS optimizes a local quadratic approximation of the objective function around this estimate, and iterates. We refer the reader to [63] for an overview of this algorithm.

In case of two unknowns $\tilde{\boldsymbol{\rho}}$ and $\tilde{\boldsymbol{z}}$, we alternate optimization over each variable, keeping the other fixed and solving the local quadratic model. This yields the following sequence of reweighted least-squares problems:

$$\tilde{\boldsymbol{\rho}}^{(k+1)} = \underset{\tilde{\boldsymbol{\rho}}\in\mathbb{R}^n}{\arg\min}\ \mathcal{E}^{(k)}_{\tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}) := \sum_{j=1}^{n}\sum_{i=1}^{m} w_j^i(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})\left(r_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}}^{(k)})\right)^2, \quad (4.15)$$

$$\tilde{\boldsymbol{z}}^{(k+1)} = \underset{\tilde{\boldsymbol{z}}\in\mathbb{R}^n}{\arg\min}\ \mathcal{E}^{(k)}_{\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{z}}) := \sum_{j=1}^{n}\sum_{i=1}^{m} w_j^i(\tilde{\boldsymbol{\rho}}^{(k+1)},\tilde{\boldsymbol{z}}^{(k)})\left(r_j^i(\tilde{\boldsymbol{\rho}}^{(k+1)},\tilde{\boldsymbol{z}})\right)^2. \quad (4.16)$$

Here the following (lagged) weight variable w is used[11]:

$$w_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}}) = \begin{cases} \phi'(r_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}}))/r_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}}) & \text{if } r_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})\neq 0,\\ 0 & \text{otherwise,}\end{cases} \quad (4.17)$$

and the functions $\mathcal{E}^{(k)}_{\tilde{\boldsymbol{\rho}}}$ and $\mathcal{E}^{(k)}_{\tilde{\boldsymbol{z}}}$ are local quadratic approximations of $\mathcal{E}$ around, respectively, $(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})$ and $(\tilde{\boldsymbol{\rho}}^{(k+1)},\tilde{\boldsymbol{z}}^{(k)})$.

Problem (4.15) can be rewritten as the following n independent linear least-squares problems, $j\in[1,n]$:

$$\tilde{\rho}_j^{(k+1)} = \underset{\tilde{\rho}_j\in\mathbb{R}}{\arg\min}\sum_{i=1}^{m} w_j^i(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})\left(r_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}}^{(k)})\right)^2 \quad (4.18)$$

Each problem (4.18) almost always admits a unique solution. When it does not, we set $\tilde{\rho}_j^{(k+1)} = \tilde{\rho}_j^{(k)}$. The update thus admits the following closed-form solution, for every pixel $j\in[1,n]$:

$$\tilde{\rho}_j^{(k+1)} = \begin{cases} \dfrac{\sum_{i=1}^m w_j^i(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})\left\{\zeta_j^i(\tilde{\boldsymbol{z}}^{(k)})\right\}_+ I_j^i}{\sum_{i=1}^m w_j^i(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})\left\{\zeta_j^i(\tilde{\boldsymbol{z}}^{(k)})\right\}_+^2} \\ \quad \text{if } \sum_{i=1}^m w_j^i(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})\left\{\zeta_j^i(\tilde{\boldsymbol{z}}^{(k)})\right\}_+^2 > 0, \\ \tilde{\rho}_j^{(k)} \text{ otherwise.}\end{cases} \quad (4.19)$$

This means that $\tilde{\boldsymbol{\rho}}^{(k+1)}$ is set to be the solution of (4.15) which has minimal (Euclidean) distance to $\tilde{\boldsymbol{\rho}}^{(k)}$. In terms of the pseudo-inverse of a matrix, this can be rephrased as[12]

$$\tilde{\boldsymbol{\rho}}^{(k+1)} = \tilde{\boldsymbol{\rho}}^{(k)} - H_\rho(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})^\dagger \partial\mathcal{E}^{(k)}_{\tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}^{(k)}), \quad (4.20)$$

where the n-by-n matrix $H_\rho(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})$ is defined, for an arbitrary $\delta\tilde{\boldsymbol{\rho}}\in\mathbb{R}^n$, by

$$\delta\tilde{\boldsymbol{\rho}}^\top H_\rho(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})\delta\tilde{\boldsymbol{\rho}} = \sum_{j=1}^{n}\sum_{i=1}^{m} w_j^i(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)}) \left(\delta\tilde{\boldsymbol{\rho}}_j\{\zeta_j^i(\tilde{\boldsymbol{z}}^{(k)})\}_+\right)^2. \quad (4.21)$$

Since $\partial\mathcal{E}^{(k)}_{\tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}^{(k)}) = \frac{\partial\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}^{(k)},\tilde{\boldsymbol{z}}^{(k)})$, Eq. (4.20) exactly yields the Newton step (4.13).

The depth update (4.16) is a nonlinear least-squares problem which is trickier to solve, due to the nonlinearity of $r_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})$ with respect to $\tilde{\boldsymbol{z}}$. We therefore follow a Gauss-Newton strategy. A first-order Taylor approximation around $\tilde{\boldsymbol{z}}^{(k)}$ yields, using (4.12), the following approximation of (4.16):

$$\tilde{\boldsymbol{z}}^{(k+1)} = \underset{\tilde{\boldsymbol{z}}\in\mathbb{R}^n}{\arg\min}\sum_{j=1}^{n}\sum_{i=1}^{m} w_j^i(\tilde{\boldsymbol{\rho}}^{(k+1)},\tilde{\boldsymbol{z}}^{(k)})\Big(r_j^i(\tilde{\boldsymbol{\rho}}^{(k+1)},\tilde{\boldsymbol{z}}^{(k)}) + \tilde{\boldsymbol{\rho}}_j^{(k+1)}\chi(\zeta_j^i(\tilde{\boldsymbol{z}}^{(k)}))(\tilde{\boldsymbol{z}}-\tilde{\boldsymbol{z}}^{(k)})^\top(\zeta_j^i)'(\tilde{\boldsymbol{z}}^{(k)})\Big)^2, \quad (4.22)$$

which can also be interpreted as an approximate solution of the $\tilde{\boldsymbol{z}}$-subproblem in (4.4), linearized around $\tilde{\boldsymbol{z}}^{(k)}$ as follows:

$$\min_{\tilde{\boldsymbol{z}}\in\mathbb{R}^n}\tilde{\mathcal{E}}(\tilde{\boldsymbol{z}};\tilde{\boldsymbol{\rho}}^{(k+1)},\tilde{\boldsymbol{z}}^{(k)}) := \sum_{j=1}^{n}\sum_{i=1}^{m}\phi\Big(r_j^i(\tilde{\boldsymbol{\rho}}^{(k+1)},\tilde{\boldsymbol{z}}^{(k)}) + \tilde{\boldsymbol{\rho}}_j^{(k+1)}\chi(\zeta_j^i(\tilde{\boldsymbol{z}}^{(k)}))(\tilde{\boldsymbol{z}}-\tilde{\boldsymbol{z}}^{(k)})^\top(\zeta_j^i)'(\tilde{\boldsymbol{z}}^{(k)})\Big). \quad (4.23)$$

Note that $\partial\tilde{\mathcal{E}}(\cdot;\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})(\tilde{\boldsymbol{z}}) = \frac{\partial\mathcal{E}}{\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})$ for all $(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})$.

[11] This variable can be used as weight since, $\forall x\in\mathbb{R}$, $\phi'(x)/x\geq 0$ and thus $w_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})\geq 0$.

[12] We use ∂ as continuous gradient operator, since ∇ is used for the discrete gradient one.

Similar to the $\tilde{\boldsymbol{\rho}}$-subproblem, $\tilde{\boldsymbol{z}}^{(k+1)}$ is taken to be of minimal distance to $\tilde{\boldsymbol{z}}^{(k)}$ whenever the non-uniqueness of the solution in (4.22) is encountered. Hence, the $\tilde{\boldsymbol{z}}$-update can be equivalently expressed as

$$\tilde{\boldsymbol{z}}^{(k+1)} = \tilde{\boldsymbol{z}}^{(k)} - H_z(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)})^{\dagger} \partial \mathcal{E}_{\tilde{\boldsymbol{z}}}^{(k)}(\tilde{\boldsymbol{z}}^{(k)}), \tag{4.24}$$

where the n-by-n matrix $H_z(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)})$ is defined, for an arbitrary $\delta\tilde{\boldsymbol{z}} \in \mathbb{R}^n$, by

$$\delta\tilde{\boldsymbol{z}}^{\top} H_z(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)})\delta\tilde{\boldsymbol{z}} = \sum_{j=1}^{n}\sum_{i=1}^{m} w_j^i(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)}) \left(\tilde{\boldsymbol{\rho}}_j^{(k+1)} \chi(\zeta_j^i(\tilde{\boldsymbol{z}}^{(k)}))\delta\tilde{\boldsymbol{z}}^{\top}(\zeta_j^i)'(\tilde{\boldsymbol{z}}^{(k)})\right)^2. \tag{4.25}$$

Since $\partial \mathcal{E}_{\tilde{\boldsymbol{z}}}^{(k)}(\tilde{\boldsymbol{z}}^{(k)}) = \frac{\partial \mathcal{E}}{\partial \tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)})$, the update (4.24) is equivalent to the Newton step (4.14).

In practice, $H_z(\tilde{\boldsymbol{\rho}}^{(k+1)}, \tilde{\boldsymbol{z}}^{(k)})^{\dagger} \partial \mathcal{E}_{\tilde{\boldsymbol{z}}}^{(k)}(\tilde{\boldsymbol{z}}^{(k)})$ in Eq. (4.24) is computed (inexactly) by preconditioned conjugate gradient iterations up to a relative tolerance of 10^{-4}. In our experiments, less than fifty iterations were enough to reach this stopping criterion. In our setup using $m = 8$ megapixel images and a recent i7 processor at 3.50 GHz with 32 GB of RAM, each depth update (the albedo one has negligible cost) required a few seconds, and 10 to 50 (depending on the initial estimate) updates were enough to reach convergence.

4.3 Convergence Analysis

In this subsection, we present two local convergence results of the proposed ARLS scheme. The proofs are provided in appendix.

When we write $A \succeq B$ (resp. $A \succ B$), this means that the difference matrix $A - B$ is positive semidefinite (resp. positive definite). The spectral radius of a matrix is denoted by $\mathrm{sr}(\cdot)$. The following lemma establishes the (local) majorization properties of H_ρ and H_z over the Hessian matrices $\frac{\partial^2 \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}}^2}$ and $\partial^2 \tilde{\mathcal{E}}$, respectively.

Lemma 1 *If the following condition holds at $(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)$:*

$$\zeta_j^i(\tilde{\boldsymbol{z}}^*) \neq 0 \quad \forall (i,j) \in [1,m] \times [1,n], \tag{4.26}$$

Then we have

$$H_\rho(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}}) \succeq \frac{\partial^2 \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}}^2}(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}}), \quad H_z(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}}) \succeq \partial^2 \tilde{\mathcal{E}}(\cdot; \tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}})(\tilde{\boldsymbol{z}}), \tag{4.27}$$

whenever $(\tilde{\boldsymbol{\rho}}, \tilde{\boldsymbol{z}})$ lies in some small neighborhood of $(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)$.

The next theorem contains the main result of our local convergence analysis.

Theorem 1 *Assume that, for some k, the iterate $(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\boldsymbol{z}}^{(k)})$ generated by Algorithm 1 is sufficiently close to some local minimizer $(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)$ where, in addition to (4.26), the following conditions hold:*

$$\frac{\partial \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) = \mathbf{0}, \qquad \frac{\partial \mathcal{E}}{\partial \tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) = \mathbf{0}, \tag{4.28}$$

$$\begin{bmatrix} \frac{\partial^2 \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}}^2}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) & \frac{\partial^2 \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}} \partial \tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) \\ \frac{\partial^2 \mathcal{E}}{\partial \tilde{\boldsymbol{\rho}} \partial \tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) & \frac{\partial^2 \mathcal{E}}{\partial \tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) \end{bmatrix} \succ \mathbf{O}, \tag{4.29}$$

$$\partial^2 \tilde{\mathcal{E}}(\cdot; \tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)(\tilde{\boldsymbol{z}}^*) \succ \mathbf{O}, \tag{4.30}$$

$$\mathrm{sr}\left(\partial^2 \tilde{\mathcal{E}}(\cdot; \tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)(\tilde{\boldsymbol{z}}^*)^{-1} \left(\frac{\partial^2 \mathcal{E}}{\partial \tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*) - \partial^2 \tilde{\mathcal{E}}(\cdot; \tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)(\tilde{\boldsymbol{z}}^*)\right)\right) < 1. \tag{4.31}$$

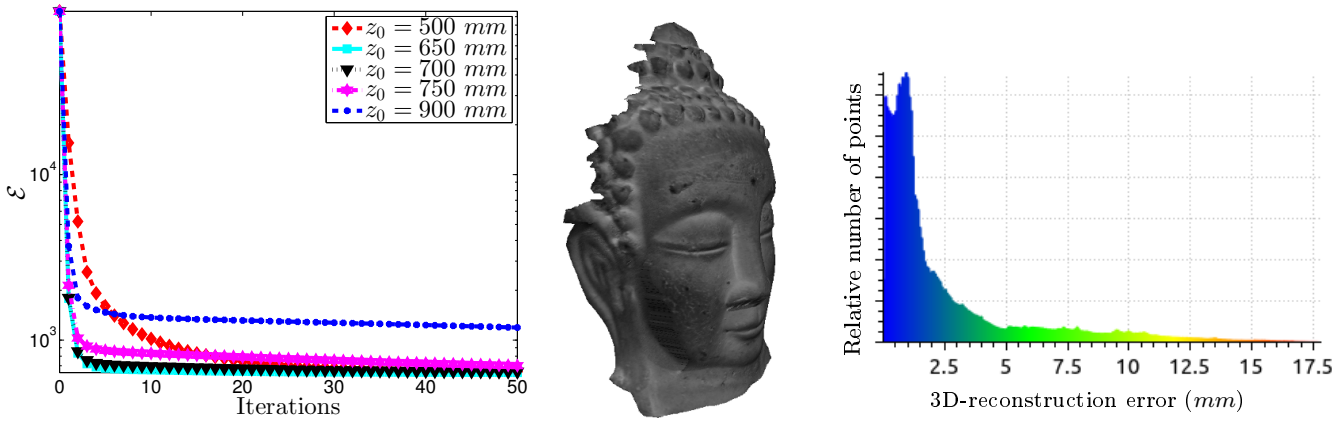
Then we have $\lim_{k\to\infty}(\tilde{\boldsymbol{\rho}}^{(k)}, \tilde{\boldsymbol{z}}^{(k)}) = (\tilde{\boldsymbol{\rho}}^*, \tilde{\boldsymbol{z}}^*)$.

As a remark, conditions (4.28) and (4.29) assumed in Theorem 1 are typically referred to as the second-order sufficient optimality conditions, while conditions (4.30) and (4.31) are similar to the local convergence criteria for Gauss-Newton method, see e.g. [20, Theorem 1]. We empirically found that they always seem satisfied i.e., the convergence of ARLS in form of Algorithm 1 is observed in all our experiments. If needed, these conditions may however be explicitly enforced by replacing $\{\cdot\}_+$ by its (smooth) proximity operator, and incorporating a line search step into ARLS, see [57].

4.4 Experimental Validation

For fair comparison with the methods discussed in Section 3, we first consider least-squares estimation without explicit self-shadows handling i.e., $\phi(x) = x^2$ and $\{x\}_+ = x$. The results in Figs. 13 and 14 show that, unlike the previous least-squares differential method from Section 3.2, the new scheme always converges towards a similar solution for a wide range of initial estimates.

Although the accuracy of the results obtained with this new scheme if not improved, convergence is guaranteed and the influence of the initialization is much reduced. Besides, it is straightforward to improve robustness by simply changing the definitions of the function ϕ and of the operator $\{\cdot\}_+$, while ensuring robustness of the ratio-based approach is not an easy task [40, 60]. Fig. 15 shows the result obtained using explicit self-shadows handling and Cauchy's M-estimator.



(a) (b) (c)

Fig. 13 (a) Evolution of the energy $\mathcal{E}$ of the proposed approach, defined in (4.4), using least-squares estimation, in function of the iterations, for the data of Fig. 2. As long as the initial scale is not over-estimated too much, the proposed scheme converges towards similar solutions for different initial estimates (cf. Fig. 14), though with different speeds. (b) 3D-model obtained at convergence, using $z_0 = 750\ mm$. (c) Histogram of point-to-point distances between (b) and the ground truth. As in the experiment of Fig. 12, the median value is 1.2 mm, yet this result is almost independent from the initialization, and is obtained using a provably convergent algorithm.

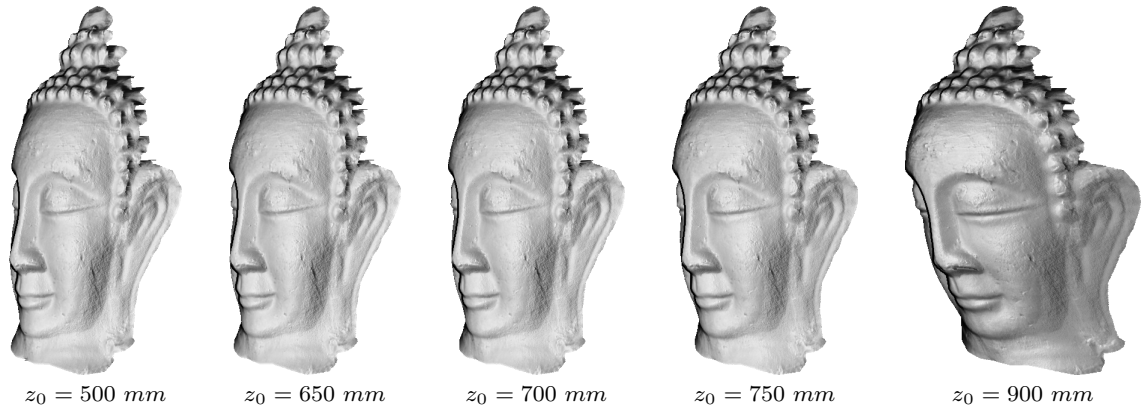
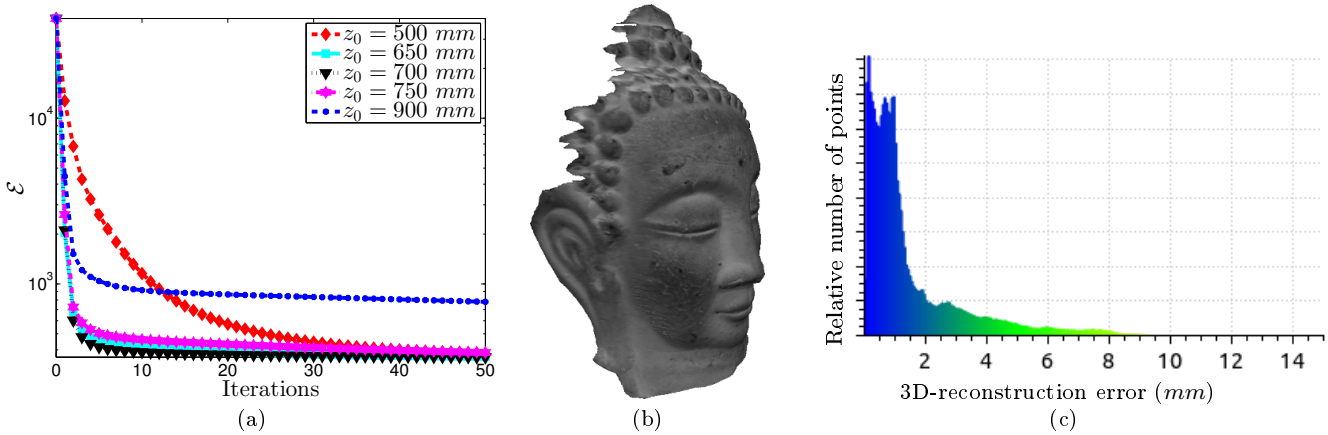


Fig. 14 3D-reconstructions after 50 iterations of the proposed scheme, taking as initial guess different fronto-parallel planes $z \equiv z_0$ and using least-squares estimation. Similar results are obtained whatever the initialization, at least as long as the initial scale is not over-estimated too much.



(a) (b) (c)

Fig. 15 Same as Fig. 13, but using explicit self-shadows handling and Cauchy's robust M-estimator. Despite the non-convexity of the estimator, convergence is similar to that obtained using least-squares. However, the median value of the 3D-reconstruction error is now 0.91 mm, which is to be compared with the value 1.2 mm obtained using least-squares (cf. Fig. 13).

5 Estimating Colored 3D-models by Photometric Stereo

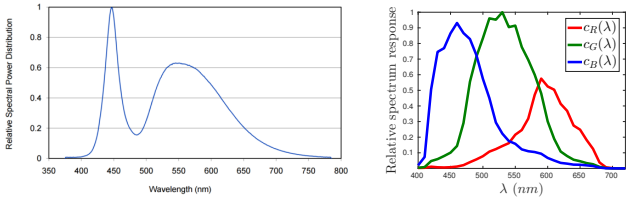
So far, we have considered gray level images. In this section, we extend our study to RGB-valued images, in order to estimate colored 3D-models using photometric stereo.

5.1 Spectral Dependency of the Luminous Flux Emitted by a LED

We need to introduce a spectral dependency in Model (2.5) to extend our study to color. It seems reasonable to limit this dependency to the intensity Φ (λ denotes the wavelength):

$$\mathbf{s}(\mathbf{x},\lambda) = \Phi(\lambda)\,\cos^{\mu}\theta\,\frac{\mathbf{x}_s-\mathbf{x}}{\|\mathbf{x}_s-\mathbf{x}\|^3}. \tag{5.1}$$

Model (5.1) is more complex than Model (2.5), because the intensity Φ_0, which is a real parameter, has been replaced by the *emission spectrum* $\Phi(\lambda)$, which is a function (cf. Fig. 16-a). The calibration of $\Phi(\lambda)$ could be achieved by using a spectrometer, but we will show how to extend the procedure from Section 2.2, which requires nothing else than a camera and two calibration patterns.



(a) (b)

Fig. 16 (a) Emission spectrum $\Phi(\lambda)$ of the LEDs used (source: http://www.lumileds.com/uploads/28/DS64-pdf). (b) Camera response functions in the three channels R, G, B, for the Canon EOS 50D camera [32] (which is similar to the Canon EOS 7D we use). Our extension to RGB images of the calibration procedure from Section 2.2 requires nothing else than a camera and two calibration patterns. We need therefore none of these diagrams in practice.

In a point $\mathbf{x}$ of a Lambertian surface with albedo $\rho(\mathbf{x})$, under the illumination described by the lighting vector $\mathbf{s}(\mathbf{x})$, we get from (2.6), (2.7) and (2.8) the expression of the illuminance $\epsilon(\mathbf{p})$ of the image plane in the pixel $\mathbf{p}$ conjugate to $\mathbf{x}$:

$$\epsilon(\mathbf{p}) = \beta\,\cos^4\alpha(\mathbf{p})\,\frac{\rho(\mathbf{x})}{\pi}\,\mathbf{s}(\mathbf{x})\cdot\mathbf{n}(\mathbf{x}). \tag{5.2}$$

This expression is easily extended to the case where $\mathbf{s}(\mathbf{x})$ and $\rho(\mathbf{x})$ depend on λ:

$$\epsilon(\mathbf{p},\lambda) = \beta\,\cos^4\alpha(\mathbf{p})\,\frac{\rho(\mathbf{x},\lambda)}{\pi}\,\mathbf{s}(\mathbf{x},\lambda)\cdot\mathbf{n}(\mathbf{x}). \tag{5.3}$$

The bijective correspondence between the points $\mathbf{x}$ and the pixels $\mathbf{p}$ allows us to denote $\rho(\mathbf{p},\lambda)$ and $\mathbf{n}(\mathbf{p})$, in lieu of $\rho(\mathbf{x},\lambda)$ and $\mathbf{n}(\mathbf{x})$. In addition, the light effectively received by each cell goes through a colored filter characterized by its *transmission spectrum* $c_\star(\lambda)$, $\star\in\{R,G,B\}$, whose maximum lies, respectively, in the red, green and blue ranges (cf. Fig. 16-b). To define the *color levels* $I_\star(\mathbf{p})$, $\star\in\{R,G,B\}$, by similarity with the expression (2.10) of the (corrected) gray level $I(\mathbf{p})$, we must multiply (5.3) by $c_\star(\lambda)$, and integrate over the entire spectrum:

$$I_\star(\mathbf{p}) = \frac{\gamma\,\beta}{\pi}\left[\int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\rho(\mathbf{p},\lambda)\,\mathbf{s}(\mathbf{x},\lambda)\,\mathrm{d}\lambda\right]\cdot\mathbf{n}(\mathbf{p}). \tag{5.4}$$

Using a Lambertian calibration pattern which is uniformly white i.e., such that $\rho(\mathbf{p},\lambda)\equiv\rho_0$, allows us to rewrite (5.4) as follows:

$$I_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho_0}{\pi}\left[\int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\mathbf{s}(\mathbf{x},\lambda)\,\mathrm{d}\lambda\right]\cdot\mathbf{n}(\mathbf{p}), \tag{5.5}$$

which is indeed an extension of (2.10) to RGB images, since (5.5) can be rewritten

$$I_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho_0}{\pi}\,\mathbf{s}_\star(\mathbf{x})\cdot\mathbf{n}(\mathbf{p}), \tag{5.6}$$

provided that the three *colored lighting vectors* $\mathbf{s}_\star(\mathbf{x})$ are defined as follows:

$$\mathbf{s}_\star(\mathbf{x}) = \int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\mathbf{s}(\mathbf{x},\lambda)\,\mathrm{d}\lambda, \quad \star\in\{R,G,B\}. \tag{5.7}$$

Plugging (5.1) into (5.7), we obtain an extension of Model (2.5) to RGB images:

$$\mathbf{s}_\star(\mathbf{x}) = \Phi_\star\,\cos^{\mu}\theta\,\frac{\mathbf{x}_s-\mathbf{x}}{\|\mathbf{x}_s-\mathbf{x}\|^3}, \qquad \star\in\{R,G,B\}, \tag{5.8}$$

where the three *colored intensities* $\Phi_\star$ are defined as follows:

$$\Phi_\star = \int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\Phi(\lambda)\,\mathrm{d}\lambda, \qquad \star\in\{R,G,B\}. \tag{5.9}$$

The spectral dependency of the lighting vector $\mathbf{s}(\mathbf{x},\lambda)$ expressed in (5.1) is thus *partially* described by Model (5.8), which contains nine parameters: three for the coordinates of $\mathbf{x}_s$, two for the unit-length vector $\mathbf{n}_s$, plus the three colored intensities Φ_R, Φ_G, Φ_B, and the anisotropy parameter μ. Nonetheless, since the definition (5.9) of $\Phi_\star$ depends on $c_\star(\lambda)$, it follows that the parameters Φ_R, Φ_G and Φ_B are not really characteristic of the LED, but of the couple camera-LED.

5.2 Spectral Calibration of the Luminous Flux Emitted by a LED

We use again the planar Lambertian calibration pattern from Section 2.2. Since it is convex, the incident light comes solely from the LED. We can thus replace $\mathbf{s}_\star(\mathbf{x})$ by its definition (5.8) in the expression (5.6) of the color level $I_\star(\mathbf{p})$. Assuming that $\mathbf{x}_s$ is estimated by triangulation and that the anisotropy parameter μ is provided by the manufacturer, we then have to solve, for each channel $\star \in \{R,G,B\}$, the following problem, which is an extension of Problem (2.17) (q is the number of poses of the Lambertian calibration pattern):

$$\min_{\mathbf{m}_{s,\star}} \sum_{j=1}^{q} \sum_{\mathbf{p}\in\Omega^j} \left[\mathbf{m}_{s,\star} \cdot (\mathbf{x}^j - \mathbf{x}_s) - \left[I^j_\star(\mathbf{p}) \, \frac{\|\mathbf{x}_s - \mathbf{x}^j\|^{3+\mu}}{(\mathbf{x}_s - \mathbf{x}^j)\cdot \mathbf{n}^j} \right]^{1/\mu} \right]^2, \tag{5.10}$$

where $\mathbf{m}_{s,\star}$ is defined by analogy with $\mathbf{m}_s$ (cf. (2.16)):

$$\mathbf{m}_{s,\star} = {\Psi_\star}^{1/\mu} \, \mathbf{n}_s, \tag{5.11}$$

and $\Psi_\star$ is defined by analogy with Ψ (cf. (2.12)):

$$\Psi_\star = \gamma \, \beta \, \frac{\rho_0}{\pi} \, \Phi_\star. \tag{5.12}$$

Each problem (5.10) allows us to estimate a colored intensity Φ_R, Φ_G or Φ_B (up to a common factor) and the principal direction $\mathbf{n}_s$, which is thus estimated three times. Table 1 groups the values obtained for one of the LEDs of our setup. The three estimates of $\mathbf{n}_s$ are consistent, but instead of arbitrarily choosing one of them, we compute the weighted mean of these estimates, using spherical coordinates to ensure that the unit-length constraint is preserved.

Red channel	Green channel	Blue channel
$\widehat{\mathbf{n}}_{s,R} = \begin{bmatrix} 0.205 \\ -0.757 \\ 0.621 \end{bmatrix}$	$\widehat{\mathbf{n}}_{s,G} = \begin{bmatrix} 0.194 \\ -0.769 \\ 0.608 \end{bmatrix}$	$\widehat{\mathbf{n}}_{s,B} = \begin{bmatrix} 0.188 \\ -0.844 \\ 0.503 \end{bmatrix}$
$\widehat{\Psi}_R = 3.10 \times 10^7$	$\widehat{\Psi}_G = 5.49 \times 10^7$	$\widehat{\Psi}_B = 3.37 \times 10^7$

Table 1 Parameters of one of the LEDs of our setup, estimated by solving (4.10) with respect to each color channel.

In Table 1, the values of $\widehat{\Psi}_R$, $\widehat{\Psi}_G$ and $\widehat{\Psi}_B$ are given without units because, from the definition (5.12) of $\Psi_\star$, only their relative values are meaningful. As it happens, the value of $\widehat{\Psi}_G$ is roughly twice as much as those of $\widehat{\Psi}_R$ and $\widehat{\Psi}_B$, but this does not mean that $\Phi(\lambda)$ is twice higher in the green range than in the red or in the blue ranges, since the definition (5.9) of a given colored intensity $\Phi_\star$ depends on the transmission spectrum $c_\star(\lambda)$ in the considered channel.

Our calibration procedure relies on the assumption that the calibration pattern is uniformly white i.e., that $\rho(\mathbf{p},\lambda) \equiv \rho_0$, which may be inexact, yet in no way does this question our rationale. Indeed, if we assume that the color of "white" cells from the Lambertian checkerboard is uniform i.e., $\rho(\mathbf{p},\lambda) = \rho(\lambda)$, $\forall \mathbf{p} \in \Omega^j$, and if we denote ρ_0 the maximum value of $\rho(\lambda)$, Eq. (5.5) is still valid, provided that $c_\star(\lambda)$ is replaced by the function $c'_\star(\lambda)$ defined as follows:

$$c'_\star(\lambda) = \frac{\rho(\lambda)}{\rho_0} \, c_\star(\lambda). \tag{5.13}$$

Of course, similar arguments can be used for the rest of the rationale, but we must remember that each colored intensity $\Phi_\star$, which depends on the transmission spectrum $c_\star(\lambda)$ by its definition (5.9), also depends on the color of the paper upon which the checkerboard is printed. We are now able to extend to RGB images our photometric stereo model from Section 2.3, but we can already predict that the color of the paper will have some influence on the estimated color and on the estimated shape of the observed scene.

5.3 Photometric Stereo under Colored Point Light Source Illumination

If we pretend to extend Model (2.10) to RGB images, then it must be possible to write the color level at $\mathbf{p}$, in each channel $\star \in \{R,G,B\}$, in the following manner:

$$I_\star(\mathbf{p}) = \gamma \, \beta \, \frac{\rho_\star(\mathbf{p})}{\pi} \, \mathbf{s}_\star(\mathbf{x}) \cdot \mathbf{n}(\mathbf{p}), \tag{5.14}$$

where the *colored albedos* $\rho_\star(\mathbf{p})$ are defined by extension to RGB images of the albedo $\rho(\mathbf{p})$. Equating both expressions of $I_\star(\mathbf{p})$ given in (5.4) and in (5.14), and using the definition (5.7) of $\mathbf{s}_\star(\mathbf{x})$, we obtain the following equality:

$$\rho_\star(\mathbf{p}) \int_{\lambda=0}^{+\infty} c_\star(\lambda) \, \mathbf{s}(\mathbf{x},\lambda) \, \mathrm{d}\lambda = \int_{\lambda=0}^{+\infty} c_\star(\lambda) \, \rho(\mathbf{p},\lambda) \, \mathbf{s}(\mathbf{x},\lambda) \, \mathrm{d}\lambda. \tag{5.15}$$

When the surface is illuminated by a LED, we can replace $\mathbf{s}(\mathbf{x},\lambda)$ in (5.15) by Model (5.1). This gives us the following expression for the colored albedos:

$$\rho_\star(\mathbf{p}) = \frac{\displaystyle\int_{\lambda=0}^{+\infty} c_\star(\lambda) \, \rho(\mathbf{p},\lambda) \, \Phi(\lambda) \, \mathrm{d}\lambda}{\displaystyle\int_{\lambda=0}^{+\infty} c_\star(\lambda) \, \Phi(\lambda) \, \mathrm{d}\lambda}, \quad \star \in \{R,G,B\}, \tag{5.16}$$

which is the mean of $\rho(\mathbf{p},\lambda)$ over the entire spectrum, weighted by the product $c_\star(\lambda)\,\Phi(\lambda)$. In addition, although the transmission spectrum $c_\star(\lambda)$ depends only

on the camera, the emission spectrum $\Phi(\lambda)$ usually varies from one LED to another. Thus, generalizing photometric stereo under point light source illumination to RGB images requires to superscript the colored albedos by the LED index i. Hence, it seems that we have to solve, in each pixel $\mathbf{p} \in \Omega$, the following problem:

$$I^i_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho^i_\star(\mathbf{p})}{\pi}\,\mathbf{s}^i_\star(\mathbf{x})\cdot\mathbf{n}(\mathbf{p}),\ i \in [1,m],\ \star \in \{R,G,B\}. \tag{5.17}$$

System (5.17) is underdetermined, because it contains $3m$ equations with $3m+3$ unknowns: one colored albedo $\rho^i_\star(\mathbf{p})$ per equation, the depth $z(\mathbf{p})$ of the 3D-point $\mathbf{x}$ conjugate to $\mathbf{p}$ (from which we get the coordinates of $\mathbf{x}$), and the normal $\mathbf{n}(\mathbf{p})$. Apart from this numerical difficulty, the dependency on i of the colored albedos is puzzling: while it is clear that the albedo is a photometric characteristic of the surface, independent from the lighting, it should go the same for the colored albedos. This shows that the extension to RGB images of photometric stereo is potentially intractable in the general case. However, such an extension is known to be possible in two specific cases [55]:

- For a white surface i.e., when $\rho(\mathbf{p},\lambda) = \rho(\mathbf{p})$, we deduce from (5.16) that $\rho_R(\mathbf{p}) = \rho_G(\mathbf{p}) = \rho_B(\mathbf{p}) = \rho(\mathbf{p})$. Problem (5.17) is thus written:

$$I^i_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho(\mathbf{p})}{\pi}\,\mathbf{s}^i_\star(\mathbf{x})\cdot\mathbf{n}(\mathbf{p}),\ i \in [1,m],\ \star \in \{R,G,B\}. \tag{5.18}$$

  The case where lighting is directional i.e., independent from $\mathbf{x}$, is very interesting, because the number of unknowns of this linear system is then equal to three: a single RGB image is thus enough to ensure that the problem is well-posed, as soon as $\mathbf{s}_R$, $\mathbf{s}_G$ and $\mathbf{s}_B$ are non-coplanar. This well-known case, which dates back to the 90's [34], has been applied by Hernandez et al. to real-time 3D-reconstruction of a deformable (white) surface [22].
- When the sources are white i.e., when $\Phi(\lambda) \equiv \Phi_0$, (5.16) gives:

$$\rho_\star(\mathbf{p}) = \frac{\displaystyle\int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\rho(\mathbf{p},\lambda)\,\mathrm{d}\lambda}{\displaystyle\int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\mathrm{d}\lambda},\ \star \in \{R,G,B\}. \tag{5.19}$$

  Since this expression is independent from i, Problem (5.17) is rewritten:

$$I^i_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho_\star(\mathbf{p})}{\pi}\,\mathbf{s}^i_\star(\mathbf{x})\cdot\mathbf{n}(\mathbf{p}),\ i \in [1,m],\ \star \in \{R,G,B\}. \tag{5.20}$$

  In (5.20), the lighting vector $\mathbf{s}^i_\star(\mathbf{x})$ really depends on the channel $\star$, although $\Phi(\lambda) \equiv \Phi_0$ since the definition (5.9) of $\Phi_\star$ depends on the transmission spectrum $c_\star(\lambda)$ and the same holds true for the definition (5.8) of $\mathbf{s}_\star(\mathbf{x})$. System (5.20), which has $3m$ equations and six unknowns, is well-posed if $m \geqslant 3$ (the case where $m = 2$ is ill-posed under directional lighting [56], but this remains unclear in the case of nearby point light sources).

Another case where the colored albedos are independent from i is when the m LEDs all share the same emission spectrum $\Phi(\lambda)$. For the setup of Fig. 2-a, the $m = 8$ LEDs probably do not satisfy exactly this constraint, although they come from the same batch, yet this assumption seems more realistic than that of "white sources", and it allows us to better justify the use of Model (5.20) in photometric stereo. Using Model (5.8) for the three colored lighting vectors $\mathbf{s}_\star(\mathbf{x})$, $\star \in \{R,G,B\}$, induced by a LED, Problem (5.20) is thus rewritten in the following way:

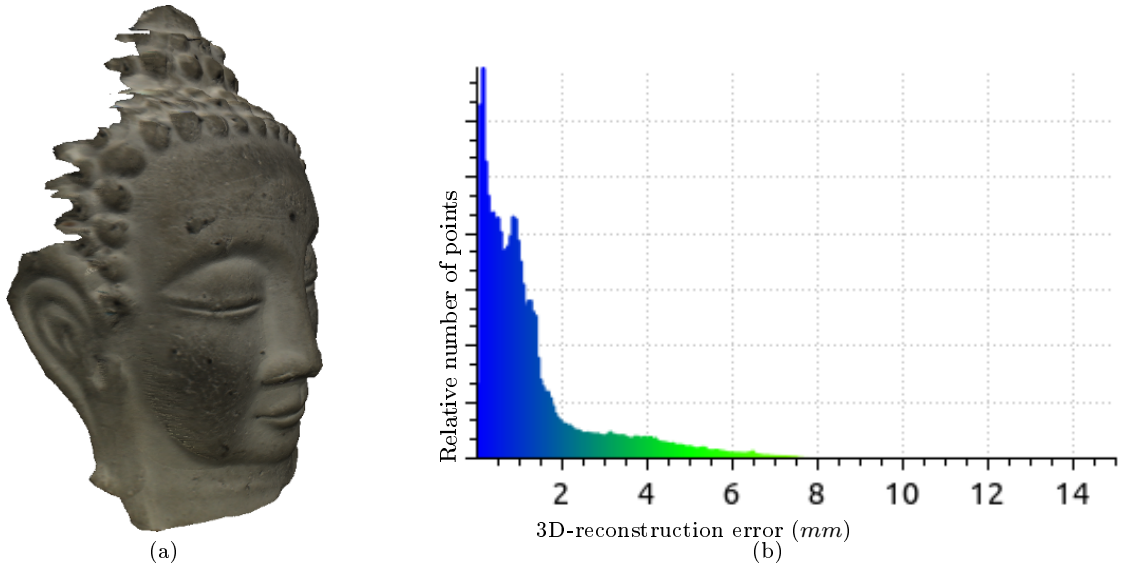
$$I^i_\star(\mathbf{p}) = \Psi^i_\star\,\frac{\rho_\star(\mathbf{p})}{\rho_0}\left[\mathbf{n}^i_s\cdot\frac{\mathbf{x}-\mathbf{x}^i_s}{\|\mathbf{x}-\mathbf{x}^i_s\|}\right]^{\mu^i}\frac{(\mathbf{x}^i_s-\mathbf{x})\cdot\mathbf{n}(\mathbf{p})}{\|\mathbf{x}^i_s-\mathbf{x}\|^3},\quad i \in [1,m],\ \star \in \{R,G,B\}. \tag{5.21}$$

which is an extension to RGB images of the photometric stereo model (2.20). In our experiments, we use (5.21) to model both the spectral dependency of the albedo and that of the luminous fluxes, although we must keep in mind that this writing is perfectly justified only if the sources share the same spectrum.

The calibration procedure described in Section 5.2 provides us with the values of the parameters $\mathbf{x}^i_s$, $\mathbf{n}^i_s$ and $\Psi^i_\star$, $i \in [1,m]$, and the parameters μ^i, $i \in [1,m]$, are provided by the manufacturer. The unknowns of System (5.21) are thus the depth $z(\mathbf{p})$ of $\mathbf{x}$, the normal $\mathbf{n}(\mathbf{p})$ and the three colored albedos $\rho_\star(\mathbf{p})$, $\star \in \{R,G,B\}$. Hence, resorting to RGB images allows us to replace the system (2.20) of m equations with four unknowns, by the system (5.21) of $3m$ equations with six unknowns. We can thus expect more accurate results.

5.4 Solving Colored Photometric Stereo under Point Light Source Illumination

The alternating strategy from Section 3.1 is not straightforward to adapt to the case of RGB-valued images, because the albedo is channel-dependent, while the normal vector is not. Principal component analysis could be employed [5], but we already know from Section 3



(a) (b)

Fig. 17 (a) 3D-model estimated from the $m = 8$ images of Fig. 2, which are RGB images. (b) Histogram of the distances between this 3D- shape and the ground truth. Using RGB images improves the result, in comparison with the experiment of Fig. 15: the median of the point-to-point distances to the ground truth is now equal to 0.85 mm.

that a differential approach should be preferred anyways. A PDE-based approach similar to that of Section 3.2 is advocated in [55]: ratios between color levels can be computed in each channel $\star \in \{R, G, B\}$, thus eliminating the colored albedos $\rho_\star(\mathbf{p})$ and obtaining a system of PDEs in z similar to (3.24). With this approach, the PDEs to solve remain quasi-linear, unlike in [29]. Yet, we know that the solution strongly depends on the initialization.

On the other hand, it is straightforward to adapt the method recommended in Section 4, by turning the discrete optimization problem (4.4) into

$$\min_{\tilde{\boldsymbol{\rho}}_R, \tilde{\boldsymbol{\rho}}_G, \tilde{\boldsymbol{\rho}}_B, \tilde{\boldsymbol{z}}} \sum_{\star \in \{R,G,B\}} \sum_{j=1}^{n} \sum_{i=1}^{m} \phi\left(r^i_{\star,j}(\tilde{\boldsymbol{\rho}}_\star, \tilde{\boldsymbol{z}})\right), \tag{5.22}$$

with the following new definition of the residual, which uses straightforward notations for the channel dependencies:

$$r^i_{\star,j}(\tilde{\boldsymbol{\rho}}_\star, \tilde{\boldsymbol{z}}) = \tilde{\rho}_{\star,j} \left\{\zeta^i_{\star,j}(\tilde{\boldsymbol{z}})\right\}_+ - I^i_{\star,j}, \tag{5.23}$$

$$\zeta^i_{\star,j}(\tilde{\boldsymbol{z}}) = \mathbf{t}^i_{\star,j}(z_j)^\top \mathbf{Q}_j^\top \begin{bmatrix} (\nabla \tilde{\boldsymbol{z}})_j \\ -1 \end{bmatrix}. \tag{5.24}$$

The actual solution of (5.22) follows immediately from the algorithm described in Section 4.2. The depth update simply uses three times more equations, which improves its robustness, while the estimation of each vector $\tilde{\boldsymbol{\rho}}_\star$, $\star \in \{R, G, B\}$, is carried out independently in each channel in exactly the same way as in Section 4.2. Eventually, the "real" albedos are recovered using (4.1).

Since the depth estimation now uses more data, the 3D-model of Fig. 17, which uses RGB images, is improved in two ways, in comparison with that of Fig. 15: it is not only colored, but also more accurate.

6 Conclusion and Perspectives

In this article, we describe a photometric stereo-based 3D-reconstruction setup using LEDs as light sources. We first model the luminous flux emitted by a LED, then the resulting photometric stereo problem. We present a practical procedure for calibrating photometric stereo under point light source illumination, and eventually, we study several numerical solutions. Existing methods are based either on alternating estimation of normals and depth, or on direct depth estimation using image ratios. Both these methods have their own advantages, but their convergence is not established. Hence, we introduce a new solution based on alternating reweighted least-squares, which is provably convergent. Finally, we extend the whole study to RGB images.

The result of Fig. 18 suggests that our goal i.e., the estimation of colored 3D-models of faces by photometric stereo, has been reached. Of course, many other types of 3D-scanners exist, but ours relies only on materials which are easy to obtain: a relatively mainstream camera, eight LEDs and an Arduino controller to synchronize the LEDs with the shutter release. Another significant advantage of our 3D-scanner is that it also estimates the albedo.

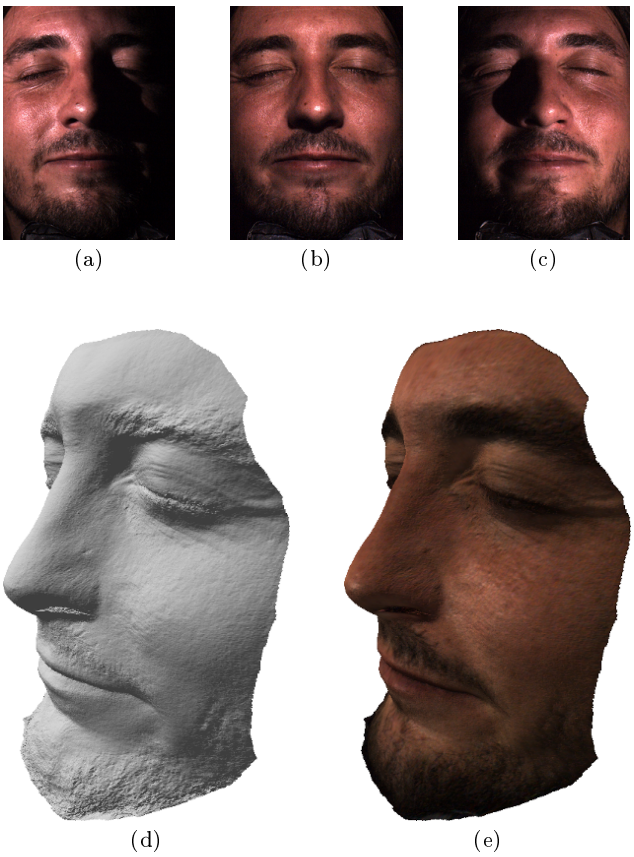


Fig. 18 (a-b-c) Three RGB images (out of $m = 8$) of a face captured by our setup. (d) Estimated 3D-shape. (e) Colored 3D-model. Since their estimation is relative to the calibration object, the colored albedos of the 3D-model may appear different from the colors of the images.

However, there may still be some points where the shape, and therefore the albedo, are poorly estimated. In the example of Fig. 19, the area under the nose, which is dimly lit, is poorly reconstructed (this problem does not appear in the example of Fig. 18, because the face is oriented in such a way that it is "well" illuminated). Although such artifacts remain confined, thanks to the robust estimation, future extensions of our work could get rid of them by resorting to an additional regularization term in the variational model.

Besides dealing with these defects, other questions arise. In particular, could we extend our 3D-scanner to full 3D-reconstruction, by coupling the proposed method with multi-view 3D-reconstruction techniques [23]? Aside from obtaining a more complete 3D-reconstruction, this would circumvent the diffult problem of handling possible discontinuities in a depth map, although Fig. 19 suggests that employing a non-convex estimator already partly allows the recovery of such sharp structures [14].

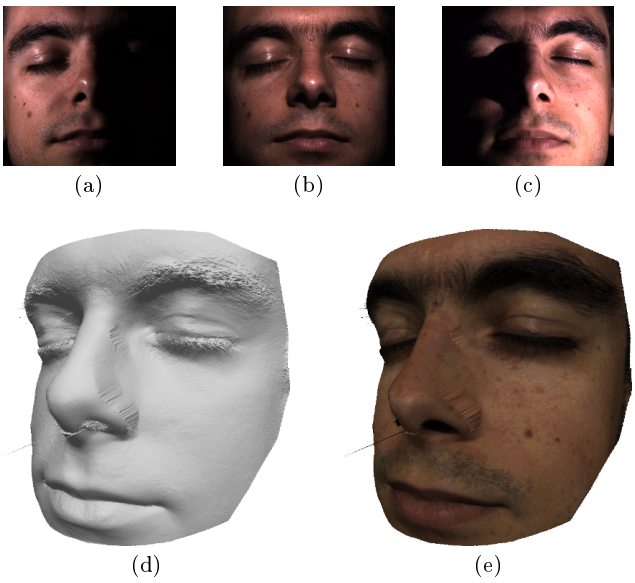


Fig. 19 (a-b-c) Three images (out of $m = 8$) of a face. (d) Estimated 3D-shape. (e) Colored 3D-model. The 3D-reconstruction is not satisfactory under the nose, which is a dimly lit area. Robustness of the proposed method to shadows could still be improved.

Eventually, the proposed numerical framework could be extended in order to automatically refine calibration. Several steps in that direction were already achieved in [37,43,50,57], but either without convergence analysis [37,43,50] or in the restricted case where only the source intensities are refined [57]. Providing a provably convergent method for uncalibrated photometric stereo under point light source illumination would thus constitute a natural extension of our work.

A Proof of Lemma 1

Proof First note that, under the condition (4.26), the function $\mathcal{E}(\cdot,\tilde{\boldsymbol{z}})$ (resp. $\tilde{\mathcal{E}}(\cdot;\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})$) is twice continuously differentiable at $\tilde{\boldsymbol{\rho}}$ (resp. $\tilde{\boldsymbol{z}}$), whenever $(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})$ is sufficiently close to $(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)$. The corresponding second-order derivatives are calculated as follows:

$$\delta\tilde{\boldsymbol{\rho}}^\top \frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}^2}(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})\delta\tilde{\boldsymbol{\rho}} = \sum_{j=1}^{n}\sum_{i=1}^{m}\phi''(r_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}}))\left(\delta\tilde{\boldsymbol{\rho}}_j\{\zeta_j^i(\tilde{\boldsymbol{z}})\}_+\right)^2, \tag{A.1}$$

$$\delta\tilde{\boldsymbol{z}}^\top\partial^2\tilde{\mathcal{E}}(\cdot;\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}})(\tilde{\boldsymbol{z}})\delta\tilde{\boldsymbol{z}} = \sum_{j=1}^{n}\sum_{i=1}^{m}\phi''(r_j^i(\tilde{\boldsymbol{\rho}},\tilde{\boldsymbol{z}}))\left(\tilde{\boldsymbol{\rho}}_j\chi(\zeta_j^i(\tilde{\boldsymbol{z}}))\delta\tilde{\boldsymbol{z}}^\top(\zeta_j^i)'(\tilde{\boldsymbol{z}})\right)^2. \tag{A.2}$$

Comparing the above two formulas with (4.21) and (4.25), the conclusion follows from condition (4.7). □

B Proof of Theorem 1

Proof First note that condition (4.29) implies that

$$\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\succ\boldsymbol{O}, \tag{B.1}$$

$$\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)-\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-1}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\succ\boldsymbol{O}. \tag{B.2}$$

Utilizing Lemma 1 in conjunction with (B.2) and (4.30), we obtain

$$H_\rho(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\succ\boldsymbol{O},\qquad H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\succ\boldsymbol{O}, \tag{B.3}$$

$$\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)-\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_\rho(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-1}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\succ\boldsymbol{O}. \tag{B.4}$$

Now consider the iteration

$$\begin{aligned}\tilde{\boldsymbol{z}}^{k+1}&=\tilde{\boldsymbol{z}}^k-H_z(\tilde{\boldsymbol{\rho}}^{k+1},\tilde{\boldsymbol{z}}^k)^{-1}\frac{\partial\mathcal{E}}{\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^{k+1},\tilde{\boldsymbol{z}}^k)\\&=\tilde{\boldsymbol{z}}^k-H_z\left(\tilde{\boldsymbol{\rho}}^k-H_\rho(\tilde{\boldsymbol{\rho}}^k,\tilde{\boldsymbol{z}}^k)^{-1}\frac{\partial\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}^k,\tilde{\boldsymbol{z}}^k),\tilde{\boldsymbol{z}}^k\right)^{-1}\\&\qquad\frac{\partial\mathcal{E}}{\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^k-H_\rho(\tilde{\boldsymbol{\rho}}^k,\tilde{\boldsymbol{z}}^k)^{-1}\frac{\partial\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}}(\tilde{\boldsymbol{\rho}}^k,\tilde{\boldsymbol{z}}^k),\tilde{\boldsymbol{z}}^k)\end{aligned} \tag{B.5}$$

as a map $\tilde{\boldsymbol{z}}^k\mapsto\tilde{\boldsymbol{z}}^{k+1}$. By the Ostrowski theorem [49, Proposition 10.1.3], the local convergence of $\{\tilde{\boldsymbol{z}}^k\}$ to $\tilde{\boldsymbol{z}}^*$ follows if the spectral radius of the Jacobian

$$\begin{aligned}\frac{\partial\tilde{\boldsymbol{z}}^{k+1}}{\partial\tilde{\boldsymbol{z}}^k}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)&=\mathrm{id}-H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-1}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\\&\quad+H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-1}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_\rho(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-1}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\end{aligned} \tag{B.6}$$

is strictly less than 1. Using the similarity transform with $H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{\frac{1}{2}}$, we derive:

$$\begin{aligned}&\mathrm{sr}\left(\frac{\partial\tilde{\boldsymbol{z}}^{k+1}}{\partial\tilde{\boldsymbol{z}}^k}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\right)\\&=\mathrm{sr}\left(H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{\frac{1}{2}}\frac{\partial\tilde{\boldsymbol{z}}^{k+1}}{\partial\tilde{\boldsymbol{z}}^k}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\right)\qquad\text{(B.7)}\\&=\mathrm{sr}\Big(\mathrm{id}-H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\\&\qquad+H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_\rho(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-1}\\&\qquad\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\Big)\qquad\text{(B.8)}\\&=\sup_{\|\mathbf{v}\|=1}\Big|\|\mathbf{v}\|^2\\&\quad-\mathbf{v}^\top H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\mathbf{v}\\&\quad+\mathbf{v}^\top H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_\rho(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-1}\\&\qquad\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\mathbf{v}\Big|.\qquad\text{(B.9)}\end{aligned}$$

It follows from condition (4.31) that

$$\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\prec2\partial^2\tilde{\mathcal{E}}(\cdot;\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)(\tilde{\boldsymbol{z}}^*)\preceq2H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*), \tag{B.10}$$

and hence

$$\mathrm{id}-H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\succ-\mathrm{id}. \tag{B.11}$$

Consequently, there exists $\epsilon_1\in(0,1)$ such that the following inequality holds for an arbitrary $\mathbf{v}$:

$$\begin{aligned}&\|\mathbf{v}\|^2-\mathbf{v}^\top H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\mathbf{v}\\&\qquad\geq-(1-\epsilon_1)\|\mathbf{v}\|^2.\end{aligned} \tag{B.12}$$

Meanwhile, condition (B.4) implies that, for some $\epsilon_2\in(0,1)$:

$$\begin{aligned}&\mathbf{v}^\top H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\mathbf{v}\\&\quad-\mathbf{v}^\top H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_\rho(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-1}\\&\qquad\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\mathbf{v}\qquad\text{(B.13)}\\&=(H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\mathbf{v})^\top\\&\quad\Big(\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{z}}^2}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)-\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)H_\rho(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-1}\frac{\partial^2\mathcal{E}}{\partial\tilde{\boldsymbol{\rho}}\partial\tilde{\boldsymbol{z}}}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\Big)\\&\quad(H_z(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)^{-\frac{1}{2}}\mathbf{v})\qquad\text{(B.14)}\\&\geq\epsilon_2\|\mathbf{v}\|^2.\qquad\text{(B.15)}\end{aligned}$$

Altogether, we conclude

$$\mathrm{sr}\left(\frac{\partial\tilde{\boldsymbol{z}}^{k+1}}{\partial\tilde{\boldsymbol{z}}^k}(\tilde{\boldsymbol{\rho}}^*,\tilde{\boldsymbol{z}}^*)\right)\leq1-\min(\epsilon_1,\epsilon_2), \tag{B.16}$$

and hence the convergence of $\{\tilde{\boldsymbol{z}}^k\}$. The convergence of $\{\tilde{\boldsymbol{\rho}}^k\}$ to $\tilde{\boldsymbol{\rho}}^*$ follows from a similar argument. □

References

1. Ackermann, J., Fuhrmann, S., Goesele, M.: Geometric Point Light Source Calibration. In: Proceedings of the 18th International Workshop on Vision, Modeling & Visualization, pp. 161–168. Lugano, Switzerland (2013) 3, 5
2. Ahmad, J., Sun, J., Smith, L., Smith, M.: An improved photometric stereo through distance estimation and light vector optimization from diffused maxima region. Pattern Recognition Letters **50**, 15–22 (2014) 3, 8
3. Angelopoulou, M.E., Petrou, M.: Uncalibrated flatfielding and illumination vector estimation for photometric stereo face reconstruction. Machine Vision and Applications **25**(5), 1317–1332 (2013) 3, 4
4. Aoto, T., Taketomi, T., Sato, T., Mukaigawa, Y., Yokoya, N.: Position estimation of near point light sources using a clear hollow sphere. In: Proceedings of the 21st International Conference on Pattern Recognition, pp. 3721–3724. Tsukuba, Japan (2012) 3, 5
5. Barsky, S., Petrou, M.: The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(10), 1239–1252 (2003) 21
6. Basri, R., Jacobs, D.W.: Lambertian reflectance and linear subspaces. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(2), 218–233 (2003) 4

7. Bennahmias, M., Arik, E., Yu, K., Voloshenko, D., Chua, K., Pradhan, R., Forrester, T., Jannson, T.: Modeling of non-Lambertian sources in lighting applications. In: Optical Engineering and Applications, *Proceedings of SPIE*, vol. 6669. San Diego, USA (2007) 3
8. Bony, A., Bringier, B., Khoudeir, M.: Tridimensional reconstruction by photometric stereo with near spot light sources. In: Proceedings of the 21st European Signal Processing Conference. Marrakech, Morocco (2013) 3, 8
9. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. Foundations and Trends in Machine Learning **3**(1), 1–122 (2011) 12, 13
10. Bringier, B., Bony, A., Khoudeir, M.: Specularity and shadow detection for the multisource photometric reconstruction of a textured surface. Journal of the Optical Society of America A **29**(1), 11–21 (2012) 8
11. Ciortan, I., Pintus, R., Marchioro, G., Daffara, C., Giachetti, A., Gobbetti, E.: A Practical Reflectance Transformation Imaging Pipeline for Surface Characterization in Cultural Heritage. In: Proceedings of the 14th Eurographics Workshop on Graphics and Cultural Heritage. Genova, Italy (2016) 3, 5
12. Clark, J.J.: Active photometric stereo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 29–34 (1992) 3
13. Collins, T., Bartoli, A.: 3D Reconstruction in Laparoscopy with Close-Range Photometric Stereo. In: Proceedings of the 15th International Conference on Medical Imaging and Computer Assisted Intervention, pp. 634–642. Nice, France (2012) 3, 8
14. Durou, J.D., Aujol, J.F., Courteille, F.: Integrating the Normal Field of a Surface in the Presence of Discontinuities. In: Proceedings of the 7th International Conference on Energy Minimization Methods in Computer Vision and Pattern Recognition, *Lecture Notes in Computer Science*, vol. 5681, pp. 261–273. Bonn, Germany (2009) 10, 12, 23
15. Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite element approximation. Computers & Mathematics with Applications **2**(1), 17–40 (1976) 12
16. Gardner, I.C.: Validity of the cosine-fourth-power law of illumination. Journal of Research of the National Bureau of Standards **39**, 213–219 (1947) 6
17. Giachetti, A., Daffara, C., Reghelin C. Gobbetti, E., Pintus, R.: Light calibration and quality assessment methods for reflectance transformation imaging applied to artworks' analysis. In: Optics for Arts, Architecture, and Archaeology V, *Proceedings of SPIE*, vol. 9527. Munich, Germany (2015) 3, 5
18. Glowinski, R., Marroco, A.: Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité d'une classe de problèmes de Dirichlet non linéaires. ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique **9**(2), 41–76 (1975) 12
19. Gotardo, P.F.U., Simon, T., Sheikh, Y., Matthews, I.: Photogeometric Scene Flow for High-Detail Dynamic 3D Reconstruction. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 846–854. Santiago, Chile (2015) 11
20. Gratton, S., Lawless, S., Nichols, N.K.: Approximate Gauss-Newton methods for nonlinear least squares problems. SIAM Journal on Optimization **18**, 106–132 (2007) 17
21. Hara, K., Nishino, K., Ikeuchi, K.: Light source position and reflectance estimation from a single view without the distant illumination assumption. IEEE Transactions on Pattern Analysis and Machine Intelligence **27**(4), 493–505 (2005) 3, 5
22. Hernández, C., Vogiatzis, G., Brostow, G.J., Stenger, B., Cipolla, R.: Non-rigid Photometric Stereo with Colored Lights. In: Proceedings of the 11th IEEE International Conference on Computer Vision. Rio de Janeiro, Brazil (2007) 21
23. Hernández, C., Vogiatzis, G., Cipolla, R.: Multiview Photometric Stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence **30**(3), 548–554 (2008) 23
24. Hinkley, D.V.: On the Ratio of Two Correlated Normal Random Variables. Biometrika **56**(3), 635–639 (1969) 12
25. Hoeltgen, L., Quéau, Y., Breuss, M., Radow, G.: Optimised photometric stereo via non-convex variational minimisation. In: Proceedings of the 27th British Machine Vision Conference. York, UK (2016) 14
26. Horn, B.K.P.: Robot Vision. The MIT Press (1986) 5
27. Horn, B.K.P., Brooks, M.J. (eds.): Shape from Shading. The MIT Press (1989) 1
28. Huang, X., Walton, M., Bearman, G., Cossairt, O.: Near light correction for image relighting and 3D shape recovery. In: Proceedings of the International Congress on Digital Heritage, vol. 1, pp. 215–222. Granada, Spain (2015) 3, 5, 8
29. Ikeda, O., Duan, Y.: Color Photometric Stereo for Albedo and Shape Reconstruction. In: Proceedings of the IEEE Winter Conference on Applications of Computer Vision. Lake Placid, USA (2008) 22
30. Ikehata, S., Wipf, D., Matsushita, Y., Aizawa, K.: Photometric Stereo Using Sparse Bayesian Regression for General Diffuse Surfaces. IEEE Transactions on Pattern Analysis and Machine Intelligence **36**(9), 1816–1831 (2014) 15
31. Iwahori, Y., Sugie, H., Ishii, N.: Reconstructing shape from shading images under point light source illumination. In: Proceedings of the 19th International Conference on Pattern Recognition, vol. 1, pp. 83–87. Atlantic City, USA (1990) 3
32. Jiang, J., Liu, D., Gu, J., Süsstrunk, S.: What is the space of spectral sensitivity functions for digital color cameras? In: Proceedings of the IEEE Winter Conference on Applications of Computer Vision, pp. 168–179. Clearwater, USA (2013) 19
33. Kolagani, N., Fox, J.S., Blidberg, D.R.: Photometric stereo using point light sources. In: Proceedings of the 9th IEEE International Conference on Robotics and Automation, vol. 2, pp. 1759–1764. Nice, France (1992) 3, 8
34. Kontsevich, L.L., Petrov, A.P., Vergelskaya, I.S.: Reconstruction of shape from shading in color images. Journal of the Optical Society of America A **11**(3), 1047–1052 (1994) 21
35. Koppal, S.J., Narasimhan, S.G.: Novel depth cues from uncalibrated near-field lighting. In: Proceedings of the IEEE International Conference on Computer Vision (2007) 3, 5
36. Liao, J., Buchholz, B., Thiery, J.M., Bauszat, P., Eisemann, E.: Indoor scene reconstruction using near-light photometric stereo. IEEE Transactions on Image Processing **26**(3), 1089–1101 (2016) 3, 5
37. Logothetis, F., Mecca, R., Cipolla, R.: Semi-calibrated Near Field Photometric Stereo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA (2017) 3, 5, 8, 11, 12, 23

38. Logothetis, F., Mecca, R., Quéau, Y., Cipolla, R.: Near-Field Photometric Stereo in Ambient Light. In: Proceedings of the 27th British Machine Vision Conference. York, UK (2016) 3, 5, 8, 11, 12
39. McGunnigle, G., Chantler, M.J.: Resolving handwriting from background printing using photometric stereo. Pattern Recognition **36**(8), 1869–1879 (2003) 3, 4
40. Mecca, R., Quéau, Y., Logothetis, F., Cipolla, R.: A Single Lobe Photometric Stereo Approach for Heterogeneous Material. SIAM Journal on Imaging Sciences **9**(4), 1858–1888 (2016) 3, 5, 8, 11, 12, 15, 17
41. Mecca, R., Rodolà, E., Cremers, D.: Realistic photometric stereo using partial differential irradiance equation ratios. Computers & Graphics **51**, 8–16 (2015) 3, 5, 11
42. Mecca, R., Wetzler, A., Bruckstein, A.M., Kimmel, R.: Near Field Photometric Stereo with Point Light Sources. SIAM Journal on Imaging Sciences **7**(4), 2732–2770 (2014) 3, 5, 11
43. Migita, T., Ogino, S., Shakunaga, T.: Direct Bundle Estimation for Recovery of Shape, Reflectance Property and Light Position. In: Proceedings of the 10th European Conference on Computer Vision, *Lecture Notes in Computer Science*, vol. 5304, pp. 412–425. Marseille, France (2008) 3, 5, 23
44. Moreno, I., Avendaño Alejo, M., Tzonchev, R.I.: Designing light-emitting diode arrays for uniform near-field irradiance. Applied Optics **45**(10), 2265–2272 (2006) 2
45. Moreno, I., Sun, C.C.: Modeling the radiation pattern of LEDs. Optics Express **16**(3), 1808–1819 (2008) 3, 4, 5
46. Nie, Y., Song, Z.: A novel photometric stereo method with nonisotropic point light sources. In: Proceedings of the 23rd International Conference on Pattern Recognition, pp. 1737–1742. Cancun, Mexico (2016) 3, 5, 8
47. Nie, Y., Song, Z., Ji, M., Zhu, L.: A novel calibration method for the photometric stereo system with non-isotropic LED lamps. In: Proceedings of the IEEE Conference on Real-time Computing and Robotics, pp. 289–294. Angkor Wat, Cambodia (2016) 3, 5
48. Oren, M., Nayar, S.K.: Generalization of the Lambertian model and implications for machine vision. International Journal of Computer Vision **14**(3), 227–251 (1995) 10
49. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York (1970) 24
50. Papadhimitri, T., Favaro, P.: Uncalibrated Near-Light Photometric Stereo. In: Proceedings of the 25th British Machine Vision Conference. Nottingham, UK (2014) 3, 5, 8, 23
51. Pătrăucean, V., Gurdjos, P., Grompone von Gioi, R.: A parameterless line segment and elliptical arc detector with enhanced ellipse fitting. In: Proceedings of the 12th European Conference on Computer Vision, pp. 572–585. Florence, Italy (2012) 5
52. Pintus, R., Ciortan, I., Giachetti, A., Gobbetti, E.: Practical Free-form RTI Acquisition with Local Spot Lights. In: Smart Tools and Applications for Graphics. Genova, Italy (2016) 5
53. Powell, M.W., Sarkar, S., Goldgof, D.: A simple strategy for calibrating the geometry of light sources. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(9), 1022–1027 (2001) 3, 5
54. Quéau, Y., Durou, J.D., Aujol, J.F.: Normal Integration – Part I: A Survey (2016). URL https://hal.archives-ouvertes.fr/hal-01334349 2, 9
55. Quéau, Y., Mecca, R., Durou, J.D.: Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach. In: Proceedings of the 29th IEEE Conference on Computer Vision and Pattern Recognition, pp. 4350–4358. Las Vegas, USA (2016) 3, 8, 11, 12, 21, 22
56. Quéau, Y., Mecca, R., Durou, J.D., Descombes, X.: Photometric Stereo with Only Two Images: A Theoretical Study and Numerical Resolution. Image and Vision Computing **57**, 175–191 (2017) 21
57. Quéau, Y., Wu, T., Cremers, D.: Semi-Calibrated Near-Light Photometric Stereo. In: Proceedings of the 6th International Conference on Scale Space and Variational Methods in Computer Vision, *Lecture Notes in Computer Science*, vol. 10302, pp. 656–668. Kolding, Denmark (2017) 3, 5, 14, 17, 23
58. Quéau, Y., Wu, T., Lauze, F., Durou, J.D., Cremers, D.: A Non-Convex Variational Approach to Photometric Stereo under Inaccurate Lighting. In: Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA (2017) 3, 14, 15
59. Shen, H.L., Cheng, Y.: Calibrating light sources by using a planar mirror. Journal of Electronic Imaging **20**(1) (2011) 3, 5
60. Smith, W., Fang, F.: Height from Photometric Ratio with Model-based Light Source Selection. Computer Vision and Image Understanding **145**, 128–138 (2016) 11, 17
61. Sun, J., Smith, M., Smith, L., Farooq, A.: Sampling Light Field for Photometric Stereo. International Journal of Computer Theory and Engineering **5**(1), 14–18 (2013) 3, 4
62. Takai, T., Maki, A., Niinuma, K., Matsuyama, T.: Difference sphere: An approach to near light source estimation. Computer Vision and Image Understanding **113**(9), 966–978 (2009) 3, 5
63. Wolke, R., Schwetlick, H.: Iteratively reweighted least squares: algorithms, convergence analysis, and numerical comparisons. SIAM Journal on Scientific and Statistical Computing **9**(5), 907–921 (1988) 16
64. Woodham, R.J.: Photometric Method for Determining Surface Orientation from Multiple Images. Optical Engineering **19**(1), 139–144 (1980) 1, 3, 14
65. Wu, L., Ganesh, A., Shi, B., Matsushita, Y., Wang, Y., Ma, Y.: Robust photometric stereo via low-rank matrix completion and recovery. In: Proceedings of the Asian Conference on Computer Vision, *Lecture Notes in Computer Science*, vol. 6494, pp. 703–717. Queenstown, New-Zealand (2010) 15
66. Wu, Z., Li, L.: A line-integration based method for depth recovery from surface normals. Computer Vision, Graphics, and Image Processing **43**(1), 53–66 (1988) 9
67. Xie, L., Song, Z., Jiao, G., Huang, X., Jia, K.: A practical means for calibrating an LED-based photometric stereo system. Optics and Lasers in Engineering **64**, 42–50 (2015) 3, 5
68. Xie, W., Dai, C., Wang, C.C.L.: Photometric Stereo With Near Point Lighting: A Solution by Mesh Deformation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA (2015) 3
69. Yeh, C.K., Matsuda, N., Huang, X., Li, F., Walton, M., Cossairt, O.: A Streamlined Photometric Stereo Framework for Cultural Heritage. In: Proceedings of the 14th European Conference on Computer Vision, pp. 738–752. Amsterdam, The Netherlands (2016) 3, 5, 8