An Operational 3D-reconstruction Device using Photometric Stereo and LEDs – Modeling, Calibration and Resolution

Yvain Quéau, Bastien Durix, Jean-Denis Durou, François Lauze

▶ **To cite this version:**

Yvain Quéau, Bastien Durix, Jean-Denis Durou, François Lauze. An Operational 3D-reconstruction Device using Photometric Stereo and LEDs – Modeling, Calibration and Resolution. 2016. hal-01399607v1

HAL Id: hal-01399607

https://hal.science/hal-01399607v1

Preprint submitted on 19 Nov 2016 (v1), 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)

An Operational 3D-reconstruction Device using Photometric Stereo and LEDs

Modeling, Calibration and Resolution

Yvain Quéau · Bastien Durix · Jean-Denis Durou · François Lauze

the date of receipt and acceptance should be inserted later

Abstract We describe a photometric stereo-based 3D-reconstruction device. In the classical formulation of the problem, the luminous fluxes are assumed to be directional, which is very difficult to achieve in practice. Rather, we suggest the use of LEDs to illuminate the scene to reconstruct. In doing so, we question the conventional photometric stereo model and must expect its resolution to be more complex. We begin by describing a calibration procedure for the luminous flux emitted by a LED. We then show how to adapt to the new photometric stereo model the classical two-stages resolution method, which involves estimating the albedo and the normals, and then integrating the normals. Then, we reformulate the problem as a system of PDEs, in order to reduce nonlinearities and to eliminate the albedo. We first suggest a fixed point resolution scheme, and then an iterative, ADMM-based scheme. Finally, we extend this study to the case of RGB images.

Keywords 3D-reconstruction · Photometric stereo · Point light sources · Variational methods · Numerical schemes · ADMM.

Yvain Quéau
Technical University Munich, Garching, Germany
E-mail: yvain.queau@tum.de

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

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

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

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 [22] and photometric stereo [41] 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 use throughout this paper. We describe a point $\mathbf{x}$ in 3D-space 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. 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. If f refers to the focal length, the relationship of conjugation between $\mathbf{x}$ and $\mathbf{p}$ is written, in perspective projection:

$$\begin{cases} x = \dfrac{z}{f}\, u \\ y = \dfrac{z}{f}\, v \end{cases} \tag{1}$$

3D-reconstruction consists in estimating, in each pixel $\mathbf{p}$ of a part Ω of the image, its conjugate point $\mathbf{x}$ in 3D-space. Relations (1) show that it suffices to find the depth z to determine $\mathbf{x}$ from $\mathbf{p}$. The only unknown of the problem is thus the depth map z, which is defined as follows:

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

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 $I^1, \ldots, I^m$ taken from the same angle, but under varying lighting. It allows to estimate not only the shape, but also the albedo, which is not the case for other computer vision techniques.

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 proportional to the light flux density. For a Lambertian surface, the classical modeling of photometric stereo is written, in each pixel $\mathbf{p} \in \Omega$, as the following system of equations[1]:

$$I^i(\mathbf{p}) = \rho(\mathbf{x})\ \mathbf{s}^i \cdot \mathbf{n}(\mathbf{x}), \qquad i \in [1, m] \tag{3}$$

where $I^i(\mathbf{p})$ denotes the gray level of $\mathbf{p}$ under lighting $\mathbf{s}^i$, $\rho(\mathbf{x})$ denotes the albedo in the point $\mathbf{x}$ conjugate of $\mathbf{p}$ and $\mathbf{n}(\mathbf{x})$ denotes the unit-length, outgoing normal to the surface in this point. Since there is a bijective correspondence between a point $\mathbf{x}$ and a pixel $\mathbf{p}$, we write for convenience $\rho(\mathbf{p})$ and $\mathbf{n}(\mathbf{p})$ in lieu of $\rho(\mathbf{x})$ and

[1] Equality (3) is in fact a proportionality relationship: see the expression (17) of $I(\mathbf{p})$.

$\mathbf{n}(\mathbf{x})$. Introducing the notation $\mathbf{m}(\mathbf{p}) = \rho(\mathbf{p})\,\mathbf{n}(\mathbf{p})$, System (3) can be rewritten in matrix form:

$$\mathbf{I}(\mathbf{p}) = \mathbf{S}\,\mathbf{m}(\mathbf{p}) \tag{4}$$

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

$$\mathbf{I}(\mathbf{p}) = \begin{bmatrix} I^1(\mathbf{p}) \\ \vdots \\ I^m(\mathbf{p}) \end{bmatrix} \qquad \text{and} \qquad \mathbf{S} = \begin{bmatrix} \mathbf{s}^{1\top} \\ \vdots \\ \mathbf{s}^{m\top} \end{bmatrix} \tag{5}$$

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 (4) is then given by:

$$\mathbf{m}(\mathbf{p}) = \mathbf{S}^+\,\mathbf{I}(\mathbf{p}) \tag{6}$$

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

$$\rho(\mathbf{p}) = \|\mathbf{m}(\mathbf{p})\| \qquad \text{and} \qquad \mathbf{n}(\mathbf{p}) = \frac{\mathbf{m}(\mathbf{p})}{\|\mathbf{m}(\mathbf{p})\|} \tag{7}$$

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 [12].

To ensure the directionality of lighting, which is required by Model (3), it is necessary to achieve a complex optical device [29]. 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 [30], and therefore in the resolution. Indeed, vectors $\mathbf{s}^i$ in Model (3) then depend on the 3D-coordinates of the point $\mathbf{x}$ of the surface, which are precisely the unknowns of the problem. The auxiliary variable $\mathbf{m}(\mathbf{p}) = \rho(\mathbf{p})\,\mathbf{n}(\mathbf{p})$, whose introduction allows to linearize the equations when the lighting is directional, since $\mathbf{m}(\mathbf{p})$ is not constrained to have unit-length, loses part of its interest in that situation.

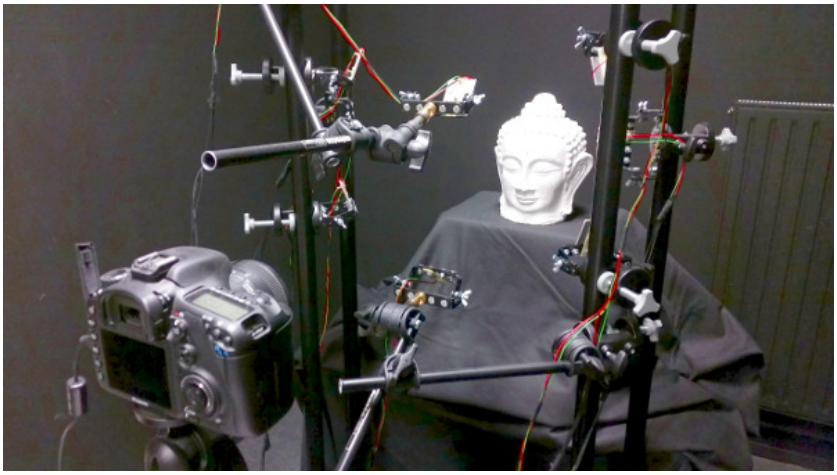
In several articles dealing with photometric stereo under non-directional lighting [7,23,33], it is proposed to use the auxiliary variable $\mathbf{m}(\mathbf{p})$, although the problem to solve is no longer linear. The aim of our work is to conduct a comprehensive and detailed study of this problem, from the modeling to the resolution, through calibration. In particular, we will show the interest of a new differential modeling allowing to no longer use the auxiliary variable $\mathbf{m}(\mathbf{p})$. Our main contributions are as follows:

- Modeling photometric stereo under point light source illumination.
- New (geometric and photometric) calibration procedure of a LED.
- Improved classical resolution method for estimating the absolute depth without a priori knowledge on the shape of the scene.
- Writing of two new numerical schemes aiming at solving the differential formulation of photometric stereo.

In Section 2, we show how to modify Model (4) when the scene is illuminated by LEDs. In Section 3, which is the core of our work, we describe three resolution schemes for this problem and propose an optimal combination of two of them. In Section 4, we extend this study to the case of RGB images. Eventually, we conclude and suggest several future research directions in Section 5.

2 Photometric Stereo under Point Light Source Illumination

Most works on photometric stereo assume that the primary luminous flux is parallel and uniform, but the directional nature of lighting is difficult to guarantee. It is much easier to illuminate a scene with LEDs. With this in mind, we have developed a photometric stereo-based device for 3D-reconstruction of human faces, which includes $m = 8$ LEDs located at about 30 cm of the scene (see Figure 1-a). The face is photographed by a Canon EOS 7D camera with focal length $f = 35$ mm. Triggering the shutter in burst mode, while sinchronizing the lighting of the LEDs, provides us with $m = 8$ images such as those of Figures 1-b, 1-c and 1-d. In this section, we model the luminous flux emitted by a LED, then we show how to estimate the parameters of this model. We then show how to model photometric stereo under m point light sources.



(a)

(b)

(c)

(d)

Fig. 1 (a) Our photometric stereo-based experimental device for 3D-reconstruction of human faces using $m = 8$ LEDs. Walls were painted in black in order to avoid reflections between the scene and the environment. (b-c-d) Three out of the $m = 8$ images obtained by this device.

2.1 Modeling the Luminous Flux Emitted by a LED

For the LEDs we use, the characteristic dimension of the illuminating volume is of the order of a millimeter. Therefore, in comparison with the scale of a human 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 luminous flux emitted by such a source is necessarily radial, that is to say collinear with the unit 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, this flow is thus written:

$$\mathbf{s}(\mathbf{x}) = -\frac{\Phi(\theta, \phi)}{r^2}\,\mathbf{u}_r \tag{8}$$

where $\Phi(\theta, \phi) \geqslant 0$ denotes the intensity of the source, and where 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 (8), which is very general, could be described by a small number of parameters, for instance by projecting $\Phi(\theta, \phi)$ on the *spherical harmonics* basis. This allowed Basri et al. to model the luminous flux in the case of uncalibrated photometric stereo [5]. The intensity $\Phi(\theta, \phi)$ can also be sampled in the vicinity of a plane using a plane with known reflectance: the classical photometric stereo approach with directional lighting can then be used, provided that the images of the surface to reconstruct are compensated by this sampling [2,26,39]. Of course, this is only justified if the surface is close enough to a plane, which greatly limits potential applications. In addition, 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 main direction indicated by a unit vector $\mathbf{n}_s$. If θ is defined relatively to $\mathbf{n}_s$, this means that $\Phi(\theta, \phi)$ is independent from ϕ. The luminous flux in $\mathbf{x}$ coming from a LED located in $\mathbf{x}_s$ is thus written:

$$\mathbf{s}(\mathbf{x}) = \frac{\Phi(\theta)}{\|\mathbf{x}-\mathbf{x}_s\|^2}\,\frac{\mathbf{x}_s-\mathbf{x}}{\|\mathbf{x}_s-\mathbf{x}\|} \tag{9}$$

The dependency on θ of the intensity Φ characterizes the anisotropy of the LED. The function $\Phi(\theta)$ is generally decreasing on $[0, \pi/2]$. An anisotropy model satisfying this constraint is that of "imperfect Lambertian sources" [6]:

$$\Phi(\theta) = \Phi_0\,\cos^{\mu}\theta \tag{10}$$

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 (10) is empirical, yet it has already been used in photometric stereo [27,43], including when all the LEDs are arranged on a plane parallel to the image plane i.e., when $\mathbf{n}_s = [0, 0, 1]^\top$ [28]. This model has proven itself and, moreover, LED manufacturers provide the angle $\theta_{1/2}$ such that $\Phi(\theta_{1/2}) = \Phi_0/2$, from which we deduce the value of μ by using (10):

$$\mu = -\frac{\log(2)}{\log(\cos\theta_{1/2})} \tag{11}$$

For the LEDs we use, the angle $\theta_{1/2}$ is $\pi/3$. From (11), we deduce that $\mu = 1$ i.e., that our LEDs are Lambertian. Plugging the expression (10) of $\Phi(\theta)$ into (9), one obtains:

$$\mathbf{s}(\mathbf{x}) = \Phi_0\,\cos^{\mu}\theta\,\frac{\mathbf{x}_s-\mathbf{x}}{\|\mathbf{x}_s-\mathbf{x}\|^3} \tag{12}$$

where we explicitly keep the parameter μ to address the most general case. Model (12) thus includes seven parameters: three for the coordinates of $\mathbf{x}_s$, two for the unit vector $\mathbf{n}_s$, plus the parameters Φ_0 and μ. Note that $\mathbf{n}_s$ appears in this model through the angle θ.

In its uncalibrated version, photometric stereo allows 3D-reconstruction of a scene without knowing the lighting. This case has been widely studied, including in the case of nearby point light sources [23,33,45], but if we can, we should rather calibrate the lighting[2].

2.2 Calibrating the Luminous Flux Emitted by a LED

Most calibration methods for point light sources [1,3,9,14,17,35,37,40] do not take into account the attenuation of the light flux 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. Xie et al. use in [43] a single partially bright and partially Lambertian surface to calibrate a LED. We intend to improve this procedure using two patterns, one bright and one Lambertian. The brilliant one will be used to determine the position of the LEDs by triangulation, and the Lambertian one to determine other parameters by minimizing the reprojection error, as recently proposed by Pintus et al. in [34].

Specular Spherical Calibration Pattern – Position $\mathbf{x}_s$ of a LED can be determined by triangulation. In [35], Powell et al. advocate the use of a spherical mirror. To estimate the positions of the $m = 8$ LEDs for our device, 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. It is then easy to determine the 3D-position and normal for any point on the surface, since the radius of the billiard ball is known. For each position 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 $\mathbf{x}_s$, but the use of ten poses improves the robustness of the estimation.

Lambertian Model – To estimate the main direction $\mathbf{n}_s$ and the intensity Φ_0 in Model (12), we use a Lambertian calibration pattern. A surface is Lambertian if the apparent clarity of any point $\mathbf{x}$ is independent from the angle under which it is observed. 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 *irradiance* $E(\mathbf{x})$, that is to say on the luminous flux received in $\mathbf{x}$ per unit of surface. The relationship between luminance and irradiance is written, for a Lambertian surface:

$$L(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\pi} E(\mathbf{x}) \tag{13}$$

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

where the albedo $\rho(\mathbf{x}) \in \mathbb{R}$, which locally characterizes the surface, is equal to the proportion of reemitted luminous energy: $\rho(\mathbf{x})$ is equal to 1 if $\mathbf{x}$ is white, and to 0 if it is black.

The parameter $\rho(\mathbf{x})$ characterizes itself the *reflectance* (or BRDF) of a Lambertian surface. In addition, the irradiance at a point $\mathbf{x}$ of any surface (Lambertian or not) with normal $\mathbf{n}(\mathbf{x})$, lit by a luminous flux $\mathbf{s}(\mathbf{x})$, is written:

$$E(\mathbf{x}) = \mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \tag{14}$$

After focusing the camera on a point $\mathbf{x}$ of the 3D-scene, the irradiance $\epsilon(\mathbf{p})$ 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 [22]:

$$\epsilon(\mathbf{p}) = \beta \, \cos^4 \alpha(\mathbf{p}) \, L(\mathbf{x}) \tag{15}$$

where β is a proportionality coefficient characterizing the clarity of the image, which depends on several factors such as the lens aperture, the exposure time, the magnification, etc. Regarding 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 that should not be confused with vignetting.

With current photosensitive receptors, the *gray level* $J(\mathbf{p})$ at pixel $\mathbf{p}$ is almost proportional[3] to its irradiance $\epsilon(\mathbf{p})$, except of course in case of saturation. Denoting γ this coefficient of quasi-proportionality, and combining equalities (13), (14) and (15), 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:

$$J(\mathbf{p}) = \gamma \, \beta \, \cos^4 \alpha(\mathbf{p}) \, \frac{\rho(\mathbf{x})}{\pi} \, \mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) \tag{16}$$

We have already mentioned that there was a bijective correspondence between a point $\mathbf{x}$ and its conjugate $\mathbf{p}$, 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, we can very easily compensate for this source of darkening and will manipulate now the "corrected gray level":

$$I(\mathbf{p}) = \frac{J(\mathbf{p})}{\cos^4 \alpha(\mathbf{p})} = \gamma \, \beta \, \frac{\rho(\mathbf{p})}{\pi} \, \mathbf{s}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{p}) \tag{17}$$

Lambertian Planar Calibration Pattern – To estimate the parameters $\mathbf{n}_s$ and Φ_0 in Model (13), that is to say 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. Figure 2), 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 luminous flux $\mathbf{s}(\mathbf{x})$ at any point $\mathbf{x}$ of the surface is purely *primary*, that is to say, it comes only from the light source, without "bouncing" on other parts of the surface of the target, whereupon the walls and surrounding objects are covered in black (see Figure 1-a). With this observation, we can replace the

[3] It is said in [7] that using RAW images is a necessary condition to achieve this proportionality relationship.

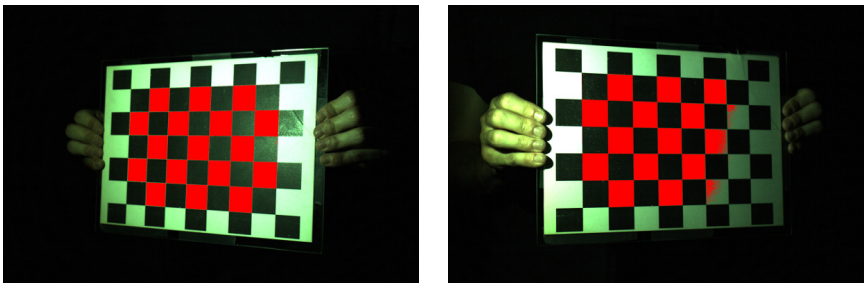


Fig. 2 Two out of the q poses of the planar Lambertian calibration pattern used for photometric calibration of the LEDs. The parts of the white cells that are used for estimation are indicated in red.

light flux $\mathbf{s}(\mathbf{x})$ of Equation (16) by the expression (12) of the flux emitted by a LED. From (12) and (16), 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:

$$I(\mathbf{p}) = \gamma\,\beta\,\frac{\rho(\mathbf{p})}{\pi}\,\Phi_0\,\cos^{\mu}\theta\,\frac{(\mathbf{x}_s-\mathbf{x})\cdot\mathbf{n}(\mathbf{p})}{\|\mathbf{x}_s-\mathbf{x}\|^3} \tag{18}$$

If $q \geqslant 3$ poses of the checkerboard are used, numerous algorithms exist for estimating unambiguously 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 on $\mathbf{p}$ of the normals, since the pattern is flat), and the intrinsic parameters of the camera[4]. 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 $\rho(\mathbf{p}) \equiv \rho_0$, that is to say, the albedo is uniform in the white cells. We can group all the multiplier coefficients of the right hand side of (18) into one coefficient:

$$\Psi = \gamma\,\beta\,\frac{\rho_0}{\pi}\,\Phi_0 \tag{19}$$

With this definition, and knowing that θ is the angle between vectors $\mathbf{n}_s$ and $\mathbf{x}-\mathbf{x}_s$, Equation (18) 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 indicated in red in the images of Figure 2):

$$I^j(\mathbf{p}) = \Psi\left[\mathbf{n}_s\cdot\frac{\mathbf{x}^j-\mathbf{x}_s}{\|\mathbf{x}^j-\mathbf{x}_s\|}\right]^{\mu}\frac{(\mathbf{x}_s-\mathbf{x}^j)\cdot\mathbf{n}^j}{\|\mathbf{x}_s-\mathbf{x}^j\|^3} \tag{20}$$

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

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

$$\min_{\Psi}\sum_{j=1}^{q}\sum_{\mathbf{p}\in\Omega^j}\left[I^j(\mathbf{p}) - \Psi\,\frac{(\mathbf{x}_s-\mathbf{x}^j)\cdot\mathbf{n}^j}{\|\mathbf{x}_s-\mathbf{x}^j\|^3}\right]^2 \tag{21}$$

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

whose solution is given by:

$$\Psi = \frac{\displaystyle\sum_{j=1}^{q} \sum_{\mathbf{p}\in\Omega^j} I^j(\mathbf{p}) \, \frac{(\mathbf{x}_s - \mathbf{x}^j)\cdot \mathbf{n}^j}{\|\mathbf{x}_s - \mathbf{x}^j\|^3}}{\displaystyle\sum_{j=1}^{q} \sum_{\mathbf{p}\in\Omega^j} \left[\frac{(\mathbf{x}_s - \mathbf{x}^j)\cdot \mathbf{n}^j}{\|\mathbf{x}_s - \mathbf{x}^j\|^3} \right]^2} \tag{22}$$

- In the other case (if $\mu > 0$), Equation (20) can be rewritten:

$$\underbrace{\Psi^{1/\mu}\,\mathbf{n}_s}_{\mathbf{m}_s} \cdot (\mathbf{x}^j - \mathbf{x}_s) = \left[I^j(\mathbf{p}) \, \frac{\|\mathbf{x}_s - \mathbf{x}^j\|^{3+\mu}}{(\mathbf{x}_s - \mathbf{x}^j)\cdot \mathbf{n}^j} \right]^{1/\mu} \tag{23}$$

The least-squares estimation of vector $\mathbf{m}_s$ is thus written:

$$\min_{\mathbf{m}_s} \sum_{j=1}^{q} \sum_{\mathbf{p}\in\Omega^j} \left[\mathbf{m}_s \cdot (\mathbf{x}^j - \mathbf{x}_s) - \left[I^j(\mathbf{p}) \, \frac{\|\mathbf{x}_s - \mathbf{x}^j\|^{3+\mu}}{(\mathbf{x}_s - \mathbf{x}^j)\cdot \mathbf{n}^j} \right]^{1/\mu} \right]^2 \tag{24}$$

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

$$\mathbf{n}_s = \frac{\mathbf{m}_s}{\|\mathbf{m}_s\|} \qquad \text{et} \qquad \Psi = \|\mathbf{m}_s\|^{\mu} \tag{25}$$

In both cases, it is impossible to deduce from the estimate of Ψ that of Φ_0, because in the definition (19) of Ψ, the product $\gamma\,\beta\,\frac{\rho_0}{\pi}$ is unknown. However, since this product is the same for all LEDs (desactivating all automatic settings of the camera makes β and γ constant), this is not a problem for photometric stereo.

Figure 3 shows two views of a schematic representation of the experimental device of Figure 1, where LEDs parameters were estimated by following the calibration procedure as described above.

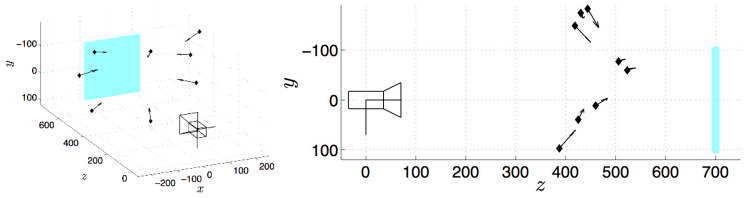


Fig. 3 Two views of a schematic representation of the experimental device of Figure 1-a, in which the arrows symbolize the LEDs (unit: millimeter). The direction of an arrow characterizes the main direction $\mathbf{n}_s$ of a LED, while its length characterizes its intensity Ψ (up to a constant factor). The plane in blue indicates the average position of the face, located about 70 *cm* away from the camera, or 30 *cm* from the LEDs.

2.3 Modeling Photometric Stereo with Point Light Sources

If the luminous flux emitted by a LED is described by (12), then we obtain from (18) and (19) the following equation for the gray level at pixel $\mathbf{p}$:

$$I(\mathbf{p}) = \Psi \frac{\rho(\mathbf{p})}{\rho_0} \left[\mathbf{n}_s \cdot \frac{\mathbf{x} - \mathbf{x}_s}{\|\mathbf{x} - \mathbf{x}_s\|} \right]^{\mu} \frac{(\mathbf{x}_s - \mathbf{x}) \cdot \mathbf{n}(\mathbf{p})}{\|\mathbf{x}_s - \mathbf{x}\|^3} \tag{26}$$

Thus, we must solve, in each pixel $\mathbf{p} \in \Omega$, the following system of equations:

$$I^i(\mathbf{p}) = \Psi^i \frac{\rho(\mathbf{p})}{\rho_0} \left[\mathbf{n}_s^i \cdot \frac{\mathbf{x} - \mathbf{x}_s^i}{\|\mathbf{x} - \mathbf{x}_s^i\|} \right]^{\mu^i} \frac{(\mathbf{x}_s^i - \mathbf{x}) \cdot \mathbf{n}(\mathbf{p})}{\|\mathbf{x}_s^i - \mathbf{x}\|^3}, \qquad i \in [1, m] \tag{27}$$

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 also introduce a new definition of the albedo, relatively to the albedo ρ_0 of the Lambertian calibration pattern:

$$\overline{\rho}(\mathbf{p}) = \frac{\rho(\mathbf{p})}{\rho_0} \tag{28}$$

Since parameters μ^i, $i \in [1, m]$, are provided by the manufacturer, the only unknowns in System (27) are the position $\mathbf{x}$ of the 3D-point conjugate to $\mathbf{p}$, its 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, $i \in [1, m]$:

$$\mathbf{t}^i(\mathbf{x}) = \Psi^i \left[\mathbf{n}_s^i \cdot \frac{\mathbf{x} - \mathbf{x}_s^i}{\|\mathbf{x} - \mathbf{x}_s^i\|} \right]^{\mu^i} \frac{\mathbf{x}_s^i - \mathbf{x}}{\|\mathbf{x}_s^i - \mathbf{x}\|^3} \tag{29}$$

System (27) is rewritten:

$$I^i(\mathbf{p}) = \mathbf{t}^i(\mathbf{x}) \cdot \overline{\mathbf{m}}(\mathbf{p}), \qquad i \in [1, m] \tag{30}$$

or, in matrix form:

$$\mathbf{I}(\mathbf{p}) = \mathbf{T}(\mathbf{x})\,\overline{\mathbf{m}}(\mathbf{p}) \tag{31}$$

where $\mathbf{I}(\mathbf{p}) \in \mathbb{R}^m$ has been defined in (5) and $\mathbf{T}(\mathbf{x}) \in \mathbb{R}^{m \times 3}$ is defined as follows:

$$\mathbf{T}(\mathbf{x}) = \begin{bmatrix} \mathbf{t}^1(\mathbf{x})^\top \\ \vdots \\ \mathbf{t}^m(\mathbf{x})^\top \end{bmatrix} \tag{32}$$

Equation (31) is similar to (4), and knowledge of the matrix field $\mathbf{T}(\mathbf{x})$ would allow us to estimate its field of pseudo-inverses in order to solve (31), just as calculating the pseudo-inverse of $\mathbf{S}$ allows us to solve (4) when the lighting is directional. However, matrix field $\mathbf{T}(\mathbf{x})$ depends on $\mathbf{x}$, and thus on the *unknown depth*. We will see that this difference induces major changes when it comes to the resolution.

The dependency on $\mathbf{x}$ of the matrix field $\mathbf{T}(\mathbf{x})$, which significantly complicates the resolution of Problem (31), 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 [7], we systematically eliminate the highest gray level, which may come from a specular highlight, as well as both lowest gray levels, which may correspond to shadows.

3 Solving Photometric Stereo under Point Light Source Illumination

This section is our main contribution. We propose three resolution schemes for photometric stereo under point light source illumination.

3.1 Scheme Inspired by the Classical Resolution of the Linear Model (4)

It seems quite natural to take inspiration from the resolution (6) of the linear model (4). To linearize Problem (31), it is enough 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 methods of resolution [7, 10, 23, 33, 45], 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 $k = 0$ and if $\mathbf{x}^{(0)}$ the initial shape, this new scheme consits in iterating:

1. Solve Problem (31) in the least-squares sense at each $\mathbf{p} \in \Omega$, by replacing $\mathbf{T}(\mathbf{x})$ by $\mathbf{T}(\mathbf{x}^{(k)})$, which provides a new estimation of $\overline{\mathbf{m}}(\mathbf{p})$:

$$\overline{\mathbf{m}}^{(k+1)}(\mathbf{p}) = \underset{\mathbf{q} \in \mathbb{R}^3}{\operatorname{argmin}} \ \|\mathbf{I}(\mathbf{p}) - \mathbf{T}(\mathbf{x}^{(k)})\,\mathbf{q}\|^2 \tag{33}$$

2. Deduce a new estimation of the normal $\mathbf{n}(\mathbf{p})$ at each $\mathbf{p} \in \Omega$:

$$\mathbf{n}^{(k+1)}(\mathbf{p}) = \frac{\overline{\mathbf{m}}^{(k+1)}(\mathbf{p})}{\|\overline{\mathbf{m}}^{(k+1)}(\mathbf{p})\|} \tag{34}$$

3. Integrate the new normal field $\mathbf{n}^{(k+1)}(\mathbf{p})$, in order to obtain an updated shape $\mathbf{x}^{(k+1)}$, up to a scale factor.
4. Estimate this scale factor by nonlinear optimization.
5. Increment k.

For this scheme to be complete, we need to specify an initial shape $\mathbf{x}^{(0)}$ and a stopping criterion. Studying convergence of this scheme towards a solution of the system of PDEs (27), which is nonlinear, is beyond the scope of this article. In this view, let us only state for now that we could observe divergence in some specific situations where the initial shape $\mathbf{x}^{(0)}$ is very far away from the real shape, or when the surface presents too many specularities.

Integration of Normals – Although Stages 1 and 2 of the scheme above are trivial and can be achieved pixelwise, that is to say in each pixel $\mathbf{p} \in \Omega$ independently from its neighbors, this is not the case of Stages 3 and 4. From the equalities in (1), it is easy to deduce that the (non-unit-length) vector:

$$\overline{\mathbf{n}}(\mathbf{p}) = \begin{bmatrix} f\,\partial_u z(\mathbf{p}) \\ f\,\partial_v z(\mathbf{p}) \\ -z(\mathbf{p}) - \mathbf{p} \cdot \nabla z(\mathbf{p}) \end{bmatrix} \tag{35}$$

is normal to the surface ($\nabla z(\mathbf{p}) = [\partial_u z(\mathbf{p}), \partial_v z(\mathbf{p})]^\top$ refers to the gradient of z at $\mathbf{p}$). Expression (35) shows that integrating the (unit-length) normal field $\mathbf{n}(\mathbf{p})$ allows to estimate the depth z only up to a scale factor $\alpha \in \mathbb{R}$, since:

$$\mathbf{n}(\mathbf{p}) \propto \begin{bmatrix} f\,\partial_u z(\mathbf{p}) \\ f\,\partial_v z(\mathbf{p}) \\ -z(\mathbf{p}) - \mathbf{p}\cdot\nabla z(\mathbf{p}) \end{bmatrix} \propto \begin{bmatrix} f\,\partial_u (\alpha\, z)(\mathbf{p}) \\ f\,\partial_v (\alpha\, z)(\mathbf{p}) \\ -(\alpha\, z)(\mathbf{p}) - \mathbf{p}\cdot\nabla(\alpha\, z)(\mathbf{p}) \end{bmatrix} \tag{36}$$

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:

$$\begin{cases} n_3(\mathbf{p})\, f\, \partial_u z(\mathbf{p}) + n_1(\mathbf{p})\ [z(\mathbf{p}) + \mathbf{p}\cdot\nabla z(\mathbf{p})] = 0 \\ n_3(\mathbf{p})\, f\, \partial_v z(\mathbf{p}) + n_2(\mathbf{p})\ [z(\mathbf{p}) + \mathbf{p}\cdot\nabla z(\mathbf{p})] = 0 \end{cases} \tag{37}$$

which is homogeneous in $z(\mathbf{p})$. Introducing the variable $\tilde{z} = \log(z)$, and denoting:

$$\overline{\mathbf{p}} = [u, v, f]^\top \tag{38}$$

it is easy to deduce from (37) the gradient of $\tilde{z}(\mathbf{p})$:

$$\nabla\tilde{z}(\mathbf{p}) = -\frac{1}{\overline{\mathbf{p}}\cdot\mathbf{n}(\mathbf{p})} \begin{bmatrix} n_1(\mathbf{p}) \\ n_2(\mathbf{p}) \end{bmatrix} \tag{39}$$

Let us now come back to Stages 3 and 4 of the proposed resolution scheme. The current normal field is $\mathbf{n}^{(k+1)}(\mathbf{p})$, from which we deduce the gradient $\nabla\tilde{z}^{(k+1)}(\mathbf{p})$ thanks to (39). By integrating this gradient between a point $\mathbf{p}_0$, chosen arbitrarily inside Ω, and point $\mathbf{p}$, and knowing that $z = \exp\{\tilde{z}\}$, we obtain:

$$z^{(k+1)}(\mathbf{p}) = z^{(k+1)}(\mathbf{p}_0)\, \exp\left\{ \int_{\mathbf{p}_0}^{\mathbf{p}} \nabla\tilde{z}^{(k+1)}(\mathbf{q}) \cdot \mathrm{d}\mathbf{q} \right\} \tag{40}$$

This integral can be calculated along one single path going from $\mathbf{p}_0$ to $\mathbf{p}$, but since the gradient field $\nabla\tilde{z}^{(k+1)}(\mathbf{p})$ is never rigorously integrable, this calculus usually depends on the choice of the path [42]. The most common parade to this well-known problem consists in using a variational approach [12].

Expression (40) confirms that the depth can be calculated from $\mathbf{n}^{(k+1)}(\mathbf{p})$ only up to a scale factor $z^{(k+1)}(\mathbf{p}_0)$. Let us determine this scale factor by minimization of the reprojection error of Model (31) over the entire domain Ω. Knowing that, from (1) and (38), we get $\mathbf{x} = \frac{z}{f}\,\overline{\mathbf{p}}$, this comes down to solving the following nonlinear least squares problem:

$$z^{(k+1)}(\mathbf{p}_0) = \underset{w\,\in\,\mathbb{R}^+}{\operatorname{argmin}} \underbrace{\sum_{\mathbf{p}\in\Omega} \left\| \mathbf{I}(\mathbf{p}) - \mathbf{T}\left(\frac{w}{f} \exp\left\{ \int_{\mathbf{p}_0}^{\mathbf{p}} \nabla\tilde{z}^{(k+1)}(\mathbf{q}) \cdot \mathrm{d}\mathbf{q} \right\} \overline{\mathbf{p}} \right) \overline{\mathbf{m}}^{(k+1)}(\mathbf{p}) \right\|^2}_{\mathcal{E}_1} \tag{41}$$

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

$$\mathbf{x}^{(k+1)} = \frac{z^{(k+1)}(\mathbf{p}_0)}{f}\, \exp\left\{ \int_{\mathbf{p}_0}^{\mathbf{p}} \nabla\tilde{z}^{(k+1)}(\mathbf{q}) \cdot \mathrm{d}\mathbf{q} \right\} \overline{\mathbf{p}} \tag{42}$$

Experimental Validation – Although studying the convergence of this scheme is out of the scope of this article, we empirically found that it always converges, provided that the surface is Lambertian and the initial shape is not too far away from the mean distance from the camera to the scene. For the curves in Figure 4, several fronto-parallel planes with equation $z \equiv z_0$ were used as initial guess. The mean distance to the scene being approximately 700 mm, it is not surprising that the fastest convergence is observed for this value of z_0.

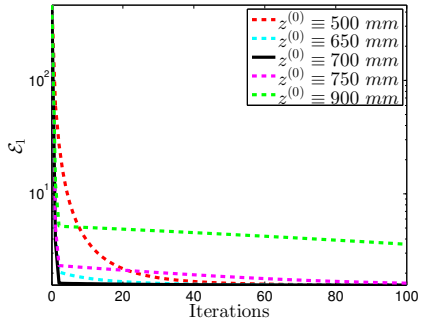


Fig. 4 Evolution of the energy $\mathcal{E}_1$ defined in (41), in function of the iterations, when the initial shape is a fronto-parallel plane with equation $z = z_0$. The used data are the $m = 8$ images of the plaster statuette of Figure 1. The proposed scheme always converges towards the same solution (at different speeds), for the five tested values of z_0.

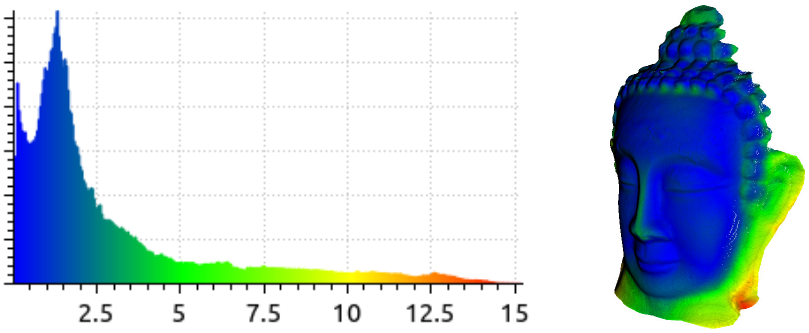
Figure 5 allows to compare the shape obtained by photometric stereo, from sub-images of size 920×1178 in full resolution (corresponding to the bounding box of the object), which contains 773794 points, with the ground truth obtained by laser scanning, which contains 1753010 points. The points density is thus almost the same on the front of the statuette, since we do 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.

Figure 6-a shows the histogram of point-to-point distances between our result (Figure 5-a) and the ground truth (Figure 5-c). The median value is worth 1.3 mm. The spatial distribution of these distances, which is illustrated in Figure 6-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 [32], would perhaps improve the accuracy of the 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.

(a) (b) (c)

Fig. 5 (a-b) Shape and albedo of the statuette obtained with the proposed resolution scheme. (c) Ground truth shape obtained by laser scanning (which does not provide the albedo).



(a) (b)

Fig. 6 (a) Histogram of point-to-point distances between our 3D-reconstruction and the ground truth. The median value is worth 1.3 *mm*. (b) Spatial distribution of these distances.

3.2 Fixed Point Resolution Scheme

Apart from the fact that we did not study its convergence, 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 must not be accomplished by the naive formulation (42), but rather by using more sophisticated methods which would allow to overcome the problem of non-integrability [11]. Also, photometric stereo indeed allows to estimate the albedo (cf. Figure 5-b), but this first scheme implicitely calculates the albedo at each iteration, since $\overline{\rho}^{(k+1)}(\mathbf{p}) = \|\overline{\mathbf{m}}^{(k+1)}(\mathbf{p})\|$. It is possible to avoid such useless computations by *a posteriori* estimating the albedo, once the shape is estimated.

This observation leads us to question the usefulness of the variable $\overline{\mathbf{m}}(\mathbf{p})$, which mixes the normal $\mathbf{n}(\mathbf{p})$ and the relative albedo $\overline{\rho}(\mathbf{p})$, that is to stay a geometric concept and a photometric one. We will show how to avoid this by a reformulation of photometric stereo which allows: to directly estimate the depth, and thus

to avoid integration; to eliminate the albedo from the unknowns, which will be a posteriori estimated, once the shape is estimated, thus avoiding useless computations. Let us note that the method proposed by Xie et al. [44] is inspired by the same ideas, yet it follows a different track: the depth is iteratively estimated by deforming a mesh representing the surface.

Differential Reformulation of Problem (31) – Let us recall that the 3D-position of the point $\mathbf{x}$ conjugate to a pixel $\mathbf{p}$ is completely characterized by the depth $z(\mathbf{p})$, since:

$$\mathbf{x} = \frac{z(\mathbf{p})}{f} \begin{bmatrix} \mathbf{p} \\ f \end{bmatrix} \tag{43}$$

The vectors $\mathbf{t}^i(\mathbf{x})$ defined in (29) can thus be denoted $\mathbf{t}^i_z(\mathbf{p})$. Knowing that the (non-unit-length) vector $\overline{\mathbf{n}}(\mathbf{p})$ defined in (35) is normal to the surface, we can rewrite System (27) in $\mathbf{p} \in \Omega$ in the following manner:

$$I^i(\mathbf{p}) = \frac{\overline{\rho}(\mathbf{p})}{\|\overline{\mathbf{n}}(\mathbf{p})\|} \mathbf{t}^i_z(\mathbf{p}) \cdot \begin{bmatrix} f\,\nabla z(\mathbf{p}) \\ -z(\mathbf{p}) - \mathbf{p} \cdot \nabla z(\mathbf{p}) \end{bmatrix}, \qquad i \in [1, m] \tag{44}$$

where the norm of $\overline{\mathbf{n}}(\mathbf{p})$ depends in a nonlinear way on $z(p)$ and on its gradient $\nabla z(p)$:

$$\|\overline{\mathbf{n}}(\mathbf{p})\| = \sqrt{f^2 \|\nabla z(\mathbf{p})\|^2 + [z(\mathbf{p}) + \mathbf{p} \cdot \nabla z(\mathbf{p})]^2} \tag{45}$$

In comparison with Equations (27), the PDEs (44) explicitely depend on the depth z, but contain two difficulties: they are nonlinear and cannot be solved locally. We can eliminate the nonlinearity due to the normalization term $\|\overline{\mathbf{n}}(\mathbf{p})\|$. Indeed, neither the relative albedo $\overline{\rho}(\mathbf{p})$, nor the norm $\|\overline{\mathbf{n}}(\mathbf{p})\|$, depend on the index i of the LED. We can thus deduce from a pair of equations such as (44), that is to say for $(i, j) \in [1, m]^2$, $i < j$, the following equalities:

$$\frac{\overline{\rho}(\mathbf{p})}{\|\overline{\mathbf{n}}(\mathbf{p})\|} = \frac{I^i(\mathbf{p})}{\mathbf{a}^i_z(\mathbf{p}) \cdot \nabla z(\mathbf{p}) - b^i_z(\mathbf{p})\, z(\mathbf{p})} = \frac{I^j(\mathbf{p})}{\mathbf{a}^j_z(\mathbf{p}) \cdot \nabla z(\mathbf{p}) - b^j_z(\mathbf{p})\, z(\mathbf{p})} \tag{46}$$

where $\mathbf{a}^i_z(\mathbf{p})$ and $b^i_z(\mathbf{p})$ are defined as follows:

$$\mathbf{a}^i_z(\mathbf{p}) = f \begin{bmatrix} t^i_{z,1}(\mathbf{p}) \\ t^i_{z,2}(\mathbf{p}) \end{bmatrix} - t^i_{z,3}(\mathbf{p})\, \mathbf{p} \qquad \text{and} \qquad b^i_z(\mathbf{p}) = t^i_{z,3}(\mathbf{p}) \tag{47}$$

where we denote $\mathbf{t}^i_z(\mathbf{p}) = [t^i_{z,1}(\mathbf{p}), t^i_{z,2}(\mathbf{p}), t^i_{z,3}(\mathbf{p})]^\top$.

From Equations (46), we obtain the following PDE in z:

$$\underbrace{\left[I^i(\mathbf{p})\, \mathbf{a}^j_z(\mathbf{p}) - I^j(\mathbf{p})\, \mathbf{a}^i_z(\mathbf{p})\right]}_{\mathbf{a}^{i,j}_z(\mathbf{p})} \cdot \nabla z(\mathbf{p}) = \underbrace{\left[I^i(\mathbf{p})\, b^j_z(\mathbf{p}) - I^j(\mathbf{p})\, b^i_z(\mathbf{p})\right]}_{b^{i,j}_z(\mathbf{p})} z(\mathbf{p}) \tag{48}$$

The fields $\mathbf{a}^{i,j}_z(\mathbf{p})$ and $b^{i,j}_z(\mathbf{p})$ defined in (48) depend on z but not on ∇z. The PDE (48) is thus *quasi-linear*. It could be solved by the characteristic strips expansion method [28] 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 $m = 8$ images, we follow the idea from [16,27,36,38]. This consists in writing $\binom{m}{2}$ PDEs such as (48), which correspond to the $\binom{m}{2}$ pairs $(i,j) \in [1,m]^2$, $i < j$. Moreover, we resort to robust estimation techniques. Forming the matrix field $\mathbf{A}_z : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^{\binom{m}{2} \times 2}$ by concatenation of the row vectors $\mathbf{a}_z^{i,j}(\mathbf{p})^\top$, and forming the vector field $\mathbf{b}_z : \Omega \subset \mathbb{R}^2 \to \mathbb{R}^{\binom{m}{2}}$ by concatenation of the scalar values $b_z^{i,j}(\mathbf{p})$, the system of PDEs to solve is written:

$$\mathbf{A}_z(\mathbf{p}) \nabla z(\mathbf{p}) = \mathbf{b}_z(\mathbf{p}) \, z(\mathbf{p}) \tag{49}$$

This new differential formulation of photometric stereo seems simpler than the original differential formulation (44), since the main source of nonlinearity, due to the denominator $\|\overline{\mathbf{n}}(\mathbf{p})\|$, has been eliminated. However, it admits more solutions than (44): it is easy to check that the system of PDEs (49) admits both trivial solutions $z \equiv 0$ and $z \equiv +\infty$, which are not solutions of (44). Moreover, we can predict that the solution which we are looking for will be approximate because of the redundancy of the PDEs system (49), while $z \equiv 0$ and $z \equiv +\infty$ may be trivial, yet they are exact solutions.

Since several solutions do exist, it is necessary to look for an approximate, *locally* optimal, solution of (49). It has been proposed in [27,36] to iteratively estimate this solution by resorting to a fixed point scheme. In order to avoid any drift towards the trivial solution $z \equiv 0$, this approach involves a regularization term penalizing the difference between z and a prior z_0 on the depth. Yet, it is not only not so frequent to have such a prior, but it is also difficult to correctly tune the relative importance of the regularization term.

We will see that one can avoid resorting to this regularization term thanks to the change of variable $\tilde{z} = \log(z)$, which is valid since $z > 0$. Dividing both sides of Equation (49) by $z(\mathbf{p})$, we get:

$$\mathbf{A}_{\exp\{\tilde{z}\}}(\mathbf{p}) \nabla \tilde{z}(\mathbf{p}) = \mathbf{b}_{\exp\{\tilde{z}\}}(\mathbf{p}) \tag{50}$$

whose approximate solution can be estimated by resorting to a variational approach.

Variational Resolution of System (50) – Exact resolution of System (50) is neither desirable, because of outliers to the Lambertian model such as noise and shadows, nor even possible, because these $\binom{m}{2}$ PDEs are usually incompatible, precisely because of these outliers.

We prefer to turn the resolution of this system into a variational problem whose formalism allows to take into account both the redundancy of the data and the outliers to the model. We assume that each of the $\binom{m}{2}$ equalities in System (50) is satisfied up to an additive noise. If we assume that this noise is Gaussian and has zero-mean[5], the optimal solution (in the maximum likelihood sense) is obtained

[5] In fact, any noise assumption should be formulated on the images, and not on Model (49), which was obtained by considering ratios of gray levels, cf. (50): if the noise on gray levels is Gaussian, then that on ratios is Cauchy-distributed [20]. Hence, the least-squares solution (51) is the best linear unbiased estimator, but it is not the optimal solution, which would require non-convex estimators [11]. Yet, we leave this as future work.

by minimizing the sum of the square residuals over the entire reconstruction domain Ω:

$$\min_{\tilde{z}:\Omega\to\mathbb{R}} \underbrace{\|\mathbf{A}_{\exp\{\tilde{z}\}}\,\nabla\tilde{z} - \mathbf{b}_{\exp\{\tilde{z}\}}\|^2_{L^2(\Omega)}}_{\mathcal{E}_2} \tag{51}$$

The nonlinearity of the least-squares problem (51) comes from the dependency on $\tilde{z}$ of the fields $\mathbf{A}_{\exp\{\tilde{z}\}}$ and $\mathbf{b}_{\exp\{\tilde{z}\}}$. It thus seems relevant to uncouple the (linear) estimation of $\tilde{z}$ from the (nonlinear) estimation of $\mathbf{A}_{\exp\{\tilde{z}\}}$ and $\mathbf{b}_{\exp\{\tilde{z}\}}$, by rewriting (51) as the following constrained optimization problem:

$$\begin{aligned} &\min_{\tilde{z}:\Omega\to\mathbb{R}} \|\mathbf{A}\,\nabla\tilde{z} - \mathbf{b}\|^2_{L^2(\Omega)} \\ &\text{s.c.} \begin{bmatrix}\mathbf{A}\\ \mathbf{b}\end{bmatrix} = \begin{bmatrix}\mathbf{A}_{\exp\{\tilde{z}\}}\\ \mathbf{b}_{\exp\{\tilde{z}\}}\end{bmatrix} \end{aligned} \tag{52}$$

which we can solve by resorting to a fixed point iterative scheme:

$$\tilde{z}^{(k+1)} = \operatorname*{argmin}_{\tilde{z}:\Omega\to\mathbb{R}} \|\mathbf{A}^{(k)}\,\nabla\tilde{z} - \mathbf{b}^{(k)}\|^2_{L^2(\Omega)} \tag{53}$$

$$\begin{bmatrix}\mathbf{A}^{(k+1)}\\ \mathbf{b}^{(k+1)}\end{bmatrix} = \begin{bmatrix}\mathbf{A}_{\exp\{\tilde{z}^{(k+1)}\}}\\ \mathbf{b}_{\exp\{\tilde{z}^{(k+1)}\}}\end{bmatrix} \tag{54}$$

The linear least-squares variational problem (53) can be solved by discretizing the gradient operator ∇ by finite differences (we used forward, first-order finite differences, with a Neumann boundary condition). The discrete optimization problem that we obtain this way admits as necessary optimality condition the following linear system:

$$\underbrace{\left(\overline{\mathbf{A}}^{(k)}\mathbf{D}\right)^\top \overline{\mathbf{A}}^{(k)}\mathbf{D}}_{\overline{\overline{\mathbf{A}}}^{(k)}}\,\tilde{\mathbf{z}} = \left(\overline{\mathbf{A}}^{(k)}\mathbf{D}\right)^\top \overline{\mathbf{b}}^{(k)} \tag{55}$$

where $\overline{\mathbf{A}}^{(k)}$, $\overline{\mathbf{b}}^{(k)}$ and $\mathbf{D}$ denote the discretizations of $\mathbf{A}^{(k)}$, $\mathbf{b}^{(k)}$ and ∇, and where the unknown vector $\tilde{\mathbf{z}}$ admits as coordinates the values of the function $\tilde{z}$ in pixels $\mathbf{p}\in\Omega$.

Matrix $\overline{\overline{\mathbf{A}}}^{(k)}$ in (55) is symmetric, positive semi-definite. Its kernel has dimension 1. This rank deficiency, which comes from the fact that $\mathbf{A}^{(k)}\,\nabla\tilde{z} - \mathbf{b}^{(k)} = \mathbf{A}^{(k)}\,\nabla(\tilde{z}+K) - \mathbf{b}^{(k)}$, for any $K\in\mathbb{R}$, would not occur with the nonlinear least-squares problem (51), because $\mathbf{A}_{\exp\{\tilde{z}\}}$ and $\mathbf{b}_{\exp\{\tilde{z}\}}$ are not invariant by translation of $\tilde{z}$. However, in practice, we found that conjugate gradient algorithms converge rapidly towards a solution of (55), provided that solution $\tilde{z}^{(k)}$ is used as initialization at iteration $(k+1)$.

Experimental Validation – It is empirically shown in [27,36] that fixed point schemes provide reasonable results, yet the question of convergence is overlooked. Yet, Figure 7 shows that, although the scheme indeed converges rapidly, its solution is not optimal in the sense of the energy $\mathcal{E}_2$ defined in (51). We will now introduce an ADMM algorithm allowing to avoid this issue.

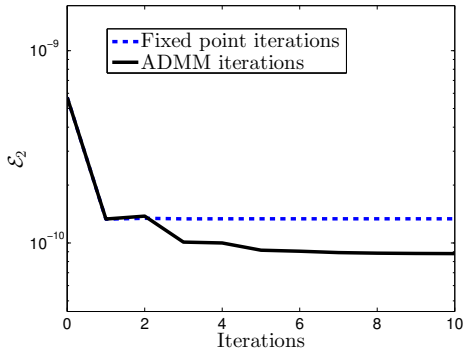


Fig. 7 Evolution of the energy $\mathcal{E}_2$ defined in (51), in function of the iterations, for the data of Figure 1 (the initial shape is a fronto-parallel plane with equation $z \equiv 700\ mm$). The fixed point scheme converges towards a non-optimal solution, contrarily to the ADMM scheme we are going to introduce.

3.3 ADMM Resolution Scheme

Because of the nonlinear dependency in $\tilde{z}$ of $\mathbf{A}_{\exp\{\tilde{z}\}}$ and $\mathbf{b}_{\exp\{\tilde{z}\}}$, it is difficult to guarantee good convergence properties for the fixed point scheme of the previous section. Instead of "freezing" the nonlinearities of the variational problem (51), we propose to estimate $\tilde{z}$ not only from the linearized parts, but also from the nonlinear ones. In this view, we introduce an auxiliary variable $\overline{z}$ and we reformulate Problem (51) in the following manner:

$$\begin{aligned} &\min_{\overline{z},\tilde{z}} \left\| \mathbf{A}_{\exp\{\overline{z}\}} \nabla\tilde{z} - \mathbf{b}_{\exp\{\overline{z}\}} \right\|^2_{L^2(\Omega)} \\ &\text{s.c. } \tilde{z} = \overline{z} \end{aligned} \tag{56}$$

In order to solve the constrained optimization problem (56), we use an augmented Lagrangian approach with alternating descent direction (ADMM).This standard procedure, which dates back to the 70s [13,15], has recently been revisited [8]. By introducing a dual variable w and a descent step ν, a local solution of (56) is obtained at convergence of the following iterative scheme:

1. Update $\tilde{z}$ by using the linear part, "while staying close to $\overline{z}$":

$$\tilde{z}^{(k+1)} = \operatorname*{argmin}_{\tilde{z}} \left\| \mathbf{A}_{\exp\{\overline{z}^{(k)}\}} \nabla\tilde{z} - \mathbf{b}_{\exp\{\overline{z}^{(k)}\}} \right\|^2_{L^2(\Omega)} + \frac{1}{2\,\nu} \left\| \tilde{z} - \overline{z}^{(k)} + w^{(k)} \right\|^2_{L^2(\Omega)} \tag{57}$$

2. Update $\overline{z}$ by using the nonlinear part, "while keeping close to $\tilde{z}$":

$$\overline{z}^{(k+1)} = \operatorname*{argmin}_{\overline{z}} \left\| \mathbf{A}_{\exp\{\overline{z}\}} \nabla\tilde{z}^{(k+1)} - \mathbf{b}_{\exp\{\overline{z}\}} \right\|^2_{L^2(\Omega)} + \frac{1}{2\,\nu} \left\| \tilde{z}^{(k+1)} - \overline{z} + w^{(k)} \right\|^2_{L^2(\Omega)} \tag{58}$$

3. Update the dual variable, in order to guarantee that the constraint $\tilde{z} = \overline{z}$ is satisfied:

$$w^{(k+1)} = w^{(k)} + \tilde{z}^{(k+1)} - \overline{z}^{(k+1)} \tag{59}$$

The first stage (57) is a linear least-squares problem which can be solved in the same way as in the previous scheme, using the normal equations. Note that the presence of the regularization term guarantees this time the positive definiteness of the matrix of the system, and therefore the convergence of the conjugate gradient algorithm for its resolution.

With the introduction of the auxiliary variable $\overline{z}$, which decouples the terms in $\nabla\tilde{z}$ and $\tilde{z}$, the second stage (58) is a *local* nonlinear least-squares problem: in fact, the gradient of $\overline{z}$ is not involved in this problem, which can be resolved pixelwise. Problem (58) 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 (58), it is unfortunately impossible to guarantee convergence of the proposed ADMM scheme, which depends on the initialization and on the descent parameter ν [8]. A reasonable initialization strategy consists in using the solution provided by the scheme of Section 3.1. As for the descent parameter, we iteratively calculate its optimal value according to the *Penalty Varying Parameter* procedure described in [8]. Finally, the iteration stops when the relative variation of the criterion of Problem (58) falls under a threshold equal to 10^{-4}.

Despite this theoretical difficulty, the proposed ADMM scheme allows quite simply to separate the difficulties, namely the global nature of the problem (presence of the gradient term) and its nonlinearity. Although this scheme should be considered as a *local* optimization method, "*The hope is that it will possibly have better convergence properties than other local optimization methods, where* better *convergence can mean faster convergence or convergence to a point with better objective value*" [8]. In practice, it allows to solve Problem (51) more efficiently than the fixed point scheme.

Experimental Validation – Figure 8 shows that this ADMM scheme guarantees better convergence properties than the fixed point scheme used in the literature. Yet, let us recall that it is a *local* optimisation method: il would be futile to seek the global minimum of energy $\mathcal{E}_2$, since we already know that it admits several minima. This is confirmed by Figure 8: even in the vicinity of the sought solution, several other local minima may exist.

The choice of an appropriate initial shape is therefore very important. The iterative scheme of Section 3.1 is a natural choice. Indeed, we have seen that, despite some flaws, it can estimate the relief unambiguously, whatever the initialization. In conclusion, the method that we recommend, and which is illustrated by the results in Figure 9, involves the following two stages:

1. Approximate estimation of shape by the iterative scheme of Section 3.1.
2. Refinement of this estimation by the above ADMM scheme.

To quantitatively evaluate the improvement due to the ADMM scheme, we tested the full method on the data of Figure 1. At first sight, the shape depicted in Figure 10-a seems hardly different from that of Figure 5-a, but the comparison of histograms in Figures 6-a and 10-b shows that bias has been significantly reduced.

In the next section, we will show that another advantage of the proposed method is that it can very easily be extended to RGB-valued images.

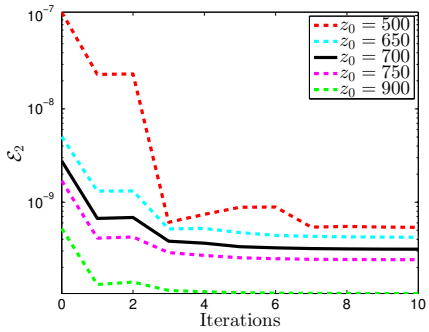


Fig. 8 Evolution of the energy $\mathcal{E}_2$ defined in (51), in function of the iterations, for the data of Figure 1. Using as initialization $\tilde{z} \equiv \log(z_0)$, the ADMM scheme always converges towards a local minimum, yet this minimum strongly depends on the value of z_0.

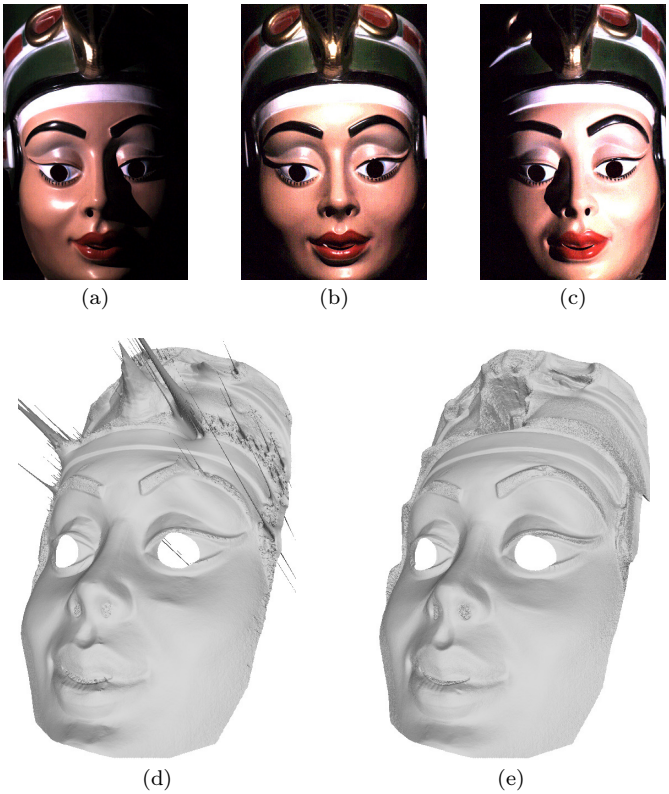
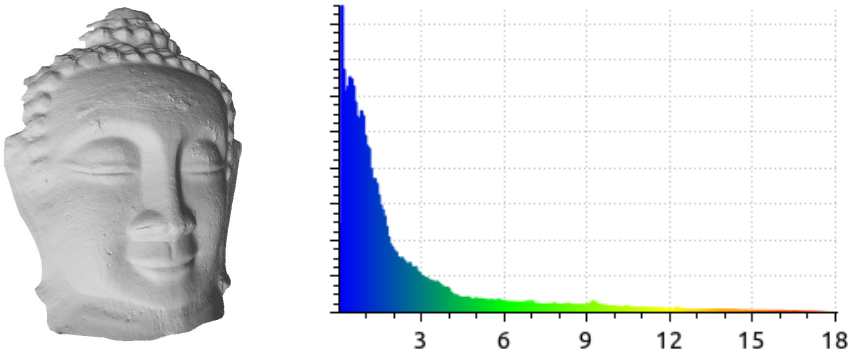


Fig. 9 (a-b-c) Three images (out of $m = 8$) of a plastic mask. (d) 3D-reconstruction obtained with the iterative scheme of Section 3.1. Artifacts appear where the estimated normal field is not integrable. (e) Refinement by the ADMM scheme from Section 3.3.



(a) (b)

Fig. 10 (a) 3D-reconstruction obtained with the ADMM scheme. (b) Histogram of point-to-point distances between this shape and the ground truth (cf. Figure 5-c). The median value is worth 1.2 *mm*. The bias one can observe in the histogram of Figure 6 has significantly been reduced.

4 Estimating Colored 3D-models by Photometric Stereo

4.1 Spectral Dependency of the Luminous Flux Emitted by a LED

To extend our study to color, we need to introduce a spectral dependency in the luminous flux model (12). It seems reasonable to limit this dependency to the intensity Φ (λ denotes the wavelength):

$$\mathbf{s}(\mathbf{x},\lambda) = \Phi(\lambda)\,\cos^{\mu}\theta\,\frac{\mathbf{x}_s-\mathbf{x}}{\|\mathbf{x}_s-\mathbf{x}\|^3} \tag{60}$$

Model (60) is more complex than Model (12), because the intensity Φ_0, which is a real parameter, has been replaced by the *emission spectrum* $\Phi(\lambda)$, which is a function. Calibration of $\Phi(\lambda)$ could be achieved by using a *spectrometer*, but we would prefer to extend the procedure from Section 2.2, which requires nothing else than a camera and calibration patterns.

For a point $\mathbf{x}$ of a Lambertian surface with albedo $\rho(\mathbf{x})$, illuminated by a luminous flux $\mathbf{s}(\mathbf{x})$, we get from (13), (14) and (15) the expression of the irradiance $\epsilon(\mathbf{p})$ of the image plane at pixel $\mathbf{p}$ conjugate to $\mathbf{x}$:

$$\epsilon(\mathbf{p}) = \beta\,\cos^4\alpha(\mathbf{p})\,\frac{\rho(\mathbf{x})}{\pi}\,\mathbf{s}(\mathbf{x})\cdot\mathbf{n}(\mathbf{x}) \tag{61}$$

It is easy to extend this expression to the more general case where the luminous flux $\mathbf{s}(\mathbf{x})$ and the albedo $\rho(\mathbf{x})$ present spectral dependencies:

$$\epsilon(\mathbf{p},\lambda) = \beta\,\cos^4\alpha(\mathbf{p})\,\frac{\rho(\mathbf{x},\lambda)}{\pi}\,\mathbf{s}(\mathbf{x},\lambda)\cdot\mathbf{n}(\mathbf{x}) \tag{62}$$

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.

Thanks to the bijective correspondance between a point $\mathbf{x}$ and a pixel $\mathbf{p}$, we denote $\rho(\mathbf{p},\lambda)$ in lieu of $\rho(\mathbf{x},\lambda)$, and $\mathbf{n}(\mathbf{p})$ in lieu of $\mathbf{n}(\mathbf{x})$. Let us define the *color level* $I_\star(\mathbf{p})$, $\star \in \{R,G,B\}$, similarly to the expression (17) of the gray level $I(\mathbf{p})$, by integrating (62) over the entire spectrum:

$$I_\star(\mathbf{p}) = \frac{\gamma\,\beta}{\pi}\left[\int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\rho(\mathbf{p},\lambda)\,\mathbf{s}(\mathbf{x},\lambda)\,\mathrm{d}\lambda\right]\cdot\mathbf{n}(\mathbf{p}) \tag{63}$$

Using a Lambertian calibration pattern which is uniformly white, that is to say such that $\rho(\mathbf{p},\lambda)\equiv\rho_0$, allows us to rewrite (63) as follows:

$$I_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho_0}{\pi}\left[\int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\mathbf{s}(\mathbf{x},\lambda)\,\mathrm{d}\lambda\right]\cdot\mathbf{n}(\mathbf{p}) \tag{64}$$

This expression of $I_\star(\mathbf{p})$ is an extension of (17) to RGB images. Indeed, it can be rewritten:

$$I_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho_0}{\pi}\,\mathbf{s}_\star(\mathbf{x})\cdot\mathbf{n}(\mathbf{p}) \tag{65}$$

if $\mathbf{s}_\star(\mathbf{x})$, $\star\in\{R,G,B\}$, is defined as follows:

$$\mathbf{s}_\star(\mathbf{x}) = \int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\mathbf{s}(\mathbf{x},\lambda)\,\mathrm{d}\lambda \tag{66}$$

Plugging the expression (60) of $\mathbf{s}(\mathbf{x},\lambda)$ into (66), we obtain the extension to RGB images of the model (12) of the luminous flux emitted by a LED:

$$\mathbf{s}_\star(\mathbf{x}) = \Phi_\star\,\cos^\mu\theta\,\frac{\mathbf{x}_s-\mathbf{x}}{\|\mathbf{x}_s-\mathbf{x}\|^3} \tag{67}$$

where the *colored intensity* $\Phi_\star$, $\star\in\{R,G,B\}$ is defined as follows:

$$\Phi_\star = \int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\Phi(\lambda)\,\mathrm{d}\lambda \tag{68}$$

The spectral dependency of the luminous flux emitted by a LED is thus described by the model (67), 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 and Φ_B and the anisotropy parameter μ. Since the definition (68) depends on function $c_\star(\lambda)$, which depends on the sensor, the intensities Φ_R, Φ_G and Φ_B are not intrinsic parameters of the LED, because they depend on the LED-camera system. Yet, we cannot obtain any other information on the function $\Phi(\lambda)$ than these three colored intensities if we use a three channel camera.

To achieve a more accurate calibration, we would need more colored filters, that is to say we should resort to multi-, or hyper-spectral imagery. Using *monochromatic* filters would even allow us to sample the function $\Phi(\lambda)$. Indeed, if the transmission spectrum of a colored filter can be written $c_{\lambda_0}(\lambda) = K_{\lambda_0}\,\delta_{\lambda_0}(\lambda)$, K_{λ_0} being a constant and $\delta_{\lambda_0}(\lambda)$ the Dirac function centered in λ_0, then the colored intensity (68) becomes:

$$\Phi_{\lambda_0} = \int_{\lambda=0}^{+\infty} c_{\lambda_0}(\lambda)\,\Phi(\lambda)\,\mathrm{d}\lambda = K_{\lambda_0}\,\Phi(\lambda_0) \tag{69}$$

but, even in this case, Φ_{λ_0} depends on the sensor through K_{λ_0}.

4.2 Spectral Calibration of the Luminous Flux Emitted by a LED

If we use again the planar Lambertian calibration pattern from Section 2.2, which is convex, the incident luminous flux comes solely from the LED. We can thus replace $\mathbf{s}_\star(\mathbf{x})$ by its definition (67) in the expression (65) of the color level $I_\star(\mathbf{p})$. If we assume 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 (q is the number of poses of the calibration pattern):

$$\min_{\mathbf{m}_{s,\star}} \sum_{j=1}^{q} \sum_{\mathbf{p}\in\Omega^j} \left[\mathbf{m}_{s,\star} \cdot (\mathbf{x}^j - \mathbf{x}_s) - \left[I_\star^j(\mathbf{p}) \, \frac{\|\mathbf{x}_s - \mathbf{x}^j\|^{3+\mu}}{(\mathbf{x}_s - \mathbf{x}^j) \cdot \mathbf{n}^j} \right]^{1/\mu} \right]^2 \tag{70}$$

where $\mathbf{m}_{s,\star}$ is defined by analogy with $\mathbf{m}_s$:

$$\mathbf{m}_{s,\star} = {\Psi_\star}^{1/\mu} \, \mathbf{n}_s \tag{71}$$

and $\Psi_\star$ is defined by analogy with Ψ:

$$\Psi_\star = \gamma \, \beta \, \frac{\rho_0}{\pi} \, \Phi_\star \tag{72}$$

Each problem (70) allows us to estimate a colored intensity Φ_R, Φ_G or Φ_B (up to a constant) and the main direction $\mathbf{n}_s$, which is thus estimated three times. Table 1 groups the values obtained for one of the LEDs of our device. The three estimates of $\mathbf{n}_s$ are consistent, but instead of arbitrarily choosing one of them, we compute their weighted mean:

$$\widehat{\mathbf{n}}_s = \frac{\pi_R \, \widehat{\mathbf{n}}_{s,R} + \pi_G \, \widehat{\mathbf{n}}_{s,G} + \pi_B \, \widehat{\mathbf{n}}_{s,B}}{\pi_G + \pi_G + \pi_B} \tag{73}$$

where the weights π_R, π_G and π_B are defined as follows:

$$\pi_\star = \sum_{j=1}^{q} \sum_{\mathbf{p}\in\Omega^j} I_\star^j(\mathbf{p}), \qquad \star \in \{R, G, B\} \tag{74}$$

then we normalize the expression (73) of $\widehat{\mathbf{n}}_s$ to ensure that it has unit-length.

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 device, estimated by resolution of Problem (70) with respect to each color channel.

In this table, the values of $\widehat{\Psi}_R$, $\widehat{\Psi}_G$ and $\widehat{\Psi}_B$ are given without units because, from the definition (72) of Ψ, only their relative values are meaningful. As it happens,

the value of $\widehat{\Psi}_G$ is roughly twice as much as that of $\widehat{\Psi}_R$ and $\widehat{\Psi}_B$, but this does not mean that the value of $\Phi(\lambda)$ is twice larger in the green range than in the red or in the blue ranges, since the definition (68) of the colored intensity $\Phi_\star$ depends on the transmission spectrum $c_\star(\lambda)$ in the considered channel.

Our calibration procedure relies on the assumption that $\rho(\mathbf{p}, \lambda) \equiv \rho_0$, which is probably not true, 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, that is to say $\rho(\mathbf{p}, \lambda) = \rho(\lambda)$, and if we denote ρ_0 the maximum value of $\rho(\lambda)$, Equation (64) is still valid, provided that $c_\star(\lambda)$ is replaced by the function $c'_\star(\lambda)$ defined as follows:

$$c'_\star(\lambda) = \frac{\rho(\lambda)}{\rho_0}\, c_\star(\lambda) \tag{75}$$

Of course, similar arguments can be used for the rest of the rationale, but we must remember that the colored intensity $\Phi_\star$, which depends on the transmission spectrum $c_\star(\lambda)$ by its definition (68), 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 under point light source illumination from Section 2.3, but we can already predict that the color of the paper will have some influence on the estimated color of the observed scene.

4.3 Extension to RGB Images of Photometric Stereo under Point Light Source Illumination

If we pretend to extend the model (17) of gray level to RGB images, then it must be possible to write the color level at $\mathbf{p}$ in the channel $\star \in \{R, G, B\}$ in the following manner:

$$I_\star(\mathbf{p}) = \gamma\, \beta\, \frac{\rho_\star(\mathbf{p})}{\pi}\, \mathbf{s}_\star(\mathbf{x}) \cdot \mathbf{n}(\mathbf{p}) \tag{76}$$

where the *colored albedo* $\rho_\star(\mathbf{p})$ is an extension of the albedo $\rho(\mathbf{p})$ to color. Comparing this expression of the color level $I_\star(\mathbf{p})$ with its other expression (63), and using the definition (66) of $\mathbf{s}_\star(\mathbf{x})$, we obtain the following equality:

$$\rho_\star(\mathbf{p}) \int_{\lambda=0}^{+\infty} c_\star(\lambda)\, \mathbf{s}(\mathbf{x}, \lambda)\, \mathrm{d}\lambda = \int_{\lambda=0}^{+\infty} c_\star(\lambda)\, \rho(\mathbf{p}, \lambda)\, \mathbf{s}(\mathbf{x}, \lambda)\, \mathrm{d}\lambda \tag{77}$$

When the surface is illuminated by a LED, we can replace $\mathbf{s}(\mathbf{x}, \lambda)$ in (77) by the luminous flux model (60), which provides us with the following expression for the colored albedo:

$$\rho_\star(\mathbf{p}) = \frac{\displaystyle\int_{\lambda=0}^{+\infty} c_\star(\lambda)\, \rho(\mathbf{p}, \lambda)\, \phi(\lambda)\, \mathrm{d}\lambda}{\displaystyle\int_{\lambda=0}^{+\infty} c_\star(\lambda)\, \phi(\lambda)\, \mathrm{d}\lambda}, \qquad \star \in \{R, G, B\} \tag{78}$$

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 subscript the colored albedo by the number of the LED. It thus seems that we have to solve, in each pixel $\mathbf{p} \in \Omega$, the following problem:

$$I^i_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho^i_\star(\mathbf{p})}{\pi}\,\mathbf{s}^i_\star(\mathbf{x})\cdot\mathbf{n}(\mathbf{p}), \qquad i \in [1,m],\ \star \in \{R,G,B\} \tag{79}$$

System (79) is ill-posed, because it contains $3m$ equations with $3m+5$ unknowns: a colored albedo $\rho^i_\star(\mathbf{p})$ per equation, the three coordinates of $\mathbf{x}$ and two values for the normal $\mathbf{n}(\mathbf{p})$. Apart from this numerical problem, the dependency on i of the colored albedo 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 albedo. This shows that the extension to RGB images of photometric stereo is potentially complex.

Photometric stereo can thus not be extended to RGB images in the general case, but this is still possible in two specific cases [36]:

- For a white surface, that is to say when $\rho(\mathbf{p},\lambda) = \rho(\mathbf{p})$, we deduce from (78) that $\rho_R(\mathbf{p}) = \rho_G(\mathbf{p}) = \rho_B(\mathbf{p}) = \rho(\mathbf{p})$. Problem (79) is thus written:

$$I^i_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho(\mathbf{p})}{\pi}\,\mathbf{s}^i_\star(\mathbf{x})\cdot\mathbf{n}(\mathbf{p}), \qquad i \in [1,m],\ \star \in \{R,G,B\} \tag{80}$$

  The case where lighting is directional, that is to say independent from $\mathbf{x}$, is very interesting, because the number of unknowns is then equal to three: a single RGB image is thus enough to ensure that the problem is well-posed. This well-known case, which dates back to the 90's [25], has been applied by Hernandez et al. to real-time 3D-reconstruction of a deformable (white) surface [18].
- When the sources are white, that is to say when $\phi(\lambda) \equiv \phi_0$, (78) gives:

$$\rho_\star(\mathbf{p}) = \frac{\displaystyle\int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\rho(\mathbf{p},\lambda)\,\mathrm{d}\lambda}{\displaystyle\int_{\lambda=0}^{+\infty} c_\star(\lambda)\,\mathrm{d}\lambda}, \qquad \star \in \{R,G,B\} \tag{81}$$

  Since this expression is independent from i, Problem (79) is rewritten:

$$I^i_\star(\mathbf{p}) = \gamma\,\beta\,\frac{\rho_\star(\mathbf{p})}{\pi}\,\mathbf{s}^i_\star(\mathbf{x})\cdot\mathbf{n}(\mathbf{p}), \qquad i \in [1,m],\ \star \in \{R,G,B\} \tag{82}$$

  In (82), the luminous flux $\mathbf{s}^i_\star(\mathbf{x})$ indeed depends on the channel $\star$, although the spectrum $\phi(\lambda)$ is assumed to be independent from λ since, according to the definition (66), the transmission spectrum $c_\star(\lambda)$ depends on $\star$. System (82), which contains $3m$ equations and eight unknowns, is well-posed as soon as $m \geqslant 3$.

Another situation where the colored albedo is independent from i is when the m LEDs all share the same emission spectrum $\phi(\lambda)$. In the case of our device (cf. Figure 1-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 the model (82) in photometric stereo.

By using the model (67) for the colored luminous flux $\mathbf{s}_\star(\mathbf{x})$ emitted by a LED, Problem (82) is rewritten in the following way:

$$I^i_\star(\mathbf{p}) = \Psi^i_\star \frac{\rho_\star(\mathbf{p})}{\rho_0} \left[\mathbf{n}^i_s \cdot \frac{\mathbf{x} - \mathbf{x}^i_s}{\|\mathbf{x} - \mathbf{x}^i_s\|} \right]^{\mu^i} \frac{(\mathbf{x}^i_s - \mathbf{x}) \cdot \mathbf{n}(\mathbf{p})}{\|\mathbf{x}^i_s - \mathbf{x}\|^3}, \quad i \in [1, m], \ \star \in \{R, G, B\} \tag{83}$$

which is an extension to RGB images of the photometric stereo model (27). In our experiments, we use the model (83) 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 abusive (let us recall that it is perfectly justified only if the sources share the same spectrum).

The calibration procedure described in Section 4.2 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 (83) are the position $\mathbf{x}$ of the 3D-point conjugate to $\mathbf{p}$, its three colored albedos $\rho_\star(\mathbf{p})$, $\star \in \{R, G, B\}$, and its normal $\mathbf{n}(\mathbf{p})$. Resorting to RGB images thus allows us to replace the system (27) of m equations with six unknowns, by the system (83) of $3m$ equations with eight unknowns. We can thus expect more accurate results.

4.4 Resolution of Colored Photometric Stereo under Point Light Source Illumination

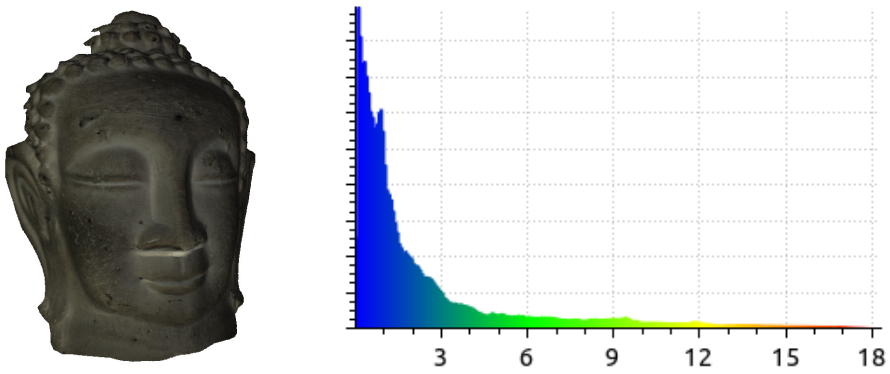
To solve Problem (83), we follow an approach similar to that of Section 3.2: we compute the ratios between color levels in each channel $\star \in \{R, G, B\}$, thus eliminating the colored albedos $\rho_\star(\mathbf{p})$ and obtaining a system of PDEs in z similar to (49). We believe that this method for separating the estimation of shape from that of the albedos is better justified than that advocated in [4], where principal component analysis is used. In addition, the PDEs to solve are quasi-linear, while simultaneous estimation of shape and reflectance is a nonlinear problem [24].

Since the new system of PDEs contains $3\binom{m}{2}$ equations, in lieu of $3m$ equations for (49), its resolution by the ADMM scheme should improve the results. Indeed, the median of the distances between the shape of Figure 11-a and the ground truth is equal to 1.1 mm when using RGB images, in lieu of 1.2 mm for gray level images (cf. Figure 10-a). As for the colored albedos, they are a posteriori estimated by plugging the estimated values of $\mathbf{x}$ and $\mathbf{n}(\mathbf{p})$ into (83).

The result of Figure 12 is improved in two ways, in comparison with that of Figure 9: colored albedos are estimated; similarly to the example of Figure 11, using RGB images should probably improve the accuracy of shape estimation, although we do not have any ground truth to support this claim.

However, the ADMM scheme has one weakness which was already mentioned in Section 3.3: it may converge slowly if the initial shape is far away from the solution. To speed up the calculations, we proceed in three stages:

1. Estimation of shape by the iterative scheme from Section 3.1 (cf. Figure 9-d), by using the gray level images.
2. Refinement of this result by the ADMM scheme, by using the RGB images.
3. Estimation of colored albedos from this second estimation of shape.



(a) (b)

Fig. 11 (a) 3D-model estimated from the $m = 8$ images (cf. Figure 1). (b) Histogram of the distances between this shape and the ground truth. Using RGB images allows to improve the result of Figure 11: the median of the distances to the ground truth is now equal to 1.1 mm.

To obtain the 3D-model depicted in Figure 12, which contains 1095396 points, the total computation time is of the order of one minute.

Fig. 12 Relighting of the 3D-model obtained from $m = 8$ RGB images (cf. Figure 9).

The result of Figure 13 shows that our initial goal i.e., the estimation of colored 3D-models of human faces by photometric stereo, has been reached. The beard shape is clearly visible in Figure 13-d, which demonstrates the high level of accuracy of the estimation.

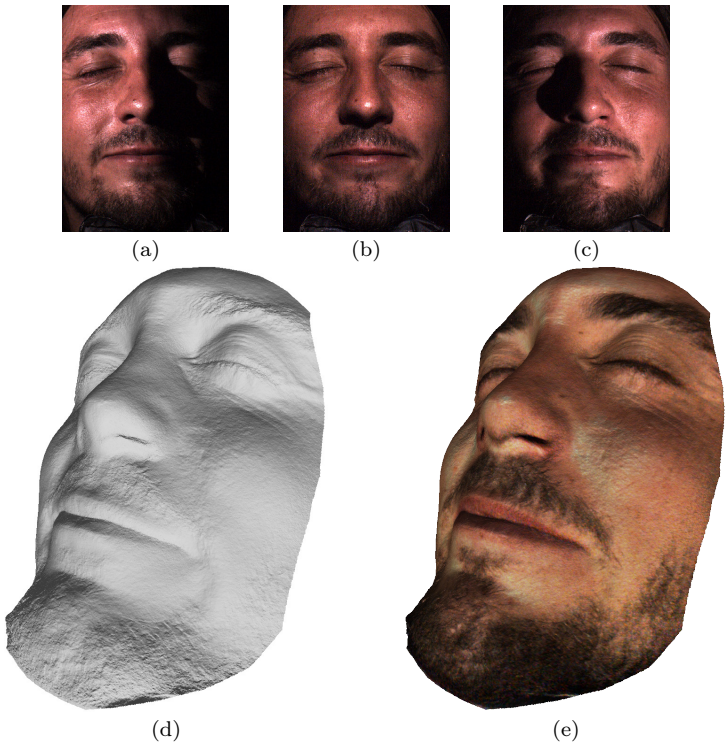


Fig. 13 (a-b-c) Three RGB images (out of $m = 8$) of a human face obtained by the device of Figure 1-a. (d) Estimated shape. (e) Estimated 3D-model.

5 Conclusion and Perspectives

In this article, we have presented a photometric stereo-based 3D-reconstruction device using LEDs as light sources. We have first modeled the luminous flux emitted by a LED, then the resulting photometric stereo problem. The calculation of ratios between gray levels makes the problem quasi-linear, which makes its resolution easier. The resolution is performed in two stages: using a fast iterative algorithm provides a visually satisfactory shape; this shape is used as initialization of an ADMM scheme, which allows to refine it. Finally, we have extended the whole study to RGB images.

The result of Figure 13 suggests that our goal 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, LEDs, a bright spherical calibration pattern, a planar Lambertian calibration pattern, and eventually 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.

There are two types of points where the shape, and therefore the albedo, are poorly estimated. The object visible in Figure 14 is a dental plaster cast. The estimated albedo should be uniform, since the material is homogeneous, which contradicts the result of Figure 14-e. The bias on the estimation of the albedo,

which corresponds to the concave areas, is clearly due to shadows, although we have taken care of eliminating, in each pixel, the two images with lower luminance. The same observation can be made about the example in Figure 15: the area under the nose, which is dimly lit, is poorly reconstructed (this problem does not appear in the example of Figure 13, because the face is oriented in such a way that it is "well" lighted).

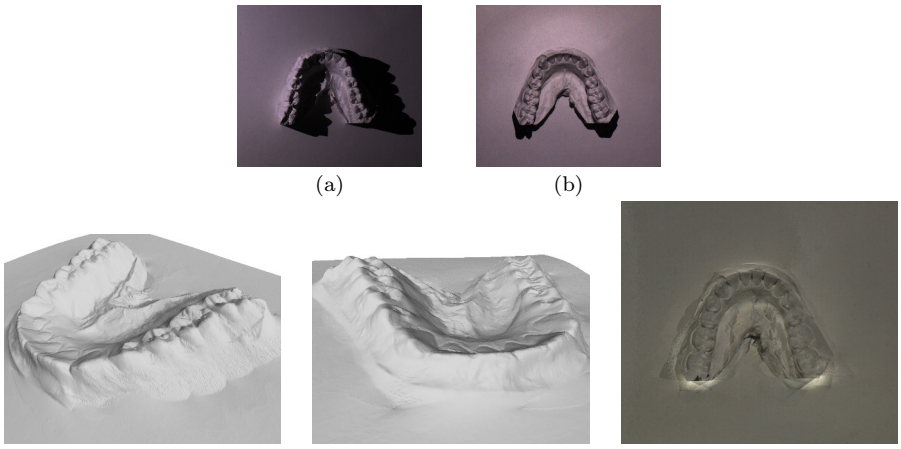


Fig. 14 (a-b) Two images (out of $m = 8$) of a dental plaster cast. (c-d) Estimated shape, seen from two angles. (e) Estimated albedo.

The example of Figure 16 shows that the estimation is incorrect when the albedo is uniformly null, that is to say, when $\rho(\mathbf{p}, \lambda) \equiv 0$, because the color levels at these points do not provide useful information. The photographed object is a comic book. While shading effects are clearly visible in the images (Figures 16-a and 16-b), the albedo appears correctly estimated (see Figure 16-d). However, the shape is badly estimated along the edges of the boxes, which are black inked (see Figure 16-c). To overcome this problem, the depth could be regularized at the points with null albedo, which are easy to detect.

Other future prospects than dealing with these defects arise. In particular, our 3D-scanner only provides 2.5D-models, which limits interest, for instance for augmented reality applications. For a complete 3D-model, it might be useful to couple the proposed method with multi-view 3D-reconstruction techniques [19].

In this article, we have shown that the differential formulation of photometric stereo is preferable to its non-differential one, but also that the reprojection error $\mathcal{E}_1$ is a more relevant criterion than the error $\mathcal{E}_2$ on image ratios, which leads to several local minima. We conclude that it would be preferable to minimize $\mathcal{E}_1$ through the differential formulation. Of course, this would be difficult because of the non-convexity due to the normalization term. Proximal methods designed for non-convex optimization could be used [31]. The extension to point light sources of the method proposed in [21] would probably allow us to achieve even greater accuracy.

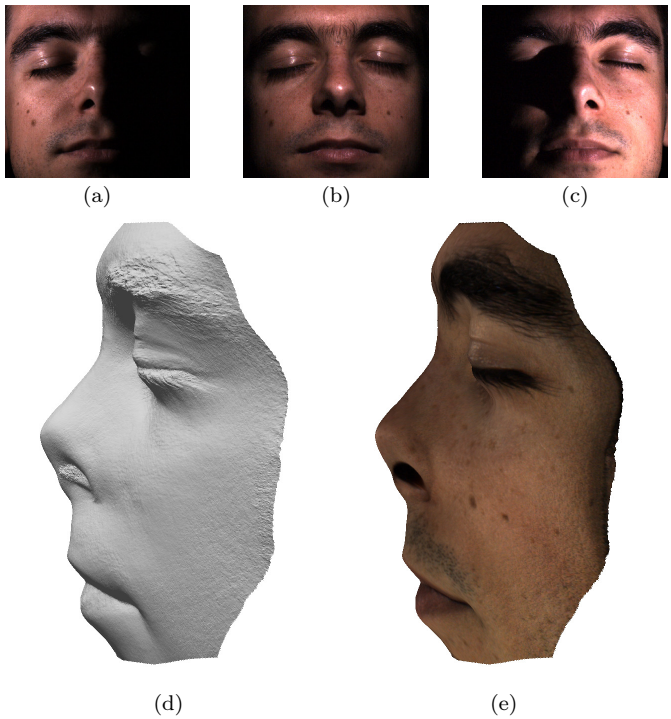


Fig. 15 (a-b-c) Three images (out of $m = 8$) of a face. (d-e) Estimated shape and 3D-model.

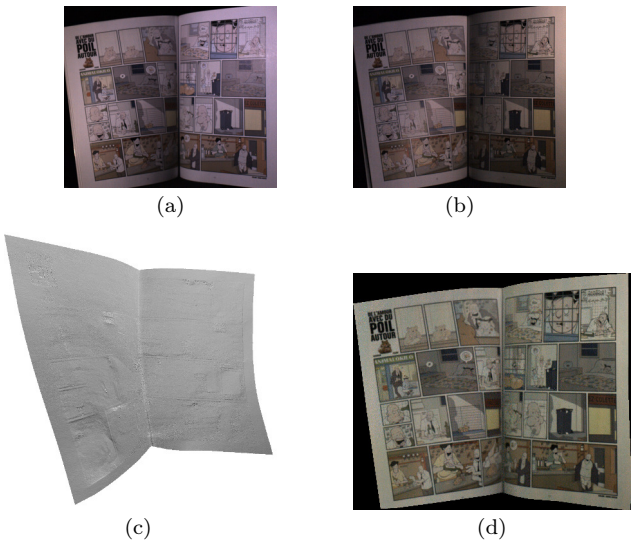


Fig. 16 (a-b) Two images (out of $m = 8$) of a comic book. (c) Estimated shape: artifacts appear around black inked points. (d) Estimated albedo.

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) 6
2. 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) 5
3. 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) 6
4. 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) 26
5. Basri, R., Jacobs, D.W.: Lambertian reflectance and linear subspaces. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(2), 218–233 (2003) 5
6. 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) 5
7. 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, 7, 10, 11
8. 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) 18, 19
9. 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) 6
10. 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) 11
11. 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) 14, 16
12. Durou, J.D., Quéau, Y., Aujol, J.F.: Normal Integration – Part I: A Survey (2016). URL https://hal.archives-ouvertes.fr/hal-01334349 3, 12
13. 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) 18
14. 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) 6
15. 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) 18
16. 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. Boston, USA (2015) 16
17. 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) 6
18. 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) 25
19. Hernández, C., Vogiatzis, G., Cipolla, R.: Multiview Photometric Stereo. IEEE Transactions on Pattern Analysis and Machine Intelligence **30**(3), 548–554 (2008) 29

20. Hinkley, D.V.: On the Ratio of Two Correlated Normal Random Variables. Biometrika **56**(3), 635–639 (1969) 16
21. Hoeltgen, L., Quéau, Y., Breuss, M., Radow, G.: Optimised photometric stereo via non-convex variational minimisation. In: Proceedings of the 27$^{\text{th}}$ British Machine Vision Conference. York, UK (2016) 29
22. Horn, B.K.P., Brooks, M.J. (eds.): Shape from Shading. MIT Press (1989) 2, 7
23. 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, 6, 11
24. 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) 26
25. 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) 25
26. McGunnigle, G., Chantler, M.J.: Resolving handwriting from background printing using photometric stereo. Pattern Recognition **36**(8), 1869–1879 (2003) 5
27. Mecca, R., Quéau, Y., Logothetis, F., Cipolla, R.: A Single Lobe Photometric Stereo Approach for Heterogeneous Material. SIAM Journal on Imaging Sciences (2016) 5, 16, 17
28. 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) 5, 15
29. 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) 3
30. Moreno, I., Sun, C.C.: Modeling the radiation pattern of LEDs. Optics Express **16**(3), 1808–1819 (2008) 3
31. Ochs, P., Chen, Y., Brox, T., Pock, T.: iPiano: Inertial Proximal Algorithm for Nonconvex Optimization. SIAM Journal on Imaging Sciences **7**(2), 1388–1419 (2014) 29
32. 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) 13
33. Papadhimitri, T., Favaro, P.: Uncalibrated Near-Light Photometric Stereo. In: Proceedings of the 25$^{\text{th}}$ British Machine Vision Conference. Nottingham, UK (2014) 3, 6, 11
34. 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) 6
35. 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) 6
36. Quéau, Y., Mecca, R., Durou, J.D.: Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach. In: Proceedings of the 29$^{\text{th}}$ IEEE Conference on Computer Vision and Pattern Recognition, pp. 4350–4358. Las Vegas, USA (2016) 16, 17, 25
37. Shen, H.L., Cheng, Y.: Calibrating light sources by using a planar mirror. Journal of Electronic Imaging **20**(1) (2011) 6
38. Smith, W., Fang, F.: Height from Photometric Ratio with Model-based Light Source Selection. Computer Vision and Image Understanding **145**, 128–138 (2016) 16
39. 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) 5
40. 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) 6
41. Woodham, R.J.: Photometric Method for Determining Surface Orientation from Multiple Images. Optical Engineering **19**(1), 139–144 (1980) 2
42. 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) 12
43. 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) 5, 6
44. 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) 15
45. 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 14$^{\text{th}}$ European Conference on Computer Vision, pp. 738–752. Amsterdam, The Netherlands (2016) 6, 11