A Non-Convex Variational Approach to Photometric Stereo under Inaccurate Lighting

Yvain QUÉAU, Tao WU
Technical University of Munich, Germany
{yvain.queau, tao.wu}@tum.de

François LAUZE
University of Copenhagen, Denmark
francois@di.ku.dk

Jean-Denis DUROU
University of Toulouse, France
durou@irit.fr

Daniel CREMERS*
Technical University of Munich, Germany
cremers@tum.de

Abstract

This paper tackles the photometric stereo problem in the presence of inaccurate lighting, obtained either by calibration or by an uncalibrated photometric stereo method. Based on a precise modeling of noise and outliers, a robust variational approach is introduced. It explicitly accounts for self-shadows, and enforces robustness to cast-shadows and specularities by resorting to redescending M-estimators. The resulting non-convex model is solved by means of a computationally efficient alternating reweighted least-squares algorithm. Since it implicitly enforces integrability, the new variational approach can refine both the intensities and the directions of the lighting.

1. Introduction

Photometric stereo (PS) is the problem of inferring the shape and, possibly, the reflectance of a surface, from a set of m images taken from a still camera under varying lighting. Assuming Lambertian reflectance, the surface normal and the albedo can be unambiguously determined in each surface point when $m = 3$ [35], but considering $m > 3$ images makes possible their robust estimation in the presence of noise, shadows and specularities (see Figure 1).

Yet, the accuracy of lighting calibration remains a bottleneck for robust methods: the effective intensities and directions of the lighting are never exactly equal to the calibrated ones. Hence, a robust PS method which can automatically refine some initial lighting estimate would be worthwile. However, existing robust methods do not propose this feature, and existing uncalibrated methods lack robustness.

This contribution fills this gap, by introducing a robust variational framework for the joint refinement of shape, albedo and lighting, given an initial lighting estimate which can be obtained by calibration or uncalibrated PS.

*Yvain QUÉAU, Tao WU and Daniel CREMERS were supported by the ERC Consolidator Grant "3D Reloaded".

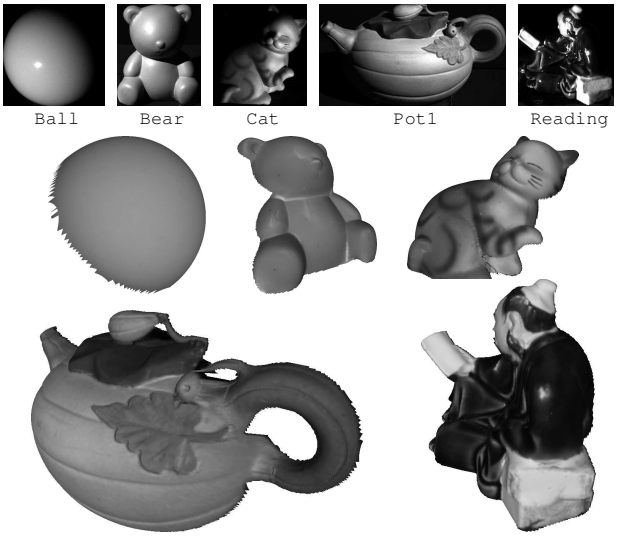


Figure 1. Top: five real-world PS datasets [31], containing self-shadows (all), cast-shadows (all except `Ball`), specular spikes (`Ball` and `Reading`), or broader specular lobes (`Bear`). Bottom: 3D-models estimated by the proposed method, taking as initial lighting the calibration from [31]. Qualitatively similar results are obtained using uncalibrated PS as initialization, see Figure 7.

1.1. Related Work

Robust Pointwise Photometric Stereo methods estimate the normal and the albedo in each pixel. Assuming Lambertian reflectance and neglecting self-shadows, PS can be formulated in each pixel as an overdetermined system of linear equations in the normal and the albedo [35]. The baseline method solves this system in the least-squares sense, which is fast, but justified only in the presence of Gaussian noise.

Yet, specularities can hardly be considered as Gaussian noise. If they consist of wide lobes, they can be explicitly handled by non-Lambertian reflectance models [9, 16, 20, 30]. The present work rather considers specular spikes, hence specularities are viewed as outliers to the Lambertian

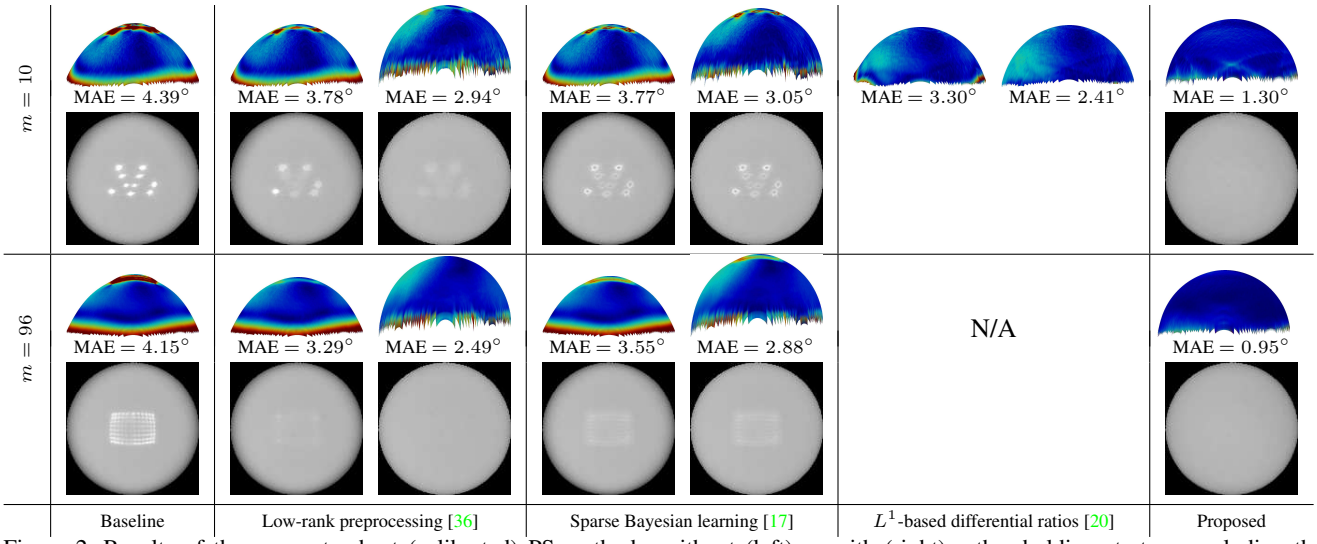


Figure 2. Results of three recent robust (calibrated) PS methods, without (left) or with (right) a thresholding strategy excluding the graylevels below the median of the 10% lowest ones, on the `Ball` dataset with $m = 10$ or $m = 96$ (ratios cannot handle this case) images. In each experiment, the top row shows the 3D-reconstruction, with angular error w.r.t. ground truth normals (red is $> 10°$, blue is $0°$, and MAE is the mean), and the bottom row shows the estimated albedo (not provided by ratios). Existing methods require thresholding to eliminate self-shadows (the proposed one explicitly handles them), remain unsatisfactory in the presence of specularities (the proposed one is more robust to such phenomena), and require either many [17, 36] or few [20] images (the proposed one handles both cases).

model [17, 36]. Cast-shadows constitute another sort of outliers, and existing methods also treat self-shadows as outliers, althouth they can be locally modeled (Equation (1)).

If outliers can be detected, the baseline method can be applied only to "inliers" [3, 33]. Yet, determining "inliers" is not straightforward [7, 34, 37]. Most recent approaches apply the baseline method to modified graylevels, in such a way that they fit a rank-3 approximation [36], or use robust estimators [17]. Yet, as shown in Figure 2, these methods must be coupled with an ad-hoc shadow thresholding strategy [30, 31], and are not very robust to specularities.

The normals estimated by pointwise methods eventually need to be integrated into a depth map [10]. *Differential PS methods* circumvent this issue by directly estimating the depth map. Assuming Lambertian reflectance, neglecting self-shadows and assuming $m = 2$ images, this yields a system of $m = 2$ nonlinear PDEs which can be linearized through image ratios [21]. It has recently been shown how to handle $m \geq 3$ RGB images [28] by variational methods, and how to ensure robustness by using L^1 norm-based methods [20]. However, a naive implementation of the ratios procedure does not provide the albedo, and only few images can be handled since a combinatorial number of equations is obtained. Although these limitations can be overcome by resorting to elaborate algorithmic procedures [32], it would be more natural to consider the nonlinearized system of PDEs. Yet, there exists no method to solve it in a robust and practical manner: existing methods are either restricted to least-squares [13], or require knowing the depth values on the boundary of the domain [27].

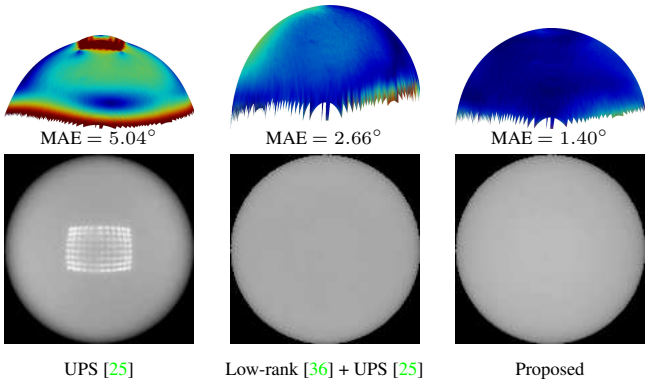


Figure 3. Uncalibrated PS method on the `Ball` dataset with $m = 96$ images. Robustness is improved by using low-rank approximation [36], but post-processing uncalibrated PS by the proposed method yields even more accurate results.

All methods presented so far assume calibrated lighting: *uncalibrated PS (UPS) methods* do not. They resort to matrix factorization to estimate shape, albedo and lighting up to a linear ambiguity [12], which is then reduced by enforcing integrability [38]. Under perspective projection, this yields a unique solution [24]. Under orthographic projection, there is a remaining (GBR) ambiguity [6] which eventually needs to be eliminated [1, 25]. The UPS framework can also be extended to non-directional lighting [4, 26], class-specific recovery of shape [39], and non-Lambertian reflectances [19, 22]. Yet, since UPS methods rely on a sequence of operations, error propagation may happen. Preprocessing the images by low-rank approximation [36] improves robustness, but only partly (Figure 3).

1.2. Contributions

This paper introduces the following innovations:

1. An accurate photometric stereo model, which precises the notions of "albedo", "noise" and "outliers", is introduced in Section 2;

2. Depth, albedo and lighting are jointly estimated within a variational framework relying on redescending M-estimators, where self-shadows are explicitly handled by the model but cast-shadows and specularities are viewed as outliers, cf. Section 3;

3. An efficient alternating reweighted least-squares algorithm is proposed in Section 4, in order to solve the resulting non-convex variational problem.

These novelties altogether yield the first robust photometric stereo method able to refine the directions and the intensities of the lighting.

2. What are "noise" and "outliers" in PS?

It is shown in this section that, if the surface reflectance is dominated by a Lambertian component, and considering m graylevel images I^i, $i \in [1,m]$, obtained under varying directional lighting characterized by vectors $\mathbf{s}^i \in \mathbb{R}^3$, $i \in [1,m]$, then the following **nonlinear image formation model** must be considered:

$$I^i(\mathbf{p}) = \rho(\mathbf{p}) \max\left\{0, \mathbf{s}^i \cdot \mathbf{n}(\mathbf{p})\right\} + \varepsilon^i(\mathbf{p}), \forall i \in [1,m], \quad (1)$$

where: $\rho(\mathbf{p}) > 0$ is the albedo and $\mathbf{n}(\mathbf{p}) \in \mathbb{R}^3$ the surface normal in point $\mathbf{x} \in \mathbb{R}^3$ conjugate to pixel $\mathbf{p} \in \mathbb{R}^2$; the max operator, which encodes self-shadows, is usually neglected in order to linearize the model; and the random variables $\varepsilon^i(\mathbf{p})$ stand for noise and outliers. Besides, these random variables are the sum of four components representing, respectively, noise and quantization, lens aberrations, cast-shadows and specularities:

$$\varepsilon^i(\mathbf{p}) = \varepsilon^i_{\mathrm{N}}(\mathbf{p}) + \varepsilon^i_{\mathrm{A}}(\mathbf{p}) + \varepsilon^i_{\mathrm{CS}}(\mathbf{p}) + \varepsilon^i_{\mathrm{S}}(\mathbf{p}) \quad (2)$$

This result is proven in the following by describing in an accurate manner the photometric phenomena which eventually yield to the formation of the image on the sensor.

If the sensor is not saturated, the graylevel $I^i(\mathbf{p})$ in pixel $\mathbf{p}$ is proportional to the sensor irradiance $e^i(\mathbf{p})$, with proportionality coefficient γ (which is an intrinsic characteristic of the camera) and up to noise and quantization, represented by a random variable $\varepsilon^i_{\mathrm{N}}(\mathbf{p})$:

$$I^i(\mathbf{p}) = \gamma e^i(\mathbf{p}) + \varepsilon^i_{\mathrm{N}}(\mathbf{p}). \quad (3)$$

If the camera is focused on a surface point $\mathbf{x} \in \mathbb{R}^3$ conjugate to pixel $\mathbf{p}$, $e^i(\mathbf{p})$ is almost proportional to the luminance $L^i(\mathbf{x})$ of the surface in $\mathbf{x}$ in the viewing direction characterized by the angle $\alpha(\mathbf{p})$ with the optical axis:

$$e^i(\mathbf{p}) = \beta \cos^4 \alpha(\mathbf{p}) L^i(\mathbf{x}) + \frac{\varepsilon^i_{\mathrm{A}}(\mathbf{p})}{\gamma}, \quad (4)$$

where β depends on the aperture, the exposure, etc., the $\cos^4 \alpha(\mathbf{p})$ factor darkens the image borders, and the random variable $\frac{\varepsilon^i_{\mathrm{A}}(\mathbf{p})}{\gamma}$ stands for lens aberrations (*e.g.*, vignetting).

If the surface is illuminated by a parallel and uniform light beam $\mathbf{s}^i \in \mathbb{R}^3$ (this vector is oriented towards the light source and its norm is equal to the luminous flux density), $L^i(\mathbf{x})$ is proportional to the surface irradiance $E^i(\mathbf{x})$, and the proportionality coefficient is the bidirectional reflectance distribution function (BRDF):

$$L^i(\mathbf{x}) = \mathrm{BRDF}^i(\mathbf{x}) E^i(\mathbf{x}), \quad (5)$$

knowing that:

$$E^i(\mathbf{x}) = b^i(\mathbf{x}) \max\left\{0, \mathbf{s}^i \cdot \mathbf{n}(\mathbf{x})\right\}, \quad (6)$$

where $\mathbf{n}(\mathbf{x})$ is the surface normal in $\mathbf{x}$, the max operator encodes self-shadows, and the binary random variable $b^i(\mathbf{x})$ is worth 0 in the presence of cast-shadows, and 1 otherwise.

A surface is called Lambertian if its luminance $L^i(\mathbf{x})$ is independent from the viewing direction. In this case: $\mathrm{BRDF}^i(\mathbf{x}) = \frac{\rho(\mathbf{x})}{\pi}$, $\forall i \in [1,m]$, where $\rho(\mathbf{x}) \in [0,1]$ is the albedo in point $\mathbf{x}$. In this work, a more general dichromatic BRDF model, accounting for an additive specular component $\rho^i_s(\mathbf{x})$, is assumed:

$$\mathrm{BRDF}^i(\mathbf{x}) = \frac{\rho(\mathbf{x}) + \rho^i_s(\mathbf{x})}{\pi}. \quad (7)$$

Physics-based models can be used for $\rho^i_s(\mathbf{x})$ [9], or empirical models which have been designed for computational feasibility [16, 20, 30], but the proposed method rather follows [17, 36] and treats specularities as outliers.

In the following, the notations $\mathbf{n}(\mathbf{p})$, $b^i(\mathbf{p})$ and $\rho^i_s(\mathbf{p})$ are abusively used for $\mathbf{n}(\mathbf{x})$, $b^i(\mathbf{x})$ and $\rho^i_s(\mathbf{x})$. The following variable is also abusively referred to as "albedo":

$$\rho(\mathbf{p}) = \frac{\gamma \cos^4 \alpha(\mathbf{p}) \rho(\mathbf{x})}{\pi}. \quad (8)$$

Following [8], it is also assumed that radiometric settings of the camera (*e.g.*, aperture and exposure) may vary between the shots. In this case the proportionality coefficient β in (4) must be denoted by β^i, $i \in [1,m]$, and it can be integrated to the intensities of the lighting vectors:

$$\mathbf{s}^i = \beta^i \mathbf{s}^i. \quad (9)$$

Combining Equations (3) to (9) yields the well-known model (1), with the definition (2) of the random variables $\varepsilon^i(\mathbf{p})$, where:

$$\varepsilon^i_{\mathrm{CS}}(\mathbf{p}) = \rho(\mathbf{p}) \max\left\{0, \mathbf{s}^i \cdot \mathbf{n}(\mathbf{p})\right\} \left(b^i(\mathbf{p}) - 1\right), \quad (10)$$

$$\varepsilon^i_{\mathrm{S}}(\mathbf{p}) = \frac{\gamma \cos^4 \alpha(\mathbf{p}) \rho^i_s(\mathbf{p}) \max\left\{0, \mathbf{s}^i \cdot \mathbf{n}(\mathbf{p})\right\} b^i(\mathbf{p})}{\pi}. \quad (11)$$

3. Proposed Variational Model

Solving the set of nonlinear equations (1) is difficult because of the nonlinearity induced by the max operator, and because in (2), the third (cast-shadows) and fourth (specularities) components are impulsive phenomena (outliers). Since the radiometric settings of the camera may vary between the shots, it is necessary to estimate the intensities of the lighting vectors $\mathbf{s}^i$ [8]. The proposed approach goes further and also refines their directions, in order to account for the inaccuracy of calibration or the non-robustness of uncalibrated PS methods, making the problem even more challenging since UPS is inherently ambiguous. To eliminate the arising ambiguities, the differential PS formulation is considered. This guarantees integrability, which is enough to disambiguate UPS under perspective projection [24].

This section shows that the **joint estimation of the shape, the albedo and the lighting** can be recast as the following variational problem:

$$\min_{\substack{\rho:\Omega\to\mathbb{R}\\ z:\Omega\to\mathbb{R}\\ \{\mathbf{s}^i\in\mathbb{R}^3\}_{i\in[1,m]}}} \iint_{\mathbf{p}\in\Omega} \sum_{i=1}^{m} \Phi\left(\frac{\rho(\mathbf{p})}{d_z(\mathbf{p})}\Psi\left(\mathbf{s}^i\cdot\left[\mathbf{J}(\mathbf{p})^\top\begin{bmatrix}\nabla z(\mathbf{p})\\ -1\end{bmatrix}\right]\right)-I^i(\mathbf{p})\right)\mathrm{d}\mathbf{p}, \tag{12}$$

with z the depth map (assumed to be differentiable), $\Omega\subset\mathbb{R}^2$ the reconstruction domain, $\mathbf{J}$ a matrix field depending on the camera's intrinsic parameters (see Equation (17)), $d_z(\mathbf{p})$ a nonlinear normalization term (see Equation (18)), Φ the maximum-likelihood estimator for the distribution of the random variables $\varepsilon^i(\mathbf{p})$ defined in (2), and:

$$\Psi(x)=\max\{0,x\}. \tag{13}$$

To achieve this variational formulation, a 3D-frame $(Oxyz)$ is attached to the camera, where O is its optical center, the z-axis is oriented towards the 3D-scene, and the (xy)-plane is parallel to the image plane (uv), whose area of interest (reconstruction domain) is denoted by $\Omega\subset\mathbb{R}^2$. Assuming perspective projection, the surface is then represented as a mapping associating a point $\mathbf{x}(u,v)\in\mathbb{R}^3$ to each pixel $\mathbf{p}=(u,v)\in\Omega$:

$$\mathbf{x}(u,v)=z(u,v)\mathbf{K}^{-1}\left[u,v,1\right]^\top, \tag{14}$$

with z the depth and $\mathbf{K}$ the (calibrated) intrinsics matrix:

$$\mathbf{K}=\begin{bmatrix}f_u & s & u_0\\ 0 & f_v & v_0\\ 0 & 0 & 1\end{bmatrix}, \tag{15}$$

with (f_u,f_v) the focal length scaled by the aspect ratio, s the skew and (u_0,v_0) the camera's principal point.

The surface normal at $\mathbf{x}(u,v)$ is the unit vector parallel to $\partial_u\mathbf{x}(u,v)\times\partial_v\mathbf{x}(u,v)$, oriented towards the camera:

$$\begin{aligned}\mathbf{n}\colon\ \Omega\ &\to\mathbb{S}^2\subset\mathbb{R}^3\\ (u,v)&\mapsto\mathbf{n}(u,v)=\frac{1}{d_{\tilde z}(u,v)}\mathbf{J}(u,v)^\top\begin{bmatrix}\nabla\tilde z(u,v)\\ -1\end{bmatrix},\end{aligned} \tag{16}$$

where the change of variable $\tilde z=\log z$ is used ($z>0$ by construction), and where:

$$\mathbf{J}(u,v)=\begin{bmatrix}f_u & -s & -(u-u_0)\\ 0 & f_v & -(v-v_0)\\ 0 & 0 & 1\end{bmatrix}, \tag{17}$$

$$d_{\tilde z}(u,v)=\left\|\mathbf{J}(u,v)^\top\left[\nabla\tilde z(u,v)^\top,-1\right]^\top\right\|. \tag{18}$$

The change of variable $\tilde z=\log z$ unifies orthographic and perspective projections [10]: setting $\mathbf{J}(u,v)=\mathbf{K}=\mathbf{K}^{-1}=\mathbf{I}_3$ and $\tilde z=z$ yields the orthographic model. Hence, the tilde notation is neglected in the following.

Using (1) and (16), the new differential PS model appears as the following system of PDEs in $(\rho,z,\{\mathbf{s}^i\}_{i\in[1,m]})$:

$$\begin{aligned}I^i(u,v)=\frac{\rho(u,v)}{d_z(u,v)}\max\left\{0,\mathbf{s}^i\cdot\left[\mathbf{J}(u,v)^\top\begin{bmatrix}\nabla z(u,v)\\ -1\end{bmatrix}\right]\right\}\\ +\varepsilon^i(u,v),\ \forall(u,v)\in\Omega,\ \forall i\in[1,m].\end{aligned} \tag{19}$$

The variational problem (12) eventually arises by solving the system of PDEs (19) in an approximate manner, introducing a robust estimator Φ chosen as the maximum likelihood estimator for the distribution of the random variables $\varepsilon^i(u,v)$. Given the definition (2) of the random variables $\varepsilon^i(\mathbf{p})$, assuming their Gaussianity is not reasonable. Heavy-tail distributions, assigning outliers (third and fourth components in (2)) non-negligible probabilities of occurrence, are better-suited.

Redescending M-estimators are a class of efficient maximum likelihood estimators for such distributions, such that $\Phi'(x)\underset{|x|\to\infty}{\longrightarrow}0$. Examples include Geman-McClure's [11], Welsh's [14] and Tukey's [5] functions, as well as least-powers and Cauchy's estimator, respectively defined as follows:

$$\begin{aligned}\Phi_{\mathrm{GM}}(x)&=\frac{x^2}{\lambda^2+x^2},\quad \Phi_{\mathrm{W}}(x)=\lambda^2\left(1-\exp\left\{-\tfrac{x^2}{\lambda^2}\right\}\right),\\ \Phi_{\mathrm{T}}(x)&=\begin{cases}\lambda^2\left(1-\left(1-\dfrac{x^2}{\lambda^2}\right)^3\right), & |x|\le\lambda,\\ \lambda^2, & |x|>\lambda,\end{cases}\\ \Phi_{\mathrm{L}^p}(x)&=|x|^p,\quad \Phi_{\mathrm{C}}(x)=\lambda^2\log\left(1+\frac{x^2}{\lambda^2}\right),\end{aligned} \tag{20}$$

where p should be set to a value within $]0,1[$ to ensure that Φ_{L^p} is redescending (p is set to 0.7 in the experiments), and $\lambda>0$ is a tunable scale parameter. Following [15], λ is set as follows:

$$\lambda=\delta\,\mathrm{median}_{i,\mathbf{p}}\left\{\left|I^i(\mathbf{p})-\mathrm{median}_{i,\mathbf{p}}\left\{I^i(\mathbf{p})\right\}\right|\right\}, \tag{21}$$

where δ is set to 0.4 for Φ_{GM} and Φ_{W}, 0.9 for Φ_{T}, and 0.15 for Φ_{C} (these values were found experimentally to be effective in all the experiments).

4. Numerical Resolution

The variational problem in (12) represents a highly non-convex, nonlinear optimization problem. The non-convexity arises not only from the denominator d_z but also from the non-convex M-estimator Φ and the coupling of ρ, z and $\{\mathbf{s}^i\}_i$, while the nonlinearity is due to the denominator d_z and to the composition of Φ and Ψ functions.

Existing differential PS methods have mostly focused on the difficulties induced by the denominator d_z. They either eliminate it through image ratios [20, 21, 28] (such approaches do not estimate the albedo and can handle only a small number of images), or explicitly handle it through variational calculus [27] (this approach requires using Dirichlet boundary conditions, hence it can hardly be applied in real-world scenarii) or proximal methods [13] (this approach does not have the previous restrictions, yet it requires very expensive computations). In this work, the nonlinear denominator d_z is neither eliminated, nor explicitly handled: the nonlinearity is simply circumvented by estimating a "**scaled albedo**" $\tilde{\rho}(u,v)$ instead of the actual one $\rho(u,v)$ (which can eventually be deduced from (22)):

$$\tilde{\rho}(u,v) = \frac{\rho(u,v)}{d_z(u,v)}. \tag{22}$$

The proposed numerical resolution for (12) adopts a "discretize-then-optimize" approach *i.e.*, replaces the $\Omega \subset \mathbb{R}^2 \to \mathbb{R}$ functions $\tilde{\rho}$ and z by $\mathbb{R}^n$ vectors $\tilde{\boldsymbol{\rho}}$ and $\mathbf{z}$ (where n is the number of pixels inside Ω) and the gradient operator by forward differences, and then solves the following discrete optimization problem:

$$\min_{\substack{\tilde{\boldsymbol{\rho}}\in\mathbb{R}^n \\ \mathbf{z}\in\mathbb{R}^n \\ \{\mathbf{s}^i\in\mathbb{R}^3\}_{i\in[1,m]}}} F(\tilde{\boldsymbol{\rho}},\mathbf{z},\{\mathbf{s}^i\}_i) = \sum_{j=1}^{n}\sum_{i=1}^{m} \Phi\left(\left(r(\tilde{\boldsymbol{\rho}},\mathbf{z},\mathbf{s}^i)\right)_j^i\right). \tag{23}$$

Here the *residual* $\left(r(\tilde{\boldsymbol{\rho}},\mathbf{z},\mathbf{s}^i)\right)_j^i$ is defined as follows, $\forall i \in [1,m]$, $\forall j \in [1,n]$:

$$\left(r(\tilde{\boldsymbol{\rho}},\mathbf{z},\mathbf{s}^i)\right)_j^i = \tilde{\rho}_j \Psi\left((\mathcal{A}\mathbf{z})_j^i - \tilde{s}_{3,j}^i\right) - I_j^i, \tag{24}$$

where I_j^i is the observed graylevel at pixel $j \in [1,n]$ in image $i \in [1,m]$, and $\mathcal{A} \in \mathcal{L}(\mathbb{R}^n;\mathbb{R}^{m\times n})$ is the linear operator such that:

$$(\mathcal{A}\mathbf{z})_j^i = \tilde{s}_{1,j}^i \mathbf{d}_{u,j}\mathbf{z} + \tilde{s}_{2,j}^i \mathbf{d}_{v,j}\mathbf{z}, \tag{25}$$

with $\mathbf{d}_{u,j}$ and $\mathbf{d}_{v,j}$ the j-th lines of the $n\times n$ matrices containing the finite-difference stencils used for approximating the operators ∂_u and ∂_v, and $[\tilde{s}_{1,j}^i, \tilde{s}_{2,j}^i, \tilde{s}_{3,j}^i]^\top$ the vector field $\tilde{\mathbf{s}}^i$ at pixel j defined by:

$$\tilde{\mathbf{s}}^i(u,v) = \mathbf{J}(u,v)\mathbf{s}^i. \tag{26}$$

In the following, a tailored **alternating reweighted least-squares** (ARLS) method for computing a stationary point (often a local minimizer) of Problem (23) is presented. Suggested by its name, ARLS handles the coupling between variables by using an alternating optimization strategy, *i.e.* minimizes F over $\tilde{\boldsymbol{\rho}}$, $\mathbf{z}$ and $\{\mathbf{s}^i\}_i$ alternatively. Given $(\tilde{\boldsymbol{\rho}}^{(k)}, \mathbf{z}^{(k)}, \{\mathbf{s}_i^{i,(k)}\})$ at iteration k, this corresponds to solving the first-order optimality conditions

$$\nabla_{\tilde{\boldsymbol{\rho}}} F(\tilde{\boldsymbol{\rho}}, \mathbf{z}^{(k)}, \{\mathbf{s}^{i,(k)}\}_i) = 0, \tag{27}$$

$$\nabla_{\mathbf{z}} F(\tilde{\boldsymbol{\rho}}^{(k+1)}, \mathbf{z}, \{\mathbf{s}^{i,(k)}\}_i) = 0, \tag{28}$$

$$\nabla_{\mathbf{s}^i} F(\tilde{\boldsymbol{\rho}}^{(k+1)}, \mathbf{z}^{(k+1)}, \{\mathbf{s}^i\}_i) = 0,\ \forall i \in [1,m], \tag{29}$$

in an iteratively reweighted least-squares (IRLS) manner [14]. More precisely, the reweighting is utilized on the following quantities, $\forall i \in [1,m],\ \forall j \in [1,n]$:

$$\left(w(\tilde{\boldsymbol{\rho}},\mathbf{z},\{\mathbf{s}^i\}_i)\right)_j^i = \frac{\Phi'\left(\left(r(\tilde{\boldsymbol{\rho}},\mathbf{z},\{\mathbf{s}^i\}_i)\right)_j^i\right)}{\left(r(\tilde{\boldsymbol{\rho}},\mathbf{z},\{\mathbf{s}^i\}_i)\right)_j^i}, \tag{30}$$

$$\left(\chi(\mathbf{z},\mathbf{s}^i)\right)_j^i = \begin{cases} 0 & \text{if } (\mathcal{A}\mathbf{z})_j^i - \tilde{s}_{3,j}^i \le 0, \\ 1 & \text{if } (\mathcal{A}\mathbf{z})_j^i - \tilde{s}_{3,j}^i > 0, \end{cases} \tag{31}$$

which brings (27)–(29) to the fully linearized systems as follows:

$$\sum_{i=1}^{m}\left(w^{(k)}\right)_j^i\left[\tilde{\rho}_j\left(\Psi^{(k)}\right)_j^i - I_j^i\right]\left(\Psi^{(k)}\right)_j^i = 0,\ \forall j \in [1,n], \tag{32}$$

$$\sum_{j=1}^{n}\sum_{i=1}^{m}\left(\bar{w}^{(k)}\right)_j^i\left[\tilde{\rho}_j^{(k+1)}\left(\bar{\chi}^{(k)}\right)_j^i\left(\left(\mathcal{A}^{(k)}\mathbf{z}\right)_j^i - \tilde{s}_{3,j}^{i,(k)}\right) - I_j^i\right] \tilde{\rho}_j^{(k+1)}\mathcal{A}_{j,l}^{i,(k)}\left(\bar{\chi}^{(k)}\right)_j^i = 0,\ \forall l \in [1,n], \tag{33}$$

$$\left(\sum_{j=1}^{n}\left(\tilde{w}^{(k)}\right)_j^i\left(\tilde{\chi}^{(k)}\right)_j^i{}^2\left(\tilde{\rho}_j^{(k+1)}\right)^2\left[\overline{\mathbf{n}}^{(k+1)}\right]_j\left[\overline{\mathbf{n}}^{(k+1)}\right]_j^\top\right)\mathbf{s}^i = \sum_{j=1}^{n}\left(\tilde{w}^{(k)}\right)_j^i\left(\tilde{\chi}^{(k)}\right)_j^i\tilde{\rho}_j^{(k+1)}I_j^i\left[\overline{\mathbf{n}}^{(k+1)}\right]_j, \forall i \in [1,m]. \tag{34}$$

Here $\left[\overline{\mathbf{n}}^{(k+1)}\right]_j \in \mathbb{R}^3$ is the linear part (*i.e.*, not normalized by the d_z denominator) of the normal vector at pixel j, computed from $\mathbf{z}^{(k+1)}$ according to (16). The weight $\left(w^{(k)}\right)_j^i = \left(w(\tilde{\boldsymbol{\rho}}^{(k)},\mathbf{z}^{(k)},\{\mathbf{s}^{i,(k)}\}_i)\right)_j^i$ is obtained by (30) using the values of the albedo, the depth and the lighting at iterations k, $\left(\bar{w}^{(k)}\right)_j^i$ is its update using the new albedo value, and $\left(\tilde{w}^{(k)}\right)_j^i$ uses the new albedo and depth values. Similar conventions are adopted for χ and Ψ.

It is worth emphasizing that the solution of each linear system above can be interpreted as the solution of a reweighted least-squares problem. For instance, the solution $\mathbf{z}^{(k+1)}$ obtained by (33) solves the reweighted least-

squares problem

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} \sum_{j=1}^{n}\sum_{i=1}^{m} \left(\bar{w}^{(k)}\right)_j^i \left[\tilde{\rho}_j^{(k+1)} \left(\tilde{\chi}^{(k)}\right)_j^i \left(\left(\mathcal{A}^{(k)}\mathbf{z}\right)_j^i - \tilde{s}_{3,j}^{i,(k)}\right) - I_j^i\right]^2, \quad (35)$$

and thus approximates the solution of the nonlinear least-squares problem:

$$\min_{\boldsymbol{z}\in\mathbb{R}^n} \sum_{j=1}^{n}\sum_{i=1}^{m} \left(\bar{w}^{(k)}\right)_j^i \left[\left(r(\tilde{\boldsymbol{\rho}}^{(k+1)}, \mathbf{z}, \{\mathbf{s}_i^{i,(k)}\})\right)_j^i\right]^2, \quad (36)$$

which arises when only the w-term is reweighted.

The linear systems (32) and (34) admit closed-form solutions. In our numerical experiments, the sparse (most of the terms in the $\mathcal{A}$-matrix are null: the summation over j in (33) is actually limited to the neighbors of pixel l) linear system (33) arising from the least-squares problem (35) was solved using preconditioned conjugate gradient [18].

5. Experimental Validation

ARLS iterates approximate resolutions of the optimality conditions (27), (28) and (29) until convergence. The convergence analysis could be conducted in a similar manner as in [29]. In this section we rather validate the practical performances of the algorithm. To select a stopping criterion, we consider in Figure 4 the `Pot1` dataset with $m = 48$ images, and using the estimator Φ_{C}. Based on this experiment, we choose as stopping criterion a threshold on the relative difference between the values of the energy $F(\tilde{\boldsymbol{\rho}}, \mathbf{z}, \{\mathbf{s}^i\}_{i\in[1,m]})$ at two successive iterations: this threshold is set hereafter to 10^{-4} (blue line in Figure 4). Yet, given the non-convexity of the problem, choosing an appropriate initialization is important. If calibrated lighting is available, the baseline method is a good candidate for shape and albedo initialization. Otherwise, the UPS method from [25] is recommended in [31], and it also provides an initial lighting estimate. In both cases, the weights are initialized using their definition (30).

The proposed linearization strategies are then evaluated in Figure 5 (without lighting refinement for fair comparison with [13]). The estimation of a "scaled" albedo containing the nonlinear denominator (Equation (22)) is compared to a recent approach estimating the "actual" albedo while explicitly handling the nonlinearity [13]. As can be seen in the first graph in Figure 5, which considers the `Cat` dataset with $m = 5$ images under the same assumptions as in [13] ($\Phi = \Phi_{\mathrm{L}^2}$, orthographic projection, $\Psi(x) = x$), the proposed strategy dramatically reduces the cost of each iteration. The second graph in Figure 5, which considers the `Ball` dataset with $m = 5$ images, shows the efficiency of the reweighting for linearizing the self-shadows operator. Each reweighted iteration decreases almost as much the energy as its non-linearized counterpart, yet it is a lot faster to compute.

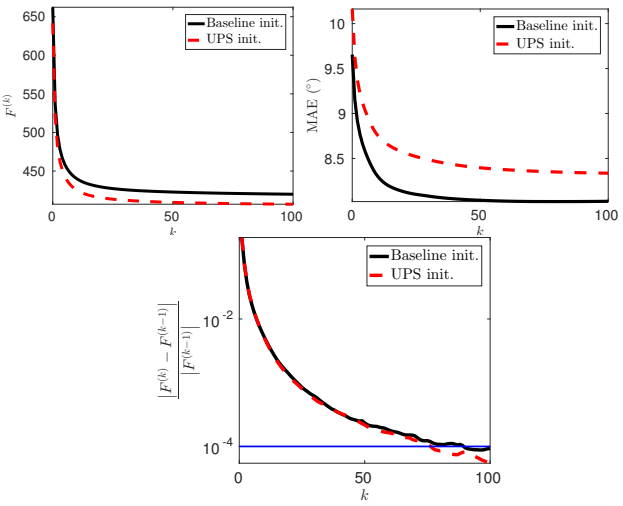
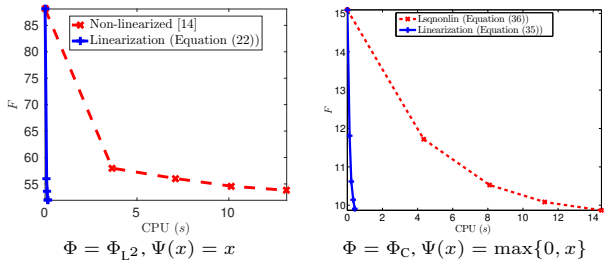


Figure 4. Evolution of the energy $F^{(k)} = F(\tilde{\boldsymbol{\rho}}^{(k)}, \mathbf{z}^{(k)}, \{\mathbf{s}^{i,(k)}\}_{i\in[1,m]})$, of the mean angular error w.r.t. ground truth normals and of the relative difference between two successive energy values, as functions of the iteration number k. Since the problem is non-convex, the effect of initialization is significant: choosing the (calibrated) baseline method or an uncalibrated PS method ([25] in this experiment) yields different energy values, and a lower final energy does not necessarily mean a more accurate shape.



$\Phi = \Phi_{\mathrm{L}^2}, \Psi(x) = x$ $\quad$ $\Phi = \Phi_{\mathrm{C}}, \Psi(x) = \max\{0, x\}$

Figure 5. Evaluation of the proposed linearization strategies (the graphs show the energy value attained after the first five iterations, and the time required to complete these iterations). Left: evaluation of the elimination of the nonlinearity (18) by estimation of a scaled albedo (22), in comparison with a recent method handling explicitly this nonlinearity by a proximal algorithm. Right: evaluation of the linearization of the self-shadows operator yielding the linear least-squares problem (35) instead of the nonlinear least-squares one (36), which requires resorting to a baseline nonlinear least-squares solver such as Matlab's "lsqnonlin". These experiments show the efficiency of the proposed linearization strategies.

The full approach (with lighting refinement) is now evaluated against state-of-the-art robust calibrated methods and uncalibrated ones, on the five sets of images shown in Figure 1. For fair comparison between pointwise PS methods (which provide normals, but not depth) and differential ones (which provide depth, but not normals), the normals estimated by pointwise methods were integrated into a depth map (using perspective camera, least-squares and the same

	Ball			Bear			Cat			Pot1			Reading		
	Mean (°)	Median (°)	CPU (s)	Mean (°)	Median (°)	CPU (s)	Mean (°)	Median (°)	CPU (s)	Mean (°)	Median (°)	CPU (s)	Mean (°)	Median (°)	CPU (s)
Baseline	4.15	2.55	0.03	8.58	6.44	0.04	9.57	7.33	0.03	9.62	7.08	0.03	21.05	15.77	0.03
LR [36]	2.49	2.50	106.43	6.85	5.38	207.49	8.68	6.91	203.19	8.29	6.36	247.15	16.64	12.80	167.12
SBL [17]	2.88	2.75	625.34	7.46	6.17	1106.17	8.82	7.05	959.66	8.71	6.76	1224.33	19.27	13.96	573.92
Ratios [20][1]	2.40	2.51	610.61	6.09	4.90	749.87	7.16	5.95	1099.89	7.95	5.98	1042.72	15.32	8.53	1912.94
LR [36] + [13]	2.22	2.31	2693.30	6.60	5.25	2535.96	7.18	6.06	17252.50	7.99	6.08	12962.26	16.80	12.36	4387.12
Φ_{GM}	1.09	0.88	67.16	6.05	4.54	105.44	6.78	5.18	319.32	8.02	5.63	236.43	13.35	7.40	239.74
Φ_{W}	1.18	0.99	64.74	6.34	4.83	104.51	**6.73**	**5.07**	373.07	8.06	5.66	240.76	13.92	7.59	284.80
Φ_{T}	1.72	1.59	**18.98**	6.21	4.77	**62.01**	7.03	5.38	**62.40**	8.19	5.96	**82.87**	14.71	8.78	**30.13**
Φ_{L^p}	1.48	1.30	69.39	6.09	4.61	93.10	6.77	5.37	294.35	**7.89**	**5.48**	261.55	15.37	9.89	88.02
Φ_{C}	**0.95**	**0.76**	81.07	**5.95**	**4.47**	114.08	6.78	5.28	336.63	7.98	5.54	265.46	**13.01**	**7.33**	184.10

Table 1. Quantitative results obtained on the five datasets presented in Figure 1, using $m = 96$ images and the calibration provided in [31] as initial lighting estimate (the runtime, the mean and median angular errors between the estimated normals and the ground truth ones are shown). The proposed approach (with lighting refinement) outperforms state-of-the-art robust calibrated methods in terms of accuracy, and is often faster. The 3D-models (3D-reconstruction and estimated albedo) corresponding to $\Phi = \Phi_{\text{C}}$ are shown in Figure 1.

Camera	Estimator	Self-shadows	Lighting refinement	Mean (°)	Median (°)	CPU (s)
Orthographic	Φ_{L^2}	$\Psi(x) = x$	No	22.69	16.75	1.05
Perspective	Φ_{L^2}	$\Psi(x) = x$	No	22.49	16.46	1.15
Perspective	Φ_{C}	$\Psi(x) = x$	No	13.71	7.90	13.28
Perspective	Φ_{C}	$\Psi(x) = \max\{0, x\}$	No	13.69	7.68	14.23
Perspective	Φ_{C}	$\Psi(x) = \max\{0, x\}$	Yes	13.51	7.47	28.71

Table 2. 3D-reconstruction error and computation time (`Reading` dataset with $m = 20$ images, using the calibrated lighting as initial estimate), for different combinations of the new features introduced in the proposed method. Using robust estimation is key for accuracy, but all the other features also contribute to improving the results.

finite-differences as in the proposed method), and then recomputed by finite-differences. On the other hand, normals were approximated from the depth maps provided by differential methods by finite-differences. Computation times were evaluated using Matlab codes and a Xeon processor of 3.50 GHz with 32 GB of RAM.

First, the proposed approach is considered while using the calibration provided in [31] as initial lighting estimate, and compared with three recent robust calibrated PS methods: low-rank approximation [36] (LR), sparse Bayesian learning [17] (SBL), and L^1-based differential ratios [20]. For completeness, the results of the recent differential approach [13] are also shown in Table 1: since this method is not robust, it is coupled with the LR method. Figure 2 shows that the proposed method is the first one which provides accurate results with both many ($m = 96$) or few ($m = 10$) images, and without resorting to any ad-hoc outlier rejection method based on thresholding. Since pointwise methods require many images, the quantitative evaluation over the whole dataset is then performed with $m = 96$ images, see Table 1 (thresholding was used for competing methods). Existing differential methods systematically outperform pointwise ones, but the proposed method is even more accurate, and significantly faster. In comparison with [20], the proposed method also provides the albedo (which is eliminated using ratios) and handles an arbitrary large number m of images (32 GB of memory was not enough to apply the differential ratios method with $m > 20$). Among the robust M-estimators defined in Equation (20), Φ_{C} is the one attaining maximum breakdown point [23], hence it should be the most robust. This is confirmed by the experiments, although if speed is an issue then Tukey's estimator Φ_{T} could be considered.

Table 2 shows the respective influence of each new feature introduced in this work. Comparing the third row of Table 2 with Table 1 shows the superiority of redescending M-estimators over existing robust estimators, as we already overcome the state-of-the-art with five times less images. The forth row proves that self-shadows should be explicitly handled instead of being treated as outliers, and the last row shows that a non-negligible improvement can be expected by refining both the intensities and the directions of the lighting, to account for inaccurate calibration. An additional experiment, combining all these features on the `Gourd` dataset from [2], is conducted in Figure 6, and shows that reasonable results can be expected with as few as $m = 10$ images.

Eventually, Table 3 evaluates the ability of the proposed method to improve UPS results. Two recent UPS methods are compared using $m = 96$ images: the minimum entropy (ME) approach [1] and the diffuse maxima (DM) method [25]. By applying the proposed method as a post-processing to the DM method, the state-of-the-art robust UPS strategy consisting in coupling UPS with low-rank preprocessing [36] is significantly outperformed (the running times of the proposed method and of this strategy are comparable). Using uncalibrated PS as initial estimate yields less accurate results than using calibration (compare Tables 1 and 3), yet their results appear qualitatively similar (compare Figures 1 and 7): the proposed approach represents an important step towards reliable uncalibrated PS.

[1] For evaluating the differential ratios approach [20], we used only $m = 20$ images, which is the maximum number of images we could handle with 32 GB of memory.

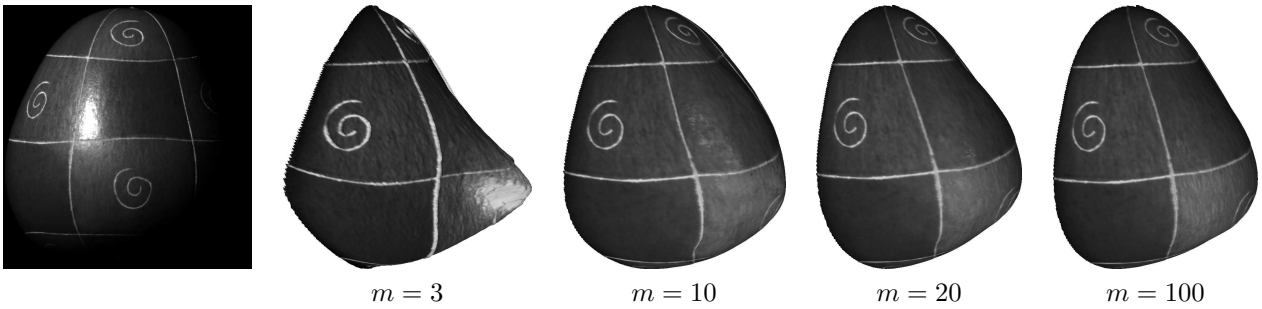


Figure 6. Results on the Gourd dataset from [2] (the left image shows one of the input images), while increasing the number of images and using the calibrated lighting as initial estimate. The reconstructed shape is much distorted using $m = 3$ images (robustness cannot be enforced), yet reasonable results are obtained with as few as $m = 10$ images.

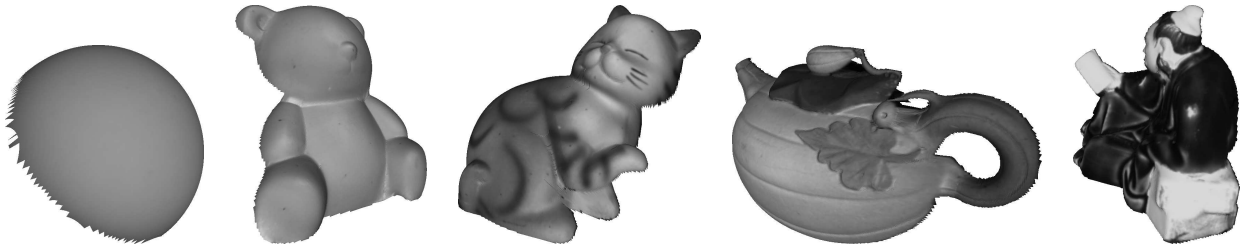


Figure 7. 3D-models estimated by the proposed uncalibrated PS method (using $\Phi = \Phi_C$ and the DM method [25] as initialization), which are very similar to those obtained using the calibrated lighting as initial estimate (see Figure 1).

	Ball	Bear	Cat	Pot1	Reading
ME [1]	6.56	15.29	19.85	16.49	82.37
DM [25]	5.04	9.20	10.62	10.27	24.49
LR [36] + DM [25]	2.66	7.69	8.89	8.94	42.23
Proposed	**1.40**	**6.66**	**7.59**	**8.46**	**20.16**

Table 3. Mean angular error (in degrees) attained by two (non-robust) uncalibrated methods [1, 25], by [25] applied to images preprocessed by low-rank approximation [36], and by the proposed uncalibrated PS approach, using the lighting estimated by the DM method [25] as initial guess for our method ($\Phi = \Phi_C$). Refining the lighting by using the proposed approach yields a significant improvement in the accuracy of the 3D-reconstructions. The 3D-models estimated by this method are shown in Figure 7.

6. Conclusion

A new variational approach to robust photometric stereo has been introduced. It explicitly handles self-shadows, and relies on redescending M-estimators for handling cast-shadows and specularities. Put together, these novelties remove the need for tedious manual identification of shadows or highlights, and yield a fully automatic robust PS method.

It has also been shown that the nonlinearities of the resulting variational model can be handled without resorting to linearization by ratios or to computationally expensive methods, by estimating a scaled albedo and appropriately reweighting the nonlinear factors appearing in the optimality conditions.

Eventually, the proposed approach is the first robust PS method which can refine both the intensities and the directions of the lighting. This important step towards reducing the importance of calibration is achieved by implicitly enforcing integrability through a differential approach.

Differential methods for solving uncalibrated (or inaccurately calibrated) PS had not been explored so far because PDE-based methods have not become a standard in PS yet. This is probably due to the computational complexity of previous methods. The proposed method, which is much more efficient, could open the door to renewed research in PS: not only the framework introduced in this paper yields the most accurate robust PS method presented so far, but it is also highly flexible. Hence, it may be used in future research as basis for other challenging PS problems *e.g.*, PS with physics-based specular reflectance models. Indeed, we treated specularities as outliers to the Lambertian model, but specular reflections (and cast-shadows) also provide clues for estimating the shape: they should thus ultimately be considered in the PS model.

References

[1] N. G. Alldrin, S. P. Mallick, and D. J. Kriegman. Resolving the Generalized Bas-Relief Ambiguity by Entropy Minimization. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, USA, 2007. 2, 7, 8

[2] N. G. Alldrin, T. Zickler, and D. J. Kriegman. Photometric Stereo With Non-Parametric and Spatially-Varying Reflectance. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, USA, 2008. 7, 8

[3] S. Barsky and M. Petrou. 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. 2

[4] R. Basri, D. W. Jacobs, and I. Kemelmacher. Photometric Stereo with General, Unknown Lighting. *International Journal of Computer Vision*, 72(3):239–257, 2007. 2

[5] A. E. Beaton and J. W. Tukey. The fitting of power series, meaning polynomials, illustrated on band-spectroscopic data. *Technometrics*, 16(2):147–185, 1974. 4

[6] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The Bas-Relief Ambiguity. *International Journal of Computer Vision*, 35(1):33–44, 1999. 2

[7] M. Chandraker, S. Agarwal, and D. J. Kriegman. ShadowCuts: Photometric Stereo with Shadows. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, USA, 2007. 2

[8] D. Cho, Y. Matsushita, Y.-W. Tai, and I.-S. Kweon. Photometric Stereo Under Non-uniform Light Intensities and Exposures. In *Proceedings of the European Conference on Computer Vision (ECCV)*, volume 9906 of *Lecture Notes in Computer Science*, pages 170–186. Amsterdam, The Netherlands, 2016. 3, 4

[9] H. J. Chung and J. Jia. Efficient photometric stereo on glossy surfaces with wide specular lobes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, USA, 2008. 1, 3

[10] J.-D. Durou, J.-F. Aujol, and F. Courteille. Integrating the normal field of a surface in the presence of discontinuities. In *Proceedings of Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*, volume 5681 of *Lecture Notes in Computer Science*, pages 261–273. Bonn, Germany, 2009. 2, 4

[11] S. Geman and D. E. McClure. Statistical methods for tomographic image reconstruction. *Bulletin of the International Statistical Institute*, LII(4):5–21, 1987. 4

[12] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America A*, 11(11):3079–3089, 1994. 2

[13] L. Hoeltgen, Y. Quéau, M. Breuss, and G. Radow. Optimised photometric stereo via non-convex variational minimisation. In *Proceedings of the British Machine Vision Conference (BMVC)*, York, UK, 2016. 2, 5, 6, 7

[14] P. W. Holland and R. E. Welsch. Robust regression using iteratively reweighted least-squares. *Communications in Statistics –Theory and Methods*, 6(9):813–827, 1977. 4, 5

[15] J. H. Huber. *Robust Statistics*. Wiley, 1981. 4

[16] S. Ikehata and K. Aizawa. Photometric Stereo Using Constrained Bivariate Regression for General Isotropic Surfaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2187–2194, Columbus, USA, 2014. 1, 3

[17] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Photometric Stereo Using Sparse Bayesian Regression for General Diffuse Surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(9):1816–1831, 2014. 2, 3, 7

[18] I. Koutis, G. L. Miller, and R. Peng. A Nearly-m log n Time Solver for SDD Linear Systems. In *IEEE Annual Symposium on Foundations of Computer Science (FOCS)*, pages 590–598, Palm Springs, USA, 2011. 6

[19] F. Lu, I. Sato, and Y. Sato. Uncalibrated photometric stereo based on elevation angle recovery from BRDF symmetry of isotropic materials. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 168–176, Boston, USA, 2015. 2

[20] R. Mecca, Y. Quéau, F. Logothetis, and R. Cipolla. A Single Lobe Photometric Stereo Approach for Heterogeneous Material. *SIAM Journal on Imaging Sciences*, 2016. 1, 2, 3, 5, 7

[21] R. Mecca, E. Rodolà, and D. Cremers. Realistic photometric stereo using partial differential irradiance equation ratios. *Computers & Graphics*, 51:8–16, 2015. 2, 5

[22] K. Midorikawa, T. Yamasaki, and K. Aizawa. Uncalibrated Photometric Stereo by Stepwise Optimization Using Principal Components of Isotropic BRDFs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4350–4358, Las Vegas, USA, 2016. 2

[23] I. Mizera and C. H. Müller. Breakdown points of Cauchy regression-scale estimators. *Statistics & Probability Letters*, 57(1):79–89, 2002. 7

[24] T. Papadhimitri and P. Favaro. A New Perspective on Uncalibrated Photometric Stereo. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1474–1481, Portland, USA, 2013. 2, 4

[25] T. Papadhimitri and P. Favaro. A Closed-Form, Consistent and Robust Solution to Uncalibrated Photometric Stereo Via Local Diffuse Reflectance Maxima. *International Journal of Computer Vision*, 107(2):139–154, 2014. 2, 6, 7, 8

[26] T. Papadhimitri and P. Favaro. Uncalibrated Near-Light Photometric Stereo. In *Proceedings of the British Machine Vision Conference (BMVC)*, Nottingham, UK, 2014. 2

[27] Y. Quéau, F. Lauze, and J.-D. Durou. A L1-TV Algorithm for Robust Perspective Photometric Stereo with Spatially-Varying Lightings. In *Scale Space and Variational Methods in Computer Vision (SSVM)*, volume 9087 of *Lecture Notes in Computer Science*, pages 498–510, Lège Cap-Ferret, France, 2015. 2, 5

[28] Y. Quéau, R. Mecca, and J.-D. Durou. Unbiased photometric stereo for colored surfaces: A variational approach. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3707–3716, Las Vegas, USA, 2016. 2, 5

[29] Y. Quéau, T. Wu, and D. Cremers. Semi-Calibrated Near-Light Photometric Stereo. In *Scale Space and Variational Methods in Computer Vision (SSVM)*, Kolding, Denmark, 2017. 6

[30] B. Shi, Y. Matsushita, and K. Ikeuchi. A biquadratic reflectance model for radiometric image analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 230–237, Providence, USA, 2012. 1, 2, 3

[31] B. Shi, Z. Wu, Z. Mo, D. Duan, S.-K. Yeung, and P. Tan. A Benchmark Dataset and Evaluation for Non-Lambertian and Uncalibrated Photometric Stereo. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4359–4368, Las Vegas, USA, 2016. 1, 2, 6, 7

[32] W. Smith and F. Fang. Height from Photometric Ratio with Model-based Light Source Selection. *Computer Vision and Image Understanding*, 145:128–138, 2016. 2

[33] F. Solomon and K. Ikeuchi. Extracting the shape and roughness of specular lobe objects using four light photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4):449–454, 1996. 2

[34] F. Verbiest and L. Van Gool. Photometric stereo with coherent outlier handling and confidence estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, USA, 2008. 2

[35] R. J. Woodham. Photometric Method For Determining Surface Orientation From Multiple Images. *Optical Engineering*, 19(1):134–144, 1980. 1

[36] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Proceedings of the Asian Conference on Computer Vision (ACCV)*, volume 6494 of *Lecture Notes in Computer Science*, pages 703–717. Queenstown, New-Zealand, 2010. 2, 3, 7, 8

[37] T.-P. Wu and C.-K. Tang. Photometric stereo via expectation maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):546–560, 2010. 2

[38] A. L. Yuille and D. Snow. Shape and albedo from multiple images using integrability. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 158–164, San Juan, USA, 1997. 2

[39] S. K. Zhou, G. Aggarwal, R. Chellappa, and D. W. Jacobs. Appearance characterization of linear lambertian objects, generalized photometric stereo, and illumination-invariant face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2):230–245, 2007. 2