HAL
open science

Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach

Yvain Quéau, Roberto Mecca, Jean-Denis Durou

To cite this version:

Yvain Quéau, Roberto Mecca, Jean-Denis Durou. Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2016), Jun 2016, Las Vegas, NV, United States. pp. 4359-4368. hal-01534802

HAL Id: hal-01534802

https://hal.science/hal-01534802

Submitted on 8 Jun 2017

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.

Open Archive TOULOUSE Archive Ouverte (OATAO)
OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 16930

The contribution was presented at CVPR 2016 :
http://cvpr2016.thecvf.com/

To cite this version : Quéau, Yvain and Mecca, Roberto and Durou, Jean-Denis *Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach.* (2016) In: IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2016), 27 June 2016 - 30 June 2016 (Las Vegas, NV, United States).

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Unbiased Photometric Stereo for Colored Surfaces: A Variational Approach

Yvain Quéau
IRIT, Université de Toulouse
France
yvain.queau@enseeiht.fr

Roberto Mecca*
Department of Engineering,
University of Cambridge, UK
roberto.mecca@eng.cam.ac.uk

Jean-Denis Durou
IRIT, Université de Toulouse
France
durou@irit.fr

Abstract

3D shape recovery using photometric stereo (PS) gained increasing attention in the computer vision community in the last three decades due to its ability to recover the thinnest geometric structures. Yet, the reliabiliy of PS for color images is difficult to guarantee, because existing methods are usually formulated as the sequential estimation of the colored albedos, the normals and the depth. Hence, the overall reliability depends on that of each subtask. In this work we propose a new formulation of color photometric stereo, based on image ratios, that makes the technique independent from the albedos. This allows the unbiased 3D-reconstruction of colored surfaces in a single step, by solving a system of linear PDEs using a variational approach.

1. Introduction

The photometric stereo (PS) technique [54] consists in inferring the shape (and, optionally, the reflectance) of a surface from a set of m images obtained from the same point of view, while changing the illumination. It is a classical computer vision task which has caught the attention of many researchers, one of the reasons being the simplicity of its formulation.

Achieving 3D-reconstruction by PS requires inverting a known photometric model. Formulating this inverse problem brings out to play important factors influencing the actual physics involved in the formation of digital images. These factors are the surface shape, its material, the lighting and the camera. To provide high-quality shape reconstruction, both latter factors are usually calibrated [55], hence PS is mainly a technique for laboratory purposes which is used in very controlled environment [51].

In most of works on PS, the surface reflectance is supposed to be Lambertian. The m lightings are usually modelled by known vectors $\mathbf{s}^i \in \mathbb{R}^3$, $i \in [1, m]$, considering the directional model. It is also frequently assumed that the camera captures graylevels. Under these assumptions, the PS problem can be formulated as the inversion of the following linear photometric model:

$$I^i(u,v) = \underbrace{\rho(u,v)\overline{\mathbf{n}}(u,v)}_{\mathbf{m}(u,v)} \cdot \mathbf{s}^i, \; i \in [1, m] \qquad (1)$$

where $I^i(u,v)$ is the graylevel, $\rho(u,v)$ the albedo, and $\overline{\mathbf{n}}(u,v)$ the unit outward normal to the surface. As long as $m \geq 3$ non-coplanar lightings are used, system (1) is a full-rank linear system in $\mathbf{m}(u,v)$, which can be estimated before deducing the albedo and the normal.

Current research on PS includes relaxing the assumptions above on the Lambertian reflectance [46], the directionality of the lightings [57], and the need for lighting calibration [38], as well as robustness enforcement [21], depth inference from the estimated normals [16], outdoor photometric stereo [23], and scattering medium modelling [33]. In this work, we propose an unbiased variational formulation of PS for *colored surfaces*, while also considering extensions to *perspective projection* and *non-directional lightings*, and avoiding the need for *normal integration*. An example of result obtained using this new approach is presented in Figure 1.

Overview of the contribution. With the aim to provide an unbiased solution to the PS problem, we derive a new framework exploiting all the photometric clues available in color images. In particular, we move beyond the classical approach to PS by suggesting a simple formulation that is independent from surface color and provides us with a single step procedure for approximating the depicted surface. After showing the limits of existing color PS methods in Section 2, we show in Section 3 how to use image ratios to eliminate the dependency on the albedo, yielding a single system of linear PDEs having the depth as unknown. We suggest in Section 4 an efficient variational method for the resolution of these PDEs. Eventually, we show in Section 5 the ductility of the new formulation, extending our framework to realistic acquisition setups such as perspective projection and non-directional lightings.

*Marie Curie fellow of the Istituto Nazionale di Alta Matematica, Italy.

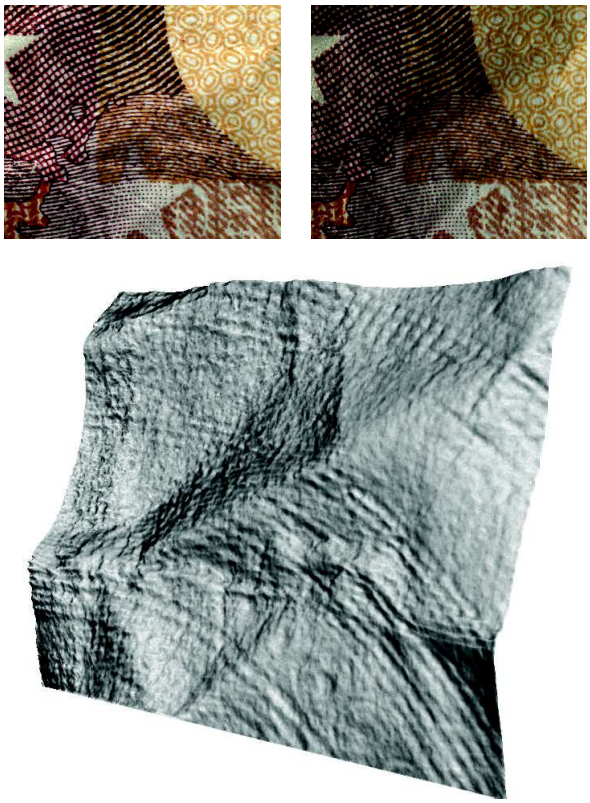


Figure 1. 3D-reconstruction of a part of a 10 euros banknote ($\approx 0.5\ cm$ wide) by variational color photometric stereo. Top: Two out of $m = 5$ RGB images of size 550×650. Bottom: 3D-reconstruction result. Our PDE-based approach to photometric stereo is able to recover the very thin surrows (around ten micrometers wide) which are part of the "unfalsifiable" structure of the banknote, contrarily to lower depth frequencies which only show that the banknote has been creased.

2. Color Photometric Stereo

The photometric model (1) describes the formation of graylevels, meaning that RGB color images provided by common cameras have to be converted into a single color level. Mallick *et al.* [28] showed that such conversion yielded by an appropriate rotation of the color space could reduce the amount of specular effects. Nevertheless, converting RGB images to single channel images consists in loosing information. To avoid this drawback, we deal with RGB images as they are, enforcing the shape reconstruction from the colored shadings. In this case, we consider Eq. (1) as wavelength-dependent, where both the reflectance parameter ρ and the lightings $\mathbf{s}^i$ may be color-dependent. For Lambertian materials, a physically plausible relation for the measured color level $\star \in \{R, G, B\}$ reads:

$$I^i_\star = \int_0^{+\infty} c_\star(\lambda)\rho(\lambda)\overline{\mathbf{n}} \cdot \mathbf{s}^i(\lambda)\,\mathrm{d}\lambda \tag{2}$$

where λ is the wavelength, $c_\star$ describes the camera response depending on the color channel $\star$, and where we omit space dependencies for better readability. Although the normal vector $\overline{\mathbf{n}}$ can obviously be taken out of the integral, the formulation (2) is too difficult to handle, hence several particular cases were suggested in the literature. They are briefly reviewed in the next paragraphs.

White surface lit by colored sources. One of the most well-known applications of color photometric stereo is the single-shot approach involving simultaneously a red, a green and a blue lightings. This idea, which dates back to the 90s [25], was popularized by real-time 3D-reconstructions of deforming surfaces [9]. Assuming $m = 1$ and the surface is white ($\rho(\lambda) = \rho$), Eq. (2) simplifies:

$$I_\star = \rho\,\overline{\mathbf{n}} \cdot \underbrace{\int_0^{+\infty} c_\star(\lambda)\mathbf{s}(\lambda)\,\mathrm{d}\lambda}_{\triangleq \mathbf{s}_\star} \tag{3}$$

where $\star \in \{R, G, B\}$, and $\mathbf{s}_\star$ combines the camera response and the lighting. Then, the PS problem simplifies to the classical model (1) with three images. To relax the assumption that the surface is white, recent studies suggested to use RGBD-sensing [3], complementary lights [45], a prior on the reflectance piecewise-uniformness [4], or to exploit the dynamicity of the scene [24]. Yet, it seems impossible to deal with a colored surface lit by colored sources, without resorting to such priors. The only case where the integral (2) can be simplified into the product of a material-dependent term and a lighting-dependent one:

$$I^i_\star = \rho_\star \overline{\mathbf{n}} \cdot \mathbf{s}^i_\star \tag{4}$$

without further assumptions, is when the camera response $c_\star(\lambda)$ is a Dirac delta function. This hypothesis may be acceptable in hyper-spectral imagery, which has interesting properties for PS related to interreflections removal [34], but it remains too empirical for consumer RGB sensors, hence we focus in this paper on the case of white sources, to guarantee theoretical validity[1].

Colored surface lit by white sources. To deal with colored surfaces without introducing a prior or an external sensor, another possibility is to consider $m \geq 3$ white sources, *i.e.* $\mathbf{s}^i(\lambda) = \mathbf{s}^i$, $i \in [1, m]$, while dealing with color-dependent material. In this case, Eq. (2) becomes:

$$I^i_\star = \left[\underbrace{\int_0^{+\infty} c_\star(\lambda)\rho(\lambda)\,\mathrm{d}\lambda}_{\triangleq \rho_\star}\right] \overline{\mathbf{n}} \cdot \mathbf{s}^i \tag{5}$$

[1] Let us note that, despite having no theoretical justification, the empirical model (4) can be numerically handled by the proposed framework.

where the $\rho_\star$, $\star \in \{R, G, B\}$, will be abusively referred to as *albedos*, though they include the camera response. Estimating the normals in an optimal manner from a set of equations such as (5) is not an easy task. Applying the classical grayscale PS approach to each color channel provides the desired color albedos, but the three estimated normals have no reason to match. Barsky and Petrou [7] showed how to use linear least-squares to simultaneously estimate the three albedos $\rho_\star$ and the shading terms $\overline{\mathbf{n}} \cdot \mathbf{s}^i$, followed by PCA to extract the normal from the shadings. Ikeda *et al*. [20] remarked that integrability is not ensured estimating normals in this way, hence they suggested an alternative two-steps procedure, approximating first the albedo triplets, and then the shape by solving a nonlinear PDE. This indeed prevents propagation of errors on the normals during the integration step, but on the other hand solving such nonlinear PDE is not straightforward. Furthermore, both these approaches suffer from the same bias: the shape estimation is reliable as long as the albedo estimation is good enough.

Hence, there is still a need for an unbiased method regarding PS for colored surfaces. We present in the next section an original solution for this problem, that simultaneously eliminates the nonlinearity of the PDEs and the need for albedo estimation.

3. PS as a System of Linear PDEs

When using a two-steps approach such as those described in the previous section, the accuracy of the shape estimation strongly depends on that of the albedo. We tackle this problem by eliminating the albedo dependency from the unknowns through photometric ratios, which are well-known to yield photometric invariants [17, 27, 53], and were recently considered by Mecca and Falcone [29] to prove uniqueness of the solution for a linearized PS problem, by Chandraker *et al*. [12] to deal with more general reflectance models, and by Alldrin and Kriegman [1] to describe the surface through its isocontours. To the best of our knowledge, the potential of dealing with ratios rather than single irradiance equations for each image has never been exploited with respect to PS for colored surfaces. We will show that such ratios allow us to extend the linearity of classical PS resolution, without having to resort to the change of variables $\mathbf{m} = \rho\mathbf{n}$ (avoiding bias due to bad albedo estimation), and without relying on the two-steps procedure consisting in estimating the normals and integrating them afterwards (avoiding bias due to integration).

Colored ratios. Assuming, for the moment, orthographic projection, we denote:

$$\overline{\mathbf{n}}(u,v) = \frac{1}{\sqrt{\|\nabla z(u,v)\|^2 + 1}} \underbrace{\begin{bmatrix} -\nabla z(u,v) \\ 1 \end{bmatrix}}_{\mathbf{n}(u,v)} \tag{6}$$

as the unit normal to the surface pointing outward, where ∇z is the gradient of the depth map. Dividing the irradiance equations in (5) coming from the ith and j^{th} light sources w.r.t. the same channel $\star \in \{R, G, B\}$, we obtain the following equation:

$$\underbrace{\frac{I^i_\star(u,v)}{\mathbf{n}(u,v)\cdot\mathbf{s}^i}}_{\text{Equation for } I^i_\star(u,v)} = \overbrace{\frac{\rho_\star(u,v)}{\|\mathbf{n}(u,v)\|} = \frac{I^j_\star(u,v)}{\mathbf{n}(u,v)\cdot\mathbf{s}^j}}^{\text{Equation for } I^j_\star(u,v)} \tag{7}$$

which gives:

$$\left[I^i_\star(u,v)\mathbf{s}^j - I^j_\star(u,v)\mathbf{s}^i\right] \cdot \mathbf{n}(u,v) = 0. \tag{8}$$

Each color channel provides $\binom{m}{2}$ linear equations such as (8), meaning that the number of equations increases quadratically w.r.t. the number m of images. In the case of graylevel images, Wu and Tang suggested in [56] to solve the homogeneous linear system (8) in $\mathbf{n}(u,v)$ by SVD. Yet, integrability of estimated normals has no reason to be satisfied, hence we suggest another route, based on PDEs.

From local to global formulation. We remark that the previous formulation (8) provides a *local* description of the shape, through normal vectors. Inferring the depth from the normals is not a trivial task, since the estimated normal field may be *non-integrable* [15].

To avoid such problems, a PDE-based PS model is preferred to the classical formulation which has the normal as unknown. Replacing the $\mathbf{n}$ vector in (8) by its definition given in (6), we obtain:

$$\begin{bmatrix} I^i_\star(u,v)s^j_1 - I^j_\star(u,v)s^i_1 \\ I^i_\star(u,v)s^j_2 - I^j_\star(u,v)s^i_2 \end{bmatrix} \cdot \nabla z(u,v) = I^i_\star(u,v)s^j_3 - I^j_\star(u,v)s^i_3 \tag{9}$$

so that, stacking these $3\binom{m}{2}$ equations, we eventually obtain the following system of linear PDEs:

$$\mathbf{A}(u,v)^\top \nabla z(u,v) = \mathbf{b}(u,v), \ (u,v) \in \Omega \tag{10}$$

where $\mathbf{A}$ is a field of $\mathbb{R}^{2\times 3\binom{m}{2}}$ matrices, $\mathbf{b}$ is a field of $\mathbb{R}^{3\binom{m}{2}}$ vectors, and Ω is the reconstruction domain.

Relevance with previous work. If only one color channel is used in the image acquisition, our approach simply consists in a new formulation of the classical PS problem [54] as a system of linear PDEs which is independent from the albedo. Such a formulation was recently considered in [30, 48]. Its advantage is to be global, thus implicitly ensuring some regularity, having the depth rather than the normals as unknown. It is simpler than other PDE-based variational formulations which do not rely on the ratio procedure [11, 20, 42], since the resulting PDEs are linear in our case.

From a theoretical side, our single-stage approach is more optimal than approaches relying on separation of albedo from shading, followed by an estimation of the shape [7, 20]. Indeed, it avoids both biases due to albedo estimation and normal integration. Note that the idea of using a single formulation for the recovery problem, instead of a sequence of estimations, was also successfuly applied to other inverse problems such as feature-specific imaging [35] or material unmixing [2].

Although the idea of coupling image ratios with a PDE formulation of PS is not novel [1, 12, 32, 31], our approach goes beyond these works for the following reasons. 1) These work deal with *graylevel* images, so color images need to be converted to grayscale (loss of information), while we show that ratios provide an unbiased solution to *color* photometric stereo. 2) To solve the resulting PDE, they rely on the knowledge of a *boundary condition*, which is rarely available in practice. Our approach does not require such boundary condition. 3) Numerical schemes used in [1, 32, 31] rely on *propagation* of information from the boundaries, accumulating errors during the propagation (see Figure 7-d in [1], and Figure 7-c in [32]). The approach presented in Section 4 solves for all depth values simultaneously, avoiding this bias. 4) These approaches do not deal with the images *as they are*: they require preprocessing. In comparison, our approach is single-stage, avoiding bias due to error propagation between pre-processing and resolution of the PDE. Indeed, Mecca et al. [32, 31] require determining a *single pair* of images rather than all the possible ratios; Alldrin et al. [1] require detecting additional cues for disambiguation, such as shadows or highlights; and Chandraker et al. [12] require pre-smoothing the ratios to estimate first two parameters (λ and κ), before actually solving a PDE. The solution we develop in the next section provides an answer to these numerical issues. Note that a similar formulation was also proposed very recently by another group, in the case of graylevel images [48].

4. Variational Resolution

Our formulation of PS for colored surfaces, resumed in the system of linear PDEs (10), requires a robust resolution method due to the high number of equations which is quadratic w.r.t. m. This suggests the use of a minimization framework, in order to reduce the discrepancy from the theoretical formulation to the real world data.

Least-squares formulation. If we assume that the real images differ from the Lambertian model (5) by a zero-mean, homoskedastic Gaussian noise (dealing with outliers such as shadows and highlights is left as future work), then the residuals between the observed ratios and the theoretical ones follow a Cauchy distribution [19]. The best linear *unbiased* estimator (BLUE) is hence the least-squares estimator, so we should solve (10) by computing:

$$\min_z \underbrace{\iint_{\Omega\subset\mathbb{R}^2} \left\|\mathbf{A}(u,v)^\top \nabla z(u,v) - \mathbf{b}(u,v)\right\|^2 \mathrm{d}u\,\mathrm{d}v}_{=\|\mathbf{A}^\top\nabla z-\mathbf{b}\|^2_{L^2(\Omega)}} \quad (11)$$

which differs from propagation schemes [1, 12, 29, 32, 31] due to the global nature of minimization, linearizes the variational approaches presented in [11, 42], and extends the recent approaches [30, 48] to RGB-valued images. In order to write Eq. (11) meaningfully, it is enough to assume that $z \in \mathcal{W}^{1,2}(\Omega)$, $\mathcal{W}^{1,2}(\Omega)$ being the Sobolev space of functions whose gradient is L^2-integrable over Ω. Hence, the recovered surface will be smooth without explicitly enforcing integrability. Eventually, let us remark that, although (11) provides the statistical BLUE estimator, this is not the *efficient* one since the variance of the estimator is not minimal. This optimal estimator would be obtained by replacing $\|\cdot\|^2$ in (11) by $\log(s^2 + \|\cdot\|^2)$, for some well-chosen $s \in \mathbb{R}$. Unfortunately, the optimization would become non-convex, hence much more difficult. Yet, robustness would be improved, and the surface would even be allowed to be only *piecewise* smooth (e.g., discontinuity jumps) [14].

Regularization. The minimization problem (11) does not admit a unique solution, because the functional to minimize is not coercive. This is a consequence of Eq. (10) where if z is a solution, then $z + k$ ($k \in \mathbb{R}$) is another one. Hence, the solution can be computed only up to an additive constant. To deal with this issue, we introduce an artificial least-squares prior z_0 on the solution, turning the initial problem (11) into its zero-order Tikhonov regularized version, as proposed recently in other PDE-based variational approaches [30, 42]. The regularized problem is given by:

$$\min_z \left\|\mathbf{A}^\top\nabla z - \mathbf{b}\right\|^2_{L^2(\Omega)} + \lambda \left\|z - z_0\right\|^2_{L^2(\Omega)}. \quad (12)$$

An interesting application of this regularization would consist in taking non-uniform functions z_0 since this could be used to handle priors on the solution, obtained for instance using a RGBD-sensor [3, 37]. In our case, we chose $z_0 \equiv 0$ and $\lambda = 10^{-9}$ if no prior information is available, which suffices to ensure a unique solution. The parameter λ should thus not be considered as critical.

Discretization. In order to avoid involving explicit boundary conditions, we discretize the functional (12) and then write the discrete optimality condition associated to each $z(u, v)$, rather than discretizing the continuous optimality conditions (Euler-Lagrange equations). This way, we obtain a sparse system of $|\Omega|$ linear equations having as unknowns the depth values ($|\Omega|$ denotes the cardinality,

i.e. the number of pixels, of the discretized image domain). The matrix of this system is symmetric, positive definite and strictly diagonal dominant as soon as $\lambda > 0$, which guarantees the existence and uniqueness of a solution. Eventually, taking $\lambda = 0$, the matrix would be rank-1 deficient, positive but only semi-definite, diagonal dominant but not strictly.

We solve the resulting linear system using either Cholesky factorization, or conjugate gradient if the reconstruction domain is too large to allow direct resolution. As a consequence, least-squares surface reconstruction (12) comes at an $O(|\Omega| \log |\Omega|)$ cost [26]. For comparison, this is the same order of complexity as least-squares surface normal integration by DCT [47]. Let us remark that no hypothesis on the shape of the reconstruction domain needs to be introduced, unlike most integration methods which are designed for rectangular domains [15, 16, 47].

Experimental validation. For quantitatively evaluating the benefits of using our approach, we considered a synthetic dataset (cf. Figure 2) consisting of a smooth surface having as albedo the *mandrill* RGB image. This surface was successively illuminated by $m = 10$ directional lightings with zenithal angles all equal to $20°$ and evenly distributed azimuthal angles (ring-light setup [59] with small angular variations [52]). A zero-mean Gaussian noise with standard deviation $\sigma = 10\%$ of the maximum RGB value was added to the simulated images. The dataset is illustrated in Figure 2. Quantitative comparison of our approach, converting the images to grayscale (using Matlab's `rgb2gray` function) or using RGB images as they are, was performed w.r.t. the classical PS approach [54], w.r.t. the two-steps procedure building on PCA from [7], and w.r.t. the estimation of surface normals using non-differential ratios [56].

Figure 2. Synthetic dataset used for quantitative evaluation. From left to right: shape, albedo, and two images of the surface under directional lightings.

We illustrate on this dataset the main advantages of the proposed PDE approach: it is more robust, requires less images and is faster than existing approaches. All experiments were run in Matlab, on a I7 processor at $2.9\,GHz$ with $32\,GB$ of memory.

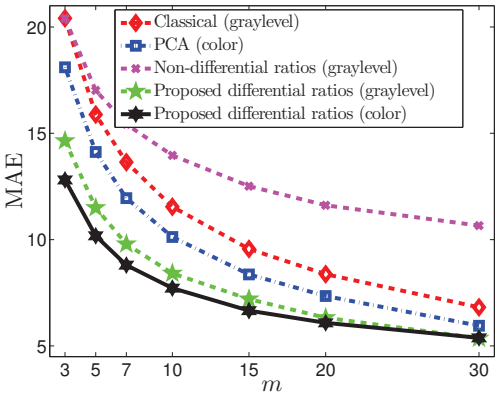
First, we compared the mean angular errors on the normals (MAE $= \mathrm{E}\left[\cos^{-1}\left(\overline{\mathbf{n}}^\top \widehat{\overline{\mathbf{n}}}\right)\right]$, $\overline{\mathbf{n}}$ being the ground truth normal and $\widehat{\overline{\mathbf{n}}}$ the estimated one), for different numbers m of images and different least-squares strategies, for graylevel PS (classical pseudo-inverse approach [54], non-differential ratios by solving (8) by SVD [56], and the proposed framework) and for color PS (two-steps PCA procedure [7] and the proposed framework). The results are shown in Figure 3-a. It is worth emphasizing, by comparing the classical approach [54] with the PCA approach [7], and both versions of our framework, that dealing with RGB images *as they are* (i.e., avoid the conversion to graylevels) provides a substantial improvement of the results. One can also remark that non-differential ratios [56] provide results which seem less accurate than the classical approach [54]: this is because the efficient estimator in the presence of additive Gaussian noise to the images is the least-squares one with the classical approach, but not with ratios (although it is the BLUE). Eventually, these results indisputably show the advantage of using a differential approach: its improved consistency regarding the regularity of the surfaces improves the results.

The experiments shown in Figure 3-b prove that, besides being more robust, the new framework is faster than existing approaches. Indeed, the PCA [7] and the non-differential ratios [56] approaches require performing, respectively, one PCA or one SVD per pixel, resulting in slow performances. Our approach is obviously slower than the classical approach [54] requiring only to compute the pseudo-inverse of a $m \times 3$ matrix, but on the other hand it is much faster than local approaches [7, 56]. This is because we need to solve a single linear system in order to estimate simultaneously all depth values, and this system is well-conditioned thanks to the regularization term. Moreover, since our discretization strategy allows us to use a fast solver (conjugate gradient or Cholesky factorization), our approach is probably also faster than the other state-of-the-art differential ratios approach from [48], which uses QR factorization and no regularization. In fact, we observed that most of the CPU time in our approach is spent constructing the $\mathbf{A}$ and $\mathbf{b}$ fields, not actually solving the resulting linear system.

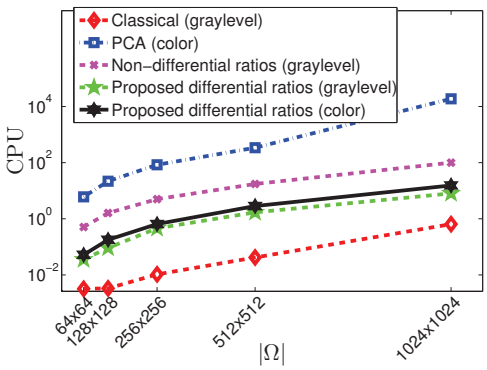
These simple experiments on synthetic datasets confirm that the proposed approach is a relevant step forward with respect to state-of-the-art. In Figure 1, we show the 3D-reconstruction of a small piece of a banknote, obtained using $m = 5$ RGB images, of size 550×650, recorded by a microgeometry capture device[2] (a HD camera with high focal length and calibrated lightings) similar to those presented in [22, 49]. This example confirms the ability of PS to separate the shape from the color of a surface without any prior, unlike single-image methods [5, 6].

In the next section we move forward the limits of our PS formulation considering orthographic viewing geometry and uniform lightings. To this end, we show how to handle more realistic assumptions related to perspective viewing geometry and non-directional lightings.

[2]We used the device developed by the Pixience company (Toulouse, France). The integration of the PS technology to this device was funded by the Toulouse Tech Transfer company.



(a)



(b)

Figure 3. (a) Mean angular error (MAE, in degrees) on the estimated normals, as a function of the number m of input images. (b) CPU time (in seconds) with respect to the size $|\Omega|$ of the reconstruction domain.

5. Extensions to Realistic Acquisition Setups

In order to derive a more realistic PS formulation, we need to move beyond the limits imposed by not considering perspective deformation and non-directional lightings. We leave as future work the important problem of dealing with shadows and non-negligible specular component in the reflectance model. A possible direction towards this goal would be to improve the robustness of our numerical solver by switching from the least-squares estimation (11) to a sparsity-enhancing one [21, 30, 42]. We rather focus on both the other aspects: the camera and the lightings, which are the parameters involved in the formation of color levels that the user can control.

Perspective projection. Perspective viewing geometry for photometric 3D-reconstruction has been introduced by Bruckstein in the shape-from-shading (SfS) context [10] and later developed by Prados and Faugeras [41] (see [8] for a recent review of perspective SfS algorithms). Regarding the classical PS problem aiming at estimating the normals [54], perspective modelling has influence neither on well-posedness, nor on the estimation method, since the estimation of normals from a set of equations such as (1) is independent from the projection model. It is only when integrating the normals that perspective must be considered, see e.g. [14, 50]. Note that perspective modelling may be used in the PS context for better constraining the uncalibrated photometric stereo problem [38], and may also be mandatory in applications involving cameras with strong perspective effects such as endoscopic imagery [13].

In our case, the PDE approach explicitly depends on the projection model which is easy to extend from the orthographic one. Considering the standard pinhole model, we assume the 3D points $\mathbf{x}(u,v)$ are conjugated to the pixels (u,v) according to:

$$\mathbf{x}(u,v) = \begin{bmatrix} \frac{uz(u,v)}{f} \\ \frac{vz(u,v)}{f} \\ z(u,v) \end{bmatrix} \tag{13}$$

where f is the *focal length* of the camera, and the (u,v) coordinates are now expressed w.r.t. the *principal point* position (the intrinsic camera parameters (f, u_0, v_0) are assumed to be calibrated). According to this parameterization, the non-unit direction of the outgoing normal to the surface is given by:

$$\mathbf{n}(u,v) = \begin{bmatrix} -\nabla z(u,v) \\ \frac{1}{f}\left(z(u,v) + [u,v]^\top \cdot \nabla z(u,v)\right) \end{bmatrix} \tag{14}$$

so that Eq. (9) must be replaced by:

$$\begin{bmatrix} I^i_\star(u,v)\left(s^j_1 - \frac{s^j_3 u}{f}\right) - I^j_\star(u,v)\left(s^i_1 - \frac{s^i_3 u}{f}\right) \\ I^i_\star(u,v)\left(s^j_2 - \frac{s^j_3 v}{f}\right) - I^j_\star(u,v)\left(s^i_2 - \frac{s^i_3 v}{f}\right) \\ -\frac{1}{f}\left(I^i_\star(u,v)s^j_3 - I^j_\star(u,v)s^i_3\right) \end{bmatrix} \cdot \begin{bmatrix} \nabla z(u,v) \\ z(u,v) \end{bmatrix} = 0 \tag{15}$$

and finally, the system of linear PDEs (10) is replaced by a new system of linear PDEs:

$$\mathbf{P}(u,v)^\top \begin{bmatrix} \nabla z(u,v) \\ z(u,v) \end{bmatrix} = \mathbf{0}, \quad (u,v) \in \Omega \tag{16}$$

with $\mathbf{P}$ a $\Omega \to \mathbb{R}^{3\times 3\binom{m}{2}}$ matrix field. As previously done while dealing with the orthographic parameterization, we use zero-order regularized least-squares to ensure robustness and fix the integration constant (Eq. (16) can be solved only up to a scale factor). This yields the following least-squares problem:

$$\min_z \left\| \mathbf{P}^\top \left[\nabla z^\top, z\right]^\top \right\|^2_{L^2(\Omega)} + \lambda \left\| z - z_0 \right\|^2_{L^2(\Omega)} \tag{17}$$

consistently with Problem (12) used in Section 4.

Spatially-varying lightings. Real-world lightings used in PS data acquisition such as pointwise [32] or extended [44] sources may not be well approximated using a directional model. The lightings $\mathbf{s}^i$ should rather be considered as tridimensional vector fields over Ω. If we adopt this strategy in the classical approach to PS, $|\Omega|$ pseudo-inverses of $m \times 3$ matrices have to be calculated rather than only one under the directional assumption. Consequently, the complexity dramatically increases. Instead, since our approach only requires additional space dependency to the lightings in Eqs. (9) and (15), computational cost does not increase.

Another issue regarding spatially-varying lightings is due to the correct parameterization depending on the depth of the surface itself. For instance, pointwise sources induce an attenuation of the luminous flux density which is proportional to the squared source-surface distance:

$$\mathbf{s}^i(u,v) = \frac{\phi^i}{\|\mathbf{x}_s^i - \mathbf{x}(u,v)\|^2} \frac{\mathbf{x}_s^i - \mathbf{x}(u,v)}{\|\mathbf{x}_s^i - \mathbf{x}(u,v)\|} \tag{18}$$

where ϕ^i is the source intensity and $\mathbf{x}_s^i$ its location.

State-of-the-art approaches dealing with nonlinear lighting models such as (18) adopt methods based on alternating estimation of normals and integration steps [39], mesh deformation [57] or nonlinear PDEs [32]. We believe that the linear PDEs solution we put forward may be an interesting alternative for solving the near-light PS problem. Yet, system (16) becomes a system of *quasilinear* PDEs, *i.e.* the $\mathbf{P}$ matrix in (16) depends on the unknown depth, hence we denote it $\mathbf{P}_z$. To deal with this issue, we suggest the following fixed-point strategy, which turns the non-linear variational problem into a series of linear problems of the form (17):

$$z^{k+1} = \underset{z}{\operatorname{argmin}} \left\| \mathbf{P}_{z^k}^\top \left[\nabla z^\top, z\right]^\top \right\|_{L^2(\Omega)}^2 + \lambda \|z - z_0\|_{L^2(\Omega)}^2 . \tag{19}$$

We experimentally observed this fixed-point scheme always converges, though we leave the proof as perspective. Theoretical results on well-posedness of PS under nearby lightings [32] would probably help understanding this convergent behavior: they would guarantee the existence of a global minimizer for the functional, while the fixed point strategy could be interpreted as a gradient descent with fixed stepsize, iteratively driving the estimate towards this minimizer.

Experimental validation. To verify the applicability of these extensions to perspective projection and spatially-varying lightings, we tested our approach on a dataset consisting of $m = 8$ RGB images, of size 1260×1600. They have been captured using a consumer camera (Canon EOS 7D) at a distance of around 50 cm (we set z_0 to this rough prior, along with $\lambda = 10^{-9}$) from a colored scene made of several colored objects: two cuddy toys, two plaster busts and a plastic Tintin character. White LEDs were placed between the camera and the scene as light sources ensuring the nearby lightings setup. The camera was calibrated using Matlab's built-in procedure, and the positions and intensities of the light sources were estimated by a procedure similar to the one described in [40]. Figure 4 shows the results of our new scheme, without and with perspective geometry and spatially-varying lightings modelling. Despite the strong shadowing effects and depth discontinuities, we were able to achieve a very satisfactory 3D-reconstruction in a reasonable time (4 iterations (19), which required approximately 30 seconds).

Figure 4. Top: two out of $m = 8$ RGB images recorded under nearby lightings. Bottom-left: reconstructed shape using the directional lightings and the orthographic projection models. Bottom-right: rendering of the shape reconstructed using the pointwise lightings (18) and the perspective projection (13) models, seen under the same angle.

Off-Lambertian reflectance. As final experiments, we question the robustness of the method with respect to off-Lambertian reflectances. To this purpose, we considered challenging datasets consisting of 5 RGB pictures of metallic coins (Figure 5) and synthetic human skin samples (Figure 6), obtained using the same device as in the experiment of Figure 1. While the approach presented in [22] introduced a chemical gel between the camera and the object, in order to make the reflectance of the object as diffuse as possible, we were able to obtain very reasonable 3D-reconstructions dealing with the images as they are. To compensate off-Lambertian reflectance by regularization, we simply set $\lambda = 100$. We also emphasize that these results were obtained in only 5 seconds using non-optimized Matlab codes without any kind of parallelization.

These experiments show that our approach, although limited to the Lambertian reflectance model, provides realistic results even in the presence of strong deviations from this model, with as few as $m = 5$ images. For comparison, state-of-the-art PS methods such as [18] use around 50 images. Of course, our results would probably be further improved by relaxing the Lambertian assumption [18, 30, 46], or by using more robust estimators [14, 21, 30, 42].

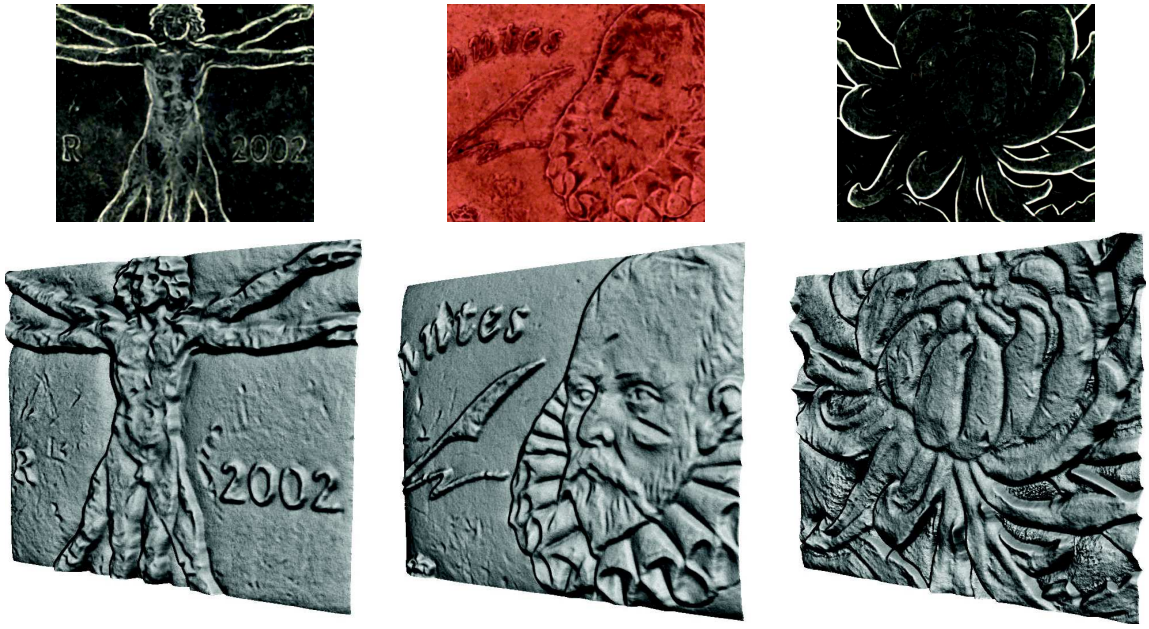


Figure 5. 3D-reconstructions of three metallic coins: an Italian euro, a Spanish 50 cents coin and a Chinese yuan. Each reconstruction domain is around 1 cm^2. Top: one of the $m = 5$ input images. Bottom: 3D-reconstruction result. Metals are successfully handled, despite the Lambertian assumption and the choice of least-squares estimation.

6. Conclusion and Perspectives

In this work we exploited important features of colored image ratios, moving beyond the state-of-the-art regarding 3D shape recovery using PS. In particular, instead of locally solving a system of graylevel irradiance equations mixing the albedo and the normal, we considered a single system of linear PDEs obtained by considering the ratios of color irradiance equations. Doing so, we take advantage of three fundamental aspects. Firstly, we avoid the loss of information due to conversion from RGB to grayscale. Secondly, formulating the recovery problem as a system of linear PDEs yields an implicit regularity assumption on the surface, avoiding integrability concerns about the normals. Thirdly, since the PDEs derived by the ratios are photometric invariants, we obtained our variational formulation while eliminating the usual accumulation of biases in color PS, due to the sequence albedo estimation / normal estimation / normal integration.

We foresee exciting extensions of this framework, regarding the PS problems where integrability plays a major role to ensure well-posedness, e.g. when $m = 2$ [36] or when the lightings are unknown (UPS, for *uncalibrated* PS). For instance, since it is enough to ensure integrability of the normals to make *perspective UPS* well-posed [38], and since integrability is implicitely granted in our (differential) framework, a rather simple extension of our approach estimating in an alternating way the shape and the lightings may actually yield a well-posed formulation of perspective UPS. Regarding *orthographic UPS* which is prone to the GBR ambiguity [58], one could introduce regularization terms penalizing the variations of the depth or of the albedo inside the objective function. Using one or the other regularization, along with integrability enforcement, ensures well-posedness of orthographic UPS [43].

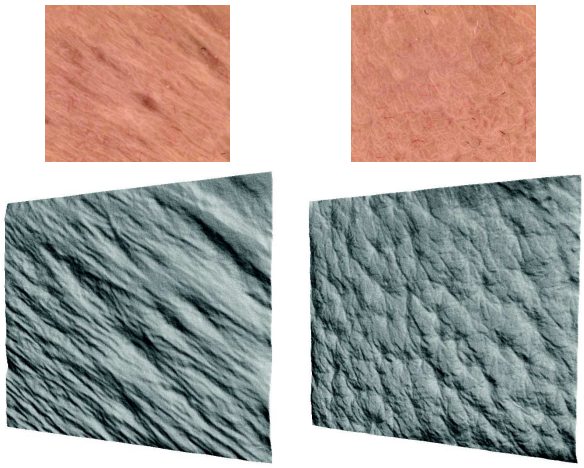


Figure 6. 3D-reconstructions of two samples of synthetic human skin. Top: one image of each dataset, out of $m = 5$. Bottom: 3D-reconstruction result. Despite the high level of subsurface scattering, the 3D-reconstruction is very satisfactory.

References

[1] N. G. Alldrin and D. J. Kriegman. Toward Reconstructing Surfaces With Arbitrary Isotropic Reflectance : A Stratified Photometric Stereo Approach. In *IEEE International Conference on Computer Vision (ICCV)*, Rio de Janeiro, Brasil, 2007. 3, 4

[2] M. Alterman, Y. Y. Schechner, and A. Weiss. Multiplexed fluorescence unmixing. In *IEEE International Conference on Computational Photography (ICCP)*, Cambridge, USA, 2010. 4

[3] R. Anderson, B. Stenger, and R. Cipolla. Augmenting Depth Camera Output Using Photometric Stereo. In *IAPR Conference on Machine Vision Applications (MVA)*, Nara, Japan, 2011. 2, 4

[4] R. Anderson, B. Stenger, and R. Cipolla. Color photometric stereo for multicolored surfaces. In *IEEE International Conference on Computer Vision (ICCV)*, Barcelona, Spain, 2011. 2

[5] J.-F. Aujol, G. Gilboa, T. Chan, and S. Osher. Structure-Texture Image Decomposition – Modeling, Algorithms, and Parameter Selection. *International Journal of Computer Vision*, 67(1):111–136, 2006. 5

[6] J. T. Barron and J. Malik. Shape, albedo, and illumination from a single image of an unknown object. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, USA, 2012. 5

[7] 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. 3, 4, 5

[8] M. Breuß, E. Cristiani, J.-D. Durou, M. Falcone, and O. Vogel. Perspective Shape from Shading: Ambiguity Analysis and Numerical Approximations. *SIAM Journal on Imaging Sciences*, 5(1):311–342, 2012. 6

[9] G. J. Brostow, C. Hernández, G. Vogiatzis, B. Stenger, and R. Cipolla. Video normals from colored lights. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(10):2104–2114, 2011. 2

[10] A. M. Bruckstein. On shape from shading. *Computer Vision, Graphic, and Image Processing*, 44(2):139–154, 1988. 6

[11] J. Chabrowski and Z. Kewei. On variational approach to photometric stereo. *Annales de l'Institut Henri Poincaré (C) Analyse non linéaire*, 10(4):363–375, 1993. 3, 4

[12] M. Chandraker, J. Bai, and R. Ramamoorthi. On Differential Photometric Reconstruction for Unknown, Isotropic BRDFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(12):2941–2955, 2013. 3, 4

[13] T. Collins and A. Bartoli. 3D Reconstruction in Laparoscopy with Close-Range Photometric Stereo. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Nice, France, 2012. 6

[14] J.-D. Durou, J.-F. Aujol, and F. Courteille. Integrating the normal field of a surface in the presence of discontinuities. In *Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR)*. Bonn, Germany, 2009. 4, 6, 7

[15] R. Frankot and R. Chellappa. A method for enforcing integrability in shape from shading algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(4):439–451, 1988. 3, 5

[16] M. Harker and P. O'Leary. Least squares surface reconstruction from gradients: Direct algebraic methods with spectral, Tikhonov, and constrained regularization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Colorado Springs, USA, 2011. 1, 5

[17] C. Hernández, G. Vogiatzis, and R. Cipolla. Overcoming shadows in 3-source photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(2):419–426, 2011. 3

[18] T. Higo, Y. Matsushita, and K. Ikeuchi. Consensus photometric stereo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Francisco, USA, 2010. 7

[19] D. V. Hinkley. On the ratio of two correlated normal random variables. *Biometrika*, 56(3):635–639, 1969. 4

[20] O. Ikeda and Y. Duan. Color Photometric Stereo for Albedo and Shape Reconstruction. In *IEEE Workshop on Applications of Computer Vision (WACV)*, Copper Mountain, USA, 2008. 3, 4

[21] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Providence, USA, 2012. 1, 6, 7

[22] M. K. Johnson, F. Cole, A. Raj, and E. H. Adelson. Microgeometry capture using an elastomeric sensor. *ACM Transactions on Graphics*, 30(4):46:1–46:8, 2011. 5, 7

[23] J. Jung, J.-Y. Lee, and I. S. Kweon. One-day outdoor photometric stereo via skylight estimation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, USA, 2015. 1

[24] H. Kim, B. Wilburn, and M. Ben-Ezra. Photometric stereo for dynamic surface orientations. In *European Conference on Computer Vision (ECCV)*. Heraklion, Greece, 2010. 2

[25] L. L. Kontsevich, A. P. Petrov, and I. S. Vergelskaya. Reconstruction of shape from shading in color images. *Journal of the Optical Society of America A*, 11(3):1047–1052, 1994. 2

[26] 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)*, Palm Springs, USA, 2011. 5

[27] S. Lee and M. Brady. Integrating stereo and photometric stereo to monitor the development of glaucoma. *Image and Vision Computing*, 9(1):39–44, 1991. 3

[28] S. Mallick, T. Zickler, D. Kriegman, and P. Belhumeur. Beyond Lambert: Reconstructing Specular Surfaces Using Color. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Diego, USA, 2005. 2

[29] R. Mecca and M. Falcone. Uniqueness and approximation of a photometric shape-from-shading model. *SIAM Journal on Imaging Sciences*, 6(1):616–659, 2013. 3, 4

[30] R. Mecca and Y. Quéau. Unifying Diffuse and Specular Reflections for the Photometric Stereo Problem. In *IEEE Workshop on Applications of Computer Vision (WACV)*, Lake Placid, USA, 2016. 3, 4, 6, 7

[31] R. Mecca, A. Tankus, A. Wetzler, and A. M. Bruckstein. A Direct Differential Approach to Photometric Stereo with Perspective Viewing. *SIAM Journal on Imaging Sciences*, 7(2):579–612, 2014. 4

[32] R. Mecca, A. Wetzler, A. Bruckstein, and R. Kimmel. Near Field Photometric Stereo with Point Light Sources. *SIAM Journal on Imaging Sciences*, 7(4):2732–2770, 2014. 4, 7

[33] Z. Murez, T. Treibitz, R. Ramamoorthi, and D. Kriegman. Photometric stereo in a scattering medium. In *IEEE International Conference on Computer Vision (ICCV)*, 2015. 1

[34] G. Nam and M. H. Kim. Multispectral photometric stereo for acquiring high-fidelity surface normals. *IEEE Computer Graphics and Applications*, 34(6):57–68, 2014. 2

[35] M. A. Neifeld and P. Shankar. Feature-specific imaging. *Applied Optics*, 42(17):3379–3389, 2003. 4

[36] R. Onn and A. Bruckstein. Integrability disambiguates surface recovery in two-image photometric stereo. *International Journal of Computer Vision*, 5(1):105–113, 1990. 8

[37] R. Or-el, G. Rosman, A. Wetzler, R. Kimmel, and A. M. Bruckstein. RGBD-Fusion: Real-Time High Precision Depth Recovery. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, USA, 2015. 4

[38] T. Papadhimitri and P. Favaro. A New Perspective on Uncalibrated Photometric Stereo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Portland, USA, 2013. 1, 6, 8

[39] T. Papadhimitri and P. Favaro. Uncalibrated Near-Light Photometric Stereo. In *British Machine Vision Conference (BMVC)*, Nottingham, United Kingdom, 2014. 7

[40] M. Powell, S. Sarkar, and D. Goldgof. A simple strategy for calibrating the geometry of light sources. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(9):1022–1027, 2001. 7

[41] E. Prados and O. Faugeras. Perspective shape from shading and viscosity solutions. In *IEEE International Conference on Computer Vision (ICCV)*, Nice, France, 2003. 6

[42] 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)*, Lège Cap-Ferret, France, 2015. 3, 4, 6, 7

[43] Y. Quéau, F. Lauze, and J.-D. Durou. Solving Uncalibrated Photometric Stereo using Total Variation. *Journal of Mathematical Imaging and Vision*, 52(1):87–107, 2015. 8

[44] Y. Quéau, R. Modrzejewski, P. Gurdjos, and J.-D. Durou. A Full Photometric and Geometric Model for Webcam + Matte Screen Devices. *Signal Processing: Image Communications*, 40:65–81, 2016. 7

[45] S. Rahman, A. Lam, I. Sato, and A. Robles-Kelly. Color Photometric Stereo Using a Rainbow Light for Non-Lambertian Multicolored Surfaces. In *Asian Conference on Computer Vision (ACCV)*, Singapore, 2014. 2

[46] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi. Bi-polynomial modeling of low-frequency reflectances. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(6):1078–1091, 2014. 1, 7

[47] T. Simchony, R. Chellappa, and M. Shao. Direct analytical methods for solving Poisson equations in computer vision problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):435–446, 1990. 5

[48] W. Smith and F. Fang. Height from Photometric Ratio with Model-based Light Source Selection. *Computer Vision and Image Understanding*, 2016. (to appear). 3, 4, 5

[49] J. Sun, M. Smith, L. Smith, L. Coutts, R. Dabis, C. Harland, and J. Bamber. Reflectance of human skin using colour photometric stereo: with particular application to pigmented lesion analysis. *Skin research and technology*, 14(2):173–179, 2008. 5

[50] A. Tankus and N. Kiryati. Photometric stereo under perspective projection. In *IEEE International Conference on Computer Vision (ICCV)*, Beijing, China, 2005. 6

[51] D. Vlasic, P. Peers, I. Baran, P. E. Debevec, J. Popovic, S. Rusinkiewicz, and W. Matusik. Dynamic shape capture using multi-view photometric stereo. *ACM Transactions on Graphics*, 28(5), 2009. 1

[52] J. Wang, Y. Matsushita, B. Shi, and A. C. Sankaranarayanan. Photometric Stereo with Small Angular Variations. In *IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 2015. 5

[53] L. B. Wolff and E. Angelopoulou. 3-d stereo using photometric ratios. In *European Conference on Computer Vision (ECCV)*, Stockholm, Sweden, 1994. 3

[54] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):134–144, 1980. 1, 3, 5, 6

[55] C. Wu, S. G. Narasimhan, and B. Jaramaz. A Multi-Image Shape-from-Shading Framework for Near-Lighting Perspective Endoscopes. *International Journal of Computer Vision*, 86(2-3):211–228, 2010. 1

[56] T.-P. Wu and C.-K. Tang. Dense photometric stereo using a mirror sphere and graph cut. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Diego, USA, 2005. 3, 5

[57] W. Xie, C. Dai, and C. C. L. Wang. Photometric Stereo With Near Point Lighting: A Solution by Mesh Deformation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, USA, 2015. 1, 7

[58] A. Yuille and D. Snow. Shape and albedo from multiple images using integrability. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, San Juan, USA, 1997. 8

[59] Z. Zhou and P. Tan. Ring-light photometric stereo. In *European Conference on Computer Vision (ECCV)*, Heraklion, Greece, 2010. 5