Image and Vision Computing 57 (2017) 175–191

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Editor's choice article

# Photometric stereo with only two images: A theoretical study and numerical resolution ☆

Yvain Quéau [a, b, *], Roberto Mecca [c, d, 1], Jean-Denis Durou [a], Xavier Descombes [e]

[a] IRIT, Université de Toulouse, France
[b] Technical University Munich, Garching, Germany
[c] Department of Engineering, University of Cambridge, United Kingdom
[d] Department of Mathematics, University of Bologna, Italy
[e] INRIA, Sophia Antipolis, France

ARTICLE INFO

*Article history:*
Received 10 July 2015
Received in revised form 13 June 2016
Accepted 15 November 2016
Available online 23 November 2016

*Keywords:*
3D-reconstruction
Shape-from-shading
Photometric stereo
PDEs
Numerical analysis
Optimization
Graph cut

ABSTRACT

This work tackles the problem of two-image photometric stereo. This problem constitutes the intermediate case between conventional photometric stereo with at least three images, which is well-posed, and shape-from-shading, which is ill-posed. We first provide a theoretical study of ambiguities arising in this intermediate case. Based on this study, we show that when the albedo is known, disambiguation can be formulated as a binary labeling problem, using integrability and a nonstationary Ising model. The resulting optimization problem is solved efficiently by resorting to the graph cut algorithm. These theoretical and numerical contributions are eventually validated in an application to three-image photometric stereo with shadows.



## 1. Introduction

In the computer vision field, 3D-shape reconstruction using digital images as input data has gained a growing importance. Interest in this task has increased even more since most mass digital devices have been equipped with cameras. Based on more than thirty years of research, such devices are potentially convertible into 3D-scanners without any hardware correction. Among all the photographic 3D-reconstruction techniques, we focus in this work on shape-from-shading (SFS) and photometric stereo (PS), which exploit shading information when one (SFS) or several (PS) sources illuminate the observed object. For a comprehensive overview on these techniques, see the reference book [1] by Horn and Brooks, but also [2] and [3,4] for up-to-date surveys on SFS and PS, respectively.

Many articles enlightened the impossibility of avoiding any ambiguity while retrieving the shape from a single image, as in the SFS problem [5]. This impossibility arises from the difficulty met in distinguishing the concave from the convex surfaces. The most natural way to solve this problem is to use more than one image: Woodham showed in [6] that three is the minimum number of images to ensure well-posedness of the PS problem.

This work focuses on the intermediate case when only two images are taken into account (this specific situation will be referred to as PS2). Besides being particularly interesting for dedicated applications as single-day outdoor PS from sun light [7,8], the PS2 problem can be seen as the degenerative case of lack of information from the three-source PS problem due to shadows [9]. In this view, we provide working tools aimed at solving the underlying ambiguities, derived from a theoretical study.

There exist a combinatorial number of normal fields which are solutions of the PS2 problem. Exhaustive search can be carried out among these normal fields, in order to find the one which best satisfies a smoothness constraint [10]. Alternatively, one may resort to the differential approach of PS, which implicitly enforces smoothness. A meaningful solution of the resulting PDE can be obtained either by specifying an explicit boundary condition [11], or by resorting to regularization [9]. Unfortunately, knowledge of the surface on the boundary is rarely available, and regularization techniques come along with parameters to tune, which might be tedious.

☆ Editor's Choice Articles are invited and handled by a select rotating 12 member Editorial Board committee. This paper has been recommended for acceptance by Enrique Dunn, Ph.D.

* Corresponding author.
*E-mail address:* yvain.queau@tum.de (Y. Quéau).

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

http://dx.doi.org/10.1016/j.imavis.2016.11.006
0262-8856/© 2016 Elsevier B.V. All rights reserved.

We put forward a new method for solving the PS2 problem, which is based on the non-differential approach of PS (normal estimation). We assume that the setup consists of orthographic viewing geometry, parallel and uniform lighting, as well as Lambertian reflectance. Under such assumptions, and provided that the albedo is known, it is shown in this paper that the number of possible solutions can be predicted beforehand. It is then demonstrated that exhaustive search of the "best" normal field can be recast as a binary labeling problem, efficiently solvable by resorting to the graph cut algorithm. In contrast with existing methods, the proposed one requires neither knowing a boundary condition, nor tuning any parameter.

In the following, we first review the general equations of SFS, PS and PS2 in Section 2. In Section 3, we compare the differential and non-differential formulations of PS2. We show why the PS2 problem usually has a unique solution in Section 4. In Section 5, a practical graph cut-based [12] algorithm to compute this solution is introduced, and it is applied in Section 6 to the problem of three-source PS with shadows.

## 2. From shape-from-shading to photometric stereo

### 2.1. Shape-from-shading

To fully describe the problem we are interested in, let us first recall some features of the SFS problem. We attach to the camera a 3D-Cartesian coordinate system $xyz$, so that $xy$ coincides with the image plane and $z$ with the optical axis. Under the assumption of orthographic projection, the visible part of a surface is a graph $z = u(x, y)$. It is well known that the SFS problem is modeled by the image irradiance equation[1]:

$$R(\mathbf{n}(x,y)) = \mathcal{I}(x,y) \tag{1}$$

where $\mathcal{I}(x, y)$ is the graylevel at image point $(x, y)$, and the reflectance function $R(\mathbf{n}(x, y))$ gives the value of the light re-emitted by the surface as a function of its orientation i.e., of the unit-length outgoing normal $\mathbf{n}(x, y)$ to the surface at surface point $[x, y, u(x, y)]^\top$. The unknown depth $u$ has to be reconstructed on a compact domain $\Omega \subset \mathbb{R}^2$ called the reconstruction domain.

Let us consider a unique parallel and uniform light beam whose direction is indicated by the unit-length vector $\mathbf{s} = [s_1, s_2, s_3]^\top = [\tilde{\mathbf{s}}^\top, s_3]^\top \in \mathbb{R}^3$, and whose intensity is denoted by $\psi$. Assuming the observed object has purely diffuse reflection, and ignoring shadows, Eq. (1) can be written as follows:

$$\rho(x,y)\,\psi\,\mathbf{s}\cdot\mathbf{n}(x,y) = \mathcal{I}(x,y) \tag{2}$$

where $\rho(x, y) \in [0, 1]$ is the albedo.

In fact, this equality is nothing more than a relation of proportionality. Knowing that the vectors $\mathbf{s}$ and $\mathbf{n}(x, y)$ have unit-length, and assuming that $\psi$ is a constant factor, it seems justified to rewrite Eq. (2) as a real equality:

$$\rho(x,y)\,\mathbf{s}\cdot\mathbf{n}(x,y) = I(x,y) \tag{3}$$

where $I(x, y) \in [0, 1]$ should now be considered as the *normalized* graylevel.

Eq. (3) is a particular non-differential formulation (among many others) of the SFS problem. Once the normal field $\mathbf{n}$ has been estimated, it has to be integrated. This means that the following equation in $u$ has to be solved [13]:

$$\mathbf{n}(x,y) = \frac{1}{\sqrt{1+\|\nabla u(x,y)\|^2}}\left[-\nabla u(x,y)^\top, 1\right]^\top \tag{4}$$

where $\nabla u(x, y) = [\partial_x u(x, y), \partial_y u(x, y)]^\top$ denotes the gradient of $u(x, y)$. From Eqs. (3) and (4), we get the following differential formulation of SFS:

$$\rho(x,y) - \frac{\tilde{\mathbf{s}}\cdot\nabla u(x,y) + s_3}{\sqrt{1+\|\nabla u(x,y)\|^2}} = I(x,y) \tag{5}$$

which is a first-order nonlinear PDE of the Hamilton–Jacobi type. We refer the interested reader to the survey presented in [2] for a presentation of recent results on the eikonal equation, which follows from Eq. (5) when $\mathbf{s} = [0, 0, 1]^\top$.

### 2.2. Photometric Stereo

Even if $\mathbf{s}$ is known, SFS is ill-posed without any additional knowledge on the surface to be reconstructed. In most papers on SFS, the albedo $\rho(x, y)$ is supposed to be known, but this is still not enough to make the problem well-posed. The simplest way to overcome SFS ill-posedness is to use $m \geq 2$ images taken from the same point of view, illuminated by $m$ light sources $(\mathbf{s}^i, \psi^i)$, $i \in [1, m]$. This new problem is called photometric stereo (PS). The classical resolution of PS is based on a local estimate of the outgoing unit-length normal to the surface [6]. For a Lambertian surface, the non-differential formulation of PS consists in solving a system of $m$ equations of type (3):

$$\rho(x,y)\,\mathbf{s}^i\cdot\mathbf{n}(x,y) = I^i(x,y), \quad i \in [1,m] \tag{6}$$

As for SFS, this formulation requires that Eq. (4) is solved afterwards. From a theoretical point of view, Eq. (4) admits a solution in $u$ only if the estimated normal field is *integrable*[14] (cf. Section 4.5). Due to estimation errors, this is rarely the case in practice, hence projection of the estimated normal field on the space of integrable fields must be achieved, resorting for instance to Fourier analysis [14] or to variational methods [13]. Alternatively, one can directly try to estimate the "most integrable" normal field. This is the approach that is followed in Section 5.

Of course, a differential formulation of PS also exists, which aims at solving a system of $m$ nonlinear PDEs of type (5):

$$\rho(x,y)\,\frac{-\tilde{\mathbf{s}}^i\cdot\nabla u(x,y) + s_3^i}{\sqrt{1+\|\nabla u(x,y)\|^2}} = I^i(x,y), \quad i \in [1,m] \tag{7}$$

In the usual case, denoted PS3, where $m \geq 3$ non-coplanar distant calibrated light sources are used [6], system (6) reduces to a full-rank linear system in $\mathbf{m}(x, y) = \rho(x, y)\mathbf{n}(x, y) \in \mathbb{R}^3$. Solving this system has several advantages, compared to SFS: it is well-posed and can be locally solved, thus parallelized. Furthermore, the albedo no longer has to be known.

We may wonder whether the differential formulation (7) would really be pertinent for PS3. The main advantage of solving Eq. (7) is that integrability is implicitly ensured, unlike solving Eq. (6), knowing that the lack of integrability of $\mathbf{n}$ complicates the resolution of Eq. (4) [14]. However, the problem (7) has two drawbacks: it is nonlinear and cannot be solved locally [15].

In this paper, we focus on the resolution of PS when the linear system (6) is not full-rank. In such cases, the non-differential formulation does not have as many advantages as for PS3, and the differential formulation might be worthwhile, at least because it is better-posed since the integrability constraint is implicitly satisfied. Indeed, practical solutions to the rank-deficient PS problem use this differential formulation. Yet, as discussed in Section 3.2, differential approaches have to resort either to a boundary condition (which is rarely available) or to regularization (which requires parameter tuning). The solution presented in Section 5, which is based on the