This is a repository copy of *Height-from-Polarisation with Unknown Lighting or Albedo*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/136316/

Version: Accepted Version

Article:

Smith, William Alfred Peter orcid.org/0000-0002-6047-0413, Ramamoorthi, Ravi and Tozza, Silvia (2019) Height-from-Polarisation with Unknown Lighting or Albedo. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2875 - 2888. ISSN 0162-8828

https://doi.org/10.1109/TPAMI.2018.2868065

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Height-from-Polarisation with Unknown Lighting or Albedo

William A. P. Smith, *Member, IEEE*, Ravi Ramamoorthi, *Fellow, IEEE*, and Silvia Tozza

Abstract—We present a method for estimating surface height directly from a single polarisation image simply by solving a large, sparse system of linear equations. To do so, we show how to express polarisation constraints as equations that are linear in the unknown height. The local ambiguity in the surface normal azimuth angle is resolved globally when the optimal surface height is reconstructed. Our method is applicable to dielectric objects exhibiting diffuse and specular reflectance, though lighting and albedo must be known. We relax this requirement by showing that either spatially varying albedo or illumination can be estimated from the polarisation image alone using nonlinear methods. In the case of illumination, the estimate can only be made up to a binary ambiguity which we show is a generalised Bas-relief transformation corresponding to the convex/concave ambiguity. We believe that our method is the first passive, monocular shape-from-x technique that enables well-posed height estimation with only a single, uncalibrated illumination condition. We present results on real world data, including in uncontrolled, outdoor illumination.

Index Terms—Polarisation, shape-from-x, bas-relief ambiguity, illumination estimation, albedo estimation

1 Introduction

When unpolarised light is reflected by a surface it becomes partially polarised [1]. This applies to both specular reflections [2] and diffuse reflections [3] caused by subsurface scattering. The angle and degree of polarisation of reflected light conveys information about the surface orientation and, therefore, provide a cue for shape recovery. There are a number of attractive properties to this 'shape-from-polarisation' (SfP) cue. It requires only a single viewpoint and illumination condition, it is invariant to illumination direction and surface albedo and it provides information about both the zenith and azimuth angle of the surface normal. Like photometric stereo, shape estimates are dense (surface orientation information is available at every pixel so resolution is limited only by the sensor) and, since it does not rely on detecting or matching features, it is applicable to smooth, featureless surfaces.

However, there are a number of drawbacks to using SfP in a practical setting. First, the polarisation cue alone provides only ambiguous estimates of surface orientation. Hence, previous work focussed on developing heuristics to locally disambiguate the surface normals. Even having done so, the estimated normal field must be integrated in order to recover surface height (i.e. relative depth) [4] or combined with a depth map from another cue [5]. This two-step approach of disambiguation followed by integration means that the integrability constraint is not enforced during disambiguation and also that errors accumulate over the two steps. Second, diffuse polarisation provides only a weak shape cue for regions of the surface with small gradient and so methods that operate locally are very sensitive to noise.

1.1 Contributions and Applicability of the Method

In this paper, we make a number of contributions to the SfP problem. After introducing notations and preliminaries in Section 3, in Section 4 we present our SfP method. This contains a number of novel ingredients. First, in contrast to prior work, we compute SfP in the height, as opposed to the surface normal, domain. Instead of disambiguating the polarisation normals, we defer resolution of the ambiguity until surface height is computed. To do so, we express the azimuthal ambiguity as a collinearity condition that is satisfied by either interpretation of the polarisation measurements. Second, we express polarisation and shading constraints as linear equations in the unknown surface height enabling efficient and globally optimal height estimation. We show an overview of our method and a sample result for unknown, outdoor illumination and uniform albedo in Fig. 1. In Sections 5 and 6 we explore what information can be obtained without disambiguating the polarisation normals. If illumination is unknown and albedo unknown but uniform then we show that illumination can be determined up to a binary ambiguity from the ambiguous normals and the unpolarised intensity. We make a theoretical contribution by showing that this ambiguity corresponds to a particular generalised Bas-relief [6] transformation (the convex/concave ambiguity). On the other hand, if illumination is known and albedo spatially varying and unknown, then we

- W. Smith is with the Department of Computer Science, University of York, York YO10 5DD, United Kingdom. E-mail: william.smith@york.ac.uk.
- R. Ramamoorthi is with the CSE Department, University of California, San Diego, La Jolla, CA 92093. E-mail: ravir@cs.ucsd.edu.
- S. Tozza is with Istituto Nazionale di Alta Matematica, Research Unit at Department of Mathematics, Sapienza - Università di Roma, Rome 00185, Italy. E-mail: tozza@mat.uniroma1.it.

Manuscript received 21 Sept. 2017; revised 9 Aug. 2018; accepted 26 Aug. 2018. Date of publication 0 . 0000; date of current version 0 . 0000.
(Corresponding author: William Smith.)
Recommended for acceptance by K. Nishino.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TPAMI.2018.2868065

This work is licensed under a Creative Commons Attribution 3.0 License. For more information, see http://creativecommons.org/licenses/by/3.0/

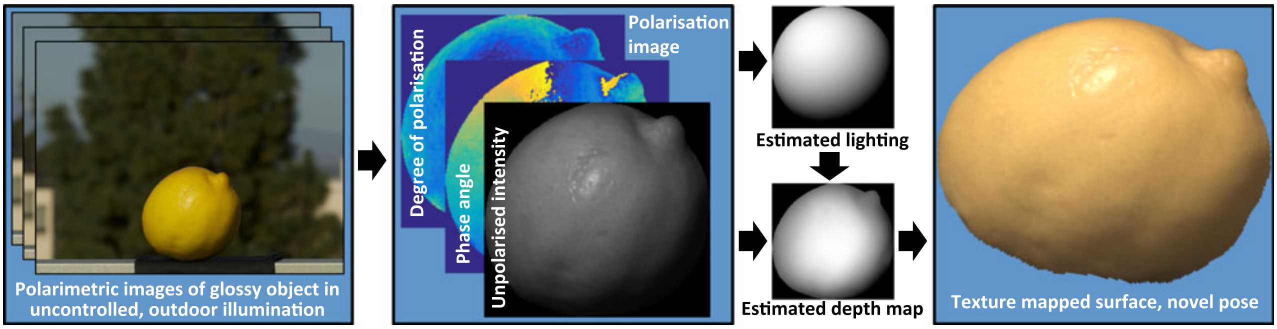


Fig. 1. Overview of method: from a single polarisation image in unknown (possibly outdoor) illumination, we estimate lighting and compute surface height directly (rightmost image shows result on real data, a piece of fruit).

show that per-pixel albedo can be determined from the ambiguous normals and the unpolarised intensity. Finally, in Section 7, we introduce a novel hybrid diffuse/specular polarisation and shading model, allowing us to handle glossy surfaces. Experimental results on synthetic and real data are reported in Sections 8 and 9 provides conclusions and future perspectives.

Although we make a variety of assumptions, the resulting methods are still useful in practice. Combining the methods in Sections 4, 5 and 7, our approach can be applied to glossy objects under uncalibrated directional illumination. In practice, this means that the method works outdoors on a sunny day (see Figs. 1 and 11) or indoors in a dark room setting (see Figs. 9, 10 and 12). In the former case, sunlight can be approximated by a point source and skylight can be neglected since it is orders of magnitude weaker. In the latter case, we require only a single uncalibrated light source and so the practical requirements are much less than for methods such as photometric stereo [7] or those that require multiple polarised light sources [8]. Other more niche applications could include polarised laparoscopy [9] or in general biomedical applications [10].

2 Related Work

Previous SfP methods can be categorised into three groups: 1. those that use only polarisation information, 2. those that combine polarisation with shading cues and 3. those that combine a polarisation image with an additional cue. Those techniques that require only a single polarisation image (of which our proposed method is one) are passive and can be considered 'single shot' methods (single shot capture devices exist using either polarising beamsplitters[1] or by combining micropolarisation filters with CMOS sensors[2]). More commonly, a polarisation image is obtained by capturing a sequence of images in which a linear polarising filter is rotated in front of the camera (possibly with unknown rotation angles [11]). SfP methods can also be classified according to the polarisation model (dielectric versus metal, diffuse, specular or hybrid models) and whether they compute shape in the surface normal or surface height domain.

Shape-from-polarisation. The earliest work focussed on capture, decomposition and visualisation of polarisation images was by Wolff [12]. Both Miyazaki et al. [4] and Atkinson and Hancock [3] used a diffuse polarisation model with assumed known refractive index to estimate surface normals from the phase angle and degree of polarisation. Disambiguation begins on the object boundary by choosing the azimuth angle that best aligns with the outward facing direction (an implicit assumption of object convexity). The disambiguation is then propagated inwards such that smoothness is maximised. This greedy approach will not produce globally optimal results, limits application to objects with a visible occluding boundary and does not consider integrability constraints. Morel et al. [13] took a similar approach but used a specular polarisation model suitable for metallic surfaces. Huynh et al. [14] also assumed convexity to disambiguate the polarisation normals; however, their approach can also estimate unknown refractive index.

Shape-from-polarisation and Shading. A polarisation image contains an unpolarised intensity channel which provides a shading cue. As in our proposed method, Mahmoud et al. [15] exploited this via a shape-from-shading cue. With assumptions of known light source direction, known albedo and Lambertian reflectance, the surface normal ambiguity can be resolved. We avoid all three of these assumptions and, by strictly enforcing integrability, impose an additional constraint that improves robustness to noise. An earlier version of the work in this paper was originally presented in [16]. Here, we have extended the method to handle unknown, spatially varying albedo and introduced an explicit specular reflectance model.

An alternative is to augment a polarisation image with additional intensity images in which the light source direction varies, providing a photometric stereo cue. Such methods are no longer passive and usually require calibrated light sources. Atkinson and Hancock [17] used Lambertian photometric stereo to disambiguate polarisation normals. Recently, Ngo et al. [18] derived constraints that allowed surface normals, light directions and refractive index to be estimated from polarisation images under varying lighting. However, this approach requires at least 4 light directions in contrast to the single direction required by our method. Atkinson [19] combines calibrated two source photometric stereo with the phase information from polarisation and resolves ambiguities via a region growing process.

Polarisation with Additional Cues. Rahmann and Canterakis [2] combined a specular polarisation model with stereo cues. Similarly, Atkinson and Hancock [20] used polarisation normals to segment an object into patches, simplifying stereo matching. Stereo polarisation cues have also been used for

1. http://www.fluxdata.com/imaging-polarimeters
2. https://www.4dtechnology.com/products/polarimeters/

transparent surface modelling [21]. Huynh et al. [22] extended their earlier work to use multispectral measurements to estimate both shape and refractive index. Drbohlav and Sara [23] showed how the Bas-relief ambiguity [6] in uncalibrated photometric stereo could be resolved using polarisation. However, this approach requires a polarised light source. Coarse geometry obtained by multi-view space carving [24], [25] has been used to resolve polarisation ambiguities. Kadambi et al. [5], [26] combined a single polarisation image with a depth map obtained by an RGBD camera. The depth map is used to disambiguate the normals and provide a base surface for integration. Cui et al. [27] used multiview stereo with a mixed polarisation model. A coarse reconstruction is provided by structure-from-motion which is used to partially disambiguate polarisation phase information. The remaining ambiguity is resolved as the phase information is propagated through a dense, multiview stereo surface reconstruction. This approach does not exploit degree of polarisation or shading information.

3 Preliminaries

In this section we list the basic assumptions common to all the following sections, we introduce the notations we will adopt throughout the whole paper and we explain how we construct our data, which is a polarisation image [12].

3.1 Assumptions

Our method relies on several assumptions. The following are assumed throughout the whole paper:

1) Orthographic camera projection
2) Smooth (i.e. C^2 continuous) object
3) Dielectric (i.e. non-metallic) material
4) Refractive index known
5) Illumination is provided by a distant point source
6) No interreflections.

Some later sections make additional assumptions. These are listed in the relevant section.

3.2 Notations

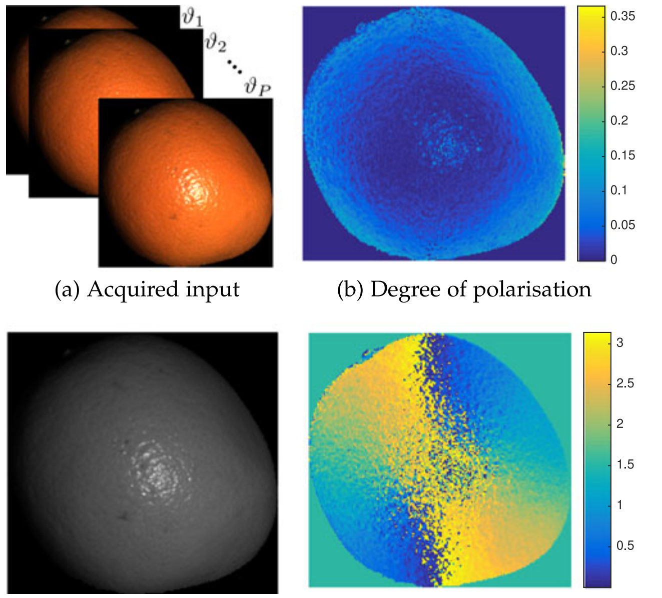
We parameterise surface height by the function $z(\mathbf{u})$, where $\mathbf{u} = (x, y)$ is an image point. Foreground pixels belonging to the surface are represented by the set $\mathcal{F}$, $|\mathcal{F}| = K$. We denote the unit surface normal by $\mathbf{n}(\mathbf{u})$. This vector can be expressed in spherical world coordinates as

$$\mathbf{n}(\mathbf{u}) = \begin{bmatrix} n_x(\mathbf{u}) \\ n_y(\mathbf{u}) \\ n_z(\mathbf{u}) \end{bmatrix} = \begin{bmatrix} \sin(\alpha(\mathbf{u}))\sin(\theta(\mathbf{u})) \\ \cos(\alpha(\mathbf{u}))\sin(\theta(\mathbf{u})) \\ \cos(\theta(\mathbf{u})) \end{bmatrix}, \tag{1}$$

where $\alpha(\mathbf{u})$ and $\theta(\mathbf{u})$ are the azimuth and zenith angle respectively. The surface normal can be formulated via the surface gradient as follows

$$\mathbf{n}(\mathbf{u}) = \frac{[-p(\mathbf{u}), -q(\mathbf{u}), 1]^T}{\sqrt{p(\mathbf{u})^2 + q(\mathbf{u})^2 + 1}}, \tag{2}$$

where $p(\mathbf{u}) = \partial_x z(\mathbf{u})$ and $q(\mathbf{u}) = \partial_y z(\mathbf{u})$, so that $\nabla z(\mathbf{u}) = [p(\mathbf{u}), q(\mathbf{u})]^T$.



(a) Acquired input (b) Degree of polarisation

(c) Unpolarised intensity (d) Phase angle

Fig. 2. Polarimetric capture (a) and decomposition to polarisation image (b-d) from captured data of a piece of fruit.

3.3 Polarisation Image

When unpolarised light is reflected from a surface, it becomes partially polarised. There are a number of mechanisms by which this process occurs. The two models that we use are described in Sections 4.3 and 7.3 and are suitable for dielectric materials. A *polarisation image* (Figs. 2b, 2c, and 2d) can be estimated by capturing a sequence of images (Fig. 2a) in which a linear polarising filter in front of the camera is rotated through a sequence of $P \geq 3$ different angles ϑ_j, $j \in \{1, \ldots, P\}$. The measured intensity at a pixel varies sinusoidally with the polariser angle

$$i_{\vartheta_j}(\mathbf{u}) = i_{\text{un}}(\mathbf{u})\big(1 + \rho(\mathbf{u})\cos\left[2\vartheta_j - 2\phi(\mathbf{u})\right]\big) + \tau. \tag{3}$$

The three parameters of the sinusoid form the three quantities of a polarisation image [12]. These are the *phase angle*, $\phi(\mathbf{u})$, the *degree of polarisation*, $\rho(\mathbf{u})$, and the *unpolarised intensity*, $i_{\text{un}}(\mathbf{u})$. The quantity τ models a stochastic process representing quantisation, sensor noise etc.

Under the assumption that τ is normally distributed, a least squares fit to the measured data provides the maximum likelihood solution for the three parameters of the sinusoid. In practice, this can be done using nonlinear least squares [3], linear methods [14] or via a closed form solution [12] for the specific case of $P = 3$, $\vartheta \in \{0°, 45°, 90°\}$.

4 Linear Height-from-Polarisation

In this section we show how to directly estimate a surface height map from a single polarisation image. Moreover, we show how this can be formulated as a sparse linear least squares problem for which the globally optimal solution can be computed efficiently.

4.1 Additional Assumptions

Throughout the whole Section 4, we require the following assumptions in addition to those introduced in Section 3.1

7) Lambertian reflectance and diffuse polarisation
8) Known or uniform albedo
9) Known point light source
10) Light and viewing directions different, i.e. $\mathbf{s} \neq \mathbf{v}$.

Assumptions 7-9 will be subsequently relaxed in Sections 5, 6 and 7.

4.2 Finite Difference Formulation

The surface gradient can be approximated numerically from the discretised surface height function by finite differences. If the surface heights are written as a vector $\mathbf{z} \in \mathbb{R}^K$, then the gradients, $\mathbf{g} \in \mathbb{R}^{2K}$, can be approximated by

$$\mathbf{g} = \begin{bmatrix} p(\mathbf{u}_1) \\ \vdots \\ p(\mathbf{u}_K) \\ q(\mathbf{u}_1) \\ \vdots \\ q(\mathbf{u}_K) \end{bmatrix} = \begin{bmatrix} \mathbf{D}_x \\ \mathbf{D}_y \end{bmatrix} \begin{bmatrix} z(\mathbf{u}_1) \\ \vdots \\ z(\mathbf{u}_K) \end{bmatrix} = \mathbf{D}\mathbf{z}, \tag{4}$$

where $\mathbf{D}_x \in \mathbb{R}^{K\times K}$ and $\mathbf{D}_y \in \mathbb{R}^{K\times K}$ evaluate the finite difference gradients in the horizontal and vertical directions respectively. Each row of $\mathbf{D}$ computes one gradient. In the simplest case, this could be done using forward differences in which case only two elements of the row are non-zero.

Hence, given a system of equations that are linear in the unknown surface gradients, $\mathbf{A}\mathbf{g} = \mathbf{b}$, this can be rewritten as a system of equations that are linear in the unknown surface height as $\mathbf{A}\mathbf{D}\mathbf{z} = \mathbf{b}$. Regardless of which finite difference approximation is used, $\mathrm{rank}(\mathbf{D}) = K - 1$. This reflects the fact that constraints on the surface gradient alone can only recover orthographic surface height up to a translation in z, i.e. the constant of integration is unknown. So, even if $\mathbf{A}$ is full rank, $\mathbf{A}\mathbf{D}$ is not and so $\mathbf{z}$ cannot be estimated from this set of equations alone. This is easily resolved by introducing an additional equation that, for example, sets the mean height to zero

$$\begin{bmatrix} \mathbf{A}D \\ \mathbf{1}_K \end{bmatrix} \mathbf{z} = \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix}, \tag{5}$$

where $\mathbf{1}_K$ is the length K row vector of ones.

4.3 Diffuse Polarisation Model

A polarisation image provides a constraint on the surface normal direction at each pixel. The exact nature of the constraint depends on the polarisation model used. We begin by assuming a diffuse polarisation model [3]. Diffuse polarisation arises due to subsurface scattering. Here, the Fresnel transmission out of the surface results in partial polarisation of the light. Exploitation of this cause of polarisation has the advantage that we do not need to assume that the illumination is unpolarised. Subsurface scattering has a depolarising effect such that the polarisation of the remitted light can be assumed to have arisen entirely due to transmission out of the surface.

For diffuse reflection, the degree of polarisation is related (Fig. 4a, red curve) to the zenith angle $\theta(\mathbf{u}) \in [0, \frac{\pi}{2}]$ of the normal in viewer-centred coordinates (i.e. the angle between the normal and viewer)

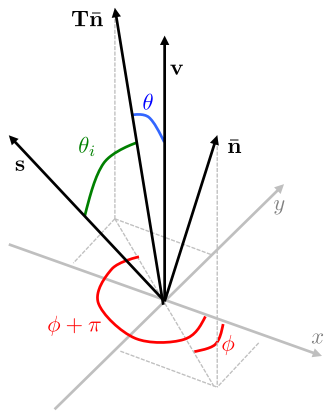


Fig. 3. Visualisation of constraints on surface normal provided by polarisation image: phase angle (red), unpolarised intensity (green) and degree of polarisation (blue). In non-degenerate cases, the three constraints uniquely determine the surface normal direction and we show how to express these constraints directly in terms of surface height.

$$\rho(\mathbf{u}) = \frac{\sin(\theta(\mathbf{u}))^2 \left(\eta - \frac{1}{\eta}\right)^2}{4\cos(\theta(\mathbf{u}))\sqrt{\eta^2 - \sin(\theta(\mathbf{u}))^2} - \sin(\theta(\mathbf{u}))^2 \left(\eta + \frac{1}{\eta}\right)^2 + 2\eta^2 + 2}, \tag{6}$$

where η is the refractive index. The dependency on η is weak [3] and typical values for dielectrics range between 1.4 and 1.6. We assume $\eta = 1.5$ for the rest of this paper. This expression can be rearranged to give a closed form solution for the zenith angle in terms of a function, $f(\rho(\mathbf{u}), \eta)$, that depends on the measured degree of polarisation and the refractive index

$$\cos(\theta(\mathbf{u})) = \mathbf{n}(\mathbf{u}) \cdot \mathbf{v} = f(\rho(\mathbf{u}), \eta) = \sqrt{\frac{\eta^4(1-\rho^2) + 2\eta^2(2\rho^2+\rho-1) + \rho^2 + 2\rho - 4\eta^3\rho\sqrt{1-\rho^2} + 1}{(\rho+1)^2(\eta^4+1) + 2\eta^2(3\rho^2+2\rho-1)}}, \tag{7}$$

where we drop the dependency of ρ on $\mathbf{u}$ for brevity. Since we work in a viewer-centred coordinate system, the viewing direction is $\mathbf{v} = [0, 0, 1]^T$ and we have simply: $n_z(\mathbf{u}) = f(\rho(\mathbf{u}), \eta)$, or, in terms of the surface gradient,

$$\frac{1}{\sqrt{p(\mathbf{u})^2 + q(\mathbf{u})^2 + 1}} = f(\rho(\mathbf{u}), \eta). \tag{8}$$

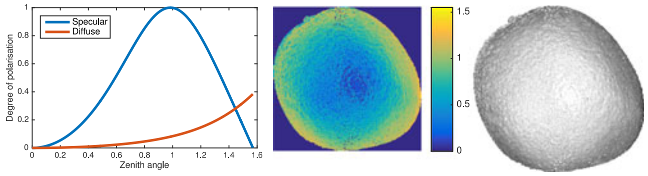
The phase angle determines the azimuth angle of the surface normal $\alpha(\mathbf{u}) \in [0, 2\pi]$ up to a 180° ambiguity: $\alpha(\mathbf{u}) = \phi(\mathbf{u})$ or $(\phi(\mathbf{u}) + \pi)$. This means that the measured degree of polarisation (via (7)) and phase angle determine the surface normal up to an ambiguity as either $\mathbf{n}(\mathbf{u}) = \bar{\mathbf{n}}(\mathbf{u})$ or $\mathbf{n}(\mathbf{u}) = \mathbf{T}\bar{\mathbf{n}}(\mathbf{u})$ where

$$\bar{\mathbf{n}}(\mathbf{u}) = \begin{bmatrix} \sin(\phi(\mathbf{u}))\sin(\theta(\mathbf{u})) \\ \cos(\phi(\mathbf{u}))\sin(\theta(\mathbf{u})) \\ \cos(\theta(\mathbf{u})) \end{bmatrix}, \tag{9}$$

and

$$\mathbf{T} = \mathbf{R}_z(180^\circ) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{10}$$

See Fig. 3 for a visualisation of these two constraints (shown in red and blue).



(a) Degree of polarisation (b) Estimated θ (c) $\cos(\theta)$

Fig. 4. (a) Relationship between degree of polarisation and zenith angle, for specular and diffuse dielectric reflectance with $\eta = 1.5$. (b) Zenith angle estimated from Fig. 2b. (c) Visualisation of the cosine of estimated zenith angle.

4.4 Shading Constraint

The unpolarised intensity provides an additional constraint on the surface normal direction via an appropriate reflectance model. Following Assumption 7, we use the Lambertian model and from Assumption 8, albedo is either: 1. known and has been divided out, or 2. uniform and factored into the light source vector $\mathbf{s} \in \mathbb{R}^3$. Hence, unpolarised intensity is related to the surface normal by

$$i_{\text{un}}(\mathbf{u}) = \cos(\theta_i(\mathbf{u})) = \mathbf{n}(\mathbf{u}) \cdot \mathbf{s}, \tag{11}$$

where $\theta_i(\mathbf{u})$ is the angle of incidence (angle between light source and surface normal). In terms of the surface gradient, this becomes

$$i_{\text{un}}(\mathbf{u}) = \frac{-p(\mathbf{u})s_x - q(\mathbf{u})s_y + s_z}{\sqrt{p(\mathbf{u})^2 + q(\mathbf{u})^2 + 1}}. \tag{12}$$

Note that if the light source and viewer direction coincide then this equation provides no more information than the degree of polarisation. This explains the need for Assumption 10. The addition of the shading cue uniquely determines the surface normal at a pixel (see Fig. 3, shading cue shown in green; in this example the solution is $\mathbf{T}\bar{\mathbf{n}}$).

4.5 Polarisation Constraints as Linear Equations

In practice, the polarisation image quantities will be noisy and an exact solution may not exist. A least squares solution at each pixel independently leads to surface normal estimates that are first noisy and second will not satisfy the integrability constraint. Both of these problems can be addressed by posing the problem in terms of estimating surface height and solving a system of equations globally. With this goal in mind, we start by showing that the polarisation shape cues can be expressed as per pixel equations that are linear in terms of the surface gradient.

First, we note that the phase angle constraint can be written as a collinearity condition. This condition is satisfied by either of the two possible azimuth angles implied by the phase angle measurement. Writing it in this way is advantageous because it means we do not have to disambiguate the surface normals explicitly. Instead, when we solve the linear system for height, the azimuthal ambiguities are resolved in a globally optimal way. Specifically, we require the projection of the surface normal into the x-y plane, $[n_x, n_y]$, and a vector in the image plane pointing in the phase angle direction, $[\sin(\phi), \cos(\phi)]$, to be collinear. These two vectors are collinear when the following condition is satisfied:

$$\mathbf{n}(\mathbf{u}) \cdot [\cos(\phi(\mathbf{u})), -\sin(\phi(\mathbf{u})), 0]^T = 0. \tag{13}$$

Substituting (2) into (13), we obtain

$$\frac{-p(\mathbf{u})\cos(\phi(\mathbf{u})) + q(\mathbf{u})\sin(\phi(\mathbf{u}))}{\sqrt{p(\mathbf{u})^2 + q(\mathbf{u})^2 + 1}} = 0. \tag{14}$$

Noting that the nonlinear term in (2) is always greater than zero, we obtain our first linear equation in the surface gradient

$$-p(\mathbf{u})\cos(\phi(\mathbf{u})) + q(\mathbf{u})\sin(\phi(\mathbf{u})) = 0. \tag{15}$$

This condition exhibits a natural weighting that is useful in practice. The phase angle estimates are more reliable when the zenith angle is large (i.e. when the degree of polarisation is high and so the signal to noise ratio is high). When the zenith angle is large, the magnitude of the surface gradient is large, meaning that disagreement with the estimated phase angle is penalised more heavily than for a small zenith angle where the gradient magnitude is small.

The second linear constraint is obtained by combining the expressions for the unpolarised intensity and the degree of polarisation. To do so, we take a ratio between (12) and (8) which eliminates the nonlinear normalisation factor

$$\frac{i_{\text{un}}(\mathbf{u})}{f(\rho(\mathbf{u}), \eta)} = -p(\mathbf{u})s_x - q(\mathbf{u})s_y + s_z, \tag{16}$$

yielding our second linear equation in the surface gradient.

4.6 Linear Least Squares Formulation

We can now write the polarisation constraints in Section 4.5 as a linear system of equations in terms of the unknown surface height, $\mathbf{ADz} = \mathbf{b}$, where

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_c & \mathbf{A}_s \\ -s_x\mathbf{I}_K & -s_y\mathbf{I}_K \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \mathbf{0}_K \\ i_{\text{un}}(\mathbf{u}_1)/f(\rho(\mathbf{u}_1), \eta) - s_z \\ \vdots \\ i_{\text{un}}(\mathbf{u}_K)/f(\rho(\mathbf{u}_K), \eta) - s_z \end{bmatrix}, \tag{17}$$

$$\mathbf{A}_c = \text{diag}(-\cos\phi(\mathbf{u}_1), \dots, -\cos\phi(\mathbf{u}_K)), \tag{18}$$

$$\mathbf{A}_s = \text{diag}(\sin\phi(\mathbf{u}_1), \dots, \sin\phi(\mathbf{u}_K)), \tag{19}$$

$\mathbf{0}_K$ is the length K zero vector and $\mathbf{I}_K$ is the $K \times K$ identity matrix. The upper half of $\mathbf{A}$ evaluates the phase angle linear Equation (15) and the lower half evaluates the shading/degree of polarisation ratio linear Equation (16).

In general, $\mathbf{A}$ is full rank and, in the presence of no noise, a unique, exact solution to (5) exists. From a theoretical perspective, $\mathbf{A}$ is rank deficient in the special case where $s_x = -s_y \neq 0$ and $\phi = \pi/4$ in at least one pixel.

In practice, the polarisation image and light source vector will be noisy. Hence, we do not expect an exact solution and formulate a least squares cost function for $\mathbf{z}$

$$\varepsilon_{\text{data}}(\mathbf{z}) = \left\| \begin{bmatrix} \mathbf{A}D \\ \mathbf{1}_K \end{bmatrix} \mathbf{z} - \begin{bmatrix} \mathbf{b} \\ 0 \end{bmatrix} \right\|^2. \tag{20}$$

For robust performance on real world data, we find it advantageous (though not essential) to include two priors on the surface height that are explained in the following sections.

4.7 Laplacian Smoothness Prior

The first prior is a Laplacian smoothness term. This takes the form of a smoothness penalty, $\varepsilon_{\mathrm{sm}}$

$$\varepsilon_{\mathrm{sm}}(\mathbf{z}) = \|w_{\mathrm{sm}}\mathbf{L}\mathbf{z}\|^2, \tag{21}$$

where w_{sm} weights the influence of the prior and $\mathbf{L} \in \mathbb{R}^{C\times K}$ is a matrix, each row of which evaluates the convolution of a 3×3 Laplacian kernel with one of the $C \leq K$ pixels whose local 3×3 neighbourhood is included in $\mathcal{F}$. This prior encourages a pixel to have a height close to the average of its neighbours. It is minimised by locally planar regions, so can lead to oversmoothing of curved regions, but has the advantage of being linear in the surface height.

4.8 Convexity Prior

The second prior (applicable only to objects with a foreground mask) is a convexity prior that encourages the azimuth angle of the surface normal to align with the azimuth of outward facing boundary normals. This is helpful for data that is noisy close to the occluding boundary, for example when some background is included in the image due to an inaccurate foreground mask.

We compute unit vectors in the image plane that are normal to the boundary and outward facing and propagate these vectors into the interior. We convert these vectors to boundary-implied azimuth angles, $\alpha_b(\mathbf{u})$. See supplementary material, which can be found on the Computer Society Digital Library at http://doi.ieeecomputersociety.org/10.1109/TPAMI.2018.2868065, for details. Now, to exploit this prior we penalise deviation in the azimuth angle of the estimated surface normals from those provided by the boundary cue, $\alpha_b(\mathbf{u})$. We wish to measure this deviation in a way that is linear in the unknown surface gradients. To achieve this, we construct a surface normal vector $\mathbf{n}_b(\mathbf{u})$ using $\alpha_b(\mathbf{u})$ and the zenith angle estimated by polarisation, $\theta(\mathbf{u})$ (using (7))

$$\mathbf{n}_b(\mathbf{u}) = [\sin\alpha_b(\mathbf{u})\sin\theta(\mathbf{u}),\ \cos\alpha_b(\mathbf{u})\sin\theta(\mathbf{u}),\ \cos\theta(\mathbf{u})]^T. \tag{22}$$

Combining (2) and (22) and rearranging, we can express the surface derivatives according to $\mathbf{n}_b(\mathbf{u})$ as

$$p(\mathbf{u}) = \frac{\sin\alpha_b(\mathbf{u})\sin\theta(\mathbf{u})}{\cos\theta(\mathbf{u})} \quad \text{and} \quad q(\mathbf{u}) = \frac{\cos\alpha_b(\mathbf{u})\sin\theta(\mathbf{u})}{\cos\theta(\mathbf{u})}. \tag{23}$$

For numerical stability, we multiply both sides of these equations by $\cos\theta(\mathbf{u})$ (this avoids the magnitude of the equation becoming very large when $\theta(\mathbf{u})$ is close to $\pi/2$). Finally, we weight this prior such that it has high influence close to the boundary but the weight falls off rapidly as distance to the boundary increases. The per-pixel weights are defined as follows:

$$w_{\mathrm{con}}(\mathbf{u}) = \left(\frac{[\max_{\mathbf{v}\in\mathcal{F}} d_b(\mathbf{v})] - d_b(\mathbf{u})}{\max_{\mathbf{v}\in\mathcal{F}} d_b(\mathbf{v})}\right)^m \in [0,1], \tag{24}$$

where $d_b(\mathbf{u})$ is the euclidean distance from $\mathbf{u}$ to the boundary pixel closest to $\mathbf{u}$. The scalar m determines how quickly the weight reduces with distance from the boundary.

We can now compute a cost that measures the discrepancy between the gradients of the reconstructed surface, $\mathbf{g} = \mathbf{D}\mathbf{z}$, and those implied by the boundary normal (23), weighted by (24)

$$\begin{aligned}\varepsilon_{\mathrm{con}}(\mathbf{z}) = \sum_{i=1}^{K} w_{\mathrm{con}}(\mathbf{u}_i)^2\big[&\big(\mathbf{g}_i\cos\theta(\mathbf{u}_i) - \sin\alpha_b(\mathbf{u}_i)\sin\theta(\mathbf{u}_i)\big)^2 \\ &+ \big(\mathbf{g}_{K+i}\cos\theta(\mathbf{u}_i) - \cos\alpha_b(\mathbf{u}_i)\sin\theta(\mathbf{u}_i)\big)^2\big].\end{aligned} \tag{25}$$

4.9 Implementation

We can now combine the height-from-polarisation cost (20) with the cost functions associated with the two priors (21), (25) to form a single system of equations in linear least squares form

$$\varepsilon(\mathbf{z}) = \varepsilon_{\mathrm{data}}(\mathbf{z}) + \varepsilon_{\mathrm{sm}}(\mathbf{z}) + \varepsilon_{\mathrm{con}}(\mathbf{z}) = \left\| \begin{bmatrix} \begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix}\mathbf{D} \\ w_{\mathrm{sm}}\mathbf{L} \\ \mathbf{1}_K \end{bmatrix} \mathbf{z} - \begin{bmatrix}\mathbf{b}\\ \mathbf{c}\\ \mathbf{0}_C\\ 0\end{bmatrix} \right\|^2, \tag{26}$$

where

$$\mathbf{B} = \begin{bmatrix} \mathrm{diag}(w_{\mathrm{con}}(\mathbf{u}_1)\cos\theta(\mathbf{u}_1),\ldots,w_{\mathrm{con}}(\mathbf{u}_K)\cos\theta(\mathbf{u}_K)) \\ \mathrm{diag}(w_{\mathrm{con}}(\mathbf{u}_1)\cos\theta(\mathbf{u}_1),\ldots,w_{\mathrm{con}}(\mathbf{u}_K)\cos\theta(\mathbf{u}_K)) \end{bmatrix}, \tag{27}$$

$$\mathbf{c} = \begin{bmatrix} w_{\mathrm{con}}(\mathbf{u}_1)\sin\alpha_b(\mathbf{u}_1)\sin\theta(\mathbf{u}_1) \\ \vdots \\ w_{\mathrm{con}}(\mathbf{u}_K)\sin\alpha_b(\mathbf{u}_K)\sin\theta(\mathbf{u}_K) \\ w_{\mathrm{con}}(\mathbf{u}_1)\cos\alpha_b(\mathbf{u}_1)\sin\theta(\mathbf{u}_1) \\ \vdots \\ w_{\mathrm{con}}(\mathbf{u}_K)\cos\alpha_b(\mathbf{u}_K)\sin\theta(\mathbf{u}_K) \end{bmatrix}. \tag{28}$$

Finally, we solve for the optimal height map using linear least squares

$$\mathbf{z}^* = \arg\min_{\mathbf{z}\in\mathbb{R}^K} \varepsilon(\mathbf{z}). \tag{29}$$

Although the system of equations is large, it is sparse and so can be solved efficiently. We use a sparse QR solver. For the height derivative operator, $\mathbf{D}$, for each row we compute a smoothed central difference approximation of the derivative equivalent to convolving the height values with a Gaussian kernel and then convolving with the central difference kernel. At the boundary of the image or the foreground mask, not all neighbours may be available for a given pixel. In this case, we use unsmoothed central differences (where both horizontal or both vertical neighbours are available) or, where only a single neighbour is available, single forward/backward differences. We use a value of $w_{\mathrm{sm}} = 0.1$ and $m = 5$ in all of our experiments.

5 Illumination Estimation from an Uncalibrated Polarisation Image

In this section, we describe how to use the polarisation image to estimate illumination, assuming uniform albedo. Hence, we retain the same assumptions as the previous section but remove Assumption 9. This means that our SfP method described in Section 4 can be applied in an uncalibrated lighting scenario. We start by showing that the problem of light source estimation is subject to an ambiguity. Next, we derive a method to compute the light source direction (up to the ambiguity) from ambiguous normals using the minimum possible number of observations. Finally, we extend this to an efficient optimisation approach that uses the whole image and is applicable to noisy data.

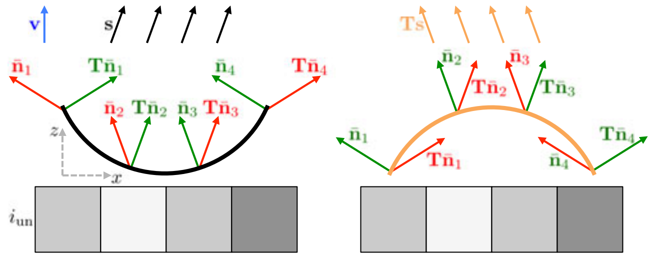


Fig. 5. Illustration of ambiguity using a 1D surface viewed from above. Polarisation normals are locally ambiguous (green versus red), leading to 2^4 possible disambiguations. With unknown lighting direction, the introduction of shading information reduces the ambiguity to a global, binary one. For the shading images at the bottom, the two possible disambiguations are black versus orange with the resulting local disambiguations shown in green.

5.1 Relationship to the Bas-relief Ambiguity

From the measured degree of polarisation and phase angle, the surface normal at a pixel can be estimated up to a *local* binary ambiguity via (9) and (10) (see green versus red in Fig. 5). Hence, there are 2^K possible disambiguations of the polarisation normals in a K pixel image. In Section 4.4, we showed how shading information can be used to resolve this ambiguity locally if the light source direction is known (see Fig. 3). We now consider the setting in which the light source direction is unknown.

For the true light source direction, $\mathbf{s}$, one of the following equalities holds:

$$i_{\text{un}}(\mathbf{u}) = \bar{\mathbf{n}}(\mathbf{u}) \cdot \mathbf{s} \quad \text{or} \quad i_{\text{un}}(\mathbf{u}) = (\mathbf{T}\bar{\mathbf{n}}(\mathbf{u})) \cdot \mathbf{s}. \tag{30}$$

Hence, the polarisation measurements for a single pixel place one of two possible linear constraints on $\mathbf{s}$, depending on which disambiguation of the surface normal is chosen.

Suppose that we know the correct disambiguation of the normals and that we stack them to form the matrix $\mathbf{N}_{\text{true}} \in \mathbb{R}^{K\times 3}$ and stack the unpolarised intensities in the vector $\mathbf{i} = [i_{\text{un}}(\mathbf{u}_1) \; \dots \; i_{\text{un}}(\mathbf{u}_K)]^T$. In this case, the light source $\mathbf{s}$ that satisfies $\mathbf{N}_{\text{true}}\mathbf{s} = \mathbf{i}$ is given by

$$\mathbf{s} = \mathbf{N}_{\text{true}}^{+}\mathbf{i}, \tag{31}$$

where $\mathbf{N}_{\text{true}}^{+}$ is the pseudoinverse of $\mathbf{N}_{\text{true}}$. However, for any invertible 3×3 linear transform $\mathbf{G} \in GL(3)$, it is also true that $\mathbf{N}_{\text{true}}\mathbf{G}^{-1}\mathbf{G}\mathbf{s} = \mathbf{i}$, and so $\mathbf{G}\mathbf{s}$ is also a solution using the transformed normals $\mathbf{N}_{\text{true}}\mathbf{G}^{-1}$. The only such $\mathbf{G}$ where $\mathbf{N}_{\text{true}}\mathbf{G}^{-1}$ would remain consistent with the zenith and phase angles implied by the polarisation image is $\mathbf{G} = \mathbf{T}$, i.e. where the azimuth angle of every true surface normal is shifted by π. Hence, if we did not know the correct disambiguation then $\mathbf{s}$ is a solution with normals $\mathbf{N}_{\text{true}}$ but $\mathbf{T}\mathbf{s}$ is also a solution with normals $\mathbf{N}_{\text{true}}\mathbf{T}$. Note that $\mathbf{T}$ is a generalised Bas-relief (GBR) transformation [6] with parameters $\mu = 0$, $\nu = 0$ and $\lambda = \pm 1$. In other words, it corresponds to the binary convex/concave ambiguity. Hence, from a polarisation image with unknown lighting, we will be unable to distinguish the true normals and lighting from those transformed by $\mathbf{T}$. Since $\mathbf{T}$ is a GBR transformation, the transformed normals remain integrable and correspond to a negation of the true surface. This is a *global*, binary ambiguity. In Fig. 5, either the black or orange interpretation corresponds to $\mathbf{N}_{\text{true}}$, but from the polarisation image alone we do not know which. To transform from black to orange or vice versa, *all* the normals are transformed by $\mathbf{T}$.

5.2 Minimal Solutions

In practice, we will not have the correct disambiguations to hand. We consider the minimum number of observations necessary to find the light source direction (up to the binary ambiguity) when only the ambiguous polarisation normals are known. Suppose that $\mathbf{N} \in \mathbb{R}^{K\times 3}$ contains one of the 2^K possible disambiguations of the K surface normals, i.e. $\mathbf{N}_j = \bar{\mathbf{n}}(\mathbf{u}_j)$ or $\mathbf{N}_j = \mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_j)$. If $\mathbf{N}$ is a valid disambiguation (i.e. $\mathbf{N} = \mathbf{N}_{\text{true}}$ or $\mathbf{N} = \mathbf{N}_{\text{true}}\mathbf{T}$), then (with no noise) we expect: $\mathbf{N}\mathbf{s} = \mathbf{N}\mathbf{N}^{+}\mathbf{i} = \mathbf{i}$. We can see in a straightforward way that three pixels will be insufficient to distinguish a valid from an invalid disambiguation. When $K = 3$, $\mathbf{N}^{+} = \mathbf{N}^{-1}$ and so $\mathbf{N}\mathbf{N}^{+} = \mathbf{I}_3$ and hence the condition is satisfied by any combination of disambiguations. The reason for this is that $\mathbf{s}$ has three degrees of freedom and so, apart from degenerate cases, any three linear equations in $\mathbf{s}$ will have a solution, i.e. any combination of transformed or untransformed normals will allow an $\mathbf{s}$ to be found that satisfies all three equations.

However, the problem becomes well posed for $K > 3$. We now require that the system of linear equations is consistent and has a unique solution. If some, but not all, of the normals are transformed from their true directions then the system of equations will be inconsistent. By the Rouché–Capelli theorem[3] [28], consistency and uniqueness requires $\text{rank}(\mathbf{N}) = \text{rank}([\mathbf{N}\;\mathbf{i}]) = 3$. This suggests an approach for simultaneous disambiguation and light source estimation for the minimal case of $K = 4$. We consider each of the 16 possible normal matrices $\mathbf{N}$ in turn until we find one satisfying the rank condition. For this $\mathbf{N}$ we find $\mathbf{s}$ by (31) and the true light source is either $\mathbf{s}$ or $\mathbf{T}\mathbf{s}$. The pseudocode for this approach is given in Algorithm 1.

Algorithm 1. Minimal Solution for Lighting

Inputs:
Vector of unpolarised intensities, $\mathbf{i} \in \mathbb{R}^4$
Ambiguous polarisation normals, $\bar{\mathbf{n}}_j \in \mathbb{R}^3$, $j \in \{1, \dots, 4\}$
Output: Estimated light source, $\mathbf{s} \in \mathbb{R}^3$

```
// Generate all binary strings⁴ of length 4
P := binaryStrings(4)
// P_{i,j} is the jth digit of the ith string
for i := 1 to 2^4 do
    // Generate ith disambiguation
    for j := 1 to 4 do
        N_j := { n̄_j    if P_{i,j} = 0
               { Tn̄_j   otherwise
    end for
    if rank(N) = rank([N i]) = 3 then
        s := N⁺i
        return s
    end if
end for
```

---

3. The Rouché–Capelli theorem states that a system of linear equations $\mathbf{Q}\mathbf{x} = \mathbf{y}$, $\mathbf{y} \in \mathbb{R}^d$, has a solution if and only if $\text{rank}(\mathbf{Q}) = \text{rank}([\mathbf{Q}\;\mathbf{y}])$ and the solution is unique if and only if $\text{rank}(\mathbf{Q}) = d$.

4 The function $\text{binaryStrings}(K)$ returns a $2^K \times K$ matrix containing all binary strings of length K such that each element of the matrix contains 0 or 1 and the ith row of the matrix contains the ith string.

5.3 Least Squares Combinatorial Lighting Estimation

With real data, we expect $\bar{\mathbf{n}}$ and $\mathbf{i}$ to be noisy. Therefore, the minimal system of equations corresponding to the correct disambiguation may not permit an exact solution. Instead, a least squares solution using all data is preferable. Following the combinatorial approach in Section 5.2, we could build all 2^K possible systems of linear equations, i.e.

$$\begin{gathered}
[\,\bar{\mathbf{n}}(\mathbf{u}_1) \quad \bar{\mathbf{n}}(\mathbf{u}_2) \quad \dots \quad \bar{\mathbf{n}}(\mathbf{u}_K)\,]^T\mathbf{s} = \mathbf{i}, \\
[\,\mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_1) \quad \bar{\mathbf{n}}(\mathbf{u}_2) \quad \dots \quad \bar{\mathbf{n}}(\mathbf{u}_K)\,]^T\mathbf{s} = \mathbf{i}, \\
[\,\bar{\mathbf{n}}(\mathbf{u}_1) \quad \mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_2) \quad \dots \quad \bar{\mathbf{n}}(\mathbf{u}_K)\,]^T\mathbf{s} = \mathbf{i}, \\
\vdots \\
[\,\mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_1) \quad \mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_2) \quad \dots \quad \mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_K)\,]^T\mathbf{s} = \mathbf{i},
\end{gathered} \tag{32}$$

solve them in a least squares sense and take the one with minimal residual as the solution. Pseudocode for this approach is given in Algorithm 2. However, this is NP-hard and impractical for any non-trivial value of K.

Algorithm 2. Least Squares Combinatorial Lighting Estimation

Inputs:
Vector of unpolarised intensities, $\mathbf{i} \in \mathbb{R}^K$, $K \geq 4$
Ambiguous polarisation normals, $\bar{\mathbf{n}}_j \in \mathbb{R}^3$, $j \in \{1, \dots, K\}$
Output: Estimated light source, $\mathbf{s}^* \in \mathbb{R}^3$

$\varepsilon^* := \infty$
$\mathbf{P} := \mathrm{binaryStrings}\,(K)$
for $i := 1$ **to** 2^K **do**
  for $j := 1$ **to** K **do**
    $\mathbf{N}_j := \begin{cases} \bar{\mathbf{n}}_j & \text{if } \mathbf{P}_{i,j} = 0 \\ \mathbf{T}\bar{\mathbf{n}}_j & \text{otherwise} \end{cases}$
  end for
  $\mathbf{s} := \mathbf{N}^{+}\mathbf{i}$
  $\varepsilon := \|\mathbf{N}\mathbf{s} - \mathbf{i}\|^2$
  if $\varepsilon < \varepsilon^*$ **then**
    $\varepsilon^* := \varepsilon$
    $\mathbf{s}^* := \mathbf{s}$
  end if
end for
return $\mathbf{s}^*$

5.4 Alternating Optimisation and Assignment

Since the unknown illumination is only 3D and we have a polarisation observation for every pixel, the systems of equations in (5.3) are highly over-constrained since $K \gg 3$, hence the least squares solutions are very robust. We can write a continuous optimisation problem whose global minima would coincide with the lowest residual system in (5.3)

$$\mathbf{s}^* = \underset{\mathbf{s}\in\mathbb{R}^3}{\arg\min} \sum_{j=1}^{K} \min\Big[r_j(\mathbf{s})^2, t_j(\mathbf{s})^2\Big], \tag{33}$$

where r_j is the residual with the untransformed normal

$$r_j(\mathbf{s}) = \bar{\mathbf{n}}(\mathbf{u}_j) \cdot \mathbf{s} - i_{\mathrm{un}}(\mathbf{u}_j), \tag{34}$$

and t_j the residual with the transformed normal

$$t_j(\mathbf{s}) = (\mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_j)) \cdot \mathbf{s} - i_{\mathrm{un}}(\mathbf{u}_j). \tag{35}$$

An expression of this form is non-convex since the minimum of two convex functions is not convex [29]. However, (33) can be efficiently optimised using alternating assignment and optimisation. We find that, in practice, this almost always converges to the global minimum even with a random initialisation. In the assignment step, given an estimate for the light source at iteration w, $\mathbf{s}^{(w)}$, we choose from each ambiguous pair of normals (i.e. $\bar{\mathbf{n}}$ or $\mathbf{T}\bar{\mathbf{n}}$) the one that yields minimal error under illumination $\mathbf{s}^{(w)}$

$$\mathbf{N}_j^{(w)} := \begin{cases} \bar{\mathbf{n}}(\mathbf{u}_j) & \text{if } r_j(\mathbf{s}^{(w)})^2 < t_j(\mathbf{s}^{(w)})^2, \\ \mathbf{T}\bar{\mathbf{n}}(\mathbf{u}_j) & \text{otherwise.} \end{cases} \tag{36}$$

At the optimisation step, we use the selected normals to compute the new light source by solving the linear least squares system via the pseudo-inverse

$$\mathbf{s}^{(w+1)} := (\mathbf{N}^{(w)})^{+}\mathbf{i}. \tag{37}$$

These two steps are iterated to convergence. In all our experiments, this converged in < 10 iterations. This approach can be extended to spherical harmonic illumination [16].

Note that the assignment step (36) disambiguates each surface normal *locally* (i.e. choosing between red and green in Fig. 5). The *global* convex/concave ambiguity described in Section 5.1 remains. To resolve this (i.e. to choose between black and orange in Fig. 5), we arbitrarily choose from the two possible light source directions the one that gives the surface height map with maximal volume.

The alternating optimisation procedure can be viewed as simultaneously estimating illumination *and shape*. Since the assignment step resolves the ambiguity at each pixel, upon convergence we have a surface normal estimate for each pixel. However, this does not perform well because the surface normal estimates use only local information, are made independently at each pixel and the integrability constraint is only imposed during surface integration. These factors motivate the global method proposed in Section 4.

6 Albedo Estimation from a Calibrated Polarisation Image

In Section 5, we assumed that albedo was uniform and estimated unknown lighting. We now present an alternative for the case of an object with spatially varying albedo. This requires that the illumination direction (but not necessarily its intensity) is known. Note that if we know only the direction of the illumination, but not its intensity, we can arbitrarily set $\|\mathbf{s}\| = 1$ and albedo is estimated up to an unknown global scale. Once albedo has been estimated, it can be divided out of the unpolarised intensity image and linear height estimation performed as in Section 4. We retain the same assumptions as Section 4.1 but can remove Assumption 8 since we now estimate spatially varying albedo.

6.1 Locally Ambiguous Albedo Estimation

Introducing a spatially varying albedo $a(\mathbf{u}) \in [0,1]$ to (30), the unpolarised intensity with no noise is given by

$$i_{\text{un}}(\mathbf{u}) = a(\mathbf{u})\bar{\mathbf{n}}(\mathbf{u}) \cdot \mathbf{s} \quad \text{or} \quad i_{\text{un}}(\mathbf{u}) = a(\mathbf{u})(\mathbf{T}\bar{\mathbf{n}}(\mathbf{u})) \cdot \mathbf{s}. \tag{38}$$

With illumination known, we can estimate the local albedo up to a binary ambiguity: $a(\mathbf{u}) = a_1(\mathbf{u})$ or $a_2(\mathbf{u})$ where

$$a_1(\mathbf{u}) = \frac{i_{\text{un}}(\mathbf{u})}{\bar{\mathbf{n}}(\mathbf{u}) \cdot \mathbf{s}}, \quad \text{and} \quad a_2(\mathbf{u}) = \frac{i_{\text{un}}(\mathbf{u})}{(\mathbf{T}\bar{\mathbf{n}}(\mathbf{u})) \cdot \mathbf{s}}. \tag{39}$$

Note that, for pixels where the light source lies on the plane bisecting the two possible surface normal directions, i.e. $\bar{\mathbf{n}}(\mathbf{u}) \cdot \mathbf{s} = (\mathbf{T}\bar{\mathbf{n}}(\mathbf{u})) \cdot \mathbf{s}$, the two expressions are equal and the albedo is well-defined. Note also that the bound can be tightened since $a(\mathbf{u}) \geq i_{\text{un}}(\mathbf{u})/\|\mathbf{s}\|$.

However, in general there will be two possible solutions. We cannot use the same approach as for lighting estimation where the unknown lighting vector is only 3D but every pixel provided a pair of possible constraints. Instead we must exploit spatial smoothness and solve an optimisation problem over the whole albedo map simultaneously. From Assumption 2 and since the diffuse shading function (11) is smooth, we can conclude that the shading itself is smooth with no further assumptions. To emphasise: we do not need to assume that the albedo itself is smooth.

6.2 Nonlinear Albedo Optimisation

The polarisation normals and, to a lesser extent, the lighting and unpolarised intensities will be noisy. Hence, neither of the two solutions in (39) may be a good estimate. For this reason, we pose albedo estimation as a nonlinear optimisation problem in which (39) is only a data term which need not be satisfied exactly

$$\varepsilon_{\text{data}}(a) = \sum_{\mathbf{u} \in \mathcal{F}} \min\Big[(a(\mathbf{u}) - a_1(\mathbf{u}))^2, (a(\mathbf{u}) - a_2(\mathbf{u}))^2\Big]. \tag{40}$$

As with the objective function for lighting estimation, this is non-convex. We augment the data term by a penalty that measures the smoothness of the shading implied by the estimated albedo, encouraging spatial smoothness of the solution. We evaluate this by convolving a Laplacian smoothing kernel with the implied shading, $\mathbf{d} \in \mathbb{R}^K$

$$\varepsilon_{\text{smooth}}(a) = \|\mathbf{L}\mathbf{d}\|^2, \quad \text{with } \mathbf{d}_i = i_{\text{un}}(\mathbf{u}_i)/a(\mathbf{u}_i), \tag{41}$$

where $\mathbf{L}$ performs the convolution, as in (21).

The overall optimisation problem is

$$\begin{aligned} a^* = \arg\min_a \varepsilon_{\text{data}}(a) + \lambda \varepsilon_{\text{smooth}}(a), \\ \text{s.t.} \quad i_{\text{un}}(\mathbf{u})/\|\mathbf{s}\| \leq a(\mathbf{u}) \leq 1, \end{aligned} \tag{42}$$

where λ is the regularisation weight. We compute the cost function gradient analytically, use sparse finite differences to compute the Hessian and solve the minimisation problem with bound constraints on the albedo using the trust region reflective algorithm. Since the data term is non-convex we require a good initialisation. This is provided by using a global convexity assumption to disambiguate the polarisation normals, as in [3], [4], and using this disambiguation to select from (39).

7 Specular Reflection and Polarisation

Many dielectric materials, including porcelain, skin, plastic and surfaces finished with gloss paint, exhibit “glossy” reflectance, i.e. in addition to subsurface diffuse reflectance, some light is reflected specularly through direct reflection at the air/surface interface. In order to allow surface height (Section 4) and albedo (Section 6) estimation to be applied to such objects, we propose some simple modifications to handle specular reflections. For lighting estimation on a glossy object, we simply apply the method in Section 5 only to diffuse-labelled pixels.

7.1 Additional Assumptions

We add the following assumptions to those listed in Section 4.1, but in so doing remove the need for Assumption 7:

11) Reflectance can be classified as diffuse dominant or specular dominant
12) Specular reflection follows the Blinn-Phong model [30] with known uniform parameters
13) Light source $\mathbf{s}$ is positioned in the same hemisphere as the viewer, i.e. $\mathbf{v} \cdot \mathbf{s} > 0$.

Assumption 11 is consistent with recent work [5], [7].

7.2 Specular Labelling

We label pixels as specular or diffuse dominant by thresholding a combination of three heuristics: 1. the degree of polarisation ($\rho > \sim 0.4$ implies specular reflection), 2. the specular coefficient estimated by the dichromatic reflectance model [31], 3. the rank order of the intensity (we consider only the top 10 percent brightest pixels). We divide the foreground mask into two sets of pixels. A pixel $\mathbf{u}$ belongs either to the set of diffuse pixels, $\mathcal{D}$, $|\mathcal{D}| = D$, or the set of specular pixels, $\mathcal{S}, |\mathcal{S}| = S$, with $\mathcal{F} = \mathcal{D} \cup \mathcal{S}, |\mathcal{F}| = D + S$. It follows from Assumptions 5 and 2 (i.e. a point source illuminating a smooth surface) that specular-labelled pixels will be sparse.

7.3 Specular Polarisation Model

For specular reflection, the degree of polarisation is again related to the zenith angle (Fig. 4a, blue curve) as follows:

$$\rho_s(\mathbf{u}) = \frac{2\sin(\theta(\mathbf{u}))^2 \cos(\theta(\mathbf{u}))\sqrt{\eta^2 - \sin(\theta(\mathbf{u}))^2}}{\eta^2 - \sin(\theta(\mathbf{u}))^2 - \eta^2 \sin(\theta(\mathbf{u}))^2 + 2\sin(\theta(\mathbf{u}))^4}. \tag{43}$$

This expression is problematic for two reasons: 1. it cannot be analytically inverted to solve for zenith angle, 2. there are two solutions. The first problem is overcome simply by using a lookup table and interpolation. The second problem is not an issue in practice. Specular reflections occur when the surface normal is approximately halfway between the viewer and light source directions. From Assumption 13, specular pixels will never have a zenith angle $> \sim 45°$. Hence, we can restrict (43) to this range and, therefore, a single solution. Based on this inversion of (43) we define the function $f_s(\rho_s(\mathbf{u}), \eta)$, similarly to (7), as

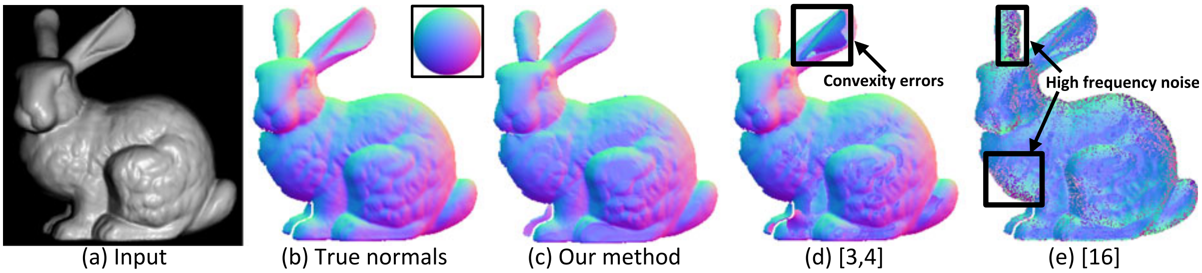


Fig. 6. Typical surface normal estimates (c-e) from noisy synthetic data (a). The inset sphere in (b) shows how surface orientation is visualised as a colour. Results obtained by [3], [4] in (d) and [15] in (e) for comparison.

$$f_s(\rho_s(\mathbf{u}), \eta) = \cos\theta(\mathbf{u}) = \frac{1}{\sqrt{p(\mathbf{u})^2 + q(\mathbf{u})^2 + 1}}. \tag{44}$$

In contrast to diffuse reflection, maximal polarisation for specular reflection occurs when the polariser's transmission axis is perpendicular to the plane of incidence/reflection. This means that the azimuth angle of the surface normal is perpendicular to the phase of the specular polarisation [32] leading to a $\frac{\pi}{2}$ shift

$$\mathbf{u} \in \mathcal{S} \Rightarrow \alpha(\mathbf{u}) = (\phi(\mathbf{u}) - \pi/2) \text{ or } (\phi(\mathbf{u}) + \pi/2). \tag{45}$$

Fig. 4b shows zenith angle estimates using the diffuse/specular model on $\mathcal{D}/\mathcal{S}$ respectively. In Fig. 4c we show the cosine of the estimated zenith angle, a visualisation corresponding to a Lambertian rendering with frontal lighting.

7.4 Specular Surface Gradient Constraints

In our earlier presentation of this work [16], we assumed that specular-labelled pixels simply had a surface normal equal to the halfway vector $\mathbf{h} = (\mathbf{s} + \mathbf{v})/\|\mathbf{s} + \mathbf{v}\|$. Here, we use an explicit specular reflectance model-the Blinn-Phong model. Although this is a non-physical model, it enables us to arrive at linear equations in the surface gradient. Accordingly, the unpolarised intensity for specular-labelled pixels is

$$\mathbf{u} \in \mathcal{S} \Rightarrow i_{\text{un}}(\mathbf{u}) = \mathbf{n}(\mathbf{u}) \cdot \mathbf{s} + k_s(\mathbf{n}(\mathbf{u}) \cdot \mathbf{h})^{\varsigma}, \tag{46}$$

where ς is the shininess, k_s the specular reflectivity and the halfway vector $\mathbf{h}$ is constant across the image. Since diffuse reflectance varies slowly with normal direction, we can use the approximation $\mathbf{n}(\mathbf{u}) \approx \mathbf{h}$ to compute and subtract the diffuse intensity from the unpolarised intensity of a specular pixel. Substituting this approximation into (46) and rewriting it in terms of the surface gradient we obtain

$$\frac{(i_{\text{un}}(\mathbf{u}) - \mathbf{h} \cdot \mathbf{s})^{\frac{1}{\varsigma}}}{k_s^{1/\varsigma}} = \frac{-p(\mathbf{u})h_x - q(\mathbf{u})h_y + h_z}{\sqrt{p(\mathbf{u})^2 + q(\mathbf{u})^2 + 1}}. \tag{47}$$

Expressing the polarisation and shading constraints for specular pixels as linear equations is very similar to the diffuse case. The phase angle provides exactly the same linear constraint as (15), though we must substitute in the $\frac{\pi}{2}$-shifted phase angles. To obtain the linear equation analogous to (16), we take a ratio between (47) and (44) yielding

$$\frac{(i_{\text{un}}(\mathbf{u}) - \mathbf{h} \cdot \mathbf{s})^{\frac{1}{\varsigma}}}{k_s^{1/\varsigma} f_s(\rho_s(\mathbf{u}), \eta)} = -p(\mathbf{u})h_x - q(\mathbf{u})h_y + h_z. \tag{48}$$

Hence, we obtain two linear equations per pixel that can be combined with the diffuse equations and solved in a single linear least squares system of the form in (29).

7.5 Diffuse Albedo Estimation in Specular Pixels

We treat diffuse albedo estimation in specular pixels as an inpainting problem. This entails making a stricter assumption about spatial smoothness than in diffuse regions where we only needed to assume that the albedo-free shading was smooth. Specifically, we use an isotropic total variation prior [33] on the estimated albedo

$$\varepsilon_{\text{TV}}(a) = \sum_{\mathbf{u} \in \mathcal{S} \cup \mathcal{D}_{\mathcal{S}}} \sqrt{[a(\mathbf{u}) - a(H(\mathbf{u}))]^2 + [a(\mathbf{u}) - a(V(\mathbf{u}))]^2}, \tag{49}$$

where $\mathcal{D}_{\mathcal{S}} \subset \mathcal{D}$ is the set of diffuse-labelled pixels that have a specular neighbour. $H(\mathbf{u})$ is the coordinate of the horizontal neighbour of pixel $\mathbf{u}$ and $V(\mathbf{u})$ is the coordinate of its vertical neighbour. Total variation minimisation has proven to be a highly effective generic prior for tasks such as denoising [33] and inpainting [34]. In our case, it amounts to encouraging the albedo to be piecewise smooth in specular regions where we cannot use the smoothness prior on the shading. We add this prior to the nonlinear albedo objective in (42). We initialise diffuse pixels as described in Section 6.2 and then initialise specular pixels with the albedo value of the diffuse pixel that is closest in terms of euclidean distance in the image plane.

8 Experimental Results

We now evaluate our illumination, albedo and surface height estimation methods on both synthetic and real data. We implement our methods in Matlab (full source code is available[5]) and run experiments on a MacBook Pro 2.7 GHz with 16 GB RAM. To construct and solve the linear system of equations required to estimate surface height takes around 1 second. The alternating optimisation to estimate illumination also takes around 1 second. Albedo estimation is the most computationally expensive part of our method, with the nonlinear optimisation taking around 20 seconds.

For synthetic data, we render images of the Stanford bunny with a physically-based reflectance model appropriate for smooth dielectrics (Fig. 6a). For diffuse reflectance we use the Wolff model [35]. For specular reflectance we use Fresnel-modulated perfect mirror reflection. We vary the light source direction $\mathbf{s} = [\sin(\alpha_l)\sin(\theta_l), \cos(\alpha_l)\sin$

5. https://github.com/waps101/depth-from-polarisation

TABLE 1
Quantitative Results on Synthetic Data ($\sigma = 0.5\%$)

	θ_l	Light (degrees)	Albedo	Method	Height (pixels)	Normal (degrees)
Uniform albedo	15°	0.62°	N/A	Ours^{gt}	10.9	8.50
				Ours^{est}	**10.8**	**8.49**
				[15]^{gt}	54.8	29.6
				[15]^{est}	48.8	26.8
				[3], [4]	44.4	9.16
	30°	1.03°	N/A	Ours^{gt}	9.80	6.86
				Ours^{est}	**9.66**	**6.81**
				[15]^{gt}	70.1	27.9
				[15]^{est}	62.0	25.1
				[3], [4]	56.3	13.3
	60°	8.14°	N/A	Ours^{gt}	9.66	**6.88**
				Ours^{est}	**8.66**	7.07
				[15]^{gt}	217	29.7
				[15]^{est}	213	28.5
				[3], [4]	205	20.3
Varying albedo	15°	N/A	0.075	Ours	**14.72**	**19.89**
				[3], [4]	69.3	24.7
	30°	N/A	0.11	Ours	**17.79**	**21.96**
				[3], [4]	176	35.6
	60°	N/A	0.17	Ours	**14.09**	**17.44**
				[3], [4]	240	38.4

$(\theta_l), \cos(\theta_l)]^T$ over $\theta_l \in \{15°, 30°, 60°\}$ and $\alpha_l \in \{0°, 90°, 180°, 270°\}$. We simulate the effect of polarisation according to (3), (6) and (43) with varying polariser angle, add Gaussian noise of standard deviation σ, saturate and quantise to 8 bits. Illumination is modelled as a dense aggregate of 1,000 point sources, distributed around $\mathbf{s}$, and we aggregate the polarisation fields over these sources. We estimate a polarisation image for each noise/illumination condition and use this as input.

In order to evaluate our method on real world images, we capture two datasets using a Canon EOS-1D X with an Edmund Optics glass linear polarising filter. The first dataset is captured in a dark room using a Lowel Prolight. We experiment with both known and unknown lighting. For known lighting, the approximate position of the light source is measured and to calibrate for unknown light source intensity and surface albedo, we use the method in Section 5.4 to compute the length of the light source vector, fixing its direction to the measured one. The second dataset is captured outdoors on a sunny day using natural illumination.

8.1 Illumination Estimation

Table 1 (uniform albedo) shows the quantitative accuracy of our light source estimate on synthetic data with $\sigma = 0.5\%$ (results with varying noise in supplementary material, available online). We report mean angular error as a function of θ_l, averaging over α_l and 100 repetitions. There is a small increase in error with the zenith angle of the light source.

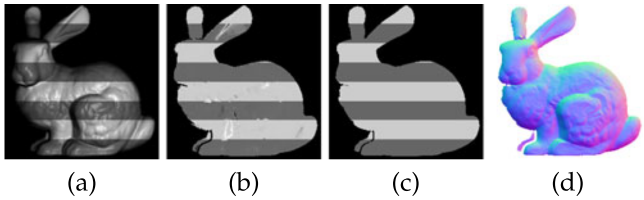


Fig. 7. From noisy synthetic data (a) we estimate a spatially varying albedo map (b). Ground truth is shown in (c). Surface normals (d) of height map estimated from (a) once estimated albedo has been divided out.

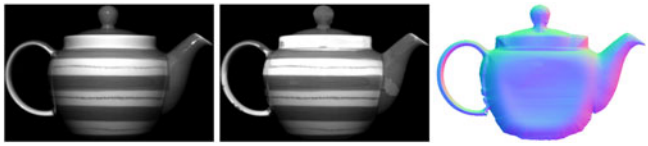


Fig. 8. Qualitative estimation results on a real teapot with varying albedo. Input (left), estimated albedo (middle), estimated surface normals (right).

8.2 Albedo Estimation

We generate synthetic data in the same way as for lighting estimation, however this time we use a simple stripe pattern as the diffuse albedo map. A sample result is in Fig. 7 where an image from the input sequence is shown in (a), our result in (b) and ground truth in (c). The result is largely devoid of shading and successfully inpaints the albedo in specular regions. Once the estimated albedo is divided out from the unpolarised intensity image, we are able to estimate a height map, the surface normals of which are shown in Fig. 7d. The edges in the albedo map cause no artefacts in the estimated surface. Table 1 (varying albedo part) shows quantitative results for albedo estimation, in terms of the Root-Mean-Square (RMS) error between estimated and ground truth albedo. We show two qualitative albedo estimation results for real images in Figs. 8 and 12. Again, the albedo maps appear largely invariant to shading and successfully inpaint texture in specular regions.

8.3 Surface Height Estimation

Finally, we evaluate surface height estimation using our method in Section 4. We compare to the only previous methods applicable to a single polarisation image: 1. boundary propagation [3], [4] and 2. Lambertian shading disambiguation [15]. The second method requires known light source direction and albedo and so for both this and for our

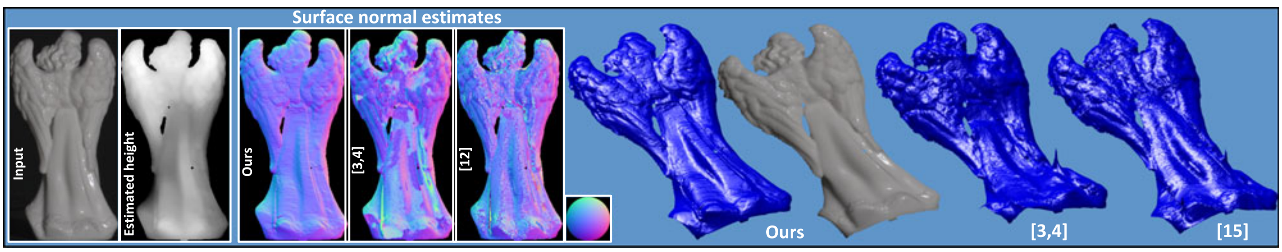


Fig. 9. Qualitative comparison against [3], [4] and [15] on real world data. Light source direction = $[2\ 0\ 7]$. For our method we show estimated surface height, normals, relit surface and texture mapped surface. For the comparison methods we show normals and relit surface.

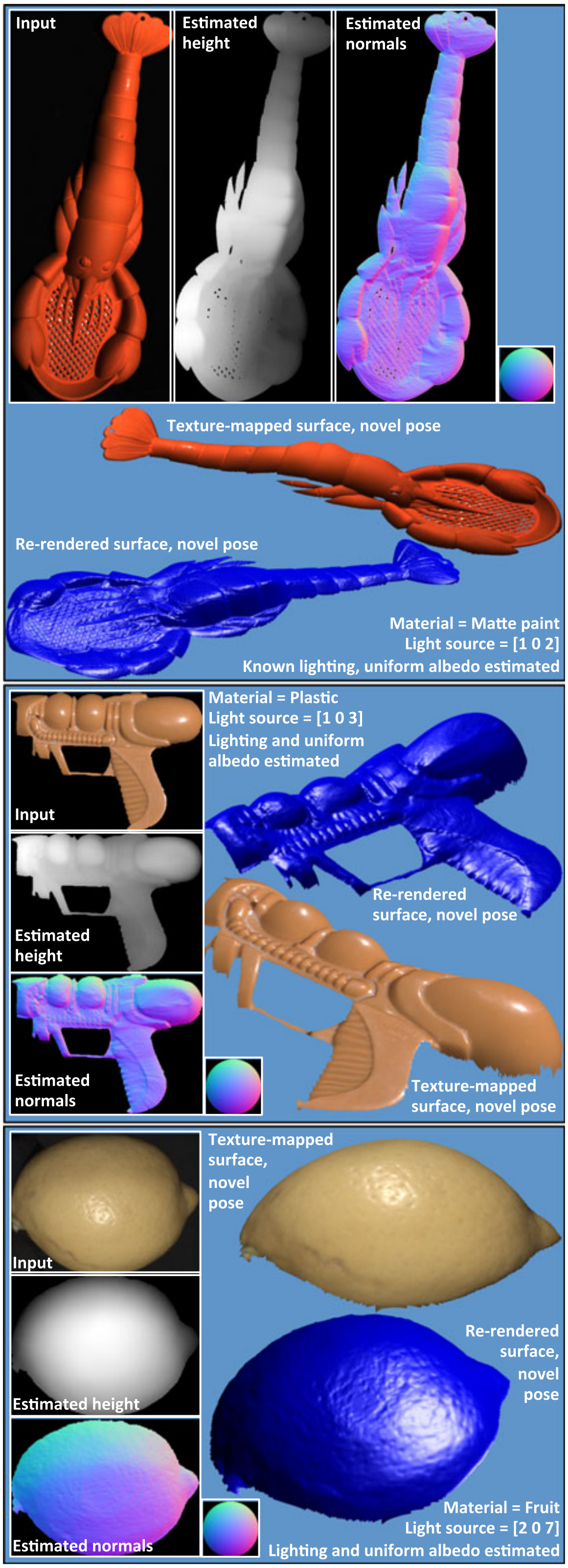


Fig. 10. Qualitative results indoors with point light source and uniform albedo.

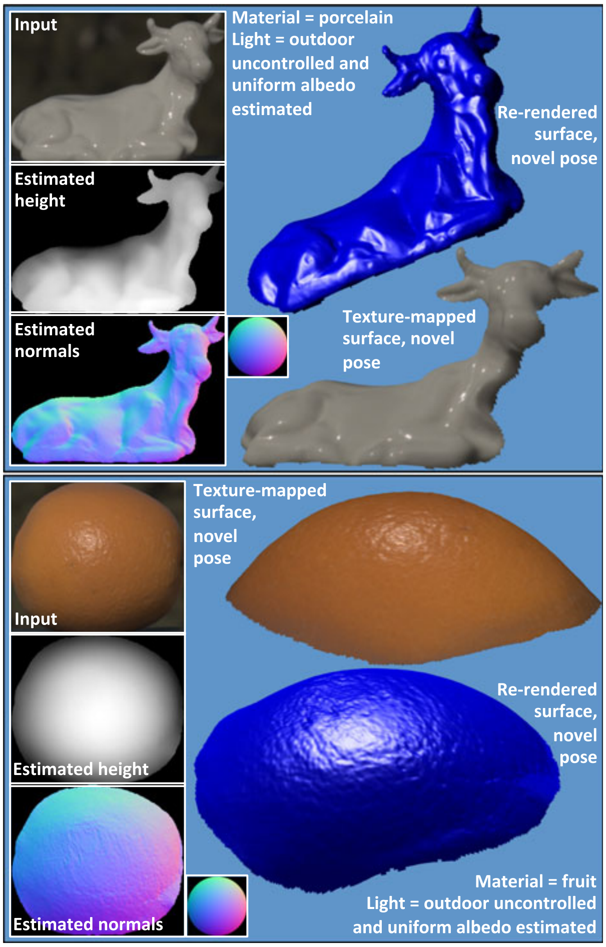


Fig. 11. Qualitative results outdoors on a sunny day and uniform albedo.

method, we provide results with ground truth lighting/albedo ("gt") and lighting/albedo estimated using the methods described in Section 5/Section 6 ("est"). For the comparison methods, we compute a height map using least squares surface integration, as in [36]. For our method, we compute surface normals using a bicubic fit to the estimated height map.

We show typical results in Figs. 6c, 6d, and 6e and quantitative results in Table 1 (RMS height error and mean angular surface normal error averaged over α_l and 100 repeats for each setting; best result for each setting emboldened). The boundary propagation methods [3], [4] assume convexity, meaning that internal concavities are incorrectly recovered. The Lambertian method [15] exhibits high frequency noise since solutions are purely local. Both methods also contain errors in specular regions and propagate errors from normal estimation into the integrated surface. Quantitatively, the result with estimated lighting is slightly better than with ground truth. We believe that this is because it enables the method to partially compensate for noise. Performance is worse in the presence of varying albedo. The flattening artefacts visible in Figs. 6c, 7d and 8 (right) is a limitation of SfP. For small zenith angles, polarisation provides only a weak cue and the smoothness prior dominates.

We show a qualitative comparison between our method and the two reference methods in Fig. 9 using known lighting. The comparison methods exhibit the same artefacts as on synthetic data. Some of the noise in the normals is removed by the smoothing effect of surface integration but concave/convex errors in [3], [4] grossly distort the overall shape, while the surface details of the wings are lost by [15]. In Figs. 10, 11 and 12 we show qualitative results of our method on a range of material types, under a variety of known or estimated illumination conditions (both indoor point source and outdoor uncontrolled) and with uniform

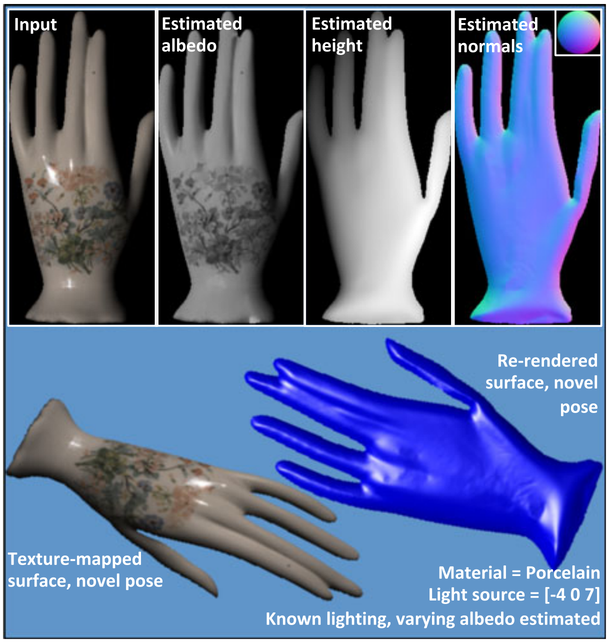


Fig. 12. Qualitative result for object with varying albedo.

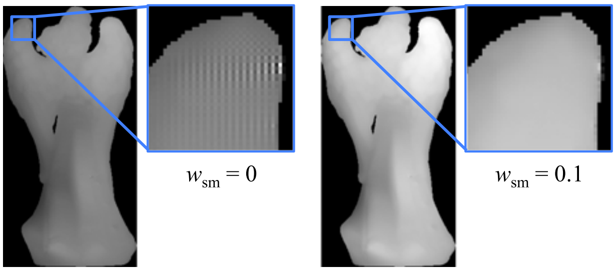


Fig. 13. Estimated depth maps without (left) and with (right) Laplacian smoothness prior for the object shown in Fig. 9. The inset zoomed region shows the "checkerboard" artefact that occurs with no smoothing.

or varying albedo. Note that our method is able to recover the fine surface detail of the skin of the lemon and orange under both point source and natural illumination. For the varying albedo example in Fig. 12, note that there are no texture transfer artefacts in the estimated shape (i.e. changes in albedo are not interpreted as changes in surface orientation).

To evaluate the influence of the priors described in Sections 4.7 and 4.8, we conducted an ablation study (see supplementary material, available online). On synthetic data, in the presence of noise, removing the smoothness prior typically increases surface normal error by around 20 percent. Removing the boundary prior increases the error by 5 percent and removing both priors increases the errors by 30 percent. See Figs. 13 and 14 for a qualitative visualisation of their influence. The smoothness prior helps reduce sensitivity to high frequency noise but also avoids a "checkerboard" effect resulting from central difference gradient approximations The convexity prior is helpful for data that are noisy close to the occluding boundary, for example when some background is included in the foreground mask. This is common with real data.

9 Conclusions

We have presented the first SfP technique in which polarisation constraints are expressed directly in terms of surface

(a) With convexity prior (b) Without convexity prior

Fig. 14. Influence of the convexity prior for the object at the bottom of Fig. 10.

height. Moreover, through careful construction of these equations, we ensure that they are linear and so height estimation is simply a linear least squares problem. The SfP cue is often described as being locally ambiguous. We have shown that, in fact, even with unknown lighting the diffuse unpolarised intensity image restricts the uncertainty to a global convex/concave ambiguity. Our method is practically useful, enabling monocular, passive surface height estimation even in outdoor lighting.

There are many ways that this work can be extended and improved. First, we would like to relax some of the assumptions. Rather than assuming that pixels are specular or diffuse dominant, we would like allow for mixtures of the two polarisation models. Instead of assuming Lambertian and Blinn-Phong reflectance models, an alternative would be to fit a BRDF model directly to the ambiguous polarisation normals, potentially allowing single shot BRDF and shape estimation. Second, linearising the objective functions by taking ratios means that we are solving a somewhat different optimisation problem to that addressed in previous literature. The linear solution could be used as an initialisation for a subsequent nonlinear optimisation over all unknowns of an objective function that can be directly related to a model of noise in the original data. Third, the minimal solution for light source estimation in Section 5.2 may lend itself to a robust light source estimation method, for example using RANSAC. This may improve robustness to outliers. Finally, we would like to explore combining our method with other cues. Since we directly compute height (or relative depth) it would be easy to combine the method with cues such as stereo or structure-from-motion that directly provide metric depth estimates.

Acknowledgments

This work was supported in part by EPSRC grant EP/N028481/1, ONR grants N000141512013 and N000141712687, GNCS-INdAM and the UC San Diego Center for Visual Computing. We thank Zak Murez and Dizhong Zhu for assistance with data collection.

References

[1] L. B. Wolff and T. E. Boult, "Constraining object features using a polarization reflectance model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 7, pp. 635–657, Jul. 1991.
[2] S. Rahmann and N. Canterakis, "Reconstruction of specular surfaces using polarization imaging," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2001, pp. I–I.
[3] G. A. Atkinson and E. R. Hancock, "Recovery of surface orientation from diffuse polarization," *IEEE Trans. Image Process.*, vol. 15, no. 6, pp. 1653–1664, Jun. 2006.
[4] D. Miyazaki, R. T. Tan, K. Hara, and K. Ikeuchi, "Polarization-based inverse rendering from a single view," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, 2003, pp. 982–987.

[5] A. Kadambi, V. Taamazyan, B. Shi, and R. Raskar, "Polarized 3D: High-quality depth sensing with polarization cues," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 3370–3378.

[6] P. N. Belhumeur, D. J. Kriegman, and A. Yuille, "The bas-relief ambiguity," *Int. J. Comput. Vis.*, vol. 35, no. 1, pp. 33–44, 1999.

[7] S. Tozza, R. Mecca, M. Duocastella, and A. Del Bue, "Direct differential photometric stereo shape recovery of diffuse and specular surfaces," *J. Math. Imag. Vis.*, vol. 56, no. 1, pp. 57–76, 2016.

[8] A. Ghosh, G. Fyffe, B. Tunwattanapong, J. Busch, X. Yu, and P. Debevec, "Multiview face capture using polarized spherical gradient illumination," *ACM Trans. Graph.*, vol. 30, no. 6, pp. 129:1–129:10, 2011.

[9] S. E. M. Herrera, A. Malti, O. Morel, and A. Bartoli, "Shape-from-polarization in laparoscopy," in *Proc. IEEE 10th Int. Symp. Biomed. Imag.*, Apr. 2013, pp. 1412–1415.

[10] V. V. Tuchin, L. Wang, and D. A. Zimnyakov, *Optical Polarization in Biomedical Applications*. Berlin, Germany: Springer, 2006.

[11] Y. Y. Schechner, "Self-calibrating imaging polarimetry," in *Proc. IEEE Int. Conf. Comput. Photography*, 2015, pp. 1–10.

[12] L. B. Wolff, "Polarization vision: A new sensory approach to image understanding," *Image Vis. Comput.*, vol. 15, no. 2, pp. 81–93, 1997.

[13] O. Morel, F. Meriaudeau, C. Stolz, and P. Gorria, "Polarization imaging applied to 3D reconstruction of specular metallic surfaces," in *Proc. EI*, 2005, pp. 178–186.

[14] C. P. Huynh, A. Robles-Kelly, and E. Hancock, "Shape and refractive index recovery from single-view polarisation images," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2010, pp. 1229–1236.

[15] A. H. Mahmoud, M. T. El-Melegy, and A. A. Farag, "Direct method for shape recovery from polarization and shading," in *Proc. 19th IEEE Int. Conf. Image Process.*, 2012, pp. 1769–1772.

[16] W. A. P. Smith, R. Ramamoorthi, and S. Tozza, "Linear depth estimation from an uncalibrated, monocular polarisation image," in *Proc. Eur. Conf. Comput. Vis.*, 2016, pp. 109–125.

[17] G. A. Atkinson and E. R. Hancock, "Surface reconstruction using polarization and photometric stereo," in *Proc. Int. Conf. Comput. Anal. Images Patterns*, 2007, pp. 466–473.

[18] T. T. Ngo, H. Nagahara, and R. Taniguchi, "Shape and light directions from shading and polarization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 2310–2318.

[19] G. A. Atkinson, "Polarisation photometric stereo," *Comput. Vis. Image Understanding*, vol. 60, pp. 158–167, 2017.

[20] G. A. Atkinson and E. R. Hancock, "Shape estimation using polarization and shading from two views," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 11, pp. 2001–2017, Nov. 2007.

[21] D. Miyazaki, M. Kagesawa, and K. Ikeuchi, "Transparent surface modeling from a pair of polarization images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 73–82, Jan. 2004.

[22] C. P. Huynh, A. Robles-Kelly, and E. R. Hancock, "Shape and refractive index from single-view spectro-polarimetric images," *Int. J. Comput. Vis.*, vol. 101, no. 1, pp. 64–94, 2013.

[23] O. Drbohlav and R. Šára, "Unambiguous determination of shape from photometric stereo with unknown light sources," in *Proc. 8th IEEE Int. Conf. Comput. Vis.*, 2001, pp. 581–586.

[24] D. Miyazaki, T. Shigetomi, M. Baba, R. Furukawa, S. Hiura, and N. Asada, "Polarization-based surface normal estimation of black specular objects from multiple viewpoints," in *Proc. 2nd Int. Conf. 3D Imag. Model. Process. Vis. Transmiss.*, 2012, pp. 104–111.

[25] D. Miyazaki, T. Shigetomi, M. Baba, R. Furukawa, S. Hiura, and N. Asada, "Surface normal estimation of black specular objects from multiview polarization images," *Opt. Eng.*, vol. 56, no. 4, pp. 041 303–1–041 303–17, 2016.

[26] A. Kadambi, V. Taamazyan, B. Shi, and R. Raskar, "Depth sensing using geometrically constrained polarization normals," *Int. J. Comput. Vis.*, vol. 125, pp. 34–51, 2017.

[27] Z. Cui, J. Gu, B. Shi, P. Tan, and J. Kautz, "Polarimetric multi-view stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 1558–1567.

[28] A. Carpinteri, *Structural Mechanics*. New York, NY, USA: Taylor and Francis, 1997, Art. no. 74.

[29] M. Grant, S. Boyd, and Y. Ye, "Disciplined convex programming," in *Global Optimization: From Theory to Implementation*. Berlin, Germany: Springer, 2006, pp. 155–210.

[30] J. F. Blinn, "Models of light reflection for computer synthesized pictures," *Comput. Graph.*, vol. 11, no. 2, pp. 192–198, 1977.

[31] S. A. Shafer, "Using color to separate reflection components," *Color Res. Appl.*, vol. 10, no. 4, pp. 210–218, 1985.

[32] A. Robles-Kelly and C. P. Huynh, *Imaging Spectroscopy for Scene Analysis*. Berlin, Germany: Springer, 2013, Art. no. 219.

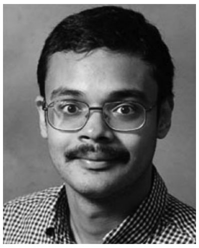
[33] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[34] P. Getreuer, "Total variation inpainting using split Bregman," *Image Process. On Line*, vol. 2, pp. 147–157, 2012.

[35] L. B. Wolff, "Diffuse-reflectance model for smooth dielectric surfaces," *J. Opt. Soc. Am. A*, vol. 11, no. 11, pp. 2956–2968, 1994.

[36] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi, "Efficiently combining positions and normals for precise 3D geometry," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 536–543, 2005.

William A. P. Smith (M'08) received the BSc degree in computer science, and the PhD degree in computer vision from the University of York, York, United Kingdom. He is currently a senior lecturer with the Department of Computer Science, University of York, York, United Kingdom. His research interests are in shape and appearance modelling, model-based supervision and physics-based vision. He has published more than 100 papers in international conferences and journals. He is a member of the IEEE.



Ravi Ramamoorthi received the BS degree in engineering and applied science the MS degrees in computer science and physics from the California Institute of Technology, in 1998, and the PhD degree in computer science from Stanford University Computer Graphics Laboratory, in 2002, upon which he joined the Columbia University Computer Science Department. He was on the UC Berkeley EECS faculty from 2009-2014. Since July 2014, he is a professor of Computer Science and Engineering, University of California, San Diego, where he holds the Ronald L. Graham chair of Computer Science. He is also the founding director of the UC San Diego Center for Visual Computing. His research interests cover many areas of computer vision and graphics, with more than 150 publications. His research has been recognized with a number of awards, including the 2007 ACM SIGGRAPH Significant New Researcher Award in computer graphics, and by the white house with a Presidential Early Career Award for Scientists and Engineers in 2008 for his work on physics-based computer vision. Most recently, he was named an IEEE and ACM fellow in 2017. He has advised more than 20 postdoctoral, PhD and MS students, many of whom have gone on to leading positions in industry and academia; and he has taught the first open online course in computer graphics on the edX platform in fall 2012, with more than 100,000 students enrolled in that and subsequent iterations. He was a finalist for the inaugural 2016 edX Prize for exceptional contributions in online teaching and learning, and again in 2017. He is a fellow of the IEEE.

Silvia Tozza received the BSc degree in mathematics from Università degli Studi di Salerno, Salerno, Italy, in 2008 the MSc degree in applied mathematics from Sapienza Università di Roma, Rome, Italy, in 2011, and the PhD degree in mathematics from Sapienza Università di Roma, in January 2015 with final judgment excellent. After one year postdoc atwith the Computer Science Department, University of York, York, United Kingdom, and then at the Department of Mathematics, Sapienza Università di Roma, currently, she is a postdoctoral research fellow of the Istituto Nazionale di Alta Matematica (INdAM), research unit at Department of Mathematics, Sapienza Università di Roma. Her research focuses mainly on analysis and approximation of nonlinear PDEs of Hamilton-Jacobi type, and applications in Image Processing, especially 3D shape recovery, segmentation and denoising. She is a member of the INdAM Research group GNCS.

▷ **For more information on this or any other computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.

Queries to the Author

Q1. There was a discrepancy in references in the PDF and the source file. We have followed the source file.
Q2. Please check whether the Ref. [13] are okay as set.