Linear Differential Constraints for Photo-polarimetric Height Estimation

Silvia Tozza
Sapienza - Università di Roma
tozza@mat.uniroma1.it

William A. P. Smith
University of York
william.smith@york.ac.uk

Dizhong Zhu
University of York
dz761@york.ac.uk

Ravi Ramamoorthi
UC San Diego
ravir@cs.ucsd.edu

Edwin R. Hancock
University of York
edwin.hancock@york.ac.uk

Abstract

In this paper we present a differential approach to photo-polarimetric shape estimation. We propose several alternative differential constraints based on polarisation and photometric shading information and show how to express them in a unified partial differential system. Our method uses the image ratios technique to combine shading and polarisation information in order to directly reconstruct surface height, without first computing surface normal vectors. Moreover, we are able to remove the non-linearities so that the problem reduces to solving a linear differential problem. We also introduce a new method for estimating a polarisation image from multichannel data and, finally, we show it is possible to estimate the illumination directions in a two source setup, extending the method into an uncalibrated scenario. From a numerical point of view, we use a least-squares formulation of the discrete version of the problem. To the best of our knowledge, this is the first work to consider a unified differential approach to solve photo-polarimetric shape estimation directly for height. Numerical results on synthetic and real-world data confirm the effectiveness of our proposed method.

1. Introduction

A recent trend in photometric [7, 13, 12, 26, 23, 20] and physics-based [24] shape recovery has been to develop methods that solve directly for surface height, rather than first estimating surface normals and then integrating them into a height map. Such methods are attractive since: 1. they only need solve for a single height value at each pixel (as opposed to the two components of surface orientation), 2. integrability is guaranteed, 3. errors do not accumulate through a two step pipeline of shape estimation and integration and 4. it enables combination with cues that provide depth information directly [10]. In both photometric stereo [13, 12, 23] and recently in shape-from-polarisation (SfP) [24], such a direct solution was made possible by deriving equations that are linear in the unknown surface gradient.

In this paper, we explore the combination of SfP constraints with photometric constraints (i.e. photo-polarimetric shape estimation) provided by one or two light sources. Photometric stereo with three or more light sources is a very well studied problem with robust solutions available under a range of different assumptions. Two source photometric stereo is still considered a difficult problem [21] even when the illumination is calibrated and albedo is known. We show that various formulations of one and two source photo-polarimetric stereo lead to the same general problem (in terms of surface height), that illumination can be estimated and that certain combinations of constraints lead to an albedo invariant formulation. Hence, with only modest additional data capture requirements (a polarisation image rather than an intensity image), we arrive at an approach for uncalibrated two source photometric stereo. We make the following novel contributions:

- We show how to estimate a polarisation image from multichannel data such as from colour images, multiple light source data or both (Sec. 2.2).
- We show how polarisation and photometric constraints (Sec. 3) can be expressed in a unified formulation (of which previous work [24] is a special case) and that various combinations of these constraints provide different practical advantages (Sec. 4).
- We show how to estimate the illumination directions in two source photo-polarimetric data leading to an uncalibrated solution (Sec. 6).

1.1. Related Work

The polarisation state of light reflected by a surface provides a cue to the material properties of the surface and, via a relationship with surface orientation, the shape. Polarisation has been used for a number of applications, in-

cluding early work on material segmentation [28] and diffuse/specular reflectance separation [18]. However, there has been a resurgent interest [24, 10, 25, 19] in using polarisation information for shape estimation.

Shape-from-polarisation The degree to which light is linearly polarised and the orientation associated with maximum reflection are related to the two degrees of freedom of surface orientation. In theory, this polarisation information alone restricts the surface normal at each pixel to two possible directions. Both Atkinson and Hancock [1] and Miyazaki et al. [15] solve the problem of disambiguating these polarisation normals via propagation from the boundary under an assumption of global convexity. Huynh et al. [8] also disambiguate polarisation normals with a global convexity assumption but estimate refractive index in addition. These works all used a diffuse polarisation model while Morel et al. [16] use a specular polarisation model for metals. Recently, Taamazyan et al. [25] introduced a mixed specular/diffuse polarisation model. All of these methods estimate surface normals that must be integrated into a height map. Moreover, since they rely entirely on the weak shape cue provided by polarisation and do not enforce integrability, the results are extremely sensitive to noise.

Photo-polarimetric methods There have been a number of attempts to combine photometric constraints with polarisation cues. Mahmoud et al. [11] used a shape-from-shading cue with assumptions of known light source direction, known albedo and Lambertian reflectance to disambiguate the polarisation normals. Atkinson and Hancock [3] used calibrated, three source Lambertian photometric stereo for disambiguation but avoiding an assumption of known albedo. Smith et al. [24] showed how to express polarisation and shading constraints directly in terms of surface height, leading to a robust and efficient linear least squares solution. They also show how to estimate the illumination, up to a binary ambiguity, making the method uncalibrated. However, they require known or uniform albedo. We explore variants of this method by introducing additional constraints that arise when a second light source is introduced, allowing us to relax the uniform albedo assumption. We also give an explanation for why the matrix they consider is full-rank except in a unique case. Recently, Ngo et al. [19] 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 lights. All of the above methods operate on single channel images and do not exploit the information available in colour images.

Polarisation with additional cues Rahmann and Canterakis [22] combined a specular polarisation model with stereo cues. Similarly, Atkinson and Hancock [2] used polarisation normals to segment an object into patches, simplifying stereo matching. Stereo polarisation cues have also been used for transparent surface modelling [14]. Huynh et al. [9] extended their earlier work to use multispectral measurements to estimate both shape and refractive index. Drbohlav and Sara [6] showed how the Bas-relief ambiguity [4] in uncalibrated photometric stereo could be resolved using polarisation. However, this approach requires a polarised light source. Recently, Kadambi et al. [10] proposed an interesting approach in which a single polarisation image is combined 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.

2. Representing Polarisation Information

We place a camera at the origin of a three-dimensional coordinate system ($Oxyz$) in such a way that Oxy coincides with the image plane and Oz with the optical axis. In Sec. 4 we propose a unified formulation for a variety of methods, all of which assume a) orthographic projection, b) known refractive index of the surface. Other assumptions will be given later on, depending on the specific problem at hand. We denote by $\mathbf{v}$ the viewer direction, by $\mathbf{s}$ a general light source direction with $\mathbf{v} \neq \mathbf{s}$. We only require the third components of these unit vectors to be greater than zero (*i.e.* all the vectors belong to the upper hemisphere). We will denote by $\mathbf{t}$ a second light source where required. We parametrise the unknown surface height by the function $z(\mathbf{x})$, where $\mathbf{x} = (x, y)$ is an image location, and the unit normal to the surface at the point $\mathbf{x}$ is given by:

$$\mathbf{n}(\mathbf{x}) = \frac{\hat{\mathbf{n}}(\mathbf{x})}{|\hat{\mathbf{n}}(\mathbf{x})|} = \frac{[-z_x, -z_y, 1]^T}{\sqrt{1 + |\nabla z(\mathbf{x})|^2}}, \tag{1}$$

where $\hat{\mathbf{n}}(\mathbf{x})$ is the outgoing normal vector and z_x, z_y denotes the partial derivative of $z(\mathbf{x})$ w.r.t. x and y, respectively, so that $\nabla z(\mathbf{x}) = (z_x, z_y)$. We now introduce relevant polarization theory, describing how we can estimate a polarisation image from multichannel data.

2.1. Polarisation image

When unpolarised light is reflected by a surface it becomes partially polarised [27]. A *polarisation image* can be estimated by capturing a sequence of images in which a linear polarising filter in front of the camera lens 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:

$$i_{\vartheta_j}(\mathbf{x}) = i_{\text{un}}(\mathbf{x}) \left(1 + \rho(\mathbf{x}) \cos(2\vartheta_j - 2\phi(\mathbf{x}))\right). \tag{2}$$

The polarisation image is thus obtained by decomposing the sinusoid at every pixel location into three quantities [27]: the *phase angle*, $\phi(\mathbf{x})$, the *degree of polarisation*, $\rho(\mathbf{x})$, and the *unpolarised intensity*, $i_{\text{un}}(\mathbf{x})$. The parameters of the sinusoid can be estimated from the captured image sequence

using non-linear least squares [1], linear methods [8] or via a closed form solution [27] for the specific case of $P = 3$, $\vartheta \in \{0°, 45°, 90°\}$.

2.2. Multichannel polarisation image estimation

A polarisation image is usually computed by fitting the sinusoid in (2) to observed data in a least squares sense. Hence, from $P \geq 3$ measurements we estimate i_{un}, ρ and ϕ. In practice, we may have access to multichannel measurements. For example, we may capture colour images (3 channels), polarisation images with two different light source directions (2 channels) or both (6 channels). Since ρ and ϕ depend only on surface geometry (assuming that, in the case of colour images, the refractive index does not vary with wavelength), then we expect these quantities to be constant over the channels. On the other hand, i_{un} will vary between channels either because of a shading change caused by the different lighting or because the albedo or light source intensity is different in the different colour channels. Hence, in a multichannel setting with C channels, we have $C + 2$ unknowns and CP observations. If we use information across all channels simultaneously, the system is more constrained and the solution will be more robust to noise. Moreover, we do not need to make an arbitrary choice about the channel from which we estimate the polarisation image. This idea shares something in common with that of Narasimhan et al. [17], though their material/shape separation was not in the context of polarisation.

Specifically, we can express the multichannel observations in channel c with polariser angle ϑ_j as

$$i^c_{\vartheta_j}(\mathbf{x}) = i^c_{\text{un}}(\mathbf{x})(1 + \rho(\mathbf{x})\cos(2\vartheta_j - 2\phi(\mathbf{x}))). \quad (3)$$

The system of equations is linear in the unpolarised intensities and, by a change of variables, can be made linear in ρ and ϕ [8]. Hence, we wish to solve a bilinear system and do so in a least squares sense using interleaved alternating minimisation. Specifically, we a) fix ρ and ϕ and then solve linearly for the unpolarised intensity in each channel and b) then fix the unpolarised intensities and solve linearly for ρ and ϕ using all channels simultaneously. Concretely, for a single pixel, we obtain the unpolarised intensities across channels by solving:

$$\min_{i^1_{\text{un}}(\mathbf{x}),\ldots,i^C_{\text{un}}(\mathbf{x})} \left\| \mathbf{C}_I \left[i^1_{\text{un}}(\mathbf{x}), \ldots, i^C_{\text{un}}(\mathbf{x}) \right]^T - \mathbf{d}_I \right\|^2, \quad (4)$$

where $\mathbf{C}_I \in \mathbb{R}^{CP \times C}$ is given by

$$\mathbf{C}_I = \begin{bmatrix} (1 + \rho(\mathbf{x})\cos(2\vartheta_1 - 2\phi(\mathbf{x})))\mathbf{I}_C \\ \vdots \\ (1 + \rho(\mathbf{x})\cos(2\vartheta_P - 2\phi(\mathbf{x})))\mathbf{I}_C \end{bmatrix}, \quad (5)$$

with $\mathbf{I}_C$ denoting the $C \times C$ identity matrix, and $\mathbf{d}_I \in \mathbb{R}^{CP}$ is given by

$$\mathbf{d}_I = \left[i^1_{\vartheta_1}(\mathbf{x}), \ldots, i^C_{\vartheta_1}(\mathbf{x}), i^1_{\vartheta_2}(\mathbf{x}), \ldots, i^C_{\vartheta_P}(\mathbf{x}) \right]^T.$$

Then, with the unpolarised intensities fixed, we solve for ρ and ϕ using the following linearisation:

$$\min_{a,b} \left\| \mathbf{C}_{\rho\phi} \begin{bmatrix} a \\ b \end{bmatrix} - \mathbf{d}_{\rho\phi} \right\|^2, \quad (6)$$

where $[a\ b]^T = [\rho(\mathbf{x})\cos(2\phi(\mathbf{x})), \rho(\mathbf{x})\sin(2\phi(\mathbf{x}))]^T$, and $\mathbf{C}_{\rho\phi} \in \mathbb{R}^{CP \times 2}$ is given by

$$\mathbf{C}_{\rho\phi} = \begin{bmatrix} i^1_{\text{un}}(\mathbf{x})\cos(2\vartheta_1) & i^1_{\text{un}}(\mathbf{x})\sin(2\vartheta_1) \\ \vdots & \vdots \\ i^1_{\text{un}}(\mathbf{x})\cos(2\vartheta_P) & i^1_{\text{un}}(\mathbf{x})\sin(2\vartheta_P) \\ i^2_{\text{un}}(\mathbf{x})\cos(2\vartheta_1) & i^2_{\text{un}}(\mathbf{x})\sin(2\vartheta_1) \\ \vdots & \vdots \\ i^C_{\text{un}}(\mathbf{x})\cos(2\vartheta_P) & i^C_{\text{un}}(\mathbf{x})\sin(2\vartheta_P) \end{bmatrix}, \quad (7)$$

and $\mathbf{d}_{\rho\phi} \in \mathbb{R}^{CP}$ is given by:

$$\mathbf{d}_{\rho\phi} = \begin{bmatrix} i^1_{\vartheta_1}(\mathbf{x}) - i^1_{\text{un}}(\mathbf{x}) \\ \vdots \\ i^1_{\vartheta_P}(\mathbf{x}) - i^1_{\text{un}}(\mathbf{x}) \\ i^2_{\vartheta_1}(\mathbf{x}) - i^2_{\text{un}}(\mathbf{x}) \\ \vdots \\ i^C_{\vartheta_P}(\mathbf{x}) - i^C_{\text{un}}(\mathbf{x}) \end{bmatrix}. \quad (8)$$

We estimate ρ and ϕ from the linear parameters using $\phi(\mathbf{x}) = \frac{1}{2}\text{atan2}(b, a)$ and $\rho(\mathbf{x}) = \sqrt{a^2 + b^2}$.

We initialise by computing a polarisation image from one channel using linear least squares, as in [8], and then use the estimated ρ and ϕ to begin alternating interleaved optimisation by solving for the unpolarised intensities across channels. We interleave and alternate the two steps until convergence. In practice, we find that this approach not only dramatically reduces noise in the polarisation images but also removes the ad hoc step of choosing an arbitrary channel to process. We show an example of the results obtained in Figure 1. The multichannel result is visibly less noisy than the single channel performance.

3. Photo-polarimetric height constraints

In this section we describe the different constraints provided by photo-polarimetric information and then show how to combine them to arrive at linear equations in the unknown surface height.

3.1. Degree of polarisation constraint

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. In this paper we will consider diffuse polarisation, due to subsurface scattering (see [1] for more details). The degree of diffuse polarisation $\rho_d(\mathbf{x})$ at each point $\mathbf{x}$ can be expressed in

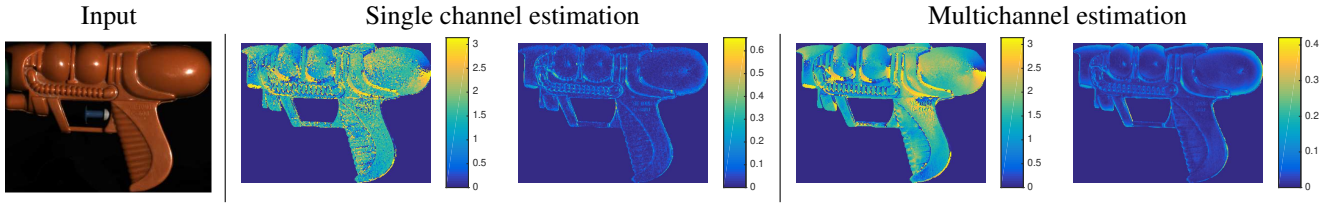

Figure 1. Multichannel polarisation image estimation. Left to right: an image from the input sequence; phase angle (ϕ) and degree of polarisation (ρ) estimated from a single channel; phase angle (ϕ) and degree of polarisation (ρ) estimated from three colour channels and two light source directions.

terms of the refractive index η and the surface zenith angle $\theta \in [0, \frac{\pi}{2}]$ as follows (Cf. [1]):

$$\rho_d(\mathbf{x}) = \frac{(\eta - 1/\eta)^2 \sin^2(\theta)}{2+2\eta^2-(\eta+1/\eta)^2 \sin^2(\theta)+4\cos(\theta)\sqrt{\eta^2-\sin^2(\theta)}}. \quad (9)$$

Recall that the zenith angle is the angle between the unit surface normal vector $\mathbf{n}(\mathbf{x})$ and the viewing direction $\mathbf{v}$. If we know the degree of polarisation $\rho_d(\mathbf{x})$ and the refractive index η (or have good estimates of them at hand), equation (9) can be rewritten with respect to the cosine of the zenith angle, and expressed in terms of the function, $f(\rho_d(\mathbf{x}), \eta)$, that depends on the measured degree of polarisation and the refractive index:

$$\cos(\theta) = \mathbf{n}(\mathbf{x}) \cdot \mathbf{v} = f(\rho_d(\mathbf{x}), \eta) = \sqrt{\frac{\eta^4(1-\rho_d^2)+2\eta^2(2\rho_d^2+\rho_d-1)+\rho_d^2+2\rho_d-4\eta^3\rho_d\sqrt{1-\rho_d^2}+1}{(\rho_d+1)^2(\eta^4+1)+2\eta^2(3\rho_d^2+2\rho_d-1)}} \quad (10)$$

where we drop the dependency of ρ_d on $(\mathbf{x})$ for brevity.

3.2. Shading constraint

The unpolarised intensity provides an additional constraint on the surface normal direction via an appropriate reflectance model. We assume that pixels have been labelled as diffuse or specular dominant and restrict consideration to diffuse shading. In practice, we deal with specular pixels in the same way as [24] and simply assume that they point in the direction of the halfway vector between $\mathbf{s}$ and $\mathbf{v}$. For the diffuse pixels, we therefore assume that light is reflected according to the Lambert's law. Hence, the unpolarised intensity is related to the surface normal by:

$$i_{\text{un}}(\mathbf{x}) = \gamma(\mathbf{x}) \cos(\theta_i) = \gamma(\mathbf{x}) \mathbf{n}(\mathbf{x}) \cdot \mathbf{s}, \quad (11)$$

where $\gamma(\mathbf{x})$ is the albedo. Writing $\mathbf{n}(\mathbf{x})$ in terms of the gradient of z as reported in (1), (11) can be rewritten as follows:

$$i_{\text{un}}(\mathbf{x}) = \gamma(\mathbf{x}) \frac{-\nabla z(\mathbf{x}) \cdot \tilde{\mathbf{s}} + s_3}{\sqrt{1 + |\nabla z(\mathbf{x})|^2}}, \quad (12)$$

with $\tilde{\mathbf{s}} = (s_1, s_2)$. This is a non-linear equation, but we will see in Sec. 3.4 and 3.5 how it is possible to remove the non-linearity by using the ratios technique.

3.3. Phase angle constraint

An additional constraint comes from the phase angle, which determines the azimuth angle of the surface normal $\alpha(\mathbf{x}) \in [0, 2\pi]$ up to a $180°$ ambiguity. This constraint can be rewritten as a collinearity condition [24], that is satisfied by either of the two possible azimuth angles implied by the phase angle measurement. Specifically, for diffuse pixels 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. This corresponds to requiring

$$\mathbf{n}(\mathbf{x}) \cdot [\cos(\phi(\mathbf{x})) \; -\sin(\phi(\mathbf{x})) \; 0]^T = 0. \quad (13)$$

In terms of the surface gradient, using (1), it is equivalent to

$$(-\cos\phi, \sin\phi) \cdot \nabla z = 0. \quad (14)$$

A similar expression can be obtained for specular pixels, substituting in the $\frac{\pi}{2}$-shifted phase angles. The advantage of doing this will become clear in Sec. 4.2.

3.4. Degree of polarisation ratio constraint

Combining the two constraints illustrated in Sec. 3.1 and 3.2, we can arrive at a linear equation, that we refer to as the DOP ratio constraint. Recall that $\cos(\theta) = \mathbf{n}(\mathbf{x}) \cdot \mathbf{v}$ and that we can express $\mathbf{n}$ in terms of the gradient of z by using (1), then isolating the non-linear term in (10) we obtain

$$\sqrt{1 + |\nabla z(\mathbf{x})|^2} = \frac{-\nabla z(\mathbf{x}) \cdot \tilde{\mathbf{v}} + v_3}{f(\rho_d(\mathbf{x}), \eta)}, \quad (15)$$

where $\tilde{\mathbf{v}} = (v_1, v_2)$. On the other hand, considering the shading information contained in (12), and again isolating the non-linearity we arrive at the following

$$\sqrt{1 + |\nabla z(\mathbf{x})|^2} = \gamma(\mathbf{x}) \frac{-\nabla z(\mathbf{x}) \cdot \tilde{\mathbf{s}} + s_3}{i_{\text{un}}(\mathbf{x})}. \quad (16)$$

Note that we are supposing $\mathbf{s} \neq \mathbf{v}$, and $i_{\text{un}}(\mathbf{x}) \neq 0$, $f(\rho_d(\mathbf{x}), \eta) \neq 0$. Inspecting Eqs. (15) and (16) we obtain

$$\frac{-\nabla z(\mathbf{x}) \cdot \tilde{\mathbf{v}} + v_3}{f(\rho_d(\mathbf{x}), \eta)} = \gamma(\mathbf{x}) \frac{-\nabla z(\mathbf{x}) \cdot \tilde{\mathbf{s}} + s_3}{i_{\text{un}}(\mathbf{x})}. \quad (17)$$

We thus arrive at the following partial differential equation (PDE):

$$\mathbf{b}(\mathbf{x}) \cdot \nabla z(\mathbf{x}) = h(\mathbf{x}), \tag{18}$$

where

$$\mathbf{b}(\mathbf{x}) := \mathbf{b}^{(f,i_{\text{un}})} = i_{\text{un}}(\mathbf{x})\tilde{\mathbf{v}} - \gamma(\mathbf{x}) f(\rho_d(\mathbf{x}), \eta)\, \tilde{\mathbf{s}}, \tag{19}$$

and

$$h(\mathbf{x}) := h^{(f,i_{\text{un}})} = i_{\text{un}}(\mathbf{x}) v_3 - \gamma(\mathbf{x}) f(\rho_d(\mathbf{x}), \eta)\, s_3. \tag{20}$$

3.5. Intensity ratio constraint

Finally, we construct an intensity ratio constraint by considering two unpolarised images, $i_{\text{un},1}, i_{\text{un},2}$, taken from two different light source directions, $\mathbf{s}, \mathbf{t}$. We construct our constraint equation by applying (11) twice, once for each light source. We can remove the non-linearity as before and take a ratio, arriving at the following equation:

$$i_{\text{un},2}(-\nabla z(\mathbf{x}) \cdot \tilde{\mathbf{s}} + s_3) = i_{\text{un},1}(-\nabla z(\mathbf{x}) \cdot \tilde{\mathbf{t}} + t_3). \tag{21}$$

The above equation is independent of albedo, light source intensity and non-linear normalisation term. Again as before, we can rewrite (21) as a PDE in the form of (18) with

$$\mathbf{b}(\mathbf{x}) := \mathbf{b}^{(i_{\text{un},1}, i_{\text{un},2})} = i_{\text{un},2}(\mathbf{x})\tilde{\mathbf{s}} - i_{\text{un},1}(\mathbf{x})\, \tilde{\mathbf{t}}, \tag{22}$$

where $\tilde{\mathbf{t}} = (t_1, t_2)$, and

$$h(\mathbf{x}) := h^{(i_{\text{un},1}, i_{\text{un},2})} = i_{\text{un},2}(\mathbf{x}) s_3 - i_{\text{un},1}(\mathbf{x})\, t_3. \tag{23}$$

4. A unified PDE formulation

Commencing from the constraints introduced in Sec. 3, in this section we show how to solve several different problems in photo-polarimetric shape estimation. The common feature is that these are all linear in the unknown height, and are expressed in a unified formulation in terms of a system of PDEs in the same general form:

$$\mathbf{B}(\mathbf{x}) \nabla z(\mathbf{x}) = \mathbf{h}(\mathbf{x}), \tag{24}$$

where $\mathbf{B} : \bar{\Omega} \to \mathbb{R}^{J \times 2}$, $\mathbf{h} : \bar{\Omega} \to \mathbb{R}^{J \times 1}$, denoting by Ω the reconstruction domain and being $J = 2, 3$ or 4 depending on the cases. (24) is a compact and general equation, suitable for describing several cases in a unified differential formulation that solves directly for surface height.

Different combinations of the three constraints described in Sec. 3 that are linear in the surface gradient can be combined in the formulation of (24). Each corresponds to different assumptions and have different pros and cons. We explore three variants and show that [24] is a special case of our formulation. We summarise the alternative formulations in Tab. 1.

Method	Phase angle	DOP ratio	Intensity ratio
[24]	✓	✓	
Proposed 1	✓		✓
Proposed 2		✓	✓
Proposed 3	✓	✓	✓

Table 1. Summary of the different formulations

4.1. Single light and polarisation formulation

This case has been studied in [24]. It uses a single polarisation image, requires known illumination (though [24] show how this can be estimated if unknown) and assumes that albedo is known or uniform. This last assumption is quite restrictive, since it can only be applied to objects with homogeneous surfaces. With just a single illumination condition, only the phase angle and DOP ratio constraints are available. This thus becomes a special case of our general unified formulation (24), where $\mathbf{B}$ and $\mathbf{h}$ are defined as

$$\mathbf{B} = \begin{bmatrix} b_1^{(f,i_{\text{un}})} & b_2^{(f,i_{\text{un}})} \\ -\cos\phi & \sin\phi \end{bmatrix}, \quad \mathbf{h} = [h^{(f,i_{\text{un}})}, 0]^T, \tag{25}$$

with $\mathbf{b}^{(f,i_{\text{un}})}$ and $h^{(f,i_{\text{un}})}$ defined by (19) and (20), with uniform $\gamma(\mathbf{x})$ and $\mathbf{v} = [0, 0, 1]^T$.

4.2. Proposed 1: Albedo invariant formulation

Our first proposed method uses the phase angle constraint (14) and two unpolarised images, taken from two different light source directions, obtained through (12) and combined as in (21). In this case the problem studied is described by the system of PDEs (24) with

$$\mathbf{B}(\mathbf{x}) = \begin{bmatrix} b_1^{(i_{\text{un},1}, i_{\text{un},2})} & b_2^{(i_{\text{un},1}, i_{\text{un},2})} \\ -\cos\phi & \sin\phi \end{bmatrix}, \mathbf{h}(\mathbf{x}) = \begin{bmatrix} h^{(i_{\text{un},1}, i_{\text{un},2})} \\ 0 \end{bmatrix}, \tag{26}$$

where $\mathbf{b}^{(i_{\text{un},1}, i_{\text{un},2})}$ and $h^{(i_{\text{un},1}, i_{\text{un},2})}$ defined as in (22) and (23). The phase angle does not depend on albedo and the intensity ratio constraint is invariant to albedo. As a result, this formulation is particularly powerful because it allows albedo invariant height estimation. Moreover, the light source directions in the two images can be estimated (again, in an albedo invariant manner) using the method in Sec. 6.

Once surface height has been estimated, we can compute the surface normal at each pixel and it is then straightforward to estimate an albedo map using (11). Where we have two diffuse observations, we can compute albedo from two equations of the form of (11) in a least squares sense. In real data, where we have specular pixel labels, we use only the diffuse observations at each pixel. To avoid artifacts at the boundary of specular regions, we introduce a gradient consistency term into the albedo estimation. We encourage the gradient of the albedo map to match the gradients of the intensity image for diffuse pixels.

4.3. Proposed 2: Phase invariant formulation

Our second proposed method uses only the DOP ratio and the intensity ratio constraints. This means that phase angle estimates are not used. The advantage of this is that phase angles are subject to a shift of $\frac{\pi}{2}$ at specular reflections when compared to diffuse reflections. So, the phase angle constraint relies upon having accurate per-pixel specularity labels, which classify reflections as either dominantly specular or diffuse (or alternatively use a mixed polarisation model [25] with a four way ambiguity). In this case we need a) two unpolarised intensity images, taken with two different light source directions, $\mathbf{s}$ and $\mathbf{t}$, obtained through (12), b) polarisation information from the function $f(\rho, \eta)$ and c) knowledge of the albedo map. We need $\mathbf{s}, \mathbf{t}, \mathbf{v}$ non-coplanar in order to have the matrix field $\mathbf{B}$ not singular. Note that the function f, obtained from polarization information (as in (10)), is the same for the two required images. The reason for this is that it does not depend on the light source directions but only on the viewer direction $\mathbf{v}$ which does not change. This formulation can be deduced starting from (21) and (17), arriving at a PDE system as in (24) with

$$\mathbf{B} = [\mathbf{b}^{(f,i_{\text{un},1})}, \mathbf{b}^{(f,i_{\text{un},2})}, \mathbf{b}^{(i_{\text{un},1},i_{\text{un},2})}]^T, \tag{27}$$

and $\mathbf{h} = [h^{(f,i_{\text{un},1})}, h^{(f,i_{\text{un},2})}, h^{(i_{\text{un},1},i_{\text{un},2})}]^T$, using (19), (20), (22), (23) to define the vector fields $\mathbf{b}$ and the scalar fields h that appear in $\mathbf{B}$ and $\mathbf{h}$.

4.4. Proposed 3: Most constrained formulation

Our final proposed method combines all of the previous constraints, leading to a problem of the form (24) with

$$\mathbf{B} = \begin{bmatrix} b_1^{(f,i_{\text{un},1})} & b_2^{(f,i_{\text{un},1})} \\ b_1^{(f,i_{\text{un},2})} & b_2^{(f,i_{\text{un},2})} \\ b_1^{(i_{\text{un},1},i_{\text{un},2})} & b_2^{(i_{\text{un},1},i_{\text{un},2})} \\ -\cos\phi & \sin\phi \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} h^{(f,i_{\text{un},1})} \\ h^{(f,i_{\text{un},2})} \\ h^{(i_{\text{un},1},i_{\text{un},2})} \\ 0 \end{bmatrix}. \tag{28}$$

This formulation uses the most information and so is potentially the most robust method. However, it requires known albedo in order to use the DOP ratio constraint. Nevertheless, it is possible to first apply proposed method 1, estimate the albedo and then re-estimate surface height using the maximally constrained formulation and the estimated albedo map. In fact, the best performance is obtained by iterating these two steps, alternately using the surface height estimate to compute albedo and then using the updated albedo to re-compute surface height.

4.5. Extension to colour images

We now consider how to extend the above systems of equations when colour information is available. If a surface is lit by a coloured point source, then each pixel provides three equations of the form in (11). In principle, this provides no more information than a grayscale observation since the surface normal and light source direction are fixed across colour channels. However, in the presence of noise using all three observations improves robustness. In particular, if the albedo value at a pixel is lower in one colour channel, the signal to noise ratio will be worse in that channel than the others. For a multicoloured object, it is impossible to choose a single colour channel that provides the best signal to noise ratio across the whole object. For this reason, we propose to use information from all colour channels where available.

We already exploit colour information in the estimation of the polarisation image in Sec. 2.2. Hence, the phase angle estimates have already benefited from the improved robustness. Both the DOP ratio and intensity ratio constraints can also exploit colour information by repeating each constraint three times, once for each colour channel. In the case of the intensity ratio, the colour albedo once again cancels if ratios are taken between the same colour channels under different light source directions.

5. Height estimation via linear least squares

We have seen that each of the variants illustrated in the previous section, each with different advantages, can be written as a PDE system (24). Denoting by M the number of pixels, we discretise the gradient in (24) via finite differences, arriving at the following linear system in $\mathbf{z}$

$$\mathbf{A}\mathbf{z} = \bar{\mathbf{h}}, \tag{29}$$

where $\mathbf{A} = \bar{\mathbf{B}}\mathbf{G}$, with $\mathbf{G} \in \mathbb{R}^{2M \times M}$ the matrix of finite difference gradients. $\bar{\mathbf{B}} \in \mathbb{R}^{JM \times 2M}$ is the discrete per-pixel version of the matrix $\mathbf{B}(\mathbf{x})$, hence $\mathbf{A} \in \mathbb{R}^{JM \times M}$, where J depends on the various proposed cases reported in Sec. 4 ($J = 2$ for (25) and (26), $J = 3$ for (27) and $J = 4$ for (28)). $\bar{\mathbf{h}}$ is the discrete per-pixel version of the function $\mathbf{h}(\mathbf{x})$, $\bar{\mathbf{h}} \in \mathbb{R}^{JM \times 1}$, and $\mathbf{z} \in \mathbb{R}^{M \times 1}$ the vector of the unknown height values. The resulting discrete system is large, since we have JM linear equations in M unknowns, but sparse, since $\mathbf{A}$ has few non-zero values for each row, and has as unknowns the height values. The per-pixel matrix $\mathbf{A}$ is a full-rank matrix, for each choice of $\bar{\mathbf{B}}$ that comes from the proposed formulations in Sec. 4, under the different assumptions specified for each case. The per-pixel matrix $\mathbf{A}$ related to [24] is full-rank except in one case: when the first two components of the light vector $\mathbf{s}$ are non-zero and $s_1 = -s_2$ and it happens that the phase angle is $\phi = \pi/4$ at least in one pixel. In that case, the matrix has a rank-deficiency (though in practice ϕ assuming a value of exactly $\pi/4$, up to numerical tolerance, is unlikely).

We want to find a solution of (29) in the least-squares sense, *i.e.* find a vector $\mathbf{z} \in \mathbb{R}^M$ such that

$$||\mathbf{A}\mathbf{z} - \bar{\mathbf{h}}||_2^2 \leq ||\mathbf{A}\mathbf{y} - \bar{\mathbf{h}}||_2^2, \quad \forall \mathbf{y} \in \mathbb{R}^M. \tag{30}$$

Considering the associate system of normal equations

$$\mathbf{A}^T(\mathbf{A}\mathbf{z} - \bar{\mathbf{h}}) = 0, \tag{31}$$

it is well-known that if there exists $\mathbf{z} \in \mathbb{R}^M$ that satisfies (31), then $\mathbf{z}$ is also solution of the least-squares problem, *i.e.* $\mathbf{z}$ satisfies (30). Since $\mathbf{A}$ is a full-rank matrix, then the matrix $\mathbf{A}^T\mathbf{A}$ is not singular, hence there exists a unique solution $\mathbf{z}$ of (31) for each data term $\bar{\mathbf{h}}$. Since neither $\mathbf{B}$ nor $\mathbf{h}$ depend on z in (24), the solution can be computed only up to an additive constant (which is consistent with the orthographic projection assumption). To resolve the unknown constant, knowledge of z at just one pixel is sufficient. In our implementation, we remove the height of one pixel from the variables and substitute its zero value elsewhere.

6. Two source lighting estimation

Our three proposed shape estimation methods require knowledge of the two light source directions. Previously, Smith et al. [24] showed that a single polarisation image can be used to estimate illumination conditions up to a binary ambiguity. However, to do so, they assumed that the albedo was known or uniform, and they also worked only with a single colour channel. In a two source setting, we show that it is possible to estimate both light source directions simultaneously, and do so in an albedo invariant manner. Moreover, we can exploit information across different colour channels to improve robustness to noise. Hence, our three methods can be used in an uncalibrated setting.

The intensity ratio (21) provides one equation per pixel relating unpolarised intensities, surface gradient and light source directions. Given two polarisation images with different light directions, we have one such equation per pixel and six unknowns in total. We assume that ambiguous surface gradient estimates are known from ρ and ϕ, and then use (21) to estimate the light source directions.

The intensity ratio (21) is homogeneous in $\mathbf{s}$ and $\mathbf{t}$ and so has a trivial solution $\mathbf{s} = \mathbf{t} = [0\ 0\ 0]^T$. If we assume that the intensity of the light source remains constant in each colour channel across the two images, then this intensity divides out when taking an intensity ratio and so the length of the light source vectors is arbitrary. We therefore constrain them to unit length (avoiding the trivial solution), and represent them by spherical coordinates (θ_s, α_s) and (θ_t, α_t), such that $[s_1, s_2, s_3] = [\cos\alpha_s \sin\theta_s, \sin\alpha_s \sin\theta_s, \cos\theta_s]$ and $[t_1, t_2, t_3] = [\cos\alpha_t \sin\theta_t, \sin\alpha_t \sin\theta_t, \cos\theta_t]$. This reduces the number of unknowns to four. We can now write the residual at each pixel given an estimate of the light source directions. There are two possible residuals, depending on which of the two ambiguous polarisation normals we use. From the phase angle and the zenith angle estimated from the degree of polarisation using (10), we have two possible surface normal directions at each pixel and therefore two possible gradients: $z_x(\mathbf{x}) \approx \pm\cos\phi(\mathbf{x})\tan\theta(\mathbf{x})$, $z_y(\mathbf{x}) \approx \pm\sin\phi(\mathbf{x})\tan\theta(\mathbf{x})$. Hence, the residuals at pixel $\mathbf{x}_j$ in channel c are given by either:

$$\begin{aligned} r_{j,c}(\theta_s, \alpha_s, \theta_t, \alpha_t) =& i^c_{\text{un},1}(\mathbf{x}_j)(-z_x(\mathbf{x}_j)t_1 - z_y(\mathbf{x}_j)t_2 + t_3) - \\ & i^c_{\text{un},2}(\mathbf{x}_j)(-z_x(\mathbf{x}_j)s_1 - z_y(\mathbf{x}_j)s_2 + s_3), \\ & \text{or} \\ q_{j,c}(\theta_s, \alpha_s, \theta_t, \alpha_t) =& i^c_{\text{un},1}(\mathbf{x}_j)(z_x(\mathbf{x}_j)t_1 + z_y(\mathbf{x}_j)t_2 + t_3) - \\ & i^c_{\text{un},2}(\mathbf{x}_j)(z_x(\mathbf{x}_j)s_1 + z_y(\mathbf{x}_j)s_2 + s_3). \end{aligned}$$

We can now write a minimisation problem for light source direction estimation by summing the minimum of the two residuals over all pixels and colour channels:

$$\min_{\theta_s,\alpha_s,\theta_t,\alpha_t} \sum_{j,c} \min[r^2_{j,c}(\theta_s, \alpha_s, \theta_t, \alpha_t), q^2_{j,c}(\theta_s, \alpha_s, \theta_t, \alpha_t)].$$

The minimum of two convex functions is not itself convex and so this optimisation is non-convex. However, we find that, even with a random initialisation, it almost always converges to the global minimum. As in [24], the solution is still subject to a binary ambiguity, in that if $(\mathbf{s}, \mathbf{t})$ is a solution then $(\mathbf{T}\mathbf{s}, \mathbf{T}\mathbf{t})$ is also a solution (with $\mathbf{T} = \text{diag}([-1, -1, 1])$), corresponding to the convex/concave ambiguity. We resolve this simply by choosing the maximal solution when surface height is later recovered.

7. Experiments

We begin by using synthetic data generated from the Mozart height map (Fig. 3). We differentiate to obtain surface normals and compute unpolarised intensities by rendering the surface using light sources $\mathbf{s} = [1, 0, 5]^T$ and $\mathbf{t} = [-1, -2, 7]^T$ according to (11). We experiment with both uniform albedo and varying albedo for which we use a checkerboard pattern. We simulate the effect of polarisation according to (2), varying the polariser angle between $0°$ and $180°$ in $10°$ increments. Next, we corrupt this data by adding Gaussian noise with zero mean and standard deviation σ, saturate and quantise to 8 bits. This noisy data provides the input to our reconstruction. First, we estimate a polarisation image using the method in Sec. 2.2, then apply each of the proposed methods or the state-of-the-art comparison method [24] to recover the height map.

In Tab. 2 we report Root-Mean-Square (RMS) error in the surface height (in pixels) and mean angular error (in degrees) in the surface normals obtained by differentiating the estimated surface height. In Fig. 3 we show a sample of qualitative results from this experiment. In all cases, more than one of our proposed methods outperform [24]. When albedo is uniform, our phase invariant (Prop. 2) or maximally constrained solution (Prop. 3) provides the best results. When albedo is non-uniform, the albedo invariant method (Prop. 1) provides much better performance. Although the combination of the albedo invariant method followed by the maximally constrained method (Prop. 1+3)

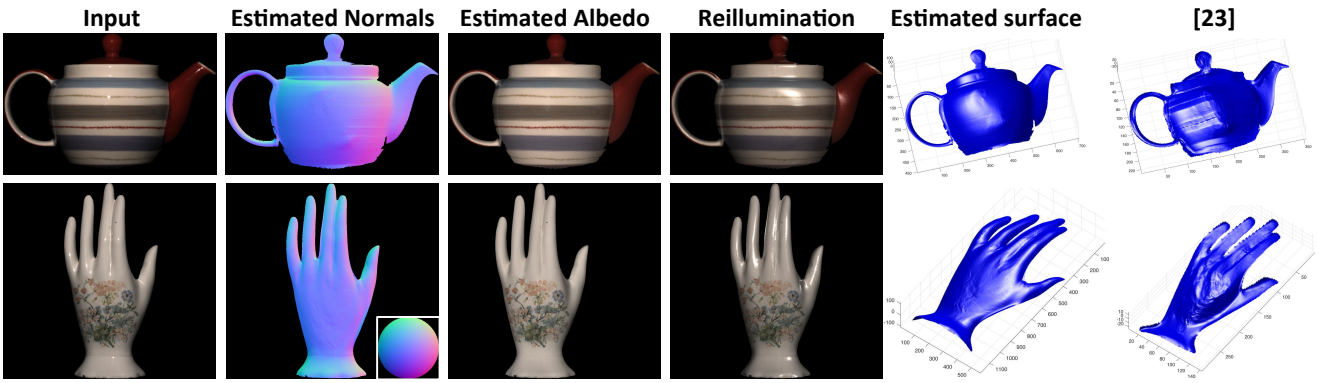

Figure 2. Qualitative results on real objects with varying albedo obtained by using Prop. 1+3 and comparison to [24] (zoom for detail).

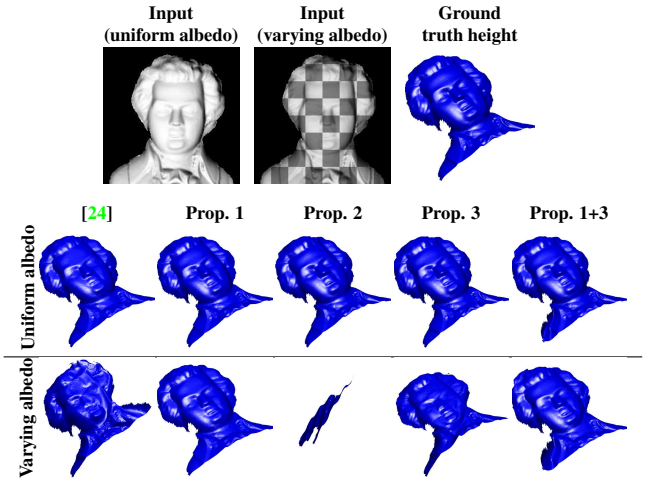

Figure 3. Qualitative results on synthetic data.

Setting	Method	$\sigma = 0\%$		$\sigma = 0.5\%$		$\sigma = 2\%$	
		Height (pix)	Normal (deg)	Height (pix)	Normal (deg)	Height (pix)	Normal (deg)
Uniform albedo, known lighting	[24]	1.12	2.85	1.68	4.48	5.06	11.28
	Prop. 1	1.78	2.52	1.94	3.30	3.49	7.22
	Prop. 2	**0.23**	1.45	0.70	**1.70**	6.50	5.33
	Prop. 3	0.42	**1.03**	**0.52**	1.74	**1.53**	**4.73**
	Prop. 1+3	3.37	3.22	3.62	4.03	5.82	9.15
Uniform albedo, estimated lighting	[24]	1.10	2.84	1.55	4.36	4.94	11.16
	Prop. 1	1.77	2.51	1.88	3.23	3.04	6.86
	Prop. 2	**0.23**	1.45	0.71	**1.71**	5.87	5.68
	Prop. 3	0.41	**1.02**	**0.49**	1.74	**1.47**	**4.88**
	Prop. 1+3	3.36	3.21	3.57	3.97	5.73	8.93
Unknown albedo, known lighting	[24]	22.50	28.03	21.63	27.76	20.76	26.74
	Prop. 1	**2.74**	**4.18**	**3.28**	**5.76**	**6.65**	**13.11**
	Prop. 2	141.19	59.69	140.04	59.49	131.16	57.69
	Prop. 3	18.62	16.76	18.58	16.85	17.33	16.82
	Prop. 1+3	5.22	9.59	5.80	11.26	7.56	16.50
Unknown albedo, estimated lighting	[24]	7.78	18.10	8.20	18.82	9.93	22.68
	Prop. 1	**2.73**	**4.17**	**3.19**	**5.62**	**6.53**	**12.98**
	Prop. 2	140.56	59.58	133.76	58.31	91.24	47.88
	Prop. 3	18.66	16.79	19.02	17.15	20.34	18.76
	Prop. 1+3	5.21	9.57	5.75	11.09	8.84	19.56

Table 2. Height and surface normal errors on synthetic data.

does not give quantitatively the best performance, we find that on real world data containing more complex noise and specular reflections, this approach is most robust.

In Fig. 2 we show qualitative results on two real objects with spatially varying albedo. From left to right we show: an image from the input sequence; the surface normals of the estimated height map (inset sphere shows how orientation is visualised as colour); the estimated albedo map; a re-rendering of the estimated surface and albedo map under novel lighting with Blinn-Phong reflectance [5]; a rotated view of the estimated surface; and, for comparison, reconstructions of the same surfaces using [24]. The results of [24] are highly distorted in the presence of varying albedo. Our approach avoids transfer of albedo details into the recovered shape, leading to convincing relighting results.

8. Conclusions

In this paper we have introduced a unifying formulation for recovering height from photo-polarimetric data and proposed a variety of methods that use different combinations of linear constraints. We proposed a more robust way to estimate a polarisation image from multichannel data and showed how to estimate lighting from two source photo-polarimetric images. Together, our methods provide uncalibrated, albedo invariant shape estimation with only two light sources. Since our unified differential formulation does not depend on a specific camera setup or a chosen reflectance model, the most obvious target for future work is to move to a perspective projection, considering more complex reflectance models, exploiting better the information available in specular reflection and polarisation. In addition, since our methods directly estimate surface height, it would be straightforward to incorporate positional constraints, for example provided by binocular stereo.

Acknowledgements

This work was supported by the "GNCS - INdAM". Part of this work was undertaken while W. Smith was a Visiting Scholar at UCSD, supported by EPSRC grant EP/N028481/1.

References

[1] G. A. Atkinson and E. R. Hancock. Recovery of surface orientation from diffuse polarization. *IEEE Transactions on Image processing*, 15(6):1653–1664, 2006. 2, 3, 4

[2] G. A. Atkinson and E. R. Hancock. Shape estimation using polarization and shading from two views. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(11):2001–2017, 2007. 2

[3] G. A. Atkinson and E. R. Hancock. Surface reconstruction using polarization and photometric stereo. In *Proc. CAIP*, pages 466–473, 2007. 2

[4] P. N. Belhumeur, D. J. Kriegman, and A. Yuille. The Bas-relief ambiguity. *Int. J. Comput. Vision*, 35(1):33–44, 1999. 2

[5] J. F. Blinn. Models of light reflection for computer synthesized pictures. *Computer Graphics*, 11(2):192–198, 1977. 8

[6] O. Drbohlav and R. Šára. Unambiguous determination of shape from photometric stereo with unknown light sources. In *Proc. ICCV*, pages 581–586, 2001. 2

[7] A. Ecker and A. D. Jepson. Polynomial shape from shading. In *Proc. CVPR*, pages 145–152. IEEE, 2010. 1

[8] C. P. Huynh, A. Robles-Kelly, and E. Hancock. Shape and refractive index recovery from single-view polarisation images. In *Proc. CVPR*, pages 1229–1236, 2010. 2, 3

[9] C. P. Huynh, A. Robles-Kelly, and E. R. Hancock. Shape and refractive index from single-view spectro-polarimetric images. *Int. J. Comput. Vision*, 101(1):64–94, 2013. 2

[10] A. Kadambi, V. Taamazyan, B. Shi, and R. Raskar. Polarized 3D: High-quality depth sensing with polarization cues. In *Proc. ICCV*, 2015. 1, 2

[11] A. H. Mahmoud, M. T. El-Melegy, and A. A. Farag. Direct method for shape recovery from polarization and shading. In *Proc. ICIP*, pages 1769–1772, 2012. 2

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

[13] R. Mecca and Y. Quéau. Unifying diffuse and specular reflections for the photometric stereo problem. In *IEEE Workshop on Applications of Computer Vision (WACV), Lake Placid, USA, 2016*, 2016. 1

[14] D. Miyazaki, M. Kagesawa, and K. Ikeuchi. Transparent surface modeling from a pair of polarization images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 26(1):73–82, 2004. 2

[15] D. Miyazaki, R. T. Tan, K. Hara, and K. Ikeuchi. Polarization-based inverse rendering from a single view. In *Proc. ICCV*, pages 982–987, 2003. 2

[16] O. Morel, F. Meriaudeau, C. Stolz, and P. Gorria. Polarization imaging applied to 3D reconstruction of specular metallic surfaces. In *Proc. EI 2005*, pages 178–186, 2005. 2

[17] S. G. Narasimhan, V. Ramesh, and S. Nayar. A class of photometric invariants: Separating material from shape and illumination. In *Proc. ICCV*, volume 2, pages 1387 – 1394, 2003. 3

[18] S. Nayar, X. Fang, , and T. Boult. Separation of reflection components using color and polarization. *Int. J. Comput. Vision*, 21(3):163–186, 1997. 2

[19] T. T. Ngo, H. Nagahara, and R. Taniguchi. Shape and light directions from shading and polarization. In *Proc. CVPR*, pages 2310–2318, 2015. 2

[20] Y. Quéau, R. Mecca, and J. D. Durou. Unbiased photometric stereo for colored surfaces: A variational approach. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4359–4368, June 2016. 1

[21] Y. Quéau, R. Mecca, J.-D. Durou, and X. Descombes. Photometric stereo with only two images: A theoretical study and numerical resolution. *Image and Vision Computing*, 57:175–191, 2017. 1

[22] S. Rahmann and N. Canterakis. Reconstruction of specular surfaces using polarization imaging. In *Proc. CVPR*, 2001. 2

[23] W. Smith and F. Fang. Height from photometric ratio with model-based light source selection. *Computer Vision and Image Understanding*, 145:128 – 138, 2016. 1

[24] W. A. P. Smith, R. Ramamoorthi, and S. Tozza. Linear depth estimation from an uncalibrated, monocular polarisation image. In *Computer Vision - ECCV 2016*, volume 9912 of *Lecture Notes in Computer Science*, pages 109–125, 2016. 1, 2, 4, 5, 6, 7, 8

[25] V. Taamazyan, A. Kadambi, and R. Raskar. Shape from mixed polarization. *arXiv preprint arXiv:1605.02066*, 2016. 2, 6

[26] S. Tozza, R. Mecca, M. Duocastella, and A. Del Bue. Direct differential photometric stereo shape recovery of diffuse and specular surfaces. *Journal of Mathematical Imaging and Vision*, 56(1):57–76, 2016. 1

[27] L. B. Wolff. Polarization vision: a new sensory approach to image understanding. *Image Vision Comput.*, 15(2):81–93, 1997. 2, 3

[28] L. B. Wolff and T. E. Boult. Constraining object features using a polarization reflectance model. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(7):635—657, 1991. 2