Uncalibrated, Two Source Photo-Polarimetric Stereo

Silvia Tozza, Dizhong Zhu, William A. P. Smith,
Ravi Ramamoorthi, *Fellow, IEEE* and Edwin R. Hancock, *Fellow, IEEE*

Abstract—In this paper we present methods for estimating shape from polarisation and shading information, i.e. photo-polarimetric shape estimation, under varying, but unknown, illumination, i.e. in an uncalibrated scenario. We propose several alternative photo-polarimetric constraints that depend upon the partial derivatives of the surface and show how to express them in a unified system of partial differential equations of which previous work is a special case. By careful combination and manipulation of the constraints, we show how to eliminate non-linearities such that a discrete version of the problem can be solved using linear least squares. We derive a minimal, combinatorial approach for two source illumination estimation which we use with RANSAC for robust light direction and intensity estimation. We also introduce a new method for estimating a polarisation image from multichannel data and provide methods for estimating albedo and refractive index. We evaluate lighting, shape, albedo and refractive index estimation methods on both synthetic and real-world data showing improvements over existing state-of-the-art.

Index Terms—polarisation, shape-from-x, Bas-relief ambiguity, illumination estimation, differential approach, photometric stereo.

1 Introduction

A Recent trend in photometric [1]–[6] and physics-based [7]–[10] 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 [11]. In both photometric stereo [2], [3], [5] and recently in shape-from-polarisation (SfP) [7], [8], 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 well-performing solutions available under a range of different assumptions. Two source photometric stereo is still considered a difficult problem [12] even when the illumination is calibrated and albedo is known. We show that various formulations of one or 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 and refractive index 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 photo-polarimetric stereo. We make the following novel contributions:

- We show how to estimate a polarisation image from multichannel data (Sec. 3.2).
- We show how polarisation and photometric constraints (Sec. 4) can be expressed in a unified formulation (of which previous works [7]–[10] are special cases) and that various combinations of these constraints provide different practical advantages (Sec. 5) including estimation of both the diffuse albedo map and the refractive index.
- Degree and phase of polarisation do not vary with wavelength. We therefore exploit the additional constraint provided by colour images both for polarisation image (Sec. 3.2) and shape estimation (Sec. 5.5).
- We show how to estimate the illumination directions in two source photo-polarimetric data leading to an uncalibrated solution (Sec. 7). Our approach is robust and simultaneously estimates a diffuse mask that can be used to exclude specular pixels.

The proposed methods can be combined into a complete shape estimation pipeline as follows. First, estimate a polarisation image using the multichannel method in Sec. 3 and colour if available. Second, if illumination is uncalibrated, estimate two source direction and intensity using the method in Sec. 7. Third, make an initial shape estimate and subsequently albedo and refractive index estimates using the method in Sec. 5.2, again using colour if available. Finally, refine the shape estimate using estimated albedo and refractive index using the method in Sec. 5.3.

Paper organization We review related state-of-the-art works in Sec. 2 then in Sec. 3 we introduce the notations, the data, and the multichannel polarisation image estimation. In Sec. 4, we illustrate the different constraints and their

S. Tozza is with Department of Mathematics and Applications "Renato Caccioppoli", Universitá degli Studi di Napoli Federico II, 80126 Naples, Italy. Member of the INdAM Research group GNCS. Email: silvia.tozza@unina.it
D. Zhu, W. Smith and E. Hancock are with the Department of Computer Science, University of York, York, UK. Email: {dz761,william.smith,edwin.hancock}@york.ac.uk
R. Ramamoorthi is with the CSE Department at the University of California, San Diego, at La Jolla, CA 92093. Email: ravir@cs.ucsd.edu

linearisation in order to arrive at a unified partial differential system (Sec. 5), solved via linear least squares directly in the unknown height (Sec. 6). In addition to the albedo and refractive index estimation (Sec. 5), in Sec. 7 we propose our two source lighting estimation, followed by experiments (Sec. 8) and final comments (Sec. 9).

2 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, including early work on material segmentation [13] and diffuse/specular reflectance separation [14]–[16]. However, there has been a resurgent interest [7], [11], [17], [18] in using polarisation information for shape estimation.

An earlier version of the work in this paper was originally presented in [19]. Here, we additionally show how to estimate refractive index, reformulate the two light source estimation problem to avoid assuming fixed light source intensity and show how to make lighting estimation robust via a RANSAC-based approach that also estimates a diffuse mask. Concurrently and independently to [19], Mecca et al. [9] also proposed a differential approach to photo-polarimetric stereo. They use a different combination of constraints which we also show here is a special case of our general formulation and which we outperform.

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 constraint alone restricts the surface normal to two possible directions.

Both Atkinson and Hancock [20] and Miyazaki et al. [21] solve the problem of disambiguating these polarisation normals via propagation from the boundary under an assumption of global convexity. Huynh et al. [22] 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. [23] use a specular polarisation model for metals. Taamazyan et al. [17] introduced a mixed specular/diffuse polarisation model. All of these methods estimate surface normals that must be integrated into a height map and 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. [24] use 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 [25] use calibrated, three source Lambertian photometric stereo for disambiguation but avoiding an assumption of known albedo. Smith et al. [7] show how to express polarisation and shading constraints directly in terms of surface height, leading to an efficient linear least squares solution. In [19] and in this work we give an explanation for why the matrix they consider is full-rank except in a unique case. They also show how to estimate the illumination, up to a binary ambiguity, making the method uncalibrated. However, they require known or uniform albedo.

This requirement was subsequently relaxed [8] via a procedure to estimate albedo directly from a polarisation image, though it requires known illumination and strong smoothness assumptions. 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, even when lighting is unknown. Ngo et al. [18] derive constraints that allow surface normals, light directions and refractive index to be estimated from polarisation images under varying lighting. However, this approach requires at least 4 lights. Yu et al. [26] pose shape-from-polarisation as an analysis-by-synthesis problem, directly optimising the recovered shape using nonlinear least squares. While flexible, this approach is prone to converging on poor local minima. Atkinson [27] combines polarisation phase information with constraints from calibrated two source photometric stereo via a region growing process.

Polarisation plus X Coarse geometry obtained by multi-view space carving [28], [29] has been used to resolve polarisation ambiguities. Huynh et al. [30] extend their earlier work to use multispectral measurements to estimate both shape and refractive index. Drbohlav and Sara [31] show how the Bas-relief ambiguity [32] in uncalibrated photometric stereo could be resolved using polarisation. However, this approach requires a polarised light source. Kadambi et al. [11], [33] combine a single polarisation image with a depth map obtained by an RGBD camera. Zhu and Smith [34] add an additional non-polarised camera giving a stereo cue from which they can estimate a rough depth map. The depth map is used to disambiguate the surface normals and provide a base surface for integration.

Multiview Polarisation Early work used a stereo pair of polarisation measurements to determine the orientation of a plane [35]. Rahmann and Canterakis [36] combine a specular polarisation model with stereo cues. Similarly, Atkinson and Hancock [37] use polarisation normals to segment an object into patches, simplifying stereo matching. Stereo polarisation cues have also been used for transparent surface modelling [38]. Berger et al. [39] use polarisation stereo for depth estimation of specular scenes. Cui et al. [40] incorporate a polarisation phase angle cue into multiview stereo enabling recovery of surface shape in featureless regions. Chen et al. [41] provide a theoretical treatment of constraints arising from three view polarisation. Yang et al. [42] propose a variant of monocular SLAM using polarisation video. Cui et al. [43] use polarisation constraints to reduce the number of correspondences required for relative pose estimation. While these methods all require multiple polarisation cameras or a moving camera, we focus on a single viewpoint but with varying illumination direction. The two approaches are likely to be complimentary.

3 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. 5 we propose a unified formulation for a variety of methods, all of which assume orthographic projection. Other assumptions will be given later on, depending on the specific problem

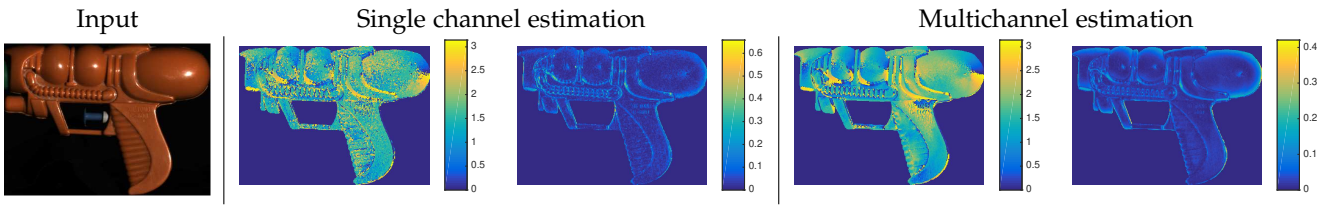

Fig. 1: Multichannel polarisation image estimation. Left to right: an image from the input sequence; phase angle (ϕ) and degree of polarisation (ρ) estimated from a single channel; phase angle (ϕ) and degree of polarisation (ρ) estimated from three colour channels and two light source directions.

at hand. We denote by $\mathbf{v} \in \mathbb{R}^3$ the viewer direction, by $\mathbf{s} \in \mathbb{R}^3$ a general light source direction with $\mathbf{v} \neq \mathbf{s}$. We only require the third components of these unit column vectors to be greater than zero (i.e. all the vectors belong to the upper hemisphere). We will denote by $\mathbf{t} \in \mathbb{R}^3$ a second light source where required. Note that the intensity of each light source can be encoded in its magnitude, $\|\mathbf{s}\|$ and $\|\mathbf{t}\|$. We parametrise the unknown surface height by the function $z(\mathbf{x})$, where $\mathbf{x} = [x, y]^T$ 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 $z(\mathbf{x})$ with respect to x and y, respectively, so that $\nabla z(\mathbf{x}) = [z_x, z_y]^T$. We now introduce relevant polarisation theory, describing how we can estimate a polarisation image from multichannel data.

3.1 Polarisation image

When unpolarised light is reflected by a surface it becomes partially polarised [44]. 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 [44]: 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 [20], linear methods [22] or via a closed form solution [44] for the specific case of $P = 3$, $\vartheta \in \{0°, 45°, 90°\}$.

3.2 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. [45], 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

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

The system of equations is linear in the unpolarised intensities and, by a change of variables, can be made linear in ρ and ϕ [22]. Hence, we wish to solve a bilinear system and do so in a least squares sense using interleaved alternating minimisation. In detail, we:

1) fix ρ and ϕ and then solve linearly for the unpolarised intensity in each channel,
2) 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:

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

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

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

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

$$\mathbf{d}_I = \left[i^1_{\vartheta_1}(\mathbf{x}), \ldots, i^C_{\vartheta_1}(\mathbf{x}), i^1_{\vartheta_2}(\mathbf{x}), \ldots, i^C_{\vartheta_P}(\mathbf{x}) \right]^T. \tag{6}$$

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

$$\min_{a,b} \left\| \mathbf{C}_{\rho\phi} \begin{bmatrix} a \\ b \end{bmatrix} - \mathbf{d}_{\rho\phi} \right\|^2, \tag{7}$$

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

$$\mathbf{C}_{\rho\phi} = \begin{bmatrix} i^1_{\text{un}}(\mathbf{x})\cos(2\vartheta_1) & i^1_{\text{un}}(\mathbf{x})\sin(2\vartheta_1) \\ \vdots & \vdots \\ i^1_{\text{un}}(\mathbf{x})\cos(2\vartheta_P) & i^1_{\text{un}}(\mathbf{x})\sin(2\vartheta_P) \\ i^2_{\text{un}}(\mathbf{x})\cos(2\vartheta_1) & i^2_{\text{un}}(\mathbf{x})\sin(2\vartheta_1) \\ \vdots & \vdots \\ i^C_{\text{un}}(\mathbf{x})\cos(2\vartheta_P) & i^C_{\text{un}}(\mathbf{x})\sin(2\vartheta_P) \end{bmatrix}, \tag{8}$$

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

$$\mathbf{d}_{\rho\phi} = \begin{bmatrix} i^1_{\vartheta_1}(\mathbf{x}) - i^1_{\text{un}}(\mathbf{x}) \\ \vdots \\ i^1_{\vartheta_P}(\mathbf{x}) - i^1_{\text{un}}(\mathbf{x}) \\ i^2_{\vartheta_1}(\mathbf{x}) - i^2_{\text{un}}(\mathbf{x}) \\ \vdots \\ i^C_{\vartheta_P}(\mathbf{x}) - i^C_{\text{un}}(\mathbf{x}) \end{bmatrix}. \tag{9}$$

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 [22], 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 Fig. 1. The multichannel result is visibly less noisy than the single channel performance. We show an additional result with ground truth in Fig. 2 where we use three colour channels and two light source directions. The left column shows the estimated degree of polarisation and the right one shows the estimated phase angle. It is clear that the multichannel result is closer to ground truth than using either light source alone.

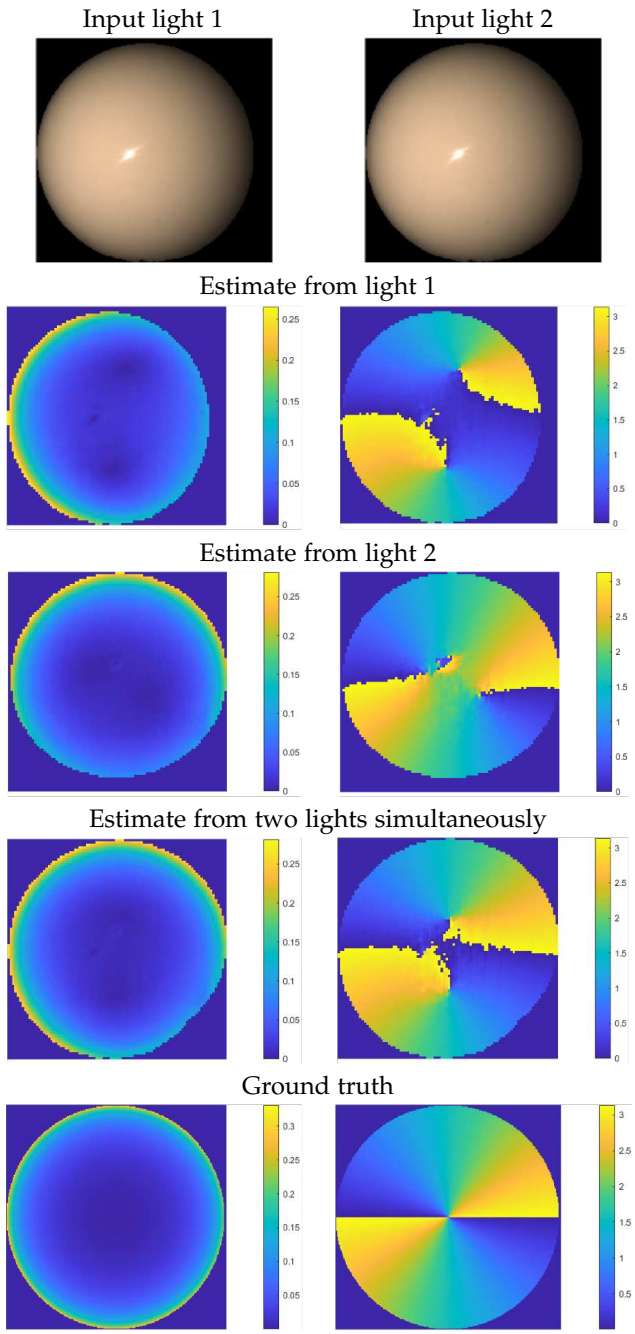

Fig. 2: Multichannel polarisation image estimation with two light sources. Row 1: input images with two different light sources. Row 2/3: Polarisation image estimation from first/second light. Row 4: Polarisation image estimation from both lights. Row 5: ground truth polarisation image.

4 Photo-polarimetric height constraints

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

4.1 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 [20] for more details). The degree of diffuse polarisation $\rho_d(\mathbf{x})$ at each point $\mathbf{x}$ can be expressed in terms of the refractive index η and the surface zenith angle $\theta \in [0, \frac{\pi}{2}]$ as follows (Cf. [20]):

$$\rho_d(\mathbf{x}) = \frac{(\eta - 1/\eta)^2 \sin^2(\theta)}{2+2\eta^2-(\eta+1/\eta)^2\sin^2(\theta)+4\cos(\theta)\sqrt{\eta^2-\sin^2(\theta)}}. \tag{10}$$

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 (10) 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 as follows:

$$\cos(\theta) = \mathbf{n}(\mathbf{x})\cdot\mathbf{v} = f(\rho_d(\mathbf{x}),\eta) = \tag{11}$$
$$\sqrt{\frac{\eta^4(1-\rho_d^2)+2\eta^2(2\rho_d^2+\rho_d-1)+\rho_d^2+2\rho_d-4\eta^3\rho_d\sqrt{1-\rho_d^2}+1}{(\rho_d+1)^2(\eta^4+1)+2\eta^2(3\rho_d^2+2\rho_d-1)}}$$

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

4.2 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 [7] 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:

$$i_{\text{un}}(\mathbf{x}) = \gamma(\mathbf{x})\cos(\theta_i) = \gamma(\mathbf{x})\mathbf{n}(\mathbf{x})\cdot\mathbf{s}, \tag{12}$$

where $\gamma(\mathbf{x})$ is the albedo. Writing $\mathbf{n}(\mathbf{x})$ in terms of the gradient of z as reported in (1), (12) can be rewritten as follows:

$$i_{\text{un}}(\mathbf{x}) = \gamma(\mathbf{x})\frac{-\nabla z(\mathbf{x})\cdot\tilde{\mathbf{s}} + s_3}{\sqrt{1+|\nabla z(\mathbf{x})|^2}}, \tag{13}$$

with $\tilde{\mathbf{s}} = [s_1, s_2]^T$. This is a non-linear equation, but we will see in Sec. 4.4 and 4.5 how it is possible to remove the non-linearity by using the ratios technique.

4.3 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 π ambiguity. This constraint can be rewritten as a collinearity condition [7], 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, $[n_x\ n_y]^T$, and a vector in the image plane pointing in the phase angle direction, $[\sin(\phi)\ \cos(\phi)]^T$, to be collinear. This corresponds to requiring

$$\mathbf{n}(\mathbf{x})\cdot[\cos(\phi(\mathbf{x}))\ -\sin(\phi(\mathbf{x}))\ 0]^T = 0. \tag{14}$$

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

$$[-\cos\phi, \sin\phi]^T\cdot\nabla z = 0. \tag{15}$$

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. 5.2.

4.4 Degree of polarisation ratio constraint

Combining the two constraints illustrated in Sec. 4.1 and 4.2, 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 (11) we obtain

$$\sqrt{1+|\nabla z(\mathbf{x})|^2} = \frac{-\nabla z(\mathbf{x})\cdot\tilde{\mathbf{v}} + v_3}{f(\rho_d(\mathbf{x}),\eta)}, \tag{16}$$

where $\tilde{\mathbf{v}} = [v_1, v_2]^T$. On the other hand, considering the shading information contained in (13) and again isolating the non-linearity, we arrive at the following

$$\sqrt{1+|\nabla z(\mathbf{x})|^2} = \gamma(\mathbf{x})\frac{-\nabla z(\mathbf{x})\cdot\tilde{\mathbf{s}} + s_3}{i_{\text{un}}(\mathbf{x})}. \tag{17}$$

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. (16) and (17) we obtain

$$\frac{-\nabla z(\mathbf{x})\cdot\tilde{\mathbf{v}} + v_3}{f(\rho_d(\mathbf{x}),\eta)} = \gamma(\mathbf{x})\frac{-\nabla z(\mathbf{x})\cdot\tilde{\mathbf{s}} + s_3}{i_{\text{un}}(\mathbf{x})}. \tag{18}$$

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

$$\mathbf{b}(\mathbf{x})\cdot\nabla z(\mathbf{x}) = h(\mathbf{x}), \tag{19}$$

where

$$\mathbf{b}(\mathbf{x}) := \mathbf{b}^{(f,i_{\text{un}})} = i_{\text{un}}(\mathbf{x})\tilde{\mathbf{v}} - \gamma(\mathbf{x})f(\rho_d(\mathbf{x}),\eta)\,\tilde{\mathbf{s}}, \tag{20}$$

and

$$h(\mathbf{x}) := h^{(f,i_{\text{un}})} = i_{\text{un}}(\mathbf{x})v_3 - \gamma(\mathbf{x})f(\rho_d(\mathbf{x}),\eta)\,s_3. \tag{21}$$

4.5 Intensity ratio constraint

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 (12) twice, once for each light source. We can remove the non-linearity as before and take a ratio, arriving at the following equation:

$$i_{\text{un},2}(\mathbf{x})(-\nabla z(\mathbf{x})\cdot\tilde{\mathbf{s}} + s_3) = i_{\text{un},1}(\mathbf{x})(-\nabla z(\mathbf{x})\cdot\tilde{\mathbf{t}} + t_3). \tag{22}$$

The above equation is independent of albedo and the non-linear normalisation term. It is also invariant to the absolute light source intensities $\|\mathbf{s}\|$ and $\|\mathbf{t}\|$ but does depend on their relative intensities, i.e. the ratio $\|\mathbf{s}\|/\|\mathbf{t}\|$. Again as before, we can rewrite (22) as a PDE in the form of (19) with

$$\mathbf{b}(\mathbf{x}) := \mathbf{b}^{(i_{\text{un},1},i_{\text{un},2})} = i_{\text{un},2}(\mathbf{x})\tilde{\mathbf{s}} - i_{\text{un},1}(\mathbf{x})\,\tilde{\mathbf{t}}, \tag{23}$$

where $\tilde{\mathbf{t}} = [t_1, t_2]^T$, and

$$h(\mathbf{x}) := h^{(i_{\text{un},1},i_{\text{un},2})} = i_{\text{un},2}(\mathbf{x})s_3 - i_{\text{un},1}(\mathbf{x})\,t_3. \tag{24}$$

4.6 Polariser angle ratio

Mecca et al. [9], [10] propose taking a ratio between observations (i.e. (2)) for polariser angles $\vartheta = 0$ and $\vartheta = \frac{\pi}{4}$. This leads to the following constraint:

$$z_x(-i_{\frac{\pi}{4}}(\mathbf{x}) + i_{\text{un}}(\mathbf{x})) + z_y(i_0(\mathbf{x}) - i_{\text{un}}(\mathbf{x}) + i_{\text{un}}(\mathbf{x})\rho(\mathbf{x})) = 0. \tag{25}$$

We now re-express this in our formulation allowing comparison to be drawn to our other constraints. By substituting (2) with the appropriate angles, we obtain:

$$\begin{aligned} -i_{\frac{\pi}{4}}(\mathbf{x}) &= -i_{\text{un}}(\mathbf{x})(1+\rho(\mathbf{x})\cos(2\frac{\pi}{4} - 2\phi(\mathbf{x})) \\ &= -i_{\text{un}}(\mathbf{x})(1+\rho(\mathbf{x})\sin(2\phi(\mathbf{x}))), \end{aligned} \tag{26}$$

and

$$\begin{aligned} &i_0(\mathbf{x}) - i_{\text{un}}(\mathbf{x}) + i_{\text{un}}(\mathbf{x})\rho(\mathbf{x}) \\ =&i_{\text{un}}(\mathbf{x})(1+\rho(\mathbf{x})\cos(-2\phi(\mathbf{x}))) - i_{\text{un}}(\mathbf{x}) + i_{\text{un}}(\mathbf{x})\rho(\mathbf{x}) \\ =&i_{\text{un}}(\mathbf{x})\rho(\mathbf{x})(1+\cos(2\phi(\mathbf{x}))). \end{aligned} \tag{27}$$

Substituting these into (25) we obtain:

$$\begin{bmatrix} -i_{\text{un}}(\mathbf{x})\rho(\mathbf{x})\sin(2\phi(\mathbf{x})) \\ i_{\text{un}}(\mathbf{x})\rho(\mathbf{x})(1+\cos(2\phi(\mathbf{x}))) \end{bmatrix}\cdot\nabla z(\mathbf{x}) = 0. \tag{28}$$

Again, we can write (28) as a PDE in the form of (19) with:

$$\mathbf{b}(\mathbf{x}) := \mathbf{b}^{(i_{\text{un}},\rho,\phi)} = \begin{bmatrix} -i_{\text{un}}(\mathbf{x})\rho(\mathbf{x})\sin(2\phi(\mathbf{x})) \\ i_{\text{un}}(\mathbf{x})\rho(\mathbf{x})(1+\cos(2\phi(\mathbf{x}))) \end{bmatrix} \quad (29)$$

and $h = 0$. We now show that this constraint is equivalent to the phase angle constraint but weighted by $i_{\text{un}}(\mathbf{x})\rho(\mathbf{x})$. Expanding $\sin(2\phi)$ and $\cos(2\phi)$ by trigonometric identity we obtain:

$$2i_{\text{un}}(\mathbf{x})\rho(\mathbf{x}) \cdot \begin{bmatrix} -\sin(\phi)\cos(\phi) \\ \cos^2(\phi) \end{bmatrix} \cdot \nabla z(\mathbf{x}) = 0, \quad (30)$$

and removing the scale and common factor $\cos(\phi)$ we obtain a weighted version of (15):

$$i_{\text{un}}(\mathbf{x})\rho(\mathbf{x}) \cdot \begin{bmatrix} -\sin(\phi) \\ \cos(\phi) \end{bmatrix} \cdot \nabla z(\mathbf{x}) = 0. \quad (31)$$

Relative to the unweighted (15), we speculate that this weighting is not helpful. When the degree of polarisation is small, the surface normal is close to the viewer direction and the elements of the gradient vector are already small. This already introduces a weighting that places more importance on pixels with high degree of polarisation (where the signal for estimating phase angle is stronger). On the other hand, weighting by i_{un} may give low weight to pixels near the occluding boundary (if the light source is close to frontal) where the azimuth cue provided by the phase angle is very important and, since the degree of polarisation is large in those regions, the phase estimate is reliable.

5 A UNIFIED PDE FORMULATION

Starting from the constraints introduced in Sec. 4, 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:

$$\mathbf{B}(\mathbf{x})\nabla z(\mathbf{x}) = \mathbf{h}(\mathbf{x}), \quad (32)$$

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 $J = 2$ or $J = 4$ depending on the cases. Note that (32) does not depend on a specific camera setup or a chosen reflectance model. It is both a very compact and general equation, suitable for describing several cases in a unified differential formulation that solves directly for surface height.

Different combinations of the constraints described in Sec. 4 that are linear in the surface gradient can be combined in the formulation of (32). Each corresponds to different assumptions and have different pros and cons. We explore two new variants and show that [7] and [9], [10] are special cases. We summarise the methods considered in Tab. 1.

Method	Phase angle	DOP ratio	Intensity ratio	Pol. ang. ratio
Smith et al. [7]	✓	✓		
Ours (invariant)	✓		✓	
Ours (alternating)	✓	✓	✓	
Mecca et al. [9], [10]			✓	✓

TABLE 1: Summary of the combinations of constraints used by the different methods. All can be expressed in the same unified PDE formulation.

5.1 Single light and polarisation formulation

This case has been studied in [7]. It uses a single polarisation image, requires known illumination (though [7] 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 (32), where $\mathbf{B}$ and $\mathbf{h}$ are defined as

$$\mathbf{B}(\mathbf{x}) = \begin{bmatrix} b_1^{(f,i_{\text{un}})} & b_2^{(f,i_{\text{un}})} \\ -\cos\phi & \sin\phi \end{bmatrix}, \quad \mathbf{h}(\mathbf{x}) = [h^{(f,i_{\text{un}})}, 0]^T, \quad (33)$$

with $\mathbf{b}^{(f,i_{\text{un}})}$ and $h^{(f,i_{\text{un}})}$ defined by (20) and (21), with uniform $\gamma(\mathbf{x})$ and $\mathbf{v} = [0,0,1]^T$.

5.2 Albedo and refractive index invariant formulation

Our first proposed method uses the phase angle constraint (15) and two unpolarised images, taken from two different light source directions, obtained through (13) and combined as in (22). In this case the problem studied is described by the system of PDEs (32) with

$$\mathbf{B}(\mathbf{x}) = \begin{bmatrix} b_1^{(i_{\text{un},1},i_{\text{un},2})} & b_2^{(i_{\text{un},1},i_{\text{un},2})} \\ -\cos\phi & \sin\phi \end{bmatrix}, \mathbf{h}(\mathbf{x}) = \begin{bmatrix} h^{(i_{\text{un},1},i_{\text{un},2})} \\ 0 \end{bmatrix}, \quad (34)$$

where $\mathbf{b}^{(i_{\text{un},1},i_{\text{un},2})}$ and $h^{(i_{\text{un},1},i_{\text{un},2})}$ are defined as in (23) and (24). 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. In addition, there is no dependence on refractive index. We now describe how both albedo and refractive index can be estimated after using this combination of constraints for height estimation.

Albedo estimation Going further, this method can be used to allow estimation of an albedo map. 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 (12). Where we have two diffuse observations, we can compute albedo from two equations of the form of (12) 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 gradient of the intensity image for diffuse pixels. Relative to the albedo estimation method in [8], our approach is more robust, offers improved performance and does not require known lighting.

Refractive index estimation A further advantage of this formulation is that it does not depend upon the degree of polarisation and therefore does not use a polarisation model with refractive index as a parameter. For this reason, it also provides a route to refractive index estimation. Again, we first estimate surface height in a refractive index invariant manner. Then, with (1), we can retrieve the zenith angle value from the surface height as $\theta(\mathbf{x}) = \arccos(n_z(\mathbf{x}))$. We also have an observed degree of polarisation, ρ_{obs}, as

computed by the method described in Sec. 3.2. The DOP equation (10) models the relationship between zenith angle θ and refractive index η. Hence, with the observed degree of polarisation and zenith angle computed from the estimated surface, an optimal refractive index can be estimated by solving the following nonlinear least squares problem:

$$\min_{\eta} \sum_{\mathbf{x}} \|\rho_{\text{obs}}(\mathbf{x}) - \rho_d(\theta(\mathbf{x}), \eta)\|_2^2, \tag{35}$$

where we assume the refractive index is uniform over the target object.

5.3 Alternating formulation

We now propose a second formulation that combines the three previous constraints described in Sections 4.3, 4.4 and 4.5, leading to a problem of the form (32) with

$$\mathbf{B} = \begin{bmatrix} b_1^{(f,i_{\text{un},1})} & b_2^{(f,i_{\text{un},1})} \\ b_1^{(f,i_{\text{un},2})} & b_2^{(f,i_{\text{un},2})} \\ b_1^{(i_{\text{un},1},i_{\text{un},2})} & b_2^{(i_{\text{un},1},i_{\text{un},2})} \\ -\cos\phi & \sin\phi \end{bmatrix}, \ \mathbf{h} = \begin{bmatrix} h^{(f,i_{\text{un},1})} \\ h^{(f,i_{\text{un},2})} \\ h^{(i_{\text{un},1},i_{\text{un},2})} \\ 0 \end{bmatrix}. \tag{36}$$

This formulation provides maximum constraint and exploits all parameters of the fitted sinusoid so is potentially more robust than the previous invariant method. However, it requires known albedo and refractive index in order to use the DOP ratio constraint. Nevertheless, it is possible to first apply the invariant formulation, estimate the albedo and refractive index 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 until convergence. This iterative approach, initialised by the invariant method, is the method that we evaluate in our experiments.

5.4 Angular and intensity ratios formulation

This case has been studied by Mecca et al. [9], [10] and combines the intensity ratio constraint in (22) with the polariser angle ratio in (28), repeated for the two unpolarised intensity measurements from the two images, arriving to a PDE system of the form (32) with

$$\mathbf{B}(\mathbf{x}) = \begin{bmatrix} b_1^{(i_{\text{un},1},i_{\text{un},2})} & b_2^{(i_{\text{un},1},i_{\text{un},2})} \\ b_1^{(i_{\text{un},1},\rho,\phi)} & b_2^{(i_{\text{un},1},\rho,\phi)} \\ b_1^{(i_{\text{un},2},\rho,\phi)} & b_2^{(i_{\text{un},2},\rho,\phi)} \end{bmatrix}, \quad \mathbf{h}(\mathbf{x}) = \begin{bmatrix} h^{(i_{\text{un},1},i_{\text{un},2})} \\ 0 \\ 0 \end{bmatrix}, \tag{37}$$

where $\mathbf{b}^{(i_{\text{un},1},i_{\text{un},2})}$ and $h^{(i_{\text{un},1},i_{\text{un},2})}$ are defined as in (23) and (24) and $\mathbf{b}^{(i_{\text{un}},\rho,\phi)}$ is defined as in (29).

5.5 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 (12). 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. 3.2. 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.

6 Linear least squares height estimation

We have seen that each of the variants illustrated in the previous section, each with different advantages, can be written as a PDE system (32). Denoting by M the number of pixels, we discretise the gradient in (32) using the method of [46] arriving at the following linear system in $\mathbf{z}$

$$\mathbf{A}\mathbf{z} = \bar{\mathbf{h}}, \tag{38}$$

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. 5 ($J = 2$ for (33) and (34), $J = 3$ for (37) and $J = 4$ for (36)). $\bar{\mathbf{h}}$ is the discrete per-pixel version of the function $\mathbf{h}(\mathbf{x})$, $\bar{\mathbf{h}} \in \mathbb{R}^{JM \times 1}$, and $\mathbf{z} \in \mathbb{R}^{M \times 1}$ is 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. 5, under the different assumptions specified for each case. The per-pixel matrix $\mathbf{A}$ related to [7] 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 (38) in the least-squares sense, i.e. find a vector $\mathbf{z} \in \mathbb{R}^M$ such that

$$||\mathbf{A}\mathbf{z} - \bar{\mathbf{h}}||_2^2 \leq ||\mathbf{A}\mathbf{y} - \bar{\mathbf{h}}||_2^2, \quad \forall \mathbf{y} \in \mathbb{R}^M. \tag{39}$$

Considering the associated system of normal equations

$$\mathbf{A}^T(\mathbf{A}\mathbf{z} - \bar{\mathbf{h}}) = 0, \tag{40}$$

it is well-known that if there exists $\mathbf{z} \in \mathbb{R}^M$ that satisfies (40), then $\mathbf{z}$ is also solution of the least-squares problem, i.e. $\mathbf{z}$ satisfies (39). 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 (40) for each data term $\bar{\mathbf{h}}$. Since neither $\mathbf{B}$ nor $\mathbf{h}$ depend on z in (32), 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.

7 Two source lighting estimation

All of the shape-from-polarisation methods in Table 1 require knowledge of the illumination (in the case of our proposed methods and that of Mecca et al. [9], [10], this amounts to two point light source directions).

Previously, Smith et al. [7] showed that a single polarisation image can be used for illumination estimation up to a binary ambiguity. However, the approach has three important weaknesses. First, they assumed that the albedo was known or uniform which is a significant practical limitation. Second, they used a least squares solution over all pixels meaning the method is not robust to outliers. Third, they assumed that a specular mask was provided.

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 solve the problem using Random Sample Consensus (RANSAC) [47] meaning the approach is robust to outliers, specifically specular pixels, so that we can automatically estimate a diffuse mask. Using this approach, two source photo-polarimetric stereo can be applied in an uncalibrated scenario.

7.1 Homogeneous system of linear equations

Suppose that we capture two images with different, unknown light directions and have noisy estimates of the surface gradient to hand. The intensity ratio (22) provides one equation per pixel relating unpolarised intensities, surface gradient and light source directions. Hence, we can form a system of linear equations in the unknown light source vectors $\mathbf{s}$ and $\mathbf{t}$:

$$\mathbf{R}\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix} = \mathbf{0}\,, \tag{41}$$

where each row of $\mathbf{R}$ is of the form:

$$[-z_x i_{\text{un},2}, -z_y i_{\text{un},2}, i_{\text{un},2}, z_x i_{\text{un},1}, z_y i_{\text{un},1}, -i_{\text{un},1}]. \tag{42}$$

Note that this system is homogeneous and so has a trivial solution $\mathbf{s} = \mathbf{t} = [0,0,0]^T$. In our earlier work [19] we reformulated this in spherical coordinates, assuming that the light sources had the same intensity and solved using nonlinear least squares. Here, we instead solve it as a homogeneous linear least squares problem:

$$\min_{\mathbf{s},\mathbf{t}\in\mathbb{R}^3} \left\|\mathbf{R}\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix}\right\|_2^2, \quad \text{s.t.} \left\|\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix}\right\|_2 = 1. \tag{43}$$

This has a closed form solution by taking the SVD of $\mathbf{R} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ with the solution given by the last column of $\mathbf{V}$. Note that there is an overall scale ambiguity between the diffuse albedo and the intensity of the light sources. The unit norm constraint here arbitrarily resolves this but does not require that the two light sources are of equal intensity. Moreover, using the SVD we guarantee to obtain the global optimum which was not the case for the optimisation in [19]. Note that there is also a sign ambiguity since:

$$\left\|\mathbf{R}\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix}\right\| = \left\|-\mathbf{R}\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix}\right\|. \tag{44}$$

This is easily resolved since we know intensities must be positive and therefore Lambertian shading must be positive. The two unknown light vectors have six degrees of freedom, though this is reduced to five by the unit norm constraint.

Algorithm 1 Minimal, brute force solution for lighting

Input: Six pixel locations, $\mathbf{x}_1, \ldots, \mathbf{x}_6$
Output: Estimated light source directions, $\mathbf{s}, \mathbf{t} \in \mathbb{R}^3$

// Generate all binary strings[1] *of length 6*
$\mathbf{P} := \text{binaryStrings}(6)$
// $P_{i,j}$ is the jth digit of the ith string
$\beta = \infty$
for $i := 1$ **to** 2^6 **do**
  // Generate ith possible system of equations
  for $j := 1$ **to** 6 **do**
    $z_x := \begin{cases}\cos\phi(\mathbf{x}_j)\tan\theta(\mathbf{x}_j) & \text{if } \mathbf{P}_{i,j} = 0\\ -\cos\phi(\mathbf{x}_j)\tan\theta(\mathbf{x}_j) & \text{otherwise}\end{cases}$
    $z_y := \begin{cases}\sin\phi(\mathbf{x}_j)\tan\theta(\mathbf{x}_j) & \text{if } \mathbf{P}_{i,j} = 0\\ -\sin\phi(\mathbf{x}_j)\tan\theta(\mathbf{x}_j) & \text{otherwise}\end{cases}$
    $\mathbf{R}_j := [i_{\text{un},2}(\mathbf{x}_j)[-z_x, -z_y, 1], i_{\text{un},1}(\mathbf{x}_j)[z_x, z_y, -1]]$
  end for
  $(\mathbf{U}, \mathbf{S}, \mathbf{V}) := \text{svd}(\mathbf{R})$
  // Minimum norm solution given by last column of V
  $(\mathbf{s}, \mathbf{t}) := \mathbf{V}_6$
  $r := \left\|\mathbf{R}\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix}\right\|_2^2$
  if $r < \beta$ **then**
    $\beta := r$
    $(\mathbf{s}_{\text{best}}, \mathbf{t}_{\text{best}}) := (\mathbf{s}, \mathbf{t})$
  end if
end for
return $\mathbf{s}_{\text{best}}, \mathbf{t}_{\text{best}}$

7.2 Brute force minimal solution

Estimating light source direction using (43) requires estimates of the surface gradient for each pixel in the system of equations. In practice, the polarisation images do not provide estimates of the surface gradient directly. Instead, the phase angle and the zenith angle, estimated from the degree of polarisation using (11), define two possible surface normal directions at each pixel and therefore two possible surface gradients. We take a combinatorial approach by considering all 2^M possible disambiguations of the polarisation measurements for a given set of M pixels. While this has exponential complexity in the number of pixels, it is feasible for a *minimal* solution of $M = 6$ pixels. In Algorithm 1 we sketch this brute force minimal solution. We consider all 2^6 disambiguations, solve for the two light source directions for each and keep track of the solution with lowest residual.

7.3 Robust lighting estimation with RANSAC

We now employ the brute force minimal solution within the RANSAC algorithm in order to robustly estimate light source direction while simultaneously identifying outlier pixels. Besides random noise, we also expect that pixels in which one or both of the two intensity observations are a specular reflection will be outliers since the observed intensity and polarisation will deviate from our diffuse shading and polarisation models. Hence, we refer to the inlier mask

1. The function 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.

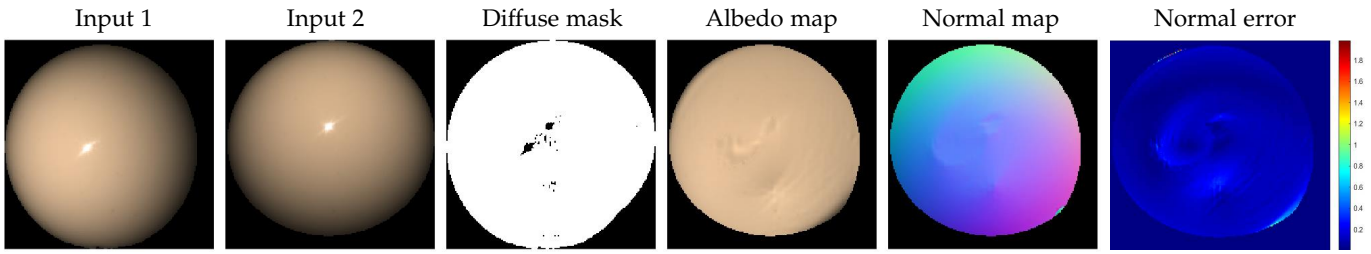

Fig. 3: An illustration of the RANSAC two source light estimation. From left to right: the input images under light source direction 1 and 2, the estimated diffuse mask, albedo map and normal map, and normal direction error map. Light estimation results are shown in Tab. 2.

as a diffuse mask, where we expect such pixels are diffuse dominant for both light source directions.

Following the RANSAC algorithm, we select a random minimal subset of 6 pixels and apply Algorithm 1 to obtain an estimate of the two light source directions. We then find the consensus set for this solution. For all pixels we compute the residual using the estimated light source directions and the disambiguation that minimises the residual. Specifically, for each pixel we define $\mathbf{r}(\mathbf{x})$ using (42) with the surface gradient arbitrarily resolved to one of the two possible disambiguations: $z_x(\mathbf{x}) = \cos\phi(\mathbf{x})\tan\theta(\mathbf{x})$ and $z_y(\mathbf{x}) = \sin\phi(\mathbf{x})\tan\theta(\mathbf{x})$. Transformation to the other possible disambiguation is achieved by $\mathbf{r}(\mathbf{x})\mathbf{T}$ where $\mathbf{T} = \text{diag}(-1,-1,1,-1,-1,1)$. Hence the residual is given by:

$$r(\mathbf{x},\mathbf{s},\mathbf{t}) = \min\left(\left\|\mathbf{r}(\mathbf{x})\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix}\right\|_2^2, \left\|\mathbf{r}(\mathbf{x})\mathbf{T}\begin{bmatrix}\mathbf{s}\\ \mathbf{t}\end{bmatrix}\right\|_2^2\right) \quad (45)$$

We now define the consensus set of inliers as those pixels whose residual is below a chosen threshold τ:

$$\mathcal{C}_{\mathbf{s},\mathbf{t}} = \{i | r(\mathbf{x}_i,\mathbf{s},\mathbf{t}) < \tau\}. \quad (46)$$

We repeat this process, keeping track of the largest inlier set observed. Finally, we use the $(\mathbf{s},\mathbf{t})$ associated with the largest inlier set to disambiguate the polarisation normals for each inlier pixel and then solve least squares for the final light source estimate, $(\mathbf{s}^*,\mathbf{t}^*)$, using all inlier pixels by solving (43). The diffuse mask is defined as $\mathcal{D} = \mathcal{C}_{\mathbf{s}^*,\mathbf{t}^*}$.

As in [7], the solution is still subject to a binary ambiguity, in that if $[\mathbf{s}^T,\mathbf{t}^T]$ is a solution then $\mathbf{T}[\mathbf{s}^T,\mathbf{t}^T]$ is also a solution, corresponding to the convex/concave ambiguity. We resolve this simply by choosing the maximal solution when surface height is later recovered.

8 Experiments

Data We now evaluate our light source, albedo, refractive index and shape estimation methods on both synthetic and real data. There are no existing benchmarks for these estimation tasks on polarisation images. For this reason, we propose to use synthetic data for which ground truth is known and real data for which ground truth can be inferred (in the case of a sphere) or measured (in the case of light source direction) or provides only a qualitative evaluation (in the case of albedo and shape on complex objects).

For synthetic data we use the 3DRFE dataset [48] comprising high quality surface normal and albedo maps captured from real faces using a lightstage. We render synthetic polarisation images from these measured normal and albedo maps. With a chosen light source direction, we first compute unpolarised intensities using (12). We simulate the effect of polarisation according to (2), varying the polariser angle between $0°$ and $180°$ in $30°$ increments. Next, we corrupt this data by adding zero mean Gaussian noise, saturate (i.e. clamp to 1) and quantise to 8 bits. This noisy data provides synthetic input. Our real data are collected in a dark room using a Canon 7D placed approximately 2m from the object with a 50mm lens and lens-mounted rotating linear polarising filter which is rotated to the same set of angles as the synthetic data. We use a collimated 200-W tungsten lamp, placed approximately 2m from the object. In order to avoid radiometric instability of lighting output over time, we were careful to allow the light sources a warm up period and verified that the variation in their output was much smaller than other uncertainties in the acquisition process. For both synthetic and real input data we estimate a polarisation image using the technique in Sec. 3.2 and provide this as input for each method. Note that for all experiments there are two polarimetric input images with different light source directions but, for the sake of space, we only show one of the two.

Dataset	Ours		Smith et al. [8]		SIRFS [49]	
	Light est.	AE (°)	Light est.	AE (°)	Light est.	AE (°)
Snooker 1	−0.42 −0.07 0.91	**7.1**	−0.40 −0.03 0.92	7.4	−0.84 −0.23 0.48	37.2
Snooker 2	−0.05 −0.44 0.90	**5.4**	−0.06 −0.37 0.93	9.6	−0.67 −0.03 0.74	48.1
3DRFE 1	−0.41 −0.13 0.90	**9.7**	−0.48 −0.19 0.86	11.0	−0.87 −0.08 0.49	35.3
3DRFE 2	−0.04 −0.50 0.87	**4.8**	−0.02 −0.57 0.82	9.2	−0.63 −0.71 0.30	54.3

TABLE 2: Light estimation on the real snooker ball and one of the 3DRFE datasets (row 2 in Fig. 4). Ground truth directions are $[-0.51,0,0.86]^T$ and $[0,-0.51,0.86]^T$, respectively, and we show angular error in degrees.

Light direction estimation We begin by evaluating estimation of the light source direction. We capture a snooker ball and render a synthetic 3DRFE dataset under light source directions $\mathbf{s} = [-0.51,0,0.86]^T$ and $\mathbf{t} = [0,-0.51,0.86]^T$. We show the input images for the snooker ball under these two light source directions in the first two columns of Fig. 3 along with the estimated diffuse mask in the third column. Note that pixels that contain specularities in either image

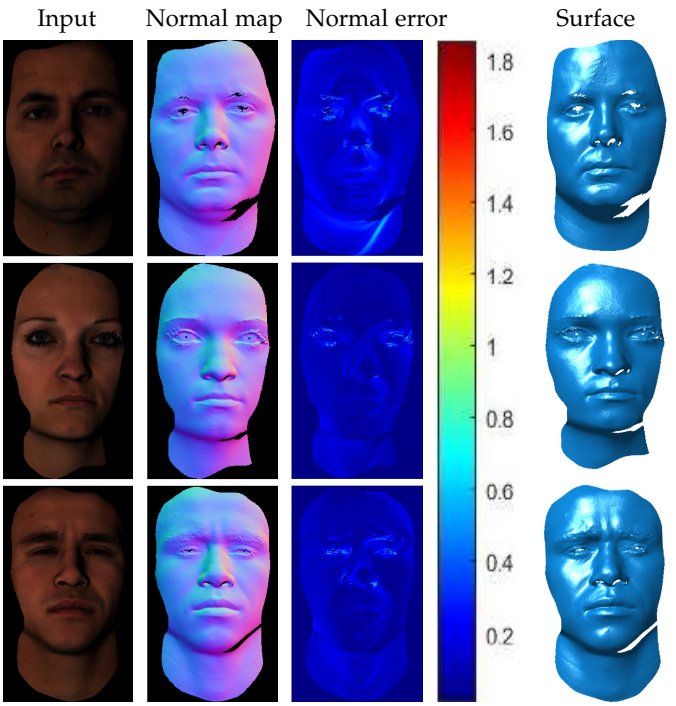

Fig. 4: Shape estimation evaluation on synthetic data from the 3DRFE dataset [48]. Left to right: one of the input lighting conditions, estimated normal map, normal map errors and shaded surface.

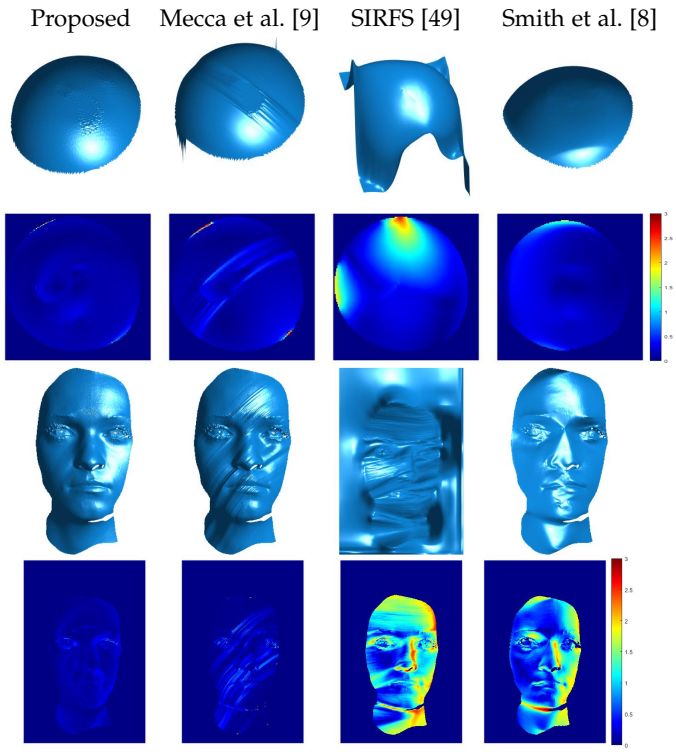

Fig. 5: Surface reconstruction and surface normal error obtained by our proposed method, Mecca et al. [9], SIRFS [49] and Smith et al. [8]. The corresponding light source estimations are reported in Tab. 2.

Method	Snooker ball				3DRFE			
	Albedo		Normals		Albedo		Normals	
	MAbsE	STD	MAngE	STD	MAbsE	STD	MAngE	STD
Ours	0.114	0.097	**0.094**	0.057	**0.037**	0.045	**0.222**	0.151
Ours (est. light)	**0.058**	0.057	0.103	0.088	0.114	0.212	0.678	0.582
Mecca et al. [9]	N/A	N/A	0.133	0.143	N/A	N/A	0.253	0.309
Mecca et al. [9] (est. light)	N/A	N/A	0.151	0.156	N/A	N/A	0.756	0.585
SIRFS	0.119	0.090	0.500	0.425	0.207	0.198	0.846	0.467
Smith et al. [8]	0.258	0.199	0.209	0.183	0.108	0.127	0.850	0.501
Smith et al. [8] (est. light)	0.230	0.178	0.182	0.133	0.115	0.138	0.860	0.497

TABLE 3: Quantitative shape and albedo estimation results. We show mean absolute error (MAbsE) for albedo and mean angular error (MAngE) in degrees for surface normals, related to the results reported in Fig. 5 and Fig. 7.

are correctly excluded by the mask. Estimated light source directions and errors for our method and two comparison methods are shown in Tab. 2. Smith et al. [8] is the only other method we are aware of for light estimation from polarisation images while SIRFS [49] is a baseline method that does not use polarisation information. Our approach provides the lowest error across all images. Averaged over the two light source directions and all 3DRFE images the angular errors are $7.7°$ (ours), $9.3°$ for [8] and $56.2°$ for [49].

Shape estimation For shape estimation we use the alternating variant of our method described in Section 5.3, initialised by the invariant method. We show shape estimation results for which we have ground truth in columns 5 (normal map) and 6 (normal map errors) of Fig. 3 for the snooker ball and in Fig. 4 for synthetic images from the 3DRFE dataset. We provide a comparison against state of the art shape-from-polarisation methods of Smith et al. [8] and Mecca et al. [9] as well as image-based baseline method SIRFS [49] in Fig. 5. For our method, [8] and [9] we compute normal maps by differentiating the estimated height maps. In all figures, error colour bars show angular errors in radians. We show quantitative results in Tab. 3. We compute the mean angular error in degrees, averaged over pixels, and the standard deviation for both datasets where we have ground truth. For methods requiring a light source estimate we provide results with both ground truth and with our estimated illumination (labelled "est. light"). Our method achieves the lowest error and also the smallest variance whilst also providing surfaces that are globally coherent, devoid of spike artefacts and retain fine detail. Using estimated illumination slightly degrades performance (except for the snooker ball where "ground truth" is only approximate), but still improves significantly on SIRFS which also estimates illumination. Finally, in Fig. 6 we show qualitative shape estimation results for more challenging real objects with albedo variations. Column 2 shows the estimated normal maps and column 4 a rotated view of the height maps.

Albedo estimation In Fig. 7 we show estimated albedo maps for which ground truth is available for our method, Smith et al. [8] and SIRFS [49]. The corresponding quantitative evaluation is in Tab. 3 where we achieve the lowest mean absolute error for both datasets. Using the estimated albedo and normal maps, in column 3 of Fig. 6 we show reillumination results under novel lighting with Blinn-Phong reflectance [50]. In Fig. 8 we show albedo estimation results on more challenging real objects for our method and the comparison methods. In all cases, our method disentangles shading from albedo while avoiding introducing artefacts or transferring albedo texture into the shape estimation.

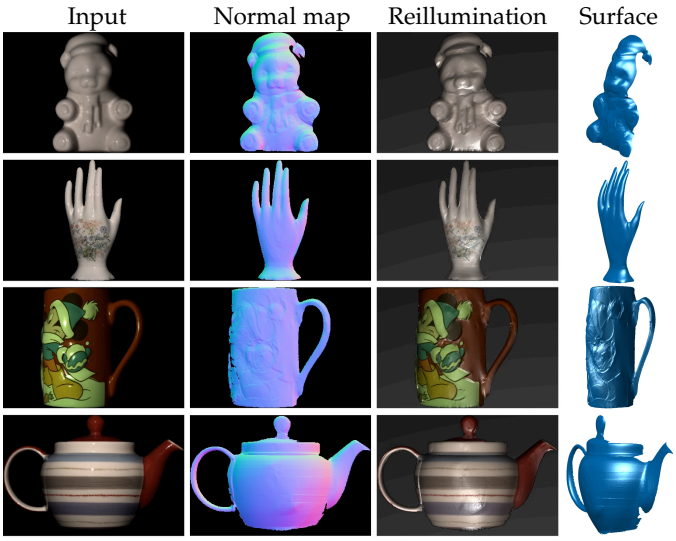

Fig. 6: Qualitative shape estimation and reillumination results on real objects with varying albedo.

	Snooker	Bear	Hand	Cup	Teapot	3DRFE
Ours	1.591	1.543	1.423	1.522	1.602	1.583
Ground Truth	1.63	1.504				1.5

TABLE 4: Refractive index estimates for various real objects and synthetic renderings of 3DRFE face data.

Refractive index estimation Finally, we evaluate performance on refractive index estimation. The synthetic 3DRFE data are rendered with a chosen refractive index of $\eta = 1.5$. The real objects are either made of porcelain (ground truth refractive index reported in [51]) or phenol formaldehyde resin (ground truth refractive index reported in [52]). We show our estimates and the ground truth values in Tab. 4. In Fig. 9 we show refractive index estimates as a function of the actual refractive index and for varying degrees of additive Gaussian noise. In all cases we initialised the nonlinear optimisation with $\eta = 1.5$. Note that there is a good agreement between estimated and actual but that as noise increases there is a bias towards overestimates. This is because noise causes the estimated surface to flatten, leading to underestimates of the zenith angle which causes overestimation of the refractive index when solving (35).

9 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 and refractive index invariant shape estimation with only two light sources. The main conclusion from this work is that the highly challenging case of two source photometric stereo becomes practical, even in an uncalibrated setting, when polarimetric images are available. We envisage our approach will find application either in industrial inspection, where the relatively modest data capture requirements and ability to estimate material properties (albedo and refractive index) will be of value, or for 3D scanning where the placement of light sources need not be calibrated (e.g. two arbitrarily placed flashes in a studio setup).

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 or multiview stereo. To be applicable outside the lab, we would also like to develop methods that work with more general lighting conditions.

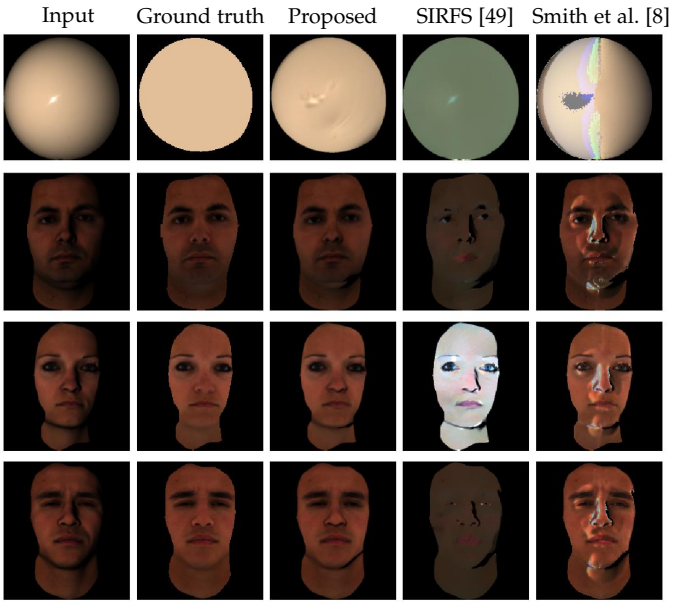

Fig. 7: Albedo estimation versus ground truth. Left to right: input, ground truth, our result and the results of SIRFS [49] and Smith et al. [8]. Quantitative results in Tab. 3.

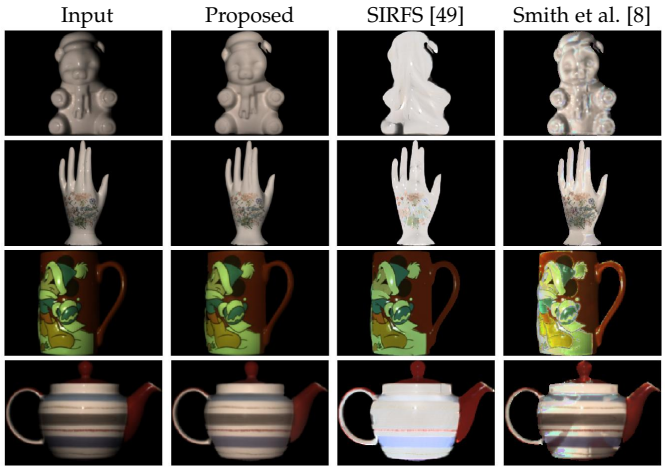

Fig. 8: Qualitative albedo estimation on real data. Left to right: input, our result, SIRFS [49] and Smith et al. [8].

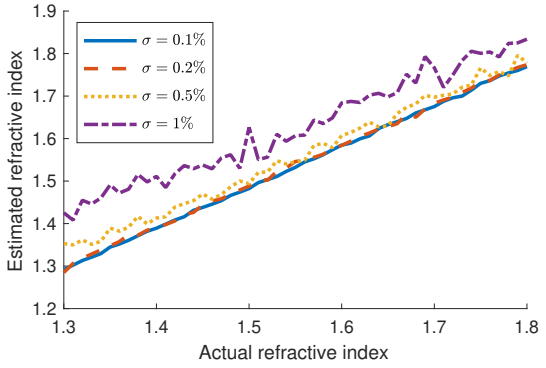

Fig. 9: Actual versus estimated refractive index for synthetic data with Gaussian noise of standard deviation σ.

This includes moving beyond the point source assumption but also considering the case of polarised or partially polarised illumination. Finally, we have ignored interreflections which would violate some of our assumptions and complicate modelling of polarised reflectance. This phenomenon could be significant for concave objects and is another avenue for future work.

Acknowledgment

This work was supported mainly by the "GNCS - INdAM", in part by ONR grants N000141512013, N000142012529 and the UC San Diego Center for Visual Computing. W. Smith was supported by EPSRC grant EP/N028481/1 and a Royal Academy of Engineering/The Leverhulme Trust Senior Research Fellowship. S. Tozza was supported by PON R&I 2014-2020 – "AIM: Attraction and International Mobility" (Linea 2.1, project AIM1834118 - 2, CUP: E61G19000050001).

References

[1] A. Ecker and A. D. Jepson, "Polynomial shape from shading," in *Proc. CVPR*. IEEE, 2010, pp. 145–152.
[2] 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.
[3] R. Mecca and M. Falcone, "Uniqueness and approximation of a photometric shape-from-shading model," *SIAM Journal on Imaging Sciences*, vol. 6, no. 1, pp. 616–659, 2013.
[4] 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*, vol. 56, no. 1, pp. 57–76, 2016.
[5] W. A. P. Smith and F. Fang, "Height from photometric ratio with model-based light source selection," *Computer Vision and Image Understanding*, vol. 145, pp. 128 – 138, 2016.
[6] 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)*, June 2016, pp. 4359–4368.
[7] W. A. P. Smith, R. Ramamoorthi, and S. Tozza, "Linear depth estimation from an uncalibrated, monocular polarisation image," in *Computer Vision - ECCV 2016*, ser. Lecture Notes in Computer Science, vol. 9912, 2016, pp. 109–125.
[8] ——, "Height-from-polarisation with unknown lighting or albedo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 12, pp. 2875–2888, 2019.
[9] R. Mecca, F. Logothetis, and R. Cipolla, "A differential approach to shape from polarization," in *Proc. BMVC*, 2017.
[10] F. Logothetis, R. Mecca, F. Sgallari, and R. Cipolla, "A differential approach to shape from polarisation: A level-set characterisation," *International Journal of Computer Vision*, vol. 127, no. 11-12, pp. 1680–1693, 2019.
[11] A. Kadambi, V. Taamazyan, B. Shi, and R. Raskar, "Polarized 3D: High-quality depth sensing with polarization cues," in *Proc. ICCV*, 2015.
[12] 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*, vol. 57, pp. 175–191, 2017.
[13] 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, 1991.
[14] S. Nayar, X. Fang, , and T. Boult, "Separation of reflection components using color and polarization," *Int. J. Comput. Vision*, vol. 21, no. 3, pp. 163–186, 1997.
[15] Y. Lyu, Z. Cui, S. Li, M. Pollefeys, and B. Shi, "Reflection separation using a pair of unpolarized and polarized images," in *Advances in Neural Information Processing Systems*, 2019, pp. 14 559–14 569.
[16] P. Wieschollek, O. Gallo, J. Gu, and J. Kautz, "Separating reflection and transmission images in the wild," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 89–104.
[17] V. Taamazyan, A. Kadambi, and R. Raskar, "Shape from mixed polarization," *arXiv preprint arXiv:1605.02066*, 2016.
[18] T. T. Ngo, H. Nagahara, and R. Taniguchi, "Shape and light directions from shading and polarization," in *Proc. CVPR*, 2015, pp. 2310–2318.
[19] S. Tozza, W. A. Smith, D. Zhu, R. Ramamoorthi, and E. R. Hancock, "Linear differential constraints for photo-polarimetric height estimation," in *Proc. ICCV*, 2017.
[20] G. A. Atkinson and E. R. Hancock, "Recovery of surface orientation from diffuse polarization," *IEEE Transactions on Image processing*, vol. 15, no. 6, pp. 1653–1664, 2006.
[21] D. Miyazaki, R. T. Tan, K. Hara, and K. Ikeuchi, "Polarization-based inverse rendering from a single view," in *Proc. ICCV*, 2003, pp. 982–987.
[22] C. P. Huynh, A. Robles-Kelly, and E. Hancock, "Shape and refractive index recovery from single-view polarisation images," in *Proc. CVPR*, 2010, pp. 1229–1236.
[23] O. Morel, F. Meriaudeau, C. Stolz, and P. Gorria, "Polarization imaging applied to 3D reconstruction of specular metallic surfaces," in *Proc. EI 2005*, 2005, pp. 178–186.
[24] A. H. Mahmoud, M. T. El-Melegy, and A. A. Farag, "Direct method for shape recovery from polarization and shading," in *Proc. ICIP*, 2012, pp. 1769–1772.
[25] G. A. Atkinson and E. R. Hancock, "Surface reconstruction using polarization and photometric stereo," in *Proc. CAIP*, 2007, pp. 466–473.
[26] Y. Yu, D. Zhu, and W. A. P. Smith, "Shape-from-polarisation: a nonlinear least squares approach," in *Proc. ICCV Workshop on Color and Photometry in Computer Vision*, 2017.
[27] G. A. Atkinson, "Polarisation photometric stereo," *Computer Vision and Image Understanding*, vol. 160, pp. 158–167, 2017.
[28] 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 *3DIMPVT*, 2012, pp. 104–111.
[29] ——, "Surface normal estimation of black specular objects from multiview polarization images," *Optical Engineering*, vol. 56, no. 4, pp. 041 303–041 303, 2017.
[30] C. P. Huynh, A. Robles-Kelly, and E. R. Hancock, "Shape and refractive index from single-view spectro-polarimetric images," *Int. J. Comput. Vision*, vol. 101, no. 1, pp. 64–94, 2013.
[31] O. Drbohlav and R. Šára, "Unambiguous determination of shape from photometric stereo with unknown light sources," in *Proc. ICCV*, 2001, pp. 581–586.
[32] P. N. Belhumeur, D. J. Kriegman, and A. Yuille, "The Bas-relief ambiguity," *Int. J. Comput. Vision*, vol. 35, no. 1, pp. 33–44, 1999.
[33] A. Kadambi, V. Taamazyan, B. Shi, and R. Raskar, "Depth sensing using geometrically constrained polarization normals," *Int. J. Comput. Vision*, 2017.
[34] D. Zhu and W. A. P. Smith, "Depth from a polarisation + RGB stereo pair," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 7586–7595.
[35] L. B. Wolff, "Surface orientation from two camera stereo with polarizers," in *Proc. SPIE Conf. Optics, Illumination, and Image Sensing for Machine Vision IV*, vol. 1194, 1990, pp. 287–298.

[36] S. Rahmann and N. Canterakis, "Reconstruction of specular surfaces using polarization imaging," in *Proc. CVPR*, 2001.
[37] 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, 2007.
[38] 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, 2004.
[39] K. Berger, R. Voorhies, and L. H. Matthies, "Depth from stereo polarization in specular scenes for urban robotics," in *Proc. ICRA*, 2017, pp. 1966–1973.
[40] Z. Cui, J. Gu, B. Shi, P. Tan, and J. Kautz, "Polarimetric multi-view stereo," in *Proc. CVPR*, 2017, pp. 1558–1567.
[41] L. Chen, Y. Zheng, A. Subpa-asa, and I. Sato, "Polarimetric three-view geometry," in *Proc. ECCV*, 2018, pp. 20–36.
[42] L. Yang, F. Tan, A. Li, Z. Cui, Y. Furukawa, and P. Tan, "Polarimetric dense monocular slam," in *Proc. CVPR*, 2018, pp. 3857–3866.
[43] Z. Cui, V. Larsson, and M. Pollefeys, "Polarimetric relative pose estimation," in *Proc. ICCV*, 2019, pp. 2671–2680.
[44] L. B. Wolff, "Polarization vision: a new sensory approach to image understanding," *Image Vision Comput.*, vol. 15, no. 2, pp. 81–93, 1997.
[45] S. G. Narasimhan, V. Ramesh, and S. Nayar, "A class of photometric invariants: Separating material from shape and illumination," in *Proc. ICCV*, vol. 2, 2003, pp. 1387 – 1394.
[46] D. Zhu and W. A. P. Smith, "Least squares surface reconstruction on arbitrary domains," in *Proc. ECCV*, 2020.
[47] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.
[48] G. Stratou, A. Ghosh, P. Debevec, and L.-P. Morency, "Effect of illumination on automatic expression recognition: a novel 3D relightable facial database," in *Proc. Face & Gesture*. IEEE, 2011, pp. 611–618.
[49] J. T. Barron and J. Malik, "Shape, illumination, and reflectance from shading," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 8, pp. 1670–1687, 2014.
[50] J. F. Blinn, "Models of light reflection for computer synthesized pictures," *Computer Graphics*, vol. 11, no. 2, pp. 192–198, 1977.
[51] *Material index of refraction*, 2004 (accessed November 25, 2020). [Online]. Available: http://www.spacekdet.com/tutorials/IOR.html
[52] F. G. Tickell, *The techniques of sedimentary mineralogy*. Elsevier, 2011.

Silvia Tozza received the B.Sc. degree in Mathematics from Università degli Studi di Salerno, Salerno, Italy, in 2008 and the M.Sc. degree in Applied Mathematics from Sapienza Università di Roma, Rome, Italy, in 2011. She received her Ph.D. degree in Applied Mathematics from Sapienza Università di Roma in January 2015 with final judgment Excellent. After years of post-doc at various institutions (the Computer Science Department of the University of York, York, UK, then at the Department of Mathematics, Sapienza Università di Roma, and at the Istituto Nazionale di Alta Matematica (INdAM)), currently, she is a Researcher at the Department of Mathematics and Applications "Renato Caccioppoli", Universitá degli Studi di Napoli Federico II, Naples, Italy. Her research focuses mainly on analysis and approximation of nonlinear PDEs of Hamilton-Jacobi type, with applications in Image Processing, especially 3D reconstruction, image segmentation and denoising.

Dizhong Zhu received the B.Sc. degree in Telecommunication from Shenzhen University, Shenzhen, China and his M.Sc. degree in computer science from University of Leicester, Leicester, U.K. He is currently a Ph.D student in the department of Computer Science, University of York, York, U.K. His research interests focus on 3D reconstruction, especially in shape from polarisation, and combine with shading, photometric stereo and multi-view stereo.

William Smith 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 Reader with the Department of Computer Science, University of York, York, United Kingdom. He holds a Royal Academy of Engineering/The Leverhulme Trust Senior Research Fellowship. 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.

Ravi Ramamoorthi received his BS degree in engineering and applied science and MS degrees in computer science and physics from the California Institute of Technology in 1998. He received his PhD degree in computer science from the 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 at the 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 100 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. He is a fellow of the IEEE, ACM, and the SIGGRAPH Academy.

Edwin R. Hancock received the B.Sc., Ph.D., and D.Sc. degrees from the University of Durham, Durham, UK. He is currently an Emeritus Professor with the Department of Computer Science, University of York, York, UK. He has published over 200 journal articles and 650 conference papers. Prof. Hancock was a recipient of the Royal Society Wolfson Research Merit Award in 2009, the Pattern Recognition Society Medal in 1991, and the BMVA Distinguished Fellowship in 2016 and the IAPR Pierre Devijver Award in 2018. He is a fellow of the IAPR, IEEE, the Institute of Physics, the Institute of Engineering and Technology, and the British Computer Society. He is currently Editor-in-Chief of the journal Pattern Recognition, and was founding Editor-in-Chief of IET Computer Vision from 2006 until 2012. He has also been a member of the editorial boards of the journals IEEE Transactions on Pattern Analysis and Machine Intelligence, Pattern Recognition, Computer Vision and Image Understanding, Image and Vision Computing, and the International Journal of Complex Networks. He has been Conference Chair for BMVC in 1994 and Programme Chair in 2016, Track Chair for ICPR in 2004 and 2016 and an Area Chair at ECCV 2006 and CVPR in 2008 and 2014, and in 1997 established the EMMCVPR workshop series. He was Second Vice President of the International Association for Pattern Recognition (2016-2018).