

## Article

# DRM-Based Colour Photometric Stereo Using Diffuse-Specular Separation for Non-Lambertian Surfaces

Boren Li <sup>1,†</sup> and Tomonari Furukawa <sup>2,\*,†,‡</sup><sup>1</sup> Beijing Institute for General Artificial Intelligence, Beijing 100124, China; liboren@vip.163.com<sup>2</sup> School of Engineering and Applied Science, University of Virginia, Charlottesville, VA 22904, USA

\* Correspondence: tomonari@virginia.edu; Tel.: +1-434-297-6827

† These authors contributed equally to this work.

‡ Current address: 675 Old Reservoir Rd, Charlottesville, VA 22903, USA.

**Abstract:** This paper presents a photometric stereo (PS) method based on the dichromatic reflectance model (DRM) using colour images. The proposed method estimates surface orientations for surfaces with non-Lambertian reflectance using diffuse-specular separation and contains two steps. The first step, referred to as diffuse-specular separation, initialises surface orientations in a specular invariant colour subspace and further separates the diffuse and specular components in the RGB space. In the second step, the surface orientations are refined by first initialising specular parameters via solving a log-linear regression problem owing to the separation and then fitting the DRM using Levenburg-Marquardt algorithm. Since reliable information from diffuse reflection free from specularities is adopted in the initialisations, the proposed method is robust and feasible with less observations. At pixels where dense non-Lambertian reflectances appear, signals from specularities are exploited to refine the surface orientations and the additionally acquired specular parameters are potentially valuable for more applications, such as digital relighting. The effectiveness of the newly proposed surface normal refinement step was evaluated and the accuracy in estimating surface orientations was enhanced around 30% on average by including this step. The proposed method was also proven effective in an experiment using synthetic input images comprised of twenty-four different reflectances of dielectric materials. A comparison with nine other PS methods on five representative datasets further prove the validity of the proposed method.

**Keywords:** photometric stereo; dichromatic reflectance model; diffuse-specular separation; non-lambertian surfaces



**Citation:** Li, B.; Furukawa, T. DRM-Based Colour Photometric Stereo Using Diffuse-Specular Separation for Non-Lambertian Surfaces. *J. Imaging* **2022**, *8*, 40. <https://doi.org/10.3390/jimaging8020040>

Academic Editor: Daisuke Miyazaki

Received: 29 November 2021

Accepted: 5 January 2022

Published: 8 February 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Photometric stereo (PS) estimates surface orientations using images captured from a fixed viewpoint under various illuminations and is especially powerful in acquisition of fine surface details at pixel level [1]. Surface orientation is important in a variety of fields, such as geometry segmentation for three-dimensional (3D) object recognition [2] and digital re-rendering in computer graphics. Surface geometries, which can be obtained via integrating surface orientations, have also been proven useful for applications, such as industrial quality control and reverse engineering. Due to the strength of PS and the importance of surface orientation acquisition, PS has drawn increasing interests since its debut [3]. However, making PS for a general real scene remains challenging due to the diverse reflectance properties of different materials that appear non-Lambertian [4]. This has given rise to the need for reliable estimation of surface orientations for a wide range of non-Lambertian reflectance, which essentially requires a proper imaging photometry model characterizing the forward problem and a subsequent PS method that inversely derives surface orientations.

Existing PS methods dealing with non-Lambertian reflectance can be classified into three categories. The first approximates surface reflectance using analytical bidirectional

reflectance distribution function (BRDF) [5] and formulates the estimation of surface orientations as a nonlinear fitting problem. Nayar et al. [6] derived surface orientations using the Torrance-Sparrow BRDF [7] and incorporated extended sources to ensure sufficient information from specularities. Georghiades [8] inverted the same BRDF to simultaneously estimate surface orientations and resolve the generalized bas-relief ambiguity. Goldman et al. [9] assumed that general material reflectance can be represented by a convex combination of fundamental materials characterized by the Ward BRDF [10] and recovered surface orientations, fundamental material BRDFs and weight maps simultaneously for further scene editing purpose. Methods in this category exploiting information from surface reflectance are capable to derive not only surface orientations but also the other parameters in the analytical BRDFs, allowing more functionalities, such as digital relighting and material classification. However, due to the nonlinearity of analytical BRDFs and larger number of parameters to be estimated, these methods are sensitive to initialisations, numerically unstable under heavily corrupted outliers (e.g., shadows), and inapplicable when the number of observations is limited.

The second category of methods infers surface orientations through adopting the general properties of BRDF, such as isotropy, monotonicity and reciprocity. Alldrin et al. [11] developed a non-parametric PS method using bi-variate approximation of the isotropy property. Higo et al. [12] analyzed the general BRDF constraints and employed the BRDF properties of monotonicity, isotropy and visibility to vote for the most possible surface orientations for single-lobed reflectance. Shi et al. [13] proposed a bi-polynomial representation for low-frequency reflectance that was especially adaptable to the inverse problem as PS, while Ikehata et al. [14] developed another general isotropic BRDF as sum of lobes with unknown center directions. Methods in this category capitalize on the most fundamental properties of BRDF and therefore have the potential to deal with a broader range of reflectance. However, these methods are only capable to derive surface orientations with limited other functionalities and require an even larger set of observations compared with the first category. To widen the applicability, recent years have seen non-parametric BRDFs based on machine learning. Santo et al. [15,16] used a deep neural network for the first time whereas Tani and Maehara [17] estimated surface normals and BRDFs by unsupervised learning. Ikehata [18] estimated surface normals more straightforward by deriving the so-called observation maps and using convolutional neural networks. Such data-driven approaches clearly improve the accuracy as observations directly create the BRDF or estimate surface normals. The use of a significantly large set of observations, which is key to the accuracy, and the expensive training effort using the observations make the approaches outside the scope of this paper.

Methods in the third category assume that non-Lambertian effects appear sparsely among observations and treat them as outliers. A substantial corpus of methods rely on robust statistical techniques for outlier rejection. Wu et al. [19] formulated the PS with outlier rejection as a global rank minimization problem, while Ikehata et al. [20] employed sparse Bayesian regression instead. Barsky et al. [21] initiated another line of PS researches using colour images where they first identified the significance of using specular colour in the dichromatic reflectance model (DRM) [22] for specular rejection. Using the cue from known specular colour under the same theoretical foundation, Zickler et al. [23] derived a PS method in a novel two-dimensional (2D) specular invariant colour subspace. The major advantages of methods in the third category are their robustness and requirement for less images, while they are inefficient in the presence of dense non-Lambertian effects.

This paper presents a colour PS method to estimate surface orientations using diffuse-specular separation dealing with non-Lambertian reflectance. The proposed method models the colour imaging photometry using DRM as the forward problem and inverts it following a two-step procedure, the diffuse-specular separation and the surface normal refinement. The first step, initialising surface orientations using known specular colour similarly to [23], separates the diffuse and specular components in the RGB space and identifies outliers to reject in the UV space. In the second step, the parameters characterizing the specular

are initialised via solving a log-linear regression problem and the surface orientations are finally refined by fitting the nonlinear DRM using Levenburg-Marquardt algorithm (LMA). The proposed method robustly initialises surface orientations in a specular-free colour subspace and further fits for the nonlinear DRM with a newly derived parameter initialisation strategy. The proposed method preserves the advantage of robustness as the third category of methods, while it also benefits from the separated specular component to tackle with dense non-Lambertian reflectance as the first category. Furthermore, DRM parameters besides the surface orientations can be additionally obtained where dense non-Lambertian reflectance appears, making more potential applications feasible, such as digital relighting and material classification.

The rest of the paper is organized as follows. The next section presents the problem formulation of colour PS incorporating DRM and the relevant PS methods. Section 3 first overviews the flow and original contribution of the proposed colour DRM-based PS method and then, elucidates the specific two steps in the subsequent subsections. It follows the comprehensive evaluations on surface orientation estimation using both synthetic and real images in Section 4, while the last section summarizes the conclusions and proposes the future works.

## 2. Colour PS Incorporating Dichromatic Reflectance Model

### 2.1. Generic Colour PS Problem Formulation

Figure 1 shows the schematic diagram of the hardware components of the sensor and the necessary data for the generic colour PS. The hardware components are an RGB camera and a set of directional point sources whose positions are priorly known. Assume that ambient light is completely blocked, the colour PS requires that the RGB camera captures a colour image for the target surface with one source lit at a time. Given the  $N$  colour images illuminated under  $N$  different sources, the objective of the colour PS is to identify surface orientations pixel-wise. More specifically, given the colour image irradiance measurements at pixel  $(i, j)$ ,  $\{e_k^{i,j*}\}_{k \in [1, N]}$ , under the unit illumination directions,  $\{\bar{\mathbf{l}}_k^{i,j}\}_{k \in [1, N]}$ , estimated from the light positions, the colour PS aims to derive the unit surface normal,  $\bar{\mathbf{n}}^{i,j}$ , by minimizing the sum of  $N$  corresponding square residuals:

$$e^{i,j} = \sum_{k=1}^N \left[ \left( e_k^{i,j*} \right)^T \mathbf{w} - f \left( \bar{\mathbf{n}}^{i,j}, \bar{\mathbf{l}}_k^{i,j} \right) \right]^2 \rightarrow \min_{\bar{\mathbf{n}}^{i,j}}. \tag{1}$$

where  $\mathbf{w}$  is a vector of weights for colour channels whereas  $f$  is a forward function of  $\bar{\mathbf{n}}^{i,j}$  having  $\bar{\mathbf{l}}_k^{i,j}$  as parameters, which corresponds to the weighted measurements. Note that  $(\bar{\cdot})$  and  $(\cdot)^*$  mean the unit version and the measurement of  $(\cdot)$  respectively throughout the paper.

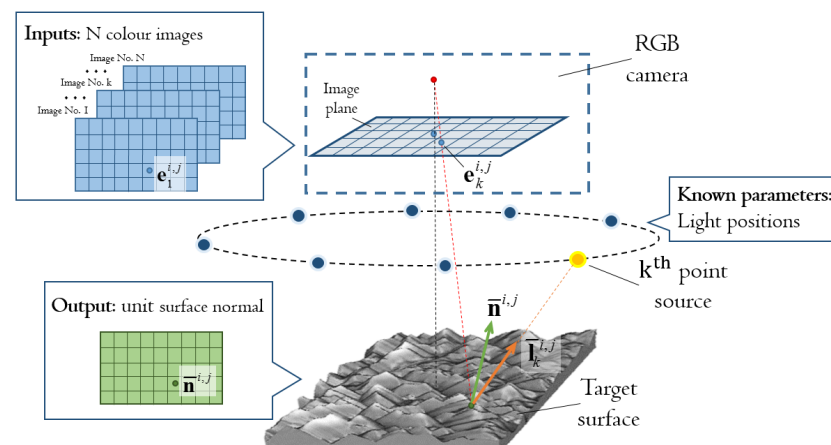


Figure 1. Schematic diagram for generic colour PS.

### 2.2. Colour Image Formation Model

Figure 2 shows the coordinate setups that are necessary for the colour PS to model the imaging geometry and the light configuration. The coordinate frames are the 3D camera frame,  $\{C\}$ , the 2D pixel frame,  $\{P\}$ , and the world frame,  $\{W\}$ , where its  $XY$  plane is known and referred to as the reference plane. These coordinate frames follow the convention given in [24]. A 3D surface point,  $\{C\}\tilde{\mathbf{X}}$ , is projected to be a 2D image point,  $\{P\}\mathbf{x}$ , in the RGB camera using the perspective projection as:

$$\{P\}\mathbf{x} = \frac{1}{\{C\}Z} \mathbf{M}_p \mathbf{K} \{C\}\tilde{\mathbf{X}} = \frac{1}{\{C\}Z} \begin{bmatrix} m_x & 0 & o_x \\ 0 & m_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \{C\}\tilde{\mathbf{X}}, \quad (2)$$

where  $m_x$  and  $m_y$  are each the number of pixels per unit distance along  $x$  and  $y$  direction respectively.  $o_x$  and  $o_y$  are the principal point location in  $\{P\}$ , and  $f$  is the effective focal length. Note that  $\{P\}\mathbf{x} = [\{P\}x, \{P\}y, 1]^\top$  and  $\{C\}\tilde{\mathbf{X}} = [\{C\}X, \{C\}Y, \{C\}Z]^\top$ .  $\{P\}\mathbf{x}$  is rasterized into a regular grid in the camera as:

$$\begin{bmatrix} j \\ i \end{bmatrix} = \text{ceil} \left( \begin{bmatrix} \{P\}x + 0.5 \\ \{P\}y + 0.5 \end{bmatrix} \right), \quad (3)$$

where  $\text{ceil}(\cdot)$  denotes the ceiling operator.  $\{W\}$  and  $\{C\}$  are related with a 3D rigid transformation:  $\{C\}\tilde{\mathbf{X}} = \mathbf{R}_W \{W\}\tilde{\mathbf{X}} + \mathbf{t}_W$ .

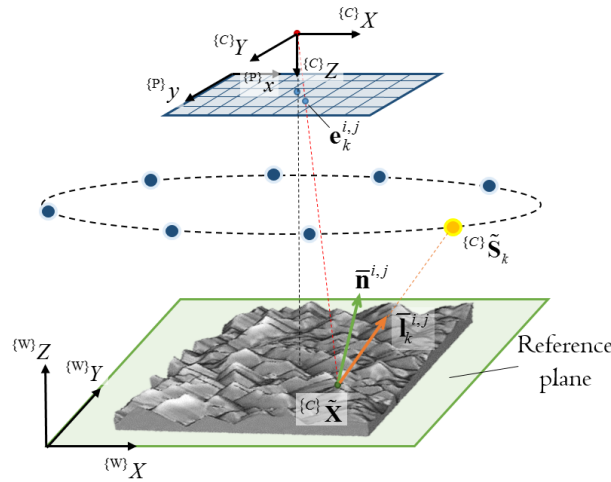


Figure 2. Coordinate setups for modeling imaging geometry and light configuration.

Let the  $k$ th light position be  $\{C\}\tilde{\mathbf{S}}_k = [\{C\}X_k, \{C\}Y_k, \{C\}Z_k]^\top$ .  $\{C\}\tilde{\mathbf{S}}_k$  is associated in the spherical space as:

$$\begin{cases} \{C\}X_k = \{C\}r_k \sin \{C\}\theta_k \cos \{C\}\phi_k \\ \{C\}Y_k = \{C\}r_k \sin \{C\}\theta_k \sin \{C\}\phi_k \\ \{C\}Z_k = \{C\}r_k \cos \{C\}\theta_k, \end{cases} \quad (4)$$

where  $\{C\}r_k$ ,  $\{C\}\theta_k$  and  $\{C\}\phi_k$  are the radius, zenith and azimuth angle, respectively. The  $N$  lights are configured such that they have the same zenith angle and radius, and uniformly distributed azimuth angles around the optical axis, i.e.,  $\{C\}\theta_k = \theta$ ,  $\{C\}r_k = r$  and  $\{C\}\phi_k = \frac{360^\circ(k-1)}{N} + \phi_0$ .

Having established the imaging geometry and light configuration, the imaging photometry is then modeled using the physically-based DRM [22]. The DRM properly characterizes reflectance for dielectric materials [25] and is represented by a linear combination of the diffuse and specular components. Assume that pixel  $(i, j)$  is shadow-free, the image

irradiance for colour channel  $c$  after light strength normalisation and vignetting correction is written as:

$$e_c^{i,j} = e_{d,c}^{i,j} + e_{s,c}^{i,j} = d_c^{i,j} f_d^{i,j} + s_c f_s^{i,j}, \tag{5}$$

where  $e_{d,c}^{i,j}$  and  $e_{s,c}^{i,j}$  are the diffuse and specular components, and  $d_c^{i,j}$  and  $s_c$  are diffuse and specular colours which are wavelength-dependent and represented by:

$$\begin{cases} d_c^{i,j} = \int_0^\infty \Phi(\lambda) R^{i,j}(\lambda) C_c(\lambda) d\lambda, \\ s_c = \int_0^\infty \Phi(\lambda) C_c(\lambda) d\lambda. \end{cases} \tag{6}$$

In the equation,  $\lambda$  is the wavelength,  $\Phi(\lambda)$  is the source spectral power density (SPD),  $R^{i,j}(\lambda)$  is the spectral body reflectance, and  $C_c(\lambda)$  is the camera spectral sensitivity for the colour channel.  $f_d^{i,j}$  and  $f_s^{i,j}$  represent the diffuse and specular geometrical scaling factors that are characterized by BRDF.

Since the scaling factors are illumination-dependent, let the image irradiance Equation (5) generalized in the RGB colour space be with the  $k$ th illumination as:

$$\mathbf{e}_k^{i,j} = \mathbf{e}_{d,k}^{i,j} + \mathbf{e}_{s,k}^{i,j} = \bar{\mathbf{d}}_{RGB}^{i,j} f_{d,k}^{i,j} + \bar{\mathbf{s}}_{RGB} f_{s,k}^{i,j} = \begin{bmatrix} \bar{d}_R^{i,j} \\ \bar{d}_G^{i,j} \\ \bar{d}_B^{i,j} \end{bmatrix} f_{d,k}^{i,j} + \begin{bmatrix} \bar{s}_R \\ \bar{s}_G \\ \bar{s}_B \end{bmatrix} f_{s,k}^{i,j} \tag{7}$$

where  $\bar{\mathbf{d}}_{RGB}^{i,j} = \{d_c^{i,j}\}_{c=R,G,B}$  and  $\bar{\mathbf{s}}_{RGB} = \{s_c\}_{c=R,G,B}$  are the unit diffuse and specular colour, respectively. The diffuse geometrical scaling factor is characterized with the  $k$ th unit illumination direction  $\bar{\mathbf{I}}_k^{i,j}$  by Lambertian model as:

$$f_{d,k}^{i,j} = k_d^{i,j} (\bar{\mathbf{n}}^{i,j})^\top \bar{\mathbf{I}}_k^{i,j}, \tag{8}$$

where  $k_d^{i,j}$  is the diffuse reflectance factor. Since the surface deviation is small compared with the light-surface distance,  $\bar{\mathbf{I}}_k^{i,j}$  can be approximately derived using a vector from the point intersected by ray  $(i, j)$  and the reference plane to  $\{C\} \tilde{\mathbf{S}}_k$ . One of the common options of the specular geometrical scaling factor, on the other hand, is a model using the Blinn-Phong BRDF [26] because of its computational efficiency and high performance in characterizing a wide range of isotropic reflectance [27]:

$$f_{s,k}^{i,j} = k_s^{i,j} \left( (\bar{\mathbf{n}}^{i,j})^\top \bar{\mathbf{h}}_k^{i,j} \right)^{\beta^{i,j}}, \tag{9}$$

where  $k_s^{i,j}$  is the specular reflectance factor,  $\bar{\mathbf{h}}_k^{i,j}$  is the unit half vector, and  $\beta^{i,j}$  is the shininess coefficient. The unit half vector is written as:

$$\bar{\mathbf{h}}_k^{i,j} = \frac{\bar{\mathbf{I}}_k^{i,j} + \bar{\mathbf{v}}^{i,j}}{\|\bar{\mathbf{I}}_k^{i,j} + \bar{\mathbf{v}}^{i,j}\|}, \tag{10}$$

where  $\bar{\mathbf{v}}^{i,j}$  is the unit viewer direction and determined by:

$$\begin{cases} \mathbf{v}^{i,j} = - \left[ \frac{j - o_x}{m_x}, \frac{i - o_y}{m_y}, f \right]^\top, \\ \bar{\mathbf{v}}^{i,j} = \mathbf{v}^{i,j} / \|\mathbf{v}^{i,j}\|. \end{cases} \tag{11}$$

### 2.3. Colour PS Methods Using DRM

Existing colour PS methods using DRM contain two major variants, representative works of which were developed by Barsky et al. [21] and by Zickler et al. [23]. The PS method of Barsky finds the unit surface normal  $\bar{\mathbf{n}}^{i,j}$  and the diffuse reflectance factor  $k_d^{i,j}$  such that the residual between the colour image irradiances weighted by  $\bar{\mathbf{d}}_{RGB}^{i,j}$  and the diffuse geometrical scaling factor is minimised:

$$\epsilon^{i,j} = \sum_{k=1}^{N_b} \left[ \left( \mathbf{e}_k^{i,j*} \right)^\top \bar{\mathbf{d}}_{RGB}^{i,j} - f_{d,k}^{i,j} \right]^2 = \sum_{k=1}^{N_b} \left[ \left( \mathbf{e}_k^{i,j*} \right)^\top \bar{\mathbf{d}}_{RGB}^{i,j} - k_d^{i,j} \left( \bar{\mathbf{n}}^{i,j} \right)^\top \bar{\mathbf{I}}_k^{i,j} \right]^2 \rightarrow \min_{\bar{\mathbf{n}}^{i,j}, k_d^{i,j}}, \quad (12)$$

where  $N_b (\leq N)$  is the number of specular-free and shadow-free images. The reason for the use of  $N_b$  instead of  $N$  is that the Barsky’s PS method assumes that non-Lambertian reflectance appears sparsely among observations, which should be rejected as outliers. As the first process for outlier rejection, the direct principal component analysis (DPCA) [28] on  $\left[ \mathbf{E}_{RGB}^{i,j*} \right]^\top \left[ \mathbf{E}_{RGB}^{i,j*} \right]$  is performed to estimate  $\bar{\mathbf{d}}_{RGB}^{i,j}$ , where  $\left[ \mathbf{E}_{RGB}^{i,j*} \right] = \left[ \mathbf{e}_1^{i,j*}, \mathbf{e}_2^{i,j*}, \dots, \mathbf{e}_N^{i,j*} \right]^\top$ . Given  $\bar{\mathbf{s}}_{RGB}$  and  $\bar{\mathbf{d}}_{RGB}^{i,j}$ , the method then determines  $f_{s,k}^{i,j}$  as:

$$f_{s,k}^{i,j} = \frac{\left( \mathbf{e}_k^{i,j*} \right)^\top \bar{\mathbf{s}}_{RGB} - \left( \left( \mathbf{e}_k^{i,j*} \right)^\top \bar{\mathbf{d}}_{RGB}^{i,j} \right) \left( \left( \bar{\mathbf{d}}_{RGB}^{i,j} \right)^\top \bar{\mathbf{s}}_{RGB} \right)}{1 - \left( \left( \bar{\mathbf{d}}_{RGB}^{i,j} \right)^\top \bar{\mathbf{s}}_{RGB} \right)^2}. \quad (13)$$

A threshold on  $f_{s,k}^{i,j}$  is finally used to determine pixels to be rejected as outliers, and this subsequently determines  $N_b$ .

The objective function (12) of this PS method is not different from that of the conventional Lambertian-based PS method [3]. Given  $\{\bar{\mathbf{I}}_k^{i,j}\}_{k \in [1, M]}$  a priori,  $\bar{\mathbf{n}}^{i,j}$  together with  $k_d^{i,j}$ , which minimizes the objective function (12), is analytically determined as:

$$\begin{cases} k_d^{i,j} = \left\| \left( \left[ \mathbf{L}^{i,j} \right]^\top \left[ \mathbf{L}^{i,j} \right] \right)^{-1} \left[ \mathbf{L}^{i,j} \right]^\top \left[ \mathbf{E}_{RGB}^{i,j*} \right] \bar{\mathbf{d}}_{RGB}^{i,j} \right\|, \\ \bar{\mathbf{n}}^{i,j} = \frac{\left( \left[ \mathbf{L}^{i,j} \right]^\top \left[ \mathbf{L}^{i,j} \right] \right)^{-1} \left[ \mathbf{L}^{i,j} \right]^\top \left[ \mathbf{E}_{RGB}^{i,j*} \right] \bar{\mathbf{d}}_{RGB}^{i,j}}{k_d^{i,j}}, \end{cases}$$

where  $\left[ \mathbf{L}^{i,j} \right] = \left[ \bar{\mathbf{I}}_1^{i,j}, \bar{\mathbf{I}}_2^{i,j}, \dots, \bar{\mathbf{I}}_{N_b}^{i,j} \right]^\top$  and  $\text{rank}(\left[ \mathbf{L}^{i,j} \right]) = 3$ . While the effect of specularly rejection through the use of  $\bar{\mathbf{s}}_{RGB}$  is exhibited, the estimation of  $\bar{\mathbf{d}}_{RGB}^{i,j}$  using DPCA is erroneous due to the existence of specularities. Furthermore, all the specularities are rejected as outliers though useful information may be contained.

Unlike Barsky’s PS method, which performs PS in the RGB colour space, the PS method of Zickler estimates surface orientations in the SUV colour space, or the UV colour space as S is not considered:

$$\epsilon^{i,j} = \sum_{k=1}^N \left[ \left( \mathbf{e}_{UV,k}^{i,j*} \right)^\top \bar{\mathbf{d}}_{UV}^{i,j} - \alpha^{i,j} \left( \bar{\mathbf{n}}^{i,j} \right)^\top \bar{\mathbf{I}}_k^{i,j} \right]^2 \rightarrow \min_{\bar{\mathbf{n}}^{i,j}}, \quad (14)$$

where  $\alpha^{i,j}$  is a scaling factor. The UV colour subspace includes more observations than the RGB colour space in surface orientation estimation. To operate in the SUV space, the colour image irradiance is transformed from RGB to SUV as:

$$\mathbf{e}_{SUV,k}^{i,j} \triangleq \mathbf{e}_k^{i,j} = \mathbf{R}_s \mathbf{e}_k^{i,j} = \bar{\mathbf{d}}_{SUV}^{i,j} f_{d,k}^{i,j} + \bar{\mathbf{s}}_{SUV} f_{s,k}^{i,j}, \quad (15)$$



where  $\bar{\mathbf{d}}_{SUV}^{ij} = [\bar{d}_U^{ij}, \bar{d}_V^{ij}, \bar{d}_S^{ij}]^T$ ,  $\bar{\mathbf{s}}_{SUV} = [0, 0, 1]^T$ , and  $\mathbf{R}_s \in SO(3)$  is any transformation that yields  $\bar{\mathbf{s}}_{SUV} = \mathbf{R}_s \bar{\mathbf{s}}_{RGB} = [0, 0, 1]^T$ . UV forms a specular-free colour subspace that is invariant to  $f_{s,k}^{ij}$ . The Zickler’s PS method first estimates  $\bar{\mathbf{d}}_{UV}^{ij}$  of unit length using DPCA on  $[\mathbf{E}_{UV}^{ij*}]^T [\mathbf{E}_{UV}^{ij*}]$ . It then becomes the conventional Lambertian-based PS problem and the solution is obtained in the similar form to Equation (14). Compared with the Barsky’s PS, the Zickler’s PS utilizes more observations and advantageously performs PS in the specular-free UV colour subspace. Shape information from specularities is however neglected again, and the method can only provide diffuse component up to the normalised RGB space. The diffuse-specular separation in the RGB space remains open.

### 3. DRM-Based Colour PS Method Using Diffuse-Specular Separation

#### 3.1. Overview of the Proposed Color PS Method

Figure 3 shows the flow and original contributions of the proposed DRM-based colour PS method using the diffuse-specular separation. The proposed method consists of two processes: Step 1, which is diffuse-specular separation using a PS, and Step 2, which is the surface normal refinement. Step 1 identifies good approximate surface normals, whereas Step 2 further refines and finalizes them with maximum accuracy. The steps are marked by red boxes.

Step 1 is composed of four sub-processes: diffuse colour estimation, diffuse-specular separation in RGB space, PS in UV space and the outlier estimation. After completing shadow rejection as a preprocess, and generating the shadow-free image irradiance matrix  $[\mathbf{E}_{RGB}^{ij*}]$ , Step 1 begins with the diffuse colour estimation using the robust principal component analysis (RPCA) proposed in this paper and derives  $\bar{\mathbf{d}}_{RGB}^{ij}$ . Note that  $(\ )$  is a quantity of  $(\ )$  after shadow rejection. The diffuse color estimation also outputs a specularity map, which describes the distribution of specular reflection. The diffuse-specular separability is then checked for each pixel by deriving the diffuse-specular chromatic angle  $\psi^{ij}$  in RGB space. For separable pixels, the proposed method performs PS in the specular-free UV colour subspace similarly to Zickler’s PS method and derives the initial guess of the unit surface normal  $(\bar{\mathbf{n}}^{ij})_1$  and the diffuse reflectance factor  $(k_d^{ij})_1$  similarly to Barsky’s PS.  $(\ )_1$  means the initial guess of  $(\ )$ . In parallel, the specular geometrical scaling factors,  $(\mathbf{f}_s^{ij})_1$ , are derived through the diffuse-specular separation process using Equation (13) for the subsequent surface normal refinement. Observations not in the specularity map are clamped to zero in  $(\mathbf{f}_s^{ij})_1$ . The surface normals and the other parameters are then refined by fitting the nonlinear DRM using LMA with regularisation.

In Step 2, the initial guess of the specular parameters,  $(k_s^{ij})_1$  and  $(\beta^{ij})_1$ , are derived from  $(\bar{\mathbf{n}}^{ij})_1$  and  $(\mathbf{f}_s^{ij})_1$ . Finally, PS for the surface normal refinement derives  $\bar{\mathbf{n}}^{ij}$  and  $k_d^{ij}$  with enhanced accuracy while additionally updating  $k_s^{ij}$  and  $\beta^{ij}$ . The strength of the proposed method lies in this PS with the diffuse-specular separation. For pixels where sparse non-Lambertian reflectances appear, the first step is sufficient to reliably estimate the surface orientations. To tackle with more general problems with dense non-Lambertian reflectances, the second step is additionally required to exploit surface normal information from specularities in the DRM with the designed parameter initialisation strategy. The rest of this section elucidates the proposed method in detail.

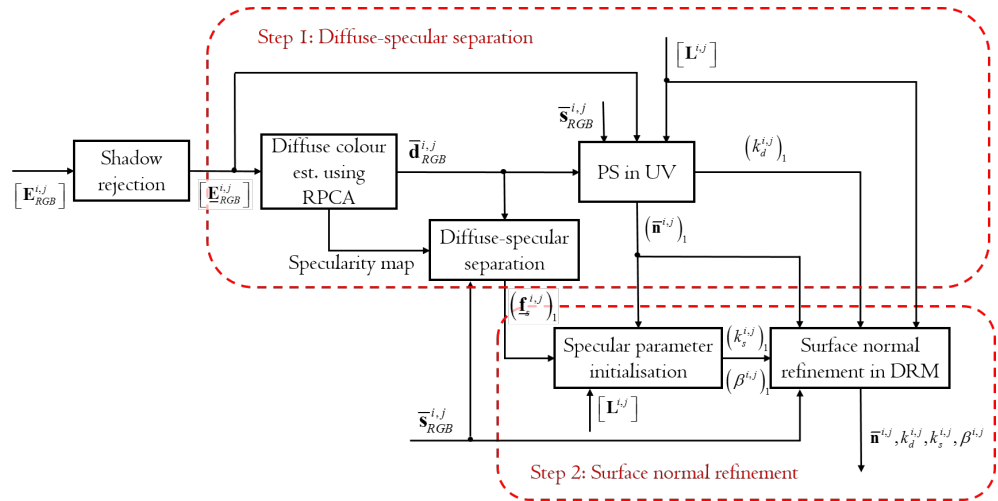


Figure 3. Flows and the original contribution of the DRM-based colour PS.

### 3.2. Diffuse-Specular Separation

#### 3.2.1. Diffuse Color Estimation

The proposed RPCA derives  $\bar{\mathbf{d}}_{RGB}^{i,j}$  and the specularity map as follows:

1. Perform principle component analysis (PCA) for  $[\mathbf{E}_{RGB}^{i,j*}]^\top [\mathbf{E}_{RGB}^{i,j*}]$  to estimate  $\bar{\mathbf{d}}_{RGB}^{i,j}$ ;
2. Compute residual matrix:  $[\mathbf{R}_d^{i,j}] = [\mathbf{E}_{RGB}^{i,j*}] \bar{\mathbf{d}}_{RGB}^{i,j} (\bar{\mathbf{d}}_{RGB}^{i,j})^\top - [\mathbf{E}_{RGB}^{i,j*}]$ .
3. Compute residual vector:  $\mathbf{r}_d^{i,j} = \sqrt{(\mathbf{r}_{d,1}^{i,j})^2 + (\mathbf{r}_{d,2}^{i,j})^2 + (\mathbf{r}_{d,3}^{i,j})^2}$ , where  $[\mathbf{R}_d^{i,j}] = [\mathbf{r}_{d,1}^{i,j}, \mathbf{r}_{d,2}^{i,j}, \mathbf{r}_{d,3}^{i,j}]$  and  $\circ 2$  denotes the hadamard square.
4. If the mean of  $\mathbf{r}_d^{i,j}$ ,  $\text{mean}(\mathbf{r}_d^{i,j})$ , is smaller than the threshold  $T_d$ , terminate RPCA and output the current estimate of  $\bar{\mathbf{d}}_{RGB}^{i,j}$  and the specularity map. Otherwise, find the element that provides the maximal value of  $\left(\frac{\mathbf{r}_d^{i,j} - \text{mean}(\mathbf{r}_d^{i,j})}{\text{std}(\mathbf{r}_d^{i,j})}\right)$  and register  $(i, j)$  in the specularity map. Remove the corresponding row vector in  $[\mathbf{E}_{RGB}^{i,j*}]$  and repeat from step 1.

The RPCA eliminates image irradiances that are inappropriate for diffuse colour estimation.

The validity of the RPCA can be explained as follows. The shadow-free image irradiance matrix  $[\mathbf{E}_{RGB}^{i,j*}]$  is expanded from Equation (7) and decomposed as

$$\begin{aligned}
 [\mathbf{E}_{RGB}^{i,j*}] &= \mathbf{f}_d^{i,j} (\bar{\mathbf{d}}_{RGB}^{i,j})^\top + \mathbf{f}_s^{i,j} (\bar{\mathbf{s}}_{RGB})^\top \\
 &= \begin{bmatrix} f_{d,1}^{i,j} \\ f_{d,2}^{i,j} \\ \dots \\ f_{d,N_p}^{i,j} \end{bmatrix} \begin{bmatrix} \bar{d}_R^{i,j} & \bar{d}_G^{i,j} & \bar{d}_B^{i,j} \end{bmatrix} + \begin{bmatrix} f_{s,1}^{i,j} \\ f_{s,2}^{i,j} \\ \dots \\ f_{s,N_p}^{i,j} \end{bmatrix} \begin{bmatrix} \bar{s}_R & \bar{s}_G & \bar{s}_B \end{bmatrix}
 \end{aligned}
 \tag{16}$$

where most entries in  $\mathbf{f}_s^{i,j}$  are near-zero since images are often near specular-free. In such a case, a non-zero residual occurs if  $\mathbf{f}_s^{i,j}$  contains non-negligible entries. The RPCA algorithms robustly and accurately estimate  $\bar{\mathbf{d}}_{RGB}^{i,j}$  by keeping rejecting specularities until the residual becomes smaller than  $T_d$ . Unlike Barsky’s PS method, which identifies specu-



larities after  $\bar{\mathbf{d}}_{RGB}^{i,j}$  estimation, the proposed algorithms derive  $\bar{\mathbf{d}}_{RGB}^{i,j}$  and the specular map simultaneously.

While the originality and superiority of the proposed RPCA algorithms have been identified, the performance of the RPCA algorithms depends on the value of  $T_d$ . If  $T_d$  is too large, the RPCA algorithms are tolerant to specularities and perform similarly to the DPCA algorithms. If  $T_d$  is too small, the RPCA algorithms reject innocent pixels as specularities and make images sensitive to noise.

### 3.2.2. Diffuse-Specular Separability Check

Having  $\bar{\mathbf{d}}_{RGB}^{i,j}$  estimated, the diffuse-specular chromatic angle, used to perform the separability check, is derived as

$$\cos \psi^{i,j} = \left( \bar{\mathbf{d}}_{RGB}^{i,j} \right)^T \bar{\mathbf{s}}_{RGB}, \tag{17}$$

Asymptotically, the chromatic angle gives

$$\lim_{\psi^{i,j} \rightarrow 0} \bar{\mathbf{d}}_{RGB}^{i,j} = \bar{\mathbf{s}}_{RGB} \tag{18}$$

This means that  $\psi^{i,j}$  should be sufficiently larger than 0 in order for  $\bar{\mathbf{d}}_{RGB}^{i,j}$  and  $\bar{\mathbf{s}}_{RGB}$  to be distinct and separable. The threshold of the chromatic angle to determine the separability is  $T_c$ . If  $\psi^{i,j} \geq T_c$ ,  $\bar{\mathbf{d}}_{RGB}^{i,j}$  and  $\bar{\mathbf{s}}_{RGB}$  are sufficiently distinct and pixel  $(i, j)$  is considered as separable. Similarly to  $T_d$ ,  $T_c$  should be chosen carefully. A small  $T_c$  results in yielding false separation. Once  $\bar{\mathbf{d}}_{RGB}^{i,j}$  and  $\bar{\mathbf{s}}_{RGB}$  have been identified,  $(\hat{\mathbf{f}}_s^{i,j})_1$  is then initialised using Equation (13).

### 3.2.3. PS in UV Space

Since the PS is performed in the UV colour space, the unit diffuse colour  $\bar{\mathbf{d}}_{RGB}^{i,j}$  should be converted into that in UV colour space  $\bar{\mathbf{d}}_{UV}^{i,j}$ . The PS in UV color space allows more observations as described in the Zickler’s PS method. With  $\mathbf{R}_s$ , the estimated  $\bar{\mathbf{d}}_{RGB}^{i,j}$  is transformed to the SUV space as  $\hat{\mathbf{d}}^{i,j} = [\hat{d}_U^{i,j}, \hat{d}_V^{i,j}, \hat{d}_S^{i,j}]^T$ . The diffuse color in UV space  $\hat{\mathbf{d}}_{UV}^{i,j}$  is then given by

$$\hat{\mathbf{d}}_{UV}^{i,j} = [\hat{d}_U^{i,j}, \hat{d}_V^{i,j}]^T = \kappa_z^{i,j} \bar{\mathbf{d}}_{UV}^{i,j} \tag{19}$$

where  $\bar{\mathbf{d}}_{UV}^{i,j} = [\bar{d}_U^{i,j}, \bar{d}_V^{i,j}]^T$  is a unit vector, and

$$\kappa_z^{i,j} = \sqrt{(\hat{d}_U^{i,j})^2 + (\hat{d}_V^{i,j})^2} > 0. \tag{20}$$

Image irradiances of separable pixels are also transformed to the SUV colour space as  $[\mathbf{E}_{SUV}^{i,j}]$  using the same  $\mathbf{R}_s$ .  $[\mathbf{E}_{UV}^{i,j}]$  is then formed by picking the first two columns of  $[\mathbf{E}_{SUV}^{i,j}]$ . Due to the presence of image noise, the PS in UV space is modified from

Equation (15) to reject noise-corrupted image irradiances as outliers using studentised residuals [29]:

$$\begin{aligned}
 \epsilon^{i,j} &= \sum_{k=1}^{N_p} \left[ \left( \mathbf{e}_{UV,k}^{i,j*} \right)^\top \bar{\mathbf{d}}_{UV}^{i,j} - \left( \kappa_z^{i,j} \right)^2 f_{d,k}^{i,j} \right]^2 \\
 &= \sum_{k=1}^{N_p} \left[ \left( \mathbf{e}_{UV,k}^{i,j*} \right)^\top \bar{\mathbf{d}}_{UV}^{i,j} - \left( \kappa_z^{i,j} \right)^2 \left( k_d^{i,j} \right)_1 \left( \bar{\mathbf{n}}^{i,j} \right)_1^\top \bar{\mathbf{I}}_k^{i,j} \right]^2 \\
 &\rightarrow \min_{\left( \bar{\mathbf{n}}^{i,j} \right)_1, \left( k_d^{i,j} \right)_1}, \tag{21}
 \end{aligned}$$

where  $N_p$  is the number of shadow-free images, so  $N_b \leq N_p \leq N$ . The derivation of  $\left( k_d^{i,j} \right)_1$  in addition to  $\left( \bar{\mathbf{n}}^{i,j} \right)_1$  similarly to the Barsky’s PS method enables full recovery of the colour diffuse component.  $N_p$  is used instead of  $N$  and  $N_b$  because shadow-free images are effective for reliability.

After the PS,  $\bar{\mathbf{n}}^{i,j}$  and  $k_d^{i,j}$  are initialised as  $\left( \bar{\mathbf{n}}^{i,j} \right)_1$  and  $\left( k_d^{i,j} \right)_1$  respectively. The diffuse component matrix,  $\left[ \mathbf{E}_{d,RGB}^{i,j} \right]$ , is then re-rendered as:

$$\left[ \mathbf{E}_{d,RGB}^{i,j} \right] = \max \left( \left( k_d^{i,j} \right)_1 \left[ \mathbf{L}^{i,j} \right] \left( \bar{\mathbf{n}}^{i,j} \right)_1, \mathbf{0}_N \right) \left( \bar{\mathbf{d}}_{RGB}^{i,j} \right)^T, \tag{22}$$

where  $\mathbf{0}_N$  is an  $N \times 1$  vector with all zero entries. The specular component matrix,  $\left[ \mathbf{E}_{s,RGB}^{i,j} \right]$ , is obtained by  $\max \left( \left( \bar{\mathbf{f}}_s^{i,j} \right)_1 \bar{\mathbf{s}}_{RGB}^T, \mathbf{0}_N \right)$ . With the separated colour diffuse and specular components, the proposed method allows the functionality of specular removal and intrinsic image decomposition.

### 3.2.4. Outlier Estimation

Let the residual vector  $\underline{\mathbf{r}}^{i,j}$  be written as:

$$\underline{\mathbf{r}}^{i,j} = \left[ \mathbf{E}_{UV}^{i,j} \right] \bar{\mathbf{d}}_{UV}^{i,j} - \left[ \mathbf{L}^{i,j} \right] \mathbf{n}^{i,j} = \left[ \mathbf{E}_{UV}^{i,j} \right] \bar{\mathbf{d}}_{UV}^{i,j} - \left[ \mathbf{L}^{i,j} \right] \rho^{i,j} \bar{\mathbf{n}}^{i,j}, \tag{23}$$

where  $\mathbf{n}^{i,j} = \rho^{i,j} \bar{\mathbf{n}}^{i,j}$  is the scaled normal,  $\left[ \mathbf{L}^{i,j} \right] = \left[ \bar{\mathbf{I}}_1^{i,j}, \bar{\mathbf{I}}_2^{i,j}, \dots, \bar{\mathbf{I}}_M^{i,j} \right]^T$ , and  $\rho^{i,j} = k_d^{i,j} \kappa_z^{i,j}$ . The hat matrix  $\left[ \mathbf{H}_a^{i,j} \right]$  [29] is represented by:

$$\left[ \mathbf{H}_a^{i,j} \right] = \left[ \mathbf{L}^{i,j} \right] \left( \left[ \mathbf{L}^{i,j} \right]^T \left[ \mathbf{L}^{i,j} \right] \right)^{-1} \left[ \mathbf{L}^{i,j} \right]^T. \tag{24}$$

Diagonal entries in  $\left[ \mathbf{H}_a^{i,j} \right]$  are leverages and the  $k$ th leverage is denoted as  $h_{a,k}^{i,j}$ . These leverages quantify the influence that the observed  $\left[ \mathbf{E}_{UV}^{i,j} \right] \bar{\mathbf{d}}_{UV}^{i,j}$  on their predicted values  $\left[ \mathbf{L}^{i,j} \right] \mathbf{n}^{i,j}$ . Each entry in the studentized residual vector  $\underline{\mathbf{r}}^{i,j}$  is then approximated using:

$$\hat{r}_k^{i,j} = \frac{r_k^{i,j}}{\sqrt{(MSE)^{i,j} (1 - h_{a,k}^{i,j})}}, \tag{25}$$

where  $(MSE)^{i,j}$  represents the mean squared error of  $\underline{\mathbf{r}}^{i,j}$ . If any entry in  $|\hat{r}^{i,j}|$  exceeds  $T_o$ , this entry is estimated as outliers. The proposed PS rejects outliers until no more can be detected or  $(MSE)^{i,j}$  is smaller than a tolerance,  $T_m$ . With the general rule of thumb in detecting outliers using studentised residuals,  $T_o$  and  $T_m$  can be chosen as  $2.5$  and  $9\sigma_n^2$ ,

where  $\sigma_n$  is the standard deviation of additive white Gaussian noise that can be estimated using the method given by [30].

### 3.3. Surface Normal Refinement

#### 3.3.1. Specular Parameter Initialisation

If dense non-Lambertian reflectance appears at pixel  $(i, j)$ , which implies that more than one entry in  $(\underline{\mathbf{f}}_s^{i,j})_1$  are in the specularity map, surface normal refinement is necessary since the sparse non-Lambertian reflectance assumption in the diffuse-specular separation is violated. With the estimated  $(\underline{\mathbf{f}}_s^{i,j})_1$ , the next objective is to initialise the specular parameters,  $(k_s^{i,j})_1$  and  $(\beta^{i,j})_1$ . The specific cost functional is given by:

$$\epsilon_s^{i,j} = \sum_{k=1}^{N_q} \left( (f_{s,k}^{i,j})_1 - (k_s^{i,j})_1 \left( (\underline{\mathbf{h}}_k^{i,j})^\top (\underline{\mathbf{n}}^{i,j})_1 \right)^{\beta^{i,j}} \right)^2 \rightarrow \min_{(k_s^{i,j})_1, (\beta^{i,j})_1}, \quad (26)$$

where  $N_q$  is the number of specularity maps showing pixel  $(i, j)$  and  $N_q \geq 2$ . Equation (26) suggests a nonlinear least-squares problem, while it can be manipulated to the natural logarithm domain as a linear problem given by:

$$\epsilon_{ls}^{i,j} = \sum_{k=1}^{N_q} \left( \ln(f_{s,k}^{i,j})_1 - \ln(k_s^{i,j})_1 - (\beta^{i,j})_1 \ln \left( (\underline{\mathbf{h}}_k^{i,j})^\top (\underline{\mathbf{n}}^{i,j})_1 \right) \right)^2 \rightarrow \min_{(k_s^{i,j})_1, (\beta^{i,j})_1}. \quad (27)$$

Let  $(\underline{\mathbf{f}}_s^{i,j})_1$  with  $N_q$  rows consist of entries in  $(\underline{\mathbf{f}}_s^{i,j})_1$  in the specularity map and  $[\underline{\mathbf{H}}]^{i,j} = [\underline{\mathbf{h}}_1^{i,j}, \underline{\mathbf{h}}_2^{i,j}, \dots, \underline{\mathbf{h}}_{N_q}^{i,j}]^\top$  be the matrix comprised of the corresponding unit half vectors. Then, the solution of  $\ln(k_s^{i,j})_1$  and  $(\beta^{i,j})_1$  to Equation (27) is given analytically by:

$$\begin{bmatrix} (\beta^{i,j})_1 \\ \ln(k_s^{i,j})_1 \end{bmatrix} = \left( [\underline{\mathbf{S}}^{i,j}]^\top [\underline{\mathbf{S}}^{i,j}] \right)^{-1} [\underline{\mathbf{S}}^{i,j}]^\top \ln(\underline{\mathbf{f}}_s^{i,j})_1, \quad (28)$$

where

$$[\underline{\mathbf{S}}^{i,j}] = \left[ \ln([\underline{\mathbf{H}}]^{i,j} (\underline{\mathbf{n}}^{i,j})_1), \mathbf{1}_{N_q} \right] \quad (29)$$

and  $\mathbf{1}_{N_q}$  is a  $N_q \times 1$  vector with all entries equal to 1.

#### 3.3.2. Surface Normal Refinement in DRM

All parameters in the DRM have been initialised up to this point. The LMA is then employed to iteratively refine these parameters by solving the optimization problem:

$$\begin{aligned} \epsilon^{i,j} = \sum_{k=1}^N & \left( (\mathbf{e}_k^{i,j})^\top \bar{\mathbf{s}}_{RGB} - k_d^{i,j} (\underline{\mathbf{n}}^{i,j})^\top \bar{\mathbf{I}}_k^{i,j} (\underline{\mathbf{d}}_{RGB}^{i,j})^\top \bar{\mathbf{s}}_{RGB} - k_s^{i,j} \left( (\underline{\mathbf{n}}^{i,j})^\top \underline{\mathbf{h}}_k^{i,j} \right)^{\beta^{i,j}} \right. \\ & \left. + T_\alpha \left( 1 - (\underline{\mathbf{n}}^{i,j})^\top \underline{\mathbf{n}}^{i,j} \right) \right)^2 \rightarrow \min_{\underline{\mathbf{n}}^{i,j}, k_d^{i,j}, k_s^{i,j}, \beta^{i,j}}, \quad (30) \end{aligned}$$

where the first three terms are from the generalised formulation, and, the fourth regularization term with constant  $T_\alpha$  is added. The regularisation term reduces the change of surface normal refinement from the initial guess of  $(\underline{\mathbf{n}}^{i,j})_1$  and prevents overfitting. Larger value of  $T_\alpha$  strengthens the robustness but limits the capability of surface normal refinement.

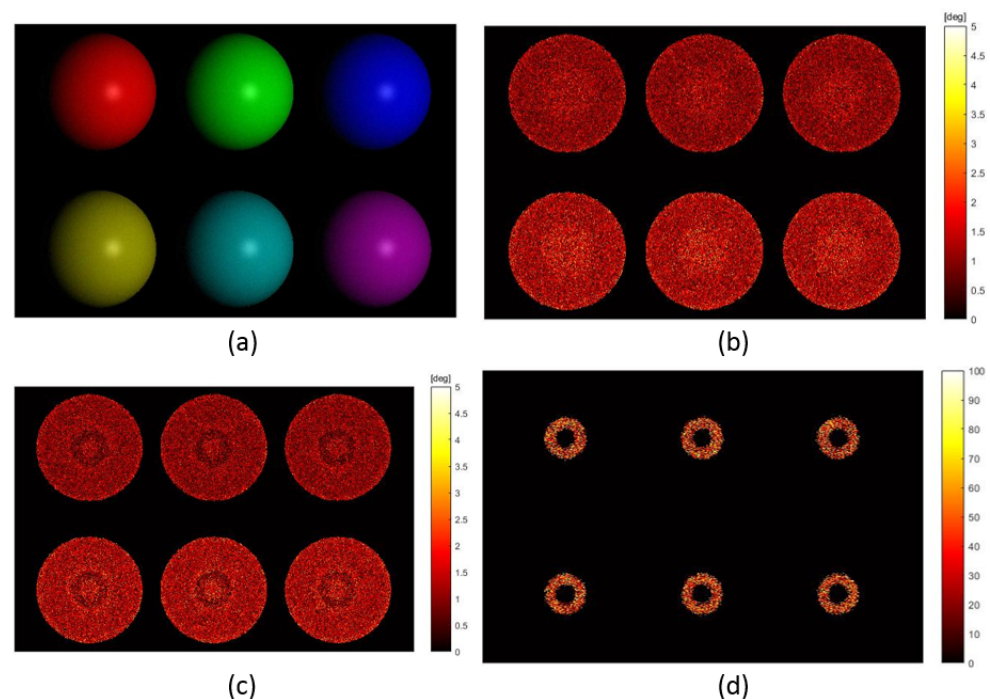
The strength of the proposed PS method lies in the determination of the surface normals while separating the diffuse and specular components in RGB space, eliminating outliers and then tuning the diffuse and specular reflectance factors simultaneously. The de-

termination considering diffuse and specular reflections simultaneously makes the surface normal estimation more accurately. In addition, the proposed method can be used for specular removal [31] and intrinsic image decomposition [32]. Parameters characterizing  $f_{s,k}$  are determined at pixels, so the proposed method could also be used for digital relighting and material classification where dense non-Lambertian reflectance appears.

#### 4. Performance Evaluation on Surface Orientation Estimation

##### 4.1. Evaluations Using Synthetic Images

The first experiment aims to evaluate the effectiveness of surface normal refinement for dense non-Lambertian reflectance using synthetic input images. A scene with six different-coloured spheres were rendered under 32 illuminants using the Blinn-Phong model. The sphere was adopted as the scene geometry since it samples all the possibilities of surface orientations for the visible surface. Different colours of the spheres were introduced for more comprehensive evaluations of the proposed PS method on surfaces with various spectral reflectances. The 32 illuminants were chosen to provide a sufficient number of pixels with dense non-Lambertian reflectance for evaluation.  $\bar{s}_{RGB}$  was set at  $[0.5774, 0.5774, 0.5774]^T$ . Three colours of the spheres were red, green and blue with the same  $\psi^{i,j}$  of  $57.74^\circ$ , while the other three spheres were yellow, cyan and magenta with  $\psi^{i,j}$  of  $35.26^\circ$ . The 32 light positions were configured with  $r = 442$  mm and  $\theta = 20^\circ$ . The reference plane was located where  $\mathbf{R}_W = \text{diag}([1, -1, -1])$  and  $\mathbf{t}_W = [0, 0, 678]^T$ .  $k_d^{i,j}$ ,  $k_s^{i,j}$  and  $\beta^{i,j}$  were set the same across the field of view as 0.4, 0.2 and 100, respectively. Additive Gaussian noise was introduced with  $\sigma_n = 0.02$ . The image irradiance under the first illuminant is shown by Figure 4a. The parameters of the proposed PS method were configured as:  $T_d = 0.01$ ,  $T_c = 5^\circ$ ,  $T_o = 2.5$ ,  $T_m = 0.0036$ ,  $T_\alpha = 3$ ,  $T_x = 0.02$  and  $T_y = 0.02$ . 100 repeated tests were conducted and the mean value was adopted for evaluation.



**Figure 4.** Evaluation of surface normal refinement: (a) Image irradiance under the first illuminant; (b) Angular error of surface orientations without surface normal refinement in degrees; (c) Angular error of surface orientations with surface normal refinement in degrees; (d) Improvement of surface orientation estimation by including surface normal refinement in percentage.

Figure 4b shows the angular error of surface orientation estimation without surface normal refinement. As is shown, the central regions of the spheres where dense non-

Lambertian reflectances appear have larger error. By including the proposed surface normal refinement step, the angular errors of surface orientations around the detected dense non-Lambertian reflectance region, shown by a ring, are significantly reduced as shown by Figure 4c. This is where the reflectance parameters are most tuned through the surface normal refinement. Errors in such regions are even smaller than those with sparse non-Lambertian reflectance, which further enhances that making specularities as meaningful signals is beneficial for surface orientation estimation. Figure 4d demonstrates the improvement percentage for quantitative evaluation. The improvement percentage is defined reduction of the error divided by the error without surface normal refinement. The accuracy of surface orientations was enhanced by 34.33% in median and 32.25% on average. The first and third quantile values of the improvement were 15.76% and 54.23%, respectively. The performance improvements onto the six different-coloured spheres are shown in Table 1. The improvements were obvious for all spheres is significant, implying that the method was applicable to a wide range of surfaces with different spectral reflectances. For spheres that had larger chromatic angle,  $\psi^{i,j}$ , the  $\bar{\mathbf{d}}_{RGB}^{i,j}$  estimations were more accurate. The better estimated  $\bar{\mathbf{d}}_{RGB}^{i,j}$  led to the more accurate recovery of  $(\mathbf{f}_s^{i,j})_1$ , resulting in the more reliable initialisation of  $(\bar{\mathbf{n}}^{i,j})_1$ . Such initialisations of  $\bar{\mathbf{n}}^{i,j}$  ultimately affected the different improvement performances. Overall, the accuracy was enhancement by including the additional surface normal refinement step and for spectral reflectances with larger chromatic angles, the improvement was more phenomenal.

**Table 1.** Improvement on different-coloured spheres

Index	Colour	$\psi^{i,j}$	Error of $\bar{\mathbf{d}}_{RGB}^{i,j}$	Mean Improvement
1	red	57.74°	1.15°	23.39%
2	yellow	35.26°	1.33°	32.85%
3	green	57.74°	1.15°	23.42%
4	cyan	35.26°	1.32°	32.04%
5	blue	57.74°	1.15°	23.15%
6	magenta	35.26°	1.33°	32.24%

In the second experiment, the performance improvement due to surface normal refinement was verified onto a wide range of reflectances. Experimental settings and parameters of the proposed method were the same as the first experiment except that different values of  $k_d^{i,j}$ ,  $k_s^{i,j}$  and  $\beta^{i,j}$  were applied to represent various surface reflectances when generating synthetic images. The ratio of  $\kappa_{ds}^{i,j} = k_d^{i,j}/k_s^{i,j}$  represents the relative strength between the diffuse and specular components, whereas  $\beta^{i,j}$  indicates the width of the specular lobe. The mean, median, first and third quantile values of the improvement were adopted for evaluation and the angular error of  $\bar{\mathbf{d}}_{RGB}^{i,j}$  estimation was used for analyses. As shown from Table 2, the improvements on all the different reflectances were significant and more obvious for larger value of  $\beta^{i,j}$  and  $\kappa_{ds}^{i,j}$ . Larger value of  $\beta^{i,j}$  suggesting narrower specular lobe made the  $\bar{\mathbf{d}}_{RGB}^{i,j}$  estimation more accurate since the sparse non-Lambertian reflectance assumption was valid. Similarly, larger value of  $\kappa_{ds}^{i,j}$  indicating stronger diffuse components made better  $\bar{\mathbf{d}}_{RGB}^{i,j}$  estimation due to stronger inliers. More accurate  $\bar{\mathbf{d}}_{RGB}^{i,j}$  estimation led to bigger improvement from the surface normal refinement. In summary, with the additional surface normal refinement step, the accuracy of surface orientation estimation is enhanced by around 30% on average and the improvement is more obvious for reflectance that has narrower specular lobe and stronger diffuse component.

**Table 2.** Evaluation of the improvement for different reflectances.

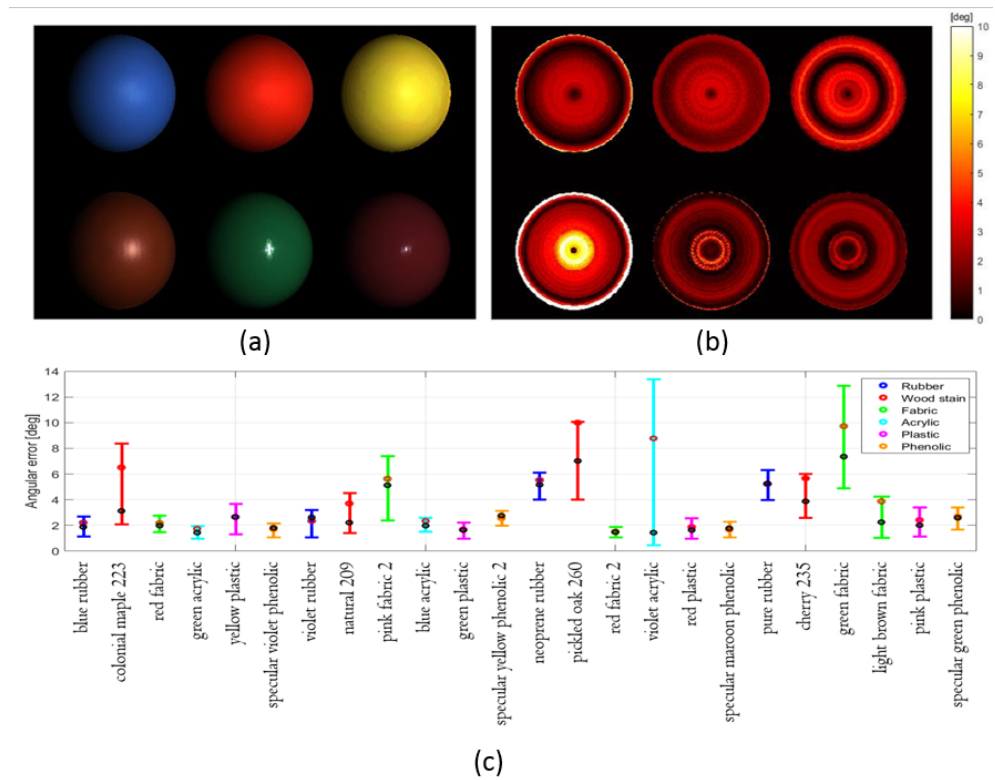
$k_d/k_s$	$\beta$	Mean	Median	First Quantile	Third Quantile	Error of $\bar{\mathbf{d}}_{RGB}$
0.4/0.2	100	32.25%	34.33%	15.76%	54.23%	1.23°
0.4/0.4	100	31.14%	32.97%	15.21%	53.75%	1.58°
0.4/0.8	100	30.48%	32.77%	14.27%	52.90%	2.34°
0.4/0.2	20	27.49%	26.70%	6.80%	51.66%	2.99°
0.4/0.4	20	25.92%	25.98%	6.67%	51.25%	5.06°
0.4/0.8	20	25.02%	24.07%	6.57%	50.81%	8.30°

The third experiment aims to evaluate the overall performance of the proposed colour PS method onto a wide variety of dielectric material reflectances. Twenty-four material BRDFs divided into six categories were evaluated, where the different BRDFs were provided from the MERL database [33]. Experimental settings and the method parameters were exactly the same as the first experiment except that the BRDFs were replaced. The image irradiance for the first six materials under the first illuminant is shown in Figure 5a. The angular errors of the surface orientation estimation are given in Figure 5b. As is shown, the error is not only larger at the near central regions where specularities overlap but also at the boundaries. Boundary pixels in the spheres have shallow grazing angles and the error is due to the lack of modeling of the fresnel reflection in the Blinn-Phong model. Figure 5c shows the performance of surface normal estimation on the 24 materials using box-and-whisker plot. The red and black dot represent the mean and median value, respectively. The lower and upper bound of the box indicates the first and third quantile values. The six colours of the boxes suggest the six material categories. As is shown, the proposed method can accurately estimate  $\bar{\mathbf{n}}^{i,j}$  for most dielectric materials within 5 degrees on average. The estimation performances on phenolic, plastic and rubber were consistent, whereas those on the wood stain, fabric and acrylic were also acceptable with two exceptions, the violet acrylic and the green fabric. The degradation for violet acrylic was due to its wide specular lobe with small  $\kappa_{ds}^{i,j}$ , while the reason for green fabric was because of its small  $\psi^{i,j}$ .

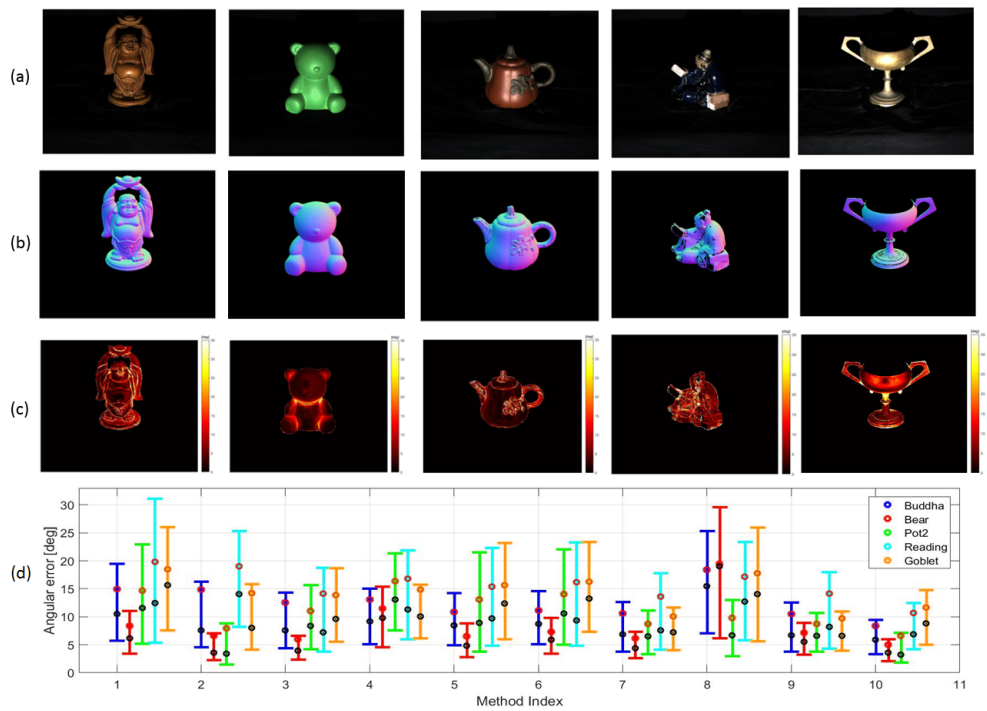
#### 4.2. Evaluations Using Real Images

The proposed method was evaluated on five different datasets comprised of real images in the DiLigenT database [4]. Parameters of the proposed method were set the same as the first experiment.  $\bar{\mathbf{s}}_{RGB}$  was estimated as  $[0.5774, 0.5774, 0.5774]^T$ . From the first to fifth column in Figure 6, the results on BUDDHA, BEAR, POT2, READING and GOBLET are respectively shown. Figure 6a shows the image irradiance under the first illuminant, while Figure 6b,c demonstrate the estimated normal map and the angular error of surface orientation estimation, respectively. As is shown from READING, the proposed method is only feasible for surfaces whose  $\bar{\mathbf{d}}_{RGB}^{i,j}$  and  $\bar{\mathbf{s}}_{RGB}$  are distinct. Figure 6c shows that the dense non-Lambertian problem where the specularity overlaps is mostly solved for BUDDHA, BEAR and POT2, while deficits still exist on GOBLET, which is because the DRM is not proper to characterize metallic reflectance.





**Figure 5.** (a) Image irradiance for the first six material BRDFs under the first illuminant; (b) Angular error of surface orientation estimation for the first six material BRDFs; (c) Angular error of surface orientation for the twenty-four dielectric materials.



**Figure 6.** Evaluations on surface orientation estimation: (a) Image irradiance under the first illuminant; (b) Estimated normal map; (c) Angular error of surface orientations; (d) Angular error of surface orientations on five datasets using ten different PS methods

The results were compared with those of nine other PS methods which are listed in Table 3. The results from the nine methods were provided from the DiLigenT database and

these methods were representative and well-recognized. It is to be noted that the results of the recent machine learning based techniques reviewed in Introduction are not shown for comparison since the effect of the physics based models is investigated in this paper. Due to the data availability, refs. [21,23] were not involved in the comparison, while the essence of their works in using DRM with known  $\bar{\mathbf{s}}_{RGB}$  was well-inherited by the proposed method.

**Table 3.** Comparative studies of PS methods

Method Index	1	2	3	4	5	6	7	8	9	10
Reference	[3]	[9]	[11]	[12]	[19]	[20]	[13]	[34]	[14]	proposed

Figure 6d shows the results from ten PS methods on the five datasets. The X-axis represents the method index and the Y-axis indicates the angular error of  $\bar{\mathbf{n}}^{i,j}$ . The five colours of the bar represent the different datasets. The angular error of  $\bar{\mathbf{n}}^{i,j}$  is also demonstrated using the box-and-whisker plot similar to Figure 5. As is shown, the proposed method performs well in estimating surface orientations for all the five datasets except GOBLET. The results suggest that modeling reflectance using DRM is proper for surfaces made of dielectric materials. The strength of the proposed method lies on the reliable estimation of  $\bar{\mathbf{d}}_{RGB}^{i,j}$  with simultaneous specular detection. The advantage is also accredited by exploiting  $\bar{\mathbf{n}}^{i,j}$  not only from the diffuse components but also from the specularities owing to the diffuse-specular separation. In spite of its advantages, the proposed method is not effective for metallic surfaces compared with methods, such as [14]. The limitation originates from the fundamental assumption of DRM that  $\bar{\mathbf{s}}_{RGB}$  are the same across the surface. This assumption is violated for metallic surfaces since the wavelength and geometry exhibit inter-dependency.

## 5. Conclusions and Future Works

A PS method using colour images dealing with non-Lambertian reflectance has been proposed. The method formulates the imaging photometry using DRM with known specular colour. It extracts surface orientations not only from the diffuse components but also specularities owing to the diffuse-specular separation. Introducing the additional surface refinement step using information from specularities, the proposed method particularly improves the accuracy for surface orientation estimation at pixels where dense non-Lambertian reflectance appear. The simultaneously acquired specular parameters can be applied for more potential functionalities, such as digital relighting and material classification.

From the experiment of validating the effectiveness of the newly proposed surface normal refinement step on surface where dense non-Lambertian reflectance appears, the results indicate that with the additional step, the accuracy is enhanced by around 30% on average and the improvement is more phenomenal for surfaces with larger chromatic angles, stronger diffuse components and narrower specular lobes due to the better estimated unit diffuse colour. The result investigating the proposed method applicability on 24 different reflectances of dielectric materials suggests that the average angular error of surface orientation estimation for most materials are within  $5^\circ$  and bigger errors occur at surface patches with shallower grazing angles due to the limitation of the Blinn-Phong model that does not incorporate the fresnel reflection. From the systematic comparison on five datasets with nine other representative methods, the proposed method shows its descent performance on reflectances of dielectric materials and degradation on metallic surface due to the limitation of DRM. The result has also demonstrated that the proposed method is only feasible for surfaces whose diffuse and specular colour are distinct.

This paper presented the first set of results to estimate surface orientations using colour images for non-Lambertian reflectance, while it can be extended in a variety of ways. First, other analytical BRDFs can be employed instead of the Blinn-Phong model to better account for the fresnel reflection. Second, machine learning based models can be incorporated to learn and reduce errors of analytical models. Further, an alternative

method can be proposed specifically for surfaces whose diffuse and specular colours are close to extend the method's applicability. Making the method adaptive is one possible solution. Last but not least, it is also of particular interest to infer surface properties from the estimated specular parameters for the purpose of material classification.

**Author Contributions:** Conceptualization, B.L. and T.F.; methodology, B.L. and T.F.; software, B.L.; validation, B.L.; formal analysis, B.L.; investigation, B.L.; resources, B.L.; data curation, B.L.; writing—original draft preparation, B.L.; writing—review and editing, T.F.; visualization, B.L.; supervision, T.F.; project administration, T.F.; funding acquisition, T.F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is sponsored in part by US Office of Naval Research (N00014-20-1-2468).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to express their sincere gratitude to Tom McKenna at Office of Naval Research for his invaluable advice.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Ackermann, J.; Goesele, M. A survey of photometric stereo techniques. *J. Found. Trends Comput. Graph. Vis.* **2015**, *9*, 149–254. [[CrossRef](#)]
2. Klasing, K.; Althoff, D.; Wollherr, D.; Buss, M. Comparison of surface normal estimation methods for range sensing applications. In Proceedings of the 2009 IEEE International Conference on Robotics and Automation, Kobe, Japan, 12–17 May 2009; pp. 3206–3211.
3. Woodham, R.J. Photometric method for determining surface orientation from multiple images. *Opt. Eng.* **1980**, *19*, 139–144. [[CrossRef](#)]
4. Shi, B.; Zhe, W.; Mo, Z.; Duan, D.; Yeung, S.; Tan, P. A benchmark dataset and evaluation for non-Lambertian and uncalibrated photometric stereo. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 3707–3716.
5. Shell, J.R. Bidirectional Reflectance: An Overview with Remote Sensing Applications & Measurement Recommendations. 2004; pp. 37–50. Available online: <http://web.gps.caltech.edu/~vijay/Papers/BRDF/shell-04.pdf> (accessed on 4 January 2022).
6. Nayar, S.K.; Ikeuchi, K.; Kanade, T. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE Trans. Robot. Autom.* **1990**, *6*, 418–431. [[CrossRef](#)]
7. Torrance, K.E.; Sparrow, E.M. Theory for off-specular reflection from roughened surfaces. *J. Opt. Soc. Am.* **1967**, *57*, 1105–1114. [[CrossRef](#)]
8. Georghiadis, A. Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. In Proceedings of the Proceedings Ninth IEEE International Conference on Computer Vision, Nice, France, 13–16 October 2003; pp. 816–823.
9. Goldman, D.B.; Curless, B.; Hertzmann, A.; Seitz, S.M. Shape and spatially-varying BRDFs from photometric stereo. *IEEE Trans. PAMI* **2010**, *32*, 1060–1071. [[CrossRef](#)] [[PubMed](#)]
10. Ward, G.J. Measuring and modeling anisotropic reflection. In *SIGGRAPH92: Proceedings of the 19th Annual Conference and Exhibition on Computer Graphics and Interactive Techniques*; Association for Computing Machinery: New York, NY, USA, 1992; pp. 265–272.
11. Alldrin, N.G.; Zickler, T.; Kriegman, D.J. Photometric stereo with non-parametric and spatially-varying reflectance. In Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 23–28 June 2008; pp. 1–8.
12. Higo, T.; Matsushita, Y.; Ikeuchi, K. Consensus photometric stereo. In Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 13–18 June 2010; pp. 1157–1164.
13. Shi, B.; Tan, P.; Matsushita, Y.; Ikeuchi, K. Bipolynomial modeling of low-frequency reflectances. *IEEE Trans. PAMI* **2014**, *36*, 1078–1091. [[CrossRef](#)] [[PubMed](#)]
14. Ikehata, S.; Aizawa, K. Photometric stereo using constrained bivariate regression for general isotropic surfaces. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Columbus, OH, USA, 23–28 June 2014; pp. 23–28.
15. Santo, H.; Samejima, M.; Sugano, Y.; Shi, B.; Matsushita, Y. Deep photometric stereo network. In Proceedings of the International Workshop on PBDL in Conjunction with IEEE ICCV, Venice, Italy, 22–29 October 2017; pp. 1–9.
16. Santo, H.; Samejima, M.; Sugano, Y.; Shi, B.; Matsushita, Y. Deep photometric stereo networks for determining surface normal and reflectances. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 114–128. [[CrossRef](#)] [[PubMed](#)]
17. Taniai, T.; Maehara, T. Neural inverse rendering for general reflectance photometric stereo. In Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 10–15 July 2018.
18. Ikehata, S. CNN-PS: CNN-based photometric stereo for general non-convex surfaces. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–18.

19. Wu, L.; Ganesh, A.; Shi, B.; Matsushita, Y.; Wang, Y.; Ma, Y. Robust photometric stereo via low-rank matrix completion and recovery. In Proceedings of the 10th Asian Conference on Computer Vision, Queenstown, New Zealand, 8–12 November 2010; pp. 703–717.
20. Ikehata, S.; Wipf, D.; Matsushita, Y.; Aizawa, K. Robust photometric stereo using sparse regression. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 318–325.
21. Barsky, S.; Petrou, M. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE Trans. PAMI* **2003**, *25*, 1239–1252. [[CrossRef](#)]
22. Shafer, S.A. Using colour to separate reflection components. *Color Res. Appl.* **1985**, *10*, 210–218. [[CrossRef](#)]
23. Zickler, T.; Mallick, S.P.; Kriegman, D.J.; Belhumeur, P.N. Color subspace as photometric invariants. *Int. J. Comput. Vis.* **2008**, *79*, 13–30. [[CrossRef](#)]
24. Hartley, R.; Zisserman, A. Chapter 6: Camera Models. In *Multiple View Geometry*; Cambridge University Press: Cambridge, UK, 2003; pp. 153–177.
25. Angelopoulou, E. Specular highlight detection based on the fresnel reflection coefficient. In Proceedings of the 2007 IEEE 11th International Conference on Computer Vision, Rio de Janeiro, Brazil, 14–21 October 2007; pp. 1–8.
26. Blinn, J.F. Models of light reflection for computer synthesized pictures. In Proceedings of the 4th Annual Conference on Computer Graphics and Interactive Techniques, San Jose, CA, USA, 20–22 July 1977; pp. 192–198.
27. Ngan, A.; Dur, F.; Matusik, W. Experimental validation of analytical BRDF models. In Proceedings of the ACM SIGGRAPH 2004 Sketches, Los Angeles, CA, USA, 8–12 August 2004; p. 90.
28. Kao, Y.H.; Roy, B.V. Directed Principal Component Analysis. *Oper. Res.* **2014**, *62*, 957–972. [[CrossRef](#)]
29. Pope, A.J. *The Statics of Residuals and the Detection of Outliers*; NOAA Technical Report; 1976; Volume 1, p. 65. Available online: [https://www.ngs.noaa.gov/PUBS\\_LIB/TRNOS65NGS1.pdf](https://www.ngs.noaa.gov/PUBS_LIB/TRNOS65NGS1.pdf) (accessed on 4 January 2022).
30. Liu, C.; Freeman, W.T.; Szeliski, R.; Kang, S.B. Noise estimation from a single image. In Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; pp. 901–908.
31. Artusi, A.; Banterle, F.; Chetverikov, D. A survey of specular removal methods. *Comput. Graph. Forum* **2011**, *30*, 2208–2230. [[CrossRef](#)]
32. Barrow, H.; Tanenbaum, J. Recovering intrinsic scene characteristic from images. *Proc. Comput. Vis. Syst.* **1978**, *2*, 3–26.
33. Matusik, W.; Pfister, H.; Br, M.; McMillan, L. A Data-Driven Reflectance Model. *ACM Trans. Graph.* **2003**, *22*, 759–769. [[CrossRef](#)]
34. Shi, B.; Tan, P.; Matsushita, Y.; Ikeuchi, K. Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances. In Proceedings of the European Conference on Computer Vision, Florence, Italy, 7–13 October 2012; pp. 455–468.