

# Simultaneous Acquisition of Polarimetric SVBRDF and Normals

SEUNG-HWAN BAEK and DANIEL S. JEON, KAIST

XIN TONG, Microsoft Research Asia

MIN H. KIM, KAIST

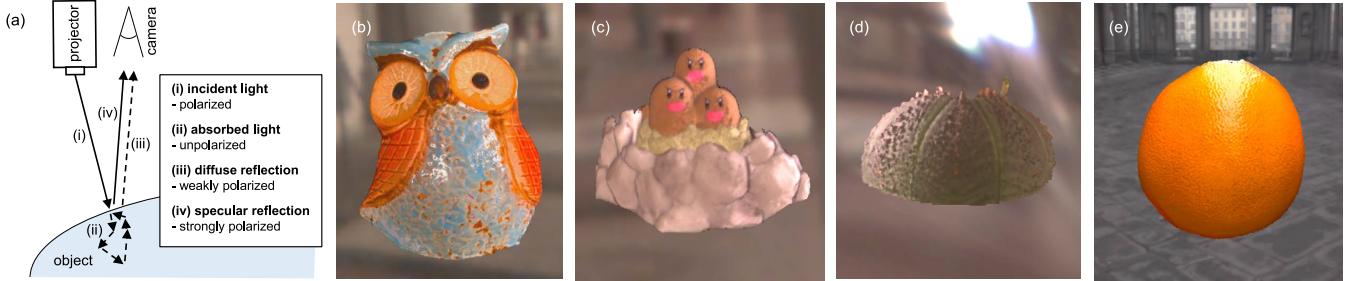


Fig. 1. (a) We propose a complete polarimetric BRDF model that defines both specular and diffuse polarimetric reflection. The model enables us to capture spatially-varying polarimetric BRDF and high-frequency normals using a compact setup with a single projector and camera. Images (b) – (e) show environment rendering results of 3D objects captured from a novel light and view.

Capturing appearance often requires dense sampling in light-view space, which is often achieved in specialized, expensive hardware setups. With the aim of realizing a compact acquisition setup without multiple angular samples of light and view, we sought to leverage an alternative optical property of light, *polarization*. To this end, we capture a set of polarimetric images with linear polarizers in front of a single projector and camera to obtain the appearance and normals of real-world objects. We encountered two technical challenges: First, no complete polarimetric BRDF model is available for modeling mixed polarization of both specular and diffuse reflection. Second, existing polarization-based inverse rendering methods are not applicable to a single local illumination setup since they are formulated with the assumption of spherical illumination. To this end, we first present a complete polarimetric BRDF (pBRDF) model that can define mixed polarization of both specular and diffuse reflection. Second, by leveraging our pBRDF model, we propose a novel inverse-rendering method with joint optimization of pBRDF and normals to capture spatially-varying material appearance: per-material specular properties (including the refractive index, specular roughness and specular coefficient), per-pixel diffuse albedo and normals. Our method can solve the severely ill-posed inverse-rendering problem by carefully accounting for the physical relationship between polarimetric appearance and geometric properties. We demonstrate how our method overcomes limited sampling in light-view space for inverse rendering by means of polarization.

CCS Concepts: • Computing methodologies → Reflectance modeling; Computational photography;

Additional Key Words and Phrases: SVBRDF acquisition, polarization

Part of this work was done during Seung-Hwan Baek's internship at Microsoft Research Asia.

Authors' addresses: Seung-Hwan Baek; Daniel S. Jeon, KAIST, School of Computing, Daejeon, South Korea, 34141; Xin Tong, Microsoft Research Asia, Beijing, China, 100080; Min H. Kim, KAIST, School of Computing, Daejeon, South Korea, 34141, minhkim@kaist.ac.kr(corresponding-author).

© 2018 Association for Computing Machinery.

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *ACM Transactions on Graphics*, <https://doi.org/10.1145/3272127.3275018>.

## ACM Reference Format:

Seung-Hwan Baek, Daniel S. Jeon, Xin Tong, and Min H. Kim. 2018. Simultaneous Acquisition of Polarimetric SVBRDF and Normals. *ACM Trans. Graph.* 37, 6, Article 268 (November 2018), 15 pages. <https://doi.org/10.1145/3272127.3275018>

## 1 INTRODUCTION

Capturing rich material appearance is critical for photorealistic rendering. Traditionally, material acquisition is often achieved by dense sampling in light-view space [Ghosh et al. 2010, 2011, 2008; Holroyd et al. 2010; Nagano et al. 2015; Nam et al. 2016; Tunwattanapong et al. 2013]. In addition to the long acquisition time for dense angular sampling, accurate optical calibration requires considerable effort and time. This limits accessibility of appearance acquisition for casual users. To address costs for material acquisition, several groups have sought alternative solutions [Nam et al. 2018; Nielsen et al. 2015; Riviere et al. 2017; Xu et al. 2016]. Our work is devised with the same objective of capturing rich material appearance with minimal cost. We identified the potential for a compact setup that can be built with a *single projector and camera*. To this end, we exploited an alternative optical property of light, *polarization*, to overcome limited sampling in light-view space.

However, we encountered two technical challenges: First, there is no complete polarimetric BRDF (pBRDF) model that can define polarization of both specular and diffuse reflection. Traditional pBRDF models [Hyde IV et al. 2009; Priest and Gerner 2000] formulate only specular polarization, assuming the polarization of diffuse reflection is negligible to avoid modeling complexity. Shape-from-polarization (SfP) methods [Cui et al. 2017; Kadambi et al. 2015] account for either specular or diffuse polarization exclusively, yielding a per-pixel binary labeling problem. Second, existing polarization-based methods for inverse rendering [Ghosh et al. 2011, 2008; Ma et al. 2007; Nagano et al. 2015] are formulated with the assumption of *spherical illumination*, such as a light stage or natural illumination

captured with a light probe. Therefore, they are inapplicable to a simple setup with single local illumination.

In this work, we present two novel contributions: First, we devise a complete pBRDF model that can define mixed polarization of both specular and diffuse reflection. Second, by leveraging our pBRDF model, we propose a novel inverse-rendering method with joint optimization of pBRDF and normals to estimate spatially-varying appearance: per-material specular properties including the refractive index, specular roughness and specular coefficient, per-pixel diffuse albedo and normals.

Since our optimization formulation carefully accounts for the physical relationship between various polarimetric appearance and geometric properties simultaneously, our method can solve the severely ill-posed inverse-rendering problem successfully in a compact setup. In particular, compared to existing SfP approaches, our method can capture surface normals with high accuracy, thanks to joint optimization of spatially-varying pBRDF and normals.

In summary, our main **contributions** are as follows:

- A complete pBRDF model that defines mixed polarization of both specular and diffuse reflection and
- A novel inverse-rendering method that can estimate spatially varying appearance and high-frequency normals simultaneously using our pBRDF model.

## 2 RELATED WORK

Focusing on polarization, we review existing pBRDF models, inverse-rendering methods using polarization, and shape-from-polarization methods. We refer the reader to recent review works on appearance acquisition [Guarnera et al. 2016; Weinmann and Klein 2015; Weyrich et al. 2009] and works that present a comprehensive background on polarization [Wilkie and Weidlich 2012].

### 2.1 Polarimetric BRDF Models

Various pBRDF models have been proposed to characterize polarization of reflection accurately [Hyde IV et al. 2009; Priest and Gerner 2000]. They define *reflection* as a mixture of *unpolarized* diffuse and *polarized* specular reflection; i.e., the diffuse component is assumed to be unpolarized for the purpose of keeping the mathematical form of the reflectance uncomplicated, although several studies have found that this assumption is not always valid [Ellis 1996; Maxwell et al. 1973; Sun 2007]. A similar simplification on diffuse polarization has also been employed in simulations of polarimetric appearance. Several bidirectional rendering works have been proposed for simulating polarized light transport [Jarabo and Arellano 2017; Mojzik et al. 2016]. They simulate polarimetric light transport of specular reflection, but diffuse reflection is also assumed as *unpolarized*.

In the previous approaches for modeling and rendering polarimetric appearance, the polarization of specular reflection has been exploited extensively, but the *polarization of diffuse reflection* has been ignored to avoid modeling complexity of polarimetric reflection. However, diffuse polarization is not negligible, as shown in the physics literature on the Fresnel wave theory [Collett 2005]. In this work, we introduce a novel pBRDF model that defines polarization of both specular and diffuse reflection.

### 2.2 Capturing Appearance from Polarization

In early works, polarization is mainly used for separating specular and diffuse reflections by cross-polarization, assuming that diffuse polarization is negligible while specular polarization is distinct [Ghosh et al. 2008; Ma et al. 2007; Nagano et al. 2015]. Subsequent polarization-based acquisition methods leverage *specular polarization* under *spherical illumination* to estimate appearance parameters [Ghosh et al. 2010, 2011; Kim et al. 2016; Nagano et al. 2015; Riviere et al. 2017]. Also, photometric stereo methods with polarization have been introduced [Miyazaki et al. 2003; Tozza et al. 2017], ignoring variation of material properties such as the refractive index and specular roughness.

There are two notable differences between our method and prior polarization-based appearance methods: First, our projector-camera approach is *compact* and beneficial in terms of the hardware cost and form factor. It does not require expensive hardware, such as a light stage, for spherical illumination, and it does not assume any outdoor illumination captured with a light probe. Second, on the other hand, lighting in our setup is *local* illumination along with a single light vector, which causes inverse rendering to be more severely ill-posed than spherical illumination in previous work. To mitigate the ill-posedness, we account for both *specular* and *diffuse polarization* in reflection, successfully capturing spatially varying polarimetric appearance and high-frequency normals with the compact setup.

### 2.3 Capturing Shape from Polarization

Polarization has been leveraged to estimate high-frequency surface normals in many SfP methods [Atkinson and Hancock 2006; Cui et al. 2017; Guarnera et al. 2012; Kadambi et al. 2015; Miyazaki et al. 2003]. Previous approaches assume that either diffuse or specular reflection is dominant for each pixel, where the separation between specular and diffuse reflections has been heuristically resolved with binary labeling. However, this assumption is impractical for real-world materials because reflection is a mixture of both specular and diffuse polarization. Separating specular and diffuse polarization from reflected light remains an open research problem [Cui et al. 2017; Kadambi et al. 2015].

There are three key differences between our method and existing SfP methods: First, our model is not bounded to the modality of specular and diffuse polarization by defining them together as a single pBRDF model. Second, we estimate per-material refractive indices, which are leveraged to resolve the inherent ambiguity of surface normals from polarization. Third, we formulate a novel optimization problem that can solve the azimuthal ambiguity of polarization normals<sup>1</sup> by means of rough base normals obtained from structured lighting. To this end, by the joint optimization of polarimetric appearance and normals, our method outperforms existing state-of-the-art SfP methods.

## 3 BACKGROUND ON POLARIZATION

Polarization describes the oscillating states of electric components of an electromagnetic wave. This section provides the foundations of polarization.

<sup>1</sup>When measuring surface normals from polarization in conventional SfP methods, the estimated surface azimuth angle could be flipped with an additional angle  $\pi$ .

*Stokes vector.* A Stokes vector  $\mathbf{s} \in \mathbb{R}^{4 \times 1}$  has been popularly used to represent the polarization state of a light wave. It is a four-dimensional coordinate vector described as  $\mathbf{s} = [s_0, s_1, s_2, s_3]^T$ , where  $s_0$  is the intensity of light  $L$ , and  $s_1$  and  $s_2$  are the powers of the  $0^\circ$  and  $45^\circ$  linear polarization components, respectively, and  $s_3$  is the power of the right circular polarization component. The degree of polarization (DoP) is defined as the ratio of the magnitude of the polarized vector elements to the intensity of the light:  $\psi = \sqrt{s_1^2 + s_2^2 + s_3^2}/s_0$ . A Stokes vector  $\mathbf{s} = [s_0, s_1, s_2, s_3]^T$  can also be defined as  $[L, L\psi \cos 2\zeta \cos 2\xi, L\psi \sin 2\zeta \cos 2\xi, L\psi \sin 2\xi]^T$ , where  $\zeta$  is the polarization angle and  $\xi$  is the ellipticity angle.

*Mueller matrices.* A transformation of a Stokes vector, e.g., a reflection event of polarized light, can be represented as a Mueller matrix, a 4-by-4 matrix for which  $\mathbf{M} \in \mathbb{R}^{4 \times 4}$ :  $\mathbf{s}_{\text{after}} = \mathbf{Ms}_{\text{before}}$ , where  $\mathbf{s}_{\text{before}}$  and  $\mathbf{s}_{\text{after}}$  are the Stokes vectors before and after the event. We employ four standard Mueller matrix transformations [Collett 2005]: Fresnel transmission/reflection, coordinate rotation, linear polarization, and depolarization to formulate our pBRDF model.

*Vector coordinate systems.* A Stokes vector of a light wave is defined with respect to a vector coordinate system, where the  $z$ -axis is aligned along the propagation direction of the light, and the orientations of the  $x$  and  $y$  axes vary depending on the polarization state of the light. Before applying any transformation to a Stokes vector, the coordinate system of the Stokes vector should be adjusted so that it is the same system as used for the transformation.

Refer to the **supplemental material** for more details on standard Mueller matrices, the conversion of polarimetric coordinate systems, and a table of entire notations used in the paper.

#### 4 COMPLETE POLARIMETRIC REFLECTANCE MODEL

This section will explain the derivation of our novel pBRDF model. Suppose we capture the polarized light of a Stokes vector  $\mathbf{s}_o$  reflected on an object surface, illuminated by the polarized light of a Stokes vector  $\mathbf{s}_i$ . Here the surface reflectance can be formulated as a Mueller transformation matrix  $\mathbf{P} \in \mathbb{R}^{4 \times 4}$  that describes a polarimetric BRDF. The general light transport of the polarized light can be defined in a vector-matrix form:

$$\mathbf{s}_o = (\mathbf{n} \cdot \mathbf{i}) \mathbf{Ps}_i, \quad (1)$$

where the pBRDF model  $\mathbf{P}$  can be formulated as the sum of the diffuse reflectance  $\mathbf{P}^d$  and the specular reflectance  $\mathbf{P}^s$ :  $\mathbf{P} = \mathbf{P}^d + \mathbf{P}^s$ . According to the traditional microfacet theory [Torrance and Sparrow 1967], the diffuse term originates from the approximated sum of both subsurface scattering and multiple reflections in microfacets [Lee et al. 2018], and the specular term originates from single-bounce mirror reflection on microfacets subject to shadowing and masking. Figure 2 depicts these two reflection phenomena of diffuse and specular polarization under polarimetric illumination.

##### 4.1 Polarization of Diffuse Reflection

In our diffuse reflectance model, the polarization state of subsurface scattering is assumed to be completely unpolarized (Figure 2(a)). However, when unpolarized light comes out of the material, it is

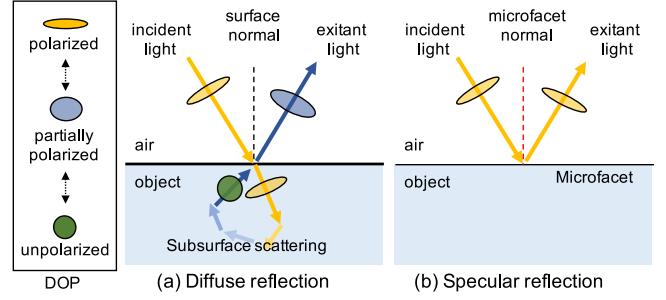


Fig. 2. Diffuse vs. specular reflectance under polarimetric illumination. (a) In diffuse reflection, light is absorbed into the object to be unpolarized by subsurface scattering. When diffused light comes out of the object, it is partially-linearly polarized again, passing through the interface between the object and air. (b) In specular reflection, light is reflected on *mirror-like* microfacet surfaces, of which the facet normal is identical to the halfway vector. Specular reflection therefore holds strong polarization.

partially polarized again due to the difference between the refractive indices of the material and air, following the Fresnel theory [Collett 2005].

To account for this diffuse polarization effect, our novel diffuse reflectance term formulates both incident and exitant polarization through light transport, different from existing polarimetric appearance models, yielding a Mueller matrix  $\mathbf{P}^d$  as follows:

$$\mathbf{P}^d = \mathbf{C}_{n \rightarrow o}(\phi_o) \mathbf{F}_o^T(\theta_o; \eta) \mathbf{D}(\rho) \mathbf{F}_i^T(\theta_i; \eta) \mathbf{C}_{i \rightarrow n}(-\phi_i), \quad (2)$$

where  $\theta_{i,o}$  is the zenith angle between the normal  $\mathbf{n}$  and the incident/exitant light and  $\phi_{i,o}$  is the azimuth angle between the plane of incident and the  $y$ -axis of the incident/exitant light.

Here,  $\mathbf{F}_{i,o}^T$  is the incident/exitant Fresnel transmission matrix between the object and air with respect to the object surface, and  $\mathbf{D}$  is a depolarization matrix, where the (0,0) element of diffuse albedo  $\rho$  is the only non-zero element in the matrix, accounting for subsurface scattering with the object pigments. We also have two standard coordinate conversion Mueller matrices: one matrix from the light to the plane of incidence (that holds  $\mathbf{n}$ )  $\mathbf{C}_{i \rightarrow n}(-\phi_i)$  and other matrix from the plane of incidence to the camera system  $\mathbf{C}_{n \rightarrow o}(\phi_o)$ . Refer to the supplemental material for more details on standard Mueller matrices,  $\mathbf{F}^T$ ,  $\mathbf{D}$  and  $\mathbf{C}$ . Note that our reflection model is defined with respect to the observation coordinate system.

Hence, combining diffuse reflectance  $\mathbf{P}^d$  with shading  $(\mathbf{n} \cdot \mathbf{i})$ , our polarimetric diffuse shading model  $\mathbf{H}^d$  can be written as a Mueller matrix by calculating Equation (2) with each component:

$$\mathbf{H}^d = \rho(\mathbf{n} \cdot \mathbf{i}) \begin{bmatrix} T_o^+ T_i^+ & T_o^+ T_i^- \beta_i & -T_o^+ T_i^- \alpha_i & 0 \\ T_o^- T_i^+ \beta_o & T_o^- T_i^- \beta_i \beta_o & -T_o^- T_i^- \alpha_i \beta_o & 0 \\ -T_o^- T_i^+ \alpha_o & -T_o^- T_i^- \alpha_o \beta_i & T_o^- T_i^- \alpha_i \alpha_o & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad (3)$$

where  $\alpha_{i,o}$  and  $\beta_{i,o}$  denote  $\sin(2\phi_{i,o})$  and  $\cos(2\phi_{i,o})$  of the incident/exitant azimuth angles of the polarized light, respectively. Here  $T_{i,o}^+$  and  $T_{i,o}^-$  denote the calculations of Fresnel transmission coefficients,  $(T_{i,o}^\perp + T_{i,o}^{\parallel})/2$  and  $(T_{i,o}^\perp - T_{i,o}^{\parallel})/2$ , respectively.  $T^\perp$  and  $T^{\parallel}$  are the Fresnel transmission coefficients for the perpendicular (denoted by  $\perp$ ) and the parallel ( $\parallel$ ) components:

$$T^\perp = \left( \frac{2\eta_1 \cos \theta_1}{\eta_1 \cos \theta_1 + \eta_2 \cos \theta_2} \right)^2, \quad T^{\parallel} = \left( \frac{2\eta_1 \cos \theta_1}{\eta_1 \cos \theta_2 + \eta_2 \cos \theta_1} \right)^2, \quad (4)$$

where  $\theta_1$  and  $\theta_2$  are the incident and exitant angles, and  $\eta_1$  and  $\eta_2$  are the refractive indices of the medium before and after the interface, respectively.  $\eta_1$  and  $\eta_2$  are set to 1.0 and the object refractive index  $\eta$ , individually, for incident Fresnel transmission coefficients and vice versa for exitant coefficients.

In summary, our polarimetric diffuse shading model is a four-dimensional function  $H^d(\theta_i, \theta_o, \phi_i, \phi_o; \rho, \eta, \mathbf{n})$ , which can predict polarimetric changes of diffuse reflection with the given diffuse albedo  $\rho$  and the refractive index  $\eta$ . We can rewrite the model with vector notations as  $H^d(y_i, y_o, \mathbf{i}, \mathbf{o}; \rho, \eta, \mathbf{n})$ , where  $\theta_i = \cos^{-1}(\mathbf{n} \cdot \mathbf{i})$ ,  $\theta_o = \cos^{-1}(\mathbf{n} \cdot \mathbf{o})$ ,  $\phi_i = \cos^{-1}(\mathbf{n}_i \cdot \mathbf{y}_i)$ ,  $\phi_o = \cos^{-1}(\mathbf{n}_o \cdot \mathbf{y}_o)$ . Here  $\mathbf{n}_{i,o}$  is the projected normal vector to the incident/exitant light frames,  $\mathbf{y}_i$  is the  $y$ -axis of the illumination coordinate system, and  $\mathbf{y}_o$  is the  $y$ -axis of the camera coordinate system.

*Discussion.* The main difference between our diffuse polarization model and the prior works [Atkinson and Hancock 2006; Kadambi et al. 2015] is that we account for both incident and exitant polarization effects by including the incident Fresnel matrix  $F_i^T(\theta_i; \eta)$  before applying the depolarization matrix. This means that our diffuse reflection model can predict the intensity change of diffuse polarization according to the polarization state of the incident light.

#### 4.2 Polarization of Specular Reflection

According to the microfacet theory [Torrance and Sparrow 1967], specular reflection is assumed as direct mirror reflections of a single bounce on microfacets, the orientation of which is the same as the halfway vector  $\mathbf{h} = \frac{\mathbf{i} + \mathbf{o}}{\|\mathbf{i} + \mathbf{o}\|}$ . Therefore, the Stokes vector of specular polarization should be calculated with respect to the halfway vector  $\mathbf{h}$ , rather than with the surface normal  $\mathbf{n}$ . See Figure 2(b). We formulate the specular reflectance term that accounts for surface roughness, geometric attenuation, and the Fresnel effect, following the traditional microfacet theory. Here, we elaborate the description of the Fresnel effect in the form of a Mueller matrix including coordinate conversion matrices, which replaces the Fresnel term in the original formulation with  $C_{h \rightarrow o} F^R C_{i \rightarrow h}$ :

$$P^s = \left( \frac{D(\theta_h; \sigma) G(\theta_i, \theta_o; \sigma)}{4 \cos(\theta_i) \cos(\theta_o)} \right) C_{h \rightarrow o}(\varphi_o) F^R(\theta_d; \eta) C_{i \rightarrow h}(-\varphi_i), \quad (5)$$

where  $\theta_h = \cos^{-1}(\mathbf{n} \cdot \mathbf{h})$  is the zenith angle between the normal  $\mathbf{n}$  and halfway vector  $\mathbf{h}$ ,  $\theta_d = \cos^{-1}(\mathbf{h} \cdot \mathbf{i})$  is the zenith angle between incident light  $\mathbf{i}$  and halfway vector  $\mathbf{h}$  [Rusinkiewicz 1998],  $\sigma$  is the surface roughness parameter for the GGX distribution  $D$  [Walter et al. 2007],  $G$  is Smith's shadowing/masking function [Heitz 2014] for geometric attenuation,  $\eta$  is the refractive index for three color channels,  $\varphi_{i,o}$  is the azimuth angle between the plane of incident (that holds the halfway vector  $\mathbf{h}$  for specularity) and the  $y$ -axis of the incident/exitant light, and  $C_{i \rightarrow h}(-\varphi_i)$  and  $C_{h \rightarrow o}(\varphi_o)$  are the coordinate conversion matrices.

Here  $F^R(\theta_d; \eta)$  is calculated with angle  $\theta_d$  with the following Fresnel reflection coefficients  $R^\perp$  and  $R^\parallel$ :

$$R^\perp = \left( \frac{\eta_1 \cos \theta_1 - \eta_2 \cos \theta_2}{\eta_1 \cos \theta_1 + \eta_2 \cos \theta_2} \right)^2, R^\parallel = \left( \frac{\eta_1 \cos \theta_2 - \eta_2 \cos \theta_1}{\eta_1 \cos \theta_2 + \eta_2 \cos \theta_1} \right)^2, \quad (6)$$

where  $\eta_1$  and  $\eta_2$  are 1.0 and the object refractive index  $\eta$ , respectively. Also,  $\cos \theta_1$  and  $\cos \theta_2$  are defined as  $\cos \theta_1 = \cos \theta_d$  and  $\cos \theta_2 = \sqrt{1 - ((1/\eta) \sin \theta_1)^2}$ , respectively, following Snell's law.

Finally, combining specular reflection  $P^s$  with shading  $(\mathbf{n} \cdot \mathbf{i})$ , our polarimetric specular shading model  $H^s$  can be derived from Equation (5) in the Mueller matrix form:

$$H^s = k_s \frac{DG}{4(\mathbf{n} \cdot \mathbf{o})(\mathbf{n} \cdot \mathbf{i})} (\mathbf{n} \cdot \mathbf{i}) \begin{bmatrix} R^+ & R^- \gamma_i & -R^- \chi_i & 0 \\ R^- \gamma_o & R^+ \gamma_i \gamma_o + R^\times \chi_i \chi_o \cos \delta & -R^+ \chi_i \gamma_o + R^\times \gamma_i \chi_o \cos \delta & \chi_o R^\times \sin \delta \\ -R^- \chi_o & -R^+ \gamma_i \chi_o + R^\times \chi_i \gamma_o \cos \delta & R^+ \chi_i \chi_o + R^\times \gamma_i \gamma_o \cos \delta & \gamma_o R^\times \sin \delta \\ 0 & -\chi_i R^\times \sin \delta & -\gamma_i R^\times \sin \delta & R^\times \cos \delta \end{bmatrix}, \quad (7)$$

where  $\chi_{i,o}$  and  $\gamma_{i,o}$  denote  $\sin(2\varphi_{i,o})$  and  $\cos(2\varphi_{i,o})$  of the incident/exitant azimuth angles of the polarized specular light, respectively,  $k_s$  is the specular coefficient as a scalar,  $\cos \delta$  is  $-1$  for a dielectric surface, when the incident angle is less than the Brewster angle;  $\cos \delta = 1$ , otherwise, and vice versa for  $\sin \delta$ . Here  $R^+$ ,  $R^-$ , and  $R^\times$  denote the calculations of the Fresnel reflection coefficients,  $(R^\perp + R^\parallel)/2$ ,  $(R^\perp - R^\parallel)/2$ , and  $\sqrt{R^\perp R^\parallel}$ , respectively.

In summary, our polarimetric specular shading model is a bidirectional function  $H^s(\theta_i, \theta_o, \theta_d, \varphi_i, \varphi_o; \sigma, \eta, \mathbf{n})$ , where  $\varphi_i = \cos^{-1}(\mathbf{h}_i \cdot \mathbf{y}_i)$ ,  $\varphi_o = \cos^{-1}(\mathbf{h}_o \cdot \mathbf{y}_o)$ . Here  $\mathbf{h}_{i,o}$  denotes the projections of  $\mathbf{h}$  to the incident/exitant polarization plane of light, respectively. We can reparameterize the specular model as  $H^s(y_i, y_o, \mathbf{i}, \mathbf{o}; \sigma, \eta, k_s, \mathbf{n})$ , which allows predictions of the polarimetric specular appearance changes of the given surface roughness  $\sigma$ , the refractive index  $\eta$ , and normals  $\mathbf{n}$ .

## 5 SURFACE APPEARANCE FROM POLARIZATION

*Overview.* As input, we capture a set of polarized images with a combination of linear polarizers on the camera and projector and a rough geometry  $\mathbf{n}_b$  of an object using structured lighting with the same projector. Based on our pBRDF model, we initially decompose a set of polarized intensities  $I$  into a polarimetric shading matrix  $H$  per pixel, which is decomposed to diffuse and specular polarization,  $H^d$  and  $H^s$ . In our joint optimization of the appearance and normals, they are then used to estimate the refractive indices of red, green and blue channels  $\eta_{1\dots 3}$  based on  $\mathbf{n}_b$ .  $\eta_{1\dots 3}$  allows us to update per-pixel surfaces normals  $\mathbf{n}$  with high-frequency details using  $H^d$ . We then estimate both the surface roughness  $\sigma$  and specular coefficient  $k_s$  from the previously estimated  $\mathbf{n}$ , which refines the normals  $\mathbf{n}$  with those in the next iteration with more details in the specular-dominant region. Finally, we estimate diffuse albedo  $\rho$  using  $\eta_{1\dots 3}$  and  $\mathbf{n}$  with consideration of the Fresnel effect on  $H^d$ .

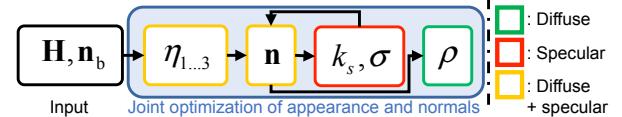


Fig. 3. Overview of our joint optimization of polarimetric appearance and high-frequency normals. Color lines of boxes indicate, which polarimetric property is used for estimating surface appearance.

### 5.1 Designing the Polarization-based Acquisition Setup

*Ideal setup.* Suppose that we have a polarized light source and a polarized camera on the same optical axis, known as the *coaxial*

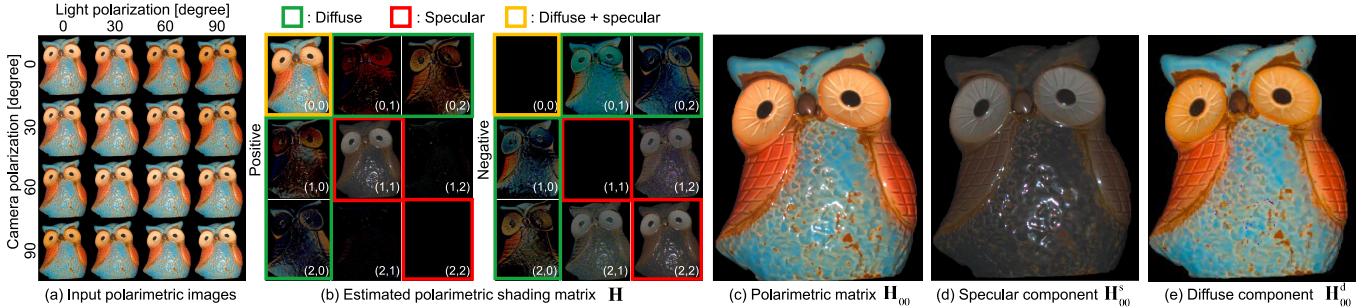


Fig. 4. Polarimetric decomposition. (a) We take polarimetric images with different combinations of linear polarization angles as input. (b) A polarimetric shading matrix  $H$  is then estimated per pixel. The positive and negative values in a wide range are visualized with different gammas (2.2 for the diagonal; 4 for the non-diagonal). Images (c), (d) & (e) show initial specular/diffuse separation of per-pixel polarimetric matrix component  $H$ .

setup, and that we can change the angles of the linear polarizers on each device individually. The ideal coaxial setup allows us to simplify the original formulation of our polarimetric appearance matrix  $H = H^d + H^s$  into a concise form. In this case, many negligible values can be approximated as zero in  $H$ . First, in the diffuse term  $H^d$  (Equation (3)), we can regard  $T_o^- T_i^-$  as close to zero due to the low diffuse degree of polarization. Second, in the specular term  $H^s$  (Equation (7)), the difference between the Fresnel reflection coefficients  $R^-$ , also approaches zero, as the parallel and perpendicular components of the Fresnel reflection coefficients become close ( $R^{\parallel} \approx R^{\perp}$ ) [Collett 2005]. This consequently leads to another simplification:  $R^+ \approx R^{\times}$ . Also,  $\gamma_i, \gamma_o$  and  $\xi_i, \xi_o$  are close to zero and one, respectively, because the Fresnel reflection effect occurs on the microfacet, for which the facet normal is close to both the view and illumination vectors. In addition,  $\cos \delta$  is  $-1$  for the dielectric surface as the incident angle to the microfacet is close to zero for the coaxial setup, where it is apparently less than the Brewster angle [Collett 2005]. In summary, the polarimetric shading matrix  $H$  can be simplified by having all the simplifications together ( $R^- \approx 0$ ,  $R^+ \approx R^{\times}$ ,  $\gamma_i \approx 1$ ,  $\gamma_o \approx 1$ ,  $\xi_i \approx 0$ ,  $\xi_o \approx 0$ ,  $\cos \delta \approx -1$ ,  $T_o^- T_i^- \approx 0$ ):

$$H \approx (\mathbf{n} \cdot \mathbf{i}) \begin{bmatrix} \rho T_o^+ T_i^+ + CR^+ & \rho T_o^+ T_i^- \beta_i & -\rho T_o^+ T_i^- \alpha_i & 0 \\ \rho T_o^- T_i^+ \beta_o & CR^+ & 0 & 0 \\ -\rho T_o^- T_i^- \alpha_o & 0 & -CR^+ & 0 \\ 0 & 0 & 0 & -CR^+ \end{bmatrix}, \quad (8)$$

where  $C$  is the coefficient of the specular term:  $C = k_s \frac{DG}{4(\mathbf{n} \cdot \mathbf{o})(\mathbf{n} \cdot \mathbf{i})}$ .

*Real setup.* However, owing to the imperfection of the non-polarized beamsplitter, it is impossible to build a perfect coaxial setup with two controllable linear polarizers. We therefore build a suboptimal coaxial setup, as shown in Figure 5. We place the camera and the projector in the shortest baseline, about 7.5 cm, where the orientations of both devices are seemingly identical and calibrated. The relatively longer distance between the object and the capture system is set to approximately  $\sim 110$  cm. After validating our coaxial assumption with a synthetic dataset and a comparison between real data and the synthetic ground truth, we could assume that our real setup can be approximated with the coaxial formulation of  $H$  in Equation (8).

*Initial normals.* Using a projector as a light source allows us to obtain additional information using the pair of the camera and the projector. This projector is used for not only illumination but also

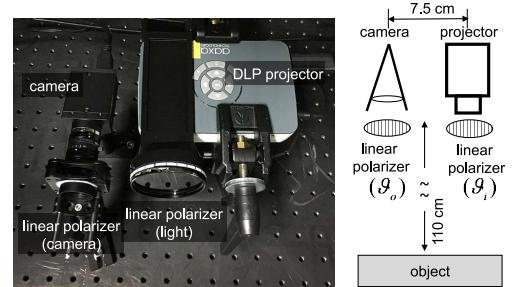


Fig. 5. Our polarimetric imaging setup. Glass-based linear polarization filters are mounted on the camera and the DLP projector (an unpolarized one). This projector is used as not only the light source but also for structured light scanning.

for structured light scanning. We capture a rough geometry using the projector with a popular structured lighting method [Moreno and Taubin 2012]. Once the point clouds are obtained, the per-point surface normals are estimated using [Hoppe et al. 1992] and this is followed by projecting the point-cloud to the image space of the camera. We propagate the sparsely projected depth and normals to the entire pixel [Levin et al. 2007] resulting in an initial rough normal map  $\mathbf{n}_b$ . For example, Figure 10(b) shows the initial rough surface normals  $\mathbf{n}_b$  from the structured lighting method compared to the ground truth.

## 5.2 Polarimetric Shading Decomposition

*Image formation.* We use the unpolarized camera and the unpolarized projector with two linear polarizers. Therefore, the Stokes vector of the light  $\mathbf{s}_i$  is  $[L, 0, 0, 0]^T$ , where  $L$  is the scalar of the maximum intensity of the projector light. The captured intensity  $I$  with the given polarization angles of the linear polarizers ( $\theta_i, \theta_o$ ) can be formulated as:  $I(\theta_i, \theta_o) = \mathbf{s}_{o,0}$ .  $I$  is normalized by the maximum radiance  $L$  of the light source. These linear polarizers can be formulated as standard linear polarization matrices,  $\mathbf{L}(\theta_i)$  and  $\mathbf{L}(\theta_o)$  (refer to the supplemental material for details on the Mueller matrix). We can extend Equation (1) as an image formation model in our setup:  $\mathbf{s}_o = \mathbf{L}(\theta_o) \mathbf{H}(\theta_i) \mathbf{s}_i$ .

Given this configuration, the per-pixel intensity  $I(\theta_i, \theta_o)$  is formed by only the intensity of light with given polarization angles  $\theta_i$  and  $\theta_o$  so that our light transport model can be simplified further for linear polarization. First, the incident light intensity  $\mathbf{s}_{i,0}$  is multiplied by only the first column of the incident linear polarimetric transmission matrix  $\mathbf{L}$ . Second, the captured light intensity  $I = \mathbf{s}_{o,0}$

is obtained from the product of the first row components in the linear polarimetric transmission matrix and the first column components in the reflection matrix  $\mathbf{H}$  per pixel. We therefore define a simple transmission matrix  $\ell(\vartheta) = \frac{1}{2} [1, \cos 2\vartheta, \sin 2\vartheta, 0]^T \in \mathbb{R}^{4 \times 1}$ . This can be used as the replacement of  $\mathbf{L}(\vartheta)$ , yielding a simpler image formation model:

$$I(\vartheta_i, \vartheta_o) = \ell(\vartheta_o)^T \mathbf{H} \ell(\vartheta_i) = \ell(\vartheta_o)^T (\mathbf{n} \cdot \mathbf{i}) \mathbf{P} \ell(\vartheta_i). \quad (9)$$

As shown in Figures 4(a), we capture a set of input images with four different angles using two linear filters on each project and each camera, setting the number of angular intervals as 4 of  $\vartheta_i, \vartheta_o \in \{0, \pi/6, \pi/3, \pi/2\}$  in Equation (9).

*Decomposition.* Suppose we have a set of multiple images with every combination of  $m$  number of incident angles  $\vartheta_i$  and  $n$  number of exitant angles  $\vartheta_o$ . In the matrix-vector form, the pixel-intensity stack of the normalized polarimetric observations with every combination of  $\vartheta_i$  and  $\vartheta_o$  can be expressed as  $\mathbf{I} \in \mathbb{R}^{n \times m}$ . The set of polarimetric observations can be expressed by means of incident and exitant linear polarization matrices:  $\Phi_i \in \mathbb{R}^{4 \times m}$  and  $\Phi_o \in \mathbb{R}^{4 \times n}$ . These matrices are made up of  $\ell_{i,o} \in \mathbb{R}^{4 \times 1}$  vectors for each combination of  $\vartheta_i$  and  $\vartheta_o$ .

Our total observation of every combination of linear polarization (Equation (9) for each observation) is now formulated in a matrix-vector form as:  $\mathbf{I} = \Phi_o^T \mathbf{H} \Phi_i$ , where  $\mathbf{H} \in \mathbb{R}^{4 \times 4}$  is the polarimetric shading matrix. By capturing at least nine combinations of linear filters of the light and the camera, we can estimate the polarimetric shading matrix  $\mathbf{H}$  by solving an overdetermined system:

$$\underset{\mathbf{H}}{\text{minimize}} \|\mathbf{I} - \Phi_o^T \mathbf{H} \Phi_i\|_2^2. \quad (10)$$

We solve it using the standard least-squares method. As the unknown matrix  $\mathbf{H}$  is surrounded by the observation matrices, we solve Equation (10) for  $\mathbf{H} \Phi_i$  first, yielding an intermediate matrix of  $\mathbf{H}' = \mathbf{H} \Phi_i$ . As a consequence, we solve another objective problem:  $\underset{\mathbf{H}'}{\text{minimize}} \|\mathbf{H}' - \mathbf{H} \Phi_i\|_2^2$  using the least-squares method again. The estimated  $\mathbf{H}$  enables the formulation of a polarimetric shading equation for inverse rendering. Note that Equation (10) is a per-pixel optimization. In practice, we solved the optimization simultaneously for every pixel by reformulating Equation (10) into matrix form concatenated for every pixel. Refer to the supplemental material for more details.

Figures 4(a) and (b) show the input image examples and decomposed  $\mathbf{H}$  matrix, respectively. Note that because we use linear polarizers, there are no values in the fourth column and row of  $\mathbf{H}$ . We therefore show only 3-by-3 images with positive and negative values. There are notable elements in  $\mathbf{H}$ . First,  $\mathbf{H}_{00}$  (Figure 4(c)) corresponds to the (0,0) element in Equation (8),  $(\mathbf{n} \cdot \mathbf{i})(\rho T_o^+ T_i^+ + CR^+)$ . Moreover, the second term of specular reflection  $(\mathbf{n} \cdot \mathbf{i})CR^+$  appears as the positive and the negative values in  $\mathbf{H}_{11}$  and  $\mathbf{H}_{22}$ , respectively, along the diagonal in Equation (8) allowing an estimate of the specular shading component  $\mathbf{H}_{00}^s$  from the average value:  $(|\mathbf{H}_{11}| + |\mathbf{H}_{22}|)/2$ . See Figure 4(d). In fact, except for  $\mathbf{H}_{00}$ , there is no diffuse dependent component along the diagonal elements in  $\mathbf{H}$  so that we can obtain diffuse shading component  $\mathbf{H}_{00}^d$  by substituting specular shading component  $\mathbf{H}_{00}^s$  from  $\mathbf{H}_{00}$  (Figure 4(e)). This allows us to initially

separate diffuse and specular shading components (including  $(\mathbf{n} \cdot \mathbf{i})$ ) from  $\mathbf{H}$ .

### 5.3 Estimating Appearance from Polarization

*Spatially-varying BRDFs.* Given the manually determined number of materials  $B$ , our per-material polarimetric BRDF  $\{\mathbf{P}^b | b \in \{1, \dots, B\}\}$  is defined on a material mask of each cluster  $\{\omega_b | b \in \{1, \dots, B\}\}$ . To estimate the mask  $\omega_b$ , we apply  $k$ -means clustering on the diffuse reflection image of  $\mathbf{P}^d$ , obtained by the structured light normal  $\mathbf{n}_b$  and the diffuse shading component  $\mathbf{H}^d$ :  $\mathbf{P}^d = \mathbf{H}^d / (\mathbf{n}_b \cdot \mathbf{i})$ . The top row in Figure 8 shows an estimated cluster map. Note that different from conventional SVBRDF works [Chen et al. 2014; Nam et al. 2016; Zhou et al. 2016], we do not include the blending weights of basis BRDFs for computational efficiency to avoid non-negative constrained optimization of blending weights.

*Refractive index.* The refractive index (RI) determines the ratio between the transmitted and reflected energy for parallel and perpendicular polarization components, according to the Fresnel theory. Existing polarization-based acquisition methods [Ghosh et al. 2010; Riviere et al. 2017] estimate the refractive index from specular reflection, assuming that the surface is illuminated under spherical lighting, which allows for devising a function that can describe the refractive index and specular intensity. However, in our setup, where the illumination is local and directional, the specular reflection information is insufficient to estimate the refractive index robustly. Instead, we utilize diffuse polarization in particular for robustly estimating the refractive index. To this end, we devise a novel optimization problem that leverages both diffuse polarization and specular shading, yielding the per-material refractive index  $\eta^{b \in \{1 \dots B\}}$ .

Because the refractive index is a wavelength-dependent property, we estimate the refractive index per color channel  $\eta_{1 \dots 3}^b$  by solving a joint optimization problem [Waltz et al. 2006] that consists of two energy terms:  $E_d$  for exitant diffuse polarization and  $E_s$  for specular shading, as follows:

$$\underset{\eta_{1 \dots 3}^b}{\text{minimize}} w_d E_d(\eta_{1 \dots 3}^b) + w_s E_s(\eta_{1 \dots 3}^b), \quad (11)$$

where  $w_d$  and  $w_s$  are the weights for  $E_d$  and  $E_s$  respectively, set to 1 and  $10^{-2}$  in our experiments.

We formulate the first term  $E_d$  by leveraging the degree of *diffuse polarization* of exitant light  $\psi_o^d$ , which is defined as the ratio of the two Fresnel terms:  $\psi_o^d = T_o^- / T_o^+$  (Equation (4)). According to the Fresnel theory [Atkinson and Hancock 2006],  $\psi_o^d$  is defined as a function  $f_o$  of the refractive index  $\eta$  and the surface zenith angle  $\theta_z$ :  $\psi_o^d = f_o(\eta; \theta_z)$ , as shown in Figure 6(a).

Given  $\psi_o^d$  and the zenith angle  $\theta_z$ ,  $f_o$  can be used to estimate the refractive index. To this end, we obtain (1) a rough estimate of zenith angle  $\theta_z$  from the structured light normal  $\mathbf{n}_b$  and (2) the degree of diffuse polarization of exitant light  $\psi_o^d$  from  $\mathbf{H}^d$  using our pBRDF model (Equation (3)):  $\psi_o^d \leftarrow \sqrt{\mathbf{H}_{10}^d{}^2 + \mathbf{H}_{20}^d{}^2} / \mathbf{H}_{00}^d = \frac{\rho T_o^-}{\rho T_o^+}$ . By having both  $\theta_z$  and  $\psi_o^d$ , we formulate an energy function  $E_d$  to estimate the

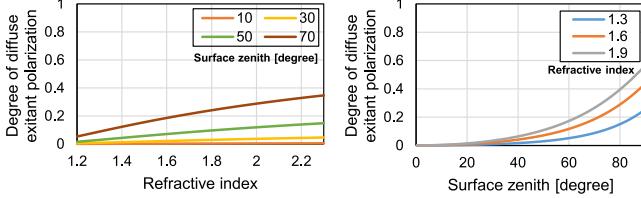


Fig. 6. Degree of diffuse extant polarization  $\psi_o^d$  is related to the refractive index and the zenith angle. (a) For the given zenith angle and with the value of  $\psi_o^d$ , we can estimate the refractive index. (b) We can also estimate the zenith angle for the given refractive index and the value of  $\psi_o^d$ .

$b$ -th material refractive index in the red ( $\eta_1^b$ ), green ( $\eta_2^b$ ) and blue ( $\eta_3^b$ ) channels, respectively:

$$E_d(\eta_{1\dots 3}^b) = \left( \prod_{p_b \in \Omega_b} \sum_{c=1}^3 \left\| f_o(\eta_c^b, \theta_z(p_b)) - \psi_o^d(p_b) \right\|_2^2 \right)^{|\Omega_b|}, \quad (12)$$

where  $p_b$  is a pixel belonging to material  $b$ , and  $\Omega_b$  is denoted as the set of pixels belonging to the material  $b \in \{1, \dots, B\}$ .  $\theta_z(p_b)$  and  $\psi_o^d(p_b)$  denote the values of  $\theta_z$  and  $\psi_o^d$  at the pixel  $p_b$ , respectively. This term  $E_d$  minimizes the difference between the observation and the simulation of the degree of diffuse polarization of exitant light as the geometric mean for statistical robustness.

We then formulate the second term  $E_s$  directly to leverage the observed Fresnel color in the specular shading component  $H_{00}^s$  in Equation (7). The refractive index affects the specularly reflected energy at the surface as the Fresnel reflection coefficient ( $R^+$ ) is determined by the refractive index  $\eta_{1\dots 3}$  and the incident angle  $\theta_d$  to the microfacet surface. To facilitate the dependency of the refractive index on the Fresnel reflection coefficient ( $R^+$ ), we use the specular intensity  $(\mathbf{n} \cdot \mathbf{i})CR^+$  from the first element of the specular shading matrix,  $H_{00}^s = (\mathbf{n} \cdot \mathbf{i})CR^+$ . Our observation is that as  $(\mathbf{n} \cdot \mathbf{i})C$  is independent of the refractive index, i.e., color, we can utilize the ratio of  $R^+$  across different color channels by removing the refractive-index-independent term from  $H_{00}^s$  in Equation (7):  $H_{00,c_1}^s / H_{00,c_2}^s = R_{c_1}^+ / R_{c_2}^+$ , where  $c_1$  and  $c_2$  are the color channels  $c_1, c_2 \in \{1, 2, 3\}$ .  $H_{00,c}^s$  and  $R_c^+$  are the values for each color channel  $c$ . We use the ratio of red and blue color channels with respect to the green channel and formulate the specular term as follows:

$$E_s(\eta_{1\dots 3}^b) = \left( \prod_{p_b \in \Omega_b} \sum_{c \neq 2} \left\| \frac{R_+(\eta_c^b, \theta_d(p_b))}{R_+(\eta_2^b, \theta_d(p_b))} - \frac{H_{00,c}^s(p_b)}{H_{00,2}^s(p_b)} \right\|_2^2 \right)^{|\Omega_b|}, \quad (13)$$

where  $\theta_d(p_b)$  and  $H_{00,\forall}^s(p_b)$  are the values of  $\theta_d$  and  $H_{00,\forall}^s$  at pixel  $p_b$ , respectively. This minimizes the difference of the observed and the reconstructed Fresnel ratio values of different channels.

Figure 7 compares four different estimates of the refractive index of the same material on the abdomen region of a porcelain owl doll (the typical RI of porcelain: 1.50). Figures 7(a) and (b) show the estimated refractive index  $\eta_{1\dots 3}$  and their Fresnel color  $R_+(\eta_{1\dots 3}, \theta_d)$  after minimizing the only incident diffuse polarization energy and the only exitant diffuse polarization energy, respectively. In this figure, for incident polarization, we exploit the diffuse degree of polarization of incident light  $\psi_i^d$  instead of  $\psi_o^d$ :

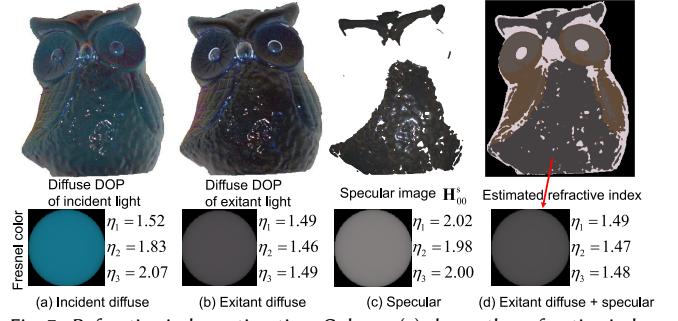


Fig. 7. Refractive index estimation. Column (a) shows the refractive index estimation from the degree of diffuse polarization of *incident* light  $\psi_i^d$  only. Column (b) presents the estimation from the degree of diffuse polarization of *exitant* light  $\psi_o^d$  only. Column (c) is the estimated refractive index from the specular shading component  $H_{00}^s$  only. Column (d) is our jointly optimized estimation of the refractive index from  $\psi_o^d$  and  $H_{00}^s$ .

$\psi_i^d \leftarrow \sqrt{H_{01}^{d,2} + H_{02}^{d,2}} / H_{00}^d = \frac{\rho T_i^-}{\rho T_i^+}$ . However, we found that the incident diffuse polarization is strongly affected by the subsurface scattering effects inside the object, resulting in inaccurate estimates of the refractive index. Therefore, the exact relationship between the degree of incident polarization and the refractive index does not hold. As shown in Figure 7(b), the Fresnel color of the exitant diffuse polarization is a better match with the Fresnel color observed in the specular image  $H_{00}^s$  shown in Figure 7(c). However, the exitant polarization still cannot provide a robust estimate that matches the Fresnel color of  $H_{00}^s$  well. Figure 7(c) presents an estimation with only the specular term, which yields a plausible reproduction of the Fresnel color, similar to the color of the specular shading component. However, the scale of the refractive index cannot be robustly estimated as we only exploit the ratio between the different channels. Figure 7(d) shows our final estimation using joint optimization with the exitant diffuse polarization and the specular shading component, achieving plausible Fresnel color in addition to an accurate scale of the refractive index per color channel.

*Specular roughness and coefficient.* To estimate both specular roughness  $\sigma$  and coefficient  $k_s$ , we require the information of the refractive index  $\eta_{1\dots 3}$  and high-quality surface normals  $\mathbf{n}$ . The input rough geometry information  $\mathbf{n}_b$  is insufficient for estimating  $\sigma$  and  $k_s$  such that we need to obtain surface normals  $\mathbf{n}$  from *diffuse polarization*, as described in Section 5.4. Once we have  $\eta_{1\dots 3}$  and  $\mathbf{n}$ , we solve the following optimization problem to determine the parameters to make our specular reflection model close to the observation:

$$\underset{\sigma^b, k_s^b}{\text{minimize}} \sum_{p_b \in \Omega_b} \left\| (W'(p_b) + W(p_b)) \left( \hat{H}_{00}^s(p_b; \sigma^b, k_s^b) - H_{00}^s(p_b) \right) \right\|_2^2, \quad (14)$$

where  $W(p_b)$  and  $W'(p_b)$  are the two different confidence values of a pixel  $p_b$ .  $W$  penalizes pixels, of which normal is close to the halfway vector:  $W = 1 - (\mathbf{n} \cdot \mathbf{h})^\lambda$ .  $W'$  weights pixels with strong specular only, of which normal is close to the halfway vector based on the reliable low-frequency normal on that region:  $W' = \left( H_{00}^s / |H_{00}^s| \right)^\lambda$ . We set  $\lambda=10$  empirically. Here  $\hat{H}_{00}^s(p_b; \sigma^b, k_s^b)$  is the reconstructed polarimetric shading matrix using Equation (7) at the pixel  $p_b$  for the given  $\sigma^b$  and  $k_s^b$ .  $H_{00}^s(p_b)$  is the value of  $H_{00}^s$  at the pixel  $p_b$ .

The reason for having these two weights  $W$  and  $W'$  is that our initial estimation of surface normals mainly relies on the degree of diffuse polarization. However, the degree of polarization is severely weak in those regions, where the zenith angle is close to zero [Atkinson and Hancock 2006]. As Equation (14) is a non-linear optimization problem, we solve it using the interior-point method [Waltz et al. 2006]. Note that the estimated roughness  $\sigma$  is used to further refine the surface normals  $\mathbf{n}$  in the next iteration using specular information. Refer to Section 5.4 for more details on estimating normals. Figure 8 shows the estimated roughness in the form of pBRDF for each material.

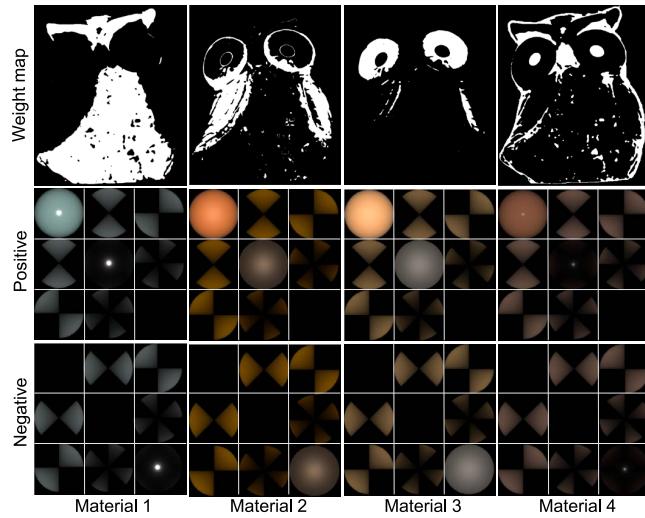


Fig. 8. Visualization of estimated pBRDF  $P$  per material (top: weight maps for each material) in the form of a sphere geometry. Note that different gamma values are applied to the diagonal ( $\gamma=2.2$ ) and non-diagonal elements ( $\gamma=4$ ) for visualization. Per-pixel diffuse albedos  $\rho$  are averaged for every pixel.

**Diffuse albedo.** The estimation of the diffuse albedo is the last step of our inverse rendering pipeline. We estimate the diffuse albedo from the polarimetric diffuse shading  $H_{00}^d$  and final surface normals  $\mathbf{n}$  (refer to Section 5.4) as follows:  $\rho = H_{00}^d / \{(\mathbf{n} \cdot \mathbf{i}) T_o^+ T_i^+\}$ , where  $T_o^+ T_i^+$  can be calculated using Equation (4) with the per-channel refractive index  $\eta_{1..3}$  as estimated (Equation (11)) to account for the diffuse Fresnel effect. Figure 9 compares polarimetric diffuse shading  $H_{00}^d$  and estimated diffuse albedo  $\rho$  with the intermediate diffuse shading  $\mathbf{n} \cdot \mathbf{i}$  and diffuse Fresnel coefficients  $T_o^+ T_i^+$ .

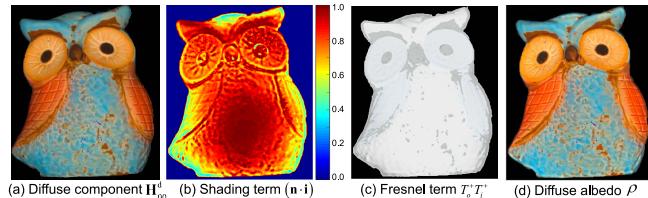


Fig. 9. (a) Polarimetric diffuse shading  $H_{00}^d$  per pixel. (b) Shading term  $\mathbf{n} \cdot \mathbf{i}$ . (c) Fresnel coefficients  $T_o^+ T_i^+$ . (d) Estimated diffuse albedo  $\rho$ .

#### 5.4 Estimating normals from polarization

**Overview.** Our polarization-based normal estimation is a two-step process. Given the estimated refractive index  $\eta$  and rough surface

normals  $\mathbf{n}_b$ , the first step estimates normals  $\mathbf{n}$  from the polarimetric diffuse shading component  $H_{00}^d$ . After estimating both the specular roughness  $\sigma$  and the coefficient  $k_s$  using the intermediate normal  $\mathbf{n}$ , the second step updates normals  $\mathbf{n}$  (estimated from diffuse polarization) in regions where the zenith angle is close to zero.

**Diffuse normals.** Existing SfP methods have two common ambiguities pertaining to the azimuth and zenith angles of estimated normals  $\mathbf{n}$ ; the azimuth angles of normals could be flipped with  $\pi$ , and the zenith angles could be overly flat. Existing SfP method solve the azimuth ambiguity problem, existing SfP methods have made efforts to solve it in various ways. Tozza et al. [2017] use multiple lights to capture multi-directional observations to resolve the ambiguity. Miyajaki et al. [2003] proposed a greedy approach, searching azimuth candidates exhaustively from the boundary to the center. Kadambi et al. [2015] estimate a binary selection mask to select one of the two candidate normals using a rough base normal from a depth sensor. However, these methods either require multiple light sources or employ a binary choice between two normal candidates, thereby resulting in notable artifacts. In contrast, we resolve the azimuth ambiguity with a single optimization of the surface normals by means of a rough base normal, instead of explicitly estimating the binary mask.

For the zenith ambiguity, previous SfP methods cannot determine the material property as the refractive index and therefore they assume a certain refractive index (such as 1.5) manually. Without knowing the accurate refractive index of the object, they encounter the zenith ambiguity. In contrast, our novel joint optimization of appearance and normals allows us to solve the zenith ambiguity problem with an accurately estimated refractive index.

To estimate high-frequency normals  $\mathbf{n}$  from the polarimetric diffuse shading matrix  $H^d$  and rough geometric normals  $\mathbf{n}_b$  from structured lighting, we formulate a joint optimization problem that consists of three terms:

$$\begin{aligned} & \underset{\mathbf{N}}{\text{minimize}} \quad \alpha_d \left\{ \|\mathbf{W}(\mathbf{ON} - \mathbf{C}_z)\|_2^2 + \|\mathbf{WAN}\|_2^2 \right\} \\ & \quad + \beta_d \|\mathbf{GN} - \mathbf{N}_b\|_2^2 + \gamma_d \|\nabla \mathbf{N}\|_2^2, \end{aligned} \quad (15)$$

where  $\mathbf{N}$  and  $\mathbf{N}_b$  are the normals  $\mathbf{n}$  and  $\mathbf{n}_b$  concatenated for every pixel in a matrix form.  $\mathbf{W}$  is a matrix of the confidence function  $W$  for handling unconfident observation of diffuse polarization used in Equation (14),  $\mathbf{O}$  is a matrix of view vectors  $\mathbf{o}$  for every pixel,  $\mathbf{C}_z$  is a matrix of the cosine values of the zenith angles,  $\cos(\theta_z)$  for every pixel,  $\mathbf{A}$  is a matrix for collinearity of azimuth cues  $\mathbf{A} = [\cos \phi_o, -\sin \phi_o, 0]$  for every pixel,  $\mathbf{G}$  is a matrix for Gaussian blurring with standard deviation (as 10 in our experiment), and  $\alpha_d$ ,  $\beta_d$  and  $\gamma_d$  are corresponding weights, set as 10, 1 and 3, respectively. Note that we pack per-pixel normals  $\mathbf{n}$  as  $\mathbf{N}$  in a matrix form to formulate the optimization problem.

In Equation (15), the first term forces the zenith angle of surface normals  $\mathbf{n}$  to be similar to the zenith angle from polarization  $\theta_z$ , which can be calculated from the refractive index  $\eta$  and the degree of exitant polarization  $\psi_o$ , similar to the case of the refractive index estimation method (Section 5.3) as shown in Figure 6(b):  $\theta_z = f^{-1}(\psi_o; \eta)$ . The second term makes the normal close to either the two different azimuth angles from polarization  $\phi_o$  or  $(\phi_o + \pi)$

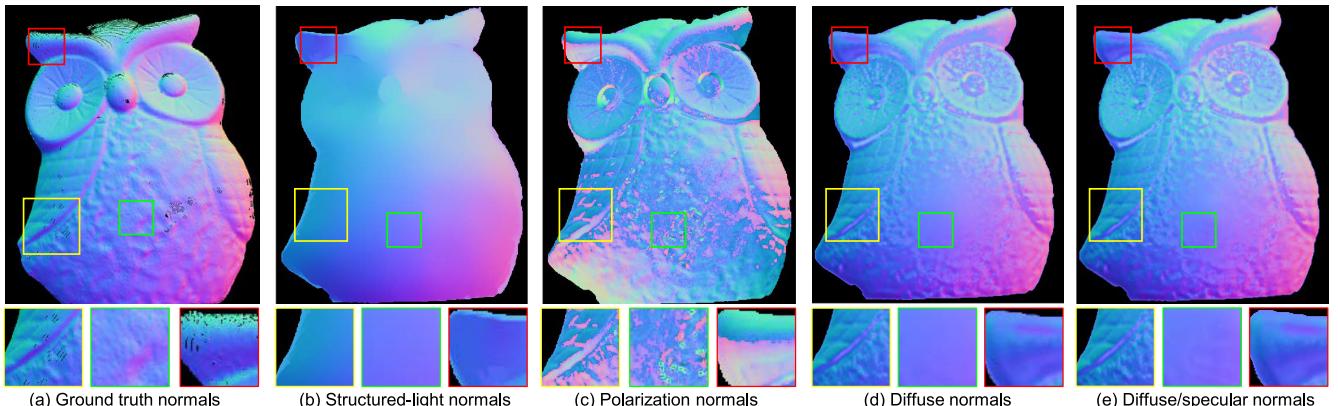


Fig. 10. Normal Estimation. Image (a) shows reference normals scanned by a 3D laser scanner (NextEngine). Structured-light normals (b) provide reliable low-frequency information of surface normals, while polarization normals (c) retain high-frequency detail but suffer from azimuthal ambiguity and noise. Image (d) shows our intermediate normal estimation result with diffuse polarization only, still missing details in the region where the zenith angle close to zero. Image (e) presents our final result with enhanced normals in the region with additional consideration of specular reflection.

with the co-linearity constraint, following [Tozza et al. 2017]. Here the azimuth angle  $\phi_o$  can be estimated from diffuse exitant polarization:

$$\tan^{-1} \left( \frac{H_{20}^d}{H_{10}^d} \right) = -\tan^{-1} \left( \frac{\sin(2\phi_o)}{\cos(2\phi_o)} \right) = -2\phi_o. \quad (16)$$

The third term forces the low-frequency components of the reconstructed normal  $\mathbf{n}$  to be similar to rough geometric normals  $\mathbf{n}_b$  from structured lighting. The last term is the local smoothness prior for the reconstructed normals  $\mathbf{n}$ . As every term is imposed in the L2 norm, we can efficiently optimize Equation (15) using the gradient decent method (refer to the supplemental material for details).

**Specular normals.** Different from existing SfP methods [Kadambi et al. 2015; Miyazaki et al. 2003; Tozza et al. 2017], our method does not require the forced choice of the modality between diffuse and specular reflection on each pixel with binary labelling. Despite that our polarimetric decomposition of mixed specular and diffuse polarization can handle the modality problem, we still require a variant approach to estimate surface normals from specular reflection in the regions where the zenith angle becomes close to zero.

In the second step, we update surface normals  $\mathbf{n}$  with additional information about the specular roughness  $\sigma$  and coefficient  $k_s$  (Section 5.3) by formulating an optimization problem as follows:

$$\begin{aligned} & \underset{\mathbf{N}}{\text{minimize}} \quad \alpha_s \|\mathbf{W}(\Psi\mathbf{N} - \mathbf{C}_h)\|_2^2 + \beta_s \|\mathbf{W}\nabla\mathbf{N}\|_2^2 \\ & \quad + \gamma_s \|(1 - \mathbf{W})(\mathbf{N} - \mathbf{N}')\|_2^2, \end{aligned} \quad (17)$$

where  $\Psi$  is a matrix of halfway vectors for every pixel,  $\mathbf{C}_h$  is a matrix of cosine values of the halfway angles,  $\cos(\theta_h)$ , concatenated for every pixel,  $\theta_h$  is the halfway angle between the normal  $\mathbf{n}$  and halfway vector  $\mathbf{h}$ , and  $\mathbf{N}'$  is a matrix of the previously estimated normals  $\mathbf{n}'$  using diffuse polarization in the first step. We also use the weighting matrix  $\mathbf{W}$  introduced in Equation (15) to identify low confidence regions, and the weights  $\alpha_s$ ,  $\beta_s$  and  $\gamma_s$  are set as 10, 1 and 1, respectively.

Here, similar to photometric stereo with specular reflection [Fyffe et al. 2016], the first term forces the halfway angle between the

reconstructed normal  $\mathbf{n}$  and the halfway vector  $\mathbf{h}$  to be close to the halfway angle  $\theta_h$  that can be estimated from specular observations in  $\mathbf{H}_{00}^s$ . To do so, we derive a function  $S(\theta_h; \sigma)$  that can describe the halfway-angle dependency in  $\mathbf{H}_{00}^s$  using the (0,0) element of the polarimetric shading matrix (Equation (7)) as follows:

$$\begin{aligned} S(\theta_h; \sigma) &= \frac{D(\theta_h; \sigma)}{4 \cos(\theta_h)} \\ &\approx (\mathbf{n} \cdot \mathbf{i}) \frac{D(\theta_h; \sigma)}{4(\mathbf{n} \cdot \mathbf{i})(\mathbf{n} \cdot \mathbf{o})} \leftarrow \frac{\mathbf{H}_{00}^s}{k_s G(\mathbf{i}, \mathbf{o}; \sigma) R^+(\mathbf{i}, \mathbf{o}; \eta)}. \end{aligned} \quad (18)$$

Using the known values of  $\eta$ ,  $k_s$  and  $\sigma$ , we first detach the influence of specular albedo  $k_s$ , the geometric term  $G(\mathbf{i}, \mathbf{o}; \sigma)$ , and the Fresnel term  $R^+(\theta_d; \eta)$  from the polarimetric specular shading  $\mathbf{H}_{00}^s$ , resulting in the NDF term  $(\mathbf{n} \cdot \mathbf{i}) D(\theta_h; \sigma)$  divided by  $4(\mathbf{n} \cdot \mathbf{i})(\mathbf{n} \cdot \mathbf{o})$ . Function  $S$  can be simplified as Equation (18) with approximation based on the coaxial setup ( $(\mathbf{n} \cdot \mathbf{i}) \approx (\mathbf{n} \cdot \mathbf{o}) \approx \cos(\theta_h)$ ).

To this end, we can observe the characteristics of the function  $S$  as shown in Figure 11: First, since  $S$  does not decrease continuously, there can be two different solutions of  $\theta_h$ , which produce the same value of  $S$ , while the first solution is close to zero and the other is close to  $\pi/2$ . However, as our goal is to utilize specular information in regions where the surface zenith angle is close to zero, we can take the solution with a smaller  $\theta_h$  value to resolve the ambiguity. Second,  $S$  is smoothly changing with respect to  $\theta_h$ . Hence, we estimate  $\theta_h$  using a lookup table approach  $S(\theta_h; \sigma)$  with linear interpolation to solve the following optimization problem  $\underset{\theta_h}{\text{minimize}} \|\hat{S}(\theta_h; \sigma) - S\|$ , where  $\hat{S}$  is the reconstructed value of the function  $S$ . Finally, we optimize Equation (17) using gradient descent with an iterative process of optimization and normalization with the unit norm constraint, similarly to Equation (15).

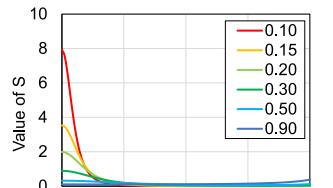


Fig. 11. Function  $S$  with respect to  $\theta_h$  for selected specular roughness  $\sigma$ .

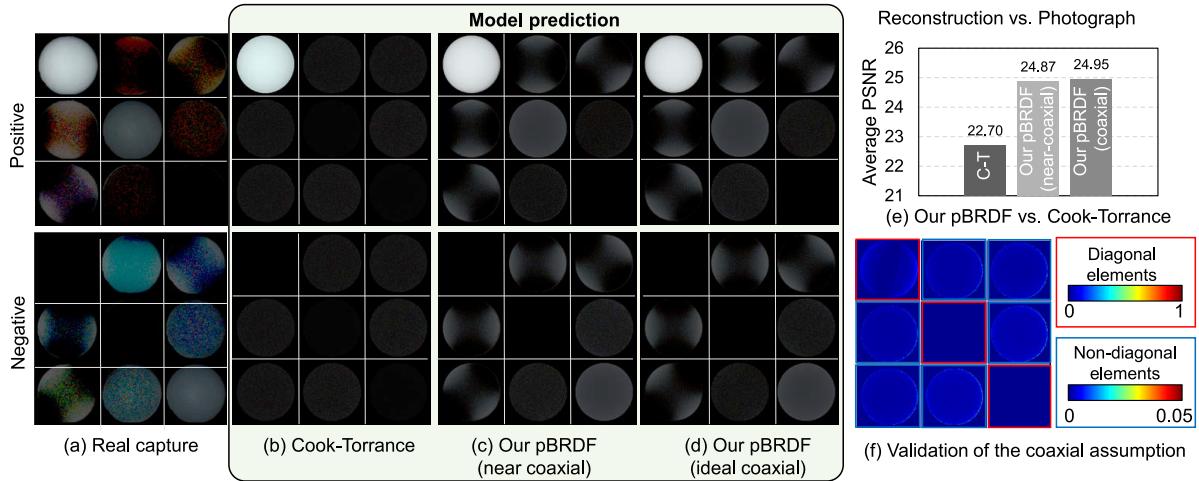


Fig. 12. Validation of our pBRDF model. (a) We captured a white sphere object and reconstructed the polarimetric shading matrix  $\mathbf{H}$  of the object. We find the best-fit parameters and reconstruct  $\mathbf{H}$  with two models: (b) the traditional C-T model [Cook and Torrance 1982] and (c) our pBRDF model. (e) Our pBRDF model can predict the decomposed polarimetric photographs better than the C-T model does both with and without coaxial assumption. In addition, we verify our coaxial assumption by comparing (c)  $\mathbf{H}$  under the near-coaxial setup and (d)  $\mathbf{H}$  under the ideal coaxial setup. They both present a good agreement in predicting polarimetric shading. (f) shows absolute differences of the estimated  $\mathbf{H}$  under the two different setups.

Figures 10(a) and (b) show the inputs, rough base normals and polarization normals with ambiguity. Our diffuse-only optimization yields high-quality normals except for the region where the zenith angle becomes close to zero, as shown in Figure 10(c). Figure 10(d) shows that our complete optimization with additional specular cues can recover surface normal details over the missing regions.

## 6 RESULTS

### 6.1 Experimental Setup

We built an experimental setup with an ordinary RGB machine-vision camera (FLIR Grasshopper GS3-U3-120S6C-C, 12MP) equipped with a 35mm lens and an unpolarized DLP projector (AAXA P450 Pro, 450 lumen), as shown in Figure 5, where the maximum intensity is calibrated using a stainless steel ball. The white balance factors are obtained by capturing a standard Spectralon tile with a known reflectance. The colors of the camera are calibrated as a 3-by-3 matrix by capturing the standard ColorChecker. The geometric relationship between the camera and the projector is calibrated using Zhang’s method [2000], yielding four vectors,  $\mathbf{y}_i$ ,  $\mathbf{y}_o$ ,  $\mathbf{i}$ , and  $\mathbf{o}$ , which are related to the extrinsic parameters of the camera and the projector. We installed two linear polarization filters in front of the camera and the projector. They are mounted firmly with 3D-printed supporting structures. The polarization angles,  $\vartheta_i$  and  $\vartheta_o$ , are controlled manually with four different measurements for each angle from 0 to  $\pi/2$  with a step size of  $\pi/6$ . We took polarimetric photographs with multiple exposures by varying the shutter time to convert raw signals into HDR radiance [Debevec and Malik 1997] to capture specular reflection without saturation.

### 6.2 Validation of the Polarimetric Reflectance Model

For validation of our pBRDF model, we captured a white sphere object and estimated the polarimetric shading matrix  $\mathbf{H}$  of the object using Equation (10), as shown in Figure 12(a). We then compare this

with the reconstructed matrix  $\mathbf{H}$ , with best-fit parameters using our pBRDF model shown in Figure 12(c). We also reconstruct  $\mathbf{H}$  using a traditional C-T BRDF model [Cook and Torrance 1982] shown in Figure 12(b). During the process of the reconstructing  $\mathbf{H}$ , we add Gaussian noise to account for noise in the captured images. Figures 12(a) and 12(c) validate that our model can characterize polarimetric reflection on the sphere object better than the non-polarimetric model in Figure 12(b). For a quantitative evaluation, we reconstruct polarimetric photographs (a) using the C-T model (b) and our pBRDF model (c). The positive and negative elements of the  $\mathbf{H}$  matrix are combined and considered as per-pixel single signed scalars. For both HDR photographs and  $\mathbf{H}$  matrices, we computed PSNR values with the peak signal value normalized to one. As shown in Figure 12(e), our model accuracy is 24.87 dB of PSNR, while the C-T model accuracy is 22.70 dB. Since our model accounts for polarimetric appearance changes, the prediction accuracy of our model is higher than that of the traditional C-T model accordingly.

### 6.3 Validation of the Coaxial Assumption

We evaluate our coaxial assumption by reconstructing a polarimetric shading matrix  $\mathbf{H}$  under the synthetic coaxial setup shown in Figure 12(d), and comparing it to the reconstruction under our real near-coaxial setup in Figure 12(c). Figure 12(f) visualizes differences between these two  $\mathbf{H}$  matrices. Numeric evaluation in terms of two metrics (average intensity difference & PSNR) shows the validity of the assumption: (0.0290 & 41.99dB) and (0.0062 & 52.46dB) for the diagonal elements and the non-diagonal elements, respectively.

### 6.4 Appearance and Normal Estimation

*Synthetic data.* We validate our algorithm on synthetic data, as shown in Figure 13. For creating synthetic input data of surface normals  $\mathbf{n}$ , we employ the Stanford dragon object. Material parameters are set as follows:  $k_s = 3.00$ ,  $\eta_{1\dots3} = \{1.50, 1.70, 1.90\}$ ,  $\sigma = 0.20$  to

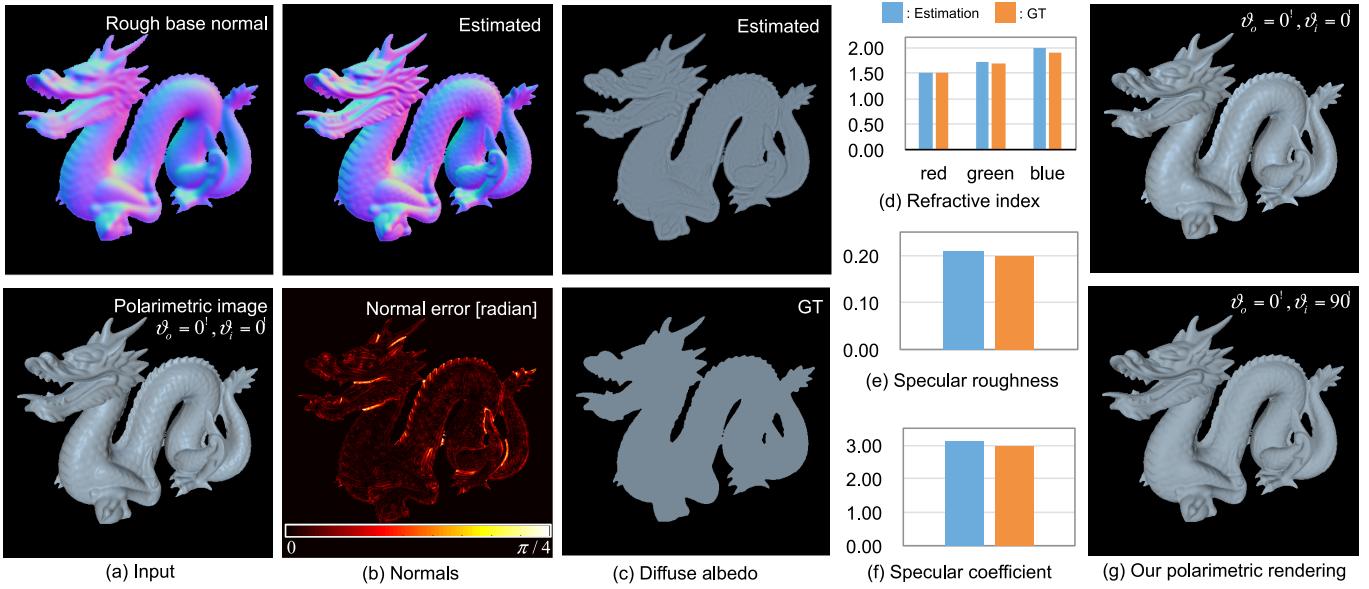


Fig. 13. Synthetic evaluation. (a) We simulate polarimetric images of a dragon object using Equations (3) and (7) for input of our method. Extrinsic parameters of our hardware setup were employed. Also, we added Gaussian noise to the images. (b) – (f) Our estimation of polarimetric appearance is very close to the ground truth with high accuracy. (g) Our method can faithfully reconstruct the original polarimetric images with the estimated pBRDF.

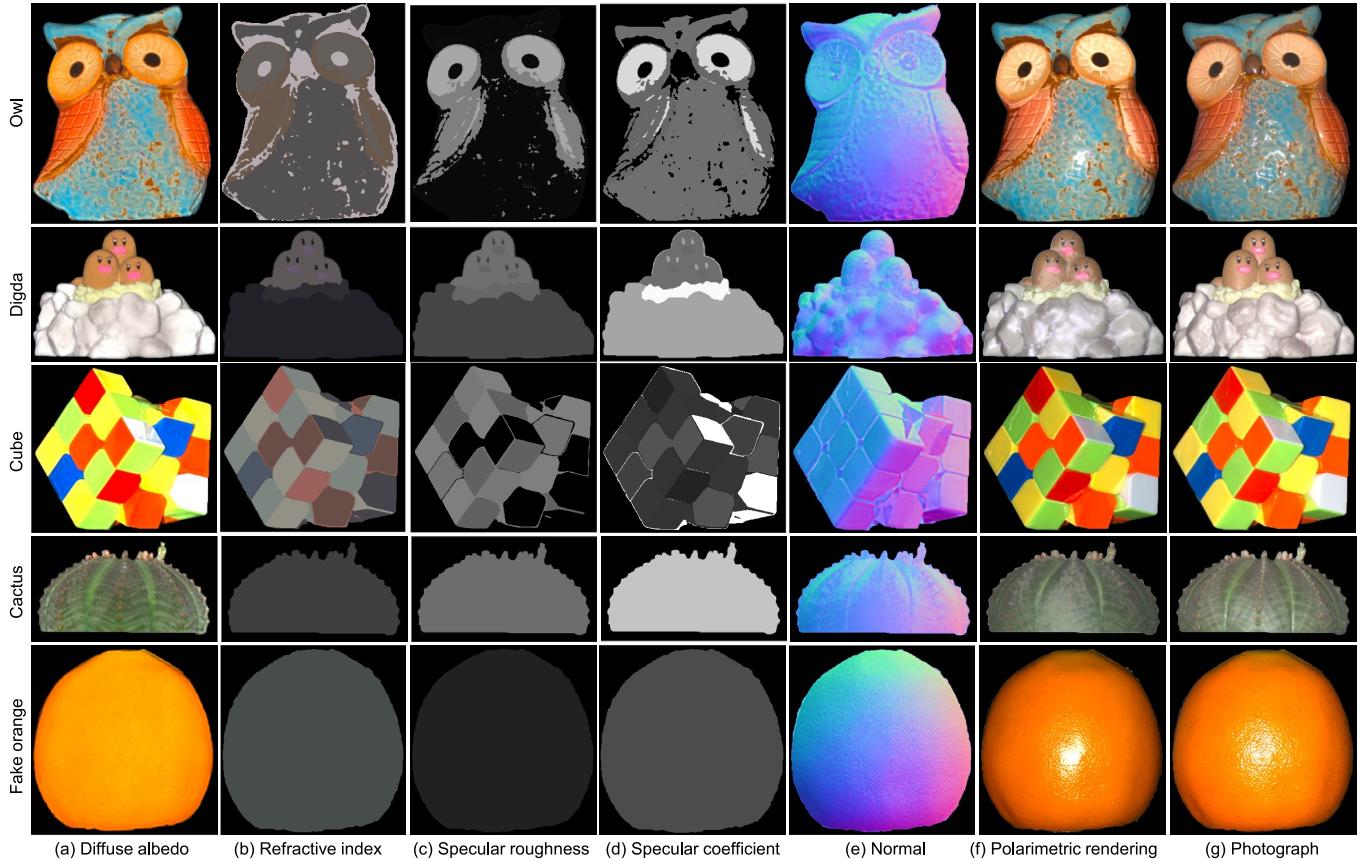


Fig. 14. Our polarimetric inverse rendering recovers (a)–(d) the surface appearance, and (e) surface normals from a single view and a single projector setup. (f) Thanks to our pBRDF model, we can provide polarization-aware rendering where the camera and the light source are both equipped with linear polarizers, with angles  $\vartheta_i = 0$  and  $\vartheta_o = 0$ . (g) Our rendering results show good agreement with ground truth photographs.

have both specular and diffuse reflection. Figure 13(a) shows the input data for our method. We create rough surface normals  $\mathbf{n}_b$  by convolving  $\mathbf{n}$  with a Gaussian kernel with standard deviation 10. Input polarimetric images are then generated using these appearance parameters and the polarimetric image formation model (Equation (9)). Note that real extrinsic parameters from a prototype are used in this simulation. Figure 13(b) shows that we significantly improve details of surface normals, compared to input normals. However, we can still observe some deviation from the ground truth near regions where normals contain high frequency details. Figure 13(c) demonstrates that diffuse albedo can be faithfully estimated except for regions where surface normals contain errors. Figures 13(d)–(f) show estimated material properties of the refractive index, specular roughness and specular coefficient, which are very close to the ground truth values, as follows: (Est. & GT):  $k_s = (3.14 \& 3.00)$ ,  $\eta_{1..3} = (\{1.51, 1.72, 1.98\} \& \{1.50, 1.70, 2.00\})$ ,  $\sigma = (0.21 \& 0.20)$ . With these estimated properties, we can faithfully reconstruct polarimetric images, such as cross-polarization, as shown in Figure 13.

*Real data.* We captured real-world objects to obtain polarimetric reflectance and surface normals (Figure 14). Our method successfully reconstructs detailed surface normals while estimating the appearance properties by means of polarization. These properties are then used for polarimetric rendering using our pBRDF model (Equations (3) and (7)). Figures 14(f) and (g) show our polarimetric inverse rendering and reference photographs captured with the linear polarizer in front of the camera and light with polarization angles of  $\vartheta_i = 0$  and  $\vartheta_o = 0$ . These results validate that our model can predict polarimetric reflection with high accuracy.

*Refractive index.* We evaluate the accuracy of the refractive index estimated by our method. We captured a white sphere object painted with white acrylic matching liquid, the refractive index of which is known as the ground truth [Kasarova et al. 2007]. Figure 15 validates that we can estimate per-channel RI of the paint liquid with high accuracy.

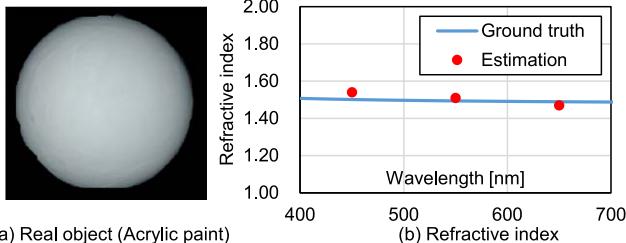


Fig. 15. Refractive index validation of a sphere painted with white acrylic matching liquid. The estimated refractive indices of the red/green/blue channels are 1.47/1.51/1.54 close to the ground truth [Kasarova et al. 2007].

*Surface normals.* Miyazaki et al. [2003] assume a single refractive index over entire surfaces and integrate polarization normals from the boundary to handle the azimuth ambiguity. Kadambi et al. [2015] resolved the azimuth and zenith ambiguity problems of polarization normals using a Boolean mask and a distortion-correction stage using rough base normals. Contrary to both existing methods, we resolve the angle ambiguity problem by estimating the refractive index and solving a single optimization (Equation (15)) that takes

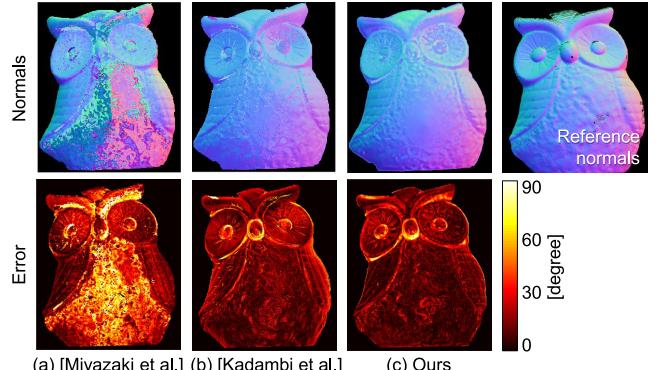


Fig. 16. Evaluation of surface normals. The top row shows surface normals of (a) Miyazaki et al., (b) Kadambi et al., (c) our method and (d) the reference 3D scanner. The bottom row presents error maps of the normals for each method in terms of angle difference, compared to the ground-truth normals.

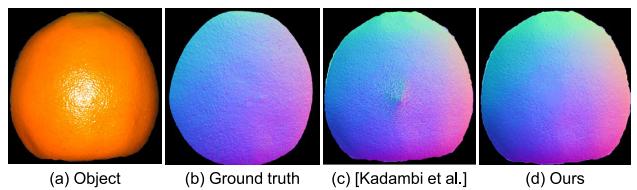


Fig. 17. Qualitative evaluation of surface normals. (b) Ground-truth normals of a plastic ball object (a) obtained from a laser scanner. (c) Normals from Kadambi et al. [2015] without a depth integration stage. (d) Our estimated normals shows fewer artifacts, compared to (c).

normal cues from diffuse polarization as well as rough base normals. For normal evaluation, we captured the 3D geometry of the owl object using a laser 3D scanner (NextEngine), which is then mapped to the screen space via correspondence mapping to obtain a reference normal map (Figure 16).

Compared to the reference normals, we calculate the accuracy of our estimated normals and those from Miyazaki et al. [2003], and Kadambi et al. [2015]. Our method outperforms the others as shown in Figure 16. The averaged normal differences with the reference normals are 10.05, 11.95, and 22.70 in degree for our method, [Miyazaki et al. 2003], and [Kadambi et al. 2015]. Figure 17 shows additional evaluation on a plastic orange toy, qualitatively demonstrating that our method results in fewer artifacts compared to Kadambi et al. [2015].

*Inverse rendering.* Miyazaki et al. [2003] proposed the use of polarization normals by integrating normals from boundary. However, they use the non-polarimetric Torrance-Sparrow BRDF model [1967] ignoring polarization for appearance estimation. Figures 18(a) and (c) validate that our method outperforms Miyazaki et al. [2003] that use the simple integration of normals. In addition, we implemented another inverse-rendering method that takes the rough structured light normal instead of the polarization normal where the appearance parameters are estimated based on non-linear fitting with the C-T model. Figures 1 and 18(b) validate that our inverse rendering can produce very realistic appearance. Refer to the supplemental video for rendering in novel view and novel light.

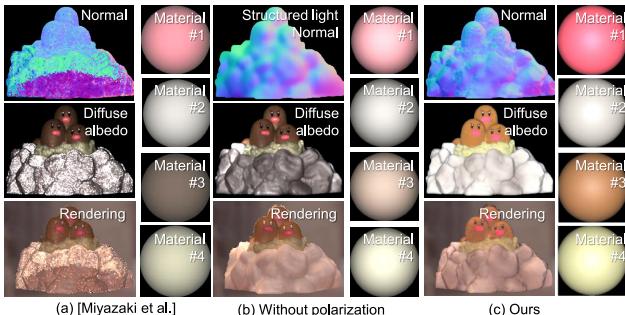


Fig. 18. Inverse rendering results for the digda scene shown in Figure 14. (a) [Miyazaki et al. 2003], (b) [Miyazaki et al. 2003] without polarization, and (c) ours. The first row shows estimated normals of each method. Note that we use the structured light normals for (b). The second row shows the estimated diffuse albedo. The third row presents rendered images of each method under the environmental lighting of St. Peter.

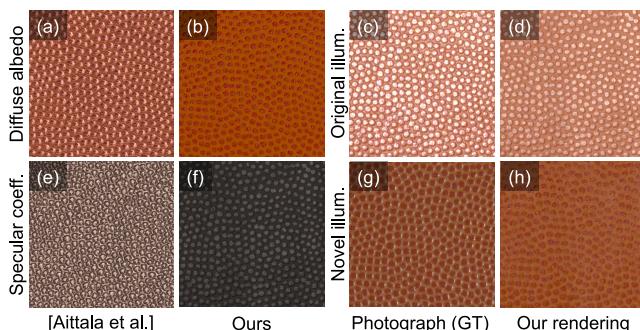


Fig. 19. Inverse-rendering results compared to Aittala et al. [2015] and rendering under novel illumination. (a) & (e) Estimated diffuse and specular albedo from Aittala et al. (b) & (f) Estimated diffuse and specular albedo from our inverse-rendering method. (c) & (d) and (g) & (h) Photographs and our rendering results under the original and the novel illumination.

We also captured a flat object with stationary patterns and compare our estimation of diffuse and specular albedo with those from Aittala et al. [2015]. Figure 19 validates that our method effectively estimates diffuse albedo using polarization, different from Aittala et al [2015]. However, since we estimate per-material parameters of specular appearance, our results present sharp edges in results.

In addition, we evaluate performance of inverse rendering under a novel illumination. We placed a light source in a different direction at a slanted degree with respect to the flat object. Figures 19(c) & (d) and (g) & (h) show photographs and rendering results under the original illumination and the novel illumination. PSNRs of the original and the novel illumination are 26.12 dB and 22.25 dB, respectively.

## 7 DISCUSSION AND FUTURE WORK

*Multiple reflection among microfacets.* According to the microfacet theory [Torrance and Sparrow 1967], diffuse reflection is believed to originate from two phenomena: subsurface scattering and multiple reflections among microfacets. The polarization behaviors of these two cases are different due to the interface traversal characteristics. As the proportion between these two factors varies significantly depending on the materials, discriminating these factors in diffuse reflection is still an open problem. However, as noted by Atkinson and Hancock [2006], the impact of multiple reflection

among facets is negligible for smooth surfaces in general. Similar to existing SfP methods [Atkinson and Hancock 2006; Kadambi et al. 2015], we assume that diffuse reflection is dominated by subsurface scattering, where the light is absorbed into the object, scattered by molecules of object materials, after which it exits the object. It would be an interesting future work to investigate a new polarimetric model that can account for multiple reflection among microfacets.

*Circular polarization.* We demonstrate the simultaneous acquisition of appearance and surface normals by analyzing linearly polarized components. Because our experimental setup requires the angle combinations of the incident and exitant polarization filters in front of the camera and the projector, the number of filters that can be attached to the system is limited to two. In addition, as the circular polarization consists of only the specular component in our reflectance model, we opt to use linear polarization only to estimate appearance parameters and normals by analyzing both diffuse and specular reflection. However, we still believe that analyzing the circular polarization may reveal hidden components of light transport, which can be explored in future work.

*Spatially-varying BRDF.* We chose the common SVBRDF acquisition approach capturing the refractive index, specular roughness and specular coefficient per material, rather than per pixel, for the following reasons: First, our inverse-rendering method assumes that the linear polarizer is perfect, which may not function identically to the linear polarization equation [Ghosh et al. 2010]. Second, we assume an ideal coaxial setup, although it is not perfect in reality, in order to estimate appearance and normal parameters using our inverse-rendering method. Lastly, the view and the light vector in our setup are fixed, sparsely capturing the specular highlights on the object. We believe that a more sophisticated hardware setup will resolve these issues, thus enabling per-pixel appearance and normal acquisition.

*Metallic surfaces.* Following Ghosh et al. [2010], we assume that the captured subjects consist of dielectric material only in order to avoid the mathematical complexity of the refractive index as a complex number. It would be interesting future work to extend our model to metallic surfaces.

*Metallic surfaces.* As mentioned by Ghosh et al. [2008], single scattering tends to maintain the polarization state of the incident light. This breaks the assumption of our pBRDF model that the absorbed light into an object is completely depolarized. It would be interesting future work to develop a pBRDF model that can take single scattering into account. Furthermore, we found that single/shallow scattering may be related to incident polarization as shown in Figure 7(a).

*Using a projector.* Different from traditional photometric stereo methods [Nam and Kim 2014], we employ a single unpolarized projector as a light source and equip it with a linear polarization filter for two main reasons. First, the projector enables us to measure rough geometry via structured lighting, which is used for estimating appearance from polarization. Second, rough geometry can be used for solving the traditional azimuthal ambiguity problem in

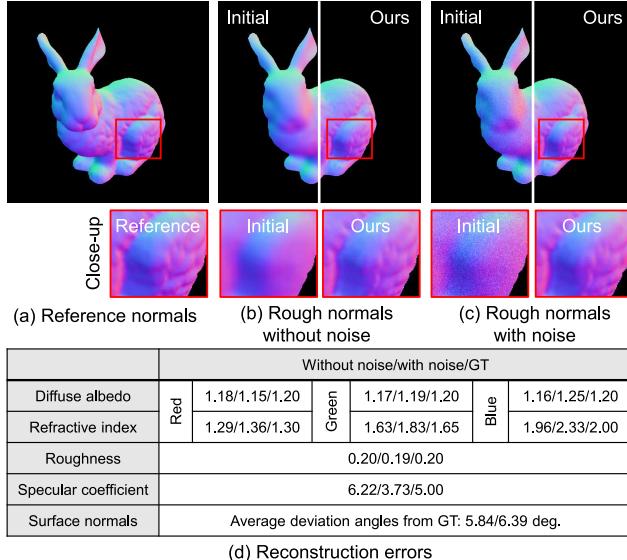


Fig. 20. Impact of initial rough normals. (a) We simulate a bunny object as input of our method. (b) Rough normals are computed by applying a Gaussian kernel, and (c) further degradation is conducted with noise. (d) Our inverse-rendering method handles noise robustly in the initial normals with subtle degradation.

SfP so that we can estimate high-quality normals from polarization. Recently, Baek et al. [2016] proposed a birefractive stereo method that uses polarization-dependent material to estimate scene depth. It would be an interesting avenue to employ birefractive stereo to remove the dependency on the projector by replacing the rough geometry from a projector with that from birefractive stereo.

*Off-center pixels.* We assume that the linear polarizer performs consistently regardless of the direction of the incident light. Korger et al. [2013] validate that conventional linear polarizers perform stably up to an incident angle of 45°. Given the long-side field-of-view of our system (21.36°), the linear polarization effects on each pixel occur consistently regardless of perspective projection.

*Impact of initial normals.* Our method takes rough normals from structured light as input to overcome the azimuthal ambiguity of polarization normals. However, our near-coaxial setup introduces noise on initial normals due to the short baseline. We designed our framework to mitigate the impact of this by utilizing multiple pixels with clustering and introducing priors for surface normals such as the fourth term in Equation (15). To validate our method, we conducted an experiment by simulating a bunny object with known appearance and normals as shown in Figure 20. Rough normals are computed by applying a Gaussian blur kernel of standard deviation 30. We also added Gaussian noise with variance of 0.01 on the rough normals, thereby making corrupted rough normals. Figure 20(d) shows that our method can handle minor noise.

*Number of materials.* We investigate the impact of the number of materials on our inverse-rendering method for a flat object. Figure 21 shows rendered images with different numbers of materials from one to three. As the object originally consists of two different materials, one cluster is not enough to faithfully reproduce the appearance

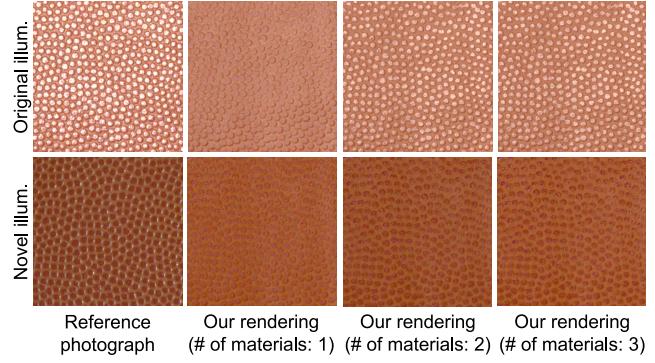


Fig. 21. We render a flat leather object with different numbers of clusters under the original illumination (top row) and a novel illumination (bottom row). When the number of clusters is higher than the number of original materials, two in this example, our method faithfully reproduces appearance under both illumination conditions.

of the object. Using two materials leads to plausible reconstruction of appearance. The use of three materials does not improve the results further. Visual similarity of the PSNR between the rendered images and the photograph is 25.22, 26.12, and 26.13 dB when we use one, two and three materials, respectively, under the original illumination. PSNR values drop for the novel illumination as 21.96, 22.25, and 22.20 dB for each.

*Tradeoff between diffuse and specular reflection.* Our compact setup of a projector and a camera introduces a tradeoff between specular and diffuse reflection. In specular dominant regions, the diffuse polarization component is relatively weak and consequently the surface normals from polarization often show suboptimal accuracy. In rough surface areas, the specular component is faint and hence it is difficult to estimate specular-dependent parameters (such as specular coefficient and roughness) with high accuracy. To mitigate this limitation, our method relies on the spatial coherence of material properties based on the clustering information. However, our method is not free from this limitation, and further investigation is necessary in future work.

## 8 CONCLUSION

We have presented a novel, complete pBRDF model that can define the mixed polarization of both diffuse and specular reflection. We also propose a novel inverse-rendering method based on our pBRDF model to estimate spatially varying appearance and normals simultaneously through joint optimization of the polarimetric appearance and geometric properties. We validate the accuracy of our pBRDF model and the results of inverse rendering on both synthetic and real data. However, extending our model toward polarimetric multiple scattering in microfacets and metallic surfaces reflections remains as future work.

Our setup requires only a pair of a camera and a projector and thus the system form factor could be optimized compactly. We anticipate that its applicability could be widened with modern polarization cameras.

## ACKNOWLEDGMENTS

Min H. Kim acknowledges Korea NRF grants (2016R1A2B2013031, 2013M3A6A6073718) and additional support by Microsoft Research Asia, KOCCA in MCST of Korea, Cross-Ministry Giga KOREA Project (GK17P0200), Samsung Electronics (SRFC-IT1402-02), and an ICT R&D program of MSIT/IITP of Korea (2017-0-00072, 2016-0-00018). The authors would like to thank Giljoo Nam and Joo Ho Lee for helpful discussions and Hyun Jin Ku for proof reading.

## REFERENCES

- Miika Aittala, Tim Weyrich, and Jaakko Lehtinen. 2015. Two-shot SVBRDF capture for stationary materials. *ACM Trans. Graph.* 34, 4 (2015), 110:1–13.
- Gary A Atkinson and Edwin R Hancock. 2006. Recovery of surface orientation from diffuse polarization. *IEEE Transactions on Image Processing* 15, 6 (2006), 1653–1664.
- Seung-Hwan Baek, Diego Gutierrez, and Min H. Kim. 2016. Birefractive stereo imaging for single-shot depth acquisition. *ACM Transactions on Graphics* 35, 6 (2016), 194.
- Guojun Chen, Yue Dong, Pieter Peers, Jiawan Zhang, and Xin Tong. 2014. Reflectance scanning: Estimating shading frame and BRDF with generalized linear light sources. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 117.
- Edward Collett. 2005. Field guide to polarization. SPIE Bellingham, WA.
- Robert L. Cook and Kenneth E. Torrance. 1982. A Reflectance Model for Computer Graphics. *ACM Transactions on Graphics (TOG)* 1, 1 (1982), 7–24. <https://doi.org/10.1145/357290.357293>
- Zhaopeng Cui, Jinwei Gu, Boxin Shi, Ping Tan, and Jan Kautz. 2017. Polarimetric Multi-View Stereo. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 1558–1567.
- Paul E. Debevec and Jitendra Malik. 1997. Recovering High Dynamic Range Radiance Maps from Photographs. In *Proc. ACM SIGGRAPH ’97*. 369–378.
- Kenneth K Ellis. 1996. Polarimetric bidirectional reflectance distribution function of glossy coatings. *JOSA A* 13, 8 (1996), 1758–1762.
- Graham Fyffe, Paul Graham, Borom Tunwattanapong, Abhijeet Ghosh, and P Debevec. 2016. Near-Instant Capture of High-Resolution Facial Geometry and Reflectance. In *Computer Graphics Forum*, Vol. 35. Wiley Online Library, 353–363.
- Abhijeet Ghosh, Tongbo Chen, Pieter Peers, Cyrus A Wilson, and Paul Debevec. 2010. Circularly polarized spherical illumination reflectometry. In *ACM Transactions on Graphics (TOG)*, Vol. 29. ACM, 162.
- Abhijeet Ghosh, Graham Fyffe, Borom Tunwattanapong, Jay Busch, Xueming Yu, and Paul Debevec. 2011. Multiview face capture using polarized spherical gradient illumination. *ACM Transactions on Graphics (TOG)* 30, 6 (2011), 129.
- Abhijeet Ghosh, Tim Hawkins, Pieter Peers, Sune Frederiksen, and Paul Debevec. 2008. Practical modeling and acquisition of layered facial reflectance. *ACM Trans. Graph.* 27, 5 (2008), 139.
- Darya Guarnera, Giuseppe Guarnera, Abhijeet Ghosh, Cornelia Denk, and Mashhud Glencross. 2016. BRDF Representation and Acquisition. *Computer Graphics Forum* 35 (2016), 625–650.
- Giuseppe Claudio Guarnera, Pieter Peers, Paul Debevec, and Abhijeet Ghosh. 2012. Estimating surface normals from spherical stokes reflectance fields. In *European Conference on Computer Vision*. Springer, 340–349.
- Eric Heitz. 2014. Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs. *J. Comput. Graph. Techn.* 3, 2 (2014), 48–107.
- Michael Holroyd, Jason Lawrence, and Todd Zickler. 2010. A coaxial optical scanner for synchronous acquisition of 3D geometry and surface reflectance. *ACM Trans. Graph.* 29, 4 (2010), 99.
- Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. 1992. Surface reconstruction from unorganized points. Vol. 26. ACM.
- MW Hyde IV, JD Schmidt, and MJ Havrilla. 2009. A geometrical optics polarimetric bidirectional reflection distribution function for dielectric and metallic surfaces. *Optics express* 17, 24 (2009), 22138–22153.
- Adrian Jarabo and Victor Arellano. 2017. Bidirectional Rendering of Vector Light Transport. *Computer Graphics Forum* (12 2017).
- Achuta Kadambi, Vage Taamazyan, Boxin Shi, and Ramesh Raskar. 2015. Polarized 3D: High-Quality Depth Sensing with Polarization Cues. In *Proc. ICCV*. IEEE Computer Society, 3370–3378.
- Stefka Nikolova Kasarova, Nina Georgieva Sultanova, Christo Dimitrov Ivanov, and Ivan Dechev Nikolov. 2007. Analysis of the dispersion of optical plastic materials. *Optical Materials* 29, 11 (2007), 1481–1490.
- Jaewon Kim, Shahram Izadi, and Abhijeet Ghosh. 2016. Single-shot Layered Reflectance Separation Using a Polarized Light Field Camera. In *Proc. Eurographics Symposium on Rendering*.
- Jan Korger, Tobias Kolb, Peter Banzer, Andrea Aiello, Christoffer Wittmann, Christoph Marquardt, and Gerd Leuchs. 2013. The polarization properties of a tilted polarizer. *Optics express* 21, 22 (2013), 27032–27042.
- Joo Ho Lee, Adrian Jarabo, Daniel S. Jeon, Diego Gutierrez, and Min H. Kim. 2018. Practical Multiple Scattering for Rough Surfaces. *ACM Transactions on Graphics (TOG)* 37, 6 (2018).
- Anat Levin, Rob Fergus, Frédéric Durand, and William T Freeman. 2007. Image and depth from a conventional camera with a coded aperture. In *ACM Transactions on Graphics*, Vol. 26. ACM, 70.
- Wan-Chun Ma, Tim Hawkins, Pieter Peers, Charles-Felix Chabert, Malte Weiss, and Paul Debevec. 2007. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Proc. Eurographics conference on Rendering Techniques*. 183–194.
- JR Maxwell, J Beard, S Weiner, Da Ladd, and S Ladd. 1973. *Bidirectional Reflectance Model Validation and Utilization*. Technical Report. Environmental Research Inst Of Michigan Ann Arbor Infrared And Optics Div.
- Daisuke Miyazaki, Robby T Tan, Kenji Hara, and Katsushi Ikeuchi. 2003. Polarization-based inverse rendering from a single view. In *null. IEEE*, 982.
- Michal Mojzik, Tomas Skrivan, Alexander Wilkie, and Jaroslav Krivanek. 2016. Bi-Directional Polarised Light Transport. In *Eurographics Symposium on Rendering*.
- Daniel Moreno and Gabriel Taubin. 2012. Simple, accurate, and robust projector-camera calibration. In *3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on*. IEEE, 464–471.
- Koki Nagano, Graham Fyffe, Oleg Alexander, Jernej Barbic, Hao Li, Abhijeet Ghosh, and Paul E Debevec. 2015. Skin microstructure deformation with displacement map convolution. *ACM Trans. Graph.* 34, 4 (2015), 109–1.
- Giljoo Nam and Min H. Kim. 2014. Multispectral Photometric Stereo for Acquiring High-Fidelity Surface Normals. *IEEE Computer Graphics and Applications* 34, 6 (2014), 57–68.
- Giljoo Nam, Joo Ho Lee, Diego Gutierrez, and Min H. Kim. 2018. Practical SVBRDF Acquisition of 3D Objects with Unstructured Flash Photography. *ACM Transactions on Graphics (TOG)* 37, 6 (2018).
- Giljoo Nam, Joo Ho Lee, Hongzhi Wu, Diego Gutierrez, and Min H. Kim. 2016. Simultaneous Acquisition of Microscale Reflectance and Normals. *ACM Transactions on Graphics (Proc. SIGGRAPH Asia 2016)* 35, 6 (2016). <https://doi.org/10.1145/2980179.2980220>
- Jannik Boll Nielsen, Henrik Wann Jensen, and Ravi Ramamoorthi. 2015. On optimal, minimal BRDF sampling for reflectance acquisition. *ACM Transactions on Graphics* 34, 6 (2015), 186.
- Richard G Priest and Thomas A Gerner. 2000. *Polarimetric BRDF in the microfacet model: Theory and measurements*. Technical Report. NAVAL RESEARCH LAB WASHINGTON DC.
- Jérémie Rivière, Ilya Reshetouski, Luka Filipi, and Abhijeet Ghosh. 2017. Polarization Imaging Reflectometry in the Wild. *ACM Trans. on Graphics* 36, 6 (2017), 206:1–206:14.
- Szymon Rusinkiewicz. 1998. A New Change of Variables for Efficient BRDF Representation. In *Proc. Eurographics Rendering Workshop*. 11–22.
- Yinlong Sun. 2007. Statistical ray method for deriving reflection models of rough surfaces. *JOSA A* 24, 3 (2007), 724–744.
- Kenneth E Torrance and Ephraim M Sparrow. 1967. Theory for off-specular reflection from roughened surfaces. *JOSA* 57, 9 (1967), 1105–1112.
- Silvia Tozza, William AP Smith, Dizhong Zhu, Ravi Ramamoorthi, and Edwin R Hancock. 2017. Linear Differential Constraints for Photo-polarimetric Height Estimation. *arXiv preprint arXiv:1708.07718* (2017).
- Borom Tunwattanapong, Graham Fyffe, Paul Graham, Jay Busch, Xueming Yu, Abhijeet Ghosh, and Paul Debevec. 2013. Acquiring Reflectance and Shape from Continuous Spherical Harmonic Illumination. *ACM Trans. Graph.* 32, 4 (2013), 109:1–12.
- Bruce Walter, Stephen R Marschner, Hongsong Li, and Kenneth E Torrance. 2007. Microfacet models for refraction through rough surfaces. In *Eurographics conference on Rendering Techniques*. Eurographics Association, 195–206.
- Richard A Waltz, José Luis Morales, Jorge Nocedal, and Dominique Orban. 2006. An interior algorithm for nonlinear optimization that combines line search and trust region steps. *Mathematical programming* 107, 3 (2006), 391–408.
- Michael Weinmann and Reinhard Klein. 2015. Advances in Geometry and Reflectance Acquisition (Course Notes). In *SIGGRAPH Asia 2015 Courses (SA ’15)*. ACM, New York, NY, USA, Article 1, 71 pages.
- Tim Weyrich, Jason Lawrence, Hendrik P.A. Lensch, Szymon Rusinkiewicz, and Todd Zickler. 2009. Principles of Appearance Acquisition and Representation. *Foundations and Trends in Computer Graphics and Vision* 4, 2 (Oct. 2009), 75–191.
- Alexander Wilkie and Andrea Weidlich. 2012. Polarised light in computer graphics. In *SIGGRAPH Asia 2012 Courses*. ACM, 8.
- Zexiang Xu, Jannik Boll Nielsen, Jiyang Yu, Henrik Wann Jensen, and Ravi Ramamoorthi. 2016. Minimal BRDF sampling for two-shot near-field reflectance acquisition. *ACM Transactions on Graphics* 35, 6 (2016), 188.
- Zhengyou Zhang. 2000. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence* 22, 11 (2000), 1330–1334.
- Zhiming Zhou, Guojun Chen, Yue Dong, David Wipf, Yong Yu, John Snyder, and Xin Tong. 2016. Sparse-as-possible SVBRDF acquisition. *ACM Transactions on Graphics (TOG)* 35, 6 (2016), 189.