Simultaneous Acquisition of Polarimetric SVBRDF and Normals

Supplemental Material

SEUNG-HWAN BAEK and DANIEL S. JEON, KAIST
XIN TONG, Microsoft Research Asia
MIN H. KIM, KAIST

Fig. 1. Polarimetric diffuse vs. specular light transport. Two linear polarization filters are installed in front of an unpolarized projector and a camera. (a) The diffuse component of the polarized light is reflected on the object surface. The coordinate basis of the incident light stoke vector is converted to the incident coordinate system and then transformed to that of the camera. (b) According to the microfacet theory, the specular component is reflected on the facet surface, of which normal is same as the halfway vector. The coordinate basis of the incident stoke vector is converted w.r.t. the incident coordinate system, where the microfacet normal stands. (Insets) The incident/exitant rotation angles show rotation transformations about the z-axis, which is the direction of the light propagating.

ACM Reference Format:

1 FOUNDATIONS OF POLARIZATION

1.1 Mueller Transformation

Converting Coordinate Systems. A stoke vector of a light ray is defined with respect to a vector coordinate system, where the

\[
\begin{pmatrix}
 y_i \\
 x_i \\
 z_i
\end{pmatrix} = \begin{pmatrix}
 0 & 0 & 1 \\
 -\sin \theta & \cos \theta & 0 \\
 -\cos \theta & -\sin \theta & 0
\end{pmatrix} \begin{pmatrix}
 x_o \\
 y_o \\
 z_o
\end{pmatrix}
\]

In case of diffuse polarization, \( C_{i-n}(\phi_i) \) is the conversion matrix from the coordinate system of the polarized light to the system of the incident polarization that is the incident coordinate system.
(with the plane of incidence that holds normal \( n \)), where the Fresnel/unpolarized Mueller matrices are defined. \( C_{\eta \rightarrow \eta} (\phi_o) \) is the conversion matrix from the exitant polarization system to the camera polarization system. Before applying the Mueller matrices, the coordinate system should be transformed from the \( y_i \) axis of the coordinate system of the incident light stoke vector to the \( y_m \) axis of the incoming surface coordinate system (by rotating it with an angle of \( -\phi_i \) about the \( z \)-axis) and then transformed from the \( y_n, o \) axis of the outgoing surface stoke vector to the \( y_o \) axis of the camera coordinate system (by rotating it with \( \phi_o \)).

In case of specular polarization, where \( C_{\eta \rightarrow \eta} (\phi_o) \) is the coordinate conversion matrix from the light ray \( (x_i, y_i) \) to the plane of incidence (that holds the halfway vector \( h \) of the microfacet normal) \( (x_h, i, y_h, i) \), where the Fresnel Mueller reflection matrix is defined, and \( C_{\eta \rightarrow \eta} (\phi_o) \) is the conversion matrix from the facet normal \( (x_h, i, y_h, i) \) to the camera \( (x_o, y_o) \).

**Fresnel Matrices.** A Fresnel matrix can be used for formulating either transmission or reflection of the polarized light in a form of a Mueller matrix \( F^{\text{Fe}[T, R]} \):

\[
F^{\text{Fe}[T, R]} = \begin{bmatrix}
\frac{\cos^2 \theta}{T^+} & \frac{\sin^2 \theta}{T^+} & 0 & 0 \\
\frac{\cos^2 \theta}{T^\perp} & \frac{\sin^2 \theta}{T^\perp} & 0 & 0 \\
0 & 0 & \sqrt{T^+ T^\perp} \cos \delta & \sqrt{T^+ T^\perp} \sin \delta \\
0 & 0 & -\sqrt{T^+ T^\perp} \sin \delta & \sqrt{T^+ T^\perp} \cos \delta
\end{bmatrix},
\]

where \( F \) can be either Fresnel transmission coefficients \( T \) or reflection coefficients \( R \), and the \( \delta \) is the retardation phase shift between the perpendicular and parallel waves, either \( \pi \) or \( 0 \). For a dielectric surface, \( \cos \delta = -1 \) when the incident angle is less than the Brewster angle; \( \cos \delta = 1 \), otherwise, and vice versa for \( \sin \delta \). Here \( T^\perp \) and \( T^\parallel \) are the Fresnel transmission coefficients for the perpendicular (denoted by \( \perp \)) and the parallel (\( \parallel \)) components, respectively. When calculating \( T^\perp \) and \( T^\parallel \), \( \theta_i \) and \( \theta_o \) 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.

In our diffuse reflectance model, when we calculate incident Fresnel transmission coefficients, \( \eta_1 \) and \( \eta_2 \) are set to 1.0 and the object refractive index \( \eta \), respectively. \( \cos \theta_i \) and \( \cos \theta_o \) are defined as \( \cos \theta_1 = n \cdot i \) and \( \cos \theta_2 = \sqrt{1 - ((1/\eta) \sin \theta_i)^2} \), respectively. In case of exitant transmission coefficients, \( \eta_1 \) and \( \eta_2 \) are switched as the object refractive index \( \eta \) and 1.0, respectively. \( \cos \theta_i \) and \( \cos \theta_o \) are defined as \( \cos \theta_1 = \sqrt{1 - ((1/\eta) \sin \theta_2)^2} \) and \( \cos \theta_2 = n \cdot o \), respectively following Snell’s law. In addition, as a consequence of the conservation of energy, both \( R^\perp \) and \( T^\parallel \) satisfy: \( T^\perp + R^\perp = 1 \) and \( T^\parallel + R^\parallel = 1 \).  

**Depolarization Matrix.** The light absorbed by the object surface is completely unpolarized due to subsurface scattering with the object pigments. The diffuse absorption is formulated as a 4-by-4 depolarization scattering Mueller matrix \( D \), where the top-left element is the diffuse albedo \( \rho \), and the rest of elements are zero.

\[
D = \begin{bmatrix}
\rho & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix},
\]

where \( \rho \) corresponds to the diffuse albedo. We also adopt the depolarization matrix \( D \) in our diffuse reflection model.

**Linear Polarizer Matrix.** A Mueller transmission matrix \( L \) with a linear polarization angle \( \vartheta \) is formulated as:

\[
L(\vartheta) = \frac{1}{2} \begin{bmatrix}
1 & \cos 2\vartheta & \sin 2\vartheta & 0 \\
\cos 2\vartheta & \cos^2 2\vartheta - \sin^2 2\vartheta & 0 & 0 \\
\sin 2\vartheta & 0 & \sin^2 2\vartheta - \cos^2 2\vartheta & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}. 
\]

**2 FACET DISTRIBUTION & GEOMETRIC ATTENUATION**

GGX distribution function \( D \) [Walter et al. 2007] is as follows:

\[
D(\theta_i; \vartheta) = \frac{1}{\pi \cos^4 \theta_i (\sigma^2 + \tan^2 \vartheta)^2},
\]

where \( \sigma \) is the roughness parameter. Smith G function [Heitz 2014] accounts for the masking-shadowing effect as follows:

\[
G(\theta_i, \vartheta; \sigma) = \frac{2}{1 + \sqrt{1 + \sigma^2 \tan^2 \vartheta}} \left( \frac{2}{1 + \sqrt{1 + \sigma^2 \tan^2 \theta_o}} \right).
\]

Since the geometric terms, both \( D \) and \( G \), determine how many microfacets are observed from the given view direction, these geometric terms are therefore independent of polarization.

**3 POLARIMETRIC DECOMPOSITION**

In advance to estimate both appearance and normals, we generate the polarimetric shading matrix \( H \) from at least nine polarimetric images with different angle of the linear filters. To this end, we solve an overdetermined system:

\[
\text{minimize} \| \bar{I} - \Phi^\top \bar{H} \Phi \|_F^2.
\]

While Equation (7) is a per-pixel optimization, we can solve it efficiently by reformulating Equation (7).

**Tensor Reformulation for Decomposition.** We denote \( \bar{H} \in \mathbb{R}^{N \times 4 \times 4} \) as a tensor consisting of polarimetric shading matrices \( H \) for every pixel, where \( N \) is the number of pixels. Captured intensities of every pixel is described as \( I \in \mathbb{R}^{N \times k \times m} \). First, we repack the tensor \( I \) as a matrix, of which dimension is \( \bar{I} \in \mathbb{R}^{N \times N \times m} \). Intermediate matrix \( \bar{H}' \in \mathbb{R}^{4 \times 4} \) is then estimated by solving the standard least-square optimization: minimize \( \| I - \Phi^\top \bar{H}' \Phi \|_F^2 \). Second, we repack the estimated \( \bar{H}' \) as the dimension of \( \mathbb{R}^{N \times 4 \times N} \) and solve another optimization problem: minimize \( \| \bar{H}' - \bar{H} \Phi \|_F^2 \), where \( \bar{H} \in \mathbb{R}^{N \times 4 \times N} \) is the polarimetric shading matrix for every pixel. Per-pixel polarimetric shading matrix \( \bar{H} \) is finally obtained by repacking \( \bar{H} \) as the dimension of \( \mathbb{R}^{N \times 4 \times 4} \).

**4 NORMAL ESTIMATION**

Our normal estimation consists of two stages where each stage exploit diffuse polarization and specular reflection respectively. In this supplemental material, we describe optimization details of each stage. For intuition and definition of terms, refer to the main paper.

**Diffuse Normals.** Diffuse polarization provides surface azimuth and zenith information with ambiguity [Kadambi et al. 2015]. Since we can estimate refractive index by jointly analyzing both the diffuse
We here analyze each term of our model stating that the (1,1) and (2,2) elements can be modeled as \( KR^+ \) and \( -KR^- \). \( H_{10,20} \) are related to the linear polarization state of the exiting light rays after the reflection as we termed exitant polarization. They exhibit low level of intensities compared to the diagonal components as expected in our modeling because specular intensities and the two elements are similar to each other while the values are negative for (2,2). This observation is aligned with our model stating that the (1,1) and (2,2) elements can be modeled as \( KR^+ \) and \( -KR^- \). \( H_{11,22} \) mainly show
of the Fresnel different term $T_0^{-}$. However, still we can see clear
dependency on the surface azimuth. Our model explains this effect
with $\beta_o$ and $\alpha_o$ which are the cosine and sine values of the surface
azimuthal angle. Also, we can observe that the intensity values be-
come low for the surface where the zenithal angle becomes zero.
This can be also explained by our model; $T_0^{-}$ becomes close to zero
when the surface zenith towards the camera. $H_{12,02}$ correspond to the
incident polarization describing how the polarization state of the
incoming light affects the intensity of the captured image. Similar
to the elements $(1,0), (2,0)$, we can observe the dependency on the
surface azimuth and zenith. However, these elements seem more
sensitive to the color of the object pigments showing clear bluish
color on the negative values. $H_{12,21}$ deviate from our approximated
model of the coaxial setup where they have visible negative intensities.
However, note that our model under the real setup explains this
phenomena well shown in Figure 2. Because of the difference, we
do not make use of these element as inputs of the inverse rendering
algorithm.

6 NOTATIONS TABLE

Table 1 provides the notations used in the main paper.

REFERENCES

Erik Heitz. 2014. Understanding the Masking-Shadowing Function in Microfacet-Based
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.
Microfacet models for refraction through rough surfaces. In Eurographics conference

Table 1. Symbols and notations used in the paper.