Polarimetric Helmholtz Stereopsis

Yuqi Ding, *Student Member, IEEE,* Yu Ji, Zhang Chen, Mingyuan Zhou, Sing Bing Kang, *Fellow, IEEE,* and Jinwei Ye, *Senior Member, IEEE*

Abstract—Helmholtz stereopsis (HS) exploits the reciprocity principle of light propagation (*i.e.*, the Helmholtz reciprocity) for 3D reconstruction of surfaces with arbitrary reflectance. In this paper, we present the polarimetric Helmholtz stereopsis (polar-HS), which extends the classical HS by considering the polarization state of light in the reciprocal paths. With the additional phase information from polarization, polar-HS requires only one reciprocal image pair. We derive the reciprocity relationship of Mueller matrix and formulate new reciprocity constraint that takes polarization state into account. We also utilize polarimetric constraints and extend them to the case of perspective projection. For the recovery of surface depths and normals, we incorporate reciprocity constraint with diffuse/specular polarimetric constraints in a unified optimization framework. For depth estimation, we further propose to utilize the consistency of diffuse angle of polarization. For normal estimation, we develop a normal refinement strategy based on degree of linear polarization. Using a hardware prototype, we show that our approach produces high-quality 3D reconstruction for different types of surfaces, ranging from diffuse to highly specular.

Index Terms—Polarization, Helmholtz Reciprocity, 3D Reconstruction

1 Introduction

Reconstructing 3D surfaces from 2D images is a long-standing ill-posed problem in computer vision. The complex surface reflectance properties of real-world objects make the problem highly challenging. All existing methods are limited to certain types of surface reflectance. For example, passive techniques examine the optical appearance of a surface under non-tightly focused illumination (e.g., a distant light source) and often assume photo-consistency or Lambertian reflectance model for 3D reconstruction.

Helmholtz Stereopsis (HS) [1] is a 3D reconstruction technique that can recover surfaces with arbitrary and unknown reflectance. HS exploits the symmetry of surface reflectance; this is accomplished by using reciprocal image pairs (minimal three) that are captured with exchanged camera and light source positions. The reciprocity property guarantees that the relationship between the intensities at corresponding pixels depends only on the surface shape, and is independent of surface reflectance.

In this paper, we present a novel method we call polarimetric Helmholtz Stereopsis (polar-HS), which extends the classical HS by *considering the polarization state of light in the reciprocal paths*. We investigate the reciprocity relationship when the polarization states of incident and outgoing light are unrestricted (in which case the original Helmholtz reciprocity property cannot be directly applied). We derive a transpositional reciprocity relationship based on the Stokes-Mueller formalism, and formulate a reciprocity constraint for depth and normal estimation. We also exploit polarimetric cues under different types of reflections. We further derive the polarimetric constraints under perspective camera projection. We propose a new polarimetric image decomposition method that allows us to apply the polarimetric constraints under different circumstances. By combining the reciprocity and polarimetric constraints, our method can recover the surface depth and normal with only one reciprocity pair, which greatly simplifies the capture process. Unlike other one-pair HS methods [2], [3] that assume continuous parametric depth functions, polar-HS works for discontinuous depth and does not require priors on surface geometry and material properties. We discuss about the degenerate cases in our one-pair solution and proposed a refinement scheme using the degree of linear polarization.

We validate our method with both synthetic and real experiments. We build a real polar-HS acquisition system with a rotating wheel to allow exchange of camera and light source positions. We perform experiments on objects with various shapes and reflectances, and on different composite scenes. Results show that our method is state-of-the-art.

Contributions. Our key contribution is the derivation of the transpositional reciprocity relationship when unrestricted polarization states are being considered. The original Helmholtz reciprocity property has strong restrictions on the polarization states of the light beams in the reciprocal paths, which is hard to satisfy in practice. The transpositional reciprocity we derive allows the classical HS to be extended to the polarimetric case.

We propose a new image decomposition formulation that consists of three components: polarized-specular, polarized-diffuse, and unpolarized-diffuse. The decomposition provides a more accurate estimation of the angle of polarization for regularizing the surface normal, because it separates diffuse and specular (the angles of polarization under these two cases have a 90° shift). Shape-from-polarization methods usually use the overall angle, and assume a dominant type of reflection. This decomposition also provides a specular map that allows us to apply the polarimetric cues under different types of reflection.

Polar-HS reduces the minimal number of image pairs to only one, without imposing any surface prior. Compared to other shape-from-polarization methods, polar-HS does not have the problem of angular ambiguity and does not require the refractive index of the surface to be known.

- *Yuqi Ding and Zhang Chen were with the Division of Computer Science and Engineering, Louisiana State University, Baton Rouge, LA 70820.*
- *Yu Ji is with Tencent America, New York, NY 10018.*
- *Mingyuan Zhou is with InnoPeak Technology Inc., Bellevue , WA 98004.*
- *Sing Bing Kang is with Zillow Group, Seattle, WA 98101.*
- *Jinwei Ye is with the Department of Computer Science, George Mason University, Fairfax, VA 22030*
- *Corresponding author: Jinwei Ye (E-mail: jinweiye@gmu.edu).*

Manuscript received August 05, 2022; revised August 05, 2022.

In this journal extension of [4], we improve Polar-HS by: (1) extending the polarimetric constraints from orthographic projection to the more general perspective projection, and (2) using the degree of linear polarization (DoLP) to refine the depth and normal estimation of degenerate case when only one pair is used. We have also added more in-depth discussions about the degenerate cases where the angle of polarization $\phi = \pm 90^\circ$, a mathematical derivation for diffuse angle of polarization (AoP), additional ablation studies, and more experimental results.

2 Related Work

We first briefly review physics-based methods for 3D shape recovery before focusing on two specific classes of methods that are most relevant to our work: reciprocity-based and polarization-based. Table 1 summarizes a comparison of our method (polar-HS) with classical methods.

2.1 Physics-based Shape Recovery

Physics-based techniques examine the optical appearance of a surface under certain illumination mode and often assume photo-consistency or Lambertian reflectance model for 3D reconstruction. We can categorize the techniques as passive or active based on the illumination mode. Passive methods use unknown and non-tightly focused illumination (*e.g.*, a distant light source). Notable examples include multi-view stereo [5], [6], [7] and structure-from-motion [8], [9], [10]. As passive methods heavily rely on the object's intrinsic appearance for feature matching, they are ineffective on textureless surfaces. Active techniques use known and controlled illumination as a probe; examples include photometric stereo [11], [12], [13], time-of-flight [14], [15], [16], and structured light [17], [18], [19], [20]. These methods can produce dense 3D reconstruction, but are usually sensitive to view-dependent specularity and the inter-reflection caused by concave surfaces. All these methods have limitations due to the complex reflectance of real-world surfaces.

2.2 Helmholtz Stereopsis (HS)

The method is first introduced by Zickler *et al*. [1]. It is an active approach that is capable of recovering surfaces with arbitrary reflectance. Much progress has subsequently been made to improve the original HS. Tu and Mendonça [3] solve HS with a single pair by assuming a piece-wise linear curve constraint. Zickler *et al*. [2] formulate a PDE constraint by assuming C^1 continuity in depth, so as to perform HS under a binocular setting. Jankó *et al*. [21] introduce a general radiometric calibration method for HS. Delaunoy *et al*. [22] extend HS to full-body scanning by using variational approach to optimize over the entire surface. Weinmann *et al*. [23] combine HS with a structured light technique to improve the reconstruction accuracy. Mori *et al*. [24] introduce an integration-based Helmholtz condition which reduces the noise sensitivity of HS. Roubtsova and Guillemaut [25], [26] derive a Bayesian framework for HS optimization, and use color multiplexing to simultaneously capture the reciprocal pair in order to handle dynamic scenes. Our method extends HS by expanding the reciprocity constraint to polarimetric reflectance and incorporating the polarimetric cues for more accurate 3D reconstruction.

TABLE 1
A comparison of polarimetric Helmholtz Stereopsis (polar-HS) with classical 3D reconstruction methods. Note: MVS - multi-view stereo; PS - photometric stereo; SL - structured light; SfP - shape-from-polarization; HS - Helmholtz stereopsis.

Method	Min # of Inputs	Surface Assumption	Accuracy
MVS	2	Lambertian	Moderate
PS	3	Lambertian	High
SL	> 10	Arbitrary	High
SfP	3	Dielectric	Low
HS	6	Arbitrary	Moderate
polar-HS	4	Arbitrary	High

2.3 Shape-from-Polarization (SfP)

This class of methods model the surface normal using the degree or angle of polarization. The surface's refractive index is usually assumed to be known. Miyazaki *et al*. [27] and Atkinson and Hancock [28] leverage the diffuse polarization for shape estimation. Rahmann and Canterakis [29] propose a specular polarization model and apply it on reflective surfaces. SfP methods usually suffer from the azimuth angle ambiguity which may cause the normal estimation being flipped. To resolve this ambiguity, additional shape priors or visual cues are combined with the polarization model. Examples of such priors and cues are convexity prior [27], [30], boundary normal prior [28], shading cues [31], photometric cues [32], [33], and multi-spectral measurements [34]. Smith *et al*. [35], [36] use SfP to solve for surface height to mitigate the angular ambiguity. Many works use SfP to recover fine surface details given a coarse shape estimated from another technique, such as multi-view stereo [37], [38], [39], photometric stereo [40], [41], space carving [42], structure-from-motion [43], or RGB-D sensors [44], [45]. Beak *et al*. [46] jointly estimate the polarimetric reflectance and the surface geometry. Ba *et al*. [47] propose a data-driven approach that estimates the surface shape from polarimetric images with a deep neural network.

3 Helmholtz Stereopsis

Helmholtz stereopsis (HS) [1] works by exploiting the symmetry of surface reflectance. It uses several reciprocal image pairs with exchanged camera and light source positions to estimate surface normal and depth. Let $O_a \in \mathbb{R}^3$ and $O_b \in \mathbb{R}^3$ be two 3D positions. A reciprocal image pair $\mathcal{I} = \{I_a, I_b\}$ is captured by swapping the camera and light source at O_a and O_b (*i.e.*, I_a is captured with the light source at O_a and the camera at O_b; I_b is captured with the light source at O_b and the camera at O_a). Given a point on the object surface, the goal is to estimate its 3D position P and normal vector $\mathbf{n}$.

Let $f(\mathbf{i}, \mathbf{o})$ be the bidirectional reflectance distribution function (BRDF) of the surface point. f is calculated as the ratio of the outgoing radiance (along the direction $\mathbf{o}$) and the incident irradiance (along the direction $\mathbf{i}$). The Helmholtz reciprocity indicates that f is symmetric about the incident and outgoing directions, *i.e.*, $f(\mathbf{i}, \mathbf{o}) = f(\mathbf{o}, \mathbf{i})$. Let $\mathbf{v}_a = (O_a - P)/\|O_a - P\|^2$ and $\mathbf{v}_b = (O_b - P)/\|O_b - P\|^2$ be two unit directions from P to O_a and from P to O_b. The two intensity images in the reciprocal pair can be formulated as

$$\begin{aligned} I_a &= f(\mathbf{v}_a, \mathbf{v}_b) E \rho_a (\mathbf{v}_a \cdot \mathbf{n}), \\ I_b &= f(\mathbf{v}_b, \mathbf{v}_a) E \rho_b (\mathbf{v}_b \cdot \mathbf{n}), \end{aligned} \tag{1}$$

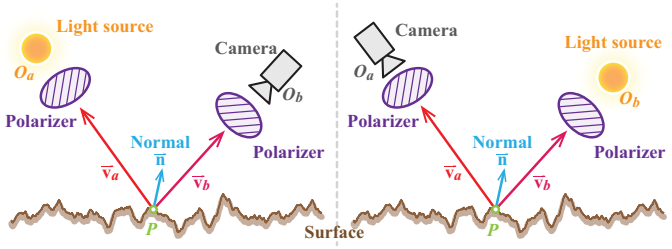


Fig. 1. Configuration of the polarimetric Helmholtz stereopsis.

where E is the light source intensity, $\rho_a = 1/||O_a - P||^2$ and $\rho_b = 1/||O_b - P||^2$ are distance attenuation factors, and $\mathbf{v}_a \cdot \mathbf{n} = \cos\theta_a$ and $\mathbf{v}_b \cdot \mathbf{n} = \cos\theta_b$ are angular fall-off factors.

By dividing the above two equations, we eliminate the light source intensity E and the surface BRDF (noting that $f(\mathbf{v}_a, \mathbf{v}_b) = f(\mathbf{v}_b, \mathbf{v}_a)$). We thus obtain the following reciprocity constraint that regularizes depth and normal with respect to the reciprocal image pair:

$$(I_a \rho_b \mathbf{v}_b^\top - I_b \rho_a \mathbf{v}_a^\top)\mathbf{n} = 0. \tag{2}$$

Given a pre-calibrated camera and light source positions (O_a and O_b), the surface position P and normal $\mathbf{n}$ can be solved with at least three reciprocal pairs. This is because given an estimated P, we need at least three equations to uniquely solve for $\mathbf{n} \in \mathbb{R}^3$.

4 Polarimetric Helmholtz Stereopsis

Our polarimetric Helmholtz stereopsis (polar-HS) uses a linearly polarized light source and a polarization camera to acquire the reciprocal image pair. Our images therefore embed the polarization state of light. Note that the original Helmholtz Reciprocity Principle restricts its applicability to corresponding polarization states for incident and outgoing light [48] (*i.e.*, when the light path is reversed, the polarization states of the two light beams should also be interchanged). It cannot be directly applied to unrestricted representations of the polarization states. We derive the reciprocity relationship under the unrestricted cases, and use it as a constraint for surface reconstruction (Section 4.1).

We introduce polarimetric constraints that are dependent on the surface reflection types (*i.e.*, specular polarized, diffuse polarized, and diffuse unpolarized), and propose a new image separation method to enable the usage of the polarimetric constraints (Section 4.2). An optimization framework is used to jointly estimate the surface normal and depth by combining the reciprocity and polarimetric constraints (Section 4.3).

4.1 Polarimetric Reciprocity

We use the Stokes vector to describe the polarization states. A Stokes vector has four components: $S = [S(0), S(1), S(2), S(3)]^\top$, where $S(0)$ specifies the radiant intensity of light (equivalent to the intensity image), $S(1)$ specifies the preference of horizontal to vertical linear polarization, $S(2)$ specifies the preference of $45°$ to $135°$ linear polarization, and $S(3)$ specifies the preference of right to left circular polarization. Additional constraints on Stokes vector values are: 1) $S(0) \in \mathbb{R}^+$, 2) $S(1), S(2), S(3) \in [-S(0), S(0)]$, and 3) $S(0)^2 \geq S(1)^2 + S(2)^2 + S(3)^2$. Note that a Stokes vector is relative to the selection of reference axes (*i.e.*, a two-dimensional orthogonal basis on the wave plane that is perpendicular to the light's propagation direction). Here we assume linearly polarized light.

We use the same configuration as the standard HS, except that the light source is linearly polarized with Stokes vector S_l and the camera is polarization-sensitive so that it can measure the Stokes vectors of the light received. Fig. 1 illustrates our configuration. Given a reciprocal Stokes vector pair $\mathcal{S} = \{S_a, S_b\}$, S_a is measured with the light source at O_a and the camera at O_b, and S_b with swapped light source and camera positions. Similar to the standard HS, we can formulate the Stokes vectors measured by the camera (S_a and S_b) with the polarimetric surface reflectance along with the distant and angular fall-off factors as

$$\begin{aligned} S_a &= M(\mathbf{v}_a, \mathbf{v}_b) S_l \rho_a (\mathbf{v}_a \cdot \mathbf{n}), \\ S_b &= M(\mathbf{v}_b, \mathbf{v}_a) S_l \rho_b (\mathbf{v}_b \cdot \mathbf{n}). \end{aligned} \tag{3}$$

The 4×4 matrix M (Mueller matrix) represents the polarimetric surface reflectance that describes how the Stokes vector is changed after reflection.

Now the question is: *does M still follows the same reciprocal relationship as the BRDF* (i.e., $M(\mathbf{v}_a, \mathbf{v}_b) = M(\mathbf{v}_b, \mathbf{v}_a)$) *when the Stokes vectors of the incident and reflected light are represented under arbitrary (or unrestricted) reference axes?* The original Helmholtz Reciprocity Principle [49] is as follows:

Theorem 1 (Helmholtz Reciprocity)**.** *Suppose a certain amount of light J leaving the point A in a given direction is polarized in a, and that of this light, the amount K arrives at point B polarized in b. Then, when the light returns over the same path, and the quantity of light J polarized in b proceeds from the point B, the amount of this light that arrives at point A polarized in a will be equal to K.*

Theorem 1 has restricted applicability on corresponding polarization states for incident and outgoing light (*i.e.*, when the light path is reversed, the polarization states of light should also be interchanged). In reality, it is hard to acquire the Stokes measurement under this restricted circumstance. For example, both the camera and light source have their own Stokes representation under local reference axes. It is impractical to exchange their Stokes reference axes when their positions are swapped. We therefore derive a new reciprocity relationship when the representation of polarization state is unrestricted (*i.e.*, the polarization states of incident and outgoing light are not interchangeable due to different reference axes). Our reciprocity relationship is stated as follows:

Lemma 1.1. *When the two light beams in a reversible path are represented by Stokes vectors with reference axes that are associated with the light beams, the reciprocity relationship can be expressed as a transposition of the Mueller matrix that correlates the two Stokes vectors.*

Proof. Consider two light beams with Stokes vectors S_i and S_o that propagate along directions $\mathbf{i}$ and $\mathbf{o}$, respectively. Both S_i and S_o are represented in their local reference axes $\mathbf{b}_i = \{\mathbf{x}_i, \mathbf{y}_i\}$ and $\mathbf{b}_o = \{\mathbf{x}_o, \mathbf{x}_o\}$ (where $\{\mathbf{x}_i, \mathbf{y}_i\}$ and $\{\mathbf{x}_o, \mathbf{x}_o\}$ are two pairs of orthogonal axes on their respective wave plane). Let M be the Mueller matrix that correlates S_i and S_o. We can write the following equations:

$$S_o = M(\mathbf{i}, \mathbf{o}) S_i, \quad S_i = M(\mathbf{o}, \mathbf{i}) S_o. \tag{4}$$

As $\mathbf{b}_i$ and $\mathbf{b}_o$ are associated with the light beams and are switched when the light path is reversed, Theorem 1 cannot be

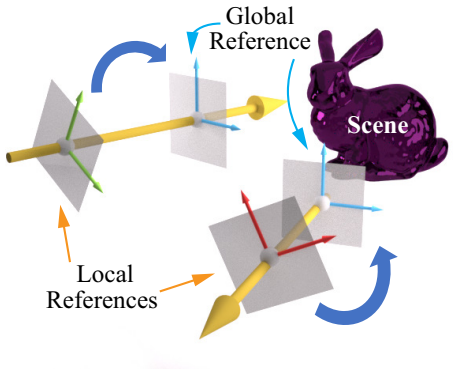


Fig. 2. We rotate the local reference axes to a global reference so that Theorem 1 can be applied to describe the reciprocity relationship between the two paths.

directly applied to describe the reciprocity relationship between $M(\mathbf{i},\mathbf{o})$ and $M(\mathbf{o},\mathbf{i})$.

In order to apply Theorem 1, we define a global reference basis $\mathbf{b}_g = \{\mathbf{x}_g, \mathbf{y}_g\}$ and transform the two Stokes vectors from their local references to the global reference (see Fig. 2). The global reference axes are defined as: $\mathbf{y}_g = \mathbf{i} \times \mathbf{o}$, $\mathbf{x}_g = \mathbf{y}_g \times \mathbf{i}$ for path $\mathbf{i}$, and $\mathbf{x}_g = \mathbf{y}_g \times \mathbf{o}$ for path $\mathbf{o}$. The global reference axes satisfy the condition of Theorem 1 as it is associated with the paths instead of the light beams.

By multiplying the Mueller matrices that rotate the local references to the global one, we obtain two new Stokes vectors S_i' and S_o' that are represented in the global reference axes: $S_i' = M_r(\Phi_i)S_i$ and $S_o' = M_r(\Phi_o)S_o$. Φ_i and Φ_o are the angles spanned by $\mathbf{y}_i$ and $\mathbf{y}_g$, and $\mathbf{y}_o$ and $\mathbf{y}_g$, respectively. Let M' be the Mueller matrix that correlates S_i' and S_o'. Thus M' satisfies the reciprocity relationship described in Theorem 1: $M'(\mathbf{i},\mathbf{o}) = M'(\mathbf{o},\mathbf{i}) = M_f$, leading to

$$S_o' = M_f S_i', \quad S_i' = M_f S_o'. \tag{5}$$

By substituting S_i' and S_o' with S_i and S_o (respectively) in Eq. 5, we have

$$\begin{aligned} S_o &= M_r^{-1}(\Phi_o) M_f M_r(\Phi_i) S_i, \\ S_i &= M_r^{-1}(\Phi_i) M_f M_r(\Phi_o) S_o. \end{aligned} \tag{6}$$

Using the definitions for S_o and S_i in Eq. 4, Eq. 6 becomes

$$\begin{aligned} M(\mathbf{i},\mathbf{o}) &= M_r^{-1}(\Phi_o) M_f M_r(\Phi_i), \\ M(\mathbf{o},\mathbf{i}) &= M_r^{-1}(\Phi_i) M_f M_r(\Phi_o). \end{aligned} \tag{7}$$

Since the rotational Mueller matrices are orthonormal (*i.e.*, $M_r^{-1} = M_r^\top$) and M_f is diagonally symmetric (*i.e.*, $M_f = M_f^\top$) [50], we can derive the following reciprocity relationship in the form of transposition:

$$\begin{aligned} M^\top(\mathbf{i},\mathbf{o}) &= (M_r^{-1}(\Phi_o) M_f M_r(\Phi_i))^\top \\ &= M_r^{-1}(\Phi_i) M_f M_r(\Phi_o) \\ &= M(\mathbf{o},\mathbf{i}). \end{aligned} \tag{8}$$

□

Sekera [51] derives a transpositional reciprocity relationship similar to Lemma 1.1 in the scattering processes.

Reciprocity Constraint. As the Stokes vectors of the camera and light source are observed in their local reference axes, the surface reflectance Mueller matrix M follows the transpositional reciprocity according to Lemma 1.1: $M^\top(\mathbf{v}_a, \mathbf{v}_b) = M(\mathbf{v}_b, \mathbf{v}_a)$. By substituting this reciprocity relationship into Eq. 3 and eliminating M, we obtain the following reciprocity constraint:

$$(S_a \otimes S_l^g \rho_b \mathbf{v}_b^\top - S_l^g \otimes S_b \rho_a \mathbf{v}_a^\top)\mathbf{n} = 0. \tag{9}$$

where S_l^g is the transpose of the pseudoinverse of S_l; $\otimes$ is the Kronecker product.

In Eq. 9, S_a, S_b, and S_l^g are of dimension 4×1 as they are Stokes vectors. The Kronecker product $\otimes$ between S_a (or S_b) and S_l^g results in a matrix of dimension 16×1. ρ_a and ρ_b are two scalars that would not affect the matrix dimension. $\mathbf{n}$, $\mathbf{v}_a$ and $\mathbf{v}_b$ are 3D vectors. The left-hand side of this equation ends up as a 16×1 matrix. Note that, this doesn't warrant solving normal using one reciprocity constraint as the 16 equations are not independent.

We use Eq. 9 to estimate surface depth ($\rho_{a,b}$ and $\mathbf{v}_{a,b}$ are derivable from depth) and normal $\mathbf{n}$ in an iterative way when given reciprocal Stokes vector pair $\mathcal{S} = \{S_a, S_b\}$ and pre-calibrated light source Stokes vector S_l. More details on the optimization algorithm can be found in Section 4.3. Given a depth estimation, we need at least three equations to uniquely solve for the $\mathbf{n} \in \mathbb{R}^3$.

The standard HS generates the three equations from three reciprocal pairs. In our polarimetric case, we obtain two equations for one reciprocal pair by changing the polarization state of light source, because our reciprocity constraint (Eq. 9) is dependent on the light source (while the standard HS is not). The maximum number of independent equations we can obtain for one pair is two. This is because two dot products (*i.e.*, $\mathbf{v}_a \cdot \mathbf{n}$ and $\mathbf{v}_b \cdot \mathbf{n}$) reduce the terms that involve the normal to scalars.

In principle, if we use only the reciprocity constraint, we would need at least two reciprocal pairs (*i.e.*, 4 equations) to solve for the depth and normal. However, because the observed polarization states are directly related to the surface geometry and reflectance type, we use the polarimetric constraints (Section 4.2) to formulate one additional independent equation on normal. This means that we can estimate depth and normal *using only one reciprocal pair*.

4.2 Polarimetric Cues

We consider the polarimetric constraints under different types of reflected light. We decompose the measured Stokes vectors into three components of different reflection and polarization characteristics (*i.e.*, specular vs. diffuse, polarized vs. unpolarized), and then derive a specific polarimetric constraint for each type of reflection.

4.2.1 Polarimetric Image Decomposition.

The problem has been studied in polarization-based specularity removal [52], [53], [54]. An image is often decomposed into a specular component and a diffuse component, depending on the polarization status. It is commonly assumed that the specular component is polarized while diffuse is unpolarized. However, as shown in [28], diffuse reflection also exhibits useful polarimetric characteristics that regularize the surface normal. Here we propose a new decomposition formulation that separates the observed Stokes vectors (S) into three components: specular-polarized (S_{sp}), diffuse-polarized (S_{dp}), and diffuse-unpolarized (S_{ud}):

$$S = S_{sp} + S_{dp} + S_{du}. \tag{10}$$

Suppose we have two linearly polarized light sources with the same intensity but perpendicular angles of polarization. Without loss of generality, we assume their angles of polarization are $0°$

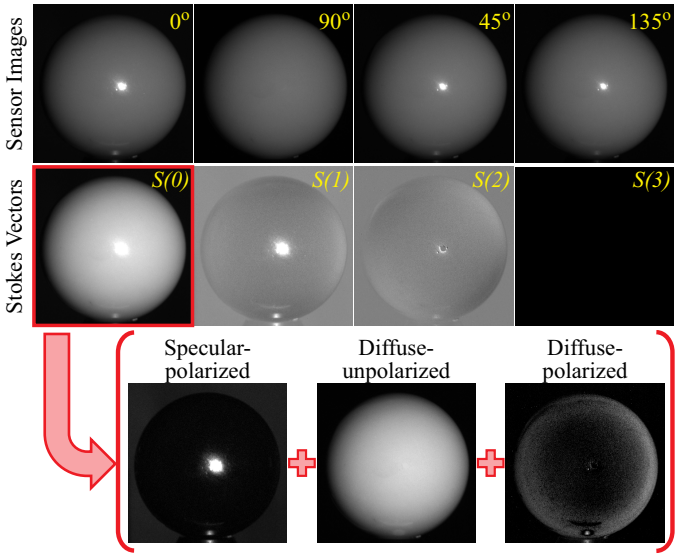


Fig. 3. The proposed decomposition. We first convert the images captured by a polarization camera to the Stokes vectors. We then decompose the Stokes vectors into three components.

and 90°. Their Stokes vectors are S_l^0 and S_l^{90}, respectively. Note that these light sources can also provide us the two reciprocity constraints. Let S^0 be the Stokes vector reflected from a surface point and observed by the camera when light source is S_l^0, and S^{90} is observed under S_l^{90} (see Section 4.1 for composition and properties of the Stokes vector). Here we assume S^0 and S^{90} are linearly polarized (*i.e.*, $S^0(3) = S^{90}(3) = 0$).

We now show how S^0 can be decomposed; S^{90} can be similarly decomposed. For notation simplicity, we drop the superscript for degree in the decomposition components (S_{sp}, S_{dp}, and S_{du}). Since the light sources' angles of polarization are crossed by 90°, the polarization parameters in S_l^0 and S_l^{90} have the following relationship:

$$S_l^0(1) + S_l^{90}(1) = 0, \quad S_l^0(2) + S_l^{90}(2) = 0. \tag{11}$$

Since the specular reflection is always fully polarized and its angle of polarization is the same as that for the light source, we use this relationship to cancel out the polarization parameters in S_{sp} by adding S^0 and S^{90}. Since S_{du} is unpolarized, it has only the intensity parameter: $S_{du} = [S_{du}(0), 0, 0, 0]^\top$. The polarization parameters in $S^0 + S^{90}$ are then solely related to S_{dp}, yielding

$$S_{dp}(1) = \frac{S^0(1) + S^{90}(1)}{2}, S_{dp}(2) = \frac{S^0(2) + S^{90}(2)}{2}. \tag{12}$$

The polarization parameters in S_{sp} are computed as

$$S_{sp}(1) = S^0(1) - S_{dp}(1), \quad S_{sp}(2) = S^0(2) - S_{dp}(2). \tag{13}$$

We calculate $S(0)$ for both S_{dp} and S_{sp} using the Stokes vector constraint $S(0)^2 = S(1)^2 + S(2)^2 + S(3)^2$, as the two components are fully polarized. Finally, we compute the intensity of S_{du} as

$$S_{du}(0) = S^0(0) - S_{sp}(0) - S_{dp}(0). \tag{14}$$

Using Eqs. 12-14, we can decompose a reflected Stokes vector into three components. Fig. 3 shows an example of our decomposition.

4.2.2 *Polarimetric Constraints.*

Both the specular and diffuse polarized reflections can regularize the surface normal. According to Fresnel's equations, the specular reflection is dominated by s-polarized light, whose angle of polarization is perpendicular to the incidence plane (*i.e.*, the plane spanned by the surface normal and the incident light). This happens when the incident light is not oblique to the local surface. In the diffuse reflection case, the angle of polarization has a 90° phase shift [28] which means that the vibration direction lies on the incidence plane. Therefore, under orthographic projection, by projecting the angle of polarization and surface normal onto the image plane, we can formulate the two constraints for the diffuse- and specular-polarized reflections:

$$[\sin(\phi), -\cos(\phi), 0]\mathbf{n} = 0, \tag{15a}$$

$$[\sin(\phi + \frac{\pi}{2}), -\cos(\phi + \frac{\pi}{2}), 0]\mathbf{n} = 0, \tag{15b}$$

where ϕ is the angle of polarization (AoP) and is calculated as

$$\phi = \frac{1}{2}\arctan(S_{*p}(2)/S_{*p}(1)), \tag{16}$$

where S_{*p} is S_{dp} for diffuse reflection and is S_{sp} for specular reflection. Eq. 15a is the constraint for diffuse-polarized reflection. Eq. 15b is for the specular-polarized case, whose angle of polarization is shifted 90°. Similar constraints are used in [35], [36]. However, most methods directly use the overall Stokes vector to compute the angle of polarization. We empirically show that we are able to estimate a more accurate angle of polarization by using the decomposed components.

Proof of Diffuse Polarimetric Constraint. It is conceptually shown in [52] that under a same set of polarization axes, the AoP of diffuse reflection lies in the incidence plane. However, to the best of our knowledge, there is lack of mathematical proof to this constraint. Here we present a Mueller matrix-based proof under the assumption of smooth dielectric surface. Within the proof, we set the polarization axes as: x axis being perpendicular to the incidence plane, y axis being parallel to the incidence plane, and the $x-y$ plane being perpendicular to the wave propagation direction.

Diffuse reflection consists of light that first transmits into the surface, be scattered and then transmits out of the surface [52]. The final outgoing light (*i.e.*, the light observed by camera) is the result of the scattered light after a one-time refraction. Hence, the Stokes vector S_T of the outgoing light satisfies:

$$S_T = M_T S_S, \tag{17}$$

where S_S is the light after scattering and before transmitting out, and M_T is the Mueller matrix for transmission. According to [55], M_T is given by:

$$M_T = \frac{\sin 2\theta_i \sin 2\theta_r}{2(\sin\theta_+ \cos\theta_-)^2} \cdot \begin{bmatrix} \cos^2\theta_- + 1 & \cos^2\theta_- - 1 & 0 & 0 \\ \cos^2\theta_- - 1 & \cos^2\theta_- + 1 & 0 & 0 \\ 0 & 0 & 2\cos\theta_- & 0 \\ 0 & 0 & 0 & 2\cos\theta_- \end{bmatrix}, \tag{18}$$

where θ_i is the incidence angle, θ_r is the refraction angle, and $\theta_\pm = \theta_i \pm \theta_r$. Note that θ_i, θ_r satisfies $n_s \sin\theta_i = n_a \sin\theta_r$ (Snell's Law), with n_s, n_a being the refractive indices of the

surface and air. Based on Eqs. 16, 17, 18, the AoP of diffuse reflection is

$$\phi = \frac{1}{2}\arctan\frac{2\cos\theta_- S_S(2)}{((\cos^2\theta_- - 1)S_S(0) + (\cos^2\theta_- + 1)S_S(1))}. \tag{19}$$

Since S_S is generally considered as unpolarized due to scattering, we have $S_S(2) = 0$ and the numerator has a value of 0. In the meantime, $S_S(1) = 0$ and $\cos^2\theta_- \leq 1$, so the denominator is ≤ 0. As a result, the diffuse AoP $\phi = \frac{1}{2}\arctan 0^- = \frac{\pi}{2}$, meaning that it is colinear with the y polarization axis. Since y polarization axis lies in the incidence plane, we prove that the diffuse AoP also lies in the incidence plane.

Extension to Perspective Projection. Note that Eq. 15a and Eq. 15b assume orthographic projection. Here we derive an extension of the polarimetric constraints for the more general perspective projection. Recall that the angle of polarization is perpendicular to the incidence plane for specular reflection while it lies in the incidence plane for diffuse reflection. Here the angle of polarization is defined based on the wave propagation direction (*i.e.*, view direction), while the ϕ calculated in Eq. 16 is the projection of the angle of polarization on the image plane. We represent the projected angle of polarization as a unit vector $\mathbf{p} = [\cos(\phi), -\sin(\phi), 0]^\top$ and denote the unit view direction as $\mathbf{v}$, both under the camera coordinate system. For diffuse reflection, all of $\mathbf{p}, \mathbf{v}$ and surface normal $\mathbf{n}$ lie in the incidence plane. Therefore, the cross product of $\mathbf{p}, \mathbf{v}$ is perpendicular to $\mathbf{n}$. For specular reflection, $\mathbf{p}$ lies in a plane that is perpendicular to the incidence plane, so an extra cross product is needed. Specifically, the polarimetric constraints under perspective projection are

$$([\cos(\phi), -\sin(\phi), 0] \times \mathbf{v}^\top)\mathbf{n} = 0, \tag{20a}$$

$$(([\cos(\phi), -\sin(\phi), 0] \times \mathbf{v}^\top) \times \mathbf{v}^\top)\mathbf{n} = 0, \tag{20b}$$

where Eq. 20a is for diffuse case and Eq. 20b is for specular.

In order to use the polarimetric constraints, we threshold $S_{sp}(0)$ to a binary mask that indicates the specular pixels. We use Eq. 20b as the additional constraint for the specular pixels, and Eq. 20a for all other pixels considered diffuse. The diffuse case (Eq. 20a) is similar to the perspective phase angle proposed in [56].

4.3 Depth and Normal Recovery

By combining the reciprocity and polarimetric constraints, we can form a linear system for the surface normal $\mathbf{n}$, *i.e.*, $\mathbf{W}(d)\mathbf{n} = 0$: the constraint equations are stacked up as rows in the linear system. The coefficient matrix $\mathbf{W}$ is a function of the surface depth d. We solve d and $\mathbf{n}$ with alternate iterations.

4.3.1 Depth and Normal Estimation

We first optimize the depth values. Once we have the depth estimation, we can solve the normal $\mathbf{n}$ using $\mathbf{W}(d)\mathbf{n} = 0$.

For each pixel in the reference view, we discretize depth into numerous candidates and estimate depth by choosing the best candidate. Under each depth candidate, we use calibrated camera extrinsics to reproject pixel to non-reference views.

Reciprocity and Polarimetric Constraints. For a reciprocal pair of pixels, the corresponding $\mathbf{W}(d)$ consists of a reciprocity constraint (Eq. 9) and two per-view polarimetric constraints (Eq. 20a or Eq. 20b). For non-reference view, the polarimetric constraint needs to take camera extrinsics into account. For example, Eq. 20a becomes

$$([\cos(\phi_i), -\sin(\phi_i), 0] \times \mathbf{v}^\top)R_i\mathbf{n} = 0 \tag{21}$$

where ϕ_i is the diffuse AoP in view i, and R_i is the rotation of coordinate system from reference view to view i.

Given the true depth d^*, the rank of $\mathbf{W}(d^*)$ should be 2 since the surface normal $\mathbf{n}$ has unit length. If the depth value is incorrect, the rank of $\mathbf{W}$ will be greater than 2. This indicates that if we apply SVD on $\mathbf{W}$: $\mathbf{W} = U\Sigma V^\top$, where $\Sigma = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$, $\sigma_1 \geq \sigma_2 \geq \sigma_3$, the ratio σ_2/σ_3 will be infinitely large at the true depth d^*. Thus, we use the exponential decay function proposed in [26] as our data term for depth estimation:

$$E_{data}(d) = \exp(-\mu\frac{\sigma_2(d)}{\sigma_3(d)}), \tag{22}$$

where $\mu = 0.2\ln(2)$ [26].

We also include a smoothness term to reduce noise in the depth estimation :

$$E_{smooth}(d) = \sum_{(p,q)\in\mathcal{N}} \min(\|d_p - d_q\|, K), \tag{23}$$

where p and q are two pixels in the neighborhood $\mathcal{N}$; K is a truncation threshold that prevents the discontinuous depths from being smoothed.

Consistency of Diffuse AoP. In addition to the above terms, we further propose to use the consistency of diffuse AoP as a matching criterion, which is especially beneficial when the input consists of only one reciprocal pair. The reason is that $\mathbf{W}(d)\mathbf{n} = 0$ is a necessary but insufficient condition [1]. Due to image noise and finite resolution, Eq. 22 can be minimized at incorrect depth. Therefore, we develop a consistency term on diffuse AoP to achieve more accurate depth estimation.

In order to apply this term, we assume the camera is only weakly perspective, so Eq. 15a is approximately correct for most object pixels. Here we set the polarization axes to be aligned with the image axes. For images at different viewpoints, Eq. 15a is invariant up to a rotation. Specifically, let ϕ and $\mathbf{n}$ be the diffuse AoP and surface normal in the reference view, then we have

$$[\sin(\phi_i), -\cos(\phi_i), 0]R_i\mathbf{n} = 0, \tag{24}$$

where ϕ_i is the diffuse AoP in view i, and R_i is the rotation of coordinate system from reference view to view i.

We can decompose R_i into rotations around the z, y, x axes as $R_{i,z}(\theta_{i,z})R_{i,y}(\theta_{i,y})R_{i,x}(\theta_{i,x})$ where $\theta_z, \theta_y, \theta_x$ are the Euler angles. By assuming $R_{i,y}(\theta_{i,y})R_{i,x}(\theta_{i,x}) \approx \mathbf{I}$, since $[\sin(\phi_i), -\cos(\phi_i), 0]R_{i,z}(\theta_z) = [\sin(\phi_i - \theta_{i,z}), -\cos(\phi_i - \theta_{i,z}), 0]$, Eq. 24 becomes

$$[\sin(\phi_i - \theta_{i,z}), -\cos(\phi_i - \theta_{i,z}), 0]\mathbf{n} = 0. \tag{25}$$

Note that the assumption of $R_{i,y}(\theta_{i,y})R_{i,x}(\theta_{i,x}) \approx \mathbf{I}$ holds when the z axis of the camera is approximately orthogonal to the rotary stage, which can be easily realized (see our setup in Fig. 6). Combining Eq. 15a and Eq. 24, we can establish the following relationship on the diffuse AoP of different views:

$$\phi = (\phi_i - \theta_{i,z}) \bmod \pi. \tag{26}$$

Based on this, we utilize the following term to measure the consistency of diffuse AoP:

$$E_{aop}(d) = \frac{1}{m}\sum_{i=1}^{m} \min(||\phi - \hat{\phi}_i||, \pi - ||\phi - \hat{\phi}_i||)/\pi, \quad (27)$$

where $\hat{\phi}_i = (\phi_i - \theta_{i,z}) \bmod \pi$, and m is the total number of views in all reciprocal pairs except the reference view. The reason for using the $\min(\cdot)$ operation is that AoP wraps around at 0 and π.

Combined Cost Function. The combined cost function for depth estimation is:

$$d^* = \underset{d}{\operatorname{argmin}} \sum (E_{data} + \lambda_1 E_{aop} + \lambda_2 E_{smooth}), \quad (28)$$

where λ_1 and λ_2 are balancing weights. In our experiments, we use $\lambda_1 = 0.2$ and $\lambda_2 = 0.01$. We use graph-cut [57] to solve the depth as a multi-labeling problem.

4.3.2 Degenerate Case for Depth and Normal Estimation

Solving depth and normal with one reciprocal pair, the coefficient matrix $\mathbf{W}(d)$ may become degenerate when AoP $\phi = \pm 90^\circ$. In this case, the rank constraint we use for solving depth becomes unstable as the singular value σ_3 is very close to zero for all depth values, resulting σ_2/σ_3 being infinitely large at all time. Therefore, when $\phi = \pm 90^\circ$, the depth and normal values estimated by this rank constraint are prone to large errors.

In order to handle this degenerate case, we propose a normal refinement strategy based on the degree of linear polarization (DoLP) of the diffuse component.

The DoLP of the diffuse component can be computed from Stokes vector as:

$$DoLP = \frac{S_{dp}(0)}{S_{dp}(0) + S_{du}(0)}. \quad (29)$$

Recall that diffuse reflection consists of light that first transmits into the surface, be scattered, and then transmits out of the surface. The light is assumed to be unpolarized after scattering and partially polarized after transmitting out. Based on the Fresnel equation for refraction, the DoLP of the diffuse component can also be formulated as [36]:

$$DoLP = \frac{(n - 1/n)^2 \sin^2\theta_r}{2 + 2n^2 - (n + 1/n)^2 \sin^2\theta_r + 4\cos\theta_r \sqrt{n^2 - \sin^2\theta_r}}, \quad (30)$$

where n is the refractive index, and θ_r is the angle between surface normal $\mathbf{n}$ and view direction $\mathbf{v}$.

Our normal refinement strategy consists of the following three steps. First, we obtain the surface normal $\mathbf{n}$ for all pixels using only reciprocity and AoP constraints. Second, we select pixels with reliable surface normal based on the angle of polarization ϕ and use these pixels to estimate the refractive index n via Eq. 30. Third, using the estimated refractive index, we refine the surface normal of the rest pixels by formulating Eq. 30 as an additional constraint on surface normal.

4.3.3 Iterative Update of Depth and Normal

We iteratively update the depth and normal with the following steps: 1) we apply Poisson integration on $\mathbf{n}$ to obtain a new set of depth d'; 2) we use d' as an additional guidance in Eq. 28 to optimize the depth with a finer depth interval; 3) we use the estimated depth to form $\mathbf{W}(d)$ and solve for the normal again. We use the normal difference to decide whether the iterative update has converged. In our experiments, it usually converges after two iterations.

5 Experiments

We validate our method with both synthetic and real experiments on scenes with various shapes and reflectance. All our experiments are run on a laptop computer with Intel Core i7-8750H processor (2.2GHz) and 16GB memory. Our surface reconstruction algorithm is implemented in Matlab without acceleration. The running time of our reconstruction algorithm is about 20 mins for one pair, 23 mins for two pairs, and 26 mins for three pairs. The execution time does not increase significantly with the number of reciprocal pairs. This is because the pair number affects only the dimension of the coefficient matrix $\mathbf{W}$. The complexity of the graph-cut-based depth optimization does not increase when more pairs are used. Instead, the computation overhead on using more numbers of pairs mainly come from accessing data and the SVD decomposition of $\mathbf{W}$.

5.1 Synthetic Experiments

We use the Mitsuba 2 renderer[1] to simulate polarimeric images as captured by a polarization camera. Specifically, we use the polarized rendering mode to simulate four directional polarization images: I^0, I^{90}, I^{45}, and I^{135} (each with resolution 500×500). In the polarized rendering mode, the renderer will track the full polarization state of light during simulation. The system configuration mirrors our real experimental setup. We use the real-captured KAIST pBRDF dataset [58] to model the polarimetric surface reflectance. We test on a variety of 3D models and surface reflectance. Fig. 4 shows our rendered images, recovered normal maps, and recovered 3D surfaces with two reciprocal pairs. We evaluate the reconstruction with per-pixel angular errors (see normal error map) and the mean angular error (MAE).

Ablation on Material Types. We use different pBRDFs provided by the KAIST dataset on a sphere object to test our performance with respect to the material types. Fig. 5 shows recovered normal maps and cross-sections of the recovered shapes. We also compare the performance of our method with respect to the number of pairs (one pair vs. two pairs). We can see that the one-pair results are sensitive to the material type as the polarized reflection of some materials (*e.g.*, mint silicone) is weak, which results in the angle of polarization being highly noisy and unreliable. The two-pair results are more robust as the reciprocity constraint alone provides sufficient regularization. Table 2 shows the mean angular errors of normal estimation for different types of materials. Note that our method also works on metallic materials, as we use the full-Stokes representation and our reciprocity constraint can circularly polarized light that might exist in metal reflections.

Ablation on Constraints. We perform ablation study to verify the effectiveness of our constraints: reciprocity constraints and polarimetric constraints. Specifically, we perform experiments using three reciprocal pairs and compare our full method with the version without using polarimetric constraint (Polar-HS*) and the standard Helmholtz Stereopsis (HS). We test on the same set of materials that are used in the material ablation. By comparing our method with HS, we can see using polarimetric reciprocity

1. https://www.mitsuba-renderer.org

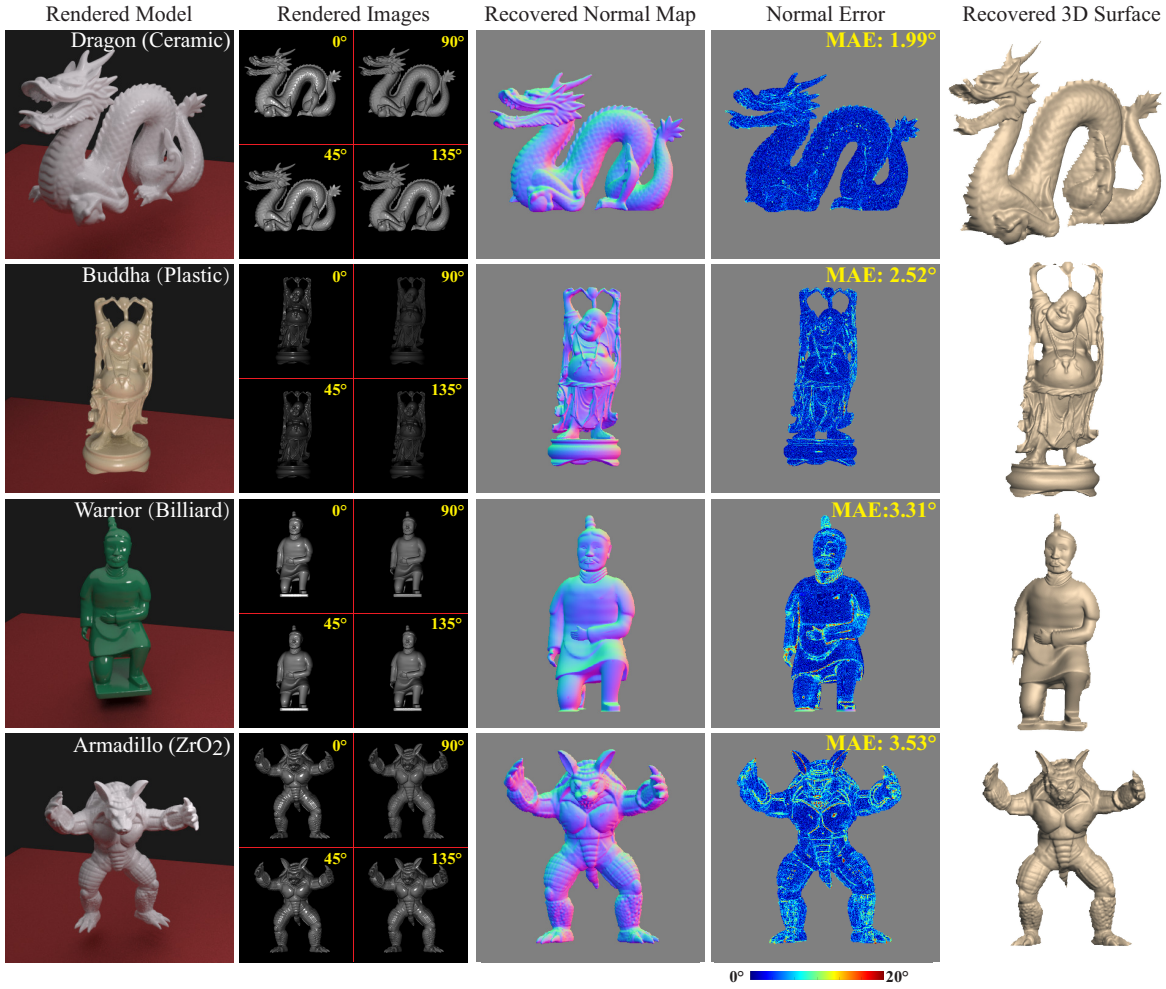


Fig. 4. Normal reconstruction on synthetic data.

TABLE 2
Mean angular error (in degree) of the normal estimations with respect to different types of materials, number of pairs, and constraints.

Materials	polar-HS (1-Pair)	polar-HS (2-Pair)	polar-HS (3-Pair)	Polar-HS* (3-Pair)	HS (3-Pair)
Spectralon	17.6963	2.2087	0.9628	2.1951	11.3151
Chrome	20.0153	9.4524	3.9634	6.3813	14.3129
Gold	18.0669	7.5412	5.1945	8.2255	13.7741
Black Billiard	18.4281	6.0594	1.2341	5.0882	15.0863
ZrO_2	8.5610	1.0912	1.0824	1.8060	10.9464
Plastic POM	8.8094	0.8978	0.8089	1.6302	9.4417

significantly helps reduce the error for all cases. By comparing with Polar-HS*, we can see that the polarimetric constraint also plays an important role on improving the reconstruction accuracy. The comparison results are shown in Table 2.

Ablation on Noise Levels. We evaluate our method with respect to different levels of noise. In this experiment, we use a sphere object with the "white billiard" material. We add Gaussian white noise to the rendered images and use the signal-noise ratio (SNR) to quantify the noise level (smaller SNR number indicates higher noise level). We evaluate the reconstruction using MAE of the normal estimation. We test our method (polar-HS) using one and two reciprocal pairs respectively, and compare with the standard HS that uses three pairs. The results are reported in Table 3. We can see that both our one-pair and two-pair methods are more accurate than the three-pair HS, and less sensitive to noise.

TABLE 3
Mean angular error (in degree) of the normal estimations with respect to the noise levels.

Method \ SNR/dB	10	20	30	40
HS (3-pair)	23.41	13.59	8.05	6.87
polar-HS (1-pair)	14.71	8.05	6.53	6.27
polar-HS (2-pair)	19.16	6.43	2.14	0.95

Ablation on Distance and Focal Length. In order to show that our polarimetric constraints formulated under the perspective project is effective, we evaluate our one-pair method on different object distances and different levels of focal length errors. In this experiment, we use a sphere object and use the "PEEK" material from the KAIST pBRDF dataset. We also compare the performance of using orthogonal polarimetric constraints (Eq. 15a, 15b) vs. perspective polarimetric constraints (Eq. 20a, 20b) during

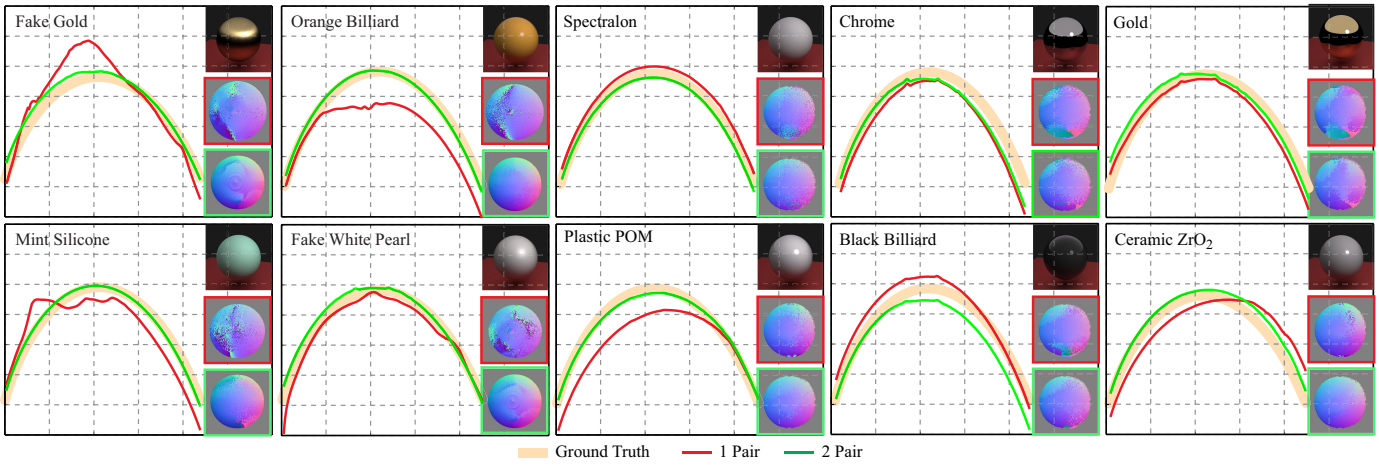


Fig. 5. Reconstruction results of our method when using one and two reciprocal pairs for different types of materials. Here we show the cross-sections of reconstructed surfaces in comparison with the ground truth, color images of the sphere rendered with the specific material, and estimated normal maps.

TABLE 4
Mean angular error (in degree) of the normal estimations with respect to different object distances. "Orthogonal cons." refers to using orthogonal polarimetric constraints (Eq. 15a, 15b). "Perspective cons." refers to using perspective polarimetric constraints (Eq. 20a, 20b). The object distances are relative to the diameter of the sphere object.

Object distance	1	2	3	4	5
Orthogonal cons.	10.97	7.83	7.28	7.01	7.17
Perspective cons.	5.48	5.25	5.77	5.71	5.97

TABLE 5
Mean angular error (in degree) of the normal estimations with respect to different relative focal length errors. The object distance is kept as 3 times the diameter of the sphere object. "Orthogonal cons." refers to using orthogonal polarimetric constraints (Eq. 15a, 15b). "Perspective cons." refers to using perspective polarimetric constraints (Eq. 20a, 20b).

Focal length error	-2%	-1%	0%	+1%	+2%
Orthogonal cons.	9.54	8.02	7.28	10.82	17.61
Perspective cons.	9.22	7.28	5.77	8.04	13.90

normal estimation. Table 4 shows the mean angular normal errors under different object distances. The distances are relative to the diameter of the sphere object. We can see that using perspective constraints consistently produces lower normal errors. The margin is larger under smaller object distances, since perspective effect is more significant under such cases. Table 5 shows the results under different focal length errors. Here the focal length errors are relative to the ground truth focal length, and the object distance is fixed at 3 times the sphere diameter. It can be seen that perspective constraints produce more accurate normal estimation and are more robust to focal length errors.

5.2 Real Experiments

5.2.1 *Prototype of Polarimetric Capture*

We implement a physical system for capturing the polar-HS image pairs. We mount a polarization camera and a pico projector on an automated rotating wheel so that their positions can be precisely exchanged.

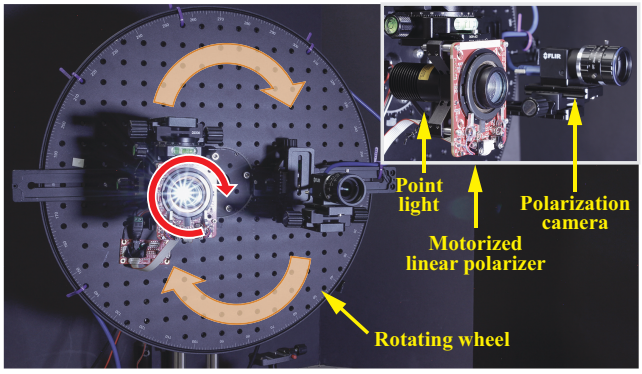


Fig. 6. Acquisition system prototype.

System Construction. As shown in Fig. 6, our acquisition system consists of a monochrome polarization camera (FLIR Blackfly S Polar-Mono) and a pico projector (Sony MP-CL1A) as point light source. The polarization camera captures four directional polarization images (*i.e.*, I^0, I^{90}, I^{45}, and I^{135}) in one shot as its sensor has on-chip polarizers[2]. The camera uses a 25mm F/5.6 lens. The exposure time we use is 800 ms. With the four directional polarization images captured by the camera, we can compute the Stokes vector as $S(0) = I^0 + I^{90}$, $S(1) = I^0 - I^{90}$, and $S(2) = I^{45} - I^{135}$. Note that since we consider only linear polarization states, $S(3) = 0$.

We mount a linear polarizer in front of the projector to generate polarized light. Note that our reciprocity constraint does not have restriction on the polarization state of light source. But we do require their state is known. We use linearly polarized light for polarimetric image decomposition. The camera and projector are mounted on a rotating wheel so that their positions can be precisely exchanged. The distance between the camera and light source is around 17.5 cm. All moving parts of our system (*i.e.*, the rotating wheel and the light source polarizer) are controlled with motorized rotators; our acquisition procedure is fully automated. The distance between our capture system and the scene is about

2. https://www.sony-semicon.co.jp/e/products/IS/industry/technology/polarization.html

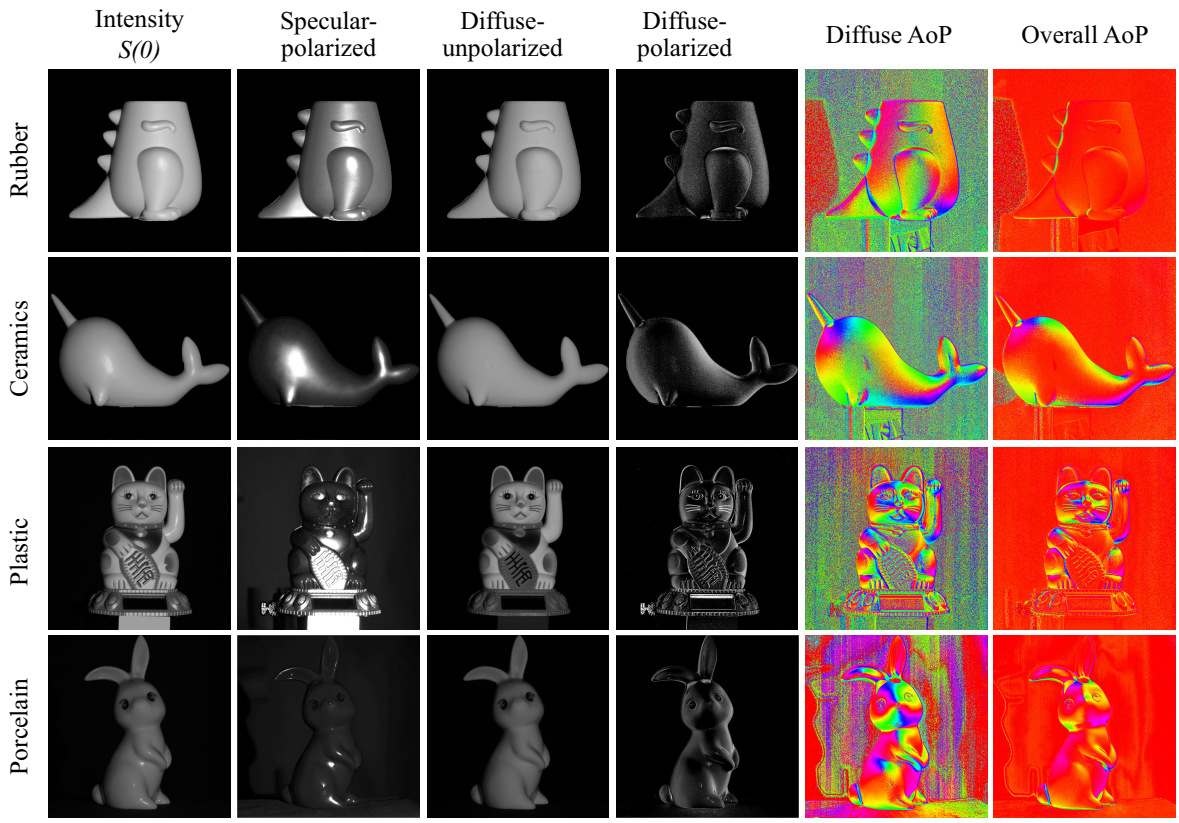


Fig. 7. Results on polarimetric image decomposition.

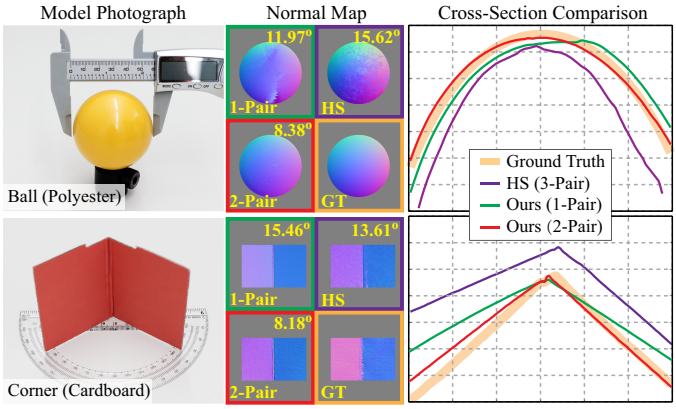


Fig. 8. Quantitative evaluation of the normal and shape reconstruction on real scenes.

50 cm. This distance is chosen based on the camera focal length. The sizes of objects captured are between 5 cm to 25 cm.

System Calibration. Both the camera and projector are calibrated geometrically and radiometrically. For geometric calibration, we measure the intrinsic and extrinsic parameters of the camera and projector [59] in order to extract their relative positions. For radiometric calibration, we compensate for light anisotropy and camera response function using the method of Jankó *et al*. [21]. We also calibrate the light source polarization state with respect to the camera's. Specifically, we place a mirror in front of the acquisition system to allow the camera to capture an image of the light source. We then turn the polarizer in front of the light source and observe the polarization image captured by the camera.

When I^{90} becomes the darkest, we consider the light source's angle of polarization to be $0°$. When I^{0} becomes the darkest, we consider the light source's angle of polarization as $90°$.

Acquisition Procedure. A reciprocal pair is captured by rotating the wheel at $180°$. At each position in the pair, we capture two polarization images under $0°$ and $90°$ polarized light by turning the polarizer in front of the light source. Hence, we have 4 polarization images in a reciprocal pair (captured under two camera positions and there are two lighting conditions in each position). Each polarization image can be decoded into four directional components (*i.e.*, I^{0}, I^{90}, I^{45}, and I^{135}). The total acquisition time of one reciprocal pair is about 2 seconds. Additional reciprocal pairs can be captured by rotating the wheel to a new position and repeating this procedure.

5.2.2 *Polarimetric Image Decomposition*

We first perform polarimetric image decomposition on the captured images. We use the diffuse-polarized and specular-polarized components to compute the polarimetric constraints. The diffuse-unpolarized component is used for computing the reciprocity constraints. Fig. 7 shows the polarimetric decomposition results of objects made with different materials. In each example, we show the intensity image (S_0 before decomposition), the three composed components (*i.e.*, specular-polarized, diffuse-unpolarized, and diffuse-polarized), as well as the angle of polarization (AoP) computed with the diffuse-polarized component ("diffuse AoP") in comparison with the AoP computed with the image before decomposition ("overall AoP"). We can see that the diffuse AoP

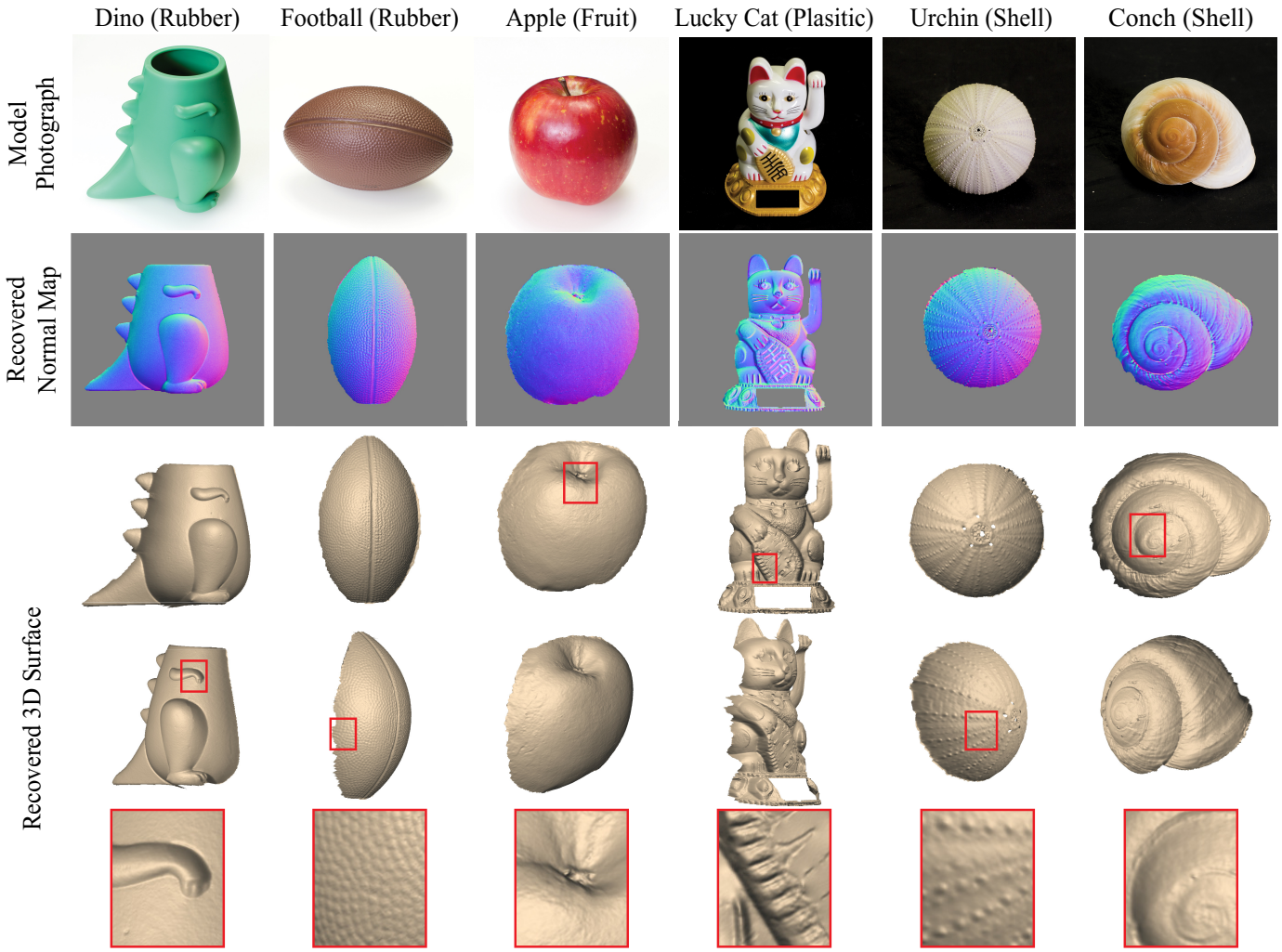


Fig. 9. Surface reconstruction results on real-world objects made with different materials. Here the results are computed using two reciprocal pairs.

Fig. 10. Surface reconstruction results on real scenes. Here the results are computed with two reciprocal pairs.

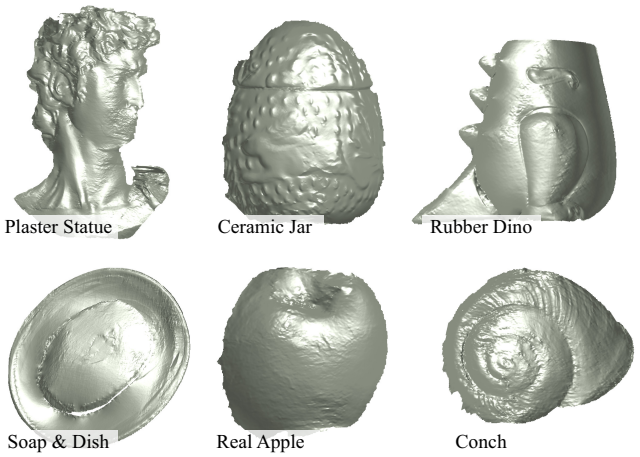


Fig. 11. One-pair reconstruction result on real scenes.

computed with diffuse-polarized component better encodes the surface normal. However, this property varies among materials.

5.2.3 Quantitative Evaluation.

We quantitatively evaluate our method on two real scenes: a billiard ball (59mm diameter) and a cardboard corner (corner angle of 110°). We compare the normal and surface reconstruction of our method with one pair and two pairs, and the standard HS with three pairs. The scene settings and the reconstruction results are shown in Fig. 8. We report the MAE of recovered normals, and compare the recovered shapes in cross-sections.

Our one-pair reconstruction results have large errors at the degenerate case (*i.e.*, $\phi = 90^\circ$). Our two-pair reconstruction results are free from such artifacts and are highly accurate in both scenes (in the cross-section comparisons, our two-pair results are the closest to the ground truth ones). Both the one-pair and two-pair results have smaller errors than the standard HS with three pairs.

5.2.4 Qualitative Evaluation.

We test our method on a variety of scenes with different types of reflectance, ranging from purely diffuse to highly specular. Some are composite scenes that contain multiple surfaces types. As we do not have ground truth geometry for these scenes, we only provide qualitative visual results here.

Fig. 9 shows our normal estimation and surface reconstruction results on objects made with different materials. We also show close-up views of the recovered surfaces to highlight details. We use the two reciprocal pairs for the reconstruction. We can see that our two-pair results are of high quality and our method works well for objects with different levels of specularity. For example, the shells are mostly diffuse and have fine structures on surface, while the plastic cat and apple have strong specularity.

We then show our results on several more challenging scenes. Some of them are composite scenes that contain surfaces of different materials. Some have intricate surface geometry. Fig. 10 shows our surface reconstruction results. We can see that our recovered surfaces preserve fine geometric details (for example, bandage on the statue and beard of the gnome). Our method also works for concave scenes that do not exhibit very strong interreflection (for example, the soap dish and bunny ears). Note that these results are recovered with two reciprocal pairs. Our one-pair reconstruction results are shown in Fig. 11. We can see that the reconstruction at degenerate cases has been largely improved.

Finally, we compare our method against other state-of-the arts 3D reconstruction methods, including shape-from-polarization (SfP) [35], photometric stereo (PS) [11], SfP+PS [41], [60], structured light (SL) [17], and standard HS [1]. Please refer to Table 1 for the number of input images needed for these methods. Fig. 13 show the comparisons of normal and surface reconstructions. In the experiment, we show both our one-pair and two-pair results. Our results visually appear better than all the other methods. Our SL system uses high-resolution projector (1920×1080) and camera (1224×1024) and is able to recover good overall geometry, but our method is able to recover finer surface details than SL. Besides, SL takes 10 input images because of using the Gray code patterns, which is a lot more than the other methods.

In this experiment, we also compare our improved method (referred to as "this work") versus our original method (ICCV '21 [4]). We can see that the one-pair reconstruction results have been improved.

5.2.5 Failure Examples

Here we show two failure examples in Fig. 12. In the "Toy Car" scene, our reconstruction fails at the car windows, which are made of transparent acrylic glass. As its diffuse-unpolarized component is too small, our reciprocity constraints cannot be effectively applied. In the "Bowl" scene, our reconstruction has artifacts at the regions with caustic effect, which is caused by strong inter-reflection. The caustic effect cannot be removed by our polarimetric image decomposition and resulting both the reciprocity constraints and polarimetric constraints to be inaccurate. In the next section, we discuss how to improve our method to handle these challenging cases.

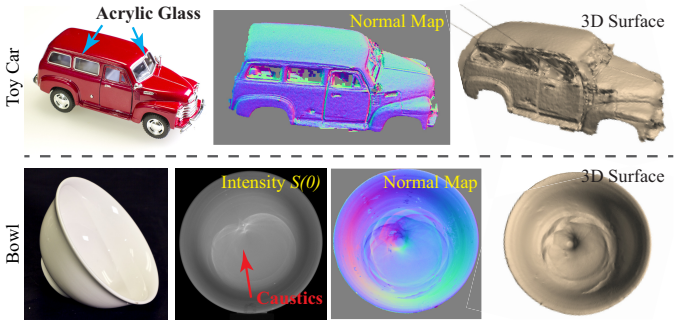


Fig. 12. Two failure examples: a toy car (with transparent glass) and a bowl (with strong inter-reflection).

6 Conclusion and Discussion

In summary, we extend the classical Helmholtz stereopsis to the polarimetric case by deriving a new transpositional reciprocity relationship. We exploit the polarimetric cues and reduce the minimal number of reciprocal pairs to perform HS to one. We show that proposed polar-HS can recover various types of surfaces with high accuracy. In this journal extension, we further generalize our polarimetric constraints to perspective projection and improved our optimization framework for better accuracy.

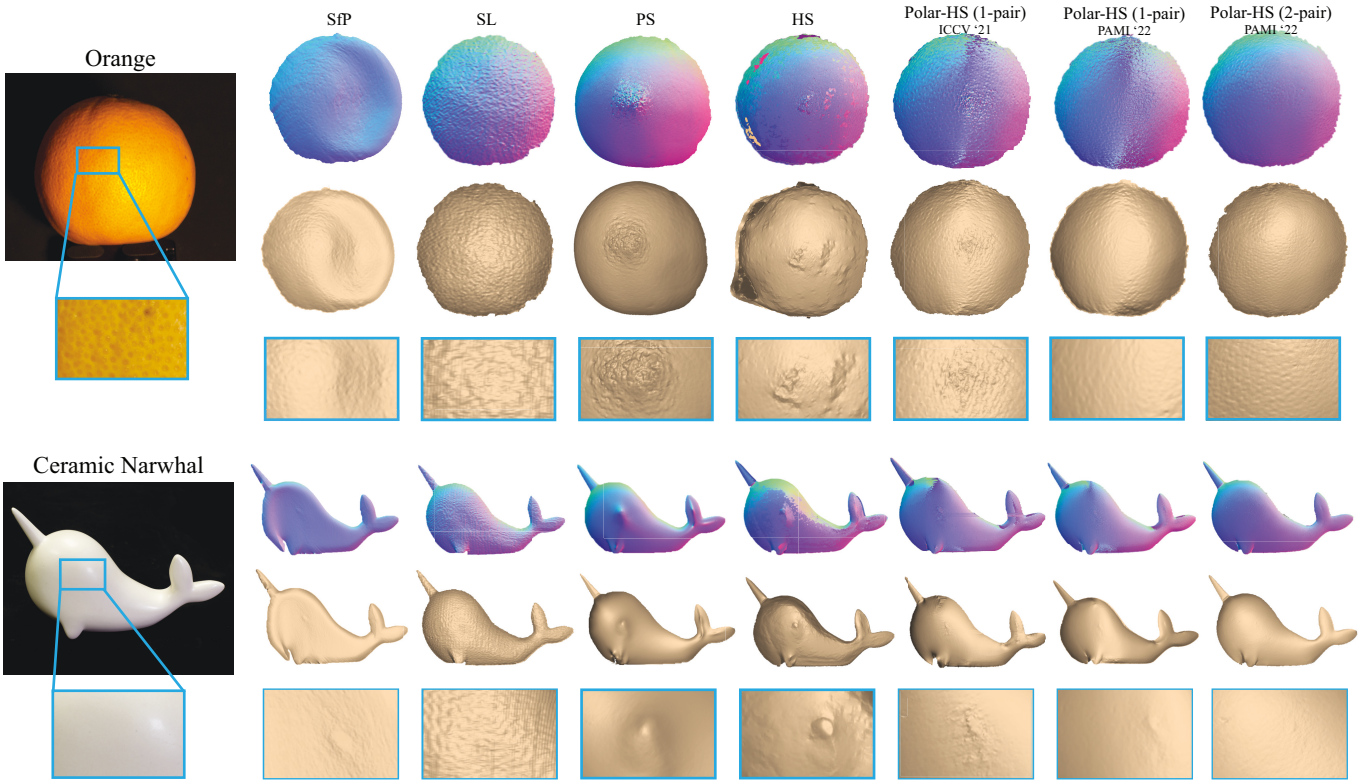


Fig. 13. Comparisons of our method with classical 3D reconstruction methods. We show the recovered normal maps (row one), relit surfaces (row two), and close-up views of the surfaces (row three).

Discussions. Although we have demonstrated successful 3D reconstruction on a variety of scenes, our method has limitations on handling strong inter-reflection and transparent scenes, as shown in the failure examples. If the inter-reflection is too strong and results in caustic effect on the surface, our method fails at the regions where the caustics occurs. There is potential to separate caustics by analyzing the polarization state of reflected light, because the change of polarization state upon each reflection is governed by the Fresnel equations. As for the transparent surfaces, the captured images are transmission dominant. However, our method relies on analyzing the reflected light for surface reconstruction. One possible solution is to separate the weak reflected image from the transmitted one.

Another limitation of our method is that it suffers from large errors at oblique polarization angles when only one reciprocal pair is used. This is because the diffuse polarization constraint becomes degenerate when the angle of polarization is $\pm 90^\circ$. We can combine additional physical constraints (*e.g.*, shading cues or multi-view constraint) to mitigate this limitation, as these other constraints could substitute the polarimetric constraint for solving the normals. Our current prototype system uses a rotating wheel to capture the reciprocal pairs, which has limitation on capturing dynamic scenes. Since our method only needs one reciprocal pair, it is possible to build a compact acquisition system without using a rotating wheel, *e.g.*, in a binocular setting with polarization-multiplexing, in order to make our method more practical.

Acknowledgement

This project is partially supported by NSF awards 2024795, 2225948, 2238141, and 2332542.

References

[1] T. E. Zickler, P. N. Belhumeur, and D. J. Kriegman, "Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction," *International Journal of Computer Vision*, vol. 49, no. 2-3, pp. 215–227, 2002.

[2] T. E. Zickler, J. Ho, D. J. Kriegman, J. Ponce, and P. N. Belhumeur, "Binocular Helmholtz stereopsis," in *IEEE International Conference Computer Vision (ICCV)*, 2003.

[3] P. Tu and P. R. S. Mendonça, "Surface reconstruction via Helmholtz reciprocity with a single image pair," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2003.

[4] Y. Ding, Y. Ji, M. Zhou, S. B. Kang, and J. Ye, "Polarimetric helmholtz stereopsis," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2021.

[5] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, pp. 7–42, 2004.

[6] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2006.

[7] C. Strecha, W. von Hansen, L. Van Gool, P. Fua, and U. Thoennessen, "On benchmarking camera calibration and multi-view stereo for high resolution imagery," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008.

[8] S. Ullman, *The Interpretation of Visual Motion*, 1st ed. MIT Press, 1979.

[9] G. Adiv, "Inherent ambiguities in recovering 3-D motion and structure from a noisy flow field," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 5, pp. 477–489, 1989.

[10] A. Chiuso, R. Brockett, and S. Soatto, "Optimal structure from motion: Local ambiguities and global estimates," *International Journal of Computer Vision*, vol. 39, p. 195–228, 2000.

[11] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Engineering*, vol. 19, no. 1, pp. 139–144, 1980.

[12] R. Basri, D. Jacobs, and I. Kemelmacher, "Photometric stereo with general, unknown lighting," *International Journal of Computer Vision*, vol. 72, p. 239–257, 2007.

[13] T. Higo, Y. Matsushita, and K. Ikeuchi, "Consensus photometric stereo," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

[14] Y. Cui, S. Schuon, D. Chan, S. Thrun, and C. Theobalt, "3D shape scanning with a Time-of-Flight camera," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

[15] F. Heide, L. Xiao, W. Heidrich, and M. B. Hullin, "Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive Time-of-Flight sensors," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.

[16] N. Naik, A. Kadambi, C. Rhemann, S. Izadi, R. Raskar, and S. B. Kang, "A light transport model for mitigating multipath interference in Time-of-Flight sensors," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[17] R. A. Morano, C. Ozturk, R. Conn, S. Dubin, S. Zietz, and J. Nissanov, "Structured light using pseudorandom codes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 322–327, 1998.

[18] M. Gupta, A. Agrawal, A. Veeraraghavan, and S. G. Narasimhan, "Structured light 3D scanning in the presence of global illumination," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.

[19] J. Gu, T. Kobayashi, M. Gupta, and S. K. Nayar, "Multiplexed illumination for scene recovery in the presence of global illumination," in *IEEE International Conference Computer Vision (ICCV)*, 2011.

[20] M. Gupta and S. K. Nayar, "Micro Phase Shifting," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.

[21] Z. Jankó, O. Drbohlav, and R. Šára, "Radiometric calibration of a Helmholtz stereo rig," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2004.

[22] A. Delaunoy, E. Prados, and P. N. Belhumeur, "Towards full 3d Helmholtz stereovision algorithms," in *Asian Conference on Computer Vision (ACCV)*, 2010.

[23] M. Weinmann, R. Ruiters, A. Osep, C. Schwartz, and R. Klein, "Fusing structured light consistency and Helmholtz normals for 3D reconstruction," in *British Machine Vision Conference (BMVC)*, 2012.

[24] H. Mori, R. Köhle, and M. Kamm, "Scene depth profiling using Helmholtz stereopsis," in *European Conference on Computer Vision (ECCV)*, 2016.

[25] Roubtsova, Nadejda and Guillemaut, Jean-Yves, "Bayesian Helmholtz stereopsis with integrability prior," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 9, pp. 2265–2272, 2017.

[26] N. Roubtsova and J.-Y. Guillemaut, "Colour Helmholtz stereopsis for reconstruction of dynamic scenes with arbitrary unknown reflectance," *International Journal of Computer Vision*, vol. 124, no. 1, pp. 18–48, 2017.

[27] D. Miyazaki, R. T. Tan, K. Hara, and K. Ikeuchi, "Polarization-based inverse rendering from a single view," in *IEEE International Conference Computer Vision (ICCV)*, 2003.

[28] G. A. Atkinson and E. R. Hancock, "Recovery of surface orientation from diffuse polarization," *IEEE Transaction on Image Processing*, vol. 15, no. 6, pp. 1653–1664, 2006.

[29] S. Rahmann and N. Canterakis, "Reconstruction of specular surfaces using polarization imaging," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.

[30] O. Morel, F. Meriaudeau, C. Stolz, and P. Gorria, "Polarization imaging applied to 3D reconstruction of specular metallic surfaces," in *SPIE Machine Vision Applications in Industrial Inspection XIII*, vol. 5679, 2005.

[31] G. A. Atkinson and E. R. Hancock, "Shape estimation using polarization and shading from two views," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 11, pp. 2001–2017, 2007.

[32] O. Drbohlav and R. Šára, "Unambiguous determination of shape from photometric stereo with unknown light sources," in *IEEE International Conference Computer Vision (ICCV)*, 2001.

[33] T. T. Ngo, H. Nagahara, and R. Taniguchi, "Shape and light directions from shading and polarization," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[34] C. P. Huynh, A. Robles-Kelly, and E. R. Hancock, "Shape and refractive index recovery from single-view polarisation images," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

[35] Smith, William A. P. and Ramamoorthi, Ravi and Tozza, Silvia, "Linear depth estimation from an uncalibrated, monocular polarisation image," in *European Conference on Computer Vision (ECCV)*, 2016.

[36] W. A. P. Smith, R. Ramamoorthi, and S. Tozza, "Height-from-polarisation with unknown lighting or albedo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 12, pp. 2875–2888, 2019.

[37] G. A. Atkinson and E. R. Hancock, "Multi-view surface reconstruction using polarization," in *IEEE International Conference Computer Vision (ICCV)*, 2005.

[38] L. Chen, Y. Zheng, A. Subpa-Asa, and I. Sato, "Polarimetric three-view geometry," in *European Conference on Computer Vision (ECCV)*, 2018.

[39] D. Zhu and W. A. P. Smith, "Depth from a polarisation + RGB stereo pair," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[40] G. A. Atkinson, "Polarisation photometric stereo," *Computer Vision and Image Understanding*, vol. 160, pp. 158–167, 2017.

[41] S. Tozza, W. A. P. Smith, D. Zhu, R. Ramamoorthi, and E. R. Hancock, "Linear differential constraints for photo-polarimetric height estimation," in *IEEE International Conference Computer Vision (ICCV)*, 2017.

[42] D. Miyazaki, T. Shigetomi, M. Baba, R. Furukawa, S. Hiura, and N. Asada, "Polarization-based surface normal estimation of black specular objects from multiple viewpoints," in *International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, 2012.

[43] Z. Cui, J. Gu, B. Shi, P. Tan, and J. Kautz, "Polarimetric multi-view stereo," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[44] Kadambi, Achuta and Taamazyan, Vage and Shi, Boxin and Raskar, Ramesh, "Polarized 3D: High-quality depth sensing with polarization cues," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[45] A. Kadambi, V. Taamazyan, B. Shi, and R. Raskar, "Depth sensing using geometrically constrained polarization normals," *International Journal of Computer Vision*, vol. 125, no. 1-3, pp. 34–51, 2017.

[46] S.-H. Baek, D. S. Jeon, X. Tong, and M. H. Kim, "Simultaneous acquisition of polarimetric svbrdf and normals," *ACM Transaction on Graphics*, vol. 37, no. 6, Dec. 2018.

[47] Y. Ba, A. Gilbert, F. Wang, J. Yang, R. Chen, Y. Wang, L. Yan, B. Shi, and A. Kadambi, "Deep shape from polarization," in *European Conference on Computer Vision (ECCV)*, 2020.

[48] F. Clarke and D. Parry, "Helmholtz reciprocity: its validity and application to reflectometry," *Lighting Research & Technology*, vol. 17, pp. 1–11, 1985.

[49] H. von Helmholtz, *Physiological Optics*. Optical Society of America, 1924.

[50] S. Chandrasekhar, *Radiative transfer*. Courier Corporation, 2013.

[51] Z. Sekera, "Scattering matrices and reciprocity relationships for various representations of the state of polarization," *Journal of the Optical Society of America*, vol. 56, no. 12, pp. 1732–1740, 1966.

[52] L. B. Wolff and T. E. Boult, "Constraining object features using a polarization reflectance model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 7, pp. 635–657, 1991.

[53] S. K. Nayar, X.-S. Fang, and T. E. Boult, "Separation of reflection components using color and polarization," *International Journal of Computer Vision*, vol. 21, no. 3, pp. 163–186, 1997.

[54] V. Müller, "Elimination of specular surface-reflectance using polarized and unpolarized light," in *European Conference on Computer Vision (ECCV)*, 1996.

[55] D. H. Goldstein, *Polarized light*. CRC press, 2017.

[56] G. Chen, L. He, Y. Guan, and H. Zhang, "Perspective phase angle model for polarimetric 3d reconstruction," in *European Conference on Computer Vision (ECCV)*. Springer, 2022.

[57] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[58] S.-H. Baek, T. Zeltner, H. J. Ku, I. Hwang, X. Tong, W. Jakob, and M. H. Kim, "Image-based acquisition and modeling of polarimetric reflectance," *ACM Transactions on Graphics (Proc SIGGRAPH 2020)*, vol. 39, no. 4, 2020.

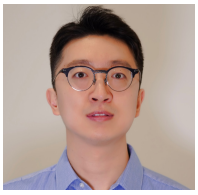
[59] D. Moreno and G. Taubin, "Simple, accurate, and robust projector-camera calibration," in *International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, 2012.

[60] S. Tozza, D. Zhu, W. Smith, R. Ramamoorthi, and E. Hancock, "Uncalibrated, two source photo-polarimetric stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

Yuqi Ding received his M.S. and B.S. degree in Geographical Information System from Wuhan University in 2013 and 2009. He is now a Ph.D. student at the Division of Computer Science and Engineering at Louisiana State University. His research interests include computational photography, computer vision, and computer graphics.

Yu Ji received his PhD degree in Computer & Information Science from the University of Delaware (UDel) in 2015. He received his M.Sc. in Digital Media from Nanyang Technological University (NTU) in 2010 and Bachelor degree in Electrical Engineering from Huazhong University of Science and Technology (HUST) in 2009. He is now a chief scientist at Peau^{x}. His research interests include computational photography, computer vision, and computer graphics.

Zhang Chen received his PhD degree in computer science from ShanghaiTech University in 2022 and his Bachelor degree in electrical and computer engineering from Shanghai Jiao Tong University in 2014. He is currently a visiting scholar at Louisiana State University. His research interests include computer vision, computer graphics and computational photography.



Mingyuan Zhou received the BE degree in intelligence science and technology from the Beijing Information Science and Technology University, China, in 2011, the ME degree in computer engineering from the Stevens Institute of Technology, in 2014, and the PhD degree in computer science from the University of Delaware, in 2019. He is now a senior staff research scientist at Innopeak Technology. His research interests span a range of topics in computer vision and computer graphics, especially computational photography, non-conventional imaging, illumination and reflectance, 3D reconstruction and rendering, and so on.

Sing Bing Kang received his Ph.D. degree in robotics from Carnegie Mellon University in 1994. He is Distinguished Scientist at Zillow Group, and was previously Principal Researcher at Microsoft Corporation. His research interests are computational photography and image-based modeling. Sing Bing has coedited two books ("Panoramic Vision" and "Emerging Topics in Computer Vision") and coauthored two books ("Image-Based Rendering" and "Image-Based Modeling of Plants and Trees"). On the community service front, he has served as Area Chair for the major computer vision conferences and as papers committee member for SIGGRAPH and SIGGRAPH Asia. He was Program Chair for ACCV 2007 and CVPR 2009, and was Associate Editor-In-Chief for IEEE Transactions on Pattern Analysis and Machine Intelligence from 2010-2014. He is a Fellow of the IEEE.

Jinwei Ye is an assistant professor in the Department of Computer Science at George Mason University. She was an assistant professor at Louisiana State University from 2017 to 2021. She received her Ph.D. from the University of Delaware in 2014 and B.Eng. in Electrical Engineering from Huazhong University of Science and Technology in 2009. Before joining academia, she was a researcher in the US Army Research Laboratory and Canon U.S.A., Inc. Her research interests include computational photography, computer vision, and computer graphics.