

Polar-Photometric Stereo Under Natural Illumination

Yuqi Ding¹ Yu Ji² Jinwei Ye^{1,3}

¹Louisiana State University ²Tencent Pixel Lab ³George Mason University

Abstract

We present a 3D shape reconstruction method that leverages both photometric and polarimetric cues. Unlike many active methods that require controlled lighting condition, our method can be used under unknown and uncontrolled natural illumination (both indoor and outdoor). We use two circularly polarized spotlights to boost the polarization cues corrupted by the environment lighting, as well as to provide photometric cues. We solve surface normals with two polarization images by combining the polarimetric and photometric constraints. To mitigate the effect of uncontrolled environment light in photometric constraints, we estimate a lighting proxy map and iteratively refine the normal and lighting estimation. We perform experiments under various natural illumination conditions and compare our results with state-of-the-arts photometric stereo and shape from polarization methods. Our method achieves good accuracy and can be used in flexible environment.

1. Introduction

Both photometric stereo and shape from polarization are vulnerable to environment lighting. Photometric stereo estimates surface normal from images captured under different lighting conditions. As lighting directions need to be known, photometric stereo is usually performed in dark room with calibrated and controlled illumination. Much effort has been made to generalize photometric stereo under uncontrolled environment light [46, 19, 28]. To perform photometric stereo in the wild, the environment light needs to be altered at least three times to provide sufficient photometric constraints. The environment maps of various lighting conditions are usually captured with a light probe for lighting estimation. It's very challenging to perform photometric stereo without knowing the environment light, or with less than three images.

Shape from polarization [27, 7, 37] estimates surface normal with shape-dependent polarimetric cue (*e.g.*, the angle or degree of polarization). One fundamental assumption is that the object is illuminated by completely unpolarized light [7]. That is to say the measured polarization is purely

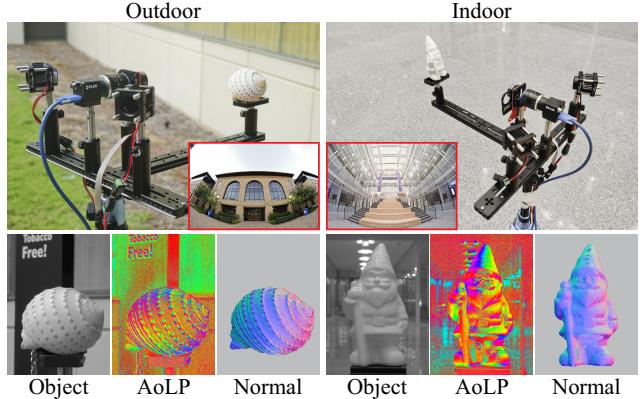


Figure 1. We introduce a method for normal estimation under natural illumination that combines polarimetric and photometric cues.

down to reflection from the object. As result, the AoLP is highly relevant to the surface geometry. Although direct illumination from many light sources (such as the sun, light bulb, *etc.*) is unpolarized, light becomes partially linearly polarized after scattering, reflection and refraction. Therefore, environment lighting usually has linearly polarized components, for instance, indirect illumination from a reflector (such as wall, floor, table top, *etc.*). The linearly polarized light will affect AoLP measurements from the object surface (see examples in Fig. 4), making them unreliable for normal estimation.

In this paper, we present method for estimating normal under unknown and uncontrolled natural illumination that combines the photometric and polarimetric cues (see Fig. 1). To boost the polarization cues corrupted by linearly polarized environment light, we illuminate the object with two controlled light sources (one at a time). We use a snap-shot polarization camera for acquisition. By analyzing the polarimetric reflectance, we show that using circularly polarized source not only restores the normal-dependent AoLP, but also allows the removal of specularity caused by the light source itself.

The two controlled sources also provide photometric constraints. To perform photometric stereo, we pre-calibrate the camera and light sources geometrically. The

Method	Category	Min Input #	Surface Type	Lighting condition	Calibration	Accuracy
Woodham <i>et al.</i> [43]	PS	3	Lambertian	controlled light sources	fully calibrated	High
Hung <i>et al.</i> [19]	PS	3	Lambertian	natural	fully calibrated	Moderate
Mo <i>et al.</i> [28]	PS	10	Lambertian	natural	uncalibrated	Moderate
Smith <i>et al.</i> [38]	SfP	1	Dielectric	natural	uncalibrated	Low
Tozza <i>et al.</i> [41]	PS + SfP	2	Dielectric	controlled light sources	uncalibrated	Moderate
Ours	PS + SfP	2	Dielectric	natural + two known sources	semi-calibrated	High

Table 1. Comparisons between our method and state-of-the-art photometric stereo (PS) and shape from polarization (SfP) methods.

calibration only needs to be performed once as the camera and light sources are rigidly mounted. The environment lighting does not need to be known. Our method thus can be considered as semi-calibrated. By combining with the polarimetric cue, we only need two photometric constraints for normal estimation, which brings down the number of input images to two. To mitigate the effect of uncontrolled environment light in photometric constraints, we estimate a lighting proxy map that emulates the complex environment light in an imaginary dark room. We iteratively refine the normal and lighting estimation until convergence. We perform experiments under various natural lighting conditions (both indoor and outdoor). By comparing our normal estimation results with state-of-the-arts photometric stereo and shape from polarization methods, we show that our method achieves good accuracy that is comparable to photometric stereo with three or more light sources. In addition, our method can be used in flexible environment.

Our technical contributions are summarized as below:

- We present a normal estimation method that combines photometric stereo and shape from polarization, while being applicable under unknown and uncontrolled lighting environment.
- We restore reliable normal-dependent angle of linear polarization by fusing the measurements under two circularly polarized light sources.
- We utilize photometric constraints under unknown environment light by introducing a lighting proxy map.

2. Related Work

Shape from polarization (SfP). This class of methods use shape-dependent polarimetric cues, such as the angle of polarization [7] and degree of polarization [27], for 3D surface reconstruction. One basic assumption is that the surface is illuminated by unpolarized light, such that the polarized light purely comes from surface reflection. Polarimetric features thus can be used for normal estimation. As polarimetric cues are subjective to angular ambiguities, many SfP methods assume additional priors, such as convexity prior [27, 29], smooth prior [32], boundary normal

prior [7], shading cues [8], and multi-spectral measurements [20], for robust normal estimation. Polarimetric cues are often integrated with other classes of methods to improve the reconstruction accuracy, for instance, multi-view stereo [6, 5, 12, 47], Helmholtz stereopsis [16], space carving [26], structure-from-motion [13], and commodity depth sensors [24, 25]. Notably, Smith *et al.* [37, 38] propose a single-image method for shape reconstruction under unknown lighting using polarimetric constraints. However, the method suffers from strong “flattening” artifact (*i.e.*, the recovered surface appears to be flattened) and cannot generalize well to arbitrary environment lighting with high degree of polarization. On the same vein as our approach, a few prior works combine SfP with photometric stereo [17, 31, 40, 5, 41]. However, all these methods require controlled lighting conditions (*e.g.*, in a dark room). In contrast, our approach can be used under natural illumination. Although recent learning-based SfP methods [9, 15] claim to extend SfP to natural environment, these methods require large number of training data and may not be robust to unseen data. Ichikawa *et al.* [21] leverage the polarized sky for SfP, but the polarization pattern of sky needs to be calibrated. In this work, we use controlled light to boost reliable polarimetric features and do not need to calibrate the environment light.

Photometric stereo (PS) in the wild. PS methods use the shading variations under different lighting conditions for normal estimation. Recovering surface normal is the traditional 3D reconstruction method that has been applied in various applications. A fundamental assumption is that the surface reflectance follows the Lambertian model. Classical PS uses controlled, calibrated (both geometrically and photometrically) directional light sources and is usually performed in a dark room [43, 30, 45, 39, 44, 22]. Much efforts have been made to extend PS to uncontrolled natural environment. Some works leverage the natural outdoor illumination change during a day to perform PS [34, 23, 18]. These methods usually take very long acquisition time (*e.g.*, more than a few hours). Some manually alter the environment lighting to create shading variations and use a light probe (*e.g.*, a chrome ball) to calibrate the environment light [46, 19]. Some directly perform PS under uncali-

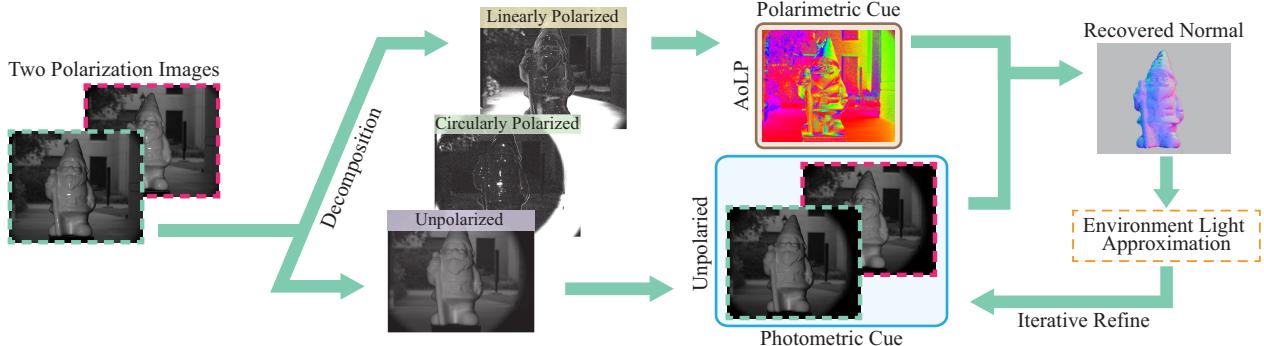


Figure 2. Processing pipeline of our proposed method.

brated natural lighting condition, but these methods need to use parametric lighting model [11, 35], coarse shape prior [35, 4, 36], or more number of images (*e.g.*, over 10) [4, 3, 2, 36] for lighting estimation. In this work, we reduce the number of shading variations to two by incorporating polarimetric cues and we do not require knowing the environment light.

Table 1 compares our method with notable state-of-the-art methods in PS and SfP.

3. In-The-Wild Polar-Photometric Stereo

In this section, we present our method for estimating normal under natural illumination. We first give an overview on our acquisition system and polarization image representation (Sec. 3.1). We then show how to obtain reliable polarimetric and photometric constraints under natural lighting conditions (Sec. 3.2 and 3.3). Finally, we iteratively refine the surface normal via a constrained optimization (Sec. 3.4).

3.1. Method Overview

The overall processing pipeline of method is shown in Fig. 2. Our acquisition setup consists of a polarization camera and two circularly polarized light sources (see Fig. 3). The setup needs a one-time geometrical calibration. The inputs to our method are two polarization images, taken with the controlled lights turned on one at a time under natural environment. By using the extra light sources, we are able obtain normal-dependent angle of linear polarization (or AoLP) map that was downgraded by linearly polarized environment light. As our controlled lights provide photometric parallax, we can formulate two photometric constraints using the Lambertian reflection model. As the overall reflection is a combinatory effect of environment lighting and controlled source, we estimate a lighting proxy map that emulates the complex environment light in the dark to reduce the reflection caused by environment light. We iteratively refine the normal and lighting proxy estimation with a constrained optimization.

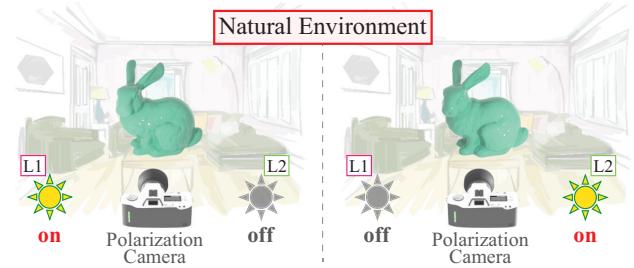


Figure 3. Our setup consists of a polarization camera and two light sources (L1 & L2). We take two images under natural environment by turning on the light sources one at a time.

In sum, our method can estimate surface normal under natural lighting condition with two polarization images. Our method doesn't require calibrating the environment light.

Polarization Images. The inputs to our method are polarization images represented in form of full-Stokes vectors: $S = [S_0, S_1, S_2, S_3]^\top$. S_0 is essentially the intensity image. S_1 to S_3 are parameters with range $[-1, 1]$ (assuming that the intensity value S_0 is normalized). They indicate the state of polarization. Specifically, S_1 specifies the preference of horizontal to vertical linear polarization; S_2 specifies the preference of 45° to -45° linear polarization; and S_3 specifies the preference of right to left circular polarization. The Stokes parameters follow the constraint: $S_0^2 \geq S_1^2 + S_2^2 + S_3^2$ (the equal sign is taken when the light is fully polarized). For linearly polarized light, $S_3 = 0$, while for circularly polarized light, $S_1 = S_2 = 0$. We'll later use these constraints for decomposing the overall image into various polarized components.

3.2. Polarimetric Constraint

Under unpolarized illumination, the polarized reflection, being solely determined by the reflector's surface, is highly relevant to the surface geometry. Specifically, we

use the angle of linear polarization (AoLP) to regularize surface normal. AoLP $\phi \in [0, \pi]$ can be computed as $\phi = (\tan^{-1}(S_2/S_1))/2$. By projecting both the surface normal and AoLP onto the image plane, Smith *et al.* [37] formulate the polarimetric constraint as linear equation. For diffuse reflection, the two projected vectors are collinear. Thus we have

$$[\sin(\phi), -\cos(\phi), 0] \mathbf{n} = 0, \quad (1)$$

where $\mathbf{n} = [n_x, n_y, n_z]^\top$ is surface normal.

For specular reflection, as the AoLP is shifted by 90° , we use $\phi + \pi/2$ in place of ϕ and have

$$[\sin(\phi + \pi/2), -\cos(\phi + \pi/2), 0] \mathbf{n} = 0. \quad (2)$$

As specular reflection is usually brighter and have higher degree of polarization, we use thresholding to separate the diffuse and specular pixels in order to apply their specific polarimetric constraint.

Although direct illumination from natural light sources (such as the sun, light bulb, *etc.*) is unpolarized, light becomes partially linearly polarized after interacting with objects in the scene. Therefore, the environment light usually has linearly polarized components, for instance, indirect illumination from some object (such as wall, floor, table top, *etc.*). As shown in Fig. 4, the linearly polarized environment light largely affect the geometry-dependent AoLP, making the AoLP measurement unreliable for normal estimation.

In order to overcome the effect of linearly polarized environment light, we shine controlled lighting on the object to obtain reliable AoLP. As our two controlled light sources are close to the target object, their reflections are dominant over that of the environment light.

Choice of Light Source. A straight-forward choice of light source would be an unpolarized one, as polarized reflection of unpolarized light is determined by surface geometry. Although being a viable option, we find a better choice is to use circularly polarized light. Same as unpolarized light, circularly polarized light wouldn't affect the geometry-dependent AoLP, as its Stokes parameters on linear polarization are zero (*i.e.*, $S_1 = S_2 = 0$). It has another advantage of being able to remove the specular highlight caused by the light source itself (see details in Polarimetric Image Decomposition). It therefore also benefit the use of photometric constraint. The downside is that circular polarization cannot be directly measured by commercial polarization cameras as they only linear polarization filters. One needs to rotate a retarder in front of the polarization camera to measure the Stokes parameter on circular polarization (*i.e.*, S_3). But with full-Stokes polarization camera being developed [33], single-shot circular polarization measurement can be made possible.

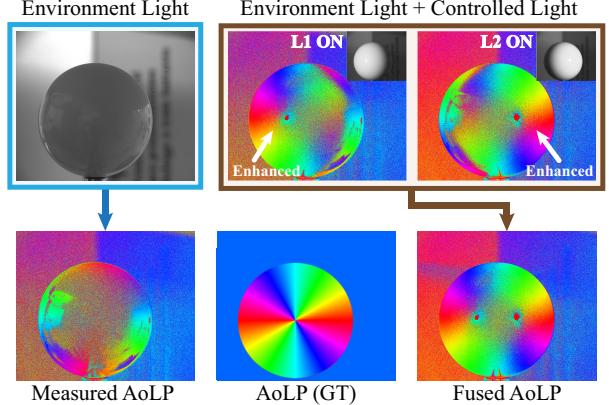


Figure 4. We compare the directly measured AoLP under environment lighting vs. the fused AoLP with our controlled light sources. We show the normal-dependent AoLP (ground truth) in the middle as reference.

AoLP Fusion. As we turn on the controlled light one at a time, we can only boost the AoLP on one side of surface facing towards the light source in the direct measurement. We therefore fuse the two AoLP maps to combine the reliable measurements. The fusion is done by comparing the intensity of the two polarization images and adopt the AoLP of the one with higher intensity value. Fig. 4 compares AoLP map of a sphere obtained in different lighting conditions. We can see AoLP map under environment light (without our controlled light) is downgraded, especially by the ground reflection, which is linearly polarized. By using external lighting, we can boost the AoLP on one side of the surface. Our fused AoLP map apparently combines the reliable AoLPs for the two direct measurements. There are regions inconsistent with the ground truth diffuse AoLP map (calculated by Eq. 1). These are caused by specular reflection of direct light sources. We consider these pixels as specular and use Eq. 2 for polarimetric constraints.

Polarimetric Image Decomposition. We decompose the polarization image into three components according to the polarization state: circularly polarized, linearly polarized, and unpolarized.

$$S = S^c + S^l + S^u, \quad (3)$$

where $S = [S_0, S_1, S_2, S_3]^\top$ is the overall polarization image; $S^c = [S_0^c, 0, 0, S_3^c]^\top$ is the circularly polarized component; $S^l = [S_0^l, S_1^l, S_2^l, 0]^\top$ is the linearly polarized component; and $S^u = [S_0^u, 0, 0, 0]^\top$ is the unpolarized component.

It's easy to see that

$$S_1 = S_1^l, S_2 = S_2^l, \text{ and } S_3 = S_3^c. \quad (4)$$

By applying the intensity constraint of the Stokes parameters, we have

$$S_0^c = |S_3^c| \text{ and } S_0^l = \sqrt{(S_1^l)^2 + (S_2^l)^2}. \quad (5)$$

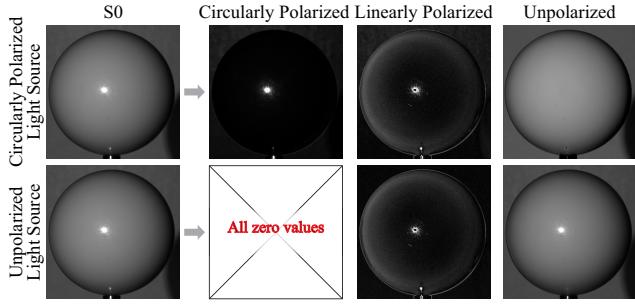


Figure 5. We compare the image decomposition results under circularly polarized lighting and unpolarized lighting. We can see that the unpolarized component under circularly polarized lighting has its own specular reflection removed.

Finally, we compute the unpolarized component as

$$S_0^u = S_0 - S_0^c - S_0^l. \quad (6)$$

As only our controlled light source is circularly polarized, S^c separates the specular reflection from the controlled source. We use the linearly polarized component S^l to compute AoLP and DoLP. The unpolarized component S^u is used for applying the photometric constraints, since the specular reflection is largely reduced by separating S^c and S^l . Fig. 5 compares the image decomposition results under circularly polarized source vs. unpolarized one. We can see that both unpolarized components have reduced specularity from the environment light. But circularly polarized source also has its own specular reflection removed by separating S^c .

3.3. Photometric Constraint

As the two controlled sources provide photometric parallax, we can also use photometric constraint derived from the Lambertian reflection model:

$$I = \rho E(\mathbf{n} \cdot \mathbf{l}), \quad (7)$$

where I is the intensity of reflection; ρ is the surface albedo; E is the light intensity; \mathbf{n} is the surface normal; and \mathbf{l} is the lighting direction.

Given a single calibrated light source, Eq. 7 can be directly applied to regularize the surface normal. Classical photometric stereo solves surface normal with three such equations established under three different lighting directions. In our problem, although our controlled light sources provide lighting variations, their intensities are mixed the environment light. Specifically, we use the unpolarized component as the image intensity: $S_0^u = I$. The overall intensity I can be considered as a combination from two types of sources: our controlled light source and the environment light.

$$I = I^c + I^e = \rho E^c(\mathbf{n} \cdot \mathbf{l}^c) + I^e, \quad (8)$$

where I^c is the reflection intensity from the known controlled light source (which can be directly modeled using Eq. 7 with lighting direction \mathbf{l}^c and intensity E^c); I^e is the reflection intensity from the unknown environment light. Then our goal is to deduct I^e from I , such that we can use the photometric constraint for normal estimation. As we also have the polarimetric constraint, the two photometric constraints provided by our controlled light sources are sufficient for solving normal. In the following, we describe how to estimate and reduce I^e .

Lighting Proxy Map. We model the environment light on a half hemi-sphere (we only consider the environment that is towards the front face of the object). Our lighting proxy map contains the light intensity uniformly sampled on the half hemi-sphere. Given the azimuth angle $\psi \in [0, \pi]$ and elevation angle $\varphi \in [0, \pi]$, we convert the spherical coordinate (ψ_i, φ_i) to lighting direction in Cartesian coordinate as $\mathbf{l}_i^e = [\cos(\varphi_i) \cos(\psi_i), \sin(\varphi_i), \cos(\varphi_i) \sin(\psi_i)]^\top$. For each surface point, we compute the reflection of environment light by integrating the light intensity from all directions. Therefor we have

$$I^e = \rho \sum_{i=1}^M E_i^e (\mathbf{n} \cdot \mathbf{l}_i^e), \quad (9)$$

where E_i^e is sampled light intensity in the lighting proxy map with i being the pixel index in the map and M the total number of pixels (we use $M = 1296$ in our experiments).

By combining Eq. 8 and Eq. 9, we can formulate the following linear equation to solve the lighting proxy map $\{E_i^e | i = 1, \dots, M\}$.

$$\rho \begin{bmatrix} \mathbf{n} \cdot \mathbf{l}_1^e & \mathbf{n} \cdot \mathbf{l}_2^e & \dots & \mathbf{n} \cdot \mathbf{l}_M^e \end{bmatrix} \begin{bmatrix} E_1^e \\ E_2^e \\ \dots \\ E_M^e \end{bmatrix} = I - \rho E^c (\mathbf{n} \cdot \mathbf{l}^c). \quad (10)$$

We first estimate a coarse normal map and albedo directly using the mixed intensity I and them plug them into Eq. 10. We formulate such equation for each pixel and stack them together to solve $\{E_i^e | i = 1, \dots, M\}$. However, it's undesirable to use all pixels, as the coarse normal map is highly inaccurate. We therefore only using those pixels with good normal estimations for solving $\{E_i^e | i = 1, \dots, M\}$. Next, we show how to use the degree of linear polarization (DoLP) to guide the selection.

DoLP as Confidence Map. According to [38], the DoLP of diffuse polarization can be modeled as

$$d = \frac{(n - 1/n)^2 \sin^2 \theta_r}{2 + 2n^2 - (n + 1/n)^2 \sin^2 \theta_r + 4 \cos \theta_r \sqrt{n^2 - \sin^2 \theta_r}}, \quad (11)$$

where n is the refractive index (we use $n = 1.5$) and θ_r is the angle of reflection. Given surface normal \mathbf{n} and viewing direction \mathbf{v} , $\theta_r = \arccos(\mathbf{n} \cdot \mathbf{v})$. DoLP is, therefore, modeled as a function of surface normal. As the DoLP measurement is less affected by various lighting conditions, it provides us guidance to select good normal estimations. Specifically, we compute a binary confidence map by comparing the DoLP computed with Eq. 11 (given the coarse normal estimation) with the DoLP computed from the polarization image ($d = \sqrt{S_1^2 + S_2^2}/S_0$). The value of the confidence map is computed as

$$C = \begin{cases} 1, & |d - \tilde{d}| < \epsilon, \\ 0, & \text{otherwise.} \end{cases} \quad (12)$$

Here d is the DoLP directly computed from the polarization image; \tilde{d} is the DoLP computed with Eq. 11 given surface normal; and ϵ is a similarity threshold. We then only use those normals whose confidence values are 1 for solving the lighting proxy map. By eliminating inaccurate normals, we can have better lighting estimation.

3.4. Optimization

We solve normal by combining the polarimetric and photometric cues. For polarimetric cue, we use Eq. 1 or Eq. 2 depending on type of reflection. For photometric cues, we reduce the effect of environment light using the estimated lighting proxy map. Specifically, we rewrite Eq. 9 as

$$\begin{aligned} I^e &= \rho E^e (\mathbf{n} \cdot \mathbf{l}^e), \\ \text{with } E^e \mathbf{l}^e &= [\sum_{i=1}^M E_i^e \mathbf{l}_{i,x}^e, \sum_{i=1}^M E_i^e \mathbf{l}_{i,y}^e, \sum_{i=1}^M E_i^e \mathbf{l}_{i,z}^e]^\top. \end{aligned} \quad (13)$$

$E^e = \|E^e \mathbf{l}^e\|$ can be considered as a weighted sum of environment light according to the lighting directions. \mathbf{l}^e is a unit vector and can be considered as the lighting direction with the environment light being mapped to a single source.

By substituting Eq. 13 into Eq. 8, we have

$$\begin{aligned} I &= \rho E^c (\mathbf{n} \cdot \mathbf{l}^c) + \rho E^e (\mathbf{n} \cdot \mathbf{l}^e), \\ &= \rho E^c (\mathbf{n} \cdot \mathbf{l}^c + \beta \mathbf{n} \cdot \mathbf{l}^e). \end{aligned} \quad (14)$$

$\beta = E^e/E^c$ is the intensity ratio between the overall environment light and our controlled light source. By combining two photometric constraints and polarimetric constraint (here we use the diffuse case), we can formulate the following linear system $A\mathbf{x} = b$:

$$\begin{bmatrix} \mathbf{l}_x^{c1} + \beta \mathbf{l}_x^e & \mathbf{l}_y^{c1} + \beta \mathbf{l}_y^e & \mathbf{l}_z^{c1} + \beta \mathbf{l}_z^e \\ \mathbf{l}_x^{c2} + \beta \mathbf{l}_x^e & \mathbf{l}_y^{c2} + \beta \mathbf{l}_y^e & \mathbf{l}_z^{c2} + \beta \mathbf{l}_z^e \\ \sin(\phi) & -\cos(\phi) & 0 \end{bmatrix} \begin{bmatrix} \gamma n_x \\ \gamma n_y \\ \gamma n_z \end{bmatrix} = \begin{bmatrix} I_1 \\ I_2 \\ 0 \end{bmatrix}, \quad (15)$$

where I_1 and I_2 are the unpolarized component of the two input polarization images (each captured under one controlled light source); \mathbf{l}^{c1} and \mathbf{l}^{c2} are the directions of the

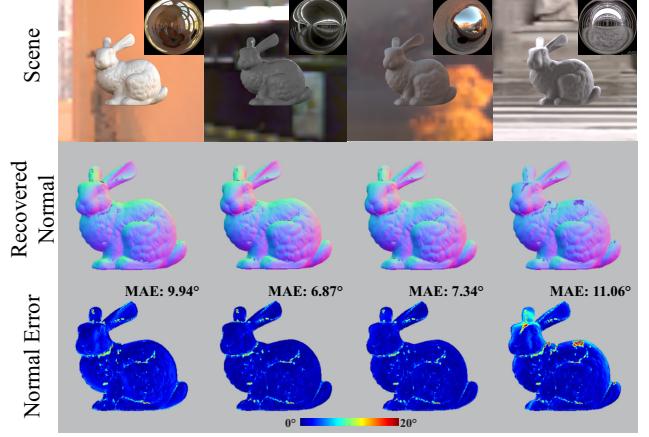


Figure 6. Normal estimation results on synthetic data.

controlled sources; $\gamma = \rho E_c$ combines the lighting intensity and surface albedo. Note that γ is estimated along with the normal \mathbf{n} . Once \mathbf{x} is solved, we have $\gamma = \|\mathbf{x}\|$ and $\mathbf{n} = \mathbf{x}/\|\mathbf{x}\|$. We formulate the linear system for each surface point and solve them together via the following objective function:

$$\operatorname{argmin}_{\{\mathbf{x}_j\}, \beta} \sum_{j=1}^N |A_j \mathbf{x}_j - b_j| + \lambda \sum_{j=1}^N |1 - \mathbf{x}_j \cdot \mathcal{N}(\mathbf{x}_j)|, \quad (16)$$

where $|\cdot|$ is the L1 norm; j is the surface point index; N is the total number of surface points; λ is a term-balancing factor (we use $\lambda = 0.04$); and $\mathcal{N}(\cdot)$ takes the four nearest neighbors of its input. The first term is the data term and the second is a smoothing term.

4. Experiments

We perform experiments on both synthetic and real data to evaluate our method. For synthetic experiments, we focus on ablation study of various influencing factors. For real experiments, we demonstrate that our method works in various indoor and outdoor environment.

4.1. Synthetic Experiments

Data Simulation. We use the Mitsuba 2 renderer [1] to simulate polarized images. Specifically, we render images with the polarized rendering mode. We use the KAIST pBRDF dataset [10] to model the polarimetric surface reflectance. We directly render images in form of Stokes vectors. Each Stokes component has resolution 500×500 . The camera center is at origin. Directions of the two controlled light sources are $\mathbf{l}^{c1} = [-0.18, 0.03, 0.98]^\top$ and $\mathbf{l}^{c2} = [0.16, 0.03, 0.98]^\top$. We use environment maps from [14] and [42] to simulate the natural environment.

Fig. 6 shows our normal estimation results under various environment maps. We use two different materials for the

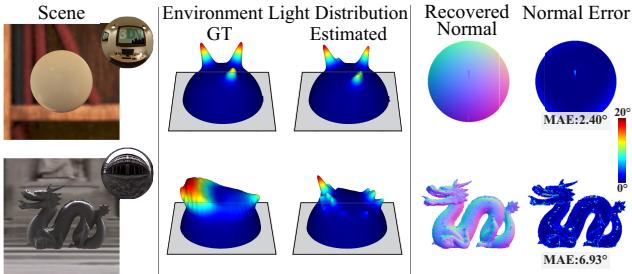


Figure 7. Results on environment light estimation. Left: Rendered image and environment map. Middle: Environment Light Distribution (ground truth vs. our estimated). Right: Normal estimation and its error.

Method \ β	10	5	1	0.67	0.5	0
Smith <i>et al.</i> [38]	36.89	29.63	28.56	28.16	27.81	16.98
Tozza <i>et al.</i> [41]	41.37	38.68	37.86	31.59	24.55	21.59
Ours	16.25	11.53	7.51	5.42	4.26	3.59

Table 2. Mean angular error (in degree) w.r.t. intensity ratio β .

object: “white billiard” (column 1 & 2) and “spectralon” (column 3 & 4). We set the ratio between overall environment light and our controlled light $\beta = 0.5$. The normal is estimated using two polarization images as input (with the controlled light source turned on one at a time). We compare the estimated normal with the ground truth normal and show the per-pixel normal error (in degree) and the mean angular error (MAE). Results on more environment maps and objects can be found in the supplementary material.

Fig. 7 shows our environment light estimation result. Note that our lighting proxy map is only an approximation of the environment map. We cannot recover a high fidelity environment, but our estimated lighting proxy map has consistent lighting directions as the ground truth environment map at least in the regions with high lighting intensity. Such environment light approximation is sufficient for reliable normal estimation.

Ablation on β . As the ratio between overall environment light and our controlled light β is important to the performance of our method, we perform ablation study on this parameter. Specifically, we test on 5 β values between 0.5 to 10. The higher the β value, the stronger the environment light. $\beta = 0$ indicates no environment light (*i.e.*, in a dark room). We test on the Bunny model with “white billiard” material. Normal estimation error (MAE) is shown in Table 2. We can see that the performance of our method downgrades as environment light gets stronger. We can obtain reasonable normal estimation under environment light 10 times of the controlled light source.

We compare the results with two SfP methods. Smith *et al.* [38] is purely polarization-based. It takes in one po-

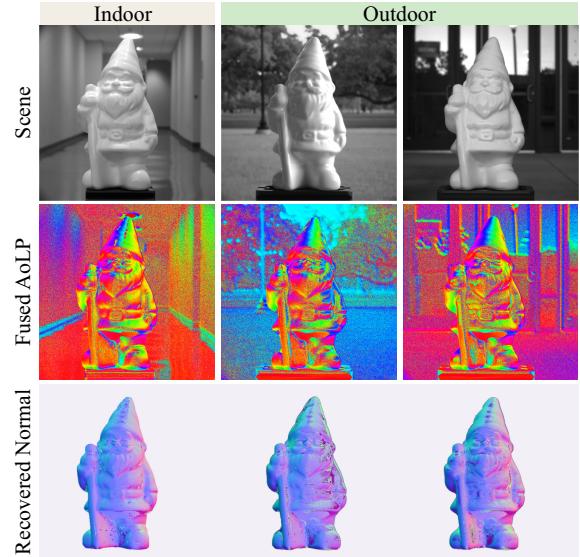


Figure 8. Real results under different environment.

larization as input. The method is claimed to be applicable under natural illumination. Tozza *et al.* [41] combines PS and SfP. Similar to our method, it takes in two polarization images. But the method requires controlled lighting condition. Note that we did not compare with PS methods, as they need three or more images. We can see that both methods have much larger errors than ours, even without environment light (since these methods are uncalibrated).

4.2. Real Experiments

Real Setup. We build a portable acquisition setup to perform real experiments (see Fig.1). Specifically, we use a monochrome polarization camera and two 36V LED lights. The luminous flux of our light source is around 1300 lumen. We mount right-handed circular polarization filter in front of each light source to generate circularly polarized light. We need to rotate a quarter wave retarder in front of the camera to capture the full-Stokes vectors. More details on data acquisition can be found in the supplementary materials. Our setup needs to be calibrated once, so we know the relative position between the camera and light sources. The object is around 50 cm away from the camera. We perform experiments under various indoor and outdoor environment.

Fig. 8 shows our normal estimation results of a same object (“Gnome”) under different environment. The indoor environment is a bright hallway. The outdoor scenes are captured in the afternoon (around 5pm). We can see that our method works well under various natural illumination. We can see that the environment light is highly polarized, which affects the normal dependent AoLP (see the supplementary material for AoLP directly measured under environment light). Our fused AoLP map largely boost the polarimetric cue to allow reliable normal estimation.

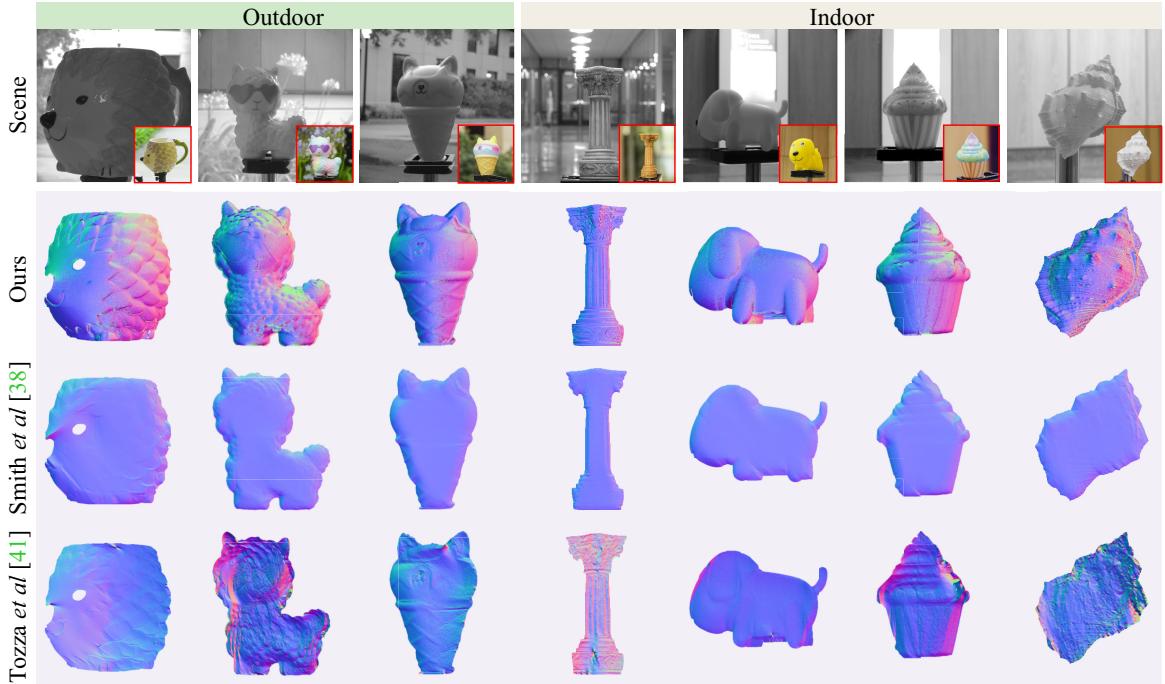


Figure 9. Real results in comparison with state-of-the-art SfP methods.



Figure 10. Surface reconstruction results.

Fig. 9 shows normal estimation results of different objects under various environment. We also compare our results with the two SfP methods ([38] and [41]). We can see that the results of [38] have strong “flattening” artifact, although estimations at boundary regions are reasonable. Results of [41] are highly inaccurate due to environment lighting. Their method assumes single directional light for photometric constraints. We have also compared our method against different versions of photometric stereo. More comparison results on real data can be found in the supplementary material.

We also perform surface integration on our estimated normal maps. Fig. 10 shows the recovered surfaces with the real objects as reference.

5. Conclusions & Discussions

We have presented a normal estimation method by combining polarimetric and photometric cues. Our method can be used under natural illumination. We setup circularly polarized light to provide photometric constraints, as well as enhance the polarimetric features corrupted by linearly polarized environment light. To mitigate the effect of environment light in photometric constraints, we estimate a lighting proxy map using a coarse normal map. We then iteratively refine the normal and lighting estimation. We have demonstrated that our method can be used in various indoor and outdoor environment for reliable normal estimation.

Limitations. As the ratio between the overall environment light and our controlled light is an influencing factor, our method can work in most indoor environment, but it does not work well in some outdoor scenes when the environment light is too bright (*e.g.*, noon time on a sunny day). We can integrate our method with [21] that directly use the polarization state of sky light as constraint. Another limitation is that we need to use a retarder in front of the camera in order to measure the circular polarization. As result, our method actually takes in four images, instead of two. This issue can be resolved with novel polarization sensor that allows single-shot full-Stokes measurement [33].

Acknowledgements

This project is supported by NSF awards CRII-1948524 and NRI-2024795.

References

- [1] Mitsuba 2, <https://mitsuba2.readthedocs.io/en/latest/index.html>.
- [2] Austin Abrams, Christopher Hawley, and Robert Pless. Helioimetric stereo: Shape from sun position. In *European Conference on Computer Vision (ECCV)*, pages 357–370. Springer, 2012.
- [3] Jens Ackermann, Fabian Langguth, Simon Fuhrmann, and Michael Goesele. Photometric stereo for outdoor webcams. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 262–269, 2012.
- [4] Jens Ackermann, Martin Ritz, André Stork, and Michael Goesele. Removing the example from example-based photometric stereo. In *European Conference on Computer Vision (ECCV)*, pages 197–210. Springer, 2010.
- [5] Gary A. Atkinson. Polarisation photometric stereo. *Computer Vision and Image Understanding*, 160:158–167, 2017.
- [6] Gary A. Atkinson and Edwin R. Hancock. Multi-view surface reconstruction using polarization. In *IEEE International Conference Computer Vision (ICCV)*, 2005.
- [7] Gary A. Atkinson and Edwin R. Hancock. Recovery of surface orientation from diffuse polarization. *IEEE Transaction on Image Processing*, 15(6):1653–1664, 2006.
- [8] Gary A. Atkinson and Edwin R. Hancock. Shape estimation using polarization and shading from two views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(11):2001–2017, 2007.
- [9] Yunhao Ba, Alex Gilbert, Franklin Wang, Jinfa Yang, Rui Chen, Yiqin Wang, Lei Yan, Boxin Shi, and Achuta Kadambi. Deep shape from polarization. In *European Conference on Computer Vision (ECCV)*, 2020.
- [10] Seung-Hwan Baek, Tizian Zeltner, Hyun Jin Ku, Inseung Hwang, Xin Tong, Wenzel Jakob, and Min H. Kim. Image-based acquisition and modeling of polarimetric reflectance. *ACM Transaction on Graphics*, 39(4), 2020.
- [11] Ronen Basri, David Jacobs, and Ira Kemelmacher. Photometric stereo with general, unknown lighting. *International Journal of Computer Vision*, 72:239–257, 2007.
- [12] Lixiong Chen, Yinqiang Zheng, Art Subpa-Asa, and Imari Sato. Polarimetric three-view geometry. In *European Conference on Computer Vision (ECCV)*, 2018.
- [13] Zhaopeng Cui, Jinwei Gu, Boxin Shi, Ping Tan, and Jan Kautz. Polarimetric multi-view stereo. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [14] Paul Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *Acm siggraph 2008 classes*, pages 1–10. 2008.
- [15] Valentin Deschaintre, Yiming Lin, and Abhijeet Ghosh. Deep polarization imaging for 3d shape and svbrdf acquisition. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 15567–15576, 2021.
- [16] Yuqi Ding, Yu Ji, Mingyuan Zhou, Sing Bing Kang, and Jinwei Ye. Polarimetric helmholtz stereopsis. In *IEEE/CVF International Conference Computer Vision (ICCV)*, pages 5037–5046, 2021.
- [17] Ondřej Drbohlav and Radim Šára. Unambiguous determination of shape from photometric stereo with unknown light sources. In *IEEE International Conference Computer Vision (ICCV)*, 2001.
- [18] Yannick Hold-Geoffroy, Jinsong Zhang, Paulo FU Gotardo, and Jean-François Lalonde. X-hour outdoor photometric stereo. In *IEEE International Conference on 3D Vision*, pages 28–36, 2015.
- [19] Chun-Ho Hung, Tai-Pang Wu, Yasuyuki Matsushita, Li Xu, Jiaya Jia, and Chi-Keung Tang. Photometric stereo in the wild. In *2015 IEEE Winter Conference on Applications of Computer Vision*, pages 302–309, 2015.
- [20] Cong Phuoc Huynh, Antonio Robles-Kelly, and Edwin R. Hancock. Shape and refractive index recovery from single-view polarisation images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.
- [21] Tomoki Ichikawa, Matthew Purri, Ryo Kawahara, Shohei Nobuhara, Kristin Dana, and Ko Nishino. Shape from sky: Polarimetric normal recovery under the sky. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14832–14841, June 2021.
- [22] Satoshi Ikehata, David Wipf, Yasuyuki Matsushita, and Kiyoharu Aizawa. Robust photometric stereo using sparse regression. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 318–325, 2012.
- [23] Jiyoung Jung, Joon-Young Lee, and In So Kweon. One-day outdoor photometric stereo via skylight estimation. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4521–4529, 2015.
- [24] Achuta Kadambi, Vage Taamazyan, Boxin Shi, and Ramesh Raskar. Polarized 3D: High-quality depth sensing with polarization cues. In *IEEE International Conference Computer Vision (ICCV)*, 2015.
- [25] Achuta Kadambi, Vage Taamazyan, Boxin Shi, and Ramesh Raskar. Depth sensing using geometrically constrained polarization normals. *International Journal of Computer Vision*, 125(1-3):34–51, 2017.
- [26] Daisuke Miyazaki, Takuya Shigetomi, Masashi Baba, Ryo Furukawa, Shinsaku Hiura, and Naoki Asada. Polarization-based surface normal estimation of black specular objects from multiple viewpoints. In *International Conference on 3D Imaging, Modeling, Processing, Visualization & Transmission*, 2012.
- [27] Daisuke Miyazaki, Robby T. Tan, Kenji Hara, and Katsushi Ikeuchi. Polarization-based inverse rendering from a single view. In *IEEE International Conference Computer Vision (ICCV)*, 2003.
- [28] Zhipeng Mo, Boxin Shi, Feng Lu, Sai-Kit Yeung, and Yasuyuki Matsushita. Uncalibrated photometric stereo under natural illumination. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2936–2945, 2018.
- [29] Olivier Morel, Fabrice Meriaudeau, Christophe Stoltz, and Patrick Gorria. Polarization imaging applied to 3D reconstruction of specular metallic surfaces. In *SPIE Machine Vision Applications in Industrial Inspection XIII*, volume 5679, 2005.
- [30] Yasuhiro Mukaigawa, Yasunori Ishii, and Takeshi Shunkaga. Analysis of photometric factors based on photometric linearization. *JOSA A*, 24(10):3326–3334, 2007.

- [31] Trung Thanh Ngo, Hajime Nagahara, and Rin-ichiro Taniguchi. Shape and light directions from shading and polarization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [32] Stefan Rahmann and Nikos Canterakis. Reconstruction of specular surfaces using polarization imaging. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2001.
- [33] Noah A Rubin, Gabriele D’Aversa, Paul Chevalier, Zhujun Shi, Wei Ting Chen, and Federico Capasso. Matrix fourier optics enables a compact full-strokes polarization camera. *Science*, 365(6448):eaax1839, 2019.
- [34] Fangyang Shen, Kalyan Sunkavalli, Nicolas Bonneel, Szymon Rusinkiewicz, Hanspeter Pfister, and Xin Tong. Time-lapse photometric stereo and applications. In *Computer Graphic Forum*, volume 33, pages 359–367. Wiley Online Library, 2014.
- [35] Li Shen and Ping Tan. Photometric stereo and weather estimation using internet images. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1850–1857, 2009.
- [36] Boxin Shi, Kenji Inose, Yasuyuki Matsushita, Ping Tan, Sai-Kit Yeung, and Katsushi Ikeuchi. Photometric stereo using internet images. In *IEEE International Conference on 3D Vision*, volume 1, pages 361–368, 2014.
- [37] William A. P. Smith, Ravi Ramamoorthi, and Silvia Tozza. Linear depth estimation from an uncalibrated, monocular polarisation image. In *European Conference on Computer Vision (ECCV)*, 2016.
- [38] William A. P. Smith, Ravi Ramamoorthi, and Silvia Tozza. Height-from-polarisation with unknown lighting or albedo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(12):2875–2888, 2019.
- [39] Kalyan Sunkavalli, Todd Zickler, and Hanspeter Pfister. Visibility subspaces: Uncalibrated photometric stereo with shadows. In *European Conference on Computer Vision (ECCV)*, pages 251–264. Springer, 2010.
- [40] Silvia Tozza, William A. P. Smith, Dizhong Zhu, Ravi Ramamoorthi, and Edwin R. Hancock. Linear differential constraints for photo-polarimetric height estimation. In *IEEE International Conference Computer Vision (ICCV)*, 2017.
- [41] Silvia Tozza, Dizhong Zhu, William Smith, Ravi Ramamoorthi, and Edwin Hancock. Uncalibrated, two source photo-polarimetric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.
- [42] Bernhard Vogl. Light probes, <http://dativ.at/lightprobes/>.
- [43] Robert J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1):139–144, 1980.
- [44] Lun Wu, Arvind Ganesh, Boxin Shi, Yasuyuki Matsushita, Yongtian Wang, and Yi Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Asian Conference on Computer Vision (ACCV)*, pages 703–717. Springer, 2010.
- [45] Tai-Pang Wu and Chi-Keung Tang. Photometric stereo via expectation maximization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(3):546–560, 2009.
- [46] Lap-Fai Yu, Sai-Kit Yeung, Yu-Wing Tai, Demetri Terzopoulos, and Tony F Chan. Outdoor photometric stereo. In *IEEE International Conference on Computational Photography (ICCP)*, pages 1–8, 2013.
- [47] Dizhong Zhu and William A. P. Smith. Depth from a polarisation + RGB stereo pair. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.