

Shape and Light Directions from Shading and Polarization

Trung Thanh Ngo * Hajime Nagahara Rin-ichiro Taniguchi
Faculty of Information Science and Electrical Engineering, Kyushu University

Abstract

We introduce a method to recover the shape of a smooth dielectric object from polarization images taken with a light source from different directions. We present two constraints on shading and polarization and use both in a single optimization scheme. This integration is motivated by the fact that photometric stereo and polarization-based methods have complementary abilities. The polarization-based method can give strong cues for the surface orientation and refractive index, which are independent of the light direction. However, it has ambiguities in selecting between two ambiguous choices of the surface orientation, in the relationship between refractive index and zenith angle (observing angle), and limited performance for surface points with small zenith angles, where the polarization effect is weak. In contrast, photometric stereo method with multiple light sources can disambiguate the surface orientation and give a strong relationship between the surface normals and light directions. However, it has limited performance for large zenith angles, refractive index estimation, and faces the ambiguity in case the light direction is unknown. Taking their advantages, our proposed method can recover the surface normals for both small and large zenith angles, the light directions, and the refractive indexes of the object. The proposed method is successfully evaluated by simulation and real-world experiments.

1. Introduction

Three-dimensional (3D) reconstruction is a key research topic in computer vision, for which various approaches have been proposed. Fundamental research works rely on the basic physical parameters of light such as the speed [10], direction [18], intensity shading [20], frequency [13], and polarization state [3] to reconstruct the 3D shape. For instance, the photometric stereo method observes an object from different light directions while the polarization-based method observes the object in different polarization states through a polarizer. Naturally, when using more parameters of light, there is more information with which to recover

the shape. In our research, we focus on the direction and polarization state of light to recover the 3D shape.

Photometric stereo method has been studied intensively and has provided very promising results recently [5, 12, 14, 15]. Conventional photometric stereo methods consider light from different directions interacting with the object surface to sense the 3D structure. In most photometric stereo methods, the light directions are known for the reconstruction. The bidirectional reflection distribution function (BRDF), which models the incoming and outgoing light, plays an important role; however, BRDF is a complex function that depends on the material. Many efforts have been made to overcome the complexity of the BRDF [5, 12, 15], yet these methods usually ignore the refractive index and polarization of the dielectric materials and hence they do not handle polarization images captured with a polarizer. Moreover, the photometric stereo approach usually faces a limitation at the edge with a large observing angle, where BRDF is not well-modeled.

Polarization-based shape reconstruction methods consider polarization states of light reflected on the material surface to sense the 3D structure. The Fresnel equations on refraction and reflection [9] in the interaction of light with a smooth dielectric surface are the main theory underlying this approach. Two types of reflection are considered in this approach: specular and diffuse reflections. In case of specular reflection, we need to obtain the specular reflection for the whole object, and therefore require many light sources which limits the practical use. In case of diffuse reflection, we do not have to use many light sources. In contrast to the photometric stereo methods, the polarization-based methods do not require manipulation of the light direction and basically does not use the information on the light source. However, the polarization-based methods only work well at the edge of an object where the polarization effect is strong. The polarization-based approach usually requires us to know the refractive index [1, 6] or use a homogeneous material with a uniform refractive index [23]. Due to these limitations, a polarization-based method is better to be combined with another method to improve the accuracy. In fact, polarization-based methods are combined with multiview stereo [2] or photometric stereo [3] to improve

*Email: trungbeo@gmail.com

the reconstruction. Although the photometric stereo and polarization-based methods have been combined, they have simply been employed as two independent steps [3, 19].

Under unknown light direction, the photometric stereo faces a serious ambiguity problem, such as the generalized bas-belief ambiguity [4], and needs complicated priors to solve it. However, since the surface normals and light directions are strongly correlated, the ambiguity can be significantly relaxed by constraining all the surface normals. Fortunately, we can use such prior from polarization of reflected light on the object.

In this paper, we propose a 3D shape estimation method that uses all the directional information (light direction and polarization state) of the light to take the advantages of two conventional (photometric stereo and polarization-based) approaches while overcoming their disadvantages. Our contributions summarized as follows. We present two constraints: the *shading-stereoscopic constraint* for a pair of polarization images associated with two light directions and the *polarization-stereoscopic constraint* for a pair of polarizer angles. Employing the two constraints, we present two algorithms. The first algorithm estimates surface normal and refractive index of an individual point on an object knowing the light directions. The second algorithm estimates the surface normals, light directions, and refractive indexes.

2. Related Works

There are photometric stereo and polarization-based fusion methods in this field. Atkinson et al. [3] first apply a standard polarization method with known refractive index to estimate zenith and ambiguous azimuth angles of the surface normals. Then, from the relationship between the known light directions and the captured image intensity, they can select the right azimuth angle for each pixel. This method assumes to know the refractive index and also light directions. In contrast, Saman et al. [19] first employ a photometric stereo method to recover the surface normals of the object assuming the Lambertian surface, then they estimate the refractive indexes of the material by utilizing the earlier estimated normals. However, actual material with polarization effect is widely known as non-Lambertian. In overall, these fusion methods employ shading and polarization independently, they do not take true advantage of both cues for shape estimation.

There also exist methods that can estimate shape and light direction for Lambertian object [11, 16] upto an ambiguous linear transformation [4]. Then additional prior is used to fix this transformation to disambiguate the shape and light directions.

Drbohlav et al [7] propose a disambiguation solution for standard photometric stereo methods by controlling two polarizers simultaneously. Hence, they have more information

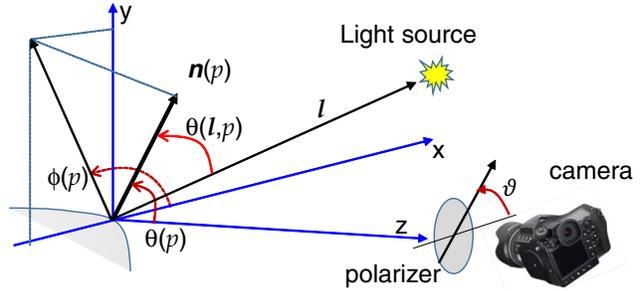


Figure 1: Normal vector $n(p)$ is defined in the world coordinate system by a zenith angle $\theta(p)$ and azimuth angle $\phi(p)$.

to disambiguate the surface normals and estimate the light directions. Compared with the standard polarization methods, this method is much stricter and not practical since they need to control both the polarization of light to the object and the polarization of reflected light to the camera simultaneously. Moreover, polarization information is not used to improve the accuracy of the surface normal or to estimate the refractive index.

3. Problem Setting and Assumption

We assume a non-homogeneous smooth dielectric object illuminated by a directional light source in different directions with the same distance to the object. A monochromatic camera observes the object through a controllable polarizer. For each light direction, images are captured with the same set of known polarizer angles. We assume the polarization for diffuse reflection on the object surface in air. We also assume an orthogonal projection model for the camera to simplify the formulation of the problem.

A world coordinate system is defined in Figure 1, which is originated at the object center, where z - axis coincides with the optical axis of the camera, and xy plane is translated from the image plane. For each pixel p on the object's image, we will recover its 3D surface normal $n(p)$ and refractive index $\eta(p)$. In this coordinate system, $n(p)$ is defined by a zenith angle $\theta(p)$ and an azimuth angle $\phi(p)$.

We propose two constraints, one for a pair of polarizer angles and one for a pair of light directions. To use these constraints, we present two algorithms that estimate the surface normals of the object. In the first algorithm, we assume we know the light directions and then estimate the surface normals and refractive indexes of the object. In the second algorithm, we do not know the light directions, and we estimate the surface normals, refractive indexes, and light directions simultaneously.

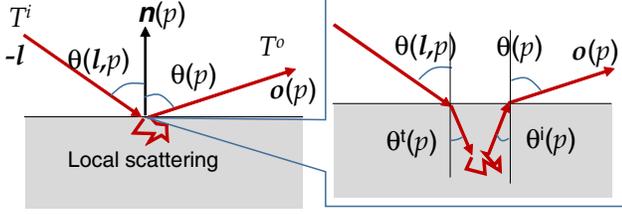


Figure 2: Diffuse reflection at surface point associated with pixel p and normal vector $\mathbf{n}(p)$. It is noted that all the 3D vectors such as \mathbf{l} , $\mathbf{n}(p)$, $\mathbf{o}(p)$ are normalized.

4. Shading-Stereoscopic Constraint

In this section, we derive a constraint for two light directions at the same polarizer angle. When a light ray from a direction \mathbf{l} is incident on a dielectric surface, it separates into two rays, a reflected ray on the surface and a refracted ray under the surface. Both rays are partially polarized obeying Fresnel theory [9]. Reflected ray can only be observed at a specific observing angle and is ignored in our case. The refracted ray is scattered and becomes depolarized and refracts back into the air. The outgoing ray is partially polarized according to Fresnel theory. This is the principle of diffuse reflection on a dielectric material, which is illustrated in Figure 2 for a surface point associated with pixel p and normal $\mathbf{n}(p)$. The camera observes the diffuse reflection from a wide range of observing directions $\mathbf{o}(p)$.

The Fresnel reflection coefficients for perpendicular and parallel polarization components of light reflected on the surface are:

$$r_{\perp}(p) = \frac{\cos \theta(\mathbf{l}, p) - \eta(p) \cos \theta^t(p)}{\cos \theta(\mathbf{l}, p) + \eta(p) \cos \theta^t(p)}, \quad (1)$$

$$r_{\parallel}(p) = \frac{\eta(p) \cos \theta(\mathbf{l}, p) - \cos \theta^t(p)}{\eta(p) \cos \theta(\mathbf{l}, p) + \cos \theta^t(p)}, \quad (2)$$

where $\theta^t(p)$ is given by Snell's law: $\sin \theta(\mathbf{l}, p) = \eta(p) \sin \theta^t(p)$, $\cos \theta(\mathbf{l}, p) = \mathbf{l} \cdot \mathbf{n}(p)$, and $\cos \theta^t(p) = \mathbf{n}(p) \cdot \mathbf{o}(p)$. Considering the total light energy or intensity, the reflection ratio is computed as:

$$F(\theta(p, \mathbf{l}), \eta(p)) = r_{\perp}^2(p) + r_{\parallel}^2(p), \quad (3)$$

and hence, the transmission ratio is $T = 1 - R$. Similarly, we can compute the transmission ratio for the diffuse light coming out of the surface and traveling to the camera at the observing angle $\theta(p)$. This diffuse reflection is modeled by Wolff reflectance model for a smooth dielectric material [21] without a polarizer in front of the camera:

$$I(\theta(\mathbf{l}, p), \theta(p), \eta(p)) = \rho(p) I(\mathbf{l}) T^i(p) T^o(p), \quad (4)$$

where

$$T^i(p) = (1 - F(\theta(\mathbf{l}, p), \eta(p))) \cos \theta(\mathbf{l}, p), \quad (5)$$

$$T^o(p) = 1 - F\left(\arcsin\left(\frac{\sin \theta(p)}{\eta(p)}\right), \eta(p)^{-1}\right), \quad (6)$$

and $\rho(p)$ is the total diffuse albedo of the surface point associated with p , which is assumed an unknown constant. $I(\mathbf{l})$ is the irradiance from the light source in direction \mathbf{l} , which is also an unknown constant according to our assumption. When we capture image with a polarizer, the rotation is defined by the polarizer angle v . The first transmission term $T^i(p)$ is for the light from the source that transmits through the surface, it does not depend on any configuration of the polarizer. In contrast, the second transmission term $T^o(p)$ is modulated by v and the image intensity $I(\theta(p, \mathbf{l}), \theta(p), \eta(p), v)$ is captured instead of $I(\theta(\mathbf{l}, p), \theta(p), \eta(p))$, which we consider it in the next section. However, if we fix the polarizer angle at v , the second term is also unchanged.

Assuming that the camera projection is orthogonal, the observation vector $\mathbf{o}(p)$ is known, and we re-parameterize the image intensity for p by $I(\mathbf{l}, v, p)$. When the polarizer angle is known and the image is captured in different light directions \mathbf{l}_1 and \mathbf{l}_2 , the ratio between image intensities at p can be obtained by removing the common factors:

$$\frac{I(\mathbf{l}_1, v, p)}{I(\mathbf{l}_2, v, p)} = \frac{[1 - F(\theta(\mathbf{l}_1, p), \eta(p))] \cos \theta(\mathbf{l}_1, p)}{[1 - F(\theta(\mathbf{l}_2, p), \eta(p))] \cos \theta(\mathbf{l}_2, p)} \quad (7)$$

or

$$I(\mathbf{l}_1, v, p) [1 - F(\theta(\mathbf{l}_2, p), \eta(p))] \cos \theta(\mathbf{l}_2, p) - \quad (8)$$

$$I(\mathbf{l}_2, v, p) [1 - F(\theta(\mathbf{l}_1, p), \eta(p))] \cos \theta(\mathbf{l}_1, p) = 0,$$

where $\theta(\mathbf{l}_i, p)$ can be derived from the dot product $\mathbf{l}_i \cdot \mathbf{n}(p)$, $i = 1, 2$. This ratio is independent of the incident irradiance and the albedo of the object. It is the first constraint on the light directions, surface normal, and refractive index of the object.

5. Polarization-Stereoscopic Constraint

In this section, we derive a constraint for two different polarizer angles of the same light direction. When we capture an image with a polarizer, the image intensity is manipulated according to the *transmitted radiance sinusoid*:

$$I(\mathbf{l}, v, p) = \frac{I_{max}(\mathbf{l}, p) + I_{min}(\mathbf{l}, p)}{2} + \frac{I_{max}(\mathbf{l}, p) - I_{min}(\mathbf{l}, p)}{2} \cos(2v - 2\phi(p)), \quad (9)$$

where $I_{max}(\mathbf{l}, p)$ and $I_{min}(\mathbf{l}, p)$ are maximum and minimum radiance of p for all angles v . To quantify the polarization effect, the degree of polarization (DOP) is computed as [9]:

$$d(p) = \frac{I_{max}(\mathbf{l}, p) - I_{min}(\mathbf{l}, p)}{I_{max}(\mathbf{l}, p) + I_{min}(\mathbf{l}, p)}. \quad (10)$$

In the case of diffuse reflection, by combining Fresnel transmission coefficients and Snell's law, we have a different formulation of the DOP depending on the refractive index and the zenith angle [2]:

$$d(p) = d(\theta(p), \eta(p)) = \frac{(\eta(p) - \eta^{-1}(p))^2 \sin^2 \theta(p)}{2 + 2\eta^2(p) - (\eta(p) + \frac{1}{\eta(p)})^2 \sin^2 \theta(p) + 4 \cos \theta(p) \sqrt{\eta^2(p) - \sin^2 \theta(p)}} \quad (11)$$

However, $I_{max}(\mathbf{l}, p) + I_{min}(\mathbf{l}, p)$ is in fact the image intensity captured without the polarizer [17], $I(\mathbf{l}, p)$. Hence, from (9), (10), and (11), we can formulate the relationship:

$$I(\mathbf{l}, v, p) = \frac{I(\mathbf{l}, p) [1 + d(\theta(p), \eta(p)) \cos(2v - 2\phi(p))]}{2}. \quad (12)$$

Therefore, if we capture an image with two different polarizer angles, v_1 and v_2 , the ratio between image intensities can be computed:

$$\frac{I(\mathbf{l}, v_1, p)}{I(\mathbf{l}, v_2, p)} = \frac{1 + d(\theta(p), \eta(p)) \cos(2v_1 - 2\phi(p))}{1 + d(\theta(p), \eta(p)) \cos(2v_2 - 2\phi(p))}, \quad (13)$$

or

$$I(\mathbf{l}, v_1, p) [1 + d(\theta(p), \eta(p)) \cos(2v_2 - 2\phi(p))] - I(\mathbf{l}, v_2, p) [1 + d(\theta(p), \eta(p)) \cos(2v_1 - 2\phi(p))] = 0 \quad (14)$$

This is the second constraint that will be used in our reconstruction algorithm. It is also independent of the incident irradiance and the albedo of the object.

6. Normal Estimation From Two Constraints

In this section, we present two algorithms that reconstruct the object using the proposed constraints. In the first algorithm, we assume that all the light directions are known, like most conventional non-Lambertian photometric stereo methods. In the second algorithm, assuming the light directions are unknown, we recover the surface normals, refractive indexes, and the light directions.

We define an error function for both constraints. First, for each pair of light directions \mathbf{l}_1 and \mathbf{l}_2 at polarizer angle v , the error function based on constraint (8) for surface point p is defined:

$$E_s(\mathbf{l}_1, \mathbf{l}_2, v, p) = \left(I(\mathbf{l}_1, v, p) [1 - F(\theta(\mathbf{l}_2, p), \eta(p))] \cos \theta(\mathbf{l}_2, p) - I(\mathbf{l}_2, v, p) [1 - F(\theta(\mathbf{l}_1, p), \eta(p))] \cos \theta(\mathbf{l}_1, p) \right)^2. \quad (15)$$

Second, for each pair of polarizer angles v_1 and v_2 with light direction \mathbf{l} , the error function based on constraint (14) is defined:

$$E_p(\mathbf{l}, v_1, v_2, p) = \left(I(\mathbf{l}, v_1, p) [1 + d(\theta(p), \eta(p)) \cos(2v_2 - 2\phi(p))] - I(\mathbf{l}, v_2, p) [1 + d(\theta(p), \eta(p)) \cos(2v_1 - 2\phi(p))] \right)^2. \quad (16)$$

Considering N surface points, we have a total error using both constraints:

$$\mathbb{E}(\{\mathbf{l}\}, \{p\}, \{v\}) = w_s \sum_{\forall \mathbf{l}_i \neq \mathbf{l}_j, v, p} \alpha(\mathbf{l}_i, p) \alpha(\mathbf{l}_j, p) E_s(\mathbf{l}_i, \mathbf{l}_j, v, p) + w_p \sum_{\forall \mathbf{l}, v_i \neq v_j, p} \alpha(\mathbf{l}, p) E_p(\mathbf{l}, v_i, v_j, p) + \lambda g(\{\mathbf{l}\}, \{p\}) \quad (17)$$

where $\alpha(\mathbf{l}, p)$ is the visibility coefficient for the light source at direction \mathbf{l} of point p . $\alpha(\mathbf{l}, p)$ is computed from the actual image intensity of p . If the image intensity is below a threshold (the image intensity is too dark), then $\alpha(\mathbf{l}, p) = 0$, otherwise $\alpha(\mathbf{l}, p) = 1$. w_s and w_p are the weights on the shading-stereoscopic and polarization-stereoscopic terms, respectively, such that $w_s + w_p = 1$. g is a prior term that stands for the prior constraints on the surface normal and light directions, and λ is the weight used to control this term. We employ a sigmoid-like function [12] to formulate these constraints. For example, to make a constraint that favors $\mathbf{n}(p) \cdot \mathbf{o}(p) > 0$, an error term is added to g :

$$E_m(\mathbf{n}(p)) = \frac{1 - \mathbf{n}(p) \cdot \mathbf{o}(p)}{1 + \exp(k\mathbf{n}(p) \cdot \mathbf{o}(p))}, \quad (18)$$

where k controls the steepness of this term, it can be as large as 100 in our experiment.

Given L light directions of the light source and P polarizer angles for each light direction, we have $N_L = L(L-1)/2$ pairs of light directions and $N_P = P(P-1)/2$ pairs of polarizer angles, which means we have N_L equations using (8) for each polarizer angle and N_P equations using (14) for each light direction for N points. However, theoretically, there are at most $L-1$ equations for the first constraint and $P-1$ for the second constraint are effective, and the remaining equations can be derived from those equations. In particular, for polarization, the polarizer angle and image intensity are modulated by (9), and we need to capture images with only three polarizer angles to recover all the parameters in (9) [22]. Images for other polarizer angles can be derived from those three images. Moreover, the polarization effect is independent of the light source. Therefore, we have a condition for the number of unknowns U

and number of point N :

$$N((L-1)\min(P,3) + (\min(P,3) - 1)) \geq U, \quad (19)$$

In practice, more than three polarization images are usually captured to improve the robustness against image noise [3, 13, 17].

6.1. Algorithm 1: Recover individual surface points

Given all light directions, the first algorithm estimates the normal vector for single point associated with pixel p with just 3 unknowns: $\theta(p), \phi(p), \eta(p)$. In this setting, the surface normal and refractive index of individual pixel p are solved directly by least square method:

$$\theta^*(p), \phi^*(p), \eta^*(p) = \arg \min_{\theta(p), \phi(p), \eta(p)} \mathbb{E}(\{\mathcal{L}\}, p, \{v\}). \quad (20)$$

In implementation, we solve the optimization using the Levenberg-Marquardt method. We employ the standard photometric stereo method [18] to initialize the surface normals, and refractive indexes are initialized to 1.

According to (19), a minimum configuration of light direction and polarization angles is $(N, L, P) = (1, 4, 1)$. However, to take advantage of both shading and polarization information, we need at least $(N, L, P) = (1, 3, 2)$.

6.2. Algorithm 2: Recover shape and light directions

The second algorithm tackles a much harder situation in which L light directions are unknown. From (19), we see that the reconstruction cannot work with only one surface point. Instead, we have to observe N points simultaneously. The number of unknowns in this case is $U = 2L + 3N$, where each light direction l can parameterized by two angles. A minimum configuration that the algorithm can be applied is $(N, L, P) = (3, 4, 2)$. Since there are a number of outliers that do not follow the diffuse reflection model, such as specular reflection from the light source or from nearby points (inter-reflection) on the object surface, and large image noise. We simply employ the random sampling method to estimate the light directions with a smaller number of surface points by optimizing (17):

$$\{\mathcal{L}^*\}, \{\theta^*(p), \phi^*(p), \eta^*(p)\} = \arg \min_{\forall \mathcal{L}, p} \mathbb{E}(\{\mathcal{L}\}, \{p\}, \{v\}), \quad (21)$$

and then use the *Algorithm1* to estimate the shape and refractive indexes of the whole object. This solution candidate is evaluated by a score which is a sum of the reconstruction error (17) and the sum of the differences in the surface normal and refractive index difference of neighboring points of the whole object. The candidate having the lowest score is the output of the reconstruction algorithm.

To initialize the optimization (21), we set all surface normals and light direction toward the camera, which is

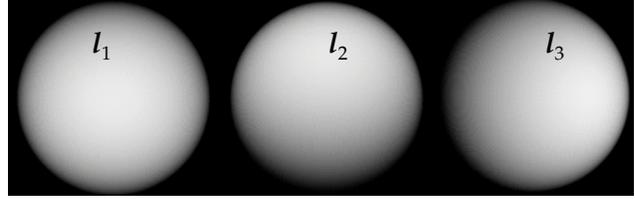


Figure 3: Three polarization images taken at a zero polarizer angle of three light directions.

$(0, 0, 1)$, and the refractive indexes to be 1. Because the optimization function (17) is nonlinear, the optimization may produce ambiguous solutions as seen for the standard polarization-based method, and we need a prior on some light directions to guide the optimization function. For instance, we assume the first light source is located in the first octant of the world coordinate system and use (18) for this prior. We can also employ a stronger prior such as knowing one light direction, which allows the *Algorithm 2* to work much more robustly. In addition, random sampling may also help to avoid the locally optimal solutions.

Compared with the factorization method in the photometric stereo method applied for Lambertian material [11, 16], the ambiguity in our case is relaxed even with a more complex BRDF. This is because fusing the proposed constraints, we have a heavily over-constrained equation system for a group of pixels. The total ambiguity is therefore significantly reduced compared with the ambiguity from each individual constraint.

7. Experiments

We evaluate the proposed algorithms with both simulation and real experiments.

7.1. Simulation Experiments with Weights on Two Constraints

To carry out the simulation experiments, we made a polarization ray tracing rendering software to simulate the diffuse polarization in the case of an orthogonal camera and distant light source for a opaque dielectric sphere with a refractive index of 1.4553. Examples of original polarization images are shown in Figure 3. Gaussian noise with deviation of 0.01 (1% of the highest normalized image intensity) was added randomly to the images. For each light source, we used three polarizer angles, $v \in \{0^\circ, 45^\circ, 90^\circ\}$, to generate polarization images of the object.

In this experiment, we analyze the effects of the weights given to the two constraints in (17). The results of the experiments are presented in Figure 4 for average errors of the surface normal and refractive index and in Figure 5 for the error distribution for all image pixels. The polarization-stereoscopic constraint is not used when the weight w_s is 1,

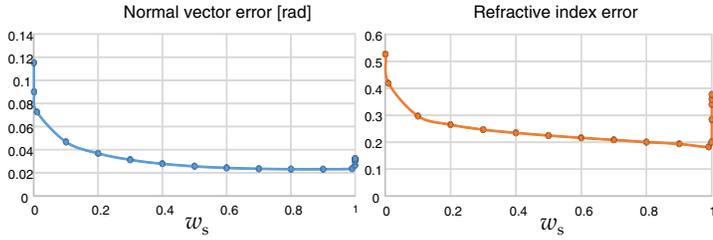


Figure 4: *Algorithm 1*: Average error for all image pixels and different weights, w_s and $w_p = 1 - w_s$, given to the two constraints.

and the shading-stereoscopic constraint is not used when w_s is zero. When the shading-stereoscopic constraint is not used, both the surface normal and refractive index are inaccurate. Standard polarization methods face ambiguity of the surface orientation and ambiguity of the relationship between the refractive index and zenith angle. In contrast, when the polarization-stereoscopic constraint is not used, the shading-stereoscopic constraint still can work to recover the surface normal but cannot recover the refractive index well. For the weight between those two extremes, $0 < w_s < 1$, the proposed method can take the advantages of the two constraints in recovering both the surface normal and refractive index robustly. Moreover, we see that the setting of the weights is relaxed since the proposed method can work for a wide range of weights; and we choose the balanced weights, $w_s = 0.5$ and $w_p = 0.5$, in the remaining experiments.

The distribution of error in Figure 5 shows a limitation of the proposed method in that the estimation of the refractive index is inaccurate at the center of the image where the zenith angle is small and DOP is weak. This limitation is not only for the proposed method but also for any method that use polarization images.

7.2. Comparison with benchmark methods

We compared proposed algorithms with the standard photometric stereo [18], standard polarization-based [1], and a shading-polarization fusion [19] methods in the similar simulated sphere as above. A large number of input datasets were generated with different deviation levels of Gaussian noise, see Figure 6, random light directions, and polarizer angles. The input requirements of these method are basically different, however, the number of input images are fixed for all the methods: 6, 8, 16, 24, 36, and 48. For examples, in case of 8 input images, proposed methods and [19] take the configuration of 4 light directions and 2 polarizer angles for each light direction, while [18] takes 8 unpolarized images for 8 light directions, and [1] takes 8 polarizer angles for only one light direction.

The average results for each noise level are shown in Fig-

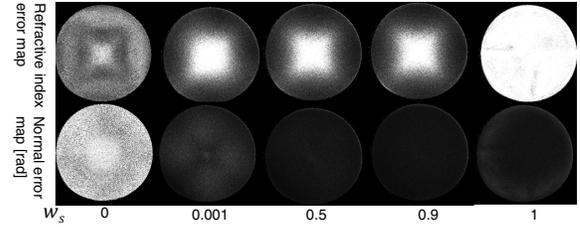


Figure 5: *Algorithm 1*: Error map of image pixels for different weights, w_s and w_p , given to the two constraints. A brighter pixel indicates larger error.

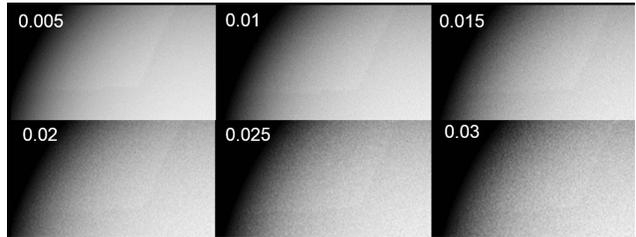


Figure 6: Variation of Gaussian noise deviation on the synthesized images. The maximum image intensity is normalized to 1.

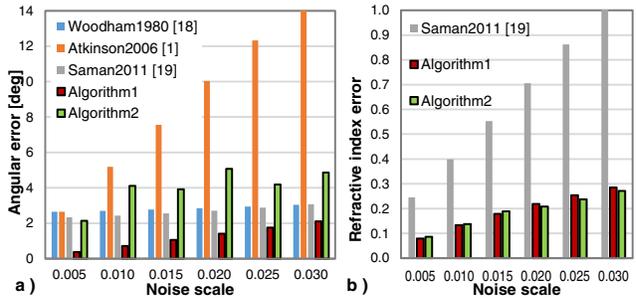


Figure 7: Normal (a) and refractive index (b) results comparison.

ure 7, which demonstrate the effectiveness of proposed fusion method through the comparison of *Algorithm 1* and state-of-the-art methods, given the light directions. On the other hand, without light directions, *Algorithm 2* could produce reasonable accuracy on the light directions and surface normals, in Figure 7.a, and similar refractive index accuracy, in Figure 7.b, compared with *Algorithm 1*. Note that [18] does not solve the refractive index, while [1] needs the true refractive index as input, hence their results for refractive index is not available in Figure 7.b.

7.3. Real Experiments

We carried out experiments in a darkroom. We employed a color Lumenera’s INFINITY-3 camera (8 bits/channel), but only red image channel was used. The polarizer was

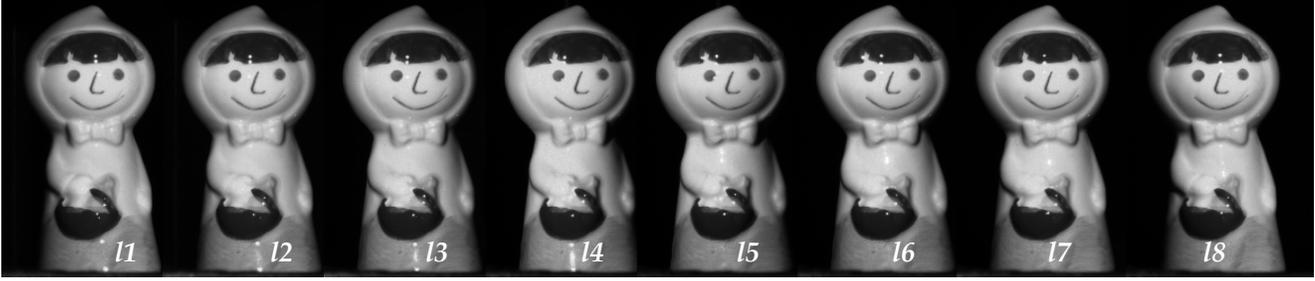


Figure 8: The ceramic doll in our experiment, images shown are with zero polarizer angle for 8 light directions.

mounted on a Sigma Koki’s SGSP-120YAW rotation stage. Both camera and rotation stage were controlled automatically with a PC, which allowed us to capture polarization images easily. A halogen light source was mounted on a tripod at a distance of about 2.8 m from the target object. All light directions were manually measured, and the ground-truth light direction is given in Table 1. The distance from the camera to a object was about 1.2 m, and the object size was about 3 cm. We assumed an orthogonal camera projection.

The first object in our experiment was a ceramic doll. Images of the doll taken with 8 light directions are shown in Figure 8. For each light direction, we captured 5 polarization images at $0^\circ, 36^\circ, 72^\circ, 108^\circ,$ and 144° . We first evaluate the results obtained with *Algorithm 1* using ground-truth light directions and then evaluate the results of *Algorithm 2* obtained without knowing the light directions.

7.3.1 Shape recovery knowing light directions

The results of the experiment are shown in Figure 9. It should be noted that the DOP map is computed from (11) when refractive index and zenith angle has already been estimated. The surface normal $\mathbf{n}(p)$ is encoded by RGB color with $(\mathbf{n}_z(p), \frac{\mathbf{n}_y(p)+1}{2}, \frac{\mathbf{n}_x(p)+1}{2})$. The dynamic ranges of the zenith, azimuth, DOP, and refractive index map are $[0, \pi/2]$, $[0, 2\pi]$, $[0, 0.4]$, and $[1, 2]$, respectively; a brighter point indicates a larger value. There is a large variation of refractive index for pixels with small zenith angle. The reason is that when DOP is small, the estimation of refractive index is not robust as discussed earlier for Figure 5. The histogram of estimated refractive index is shown in Figure 11, from which we can see that the dominant refractive index of the doll is about 1.3.

These results also show that the performance of the proposed method can be affected by specularity, although there are not many of such points.

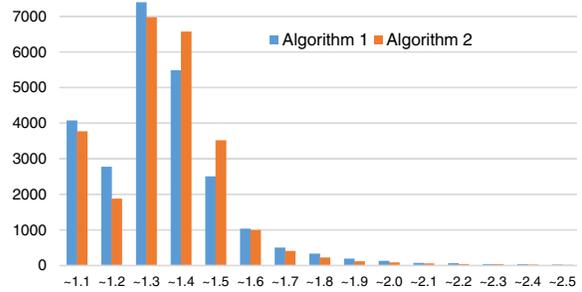


Figure 11: Histogram of estimated refractive indexes for the doll.

	Groundtruth direction			Estimated direction			Error [deg]
I1	0.2033	0.1545	0.9668	0.2183	0.0942	0.9713	3.567
I2	-0.0144	0.1561	0.9876	-0.0031	0.0873	0.9962	4.023
I3	-0.1710	0.1498	0.9738	-0.1703	0.0863	0.9816	3.663
I4	-0.3448	0.1478	0.9270	-0.3290	0.1056	0.9384	2.661
I5	-0.3167	0.3146	0.8949	-0.3036	0.2659	0.9150	3.111
I6	-0.1637	0.3023	0.9390	-0.1421	0.2451	0.9590	3.688
I7	0.0015	0.3360	0.9419	0.0144	0.2496	0.9682	5.231
I8	0.1821	0.3307	0.9260	0.2098	0.2406	0.9477	5.541
Avg.							3.936

Table 1: *Algorithm 2*: Estimated light directions compared with the ground-truth for the doll.

7.3.2 Shape and light direction recovery

We evaluated ability of *Algorithm 2* to recover the 3D shape and light directions. We employed two priors on two light directions to help the disambiguation: first light direction, l_1 , is toward the first octant, ($l_{1,x} > 0, l_{1,y} > 0, l_{1,z} > 0$), and the fourth light direction is toward the second octant, ($l_{4,x} < 0, l_{4,y} > 0, l_{4,z} > 0$). The experiment results are presented in Figure 10 for the summary of overall performance, Figure 11 for refractive index, and Table 1 for light direction. In Table 1, the average angle error is 3.936° , which is reasonable since the object is small and the resolution of the object image is low (less than 250×250 pixels). Because the light directions and surface normals are strongly correlated, the accuracy of surface normals is on

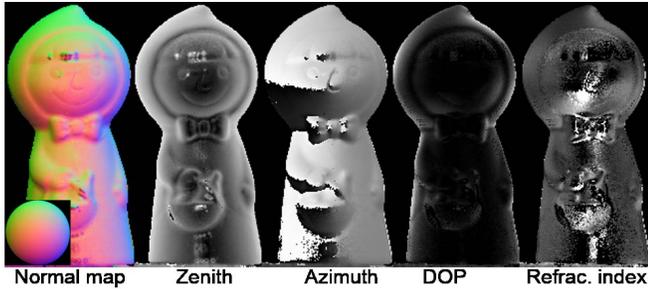


Figure 9: Performance of *Algorithm 1* for the doll.

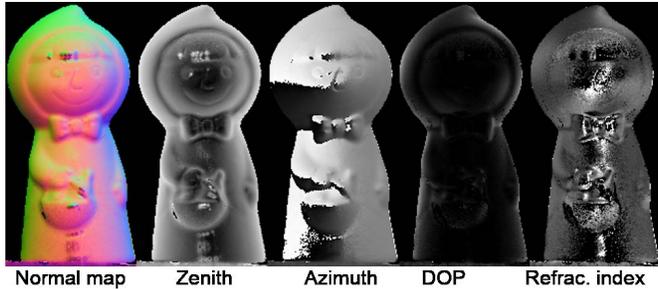


Figure 10: Performance of *Algorithm 2* for the doll.

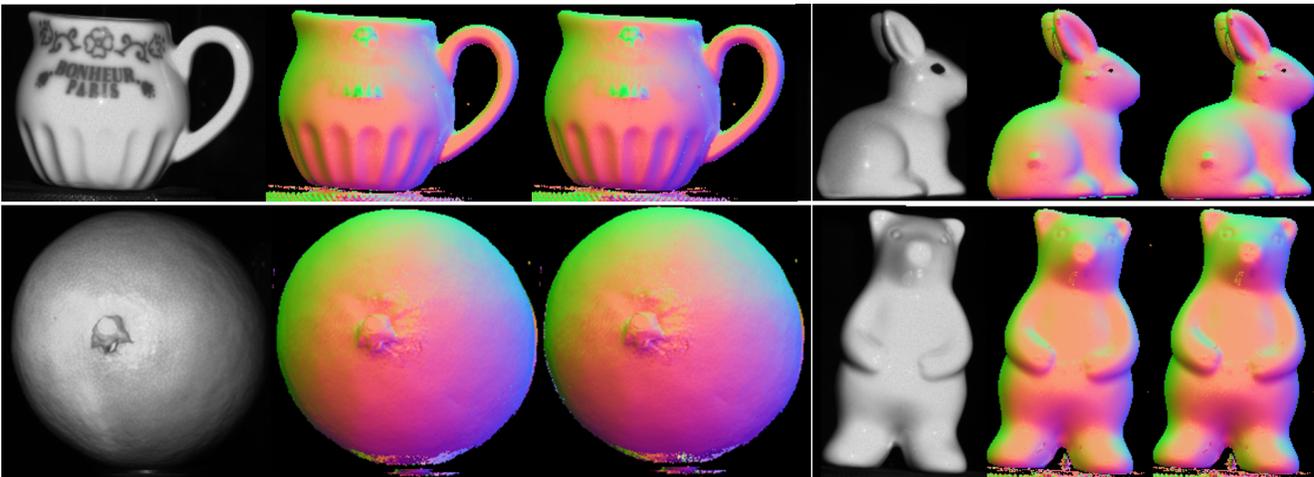


Figure 12: Performances of proposed methods for a cup, a rabbit, a fresh lemon, and a bear. For each object, a captured image, and results of *Algorithm 1* and *Algorithm 2* are shown from left to right, respectively.

par with the accuracy of light directions. Therefore, we also can predict the surface normal accuracy.

More experiments and comparisons between the two algorithms are shown in Figure 12 for various objects with the same light sources for the doll. In these objects, the actual surface normals are distributed more uniformly than that in the doll. As the result, the *Algorithm 2* gave better performances for the light direction estimation, and intuitively better surface normals. The average angular errors for light directions are 3.17° , 2.60° , 2.84° , and 3.45° , respectively. The surface normal reconstruction performances of both algorithms are very close. Although we do not have a complete solution for disambiguation of the light directions and surface normals, these results are promising compared with [7] which controls two polarizers to disambiguate the light directions and surface normals. Their accuracy is about 10° for even 14 bit images.

8. Conclusion and Future Works

We presented a shape reconstruction method that integrates shading and polarization to take the advantage and

overcome the disadvantages of these two cues. We present two constraints, one constraint for shading with a pair of light directions and one constraint for polarization with a pair of polarizer angles. As the result, the surface normals, refractive indexes, and light directions can be reconstructed. The proposed algorithms were positively evaluated in both simulation and real-world experiments. We can recovery the shape, refractive indexes, and light directions.

For light direction estimation, we currently use a simple method to solve the ambiguity since it is relaxed significantly by incorporating polarization compared with conventional methods. However, in the future work, we would like to improve this solution by employing a stronger technique such as smoothness constraint or specular cue [8]. A solution for inter-reflection is also addressed as another future work.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number 25240027

References

- [1] G. A. Atkinson and E. R. Hancock. Recovery of surface orientation from diffuse polarization. *IEEE transactions on image processing*, 15(6):1653–64, June 2006. 1, 6
- [2] G. A. Atkinson and E. R. Hancock. Shape estimation using polarization and shading from two views. *IEEE transactions on pattern analysis and machine intelligence*, 29(11):2001–17, Nov. 2007. 1, 4
- [3] G. A. Atkinson and E. R. Hancock. Surface Reconstruction Using Polarization and Photometric Stereo. *Computer Analysis of Images and Patterns Lecture Notes in Computer Science*, pages 466–473, 2007. 1, 2, 5
- [4] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The Bas-Relief Ambiguity. *International Journal of Computer Vision*, 35(1):33–44, 1999. 2
- [5] M. Chandraker. What Camera Motion Reveals About Shape With Unknown BRDF. In *CVPR*, 2014. 1
- [6] D. Miyazaki and M. Saito and Y. Sato and K. Ikeuchi. Determining surface orientations of transparent objects based on polarization. *The Journal of the Optical Society of America A*, 19(4):687–694, 2002. 1
- [7] O. Drbohlav and R. Sara. Unambiguous determination of shape from photometric stereo with unknown light sources. In *ICCV*, pages 581–586, 2001. 2, 8
- [8] O. Drbohlav and R. Sara. Specularities Reduce Ambiguity of Uncalibrated Photometric Stereo. In *ECCV*, pages 46–60, 2002. 8
- [9] Eugene Hecht. *Optics*. Addison Wesley, 4th edition, 2001. 1, 3
- [10] S. Foix and G. Aleny. Lock-in Time-of-Flight (ToF) Cameras : A Survey. *IEEE Sensors Journal*, 11(3):1–11, 2011. 1
- [11] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America A*, 11(11):3079, Nov. 1994. 2, 5
- [12] T. Higo, Y. Matsushita, and K. Ikeuchi. Consensus photometric stereo. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1157–1164. Ieee, June 2010. 1, 4
- [13] C. P. Huynh, A. Robles-Kelly, and E. R. Hancock. Shape and Refractive Index from Single-View Spectro-Polarimetric Images. *International Journal of Computer Vision*, 101(1):64–94, June 2013. 1, 5
- [14] C. Inoshita, Y. Mukaigawa, Y. Matsushita, and Y. Yagi. Surface Normal Deconvolution : Photometric Stereo for Optically Thick Translucent Objects. In *ECCV*, pages 1–15, 2014. 1
- [15] F. Lu, Y. Matsushita, I. Sato, T. Okabe, and Y. Sato. Uncalibrated Photometric Stereo for Unknown Isotropic Reflectances. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1490–1497, June 2013. 1
- [16] D. Miyazaki and K. Ikeuchi. Photometric stereo under unknown light sources using robust SVD with missing data. In *2010 IEEE International Conference on Image Processing*, pages 4057–4060, Sept. 2010. 2, 5
- [17] S. K. Nayar, X.-s. Fang, and T. Boulton. Separation of Reflection Components Using Color and Polarization. *International Journal of Computer Vision*, 21(3):163–186, 1997. 4, 5
- [18] Robert J. Woodham. Photometric Method For Determining Surface Orientation From Multiple Images. *Optical Engineering*, 19(1), 1980. 1, 5, 6
- [19] G. Saman, E. Hancock, and C. Science. Refractive index estimation using photometric stereo. In *Image Processing (ICIP), 2011 18th IEEE International Conference on*, pages 1925–1928, 2011. 2, 6
- [20] I. Sato, T. Okabe, Q. Yu, and Y. Sato. Shape Reconstruction Based on Similarity in Radiance Changes under Varying Illumination. *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8, 2007. 1
- [21] L. B. Wolff, N. York, and M. Oren. Improved Diffuse Reflection Models for Computer Vision. *IJCV*, 30(1):55–71, 1998. 3
- [22] T. Wolff, L.B. ; Boulton. Constraining Object Features Using a Polarization Reflectance Model. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 13(7):635 – 657, 1991. 4
- [23] L. Zhang and E. R. Hancock. Robust estimation of shape and polarisation using blind source separation. *Pattern Recognition Letters*, 34(8):856–862, June 2013. 1