Polarimetric Multi-View Stereo

Zhaopeng Cui[1] Jinwei Gu[2] Boxin Shi[3] Ping Tan[1] Jan Kautz[2]
[1]Simon Fraser University [2]NVIDIA Research
[3]Artificial Intelligence Research Center, National Institute of AIST

Abstract

Multi-view stereo relies on feature correspondences for 3D reconstruction, and thus is fundamentally flawed in dealing with featureless scenes. In this paper, we propose ***polarimetric multi-view stereo****, which combines per-pixel photometric information from polarization with epipolar constraints from multiple views for 3D reconstruction. Polarization reveals surface normal information, and is thus helpful to propagate depth to featureless regions. Polarimetric multi-view stereo is completely passive and can be applied outdoors in uncontrolled illumination, since the data capture can be done simply with either a polarizer or a polarization camera. Unlike previous work on shape-from-polarization which is limited to either diffuse polarization or specular polarization only, we propose a novel polarization imaging model that can handle real-world objects with mixed polarization. We prove there are exactly two types of ambiguities on estimating surface azimuth angles from polarization, and we resolve them with graph optimization and iso-depth contour tracing. This step significantly improves the initial depth map estimate, which are later fused together for complete 3D reconstruction. Extensive experimental results demonstrate high-quality 3D reconstruction and better performance than state-of-the-art multi-view stereo methods, especially on featureless 3D objects, such as ceramic tiles, office room with white walls, and highly reflective cars in the outdoors.*

1. Introduction

Multi-view stereo reconstructs dense 3D models from multiple images. It has been intensively studied in computer vision [34, 35, 10, 11], with wide applications in computer graphics, robotics, computer-aided design, and human-computer interactions. Multi-view stereo relies on finding feature correspondences with epipolar constraints, and thus it is fundamentally flawed in dealing with featureless objects, since correspondence cannot be reliably found.

To deal with featureless regions, one approach is to combine multi-view and photometric cues, either from photometric stereo [41, 12, 29] or structured light [26]. These methods, however, require active/controlled illumination and cannot operate in the outdoors. Another approach is to refine 3D geometry with shape-from-shading [40, 18, 28], which requires the knowledge of illumination (captured or estimated), and is generally limited to Lambertian surfaces or surfaces with uniform reflectance.

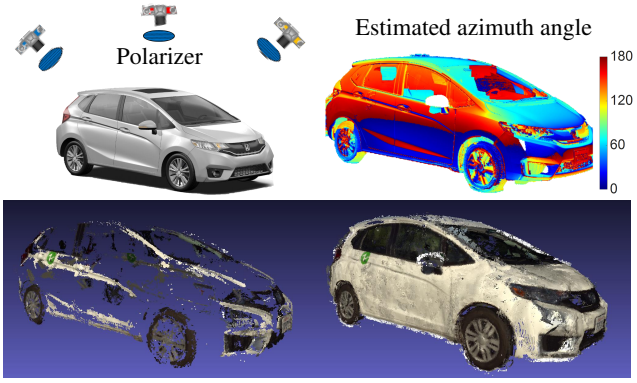


Figure 1: Polarimetric multi-view stereo exploits per-pixel photometric information from polarized images to significantly improve 3D reconstruction quality for challenging scenes such as highlight and featureless regions. Please refer to Figure 4 and Figure 9 for more details and results.

In this paper, we propose a novel method called **polarimetric multi-view stereo** that exploits per-pixel polarization information for dense 3D reconstruction. Polarized images provide information about surface normals for a wide range of materials (*e.g.*, specular, diffuse, glass, metal, dielectrics) [22, 14, 3, 25], which are additional constraints for 3D depth recovery for featureless regions. Polarimetric multi-view stereo is a completely passive approach, and can be applied to a wide variety of objects in the outdoors with uncontrolled illumination. Polarized images can be captured in a single shot with a polarization camera [1, 9, 31, 33] or a rotating polarizer [38], which makes this approach as simple and practical as using ordinary cameras.

Polarization has been studied previously for 3D reconstruction. Two recent work includes Polarized3D [16] which requires RGB-D sensors and thus is limited to indoors, and [36] which assumes uniform albedo and requires lighting

estimation (which itself is a challenging problem). In addition, almost all prior work assumes either pure diffuse polarization reflection (*e.g.*, [3]) or pure specular polarization reflection (*e.g.*, [23]) – this assumption is impractical for many real-world objects with mixed polarization reflection. To our knowledge, *the proposed polarimetric multi-view stereo approach is the first passive method that can deal with mixed polarization for non-uniform albedo, without the need for lighting estimation*.

Specifically, in this paper, we prove that polarized images can determine the azimuth angles of surface normals with exactly two types of ambiguities — the π-ambiguity and the $\pi/2$-ambiguity. Since azimuth angles provide strong geometry constraint on surface shapes, our key idea is to simultaneously resolve these ambiguities in azimuth angle estimation, and to use the azimuth angles to propagate depth estimated from sparse points with sufficient features to featureless regions for dense 3D reconstruction. We resolve the $\pi/2$-ambiguity with graph optimization and bypass the π-ambiguity with iso-depth contour tracing. This step significantly improves the initial depth maps estimated from classical multi-view stereo, which are later fused together for a complete 3D reconstruction.

We performed extensive experiments on challenging scenarios for classical multi-view stereo, such as ceramic tile, vase with vanish paints, office room with white walls, and highly-reflective cars in the outdoors. Our method achieves high-quality 3D reconstruction and outperforms the state-of-the-art multi-view stereo methods.

2. Related Work

Multi-view Stereo for Textureless Surfaces Multi-view stereo (MVS) [34, 35, 10, 11] relies on finding feature correspondences across multiple images for 3D reconstruction. To deal with featureless objects, one approach is to enforce priors on 3D shapes (*e.g.*, smoothness) for robust correspondence matching [4, 11, 13]. Another approach is to combine multi-view stereo and photometric cues, *i.e.*, dense surface normals estimated either from photometric stereo [12, 29, 41, 26] or from shape-from-shading [40, 18, 28]. The former requires active/controlled illumination and cannot operate in the outdoors, and the latter requires the knowledge of illumination (captured [28] or estimated [40, 18]) and is generally limited to Lambertian surfaces or surfaces with uniform reflectance.

The proposed polarimetric multi-view stereo uses photometric cues from polarization and is a completely passive approach. It does not require the estimation of illumination, and is generally applicable to non-Lambertian surfaces with spatially-varying reflectance properties.

3D Reconstruction with Polarization Polarimetric properties of reflected light reveals information about surface normals (*i.e.*, azimuth angles and zenith angles), and thus useful for 3D reconstruction. There are three categories of prior work on shape-from-polarization.

The *polarization-only approaches* use geometric priors such as the surface normals on the boundary and convexity of the objects to constrain the shape estimation [24, 3, 25]. Such heuristic constraints limit the application to smooth objects without visible occlusion boundaries.

The *photo-polarimetric approaches* combine photometric cues from shape-from-shading [20, 36] or photometric stereo [8, 27] to resolve ambiguities for surface normal estimation. These methods often make strong assumptions on materials or illumination (*e.g.*, uniform albedo [36], controlled illumination, known light direction [20]). In addition to shading, multi-spectral measurements are shown to be effective for both shape and refractive index estimation [14, 15].

The *geo-polarimetric approaches* integrate coarse depth estimation from multi-view stereo or low-cost depth sensor with polarization cues. Two-view stereo constraints have been adopted for specular and transparent surface modeling [32, 21], as well as correspondence matching [2]. Space carving [22, 23] or RGB-D sensors [16] have been employed to obtain initial 3D shape, from which the ambiguities in shape-from-polarization are resolved.

Compared to these prior work, our approach has the most general operating environment in terms of illumination and applicable objects. Our approach is completely passive, and can work under uncontrolled illumination in the outdoors, instead of active illumination [16], diffuse lighting [40, 18] or distant lighting [26, 12, 29, 41]. The method is applicable to a wide variety of objects with mixed polarized diffuse and specular reflections as well as non-uniform albedo, instead of being limited to either diffuse reflection only [3, 40, 18, 26, 12, 29] or specular reflection only [23], or only objects with uniform albedo [36].

3. A Model for Mixed Polarization Reflection

The polarization of the reflected light from a surface is determined by the polarization BRDF of the surface material [19, 5, 7]. For most surfaces, as shown in Figure 2, the reflected light includes three components: (1) the polarized specular reflection (*i.e.*, highlight), (2) the polarized diffuse reflection (due to subsurface scattering and refraction), and (3) the unpolarized diffuse reflection (due to micro-facet rough surface reflection). As we show in the appendix with Muller calculus [6, 30], the polarized specular reflection and the polarized diffuse reflection have a $\pi/2$ difference in their phase angles (defined in Equation (3)), because there is an additional refraction between the air-surface interface.

When imaging a surface under unpolarized illumination through a polarizer, the image intensity (*i.e.*, scene radiance) varies with the angle ϕ_{pol} of the polarizer sinusoidally. More specifically, for polarized diffuse reflection [16, 36, 37],

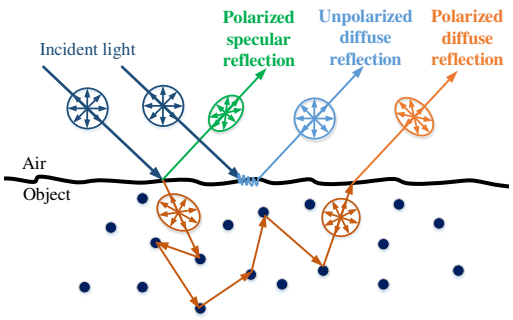


Figure 2: A diagram of surface reflection with mixed polarization. The circles with arrows show the polarization status: round circles – unpolarized, elliptic circles – partially polarized.

$$I^{dp}(\phi_{pol}) = \frac{I^{dp}_{max} + I^{dp}_{min}}{2} + \frac{I^{dp}_{max} - I^{dp}_{min}}{2} \cos(2(\phi_{pol} - \varphi)), \tag{1}$$

where φ is the azimuth angle of surface normal, I^{dp}_{max} and I^{dp}_{min} are the maximum and minimum values of the observed intensity. For the polarized specular reflection, we have a similar equation with a $\pi/2$ difference in the phase [22]:

$$I^{sp}(\phi_{pol}) = \frac{I^{sp}_{max} + I^{sp}_{min}}{2} + \frac{I^{sp}_{max} - I^{sp}_{min}}{2} \cos(2(\phi_{pol} - \varphi + \frac{\pi}{2})). \tag{2}$$

The unpolarized diffuse reflection contributes only to the DC component, and thus can be subsumed in both equations.

Equations (1) and (2) provide the foundation for many shape-from-polarization methods, since by capturing images with multiple polarization angles ϕ_{pol}, one can potentially estimate the azimuth angle φ of the surface normal. However, as shown, there are two types of ambiguities in solving for φ. The first type is called the **π-ambiguity**, since both φ and $\varphi + \pi$ satisfy the equations. The second is called the **$\pi/2$-ambiguity**, since real world objects have both the polarized specular reflection and the polarized diffuse reflection, $I = I^{sp} + I^{dp}$. The π-ambiguity is relatively easier to resolve. Prior work used cues such as rough depth from RGB-D sensor [16] or lighting direction [36] to resolve this ambiguity. The $\pi/2$-ambiguity, due to the mixed polarization in reflection, is more challenging to resolve.

As mentioned previously, many prior methods simplify the $\pi/2$-ambiguity, by assuming there is either only polarized diffuse reflection (*e.g.*, [3]) or only polarized specular reflection (*e.g.*, [22]). However, this assumption is impractical for real-world objects [37]. Figure 3 shows an example of a vase. A checkerboard is reflected on the vase surface due to specular reflection from the varnish. If we use Equation (1) to compute the azimuth angle, we can see the checkerboard pattern appear. This is because for pixels in the white squares, the polarized specular reflection dominates, while for pixels in the black squares, the polarized diffuse reflection dominates. For real-world objects, the mixture of specular and diffuse reflections, I^{sp} and I^{dp}, varies also with the material properties and surface geometry, which further complicates the problem.

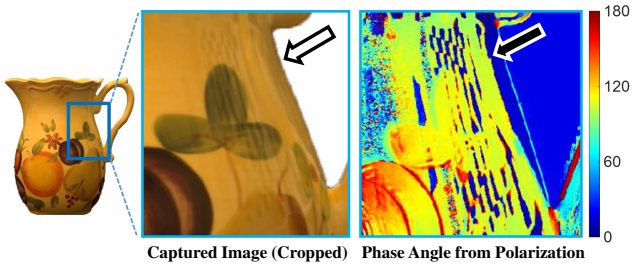


Figure 3: A real example of mixed polarization. **Left**: A checkerboard is reflected on the vase surface in the captured image (specular reflection). **Right**: The estimated phase angle (defined in Eq. 3) from polarization shows the checkerboard pattern due to the mix polarization, *i.e.*, the polarized diffuse reflection dominates in the black squares and the polarized specular reflection dominates in the white squares.

One of our key findings in this paper, is that *under unpolarized illumination, for any point on an object surface, no matter what the relative proportion of the polarized specular reflection and the polarized diffuse reflection is, the measured radiance at this point always varies sinusoidally, and there can only be the two types of ambiguities – the π-ambiguity and $\pi/2$-ambiguity – for the azimuth angle estimated from polarized images.* Proposition 1 gives a detailed description. Please refer to the appendix for the proof.

Proposition 1. *Under unpolarized illumination, the measured scene radiance from a reflective surface through a linear polarizer at a polarization angle ϕ_{pol} is*

$$I(\phi_{pol}) = \frac{I_{max} + I_{min}}{2} + \frac{I_{max} - I_{min}}{2} \cos(2(\phi_{pol} - \phi)), \tag{3}$$

where I_{max} and I_{min} is the maximum and minimum measured radiance. ϕ is defined as the phase angle, which relates to the azimuth angle φ as follows:

$$\phi = \begin{cases} \varphi & \text{if polarized diffuse reflection dominates} \\ \varphi - \frac{\pi}{2} & \text{otherwise} \end{cases}. \tag{4}$$

Proposition 1 shows that resolving the $\pi/2$-ambiguity becomes a binary labeling problem for each pixel for any given view, depending on whether the reflection at that pixel is dominated by the specular reflection or diffuse reflection. This conclusion inspires our solution for polarimetric multi-view stereo in the next section.

4. Polarimetric Multi-View Stereo

Figure 4 summarizes the proposed polarimetric multi-view stereo algorithm. The input is the polarized images captured at multiple viewpoints, either with polarization cameras [1, 9, 31, 33] or with a linear polarizer rotated at

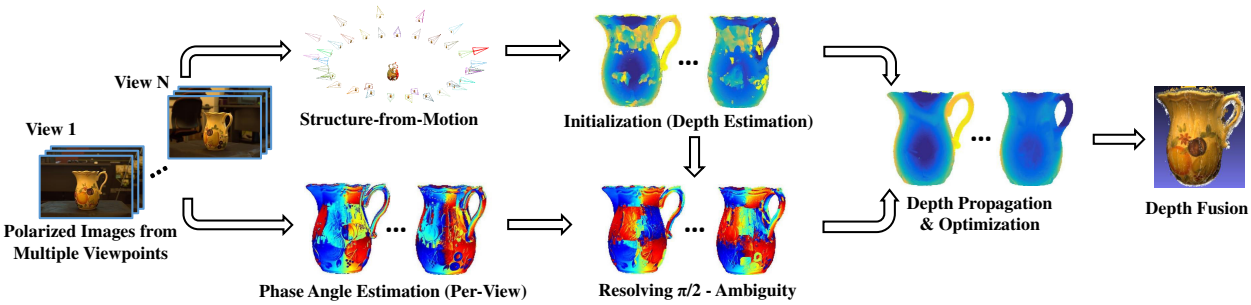


Figure 4: Flowchart of the proposed polarimetric multi-view stereo algorithm. Please refer to Section 4 for details.

multiple angles (≥ 3). In all our experiments, we capture seven images per view. Classical structure-from-motion [39] and multi-view stereo [11] methods are used to first recover the camera positions as well as an initial 3D shape for well-textured regions. On the other hand, we compute the phase angle maps ϕ for each view from the corresponding polarized images, from which we resolve the ambiguities to estimate azimuth angles φ to recover depth for featureless regions. Finally depth maps from multiple views are fused together to recover the complete 3D shape similar to [11].

4.1. Initialization and Pre-processing

We take one image per view to compute the camera poses with VisualSFM [39], and reconstruct the initial 3D shape with a recent GPU-enabled multi-view stereo method [11]. We also compute the phase angle maps ϕ for each view from Equation (3) with least square methods.

In order to use the phase angles ϕ estimated from polarization for 3D reconstruction, we need to solve the π-ambiguity and the $\pi/2$-ambiguity for each view. In Section 4.3, we explain that we can actually bypass the π-ambiguity by iso-depth contour tracing. In the next section, we present our method to resolve the $\pi/2$-ambiguity.

4.2. Resolving the $\pi/2$-ambiguity

As mentioned previously, based on Proposition 1, resolving the $\pi/2$-ambiguity can be formulated as a binary labeling problem in a graph optimization:

$$E(\{f_p\}) = \sum_{p\in\mathcal{P}} D(f_p) + \lambda \sum_{p,q\in\mathcal{N}} V(f_p, f_q), \qquad (5)$$

where f_p is a binary label at pixel p indicating whether the polarized diffuse reflection dominates ($f_p = 1$) or the polarized specular reflection dominates ($f_p = 0$) at this pixel, $\lambda = 1$, $\mathcal{P}$ is the set of all pixels, and $\mathcal{N}$ is the set of all neighboring pixel pairs.

The data term, $D(f_p)$, is defined as follows. For pixels in well-textured regions (*i.e.*, with consistent depth reconstruction at initialization), its azimuth angle from polarization should be close to the azimuth angle computed from the initial 3D shape recovered from initialization. Suppose $\mathcal{P}_+$ is the set of pixels with reliable depth recovered from initialization, we use these pixels to guide the disambiguation. For a pixel $p \in \mathcal{P}_+$, we have

$$D(f_p) = \begin{cases} g(\phi_p + \frac{\pi}{2}, \hat{\varphi}_p)/A & \text{if} \quad f_p = 0 \\ g(\phi_p, \hat{\varphi}_p)/A & \text{if} \quad f_p = 1 \end{cases}, \qquad (6)$$

where ϕ_p is the phase angle at pixel p computed from the polarized images, $\hat{\varphi}_p$ is its azimuthal angle estimated from the initial 3D shape, and $g(\phi_p, \phi_q)$ is a function that computes the distance between the two angles ϕ_p and ϕ_q, considering the cycle of π,

$$g(\phi_p, \phi_q) = \min(|\phi_p - \phi_q + \pi|, |\phi_p - \phi_q|, |\phi_p - \phi_q - \pi|), \qquad (7)$$

and $A = g(\phi_p + \frac{\pi}{2}, \hat{\varphi}_p) + g(\phi_p, \hat{\varphi}_p)$ is for normalization. The set $\mathcal{P}_+$ is determined by depth consistency check [11].

For pixels in texture-less regions (*i.e.*, with no consistent depth recovered from the initial multi-view stereo step), we can only rely on statistical priors for the data term. Suppose $\mathcal{P}_- = \mathcal{P}/\mathcal{P}_+$ is the set of pixels with no reliable depth, for a pixel $p \in \mathcal{P}_-$, we have

$$D(f_p) = \begin{cases} \rho_p & \text{if} \quad f_p = 0 \\ 1 - \rho_p & \text{if} \quad f_p = 1 \end{cases}, \qquad (8)$$

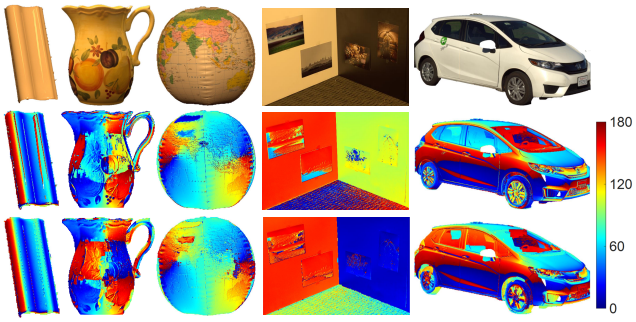
where $0 < \rho_p < 1$ is the prior probability that at pixel p it is polarized diffuse reflection dominates. ρ_p can be set according to prior knowledge of target scenes or viewing/illumination conditions. For all the experiments in this paper, we simply set $\rho_p = 0.4$ if $I_p < 0.1$, and $\rho_p = 0.55$ otherwise, where I_p is the intensity at pixel p (maximum 1.0). This setting assumes the black objects in the scene are slightly more likely to be specular reflection dominated.

The smoothness term $V(f_p, f_q)$ is designed to encourage neighboring pixels to have similar azimuth angles,

$$V(f_p, f_q) = \begin{cases} g(\phi_p, \phi_q)/B & \text{if} \quad f_p = f_q \\ g(\phi_p + \frac{\pi}{2}, \phi_q)/B & \text{if} \quad f_p \neq f_q \end{cases}, \qquad (9)$$

where $B = g(\phi_p, \phi_q) + g(\phi_p + \frac{\pi}{2}, \phi_p)$ is for normalization.

We solve this binary labeling problem with tree-reweighted belief propagation [17]. Figure 5 shows results for five examples. As shown, our algorithm can effectively resolve the $\pi/2$-ambiguity from the estimated phase angles and correct the estimated azimuth angles in the regions where the polarized specular reflection dominates (*e.g.*, the middle of a ceramic tile, the top-left of a plastic balloon, the body of a vase, the black wall and the carpet in the corner of an office, and the side windows of a car).



TILE VASE BALLOON CORNER CAR

Figure 5: Results for resolving the $\pi/2$-ambiguity for azimuthal angles. **Top**: captured images. **Middle**: phase angle maps ϕ computed from polarization. **Bottom**: azimuth angles φ after resolving the $\pi/2$-ambiguity.

4.3. Depth Propagation: Bypassing the π-ambiguity

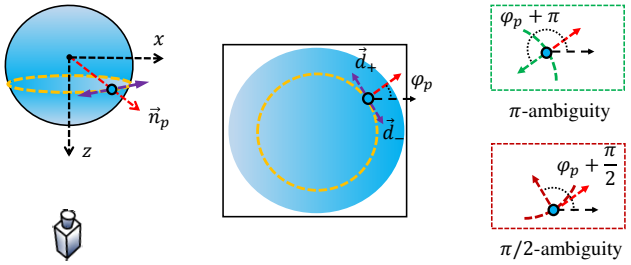
After the $\pi/2$-ambiguity is resolved, there is only the π-ambiguity between the azimuth angles estimated from polarization and the true azimuth angles of 3D objects. To resolve the π-ambiguity, prior work requires either depth [16] or lighting direction [36]. In this section, we show that without resolving the π-ambiguity, the estimated azimuth angles can still be used to improve depth estimation.

Note that even with the π-ambiguity, azimuth angles determine iso-depth contours, *i.e.*, points with the same distance to the image plane. As shown in the middle of Figure 6(a), on the camera image plane, the azimuth angle of the surface normal points in the direction (shown as the red arrow) perpendicular to the iso-depth contour (shown as the orange dash circle). Even with the π-ambiguity unresolved (shown at the top-right in Figure 6(a)), we can still trace the iso-depth contour perpendicular to the azimuth angle direction.

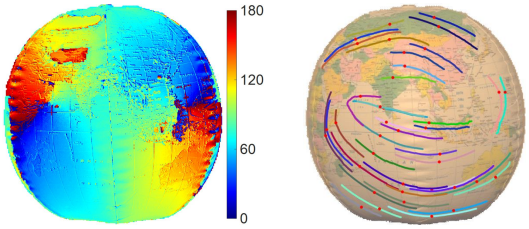
Therefore, in order to use the azimuth angles (with the π-ambiguity) for depth estimation, we trace the iso-depth contours from a set of sparse points with reliable depth estimated from initialization (*i.e.*, $p \in \mathcal{P}_+$). This will propagate depth from sparse points to featureless regions. Prior work [41] also exploited the idea of tracing iso-depth contours for 3D reconstruction in the context of photometric stereo.

More specifically, as shown in Figure 6 (b), to trace iso-depth contours, we select $N = 2000$ pixels from $\mathcal{P}_+$ (*i.e.*, with reliable depth) as seed points and trace them along the two directions perpendicular to the azimuth angle φ_p: $\vec{d}_+ = [\cos(\varphi_p + \frac{\pi}{2}), \sin(\varphi_p + \frac{\pi}{2})]$ and $\vec{d}_- = [\cos(\varphi_p - \frac{\pi}{2}), \sin(\varphi_p - \frac{\pi}{2})]$. We used a step size of 0.5 pixel for tracing. Since the tracing is imprecise at depth discontinuities, we stop the tracing once the change in the azimuth angles between two neighboring pixels is greater than a threshold ($\pi/6$ for all the experiments).

The tracing can propagate depth for most pixels in $\mathcal{P}_-$ in many cases. However, for scenes with large featureless regions or complex geometry, there may still be pixels left



(a) Iso-depth contour tracing with azimuth angles



(b) An example of tracing iso-depth contours

Figure 6: (a) The π-ambiguity in azimuth angles does not affect iso-depth contour tracing. (b) An example of tracing iso-depth contours for BALLOON.

with unreliable depth after tracing. To further optimize the depth for all pixels, we employ a similar approach as [26] to solve the depth map $d(x, y)$ by minimizing

$$\sum_{(x,y)\in\mathcal{P}} E_p(d(x,y)) + \gamma E_d(d(x,y)) + |\Delta d(x,y)|, \quad (10)$$

where $E_p(d)$ is the constraint derived from the azimuth angles, $E_d(d)$ is to discourage depth update for $(x, y) \in \mathcal{P}_+$, Δd is the Laplacian of the depth map, and $\gamma = 0.1$ in our experiments. Specifically, based on the definition of azimuth angles, we have $\tan(\varphi) = \frac{\partial d}{\partial y} / \frac{\partial d}{\partial x}$, *i.e.*,

$$\frac{\sin\varphi(x,y)}{\cos\varphi(x,y)} = \frac{d(x,y+1) - d(x,y)}{d(x+1,y) - d(x,y)}, \quad (11)$$

and thus

$$E_p(d(x,y)) = |\sin\varphi(x,y)\,(d(x+1,y) - d(x,y)) - \cos\varphi(x,y)\,(d(x,y+1) - d(x,y))\,|. \quad (12)$$

For point $(x, y) \in \mathcal{P}_+$, it has reliable initial depth estimate $\hat{d}(x, y)$, and $E_d(d(x, y))$ is to discourage depth update

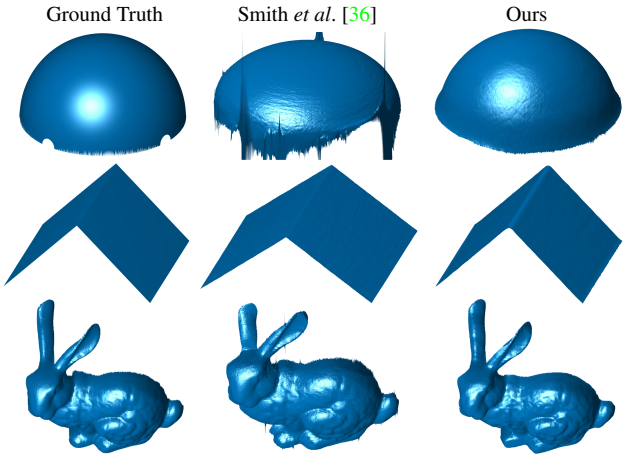
$$E_d(d(x,y)) = \begin{cases} |d(x,y) - \hat{d}(x,y)| & \text{if} \quad (x,y) \in \mathcal{P}_+ \\ 0 & \text{otherwise} \end{cases}. \quad (13)$$

$\Delta d(x, y)$ is computed as a 2D convolution with the 3×3 Lapacian filter $\frac{1}{12}\begin{pmatrix} 1 & 2 & 1 \\ 2 & -12 & 2 \\ 1 & 2 & 1 \end{pmatrix}$. All these are linear constraints, and thus the optimization problem can be solved efficiently with linear programming.

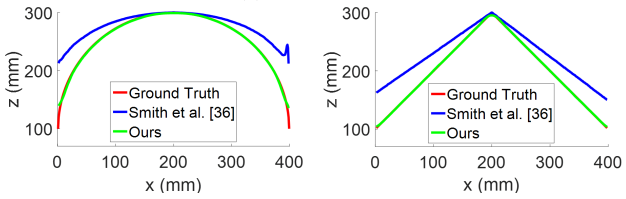
5. Experimental Results

5.1. Evaluation on Simulated Data

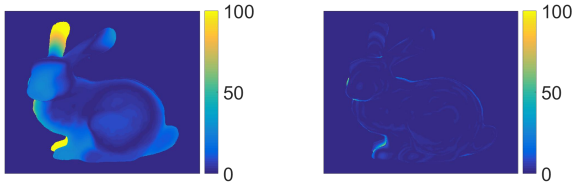
Three synthetic shapes, SPHERE, ROOF (with two planes) and BUNNY, are used for quantitative evaluation and comparison with [36]. Both the azimuth angles and the zenith



(a) 3D view



(b) 2D cross section view for SPHERE (left) and ROOF (right)



(c) Depth error (in mm) of [36] (left) and ours (right) for BUNNY

Figure 7: Quantitative evaluation and comparison with [36] on the synthetic examples.

angles are calculated from the ground truth shapes, with additive Gaussian noise ($\sigma_{azimuth} = 6°, \sigma_{zenith} = 3°$). For [36] we further provide the ground truth lighting direction and unpolarized intensities. For SPHERE and ROOF , we randomly select 50 pixels from a 400×400 image (less than 0.04%) with noisy depth (*i.e.*, ground truth depth plus 1% Gaussian noise) as seed points for tracing, and for BUNNY, we randomly select 1000 pixels from a 800×800 image because of its geometric details.

The estimation results by [36] and our method are shown in Figure 7. Figure 7(a) visualizes the recovered shape in a novel view. Figure 7(b)(c) show the depth errors of [36] and our method. For all the three examples, our method reconstructs more accurate depth than [36]. We also conducted noise analysis (please refer to the supplementary). We found that the linear constraints for integration in [36] seems sensitive to noise, while our method is more robust thanks to the iso-depth contour tracing.

5.2. Results on Real Data

We captured five scenes under both natural indoor and outdoor illumination — VASE (36 views), TILE (10 views), BALLOON (24 views), CORNER (6 views), and CAR (36 views). All images were captured using a Cannon EOS 7D camera with a $50mm$ lens. We mounted a Hoya linear polarizer in the front of the camera lens. For each view, seven images were captured with the polarizer angle spaced $30°$ degrees apart. Except for CORNER, we segmented the foreground objects from the background. Exemplar images and the camera poses recovered from VisualSFM are shown in the leftmost column of Figure 9.

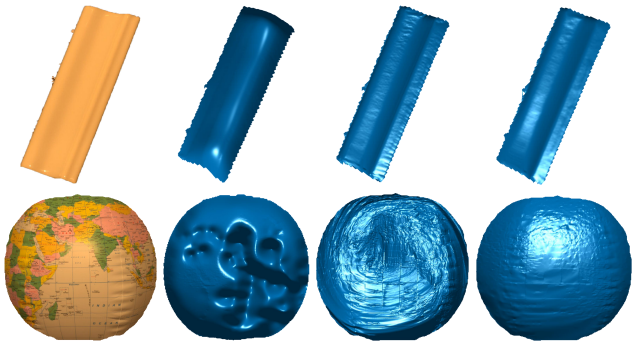


Figure 8: Comparison of depth estimation with Polarized3D [16] and Smith *et al.* [36].

Comparison for Depth Map Estimation We compare with two recent methods — Polarized3D [16], which requires rough depth from a RGB-D sensor, and Smith *et al.* [36], which requires lighting estimation. In addition, neither of two methods can handle mixed polarization. In order to make fair comparisons, we manually mask out the specular regions in the images. For Polarized3D [16], we used the initial depth map from MVS as the rough depth input. Results on the TILE and BALLOON datasets are shown in Figure 8.

As shown, Polarized3D [16] shows some artifacts in the estimated depth mainly due to the initial MVS depth is noisy in textureless regions. With a reliable depth input (*e.g.*, from a RGBD sensor as shown in [16]), it can achieve high accuracy. Moreover, the mixed polarization of the object still affects its performance even though we have masked out the specular highlight regions. Smith *et al.* [36] has high quality for the TILE as the azimuth and zenith angles can be accurately estimated for this data. However, their result for BALLOON is worse, as the estimated azimuth and zenith angles are noisy given the complicated texture. Our method shows superior performance on both two examples.

Comparison for Multi-View Stereo Finally, we show our final reconstructed 3D models with two state-of-the-art MVS methods — MVE [10] and Gipuma [11]. For fair comparison, we show the results after depth fusion for all the methods. The complete results are shown in Figure 9.

For the VASE and BALLOON, MVE [10] succeeds in reconstructing the whole shape, but the reconstructed point

clouds are quite noisy, due to inconsistent intensities across different views caused by specular highlights. Gipuma [11] cannot reconstruct the textureless areas near the neck of the vase or around the white part of the balloon. Similar artifacts can be observed for the middle part of TILE, and the white and black walls of CORNER, where both MVE [10] and Gipuma [11] show large holes. Our method generates more complete and smoother result for all these examples.

The outdoor data CAR is more challenging, since the car body is almost pure white with strong specularity. MVE [10] produces a reasonable reconstruction, but with many outliers. Gipuma [11] can only reconstruct a skeleton of the car. Our proposed method achieves the most complete and accurate 3D reconstruction of the car, thanks to the estimated phase angle from polarization and depth propagation. Nevertheless, none of the three methods can recover the car windows, because these parts are transparent — this is an interesting direction for further investigation.

We show additional comparison with a scanned ground truth shape and analyze the running time in the supplementary.

6. Conclusions and Discussions

We presented *polarimetric multi-view stereo*, a completely passive, novel approach for dense 3D reconstruction. Polarimetric multi-view stereo shows its strength especially for featureless regions and non-Lambertian surfaces, where it propagates depth estimated from well-textured regions to featureless regions guided by the azimuth angles estimated from polarized images. Extensive experimental results demonstrate high-quality 3D reconstruction and better performance than state-of-the-art multi-view stereo methods.

In addition, we also proposed a novel polarization imaging model that can handle real-world objects with mixed polarization. We proved that there are exactly two types of ambiguities for azimuth angle estimation from polarized images. These theoretical results are useful for further studies on shape-from-polarization.

The proposed polarimetric multi-view stereo has its limitations. First, it still requires a few points with reliable depth as seeds for depth propagation. Second, our current algorithm cannot recover transparent objects, despite some information being recovered from polarized images. We plan to investigate these directions in the future.

Acknowledgement

This study is partially supported by NVIDIA, Canada NSERC Discovery Grant 31-611664, Discovery Accelerator Supplement 31-611663, and a project commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

Appendix: Proof of Proposition 1

We briefly show the proof for Equations (1)(2) and Proposition 1 based on Muller calculus [6]. Due to page limit, please refer to the supplementary for a complete proof.

The polarization state of light can be represented by a 4×1 Stokes vector $\mathbf{S} = [S_0, S_1, S_2, S_3]^\top$ where S_0 is the energy of the light [6]. The effect of light-matter intersections (*e.g.*, reflection, transmission, polarizer) to the polarization state is represented with a 4×4 Muller matrix $\mathbf{M}$, which updates the Stokes vector from $\mathbf{S}$ to $\mathbf{MS}$.

As shown in Figure 2, there are two polarized components in the reflected light. The polarized specular reflection is from the air-object surface, denoted by $\mathbf{S}_{sp}$. The polarized diffuse reflection is from the refraction from the depolarized subsurface scattered light to air, denoted by $\mathbf{S}_{dp}$. Both components will be measured by the camera via a linear polarizer Let $\mathbf{S}_i$ be the Stokes vector for the illumination, $\mathbf{M}_{pol}(\theta)$ be the Muller matrix for the linear polarizer at angle θ, $\mathbf{M}_R$ and $\mathbf{M}_T$ denote the Muller matrices for Fresnel reflection and transmission, respectively. We have

$$\mathbf{S}_{sp} = \mathbf{M}_{pol}(\theta)\mathbf{M}_R\mathbf{S}_i, \qquad \mathbf{S}_{dp} = \mathbf{M}_{pol}(\theta)\mathbf{M}_T\mathbf{S}_d, \quad (14)$$

where $\mathbf{S}_d$ is the Stokes vector for the depolarized scattered light under surface. For unpolarized illumination, $\mathbf{S}_i = L_i[1, 0, 0, 0]$. $\mathbf{S}_d$ is also unpolarized due to random subsurface scattering, $\mathbf{S}_d = L_d[1, 0, 0, 0]$. $\mathbf{M}_R$ and $\mathbf{M}_T$ are the Muller-Stokes matrices for Fresnel equations [6]. For $\mathbf{M}_{pol}(\theta)$, we note θ is related to the polarization angle ϕ_{pol} of the linear polarizer and the azimuth angle φ of the surface normal by $\theta = \phi_{pol} + \frac{\pi}{2} - \varphi$. By definition, the measured radiance for both polarized specular reflection and polarized diffuse reflection are

$$I_{sp}(\phi_{pol}) = \mathbf{S}_{sp}(0), \qquad I_{dp}(\phi_{pol}) = \mathbf{S}_{dp}(0). \quad (15)$$

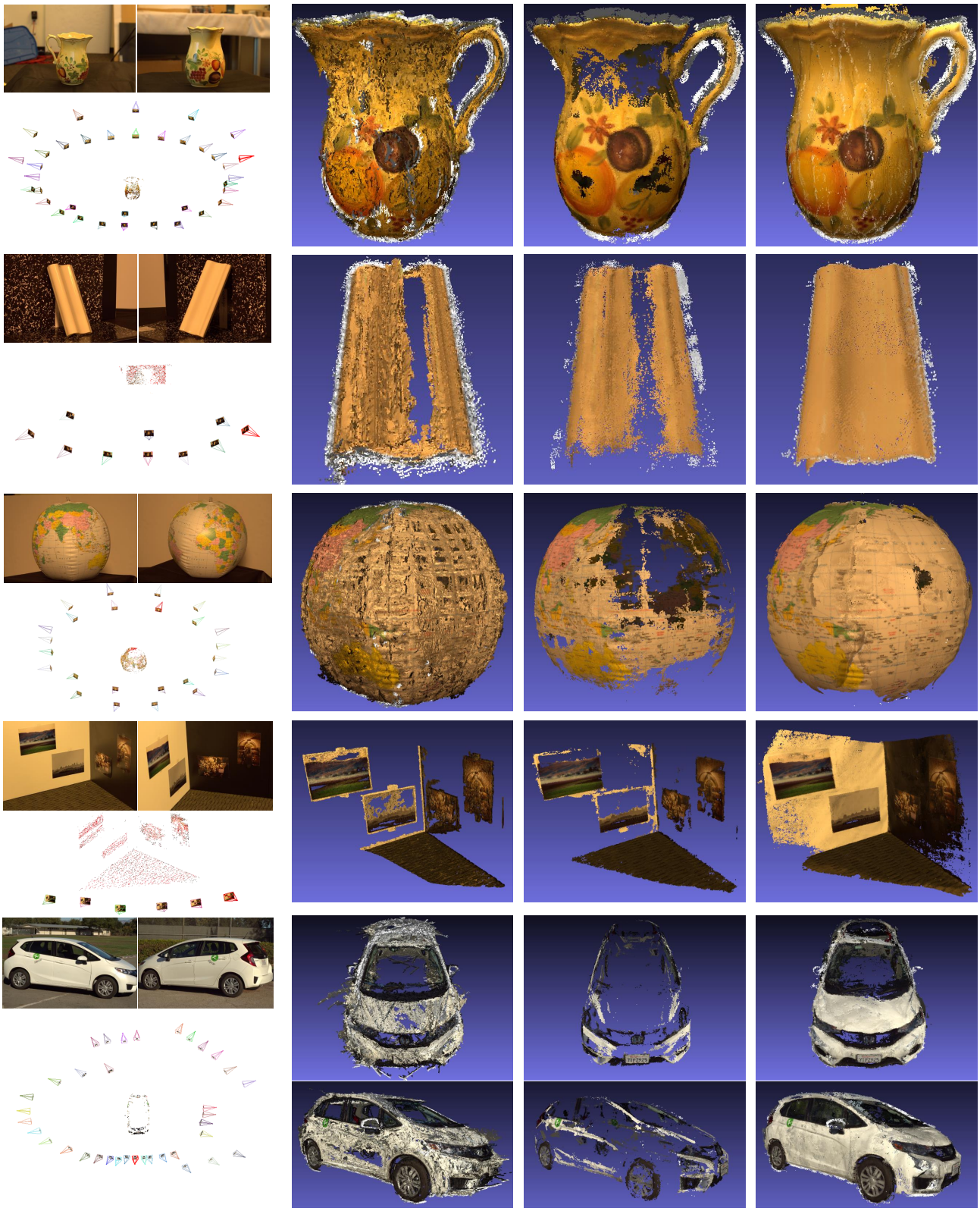
From the above two equations, together with the Muller matrices ($\mathbf{M}_R$, $\mathbf{M}_T$, $\mathbf{M}_{pol}(\theta)$) defined in [6], we can derive Equations (1)(2). The $\pi/2$ phase difference between I_{sp} and I_{dp} is caused by the $+/-$ sign for the sinusoid term $\cos(2(\phi_{pol} - \varphi))$ in Equation (15).

Many real-world objects, as shown in Figure 2, have both polarized specular reflection and polarized diffuse reflection, as well as unpolarized diffuse reflection.

$$I(\phi_{pol}) = I_d + I_{dp}(\phi_{pol}) + I_{sp}(\phi_{pol}), \quad (16)$$

where I_d is the unpolarized diffuse reflection that does not vary with the polarization angle ϕ_{pol}. By inserting Equation (1) and Equation (2) in Equation (16), we can derive the equations in Proposition 1.

The π-ambiguity is easy to see, since the period of $I(\phi_{pol}$ is π because of $\cos(2(\phi_{pol} - \varphi))$. The $\pi/2$-ambiguity appears, because the relative amount of I_{dp} and I_{sp} is unknown which can flip the $+/-$ sign of the sinusoid term $\cos(2(\phi_{pol} - \varphi))$ in Equation (16).



(a) Sample images (b) MVE [10] (c) Gipuma [11] (d) Ours

Figure 9: Comparison with state-of-the-art MVS methods [10, 11] for complete reconstruction. From **top** to **bottom** rows: VASE (36 views), TILE (10 views), BALLOON (24 views), CORNER (6 views), and CAR (36 views).

References

[1] 4D Technology polarizatoin camera. http://www.4dtechnology.com/products/polarimeters/polarcam/. 1, 3

[2] G. A. Atkinson and E. R. Hancock. Multi-view surface reconstruction using polarization. In *Proc. of Internatoinal Conference on Computer Vision*, 2005. 2

[3] G. A. Atkinson and E. R. Hancock. Recovery of surface orientation from diffuse polarization. *IEEE Transactions on Image Processing*, 15(6):1653–1664, 2006. 1, 2, 3

[4] M. Bleyer, C. Rhemann, and C. Rother. Patchmatch stereo - stereo matching with slanted support windows. In *Proc. of British Machine Vision Conference*, 2011. 2

[5] D. BrayfordD, M. Turner, and W. T. Hewitt. A Physical Model for the Polarized Scattering of Light. In *Theory and Practice of Computer Graphics*. The Eurographics Association, 2008. 2

[6] E. Collett. *Field Guide to Polarization*. SPIE, 2005. 2, 7

[7] C. Collin, S. Pattanaik, P. LiKamWa, and K. Bouatouch. Computation of polarized subsurface brdf for rendering. In *Graphics Interface*, 2014. 2

[8] O. Drbohlav and R. Šára. Unambiguous determination of shape from photometric stereo with unknown light sources. In *Proc. of Internatoinal Conference on Computer Vision*, 2001. 2

[9] FluxData polarizatoin camera. http://www.fluxdata.com/imaging-polarimeters. 1, 3

[10] S. Fuhrmann, F. Langguth, and M. Goesele. MVE-a multiview reconstruction environment. In *Proc. of the Eurographics Workshop on Graphics and Cultural Heritage*, 2014. 1, 2, 6, 7, 8

[11] S. Galliani, K. Lasinger, and K. Schindler. Massively parallel multiview stereopsis by surface normal diffusion. In *Proc. of Internatoinal Conference on Computer Vision*, 2015. 1, 2, 4, 6, 7, 8

[12] C. Hernández, G. Vogiatzis, and R. Cipolla. Multiview photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(3):548–554, 2008. 1, 2

[13] H. Hirschmüller. Stereo processing by semi-global matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):328–341, 2008. 2

[14] C. P. Huynh, A. Robles-Kelly, and E. R. Hancock. Shape and refractive index recovery from single-view polarisation images. In *Proc. of Computer Vision and Pattern Recognition*, 2010. 1, 2

[15] 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, 2013. 2

[16] A. Kadambi, V. Taamazyan, B. Shi, and R. Raskar. Polarized 3D: High-quality depth sensing with polarization cues. In *Proc. of Internatoinal Conference on Computer Vision*, 2015. 1, 2, 3, 5, 6

[17] V. Kolmogorov. Convergent tree-reweighted message passing for energy minimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(10):1568–1583, 2006. 4

[18] F. Langguth, K. Sunkavalli, S. Hadap, and M. Goesele. Shading-aware multi-view stereo. In *Proc. of European Conference on Computer Vision*, 2016. 1, 2

[19] D. A. LeMaster and M. T. Eismann. *Multi-dimensional Imaging*, chapter Passive Polarimetric Imaging. John Wiley & Sons, Ltd., 2014. 2

[20] A. H. Mahmoud, M. T. El-Melegy, and A. A. Farag. Direct method for shape recovery from polarization and shading. In *Proc. of International Conference on Image Processing*, 2012. 2

[21] D. Miyazaki, M. Kagesawa, and K. Ikeuchi. Transparent surface modeling from a pair of polarization images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(1):73–82, 2004. 2

[22] 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 *Proc. of 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT)*, 2012. 1, 2, 3

[23] D. Miyazaki, T. Shigetomi, M. Baba, R. Furukawa, S. Hiura, and N. Asada. Surface normal estimation of black specular objects from multiview polarization images. *Optical Engineering*, 56(4):041303, 2016. 2

[24] D. Miyazaki, R. T. Tan, K. Hara, and K. Ikeuchi. Polarization-based inverse rendering from a single view. In *Proc. of Internatoinal Conference on Computer Vision*, 2003. 2

[25] O. Morel, F. Meriaudeau, C. Stolz, and P. GorriaK. Polarization imaging applied to 3D reconstruction of specular metallic surfaces. In *Proc. of SPIE 5679, Machine Vision Applications in Industrial Inspection XIII*, 2005. 1, 2

[26] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi. Efficiently combining positions and nnormal for precise 3d geometry. *ACM Transactions on Graphics (Proc. of ACM SIGGRAPH)*, 24(3):536–543, 2005. 1, 2, 5

[27] T. T. Ngo, H. Nagahara, and R. Taniguchi. Shape and light directions from shading and polarization. In *Proc. of Computer Vision and Pattern Recognition*, 2015. 2

[28] G. Oxholm and K. Nishino. Multiview shape and reflectance from natural illumination. In *Proc. of Computer Vision and Pattern Recognition*, pages 2155–2162, 2014. 1, 2

[29] J. Park, S. N. Sinha, Y. Matsushita, Y.-W. Tai, and I. S. Kweon. Multiview photometric stereo using planar mesh parameterization. In *Proc. of Internatoinal Conference on Computer Vision*, 2013. 1, 2

[30] N. G. Parke. Optical algebra. *Journal of Mathematics and Physics*, 28, 1949. 2

[31] Photonic Lattice polarizatoin camera. https://www.photonic-lattice.com/en/products/polarization_camera/pi-110/. 1, 3

[32] S. Rahmann and N. Canterakis. Reconstruction of specular surfaces using polarization imaging. In *Proc. of Computer Vision and Pattern Recognition*, 2001. 2

[33] Ricoh polarizatoin camera. https://www.ricoh.com/technology/tech/051_polarization.html. 1, 3

[34] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multiview stereo reconstruction algorithms. In *Proc. of Computer Vision and Pattern Recognition*, 2006. 1, 2

[35] B. Semerjian. A new variational framework for multiview surface reconstruction. In *Proc. of European Conference on Computer Vision*, 2014. 1, 2

[36] W. A. P. Smith, R. Ramamoorthi, and S. Tozza. Linear depth estimation from an uncalibrated, monocular polarisation image. In *Proc. of European Conference on Computer Vision*, 2016. 1, 2, 3, 5, 6

[37] V. Taamazyan, A. Kadambi, and R. Raskar. Shape from mixed polarization. In *arXiv:1605.02066*, 2016. 2, 3

[38] L. B. Wolff. Polarization vision: a new sensory approach to image understanding. *Image Vision Computing*, 15(2):81–93, 1997. 1

[39] C. Wu. Towards linear-time incremental structure from motion. In *Proc. of International Conference on 3D Vision*, 2013. 4

[40] C. Wu, B. Wilburn, Y. Matsushita, and C. Theobalt. High-quality shape from multi-view stereo and shading under general illumination. In *Proc. of Computer Vision and Pattern Recognition*, 2011. 1, 2

[41] Z. Zhou, Z. Wu, and P. Tan. Multi-view photometric stereo with spatially varying isotropic materials. In *Proc. of Computer Vision and Pattern Recognition*, 2013. 1, 2, 5

Polarimetric Multi-View Stereo Supplementary Material

Zhaopeng Cui[1] Jinwei Gu[2] Boxin Shi[3] Ping Tan[1] Jan Kautz[2]
[1]Simon Fraser University [2]NVIDIA Research
[3]Artificial Intelligence Research Center, National Institute of AIST

The supplementary material contains the complete proof of the appendix in our paper and more experimental results as follows: Section 1 describes the detailed proof of Equations (1) and (2) as well as Proposition 1 in the main paper based on the Muller calculus [1]; Section 2 presents more quantitative evaluation on the synthetic data, the running time and more experiment on real data with the laser scanned ground truth.

1. Proof of Proposition 1

The polarization state of light can be represented by a 4×1 Stokes vector $\mathbf{S} = [S_0, S_1, S_2, S_3]^\top$ where S_0 describes the total intensity of the light, S_1 is the intensity difference between polarized components of electromagnetic wave parallel and perpendicular to the reference plane, S_2 indicates the intensity difference between polarized components in planes $45°$ and $-45°$ to the reference plane, and S_3 describes the circularly polarized radiation [1]. The effect of light-matter intersections (*e.g.*, reflection, transmission, polarizer) to the polarization state is represented with a 4×4 Muller matrix $\mathbf{M}$. When a beam passes through a polarizing element, its polarization state changes from $\mathbf{S}$ to $\mathbf{MS}$.

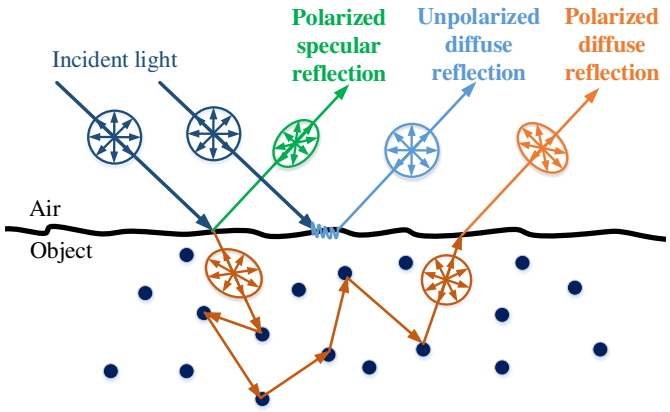


Figure 1: A diagram of surface reflection with mixed polarization. Reflected radiance for many surfaces includes three parts: **(1)** the polarized specular reflection (*i.e.*, highlight), **(2)** the polarized diffuse reflection (due to subsurface scattering and refraction), and **(3)** the unpolarized diffuse reflection (due to micro-facet rough surface reflection). The polarized specular reflection and the polarized diffuse reflection have a $\pi/2$ difference in phase angle. The circles with arrows show the polarization status: round circles – unpolarized, elliptic circles – partially polarized.

As shown in Figure 1, there are two polarized components in the reflected light. The polarized specular reflection is from the air-object surface, denoted by $\mathbf{S}_{sp}$. The polarized diffuse reflection is from the refraction from the depolarized subsurface scattered light to air, denoted by $\mathbf{S}_{dp}$. Both components will be measured by the camera via a linear polarizer. Let $\mathbf{S}_i$ be the Stokes vector for the illumination, $\mathbf{M}_{pol}(\theta)$ be the Muller matrix for the linear polarizer at angle θ, $\mathbf{M}_R$ and $\mathbf{M}_T$ denote the Muller matrices for Fresnel reflection and transmission, respectively. We have

$$\mathbf{S}_{sp} = \mathbf{M}_{pol}(\theta)\mathbf{M}_R\mathbf{S}_i, \qquad \mathbf{S}_{dp} = \mathbf{M}_{pol}(\theta)\mathbf{M}_T\mathbf{S}_d, \tag{1}$$

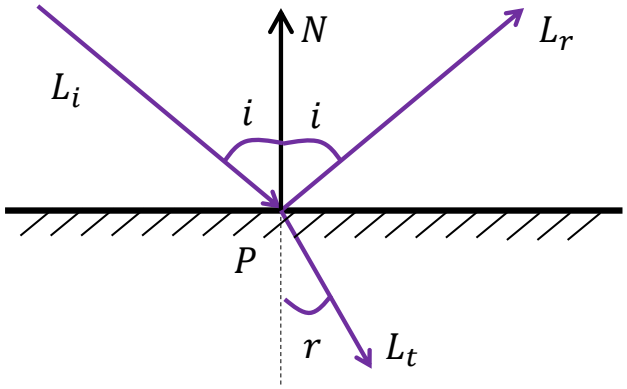


Figure 2: Reflection and transmission for polarized light.

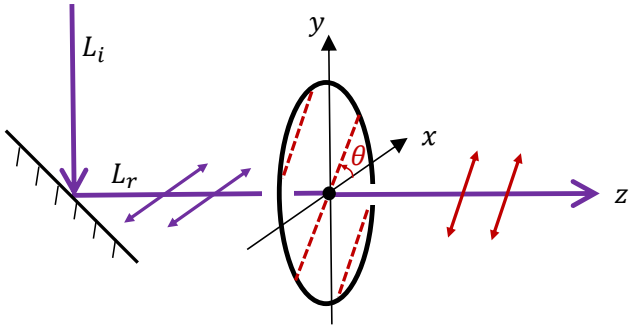


Figure 3: The rotation angle θ of the polarizer in Equation (4) is defined as the angle between the polarization direction of the linear polarizer (*i.e.*, the red dash line) and the direction perpendicular to the plane of the incident illumination (*i.e.*, the x axis in this figure).

where $\mathbf{S}_d$ is the Stokes vector for the depolarized scattered light under surface.

For unpolarized illumination, $\mathbf{S}_i = L_i[1,0,0,0]$. $\mathbf{S}_d$ is also unpolarized due to random subsurface scattering, $\mathbf{S}_d = L_d[1,0,0,0]$. $\mathbf{M}_R$ and $\mathbf{M}_T$ are the Muller-Stokes matrices for Fresnel equations [1]. As shown in Figure 2, we have [1]

$$\mathbf{M}_R = f_R \begin{bmatrix} \cos^2\alpha_- + \cos^2\alpha_+ & \cos^2\alpha_- - \cos^2\alpha_+ & 0 & 0 \\ \cos^2\alpha_- - \cos^2\alpha_+ & \cos^2\alpha_- + \cos^2\alpha_+ & 0 & 0 \\ 0 & 0 & -2\cos\alpha_-\cos\alpha_+ & 0 \\ 0 & 0 & 0 & -2\cos\alpha_-\cos\alpha_+ \end{bmatrix}, \tag{2}$$

where $\alpha_\pm = i \pm r$ and $f_R = \frac{1}{2}\left(\frac{\tan\alpha_-}{\sin\alpha_+}\right)^2$, and

$$\mathbf{M}_T = f_T \begin{bmatrix} \cos^2\alpha_- + 1 & \cos^2\alpha_- - 1 & 0 & 0 \\ \cos^2\alpha_- - 1 & \cos^2\alpha_- + 1 & 0 & 0 \\ 0 & 0 & -2\cos\alpha_- & 0 \\ 0 & 0 & 0 & -2\cos\alpha_- \end{bmatrix}, \tag{3}$$

where $f_T = \frac{1}{2}\frac{\sin 2i \sin 2r}{(\sin\alpha_+\cos\alpha_-)^2}$.

$\mathbf{M}_{pol}(\theta)$ is the Muller matrix for a rotated linear polarizer with angle θ. From [1], for an ideal rotated linear polarizer, we have

$$\mathbf{M}_{pol}(\theta) = \frac{1}{2}\begin{bmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \\ \cos 2\theta & \cos^2 2\theta & \sin 2\theta\cos 2\theta & 0 \\ \sin 2\theta & \sin 2\theta\cos 2\theta & \sin^2 2\theta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{4}$$

As shown in Figure 3, the angle θ is defined as the rotation angle between the polarization direction of the linear polarizer (*i.e.*, the red dash line) and the direction perpendicular to the plane of the incident illumination (*i.e.*, the x axis in Figure 3).

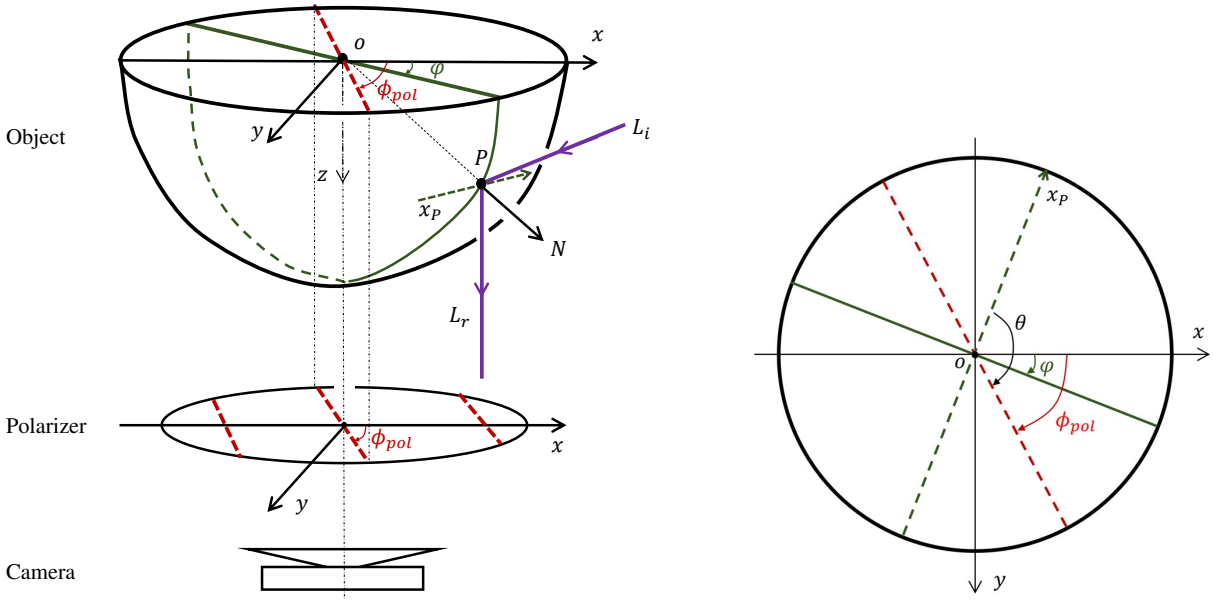


Figure 4: **Left**: A diagram of imaging a 3D object through a linear polarizer. The light reflected from a point P with surface normal N has two polarized reflection components, defined in Equation (9) and Equation (11) (*i.e.*, Equations (2) and (1) in the main paper), respectively. **Right**: Definitions of the three angles, φ, ϕ_{pol}, and θ, in the $x-y$ plane.

Consider imaging an object through a linear polarizer, as shown in Figure 4. For a point P on the object surface, suppose its surface normal is N. Let φ denote the azimuth angle of P, and ϕ_{pol} denote the angle between the polarization direction (*i.e.*, red dash line) of the polarizer and the x axis. Note that for the reflection and refraction at point P, the plane of the incident illumination is defined by surface normal N and the incident illumination L_i. The direction perpendicular to this plane is x_P. Thus, based on the definition in Figure 3, the rotation angle θ at point P is the angle between the polarization direction (*i.e.*, red dash line) and the direction x_P, and thus θ is given by

$$\theta = \phi_{pol} + \frac{\pi}{2} - \varphi. \tag{5}$$

The right side of Figure 4 shows a 2D view of the $x-y$ plane, with clear definitions of these three angles.

By the definition of the Stokes vector, the measured radiance for both polarized specular reflection and polarized diffuse reflection are the first element in the Stokes vectors,

$$I_{sp}(\phi_{pol}) = \mathbf{S}_{sp}(0), \qquad I_{dp}(\phi_{pol}) = \mathbf{S}_{dp}(0). \tag{6}$$

By putting Equation (2), Equation (3), Equation (4), and Equation (5) into Equation (1) and Equation (6), we can derive Equations (1) and (2) in the main paper. More specifically, we have

$$\mathbf{S}_{sp} = \mathbf{M}_{pol}(\theta)\mathbf{M}_R\mathbf{S}_i = \frac{1}{2}\begin{pmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \\ \cos 2\theta & \cos^2 2\theta & \sin 2\theta \cos 2\theta & 0 \\ \sin 2\theta & \sin 2\theta \cos 2\theta & \sin^2 2\theta & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \frac{L_i}{2}\left(\frac{\tan\alpha_-}{\sin\alpha_+}\right)^2 \begin{pmatrix} \cos^2\alpha_- + \cos^2\alpha_+ \\ \cos^2\alpha_- - \cos^2\alpha_+ \\ 0 \\ 0 \end{pmatrix} \tag{7}$$

and

$$\mathbf{S}_{dp} = \mathbf{M}_{pol}(\theta)\mathbf{M}_T\mathbf{S}_d = \frac{1}{2}\begin{pmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \\ \cos 2\theta & \cos^2 2\theta & \sin 2\theta \cos 2\theta & 0 \\ \sin 2\theta & \sin 2\theta \cos 2\theta & \sin^2 2\theta & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \frac{L_d}{2}\frac{\sin 2i \sin 2r}{(\sin\alpha_+ \cos\alpha_-)^2} \begin{pmatrix} \cos^2\alpha_- + 1 \\ \cos^2\alpha_- - 1 \\ 0 \\ 0 \end{pmatrix} \tag{8}$$

From Equation (5), we have $\theta = \phi_{pol} + \frac{\pi}{2} - \varphi$. Thus, we have

$$\begin{aligned} I_{sp}(\phi_{pol}) = \mathbf{S}_{sp}(0) &= \frac{L_i}{4}\left(\frac{\tan\alpha_-}{\sin\alpha_+}\right)^2 \left(\left(\cos^2\alpha_- + \cos^2\alpha_+\right) + \left(\cos^2\alpha_- - \cos^2\alpha_+\right)\cos 2\theta\right) \\ &= \frac{I^{sp}_{max} + I^{sp}_{min}}{2} + \frac{I^{sp}_{max} - I^{sp}_{min}}{2}\cos(2\theta) \\ &= \frac{I^{sp}_{max} + I^{sp}_{min}}{2} + \frac{I^{sp}_{max} - I^{sp}_{min}}{2}\cos(2(\phi_{pol} - \varphi + \frac{\pi}{2})), \end{aligned} \tag{9}$$

where

$$I^{sp}_{max} = \frac{L_i}{2}\left(\frac{\tan\alpha_-}{\sin\alpha_+}\right)^2 \cos^2\alpha_-, \qquad I^{sp}_{min} = \frac{L_i}{2}\left(\frac{\tan\alpha_-}{\sin\alpha_+}\right)^2 \cos^2\alpha_+. \tag{10}$$

Similarly, we have

$$\begin{aligned} I_{dp}(\phi_{pol}) = \mathbf{S}_{dp}(0) &= \frac{L_d}{4}\frac{\sin 2i \sin 2r}{\left(\sin\alpha_+ \cos\alpha_-\right)^2}\left(\left(\cos^2\alpha_- + 1\right) + \left(\cos^2\alpha_- - 1\right)\cos 2\theta\right) \\ &= \frac{I^{dp}_{max} + I^{dp}_{min}}{2} + \frac{I^{dp}_{max} - I^{sp}_{min}}{2}\cos(2(\theta - \pi/2)) \\ &= \frac{I^{dp}_{max} + I^{dp}_{min}}{2} + \frac{I^{dp}_{max} - I^{dp}_{min}}{2}\cos(2(\phi_{pol} - \varphi)), \end{aligned} \tag{11}$$

where

$$I^{dp}_{max} = \frac{L_d}{2}\frac{\sin 2i \sin 2r}{\left(\sin\alpha_+ \cos\alpha_-\right)^2}, \qquad I^{dp}_{min} = \frac{L_d}{2}\frac{\sin 2i \sin 2r}{\left(\sin\alpha_+ \cos\alpha_-\right)^2}\cos^2\alpha_-. \tag{12}$$

Note that Equation (9) and Equation (11) are exactly Equation (2) and Equation (1) in the main paper. Many real-world objects have both the polarized specular reflection and the polarized diffuse reflection, as well as an unpolarized diffuse reflection. So we have

$$I(\phi_{pol}) = I_d + I_{dp}(\phi_{pol}) + I_{sp}(\phi_{pol}), \tag{13}$$

where I_d is the unpolarized diffuse reflection that does not vary with the polarization angle ϕ_{pol}. By inserting Equation (9) and Equation (11) in Equation (13) and considering $\cos(x \pm \pi) = -\cos(x)$, we have

$$\begin{aligned} I(\phi_{pol}) &= I_d + \frac{I^{dp}_{max} + I^{dp}_{min}}{2} + \frac{I^{dp}_{max} - I^{dp}_{min}}{2}\cos(2(\phi_{pol} - \varphi)) + \frac{I^{sp}_{max} + I^{sp}_{min}}{2} + \frac{I^{sp}_{max} - I^{sp}_{min}}{2}\cos(2(\phi_{pol} - \varphi + \frac{\pi}{2})), \\ &= \frac{I_{max} + I_{min}}{2} + \frac{I_{max} - I_{min}}{2}\cos(2(\phi_{pol} - \phi)), \end{aligned} \tag{14}$$

where ϕ is defined as the phase angle, I_{max} and I_{min} are the maximum and minimum observed intensities. When polarized diffuse reflection dominates ($\frac{I^{dp}_{max} - I^{dp}_{min}}{2} > \frac{I^{sp}_{max} - I^{sp}_{min}}{2}$), we have

$$\phi = \varphi, \quad I_{max} = I_d + I^{dp}_{max} + I^{sp}_{min}, \quad I_{min} = I_d + I^{dp}_{min} + I^{sp}_{max}. \tag{15}$$

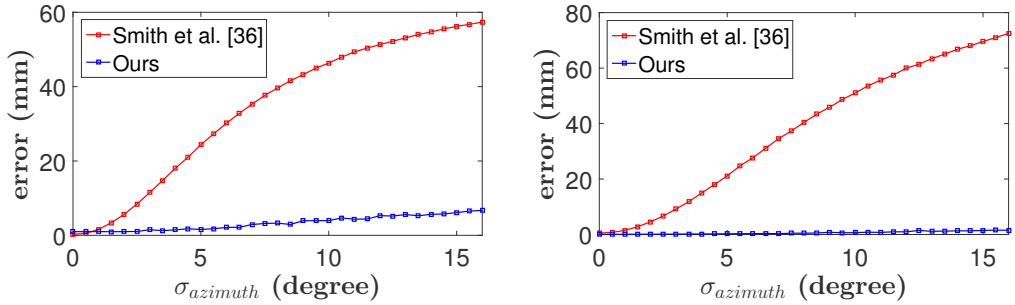
When polarized specular reflection dominates ($\frac{I^{sp}_{max} - I^{sp}_{min}}{2} > \frac{I^{dp}_{max} - I^{dp}_{min}}{2}$), we have

$$\phi = \varphi - \frac{\pi}{2}, \quad I_{max} = I_d + I^{sp}_{max} + I^{dp}_{min}, \quad I_{min} = I_d + I^{sp}_{min} + I^{dp}_{max}. \tag{16}$$

From Equations (14), (15) and (16), we have Proposition 1 in the main paper.

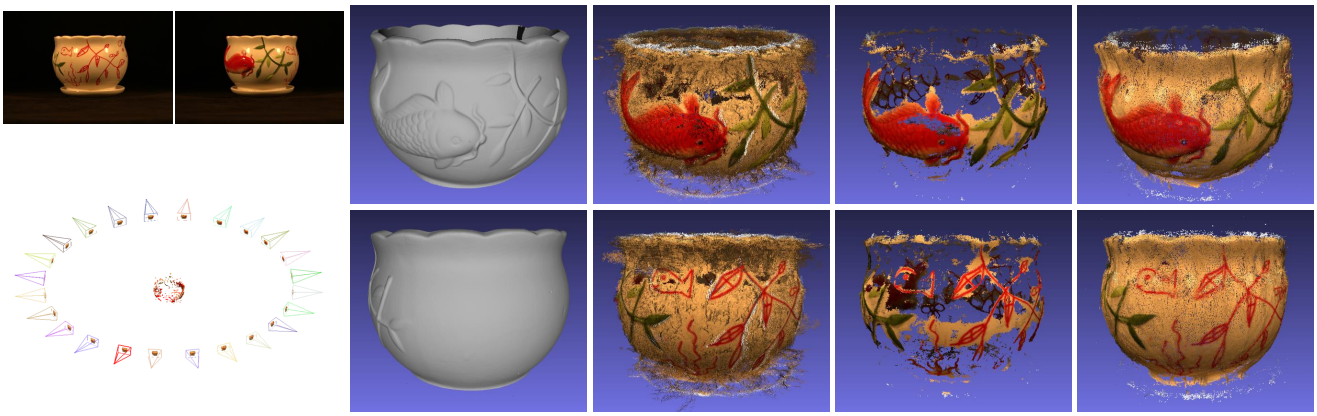
Structure-from-Motion:	23
Phase Angle Estimation:	25
Initialization (Depth Estimation):	352
(*)Resolving $\pi/2$-Ambiguity:	1096
(*)Depth Propagation:	132
(*)Depth Optimization:	1475
Depth Fusion:	215
Total:	3318

Table 1: Running time (in seconds) on VASE. Note that the steps with (*) are not optimized as we currently computed them sequentially for each view. They can be easily parallelized.



(a) SPHERE (b)ROOF

Figure 5: Average reconstruction errors of the synthetic examples with varying noise in the azimuth angle.



(a) Sample images (b) Ground truth (c) MVE [10] (d) Gipuma [11] (d) Ours

Figure 6: Comparison with state-of-the-art MVS methods [10, 11] for complete reconstruction.

2. More experimental results

We conducted the quantitative evaluation against noises on the synthetic data SPHERE and ROOF. As our method maintains only one common parameter the azimuth angle with [36], for fair comparison, we set all other parameters to be the ground truth and add a Gaussian noise to the azimuth angle map with varying $\sigma_{azimuth}$ from 0 to 16 degrees with a step of 0.5 degree. The mean error is used for comparison. As it is shown in Figure 5, our method shows more stable performance under various noises than [36].

Figure 6 shows the comparison of our method and two state-of-the-art MVS methods [10, 11] on a real dataset with the laser scanned ground truth. As there are obvious specular highlights and featureless parts on the body of the flowerpot, Gipuma [11] fail to reconstruct large parts of the flowerpot. The result of MVE [10] is noisy and it also has larges holes in some featureless parts (e.g. the parts above the red fish). In contrast, our method generates more complete and smooth model.

We tested the running time of our method for one of our datasets (VASE) on a desktop PC (two 2.3GHz Intel Xeon E5-2650 CPUs and one NVIDIA Quadro K5200 GPU). The result is listed in Table 1.

References

[1] E. Collett. *Field Guide to Polarization*. SPIE, 2005. 1, 2