# SHAPE FROM POLARIZATION FOR FEATURELESS AND SPECULAR OBJECTS

Fuduo Xue[a,b*], Sagi Filin[b], Bashar Elnashef[b], Weiqi Jin[a]

[a] MOE Key Laboratory of Optoelectronic Imaging Technology and System, Beijing Institute of Technology, Beijing, China
[b] Mapping and Geoinformation Engineering, Technion – Israel Institute of Technology, Haifa, Israel
(fuduo.xue, basharnashef)@campus.technion.ac.il, filin@technion.ac.il, jinwq@bit.edu.cn

**Technical Commission II**

**KEY WORDS:** 3-D reconstruction, Shape from Polarization, Textureless surface, Specular reflection, Normal estimation.

**ABSTRACT:**

Objects with textureless and specular surfaces are commonplace in manmade environments. Their reconstruction challenges optical 3-D metrology methods typically yielding incomplete and noisy results. Existing solutions, including structured light, shape from shading, and learning-based specular removal methods, place a high demand on projection patterns, the number of light sources and cameras, or training data. While requiring elaborate setups, these solutions are not general enough, and their success in handling these surfaces is only partial. To address this reconstruction challenge, we study in this paper the application of shape from polarization for the 3-D reconstruction of textureless objects. Using a single view and a known light source, polarization-based constraints can be expressed as a set of linear equations for the unknown depth. We then estimate depth directly as an optimization problem constrained by the linear polarization equations. Results demonstrate complete and accurate 3-D reconstructions of typical glossy featureless objects, suggesting that shape from polarization is a valuable strategy for generating dense 3-D surface models.

## 1. INTRODUCTION

The reconstruction of objects with textureless and specular surfaces is an open optical 3-D metrology problem. As objects of this kind inherently lack texture or sharp geometric shapes, their reconstruction by traditional image-based strategies, is partial often leaving large voids in and around these regions (Fig. 1, c.f. Li et al., 2015; Cui et al., 2017, 2019; Cheng et al., 2021; Fukao et al., 2021). An additional factor affecting the quality of the reconstruction are the specular reflections that generate highlights or fake textured regions (Zhang et al., 2021; Li et al., 2022). The application of active sensing methods, e.g., laser scanners and time-of-flight cameras, cannot secure reconstruction in these specular areas as the light is reflected in the specular direction rather than being scattered to the sensor (Kadambi et al., 2015; Lei et al., 2021). As modeling such objects finds use in various fields, e.g., manufacturing, cultural heritage-related studies, and graphics, solving this challenging reconstruction-related problem is of value.

A long-standing reconstruction solution for featureless object has relied on projecting matchable image patterns on the objects surface as a means to generate radiometric related cues (Koutsoudis et al., 2015; Kurth et al., 2018; Alzuhiri et al., 2021). However, as specular reflection produces highlights, the projection over such regions may lead to image saturation due to overexposure. To mitigate the consequential saturation, Wu et al. (2021) and Li et al. (2022) have proposed generating a well-exposed image over the entire object surface by acquiring image sequences with different exposure times, but remaining problems exist with the characteristics of the projected pattern, which have been demonstrated to affect the reconstruction results. As Hafeez et al. (2018, 2020) indicated, for the feature pattern to be detected, the type, gray hue, and density of the projected pattern need to be carefully designed in accordance with
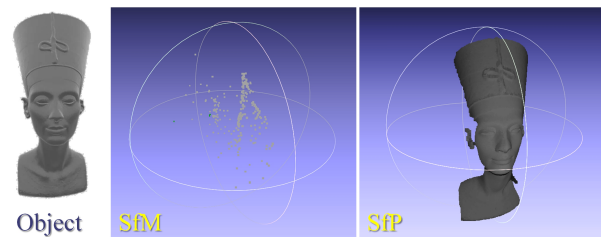


Figure 1. 3-D reconstructions using SfM and SfP-based models. Despite using 50 images, the SfM could not effectively extract and match object surface features. In contrast, the SfP completely reconstructed this specular featureless object.

the object shape and reflective properties, making the overall process empirical.

An alternative technique for reconstructing featureless surfaces lies in applying shape-from-shading (SfS) based solutions (Wu and Narasimhan, 2010; Barron and Malik, 2015). Based on the photometric equations and BRDF models, the relations between the image brightness and surface normal/depth are utilized to facilitate this reconstruction. Nonetheless, as SfS is an inherently under-constrained problem, strong assumptions are traditionally being made about the characteristics of the diffused reflections, such as the albedo (Bednarik et al., 2018). To handle specular surfaces, Wang et al. (2020) introduced a Blinn-Phong reflectance model describing the relationship between the specularly reflected radiance and surface depth under an orthographic camera projection assumption. Cheng et al. (2021) reformulated the spatially uniform BRDF describing the Lambertian surface in a logarithmic form to express the relationship between the specular reflection and surface normal, enabling to solve the depth using three cameras with a known rigidly attached light source. A limitation of SfS is that the BRDF capable of describing specular reflection has multiple underdetermined parameters. To apply the SfS, these

---

* Corresponding author

parameters should be estimated, and for a sufficient number of constraints, multiple camera positions or known light sources are required.

Learning-based solutions addressing specularity have been another avenue of research. Wu et al. (2018) proposed a generative adversarial network (GAN) to transfer a specular object image into the same appearance as the regular diffuse form, thereby facilitating the application of SfM and SfS methods. As the training images were synthetic and used the same albedo, the proposed network could not consistently convert specular materials with spatially varying albedos to diffuse surfaces. Furthermore, misclassifying the foreground as part of the background also led to incomplete reconstruction. To address false appearances generated by GAN-based methods, recent research attempted to analyze the difference between diffuse surfaces and specular highlights and separate the image into these two components. In Xia et al. (2019), the differences in hue and saturation in the diffuse and specular regions were analyzed, and then a training-based global separation algorithm was proposed. Fu et al. (2021) proposed a dataset of about 16K real images, each with a corresponding ground truth of highlight mask, aiming to train a CNN for the highlight removal in complex scenes. Despite the large volume of data, such an approach is still too difficult to generalize to different shapes and illuminations.

In this paper, we study the 3-D reconstruction of textureless and specular objects using a shape from polarization (SfP) related modeling form. SfP utilizes the observation that light information can reveal the relationships between reflection and surface shape independent of its properties through polarization constraints. Therefore, it provides means for 3-D reconstruction of featureless and specular objects (Cui et al., 2017; Ba et al., 2020; Lei et al., 2021). An inherent challenge associated with the reconstruction by the polarimetric constraints is the normal ambiguity that the polarimetric forms generate. To address this ambiguity, Tozza et al. (2017); Smith et al. (2019); Fukao et al. (2021) combined the polarimetric and photometric constraints, while Kadambi et al. (2015) and Zhu and Smith (2019) introduced a coarse geometric normal map. We follow here Smith et al. (2019) approach that addresses this ambiguity through a setup consisting of a polarization camera and a light source with a known position. We modify the orthographic projection that was considered there with the perspective camera model. Using this setup, full polarimetric and photometric constraints can be imposed on the polarization parameters. These constraints can then be formulated as a system of linear equations for the surface depth and normals. The outcome is a pixel-level dense monocular 3-D reconstruction model independent of surface texture, which exhibits stability in featureless regions and ones characterized by specular reflection. Both simulated and real-world results demonstrate that SfP can achieve complete and accurate reconstructions of glossy featureless objects.

## 2. METHODOLOGY

### 2.1 Camera-geometry-related relations

Our developed forms are in defined in reference to the calibrated polarization camera frame and all image measurements are referenced to the principal point. Following Zhu and Smith (2019) our aim is to express the surface normal using image quantities. To begin, we define the normalized viewing direction $\bar{\mathbf{v}} = [v_x, \ v_y, \ v_z]^{\mathrm{T}}$,

$$\bar{\mathbf{v}} = [x/f, \ y/f, \ 1]^{\mathrm{T}} / \, \| [x/f, \ y/f, \ 1] \| \qquad (1)$$

where $(x, y)$ is the image point coordinate, and $f$ is the principal distance. The surface point $\mathbf{P}$ in a calibrated camera frame can be defined as a function of its unknown depth $Z$ by:

$$\mathbf{P} = \begin{bmatrix} xZ/f & yZ/f & Z \end{bmatrix}^{\mathrm{T}}, \qquad (2)$$

Assuming a relatively narrow field of view (so that the image rays are paraxial), and a largely frontoparallel imaging setup, the surface normal $\mathbf{n} = [n_x, \ n_y, \ n_z]^{\mathrm{T}}$ at $\mathbf{P}$ can be expressed by the cross product of its partial derivatives with respect to $x$ and $y$ (Graber et al., 2015; Zhu and Smith, 2019):

$$\mathbf{n} = \frac{\partial \mathbf{P}}{\partial x} \times \frac{\partial \mathbf{P}}{\partial y} = \left[ -f \frac{\partial Z}{\partial x}, -f \frac{\partial Z}{\partial y}, x \frac{\partial Z}{\partial x} + y \frac{\partial Z}{\partial y} + Z \right]^{\mathrm{T}} (3)$$

while $\bar{\mathbf{n}} = \mathbf{n} / \|\mathbf{n}\|$ defines the normalized surface normal. Eq. (3) establishes the relation between the image data and the surface normal. It is used in Sec. (2.6) for the linear reconstruction solution, and with its definition, we turn to describe the polarization image formation process and demonstrate how it integrates with the normal constraints to yield the observation for the SfP equations.

### 2.2 Polarization imaging

According to Fresnel equations, an unpolarized light becomes partially linearly polarized when reflected from a surface (Goldstein, 2011). Consequently, a polarization image can inform on the object surface shape at each pixel. To compute these surface-related parameters, a sequence of images is acquired in which a linear polarizer positioned in front of the camera is rotated to yield $n \geq 3$ images, each with polarizer angle $\phi_i$, $i \in \{1, ..., n\}$. The intensity, $I_i$, at a single image pixel, varies sinusoidally as a function of $\phi_i$ according to:

$$I_i(\phi_i) = \frac{I_{max} + I_{min}}{2} + \frac{I_{max} - I_{min}}{2} \cos\left[2\left(\phi_i - \phi_{aop}\right)\right]$$
$$= \bar{I} + \bar{I} \cdot \rho \cos\left[2\left(\phi_i - \phi_{aop}\right)\right] \qquad (4)$$

where $I_{max}$ and $I_{min}$ are the respective maximum and minimum intensities obtained by rotating the polarizer; $\bar{I} = (I_{max} + I_{min})/2$ is the average intensity; $\rho = (I_{max} - I_{min})/(I_{max} + I_{min}) \in [0, 1]$ is the degree of polarization (DoLP); and $\phi_{aop} \in [0, \pi)$ is the angle of polarization (AoP), where $I(\phi_{aop}) = I_{max}$ (Goldstein, 2011). The three parameters, $\bar{I}$, $\rho$, and $\phi_{aop}$, completely describe the polarization state of the incoming light and can be estimated by Eq. (4). The SfP task is to parameterize the object surface shape as a function of these parameters.

### 2.3 Polarimetric surface constraints

The polarimetric surface constraints associate the surface normal parameters with the observed polarization state, specifically the DoLP and the AoP. They are expressed now for both specular and diffuse dominant surface facets.

**DoLP constraint** Substituting Fresnel equations (Goldstein, 2011) into the definition of the DoLP, we can obtain the DoLP at each point as a function of the refractive index $\eta$ and the angle of incidence between the viewing direction $\bar{\mathbf{v}}$ and the surface normal $\bar{\mathbf{n}}$, such that $\theta_i = \arccos(\bar{\mathbf{n}} \cdot \bar{\mathbf{v}}) \in [0, \frac{\pi}{2}]$. The DoLP for specular and diffuse dominant surface facets is different, and as demonstrated in Atkinson and Hancock (2006),

for a diffuse dominated surface its expression reads:

$$\rho_d = \frac{(\eta - 1/\eta)^2 \sin^2\theta_i}{2 + 2\eta^2 - (\eta - 1/\eta)^2 \sin^2\theta_i + 4\cos\theta_i\sqrt{\eta^2 - \sin^2\theta_i}} \tag{5}$$

and has only a single solution for $\theta_i$, given $\rho_d$. Therefore, the incident angle of a diffused surface can be uniquely solved. Following manipulations of Eq. (5) to express $\cos\theta_i$ as a function of the observed $\rho_d$ and $\eta$, we have:

$$\cos\theta_i = \tag{6}$$
$$\sqrt{\frac{\eta^4(1-\rho_d^2)+2\eta^2(2\rho_d^2+\rho_d-1)+\rho_d^2+2\rho_d-4\eta^3\rho_d\sqrt{1-\rho_d^2}+1}{(\rho_d+1)^2(\eta^4+1)+2\eta^2(3\rho_d^2+2\rho_d-1)}}$$

and as $\cos\theta_i = \bar{\mathbf{n}} \cdot \bar{\mathbf{v}}$ we obtain a normal-related constraint for a diffuse dominated surface facet.

The DoLP for specular surface reflection has the form (Tozza et al., 2017):

$$\rho_s = \frac{2\sin^2\theta_i \cos\theta_i \sqrt{\eta^2 - \sin^2\theta_i}}{\eta^2 - \sin^2\theta_i - \eta^2\sin^2\theta_i + 2\sin^4\theta_i}. \tag{7}$$

and note that two possible solutions exist here for $\theta_i$ given an observed $\rho_s$. Therefore, the incident angle of a specular surface is ambiguous when solved from the DoLP.

**AoP constraint** Fresnel equations show that the AoP at each point corresponds to the azimuth angle of the plane of incidence (also the azimuth angle of surface normal). However, as $\phi_{aop} \in [0, \pi)$ whereas the azimuth angle is defined in the range of $\phi_n \in [0, 2\pi)$, there is a $\pi$ ambiguity in the surface azimuth, yielding for diffused dominated surface facet normals, the two following values:

$$\phi_n^d = \phi_{aop} \quad \text{or} \quad \phi_n^d = \phi_{aop} + \pi \tag{8}$$

We also note that there is a $\pi/2$ difference in the normal azimuth angle between diffuse dominant surface facets and specular ones (Goldstein, 2011). Therefore,

$$\phi_n^s = \phi_{aop} \pm \frac{\pi}{2} \tag{9}$$

where the $s$ and $d$ superscripts relate to the diffuse and specular domination of the surface facets, respectively.

**Surface normal ambiguity** Considering the DoLP and AoP related ambiguities, there are two possible surface normals for a diffuse dominant facet, while there are four possible normals for a specular dominant one. To resolve these ambiguities a photometric constraint is now introduced.

## 2.4 Surface normal disambiguation using photometric constraint

An added photometric constraint can disambiguate the polarization-derived normals. Note that $\bar{I}$ provides an additional constraint on diffuse dominant surface regions. For a diffuse dominant facet, we follow Lambert's reflection law which describes $\bar{I}$ as a function of the surface normal $\bar{\mathbf{n}}$ and the normalized point light source direction, $\bar{\mathbf{s}} = [s_x, s_y, s_z]$:

$$\bar{I}^d = a\bar{\mathbf{n}} \cdot \bar{\mathbf{s}} \tag{10}$$

where $a$ is the albedo of the diffuse facet. The estimation of $a$ can be carried out using two or more diffuse-related pixels from which the normal ambiguity can be resolved, and $a$ can be computed. To disambiguate the polarization-derived diffused-surface-related normal direction, one can select the candidate that best agrees with Eq. (10). We also assume that a specular dominant facet is ideally mirror-reflective. In such a case, the surface normal can be expressed as the halfway vector between the viewing and light source directions:

$$\bar{\mathbf{n}} = (\bar{\mathbf{v}} + \bar{\mathbf{s}})/\|\bar{\mathbf{v}} + \bar{\mathbf{s}}\| \tag{11}$$

To disambiguate the polarization-derived specular normal, of the four options, one can select the one agreeing the most with Eq. (11).

## 2.5 Normal related constraints

The disambiguation of the surface normals provides object shape information. For the reconstruction, the derived forms are also used to generate surface normal constraints, which are utilized in Sec. (2.6) to estimate both depth and normals. There are two constraints that can be written, *i*) between the surface normal and the light intensity and the DoLP of the diffused facet, and *ii*) between the AoP and the surface normal. Both are introduced into the surface depth equations, and their linear form makes the estimation direct.

**Diffuse DoLP ratio constraint** Here, the combination of the DoLP and the photometric constraints for a diffuse dominant facet (Eqs. 6 & 10, respectively) yields the linear constraint:

$$\|\mathbf{n}\| = \frac{\bar{\mathbf{v}} \cdot \mathbf{n}}{\cos\theta_i} = \frac{a\bar{\mathbf{s}} \cdot \mathbf{n}}{\bar{I}^d} \tag{12}$$

which can be rearranged in a matrix form as a function of $\mathbf{n}$ (Tozza et al., 2017):

$$\begin{bmatrix} as_x \cos\theta_i - \bar{I}^d v_x \\ as_y \cos\theta_i - \bar{I}^d v_y \\ as_z \cos\theta_i - \bar{I}^d v_z \end{bmatrix}^{\mathrm{T}} \mathbf{n} = 0 \tag{13}$$

where $\cos\theta_i$ is obtained by Eq. (6).

**AoP normal-related constraint** As the projection of the surface normal onto the image plane $(n_x, n_y)$ should be collinear with the direction of AoP, one can establish the constraint (Zhu and Smith, 2019):

$$\begin{bmatrix} \sin\phi_n & -\cos\phi_n & 0 \end{bmatrix} \mathbf{n} = 0 \tag{14}$$

where $\phi_n$ refers to the ones obtained for either diffused dominated or a specular dominated facet. This linear constraint between the AoP and the surface normal can then be incorporated into the surface reconstruction solution.

**Disambiguated normal constraint** An additional constraint for the estimated normal can be written in reference to the disambiguated form, $\mathbf{n}' = (n'_x, n'_y, n'_z)^{\mathrm{T}}$, obtained by Sec. (2.4). This constraint is of collinearity nature and can be expressed as:

$$\begin{bmatrix} 0 & -n'_z & n'_y \\ n'_z & 0 & -n'_x \\ -n'_y & n'_x & 0 \end{bmatrix} \mathbf{n} = \mathbf{0} \tag{15}$$
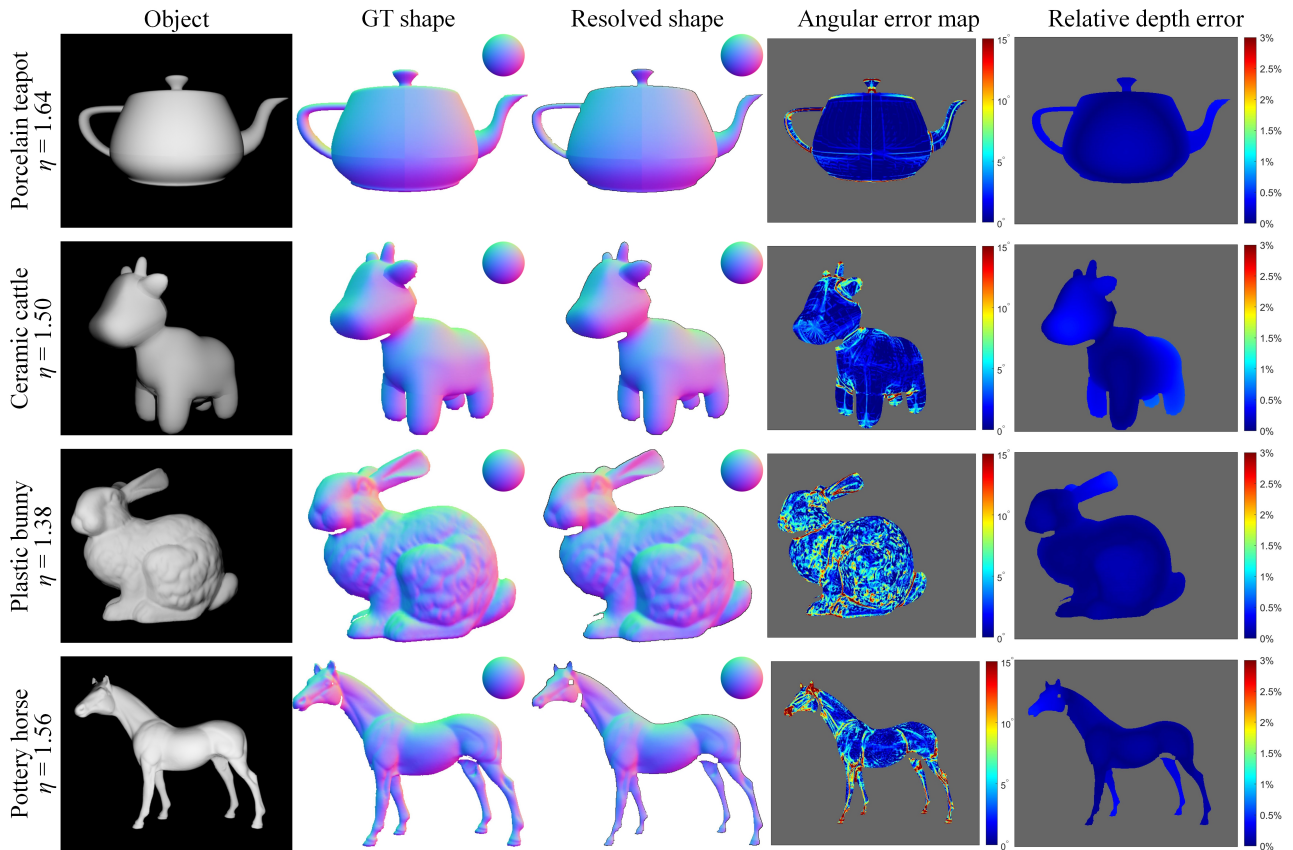
Figure 2. Qualitative normal and depth estimation results on synthetic data of different materials with uniform albedo.

## 2.6 Depth estimation via linear least squares adjustment

The surface normal can be expressed as linear equations of depth by using finite difference approximation. For a $C^2$ continuous surface, the standard way of integrating its normal $\mathbf{n}$ to obtain the surface depth $Z$ follows Poisson's equation. Poisson's equation can be transformed into a linear form by discretizing the gradient (Tozza et al., 2017):

$$\mathbf{L} \cdot \text{vec}(Z) = \mathbf{n}, \tag{16}$$

where $\mathbf{L}$ is the sparse Laplacian matrix, and $\text{vec}(\cdot)$ is the vector operator. For utilizing the normal constraint to compute the surface depth, we use Zhu and Smith (2019) solution of adapting Eq. (3) to obtain the Poisson equation in the form Eq. (16):

$$\mathbf{L} \cdot \text{vec}(Z) = \mathbf{n}, \quad \mathbf{L} = \begin{bmatrix} -f\mathbf{I} & 0 & 0 \\ 0 & -f\mathbf{I} & 0 \\ DX & DY & \mathbf{I} \end{bmatrix} \begin{bmatrix} D_x \\ D_y \\ \mathbf{I} \end{bmatrix} \tag{17}$$

where $\mathbf{n} = \left[ n_x^1, \ldots, n_x^N, n_y^1, \ldots, n_y^N, n_z^1, \ldots, n_z^N \right]^{\mathrm{T}}$, $\text{vec}(Z) = \left[ Z^1, \cdots Z^N \right]^{\mathrm{T}}$, $N$ is the number of object surface pixels, the superscripts $1 \cdots N$ indicate the pixels $(x_1, y_1), \ldots, (x_N, y_N)$, $DX = \text{diag}(x_1, \cdots, x_N)$, $DY = \text{diag}(y_1, \cdots, y_N)$, and $D_x, D_y \in \mathbb{R}^{N \times N}$ act on $\text{vec}(Z)$ to compute finite difference approximations to the derivative of $Z$ in the $x$ and $y$ directions, respectively. Combining the polarimetric, photometric, and normal constraints in Eqs. (13), (14), & (15) yield a linear set of equations of the surface depth:

$$\mathbf{A} \cdot \mathbf{L} \cdot \text{vec}(Z) = \mathbf{0}, \tag{18}$$

where $\mathbf{A}$ consists of the three matrix-form constraints. Note that by this form, depth can be estimated up to scale. Therefore, constraining one depth value into a fixed value allows for solving this system of equations. This system of linear equations can be solved by sparse linear least squares adjustment. Additional polarization images from different viewing angles can increase the number of constraints for better results, and so is the increase of point light sources.

## 3. RESULTS

Experimental results are presented on both synthetic and real-world data. The synthetic polarization images (Fig. 2) were rendered using the polarized ray tracing software Mitsuba 2, which generates high-fidelity polarization images (Cui et al., 2017). Both simulated objects are smooth and featureless, and to model their different materials, we specify their pBRDF and refractive index $\eta$. Since each object is made of a single material, we also assume a uniform albedo. The simulated measurement system consists of a polarization camera and a point light source. We also corrupt the polarization image by adding Gaussian noise with zero mean and standard deviation $\sigma$ to analyze the influence of different noise levels of the polarization image on the reconstruction result.

In Fig. (2) we show the qualitative results on a porcelain teapot, a ceramic cattle, a coated plastic bunny, and a pottery horse. All of them have different refractive indices. The SfP completely reconstructed the surface shape without distortions, and the surface normals are also smooth. From the angular error maps for the estimated normals, we see that the normals of the SfP are

| Object | $\sigma = 0$ | | $\sigma = 1\%$ | | $\sigma = 2\%$ | |
|--------|-----------|-------------|-----------|-------------|-----------|-------------|
| | Depth (%) | Normal (deg) | Depth (%) | Normal (deg) | Depth (%) | Normal (deg) |
| Teapot | 0.11 | 1.06 | 0.18 | 2.41 | 0.22 | 3.70 |
| Cattle | 0.16 | 1.67 | 0.19 | 2.67 | 0.58 | 3.75 |
| Bunny | 0.81 | 2.34 | 1.20 | 3.04 | 1.85 | 4.17 |
| Horse | 0.70 | 2.10 | 1.13 | 2.85 | 1.60 | 3.94 |

Table 1. Mean relative depth error and surface normal mean angular error on synthetic data at different noise levels.

overall accurate. The angular error mainly exists along the objects edges, e.g., the teapot handle, the cattle and bunny ears, and the horse limbs. At these positions, where the object surfaces occupy few image pixels, the normals change drastically, and there are multiple reflections between the close surfaces, all leading to some errors. The relative depth error maps show that the SfP can accurately reconstruct the surface depth. The sharp change of the normals at the edge of the object has little effect on the depth reconstruction. Table (1) lists the mean relative error in the reconstructed surface depth and mean angular error in the estimated surface normals at different noise levels. As the noise level increases, the angular and relative depth errors increase but remain low, demonstrating that this SfP method is robust to increased Gaussian noise.

For our real-world test we use a cylindrical wooden object and a porcelain mug, both featuring a smooth surface. Both objects were imaged by the LUCID TRI050S1-PC monocular polarization camera with a Nikkor 35mm lens, and a fixed-point light source. In addition, 75 RGB images of each of the objects were acquired and introduced in the Pix4D software for MVS reconstruction. Fig. (3) presented the reconstruction results using the MVS and the SfP methods. Despite the random texture on the cylindrical surface (Fig. 3a), the MVS reconstruction is relatively sparse, making it difficult to observe the curvature of the cylinder surface. The colorized normals allude to the noisy reconstruction. In contrast, the SfP completely reconstructed the cylinder part within the camera FOV and provided smooth normal estimates (c.f. the smooth color map). The mug (Fig. 3b) has a more complex structure than the cylinder, it is featureless and also features specular reflections. Here again, the MVS solution could not reconstruct the mug surface well, with a noisy point cloud and many holes. The SfP reconstruction yielded a dense and smooth point cloud of both the mug and handle. One can also clearly observe the shape variation of the mug surface from the point cloud. Both results demonstrate the value in SfP, as a feasible and easy-to-implement solution for reconstructing featureless or specular surfaces.

## 4. CONCLUSIONS

This paper proposed a generalized version of the shape-from-polarization model for the 3-D reconstruction of textureless and specular objects. The model considered the relationship between specular/diffuse reflected polarized light and surface normal, regardless of the presence of features on the surface. It demonstrated that polarization constraints can be expressed linearly in terms of depth, so depth estimation is essentially a linear least squares problem. By solving this linear least-squares adjustment problem with the polarization observations, we obtain the surface depth of the specular and featureless object. The outcome is a pixel-level dense monocular 3-D reconstruction model independent of surface texture, demonstrating stability



Wooden cylinder    MVS    SfP
(a)

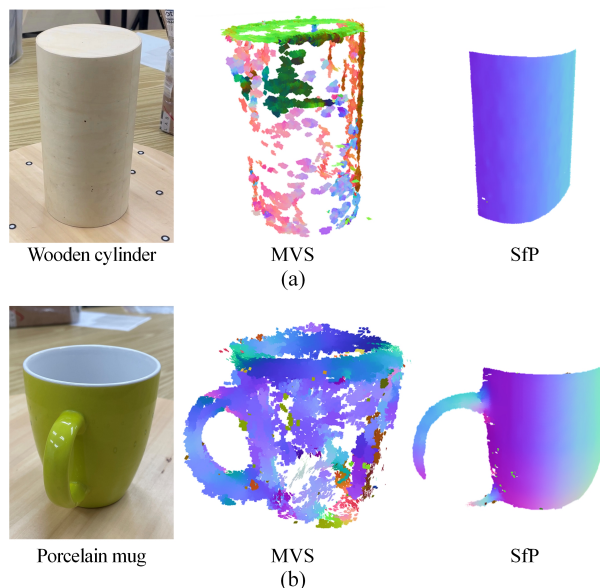Porcelain mug    MVS    SfP
(b)

Figure 3. Real-world testing on a wooden cylindrical form object and a porcelain mug.

in featureless regions and one characterized by specular reflection. Our experiments demonstrate that the proposed algorithm is general and performs well on featureless objects of different materials and shapes. The visualization of 3-D reconstruction results also showed rich details and minute structures/edges, which are usually unobservable in an ordinary RGB image. Both synthetic and real-world results demonstrated that our method is practically useful, enabling monocular, passive depth estimation.

## 5. ACKNOWLEDGMENT

### References

Alzuhiri, M., Farrag, K., Lever, E., Deng, Y., 2021. An Electronically Stabilized Multi-Color Multi-Ring Structured Light Sensor for Gas Pipelines Internal Surface Inspection. *IEEE Sensors Journal*, 21(17), 19416–19426.

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

Ba, Y., Gilbert, A., Wang, F., Yang, J., Chen, R., Wang, Y., Yan, L., Shi, B., Kadambi, A., 2020. Deep Shape from Polarization. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 12369 LNCS, 554–571.

Barron, J. T., Malik, J., 2015. Shape , Illumination , and Reflectance from Shading. 37(8), 1670–1687.

Bednarik, J., Fua, P., Salzmann, M., 2018. Learning to reconstruct texture-less deformable surfaces from a single view. *Proceedings - 2018 International Conference on 3D Vision, 3DV 2018*, 606–615.

Cheng, Z., Li, H., Asano, Y., Zheng, Y., Sato, I., 2021. Multi-view 3D Reconstruction of a Texture-less Smooth Surface of Unknown Generic Reflectance. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 16221–16230.

Cui, Z., Gu, J., Shi, B., Tan, P., Kautz, J., 2017. Polarimetric Multi-View Stereo. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1558–1567.

Cui, Z., Larsson, V., Pollefeys, M., 2019. Polarimetric relative pose estimation. *Proceedings of the IEEE International Conference on Computer Vision*, 2019-Octob, 2671–2680.

Fu, G., Zhang, Q., Zhu, L., Li, P., Xiao, C., 2021. A Multi-Task Network for Joint Specular Highlight Detection and Removal. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 7748–7757.

Fukao, Y., Kawahara, R., Nobuhara, S., Nishino, K., 2021. Polarimetric Normal Stereo. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 682–690.

Goldstein, D., 2011. *Polarized Light*. 3rd edition edn, CRC Press, Boca Raton.

Graber, G., Balzer, J., Soatto, S., Pock, T., 2015. Efficient minimal-surface regularization of perspective depth maps in variational stereo. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 07-12-June-2015, 511–520.

Hafeez, J., Jeon, H. J., Hamacher, A., Kwon, S. C., Lee, S. H., 2018. The effect of patterns on image-based modelling of texture-less objects. *Metrology and Measurement Systems*, 25(4), 755–767.

Hafeez, J., Lee, J., Kwon, S., Ha, S., Hur, G., Lee, S., 2020. Evaluating feature extraction methods with synthetic noise patterns for image-based modelling of texture-less objects. *Remote Sensing*, 12(23), 1–19.

Kadambi, A., Taamazyan, V., Shi, B., Raskar, R., 2015. Polarized 3D: High-quality depth sensing with polarization cues. *Proceedings of the IEEE International Conference on Computer Vision*, 2015 Inter, 3370–3378.

Koutsoudis, A., Ioannakis, G., Vidmar, B., Arnaoutoglou, F., Chamzas, C., 2015. Using noise function-based patterns to enhance photogrammetric 3D reconstruction performance of featureless surfaces. *Journal of Cultural Heritage*, 16(5), 664–670.

Kurth, P., Lange, V., Siegl, C., Stamminger, M., Bauer, F., 2018. Auto-calibration for dynamic multi-projection mapping on arbitrary surfaces. *IEEE Transactions on Visualization and Computer Graphics*, 24(11), 2886–2894.

Lei, C., Qi, C., Xie, J., Fan, N., Koltun, V., Chen, Q., 2021. Shape from Polarization for Complex Scenes in the Wild. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 12632–12641.

Li, B., Xu, Z., Gao, F., Cao, Y., Dong, Q., 2022. 3D Reconstruction of High Reflective Welding Surface Based on Binocular Structured Light Stereo Vision. *Machines*, 10(2).

Li, F., Tian, J., Tang, Y., Wang, Y., 2015. An Image Highlights Removal Method with Polarization Principle. *Proceedings of the 2015 3rd International Conference on Machinery, Materials and Information Technology Applications*, 35number Icmmita, 402–407.

Smith, W. A., Ramamoorthi, R., Tozza, S., 2019. Height-from-Polarisation with Unknown Lighting or Albedo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(12), 2875–2888.

Tozza, S., Smith, W. A., Zhu, D., Ramamoorthi, R., Hancock, E. R., 2017. Linear Differential Constraints for Photo-Polarimetric Height Estimation. *Proceedings of the IEEE International Conference on Computer Vision*, 2017-Octob, 2298–2306.

Wang, G., Zhang, X., Cheng, J., 2020. A Unified Shape-From-Shading Approach for 3D Surface Reconstruction Using Fast Eikonal Solvers. *International Journal of Optics*, 2020, 9–16.

Wu, C., Narasimhan, S. G., 2010. A Multi-Image Shape-from-Shading Framework for Near-Lighting Perspective Endoscopes. 211–228.

Wu, K., Tan, J., Xia, H. L., Liu, C. B., 2021. An Exposure Fusion-Based Structured Light Approach for the 3D Measurement of a Specular Surface. *IEEE Sensors Journal*, 21(5), 6314–6324.

Wu, S., Huang, H., Portenier, T., Sela, M., Cohen-or, D., Kimmel, R., Zwicker, M., 2018. Specular-to-Diffuse Translation for Multi-View Reconstruction. *Proceedings of the European Conference on Computer Vision (ECCV)*, 183–200.

Xia, W., Chen, E. C., Pautler, S. E., Peters, T. M., 2019. A Global Optimization Method for Specular Highlight Removal from a Single Image. *IEEE Access*, 7, 125976–125990.

Zhang, Z., Chang, C., Liu, X., Li, Z., Shi, Y., Gao, N., Meng, Z., 2021. Phase measuring deflectometry for obtaining 3D shape of specular surface: a review of the state-of-the-art. *Optical Engineering*, 60(02), 1–17.

Zhu, D., Smith, W. A., 2019. Depth from a polarisation + RGB stereo pair. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019-Junenumber September, 7578–7587.