

Supplementary Material for DiLiGenRT: A Photometric Stereo Dataset with Quantified Roughness and Translucency

Anonymous CVPR submission

Paper ID 2461

This supplementary material provides additional technical details about the light calibration and data capture process, followed by analyzing the influence of light distribution on the photometric stereo when dueling with diverse materials. After that, we provide the evaluation of additional photometric stereo methods, including UniPS [6], GPS-Net [20], PX-Net [11], SPLINE-Net [21], and DeepPS2 [18]. Finally, we provide the complete benchmark results containing estimated surface normals and the corresponding angular error maps on our DiLiGenRT dataset for each method. We also release a preview subset of our DiLiGenRT at this [anonymous link](#).

A. Light calibration and capture settings

Light calibration As shown in Fig. S1 (left), we place 6 mirror spheres near the target object. Following the practice of existing photometric stereo datasets (*e.g.*, DiLiGenT [17], DiLiGenT10² [14]), the incident light directions can be calibrated via the specular spots on the mirror balls. Specifically, taking one mirror sphere as an example, we first use circle Hough transform [8] to detect the projected circles on the image plane, extracting radius r in pixel unit and circle center location $\mathbf{c} = (u_0, v_0)$, as shown in the middle column of Fig. S1. Then we manually label specular highlight position $\mathbf{p} = (u_p, v_p)$. Assuming that the world center is aligned with the sphere, the 3D coordinates in pixel unit of the specular spot can be represented as

$$\mathbf{s} = \begin{pmatrix} u_p - u_0 \\ v_p - v_0 \\ \sqrt{r^2 - (u_p - u_0)^2 - (v_p - v_0)^2} \end{pmatrix} \quad (1)$$

As shown in Fig. S1 (right), the surface normal direction at point \mathbf{s} on the sphere can be calculated as

$$\mathbf{n} = \frac{\mathbf{s} - \mathbf{o}}{\|\mathbf{s} - \mathbf{o}\|} = \begin{pmatrix} (u_p - u_0)/r \\ (v_p - v_0)/r \\ \sqrt{r^2 - (u_p - u_0)^2 - (v_p - v_0)^2}/r \end{pmatrix} \quad (2)$$

Based on the law of reflection, light direction \mathbf{l} , camera view direction \mathbf{v} , and surface normal vector \mathbf{n} at specular spot \mathbf{s} follows

$$\mathbf{l} = 2(\mathbf{n}^\top \mathbf{v})\mathbf{n} - \mathbf{v}. \quad (3)$$

As we fix the view direction $\mathbf{v} = [0, 0, 1]^\top$ and know the surface normal \mathbf{n} from the specular spot at \mathbf{s} , light direction \mathbf{l} can be calibrated following Eq. (3) from a single sphere ball. In a similar manner, we can obtain 6 light directions $\{\mathbf{l}_1, \mathbf{l}_2, \dots, \mathbf{l}_6\}$ from the image observations of the 6 mirror spheres. The optimized light direction $\hat{\mathbf{l}}$ can be estimated by solving the following constrained linear system:

$$\begin{bmatrix} \mathbf{l}_1^\top \\ \mathbf{l}_2^\top \\ \vdots \\ \mathbf{l}_6^\top \end{bmatrix} \hat{\mathbf{l}} = \mathbf{1}, \quad s.t. \|\hat{\mathbf{l}}\|_2^2 = 1. \quad (4)$$

By applying existing solver [15] for Eq. (4), the light direction at each image capture can be calibrated.

On the other hand, similar to DiLiGenT10² [14], our capture setup changes the illumination by shifting a single point light source mounted on a robot arm across a hemisphere, with the target object placed at the hemisphere center. This arrangement keeps the distance between the point light and the object approximately constant. As a result, we make the assumption that the light intensities received from various light positions on the object remain uniform.

Capture settings To capture our DiLiGenRT dataset, we adopt DaHeng Image MER-503-36U3C¹ camera equipped with a 50 mm lens, producing raw images at a resolution of 2448 × 2048, as shown in Fig. S1. We crop the images to 960 × 960 resolution to focus on the central valid areas. For each target object, we first position the point light source

¹DaHeng camera: <https://en.daheng-imaging.com/show-107-2044-1.html>

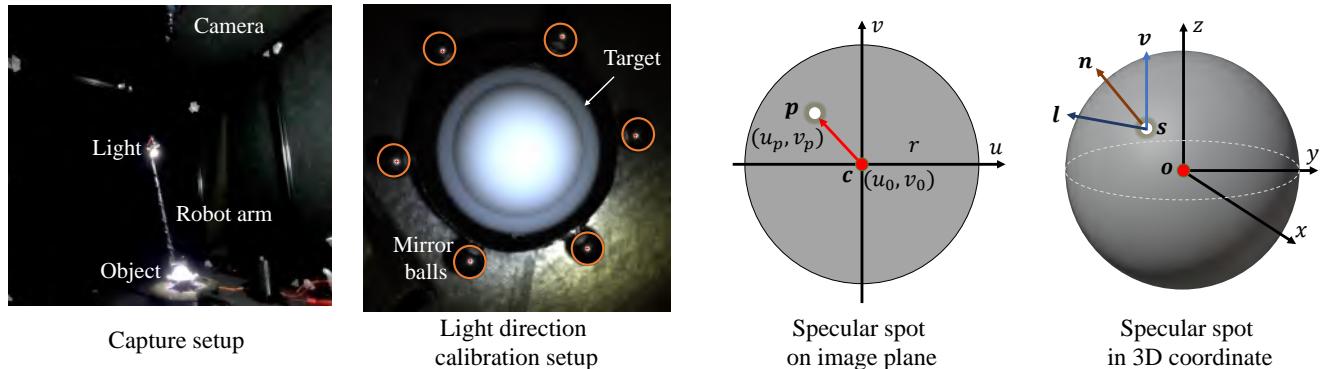


Figure S1. Our capture and calibration setups include 6 mirror balls for light calibration.

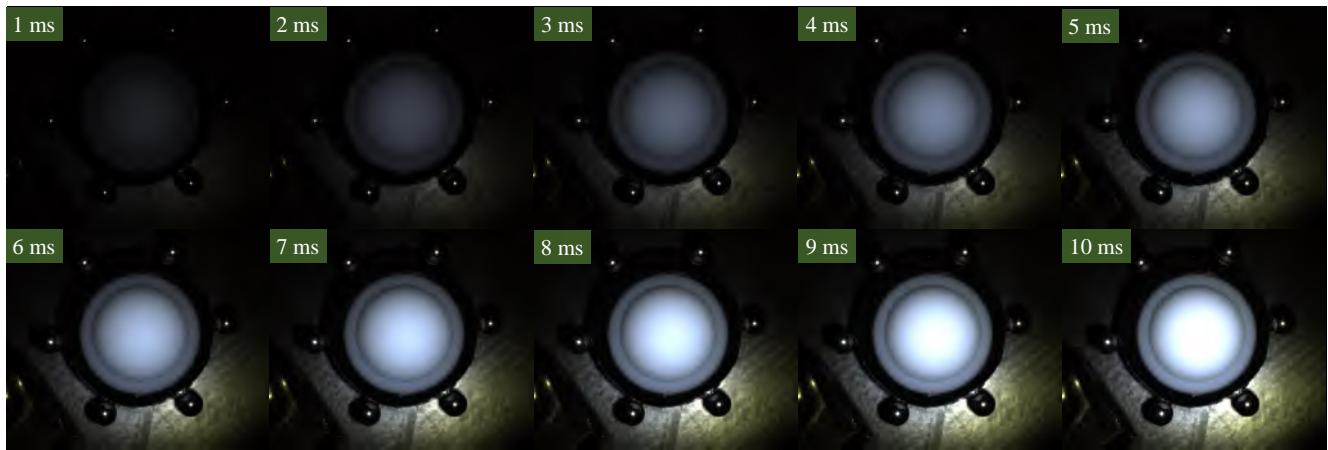
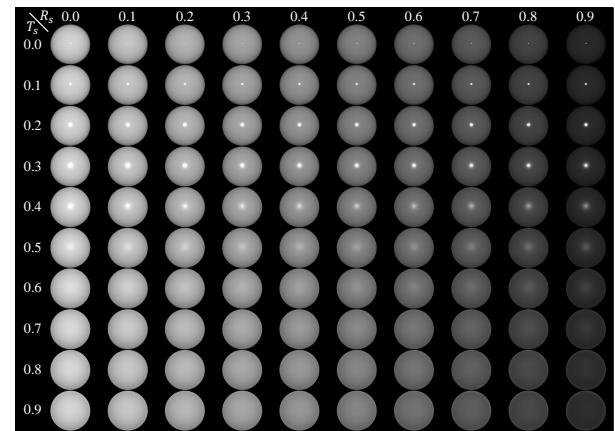


Figure S2. The target scene is captured under multiple exposure times from 1ms to 10ms to composite a HDR image.



Figure S3. Subset of the 3000 surface normals of our synthetic dataset PS-SSS.

Figure S4. Synthetic dataset PS-Sphere indexed by the roughness R_s and transmittance T_s defined in the Principled BSDF model.

via the robot arm to illuminate the scene, followed by capturing images at 10 distinct exposure times that range from 1ms to 10ms, as illustrated in Fig. S2. Subsequently, these low dynamic range images are amalgamated to compose a

059
060
061
062

LS [19] (22.42/18.96)				TH28 [17] (16.26/10.05)				TH46 [17] (15.65/5.48)				ST14 [16] (17.69/13.09)				PF14 [12] (25.41/11.05)				CNN-PS [5] (16.12/5.89)					
0.0	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	4.3	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	1.1	
0.1	4.5	4.5	4.5	4.5	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	1.1	1.3	2.5	2.6	2.7	2.7	2.8	2.9	3.0	3.0	3.0	3.0	3.0
0.2	4.5	4.5	4.5	4.5	4.4	4.4	4.4	4.4	4.4	4.4	4.4	4.4	1.1	2.9	3.2	3.3	3.4	3.5	3.6	3.7	3.9	3.9	3.9	3.9	3.9
0.3	4.5	4.5	4.5	4.5	4.5	4.5	4.6	4.6	4.8	5.3	5.3	5.3	1.3	3.6	3.8	3.9	4.0	4.1	4.3	4.6	5.1	6.4	6.4	6.4	
0.4	4.6	4.6	4.6	4.7	4.7	4.7	4.8	4.9	5.1	5.6	5.9	6.9	1.0	1.0	4.0	4.2	4.4	4.6	4.8	5.2	5.8	7.0	10.4	10.4	
0.5	4.7	4.8	4.8	4.9	4.9	5.1	5.3	5.7	6.5	8.5	1.2	1.2	1.2	1.2	1.3	1.3	1.4	1.6	1.8	3.7	4.5	5.7	6.5	7.0	
0.6	5.1	5.1	5.1	5.2	5.4	5.6	5.8	6.5	7.6	10.2	1.8	4.5	4.8	5.2	5.7	6.5	7.6	9.4	12.8	12.8	12.8	12.8	12.8	12.8	
0.7	5.5	5.6	5.8	5.9	6.1	6.5	7.0	7.8	9.4	12.9	1.8	4.5	4.9	5.3	6.0	7.0	8.4	10.9	15.1	23.9	1.5	1.5	1.5	1.5	
0.8	6.1	6.2	6.5	6.8	7.2	7.7	8.5	9.8	12.1	16.3	2.2	4.8	5.2	5.7	6.4	7.5	9.3	12.0	17.0	23.0	1.5	1.5	1.5	1.5	
0.9	6.5	6.8	7.1	7.5	8.1	8.6	9.9	11.7	15.0	19.7	2.7	5.4	5.8	6.5	7.4	8.7	10.8	14.0	19.4	16.3	1.7	1.7	1.8	1.8	
T_S	PS-FCN [2] (16.38/12.62)				Att-PSN [10] (14.52/9.22)				SPLINE-Net [21] (13.66/7.63)				PX-Net [11] (13.97/14.48)				SDPS-Net [3] (14.88/10.37)				SDPS-UnIPs [7] (3.61/3.69)				
	10.2	10.2	10.2	10.2	10.2	10.2	10.2	10.2	10.2	10.2	10.2	10.2	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7	
	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.3	10.2	10.2	10.2	1.8	1.8	1.8	1.8	1.8	1.8	1.7	1.7	1.6	1.6	1.6	1.6	
	10.4	10.4	10.4	10.4	10.3	10.3	10.3	10.3	10.2	10.1	9.6	9.6	1.8	1.7	1.7	1.7	1.6	1.5	1.5	1.4	1.4	1.4	1.4	1.4	
	10.4	10.4	10.4	10.4	10.3	10.3	10.3	10.3	10.2	10.1	9.6	9.6	1.8	1.7	1.7	1.7	1.6	1.5	1.5	1.4	1.4	1.4	1.4	1.4	
	10.4	10.4	10.4	10.3	10.3	10.3	10.3	10.2	10.1	9.6	9.6	9.6	1.7	1.7	1.6	1.6	1.6	1.5	1.5	1.4	1.4	1.4	1.4	1.4	
	10.4	10.4	10.3	10.2	10.2	10.2	10.2	10.2	10.2	10.2	10.2	10.2	1.7	1.7	1.6	1.6	1.6	1.6	1.5	1.5	1.4	1.4	1.4	1.4	
	10.5	10.5	10.3	9.9	9.6	9.3	9.0	8.6	7.9	7.5	7.0	7.0	1.8	1.7	1.7	1.7	1.7	1.7	2.2	3.5	3.5	3.5	3.5	3.5	
	10.6	10.8	10.7	10.3	9.9	9.7	9.5	9.3	9.1	8.5	7.4	7.4	2.2	2.1	1.9	1.8	1.6	1.7	2.4	4.4	9.8	4.4	4.0	5.5	
	10.7	11.0	10.8	10.3	10.1	9.9	9.8	9.6	9.2	8.6	7.8	7.8	2.7	2.5	2.3	2.1	2.0	2.3	3.8	7.5	16.0	6.6	5.6	5.0	
	10.8	11.2	10.9	10.5	10.3	10.2	10.0	9.7	9.3	8.8	7.7	7.7	3.3	3.2	3.1	2.9	2.8	2.7	3.2	4.5	7.4	13.0	4.2	4.8	
	10.9	11.4	11.0	10.6	10.3	10.1	9.8	9.4	8.8	7.8	7.8	7.8	3.8	3.8	3.6	3.6	3.6	3.7	4.7	7.1	10.1	10.1	10.1	10.1	
R_p	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	
	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	
	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.9	0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	
	0.4	0.5	0.6	0.7	0.8	0.9	0.9	0.9	0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	
	0.5	0.6	0.7	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	
	0.6	0.7	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	
	0.7	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	
	0.8	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	
	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.0	0.1	0.2	0.3	

Figure S5. Roughness-translucency MAE matrices for 16 photometric stereo methods, where the ticks of row and column are σ_t and Sa . The mean and median of the MAE matrix are presented near the method name, showing method's performance profile under different level of reflectance properties. ‘Att-PSN’ and ‘SDPS’ are the abbreviations of NormAttention-PSN [10] and SDPS-Net [3].

high dynamic range (HDR) image [13]. In this way, the images captured under various illuminations within our DiLi-GenRT are in HDR format, avoiding the impact of image saturation and low-albedo pixels on photometric stereo.

B. Diverse shapes in PS-SSS

As shown in Fig. S3, we select 30 objects from Sketchfab². For each object, we randomly rotate it for 100 times, leading to 3,000 diverse surface normal maps. Given the rotated shapes and diverse materials controlled by Principled BSDF model [1], we render 3,000 sets of images to create PS-SSS.

C. DiLiGenRT vs synthetic sphere dataset

Compared to the labor-intensive manufacturing process of DiLiGenRT, an alternative way is rendering a synthetic sphere dataset (denoted as PS-Sphere) by adjusting the roughness and transmission metrics Rs and Ts defined in the Disney Principle BSDF model [1], ranging from 0 to 1. We provide such a synthetic dataset and conduct a similar benchmark evaluation like DiLiGenRT, as shown in Figs. S4 and S5, respectively.

To the best of our knowledge, there is no mapping between the synthetic roughness R_s and the real-world *measurable* roughness Sa . Therefore, the evaluation results shown in Fig. S5 cannot be used to select best-fit photometric stereo methods as we have no device to access R_s of a real-world object. This also applies to the case of T_s . On the other hand, we observed that the mean angular errors (MAE) of photometric stereo methods evaluated on PS-

Sphere are much smaller than those on DiLiGenRT shown in Fig. 5 of the main paper, despite that the observed images from Fig. S4 and Fig. 1 of the main paper are similar. Therefore, there could be a domain gap between real-world reflectance and the parametric reflectance model, which highlights the necessity of DiLiGenRT for accurately assessing photometric stereo performance in real-world scenarios.

D. Analysis on the light distribution

As shown in Fig. 7 of the main paper, we provide the evaluation results of photometric stereo methods under 10 and 100 lights. In Fig. S6, we further show the evaluation results under 20 and 50 uniformly distributed lights and present the best-performing method under different reflectance settings.

Increasing the number of input lights generally reduces the MAEs, as supported by the summarized mean and median MAE values in Fig. S6. However, we find 50 uniformly distributed lights is optimal on the DiLiGenRT dataset. Adding lights to 100 shows only a marginal improvement for opaque and semi-translucent surfaces across various roughness levels, but could be even harmful for surfaces with higher translucency levels. For instance, CNN-PS [5] achieves much smaller MAE on surfaces whose translucency measurement (σ_t) is 0.13, if reducing the number of input lights from 100 to 50. This effect could be attributed not only to the decrease in the amount of light but also to alterations in the distribution of light directions.

To demonstrate this, we conducted an experiment with CNN-PS [5] on two target spheres with differing degrees of roughness and translucency, as shown in Fig. S7. The num-

²<https://sketchfab.com>. Retrieved Nov. 22, 2023.

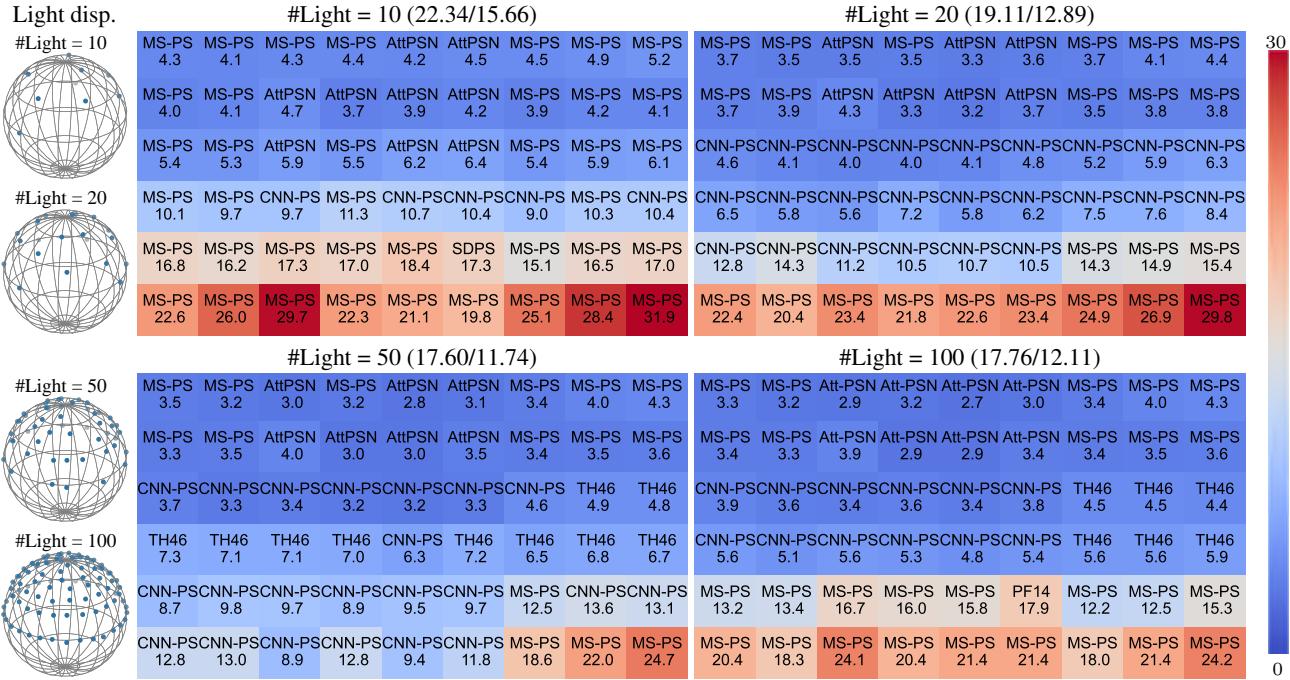


Figure S6. Benchmark evaluation on DiLiGenRT under different number of lights (#10 to #100) distributed uniformly, summarized by mean/median MAE values. Each cell records the best performing method for the material in that cell, along with the associated lowest MAE value.

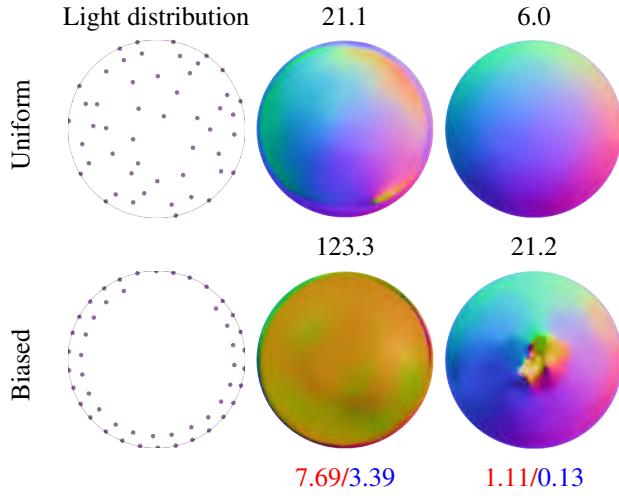


Figure S7. The influence of light distribution on surface normal estimation, tested on two spheres whose roughness and translucency measured by Sa/σ_t are shown at the bottom. The normal estimation error measured by MAE is displayed at the top of the surface normal estimates.

ber of input lights was fixed at 50, but their distribution was manipulated to be either uniform or biased, as illustrated on the left side of Fig. S7. We set the hyper-parameter K in CNN-PS [5] as 1 to avoid the influence of averaging rotated

surface normals.

As shown in Fig. S7, when more light directions are concentrated near the equator, the estimated surface normals from CNN-PS [5] exhibit greater MAEs compared to those achieved under uniformly distributed light directions. Furthermore, the sensitivity of surface normal estimation w.r.t. the amount and distribution of the incident lights increases when surfaces exhibit a higher level of translucency and a lower level of roughness. Therefore, the amount and distribution of lights serve as an important role in improving the accuracy of photometric stereo methods. It is desired to conduct adaptive illumination planning corresponding to varying reflectance.

E. Performance profiles for additional photometric stereo methods

Besides the 12 photometric stereo methods evaluated in the main paper, this supplementary material offers evaluations on 5 additional cutting-edge photometric stereo methods: PX-Net [11], SPLINE-Net [21], UniPS [6], DeepPS2 [18], and GPS-Net [20], along with their performance results on DiLiGenRT. GPS-Net [20] published on NeurIPS 2020 combines the merits of all-pixel-based and per-pixel-based photometric stereo method via a graph neural network. SPLINE-Net [21] and PX-Net [11], presented at ICCV 2019 and 2021 respectively, are enhancements to the per-

120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136

137
138
139
140
141
142
143
144
145
146
147
148

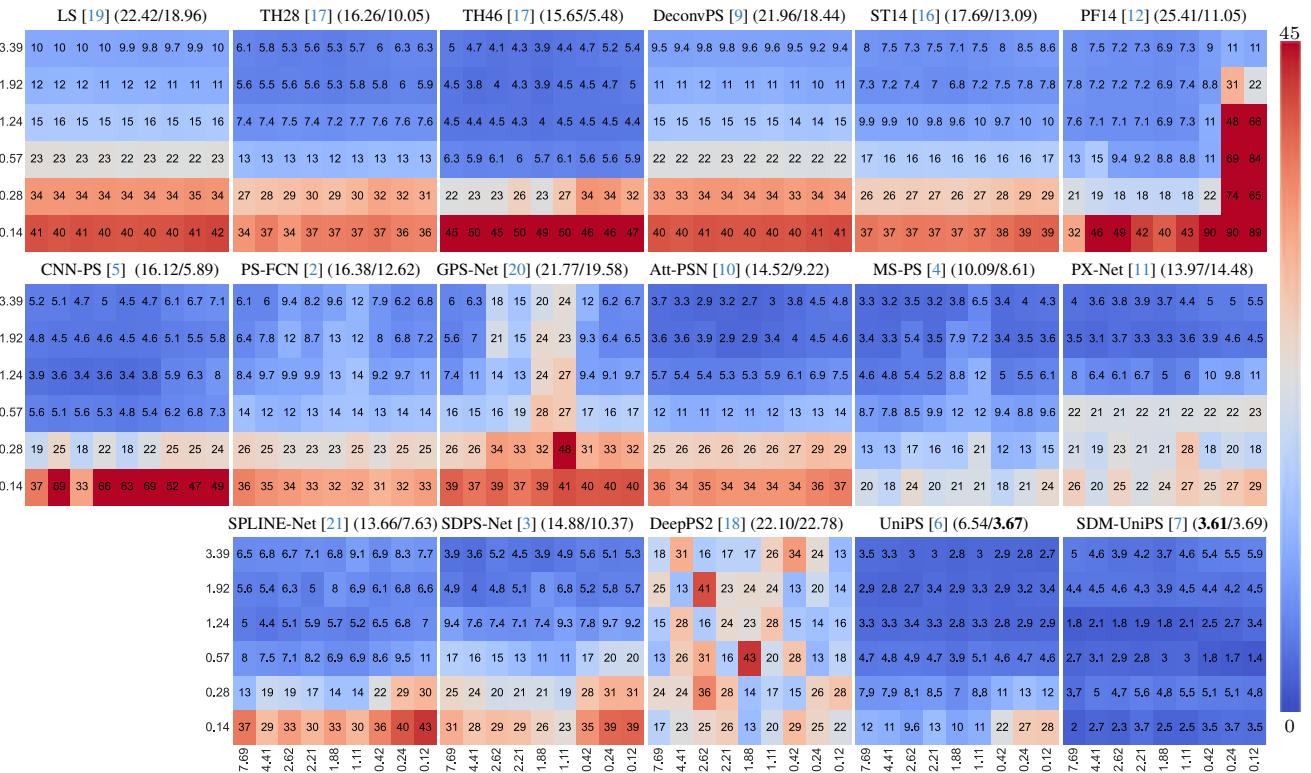


Figure S8. Roughness-translucency MAE matrices for 16 photometric stereo methods, where the ticks of row and column are σ_t and Sa . The mean and median of the MAE matrix are presented near the method name, showing method's performance profile under different level of reflectance properties. 'Att-PSN' and 'SDPS' are the abbreviations of NormAttention-PSN [10] and SDPS-Net [3].

pixel based method CNN-PS [5], focusing on sparse inputs and global illuminations. DeepPS2 [18] published at ECCV 2022 further reduces the sparse light input to only 2 distinct directional lights. UniPS [6] introduced at CVPR 2022 are built under uncalibrated universal illumination. As illustrated in Fig. S8, we display the performance profiles of all 17 photometric stereo methods.

F. Complete benchmark results

From Fig. S9 to Fig. S25, we present the complete benchmark evaluations for 17 photometric stereo methods using our DiLiGenRT dataset, including the 12 methods outlined in the main paper, as well as 4 additional methods detailed in the supplementary material. For each method, we provide a 6×9 matrix format of their estimated surface normal map alongside their corresponding angular error distribution map. For better visualization, the maximal MAE is truncated at 10° for UniPS [6] and SDM-UniPS [7] and 45° for other methods. The x and y axes in the matrix denote the translucency and roughness measurements, measured by σ_t and Sa respectively.

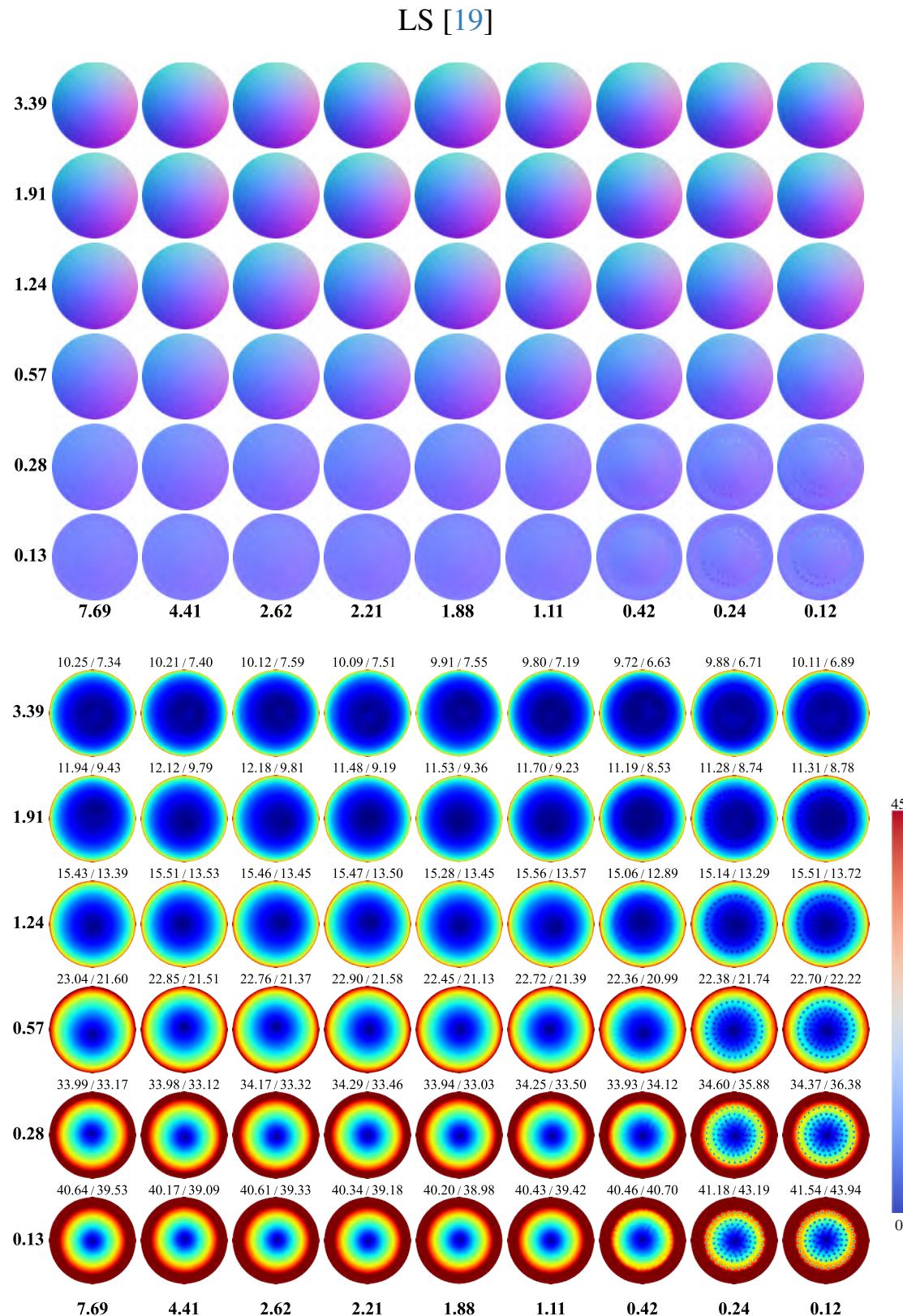


Figure S9. Estimated normal maps (top) and the corresponding angular error maps (bottom) of LS [19]. The mean and median errors for each material are displayed at the top of each error map.

TH28 [17]

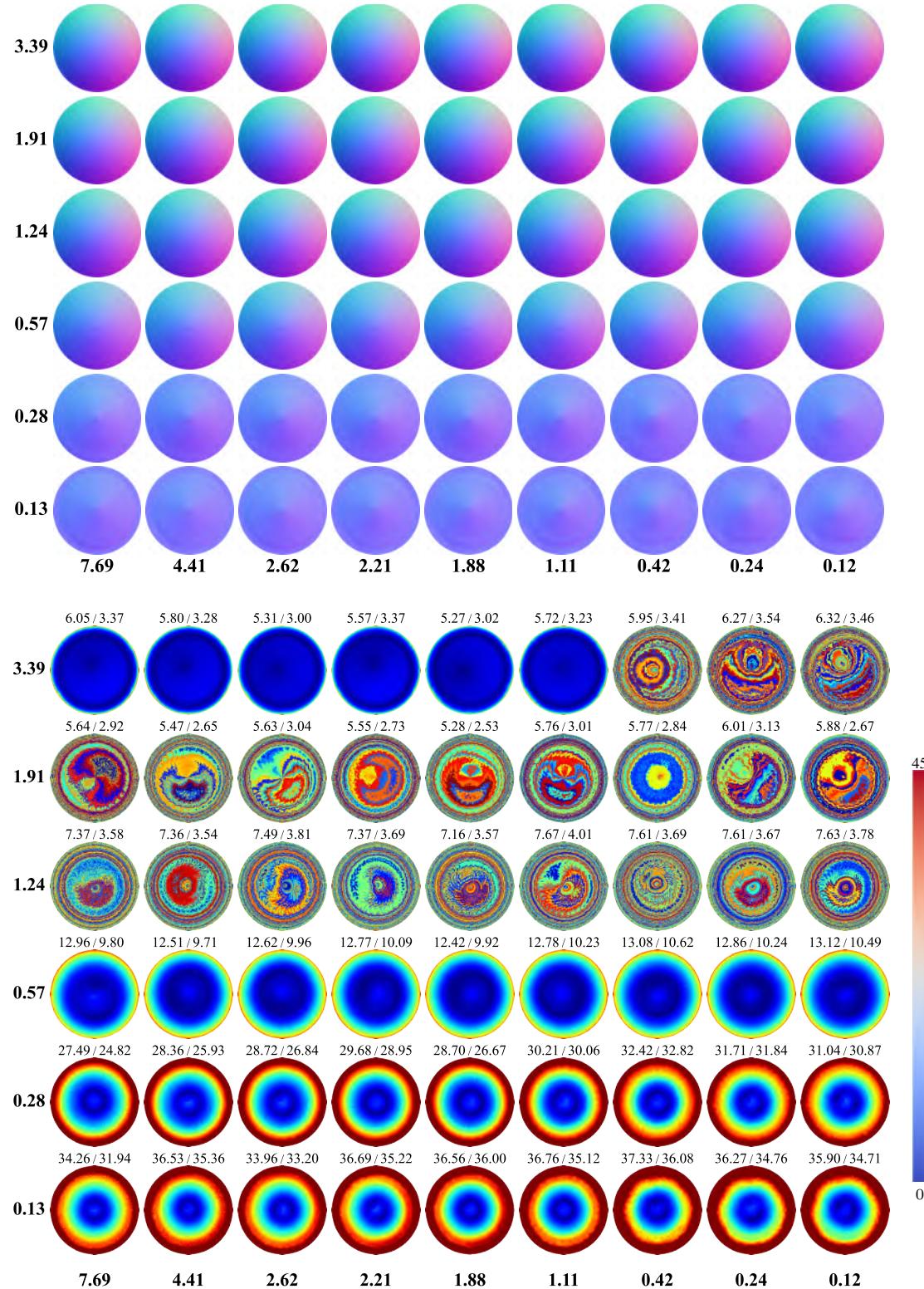


Figure S10. Estimated normal maps (top) and the corresponding angular error maps (bottom) of TH28 [17]. The mean and median errors for each material are displayed at the top of each error map.

TH46 [17]

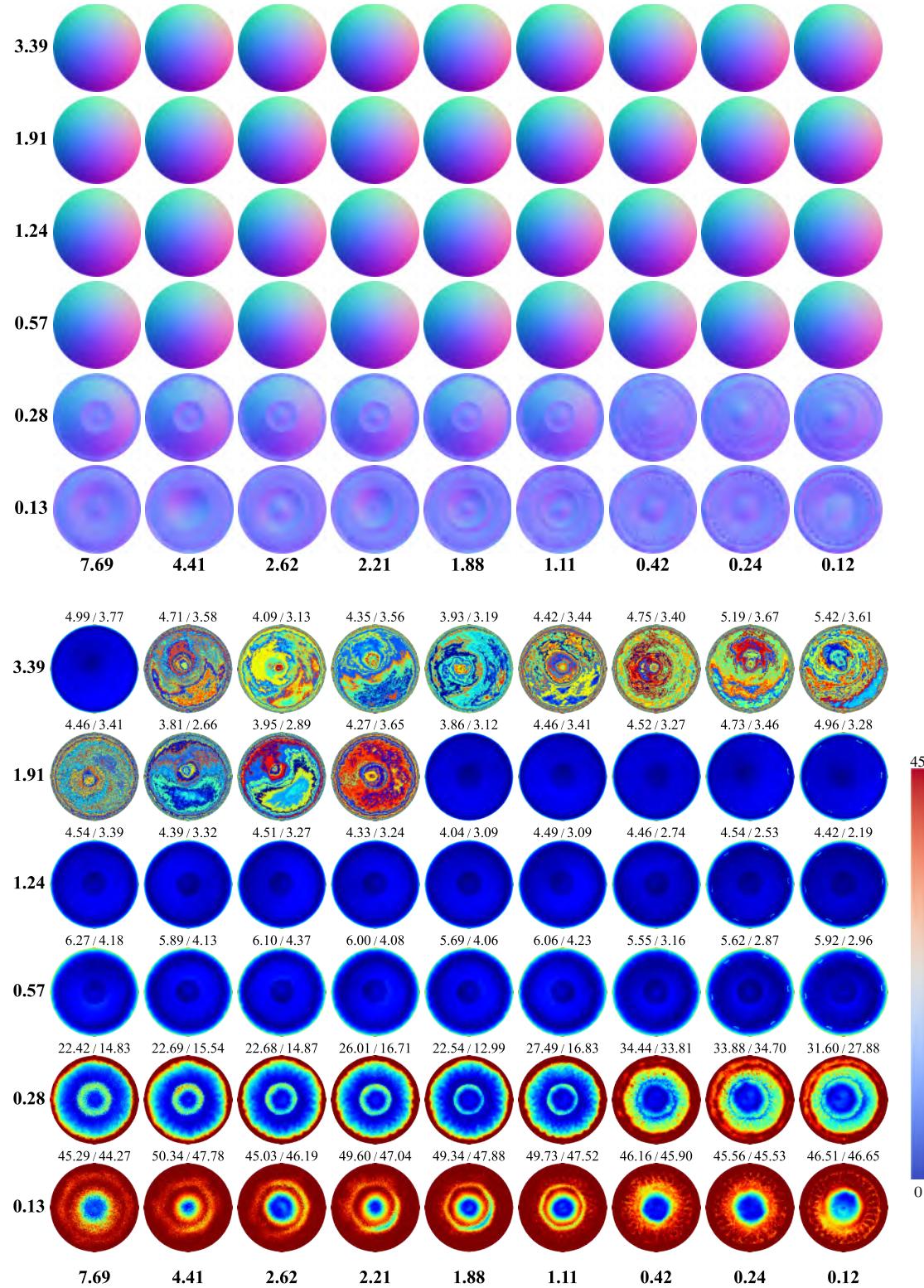


Figure S11. Estimated normal maps (top) and the corresponding angular error maps (bottom) of TH46 [17]. The mean and median errors for each material are displayed at the top of each error map.

DeconvPS [9]

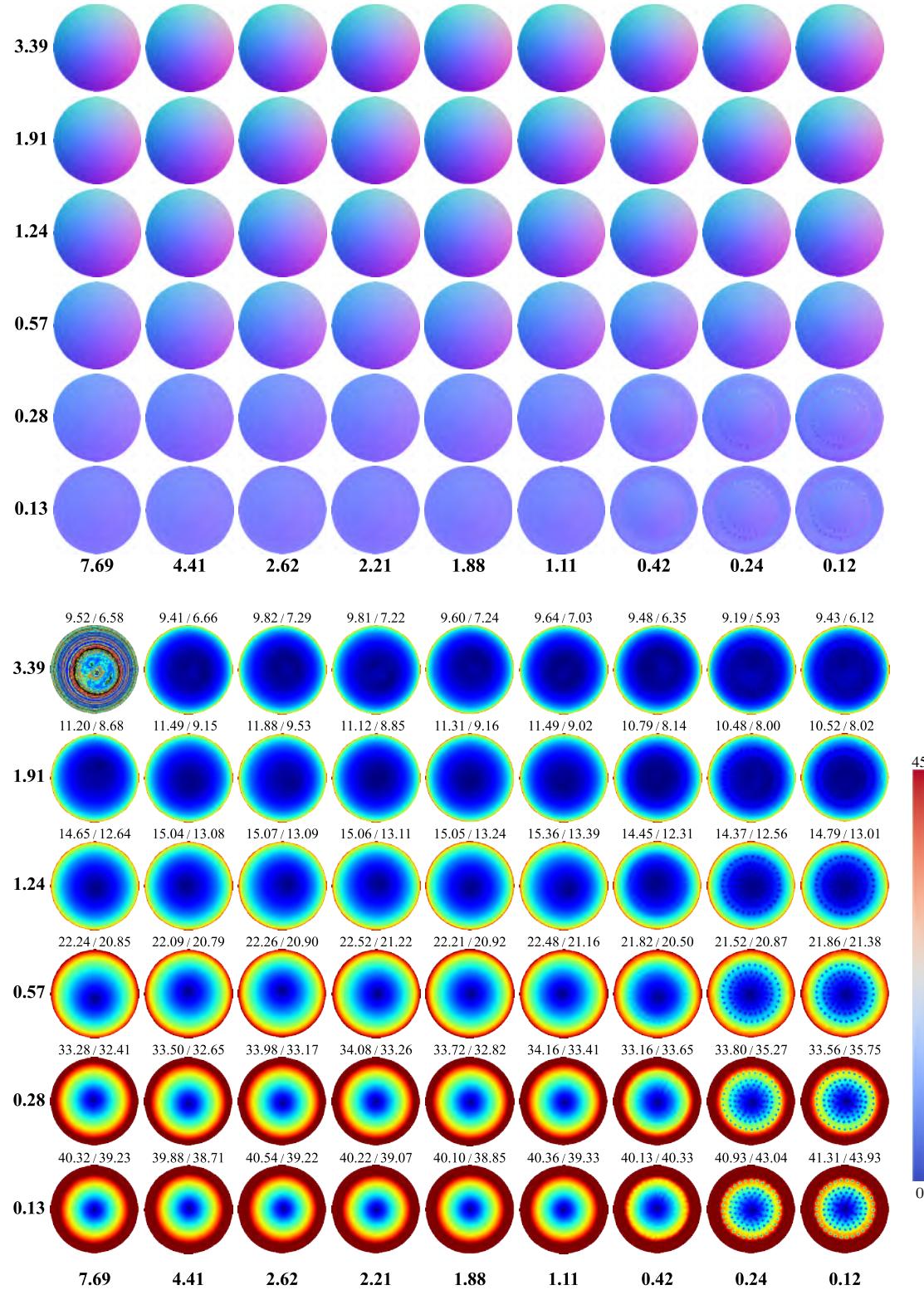


Figure S12. Estimated normal maps (top) and the corresponding angular error maps (bottom) of DeconvPS [9]. The mean and median errors for each material are displayed at the top of each error map.

ST14 [16]

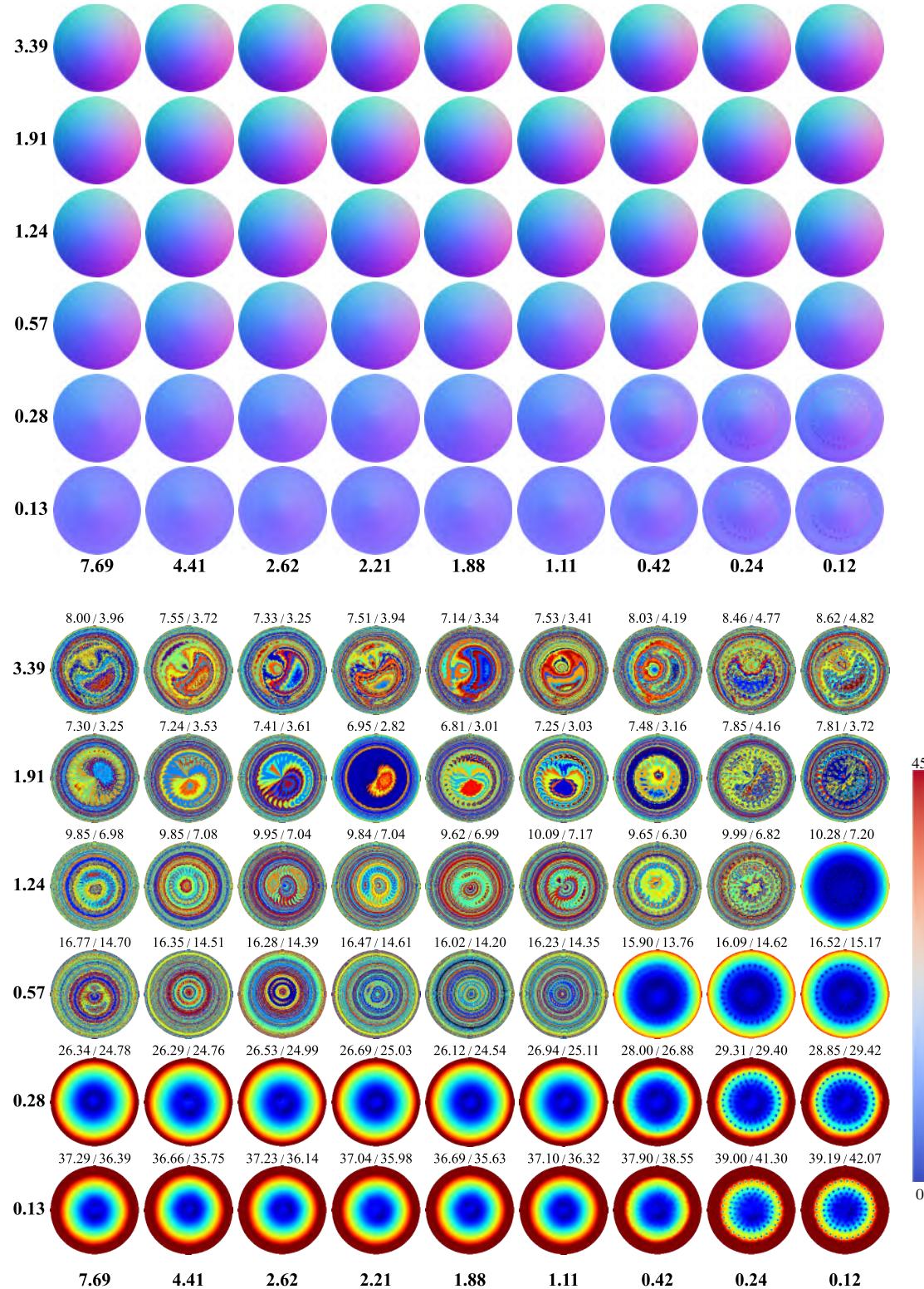


Figure S13. Estimated normal maps (top) and the corresponding angular error maps (bottom) of ST14 [16]. The mean and median errors for each material are displayed at the top of each error map.

CNN-PS [5]

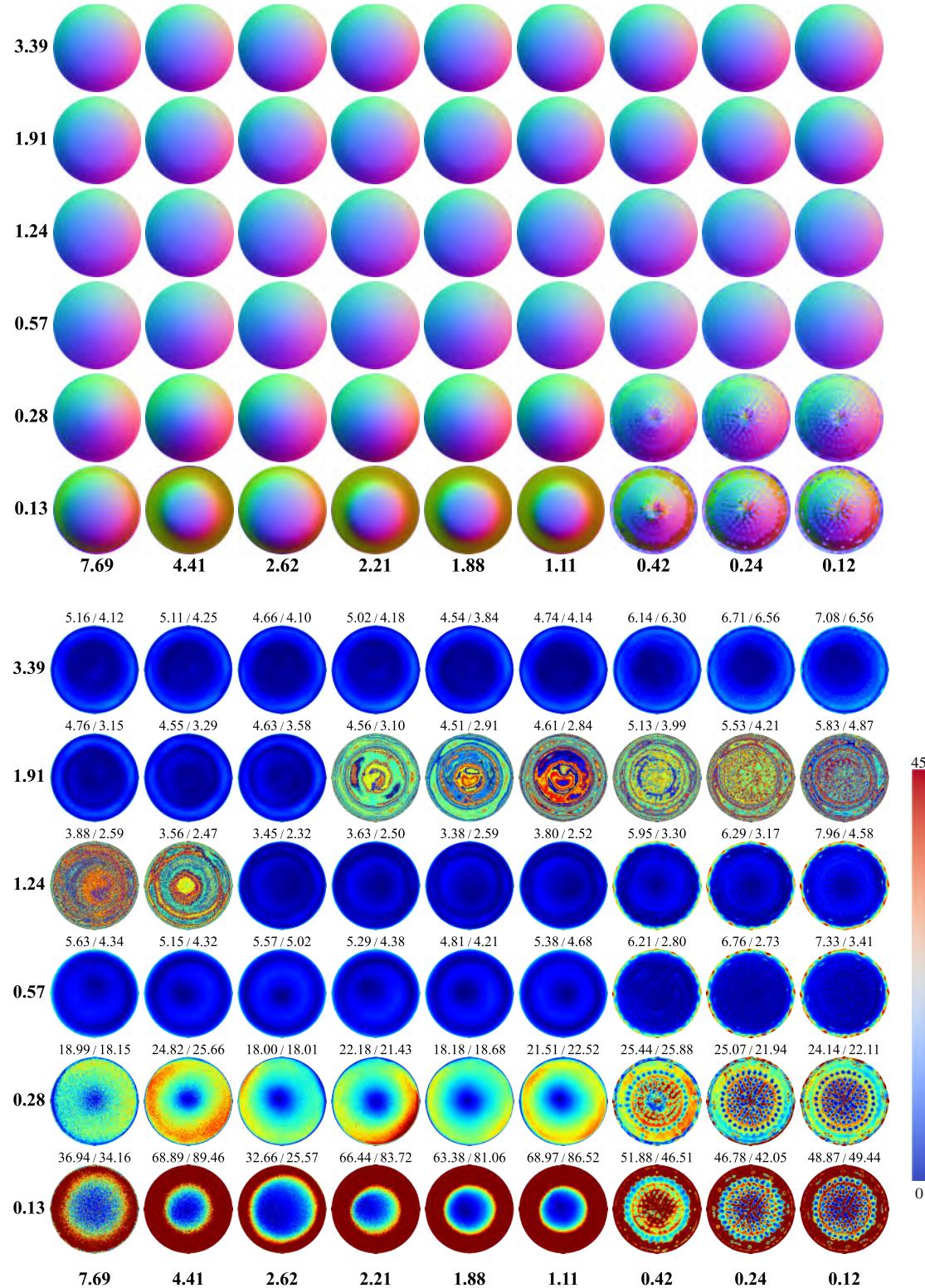


Figure S14. Estimated normal maps (top) and the corresponding angular error maps (bottom) of CNN-PS [5]. The mean and median errors for each material are displayed at the top of each error map.

PS-FCN [2]

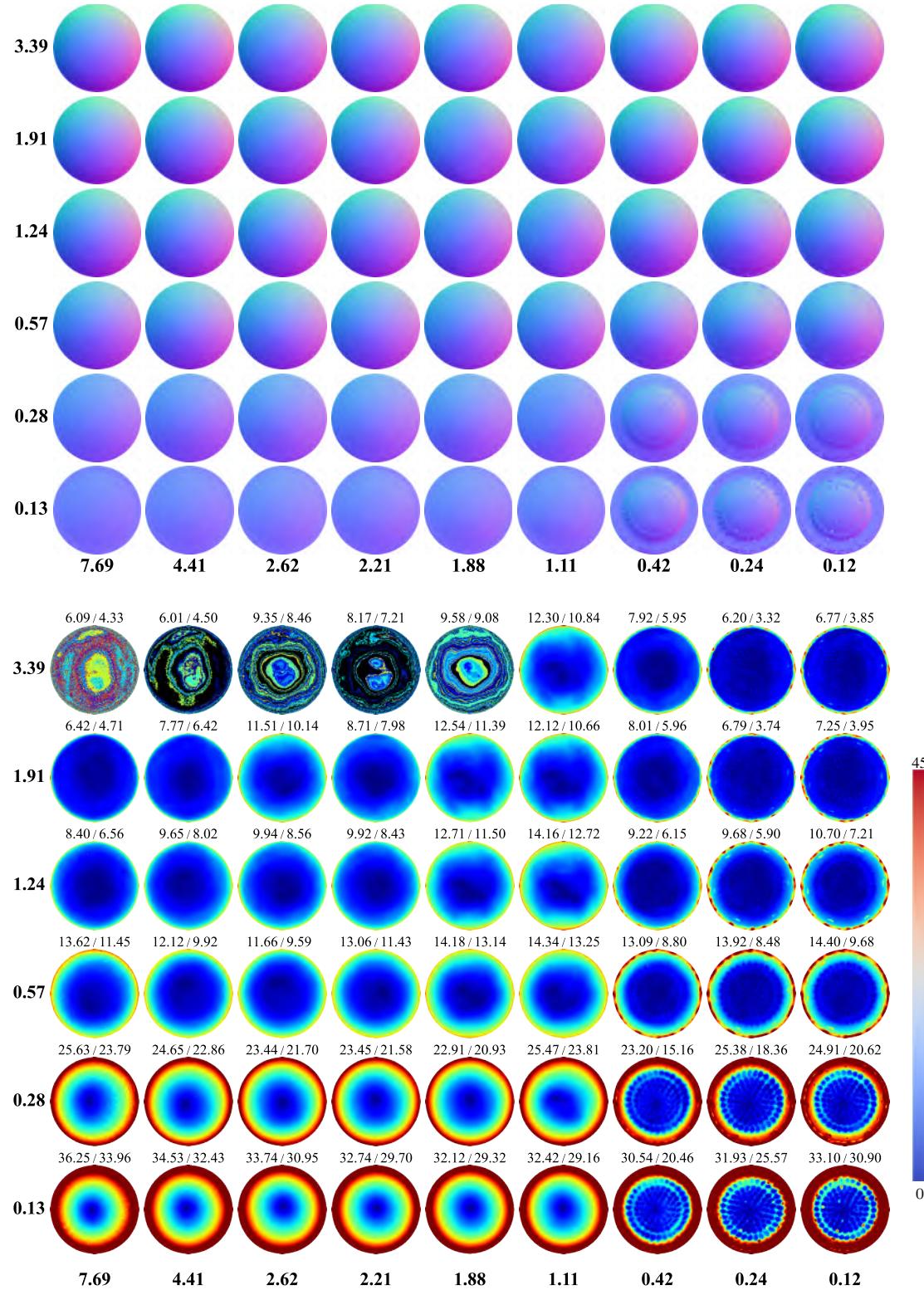


Figure S15. Estimated normal maps (top) and the corresponding angular error maps (bottom) of PS-FCN [2]. The mean and median errors for each material are displayed at the top of each error map.

GPS-Net [20]

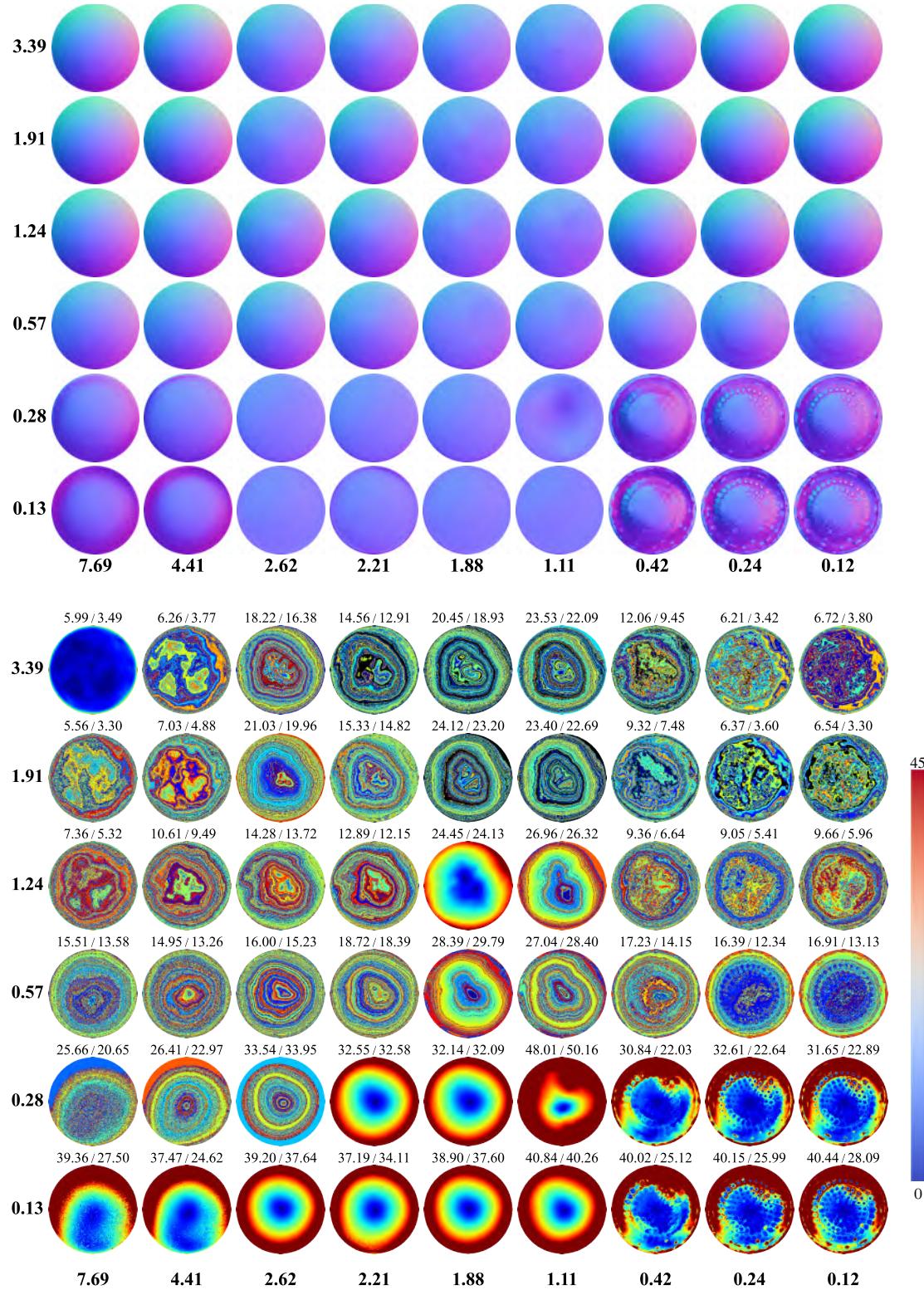


Figure S16. Estimated normal maps (top) and the corresponding angular error maps (bottom) of GPS-Net [20]. The mean and median errors for each material are displayed at the top of each error map.

NormAttention-PSN [10]

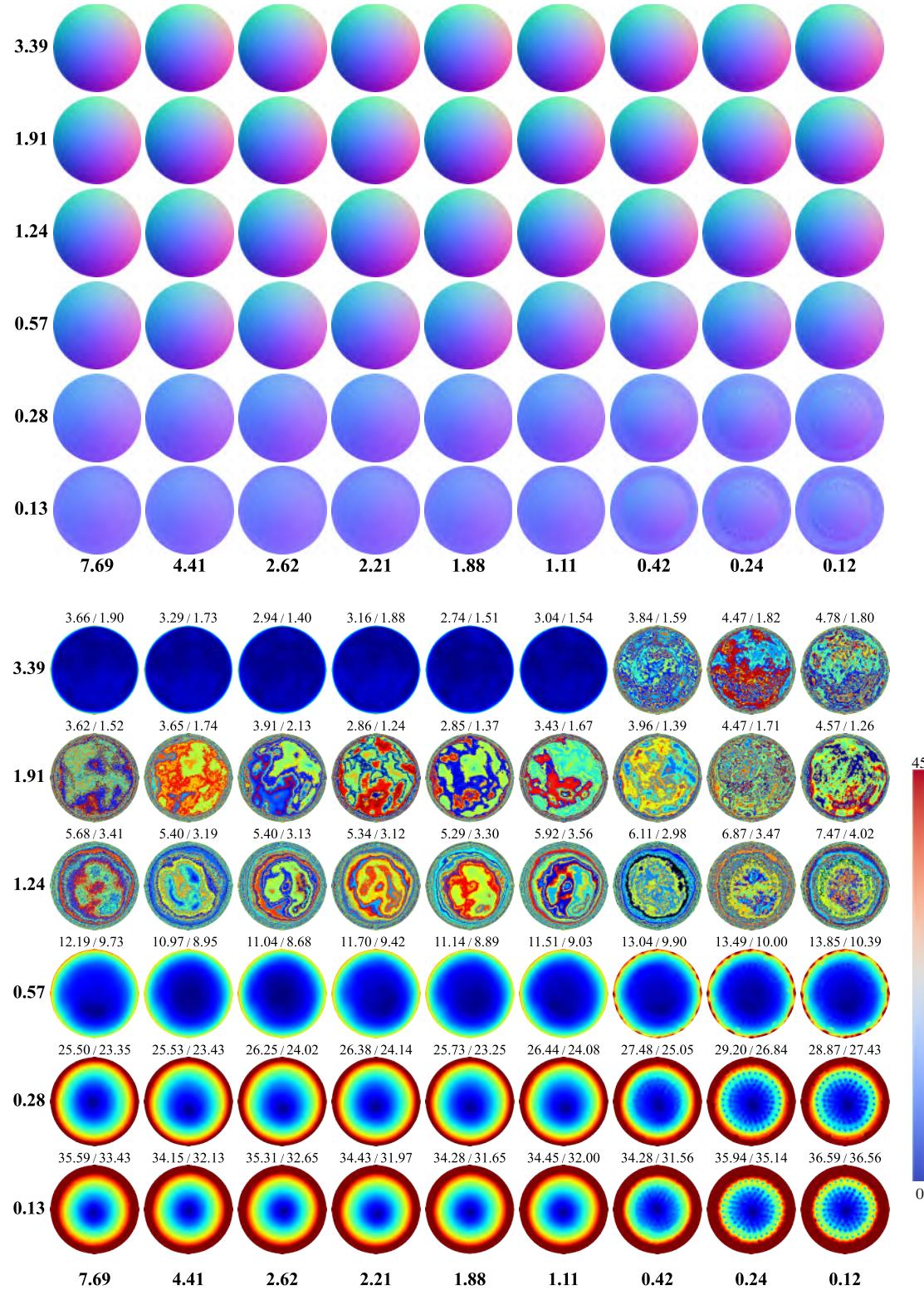


Figure S17. Estimated normal maps (top) and the corresponding angular error maps (bottom) of NormAttention-PSN [10]. The mean and median errors for each material are displayed at the top of each error map.

SPLINE-Net [21]

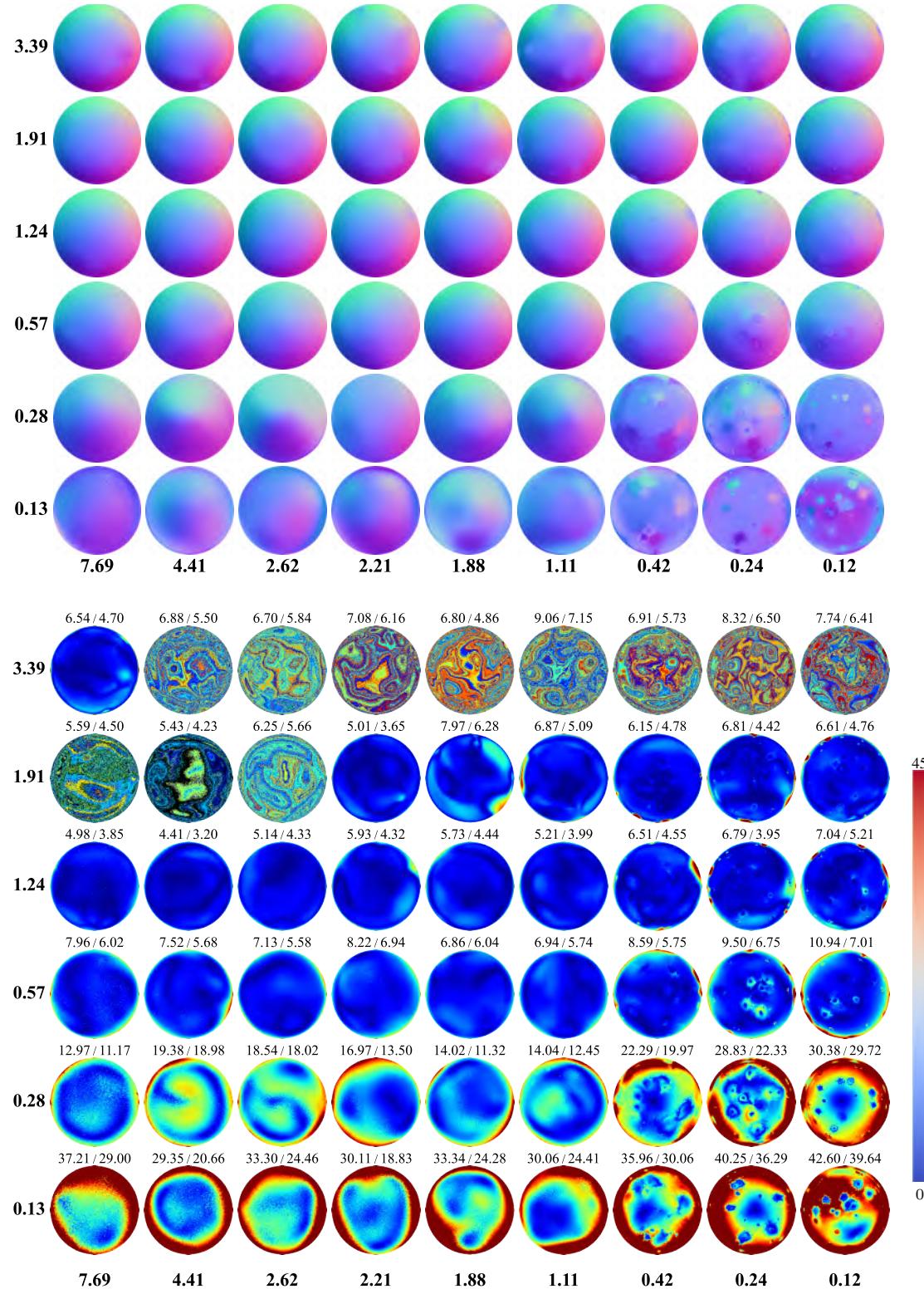


Figure S18. Estimated normal maps (top) and the corresponding angular error maps (bottom) of SPLINE-Net [21]. The mean and median errors for each material are displayed at the top of each error map.

PX-Net [11]

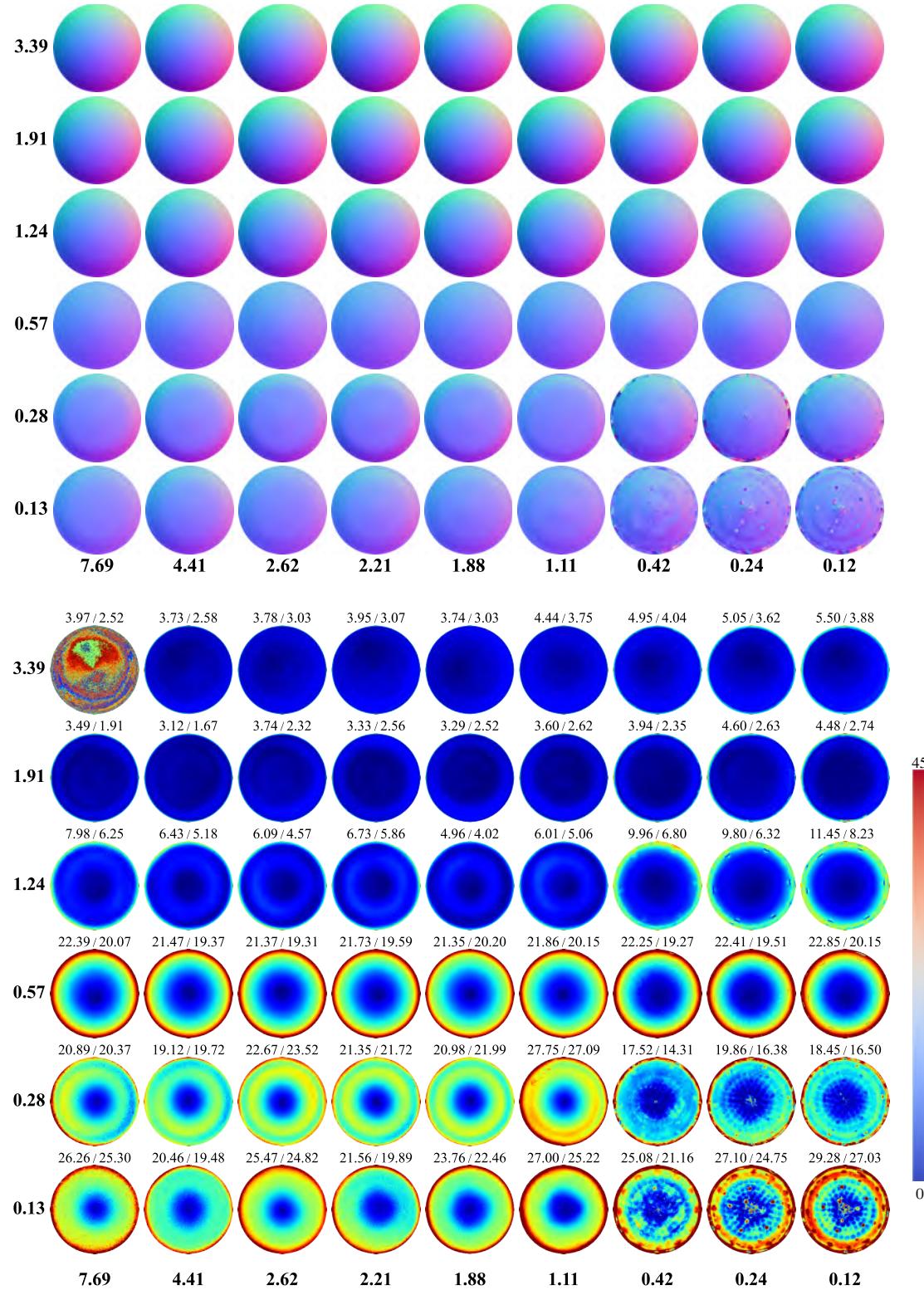


Figure S19. Estimated normal maps (top) and the corresponding angular error maps (bottom) of PX-Net [11]. The mean and median errors for each material are displayed at the top of each error map.

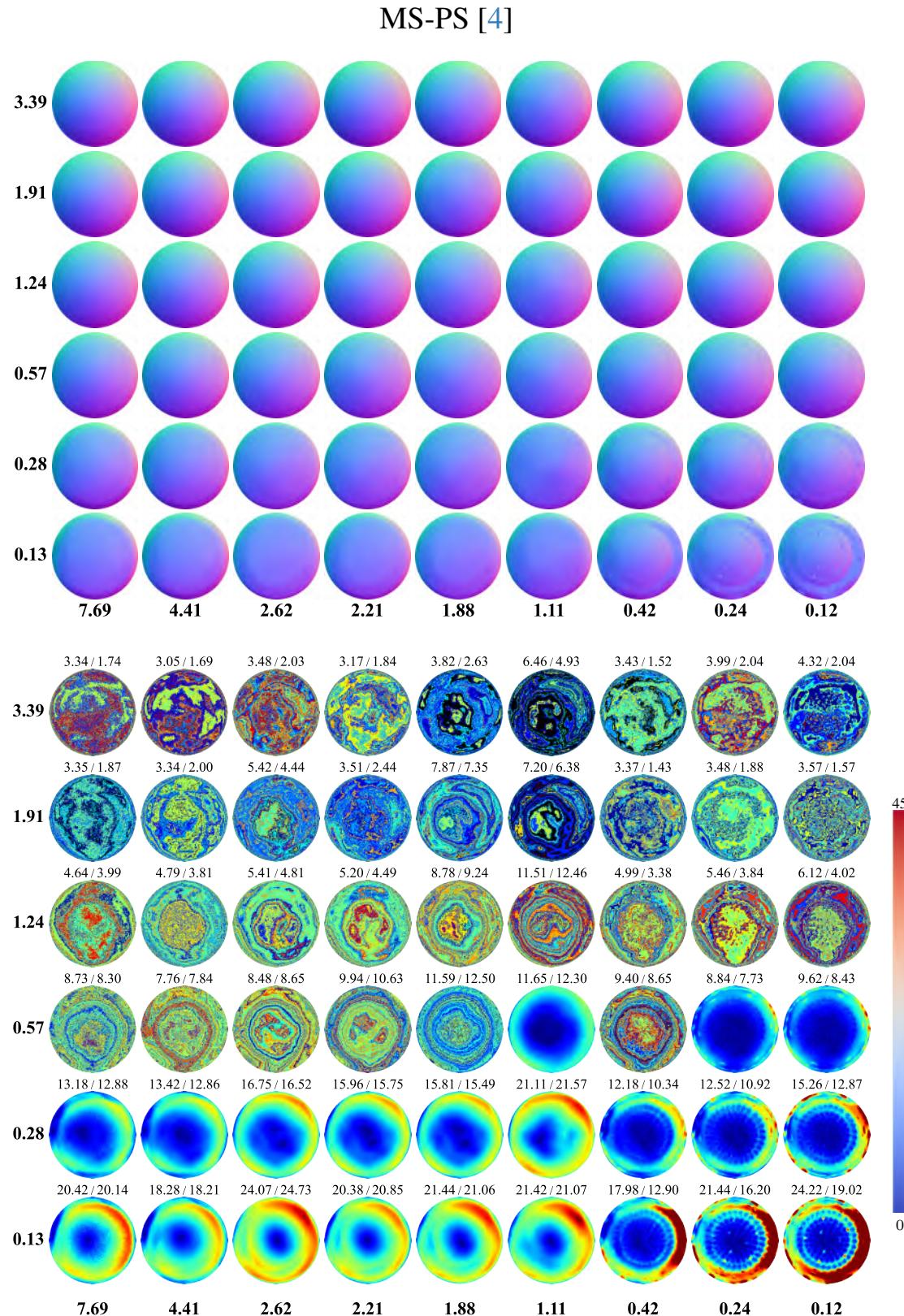


Figure S20. Estimated normal maps (top) and the corresponding angular error maps (bottom) of MS-PS [4]. The mean and median errors for each material are displayed at the top of each error map.

PF14 [12]

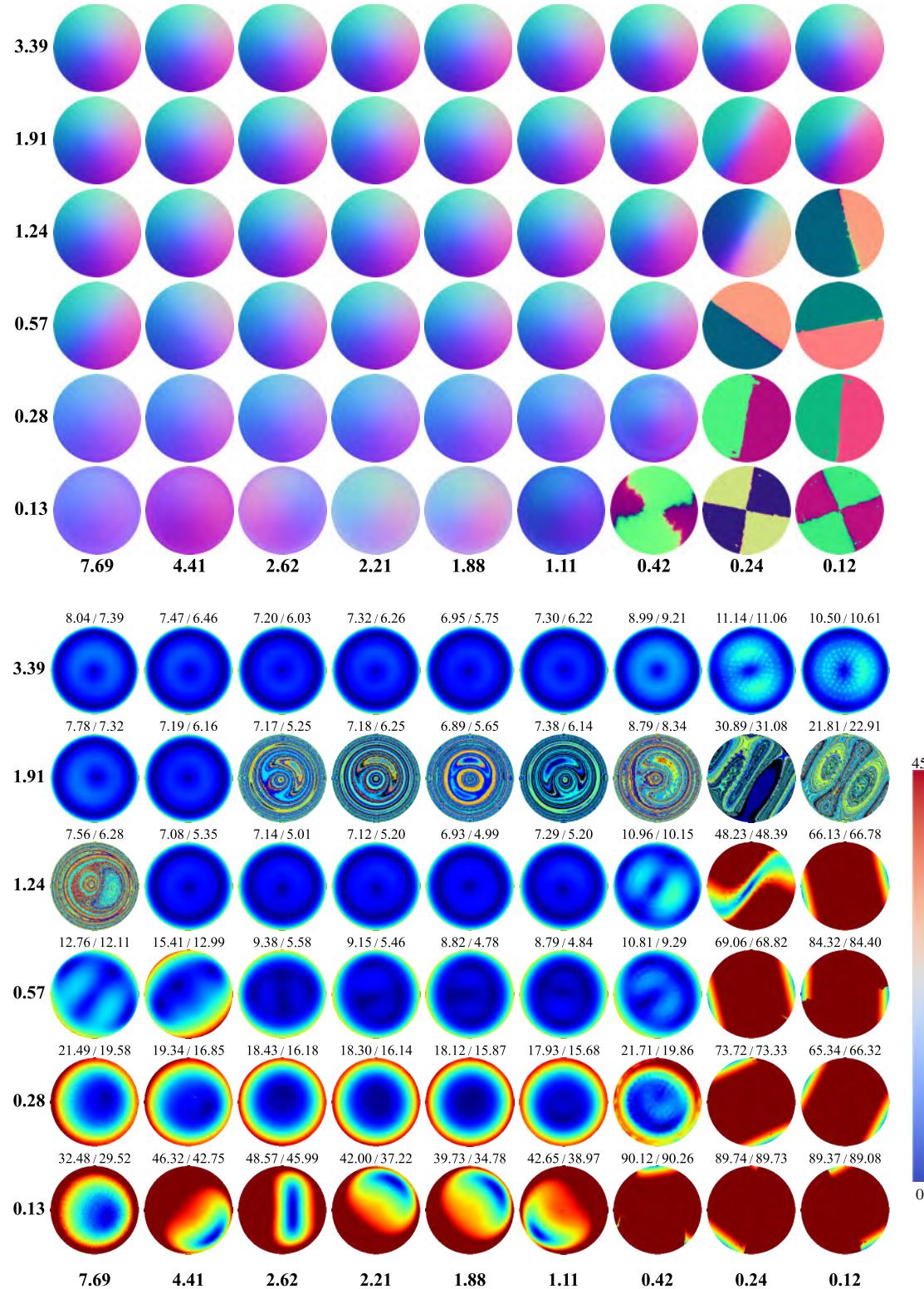


Figure S21. Estimated normal maps (top) and the corresponding angular error maps (bottom) of PF14 [12]. The mean and median errors for each material are displayed at the top of each error map.

SDPS-Net [3]

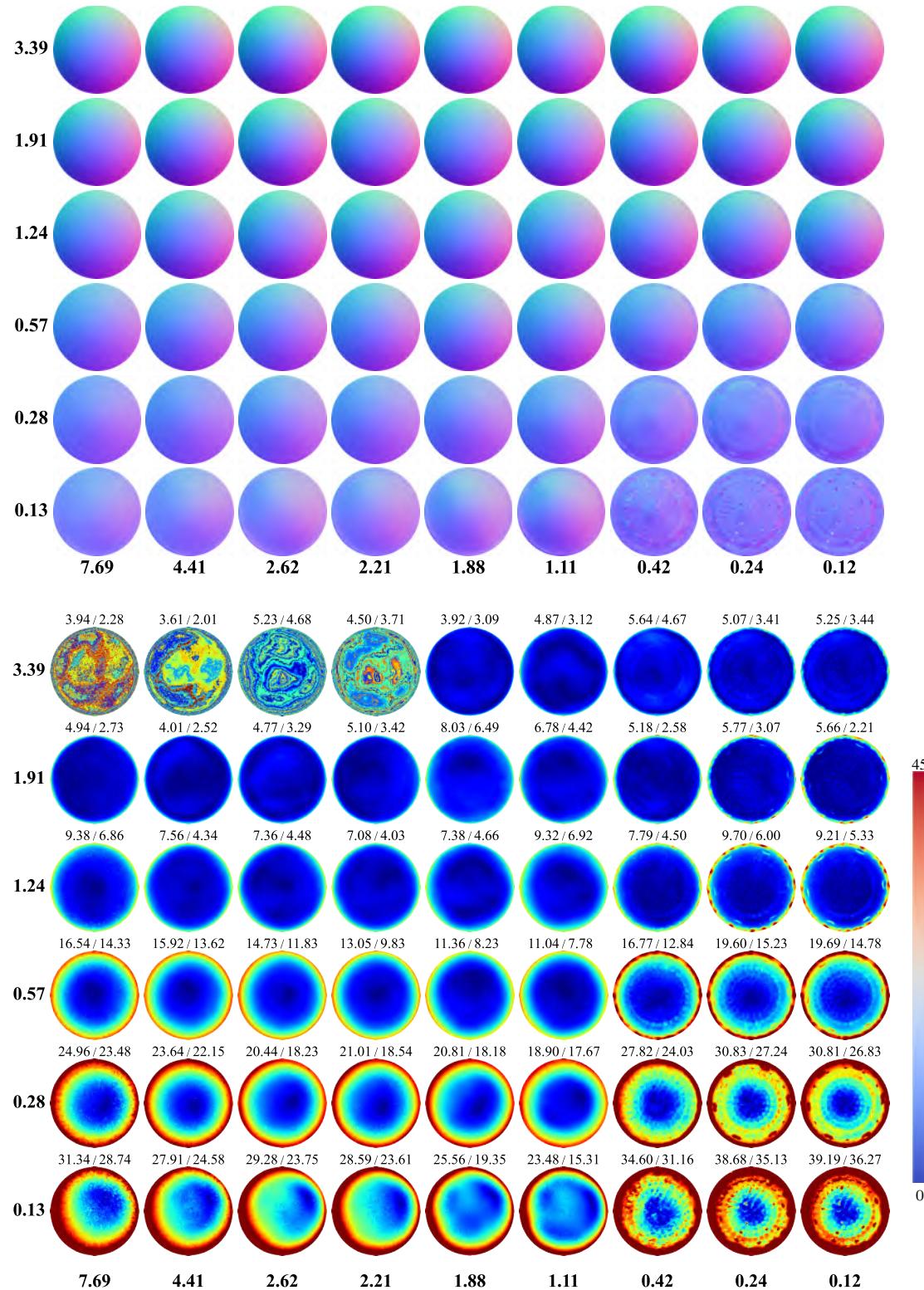


Figure S22. Estimated normal maps (top) and the corresponding angular error maps (bottom) of SDPS-Net [3]. The mean and median errors for each material are displayed at the top of each error map.

DeepPS2 [18]

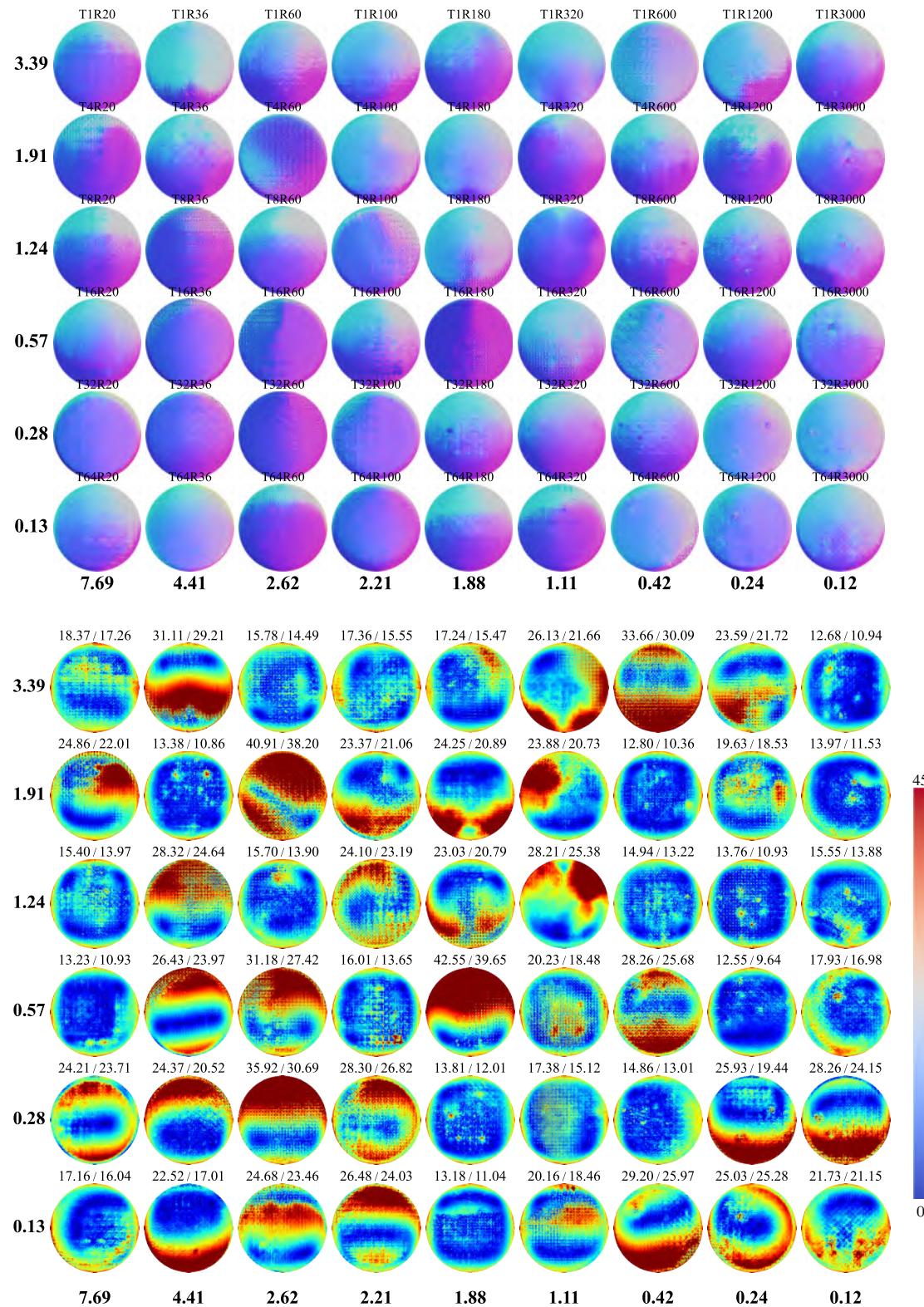


Figure S23. Estimated normal maps (top) and the corresponding angular error maps (bottom) of DeepPS2 [18]. The mean and median errors for each material are displayed at the top of each error map.

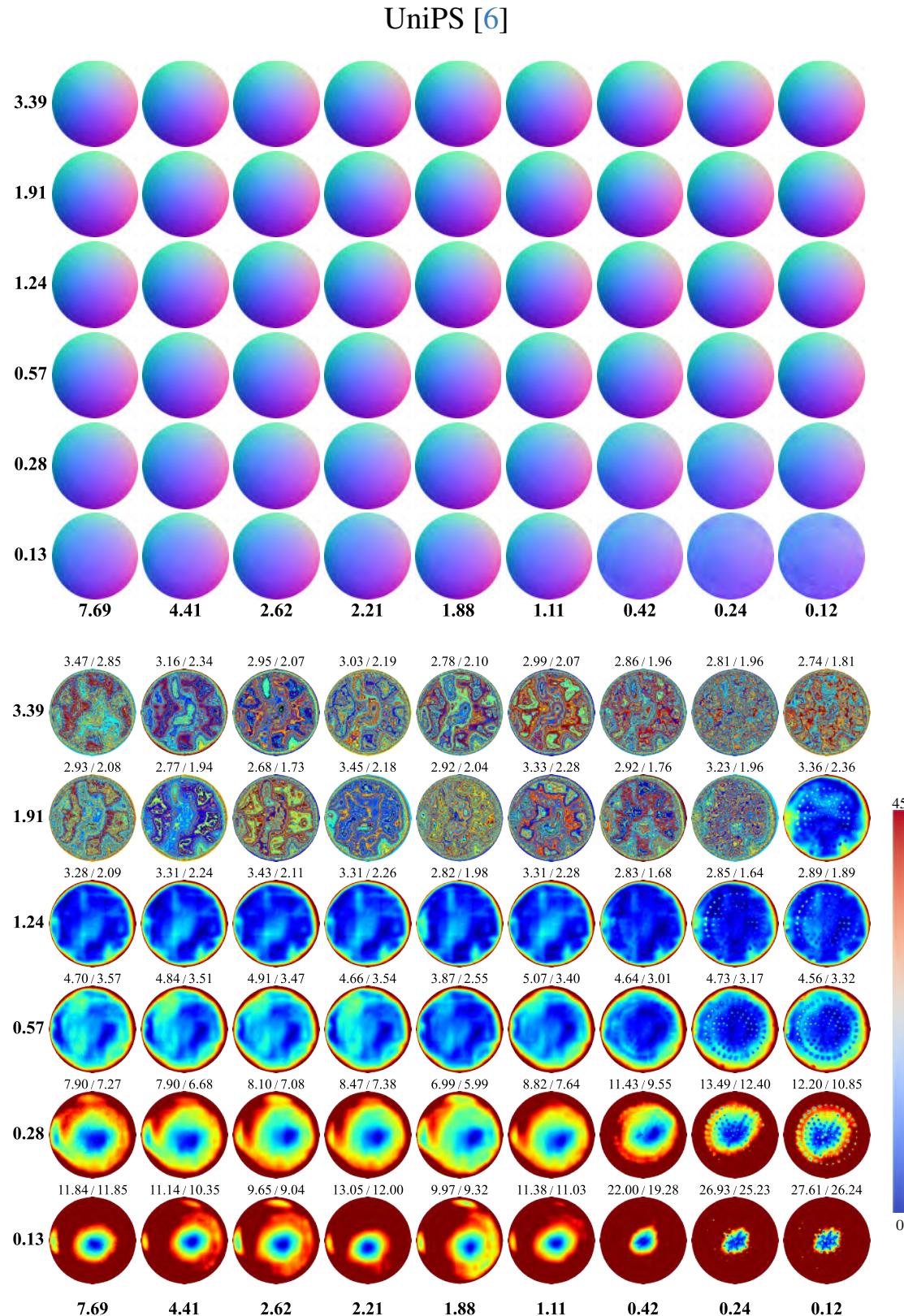


Figure S24. Estimated normal maps (top) and the corresponding angular error maps (bottom) of UniPS [6]. The mean and median errors for each material are displayed at the top of each error map.

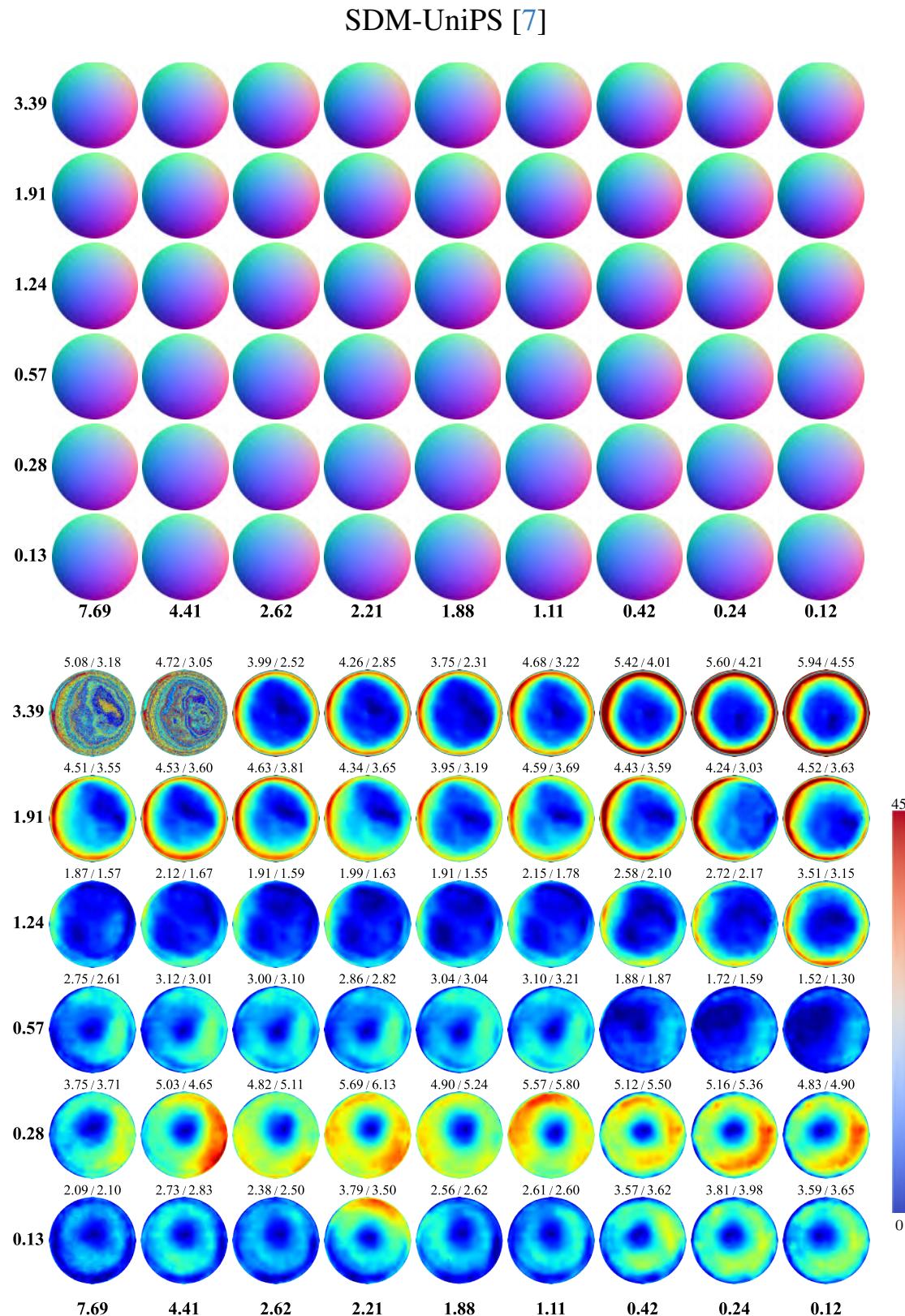


Figure S25. Estimated normal maps (top) and the corresponding angular error maps (bottom) of SDM-UniPS [7]. The mean and median errors for each material are displayed at the top of each error map.

169 **References**

- 170 [1] Brent Burley and Walt Disney Animation Studios.
171 Physically-based shading at disney. In *Proc. of SIG-
172 GRAPH*, pages 1–7. vol. 2012, 2012. 3
- 173 [2] Guanying Chen, Kai Han, and Kwan-Yee K. Wong. PS-
174 FCN: A flexible learning framework for photometric stereo.
175 In *Proc. of European Conference on Computer Vision
(ECCV)*, 2018. 3, 5, 12
- 176 [3] Guanying Chen, Kai Han, Boxin Shi, Yasuyuki Matsushita,
177 and Kwan-Yee K. Wong. Self-calibrating deep photometric
178 stereo networks. In *Proc. of IEEE Conference on Computer
179 Vision and Pattern Recognition (CVPR)*, 2019. 3, 5, 19
- 180 [4] Clément Hardy, Yvain Quéau, and David Tschumperlé.
181 MS-PS: A multi-scale network for photometric stereo with
182 a new comprehensive training dataset. *arXiv preprint
183 arXiv:2211.14118*, 2022. 5, 17
- 184 [5] Satoshi Ikehata. CNN-PS: CNN-based photometric stereo
185 for general non-convex surfaces. In *Proc. of European
186 Conference on Computer Vision (ECCV)*, 2018. 3, 4, 5, 11
- 187 [6] Satoshi Ikehata. Universal photometric stereo network us-
188 ing global lighting contexts. In *Proc. of IEEE Conference
189 on Computer Vision and Pattern Recognition (CVPR)*, pages
190 12591–12600, 2022. 1, 4, 5, 21
- 191 [7] Satoshi Ikehata. Scalable, detailed and mask-free universal
192 photometric stereo. *arXiv preprint arXiv:2303.15724*, 2023.
193 3, 5, 22
- 194 [8] John Illingworth and Josef Kittler. The adaptive hough trans-
195 form. *IEEE Transactions on Pattern Analysis and Machine
196 Intelligence*, (5):690–698, 1987. 1
- 197 [9] Chika Inoshita, Yasuhiro Mukaigawa, Yasuyuki Matsushita,
198 and Yasushi Yagi. Surface normal deconvolution: Photo-
199 metric stereo for optically thick translucent objects. In *Proc.
200 of European Conference on Computer Vision (ECCV)*, pages
201 346–359, 2014. 5, 9
- 202 [10] Yakun Ju, Boxin Shi, Muwei Jian, Lin Qi, Junyu Dong, and
203 Kin-Man Lam. Normattention-PSN: A high-frequency re-
204 gion enhanced photometric stereo network with normalized
205 attention. *International Journal of Computer Vision*, 130
206 (12):3014–3034, 2022. 3, 5, 14
- 207 [11] Fotios Logothetis, Ignas Budvytis, Roberto Mecca, and
208 Roberto Cipolla. PX-NET: Simple and efficient pixel-wise
209 training of photometric stereo networks. In *Proc. of Inter-
210 national Conference on Computer Vision (ICCV)*, pages
211 12757–12766, 2021. 1, 3, 4, 5, 16
- 212 [12] Thoma Papadimitri and Paolo Favaro. A closed-form, con-
213 sistent and robust solution to uncalibrated photometric stereo
214 via local diffuse reflectance maxima. *International Journal
215 of Computer Vision*, 2014. 3, 5, 18
- 216 [13] Erik Reinhard, Wolfgang Heidrich, Paul Debevec, Sumanta
217 Pattanaik, Greg Ward, and Karol Myszkowski. *High dy-
218 namic range imaging: acquisition, display, and image-based
219 lighting*. Morgan Kaufmann, 2010. 3
- 220 [14] Jieji Ren, Feishi Wang, Jiahao Zhang, Qian Zheng, Mingjun
221 Ren, and Boxin Shi. DiLiGenT10²: A photometric stereo
222 benchmark dataset with controlled shape and material varia-
223 tion. In *Proc. of IEEE Conference on Computer Vision and
224 Pattern Recognition (CVPR)*, 2022. 1
- 225 [15] Masaki Samejima and Yasuyuki Matsushita. Fast general
226 norm approximation via iteratively reweighted least squares.
227 In *Proceedings of Asian Conference on Computer Vision
228 Workshops (ACCVW)*, 2016. 1
- 229 [16] Boxin Shi, Ping Tan, Yasuyuki Matsushita, and Katsushi
230 Ikeuchi. Bi-polynomial modeling of low-frequency re-
231 flectances. *IEEE Transactions on Pattern Analysis and Ma-
232 chine Intelligence*, 2014. 3, 5, 10
- 233 [17] Boxin Shi, Zhipeng Mo, Zhe Wu, Dinglong Duan, Sai-Kit
234 Yeung, and Ping Tan. A benchmark dataset and evaluation
235 for non-Lambertian and uncalibrated photometric stereo.
236 *IEEE Transactions on Pattern Analysis and Machine Intel-
237 ligence*, 2019. 1, 3, 5, 7, 8
- 238 [18] Ashish Tiwari and Shanmuganathan Raman. DeepPS2: Re-
239 visiting photometric stereo using two differently illuminated
240 images. In *Proc. of European Conference on Computer Vi-
241 sion (ECCV)*, pages 129–145. Springer, 2022. 1, 4, 5, 20
- 242 [19] Robert J. Woodham. Photometric method for determining
243 surface orientation from multiple images. *Optical engineer-
244 ing*, 1980. 3, 5, 6
- 245 [20] Zhuokun Yao, Kun Li, Ying Fu, Haofeng Hu, and Boxin Shi.
246 GPS-Net: Graph-based photometric stereo network. *Proc. of
247 Annual Conference on Neural Information Processing Sys-
248 tems (NeurIPS)*, 33:10306–10316, 2020. 1, 4, 5, 13
- 249 [21] Qian Zheng, Yiming Jia, Boxin Shi, Xudong Jiang, Ling-Yu
250 Duan, and Alex C. Kot. SPLINE-Net: Sparse photometric
251 stereo through lighting interpolation and normal estimation
252 networks. In *Proc. of International Conference on Computer
253 Vision (ICCV)*, 2019. 1, 3, 4, 5, 15