

# DiLiGenRT: A Photometric Stereo Dataset with Quantified Roughness and Translucency

Anonymous CVPR submission

Paper ID 2461

## Abstract

Photometric stereo faces challenges from non-Lambertian reflectance in real-world scenarios. Systematically measuring the reliability of photometric stereo methods in handling such complex reflectance necessitates a real-world dataset with quantitatively controlled reflectances. This paper introduces DiLiGenRT, the first real-world dataset for evaluating photometric stereo methods under quantified reflectances by manufacturing 54 hemispheres with varying degrees of two reflectance properties: **Roughness** and **Translucency**. Unlike qualitative and semantic labels, such as “diffuse” and “specular,” that have been used in previous datasets, our quantified dataset allows comprehensive and systematic benchmark evaluations. In addition, it facilitates selecting best-fit photometric stereo methods based on the quantitative reflectance properties. Our dataset and benchmark results will <http://be.available.upon.acceptance>.

001  
002  
003  
004  
005  
006  
007  
008  
009  
010  
011  
012  
013  
014  
015  
016

036  
037  
038  
039  
040  
041  
042  
043  
044  
045  
046  
047  
048  
049

metallic, steel, and wood. These semantic labels are intuitive but at the same time ambiguous. For example, a steel surface can exhibit a rough or mirror-like reflectance depending on the manufacturing process. Furthermore, these semantic labels are not explicitly ordered and non-metric; therefore, it is difficult to predict the accuracy of photometric stereo for “in-between” reflectances, *e.g.*, predicting accuracy for the paper reflectance from the evaluations based on steel and wood reflectances. In addition, existing real-world datasets [34, 36, 44, 48, 52] primarily focus on opaque surfaces, like plaster and ceramic. Since a wide range of real-world objects exhibit significant translucency, such as human skin and plant leaves, it is desired to have a dataset with thorough coverage of varying levels of translucency.

To address these problems, we build a new real-world dataset, DiLiGenRT<sup>1</sup>, with reflectances controlled in a *quantified* manner by carefully manufacturing materials. DiLiGenRT consists of **54** hemispheres made of materials having **9** levels of roughness and **6** levels of translucency, as shown in Fig. 1. The design of the shared geometry allows us to isolate the error source from shape variations and solely attribute normal estimation errors to varying reflectances. At the same time, for the test purpose on shape diversity, we further introduce DiLiGenRT-S, which contains three distinct Shapes, as depicted in Fig. 1 (right).

Using the DiLiGenRT dataset, we conduct benchmark evaluations to provide each photometric stereo method a performance profile matrix under various roughness and translucency levels. As reflectance properties are quantifiable, continuous, and common to various materials, it enables a comprehensive analysis of photometric stereo’s working range w.r.t. reflectance, and facilitates selecting best-fit photometric stereo methods based on the reflectance properties. In addition, it offers the capability to infer the performance of photometric stereo methods for materials that are not directly covered in the dataset based on the numerical interpolation of reflectance properties.

050  
051  
052  
053  
054  
055  
056  
057  
058  
059  
060

017

## 1. Introduction

018  
019  
020  
021  
022  
023  
024  
025  
026  
027  
028  
029  
030  
031  
032  
033  
034  
035

Photometric stereo [53] aims at recovering detailed surface normals and reflectances from images captured under varying illuminations from a fixed camera. To measure the applicability of photometric stereo methods in the real world, comprehensive benchmark evaluations on real captured datasets are desired. General reflectance is one of the key challenges for photometric stereo, making it a central consideration when building real-world photometric stereo datasets. By evaluating a photometric stereo method on a real-world dataset, we seek to know the exact reflectance conditions where the method would become trustworthy. Achieving this necessitates a dedicated design of the real-world dataset with quantitatively controlled reflectances.

Previous datasets for photometric stereo are constructed with reflectances defined in a *qualitative* and *semantic* manner. Datasets like DiLiGenT [48], LUCES [36, 40], DiLiGenT10<sup>2</sup> [44], and DiLiGenT-II [52] categorize reflectances using a set of semantic labels, such as specular,

061  
062  
063  
064  
065  
066  
067  
068  
069  
070  
071  
072

<sup>1</sup>DiLiGenRT = Directional Lighting with Generalized Roughness and Translucency

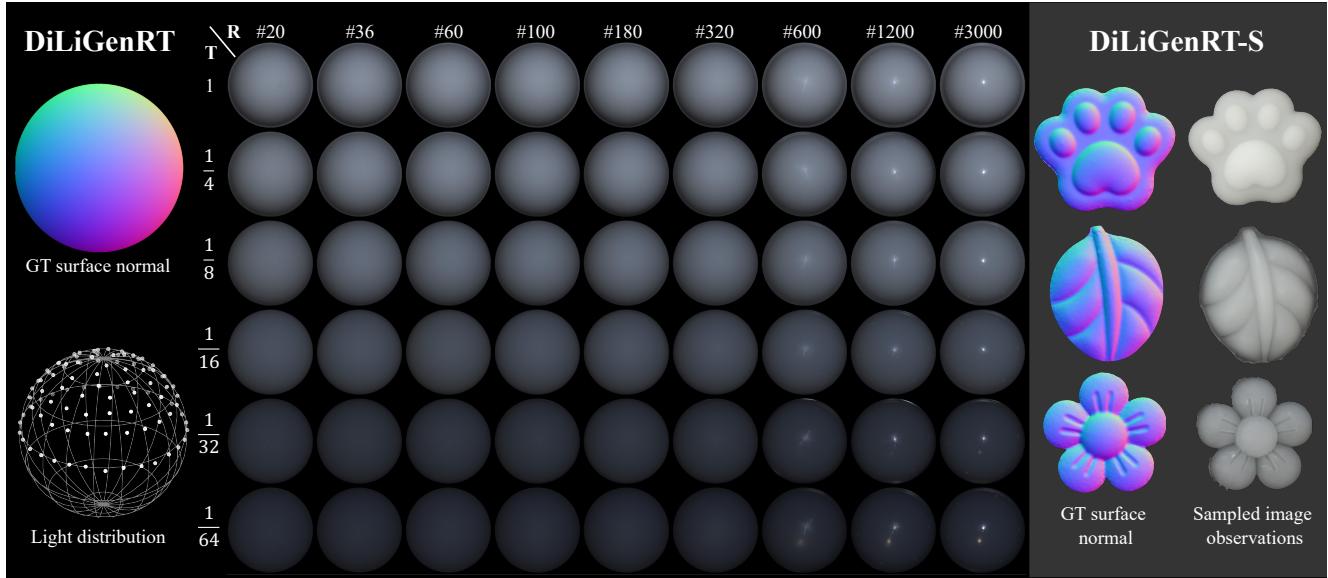


Figure 1. Overview of DiLiGenRT represented as a **translucency-roughness** ( $6 \times 9$ ) matrix and DiLiGenRT-S including 3 distinct shapes. **(Left)** All objects in DiLiGenRT share the same ground truth sphere surface normal and are illuminated by 100 uniformly distributed directional lights. **(Middle)** Image observations of DiLiGenRT under varying degrees of roughness (R) and translucency (T) indexed by the solution concentration and grit size, respectively. The larger the solution concentration or grit size, the lesser the roughness or translucency. From top-left to bottom-right, the surface becomes shinier as roughness decreases, while the image appearance grows darker as translucency increases. **(Right)** The ground truth surface normals and selected image samples are included in DiLiGenRT-S.

**Contributions** In this paper, we present DiLiGenRT, the first real-world photometric stereo dataset manufactured with different levels of quantified roughness and translucency. Based on DiLiGenRT, we provide a benchmark result presented as performance profiles of photometric stereo methods, showing their working ranges over quantitative reflectances.

## 2. Related works

This paper’s main goal is to provide a benchmark dataset for photometric stereo. We begin by reviewing photometric stereo methods categorized by their reflectance assumptions. After that, we summarize existing synthetic and real-world datasets for photometric stereo.

### 2.1. Photometric stereo methods

We briefly summarize photometric stereo methods based on their reflectance assumptions: Lambertian and more general Bidirectional Reflectance Distribution Functions (BRDFs) for opaque surfaces, and subsurface scattering for translucent surfaces.

**Lambertian reflectance** The Lambertian reflectance model is an ideal diffuse reflectance model that has been extensively used in the literature because of its simplicity and capability of representing diffuse reflectances. Lambertian photometric stereo is well-posed with calibrated distant

lights and can be solved in a closed form given three or more images under different light directions [49, 53]. In an uncalibrated light setting, surface normal can be recovered up to a linear ambiguity [17]. With the surface integrability [19], the linear ambiguity can be reduced to a General Bas-Relief (GBR) ambiguity [3]. Existing uncalibrated photometric stereo methods resolve this GBR ambiguity by clustering uniform albedo areas [46], considering specific light distributions (e.g., symmetric light [41], ring light [60], differential light [5]) or perspective camera projections [42]. Papadimitriou *et al.* [43] proposes to detect local diffuse reflectance maxima from image observations, which infers surface normals coinciding with the light direction.

**General BRDFs for opaque surfaces** Photometric stereo for surfaces with general BRDFs is more challenging than the ideal Lambertian case. Under calibrated settings, existing optimization-based methods assume dominant diffuse reflectance and treat specular highlights in BRDF as sparse outliers, developing robust photometric stereo approaches based on sparse Bayesian regression [25], low-rank matrix completion [54], and positional thresholding strategies [48]. Other methods explicitly represent general BRDFs via bi-polynomial approximation [47], Microfacet reflectance models [9, 14], or linear bases [12, 13, 20].

Since the work of Santo *et al.* [45], learning-based photometric stereo methods have achieved significant progress

097  
098  
099  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109

110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122

Table 1. Summary of real (top) and synthetic (bottom) photometric stereo datasets organized by number (#) of distinct translucent materials, surface normal maps, lights, and sets (one set means a sequence of photometric stereo images under varying lights).

| Dataset                      | Reflectance control   | # Translucency | # Normal | # Lights | # Sets |
|------------------------------|-----------------------|----------------|----------|----------|--------|
| DiLiGenT [48]                | Semantic labels       | 0              | 10       | 96       | 10     |
| DiLiGenT-MV [34]             | Semantic labels       | 0              | 100      | 100      | 100    |
| LUCES [36, 40]               | Semantic labels       | 0              | 14       | 52       | 14     |
| Harvard [56]                 | Semantic labels       | 0              | 7        | 20       | 7      |
| ETHz [31]                    | Semantic labels       | 0              | 3        | 260      | 3      |
| Gourd&Apple [2]              | Semantic labels       | 0              | 2        | 102/112  | 2      |
| DiLiGenT-II [52]             | Semantic labels       | 1              | 30       | 100      | 30     |
| DiLiGenT10 <sup>2</sup> [44] | Semantic labels       | 1              | 10       | 100      | 100    |
| <b>DiLiGenRT</b>             | Quantified properties | 6              | 1        | 100      | 54     |
| <b>DiLiGenRT-S</b>           | Quantified properties | 3              | 3        | 100      | 3      |
| BlobbyPS [45]                | Semantic labels       | 0              | 8        | 96       | 800    |
| PS-Sculpture [6]             | Semantic labels       | 0              | 59,292   | 64       | 59,292 |
| CyclesPS [21]                | Quantified properties | 0              | 15       | 1,300    | 45     |
| SymPS [51]                   | Quantified properties | 0              | 5        | 77       | 9      |
| PS-Relief [52]               | Quantified properties | 0              | 127      | 100      | 3,429  |
| PS-Wild [23]                 | Quantified properties | 0              | 410      | 31       | 10,099 |
| PS-Mix [24]                  | Quantified properties | 0              | 410      | 10       | 34,921 |

in handling general BRDFs by finding a direct mapping between input images to the surface normal map. The proposed network structures can be divided into the all-pixel branch [6, 30] and the per-pixel branch [21, 33, 35, 59]. Follow-up works further improve photometric stereo by addressing global illumination effects [16, 35], reducing the number of inputs [33, 50, 51, 59], and combining the merits of per-pixel-based and all-pixel-based methods [22, 57].

Beginning from SDPS-Net [7], learning-based uncalibrated photometric stereo methods under general BRDFs are proposed. Follow-up works focus on analyzing the feature map of deep uncalibrated photometric stereo [8] or applying an inverse rendering module to solve uncalibrated photometric stereo in a self-supervised manner [31, 32]. The most recent work, UniPS [23] and SDM-UniPS [24], introduce neural global light context to recover surface normals under unknown, spatially-varying illumination, which is shown to be effective for surfaces with general BRDFs.

**Subsurface scattering** Subsurface scattering is non-local light transport beneath a surface, which is observed on translucent surfaces. Due to the complexity of subsurface scattering, only a few photometric stereo methods have been developed for translucent surfaces. Inoshita *et al.* [26] observe that subsurface scattering in optically thick translucent objects can be approximated as a convolution with a blur kernel, thus recasting photometric stereo as a deconvolution problem. Similarly, Dong *et al.* [11] frame photometric stereo with subsurface scattering as a blind deconvolution problem, jointly estimating subsurface scattering parameters and surface normals through an inverse rendering based on the dipole-diffusion model [28]. However, these methods either assume a restrictive Lambertian reflectance [26] or near-planar surface shapes [11], limiting their application scope.

## 2.2. Photometric stereo datasets

Table 1 lists existing synthetic and real-world photometric stereo datasets. Here we discuss their properties in reflectance, shape, and illumination in detail.

**Synthetic dataset** utilizes physics-based rendering engines, like Mitsuba [27] and Blender [10], to produce image observations and corresponding surface normal maps for diverse synthetic scenes. During the rendering, factors like surface normal, illumination, and material are completely controlled. For example, the BlobbyPS dataset [45] contains 8 smooth Blobby shapes [29] with reflectances of measured MERL BRDF dataset [38]. PS-Sculpture [6] extends BlobbyPS by adding 59, 292 diverse surface normals. Both datasets control the reflectance via semantic labels of the MERL dataset [38], such as blue-paper and alum-bronze.

CyclesPS [21] controls continuous parameters of reflectance properties (*e.g.*, roughness, metallic, and specular) based on the Disney Principled BSDF model [4]. Unlike semantic labels, reflectance properties can be quantitatively assigned, making it possible to create a large number of unique materials (*e.g.*, 30,000 materials are generated in CyclesPS [21]). The follow-up datasets, such as SymPS [51], PS-Wild [23] and PS-Mix [24] build upon CyclesPS [21], increase the dataset scale by including more diverse shapes, materials, and illuminations.

All these synthetic datasets assume an opaque surface without considering the translucency. In reality, translucent surfaces are relevant everywhere, such as human skin and plant leaves; thus, it is desired to have a dataset with controlled translucency.

**Real-world dataset** Real-world datasets complement synthetic ones by bridging the gap between computer graphics rendering and actual image formation. For instance, the Gourd&Apple dataset [1] offers image observations of two objects with spatially-varying isotropic BRDFs. The Harvard dataset [55] includes 7 surfaces with a uniform diffuse reflectance. However, these two datasets do not provide the ground truth surface normals. The DiLiGenT [48] dataset includes 10 objects of different shapes and general reflectances. It introduces benchmark evaluation of photometric stereo for the first time based on the ground truth surface normals from scanned meshes. Subsequent datasets further expand DiLiGenT [48], introducing aspects like multi-views (DiLiGenT [48]-MV [34]), near-field illumination (LUCES [40]), environment illumination [15, 18], planar surfaces with rich details (DiLiGenT-II [52]), and global illumination effects [31].

Closest to our work, to disentangle the error from shape and reflectance, DiLiGenT10<sup>2</sup> [44] was proposed by fabricating 10 objects from CAD models with 10 carefully chosen

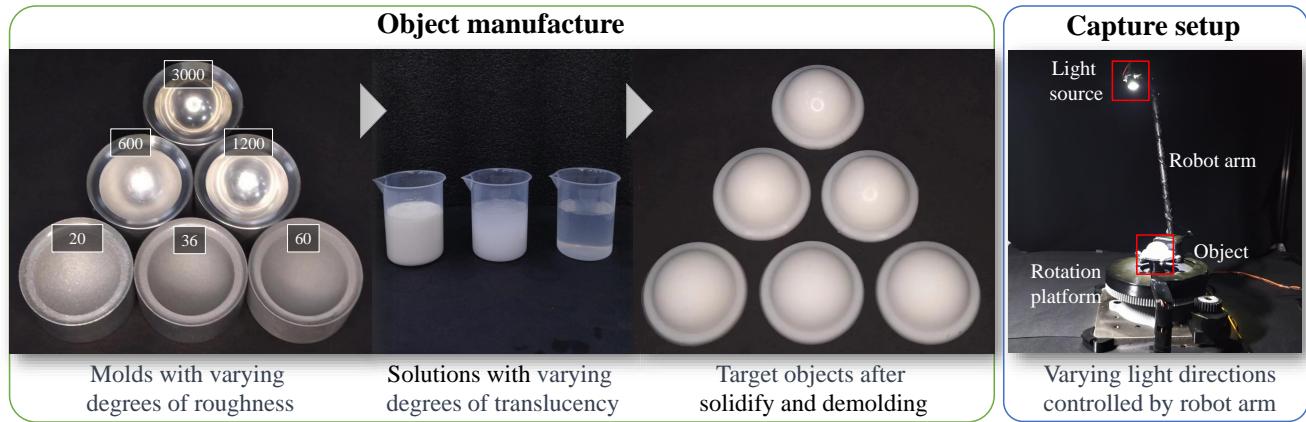


Figure 2. (Left) Objects in DiLiGenRT are constructed by first manufacturing molds with varying degrees of roughness through sandblasting and then injecting solutions of different concentrations into the molds, followed by solidifying and de-molding. (Right) Images in DiLiGenRT are captured by moving a point light source bundled on a robot arm.

208 materials (e.g., polyoxymethylene and steel). Unlike  
209 previous datasets where the reflectance is controlled via  
210 semantic labels, our DiLiGenRT quantitatively controls the  
211 reflectance properties by manufacturing 54 objects with 6  
212 translucency levels and 9 roughness levels. In this way, we  
213 provide a more comprehensive evaluation of photometric  
214 stereo from the reflectance aspect. Besides, compared with  
215 DiLiGenT10<sup>2</sup> [44] and DiLiGenT-II [52] providing only one  
216 translucent material, DiLiGenRT is more capable of evaluating  
217 the influence of subsurface scattering in photometric  
218 stereo by including different levels of translucency.

### 219 3. DiLiGenRT dataset

220 This section details the manufacturing procedure and cap-  
221 turing process of our real-world photometric stereo dataset,  
222 DiLiGenRT. We also introduce two setups for measuring the  
223 translucency and roughness of the materials in DiLiGenRT.

#### 224 3.1. Object manufacturing and capturing

225 As shown in Fig. 1, our DiLiGenRT consists of 54 spheres  
226 sharing the same shape but different materials. The objects  
227 on the same row share the same translucency, while those in  
228 the same column have the same roughness.

229 **Object manufacturing** To create DiLiGenRT, we begin  
230 with manufacturing 9 identical aluminum alloy molds. For  
231 creating varying roughness levels, the mold surfaces are  
232 sandblasted by grit with 9 different sizes ranging from #3000  
233 to #20 (FEPA standard). Figure 2 (left) shows 6 out of the  
234 9 molds, where the mold sandblasted with grit size #3000  
235 exhibits a shinier appearance than the one with grit size  
236 #20. To produce distinct levels of translucency, we blend  
237 transparent silicone with color paste in different proportions,  
238 resulting in 6 solutions with varying concentrations between

239 1/1 and 1/64 as shown in the second column of Fig. 2. By  
240 injecting one of these solutions into the molds of different  
241 roughness levels and solidifying them, we obtain objects  
242 with identical translucency but different roughness levels,  
243 as shown in the third column of Fig. 2. Objects with other  
244 levels of translucency are also created in a similar manner.  
245 In total, we manufacture 54 spherical objects, each having  
246 one of 6 translucency levels and one of 9 roughness levels.

247 **Capturing setup** As shown on the right side of Fig. 2,  
248 we use a capture setup similar to DiLiGenT10<sup>2</sup> [44], where  
249 light directions are controlled by moving a single point light  
250 source attached to a robot arm. A lens is mounted in front  
251 of the point light source to enhance the directionality and  
252 uniformity of illumination. We capture 100 images for each  
253 object under different light directions, uniformly distributed  
254 over a hemisphere as shown in the left bottom of Fig. 1.  
255 Please refer to our supplementary material for further details  
256 regarding light calibration and camera settings.

#### 257 3.2. Roughness and translucency measurement

258 This section describes the procedure for measuring the rough-  
259 ness and translucency of the manufactured materials.

260 **Roughness measurement** In the mechanical industry,  
261 measuring roughness is a well-explored field, leading to various  
262 roughness metrics. We choose surface area roughness  
263  $S_a$  as our roughness metric, which averages the absolute  
264 differences between each point's surface height and the local  
265 area's mean surface height. As depicted in Fig. 3, we assume  
266 the surface roughness across the entire sphere is uniform and  
267 conduct surface height measurements at a highlighted local  
268 area. The surface heights are measured using a white light  
269 interferometer Zyglo Nexview NX2<sup>2</sup> with vertical resolution

<sup>2</sup><https://www.zygo.com/products/metrology-systems/3d-opticalprofilers/nexview-nx2>. Retrieved

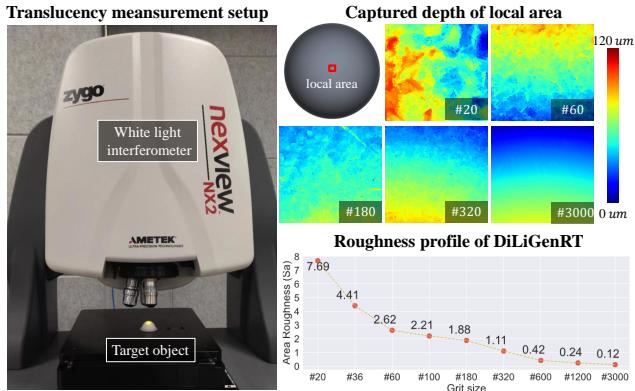


Figure 3. Roughness measurement at varying levels of grit size labeled in gray boxes. The measured roughness profile is shown in the bottom right.

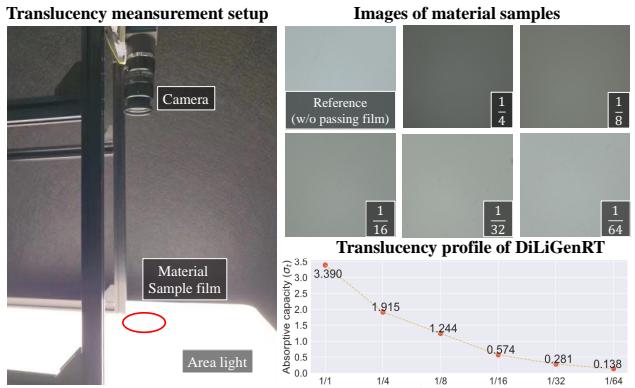


Figure 4. Translucency measurement at varying levels of solution concentration labeled in gray boxes. Observed radiance without a material sample is shown as a reference image. The translucency profile is shown in the bottom right.

of 0.15 [nm]. As the grit size decreases, we observe the reduction in the variance of measured surface heights within the local area, showing that our objects have different levels of roughness. The roughness profile of our DiLiGenRT, based on the  $Sa$  metric w.r.t. different grit sizes, is shown in Fig. 3 (bottom-right).

**Translucency measurement** Based on the Beer-Lambert law [39], we use the absorptive capacity defined as  $\sigma_t = \varepsilon c$  as the metric of translucency, which jointly considers the solution concentration  $\varepsilon$  and molar extinction coefficient  $c$  within the medium. Smaller  $\sigma_t$  corresponds to a higher translucency level. As shown in Fig. 4, we manufacture 6 thin disks with a thickness of 3 [mm], solidified from the same 6 solutions used for DiLiGenRT. We put each thin film onto a screen light source and record the images by a camera with and without the film, whose average pixel

Nov. 17, 2023.

values are recorded as  $I_0$  and  $I_1$ , respectively. Due to the subsurface scattering, the received radiance decreases as the concentration of the solution increases. The absorptive capacity is calculated as  $\sigma_t = -\log_{10}(\frac{I_0}{I_1})/d$ , where  $d$  is the thickness of the thin medium [39]. The translucency profile of our DiLiGenRT measured by metric  $\sigma_t$  w.r.t. different solution concentrations is shown in Fig. 4 (bottom-left).

## 4. Benchmark evaluations

Similar to the latest photometric stereo benchmark DiLiGenT10<sup>2</sup> [44], we choose non-learning-based methods LS [53] and its robust version via position thresholding strategy TH28 [48] and TH46 [48], and representative work ST14 [47] as the baselines. For learning-based methods, we select classical photometric stereo methods PS-FCN [6] and CNN-PS [21], as well as the latest photometric stereo methods MS-PS [16] and NormAttention-PSN [30] (abbreviated as Att-PSN) as the baselines. Besides, we add a photometric stereo method tailored for translucent surfaces: DeconvPS [26]. We also include three uncalibrated photometric stereo methods: PF14 [43], SDPS-Net [7] (abbreviated as SDPS), and SDM-UniPS [24] in our benchmark.

The evaluations of other photometric stereo methods: UniPS [23], GPS-Net [57], PX-Net [35], and SPLINE-Net [59], DeepPS2 [50] are provided in our supplementary material. We implemented DeconvPS [26] based on their paper by utilizing 2D Gaussian functions as the convolution kernel and assigning the kernel size with the inverse of absorptive capacity value empirically. For the remaining baselines, we adopt the code and pre-trained models released by the authors. Following the practice of previous benchmark datasets, we use mean angular error (MAE) in degrees as the metric to evaluate estimated surface normals.

### 4.1. Roughness-translucency error matrix

Figure 5 presents the roughness-translucency error matrices as a heat map for each method evaluated on DiLiGenRT, where the non-learning-based and learning-based photometric stereo methods are shown in the top and bottom rows, respectively. Each cell in the matrix shows the MAE for the material indexed by the translucency and roughness levels based on the  $\sigma_t$  and  $Sa$  metrics. These matrices provide a performance profile of each photometric stereo method across the quantified reflectances.

**Non-learning-based methods** Opaque materials with a high degree of roughness, as in the matrix's top left corner, exhibit satisfactory results by all photometric stereo methods because the reflectance is close to Lambertian. The surface normal estimation error tends to grow as translucency increases and roughness decreases. However, translucency appears to have a more significant influence than roughness.

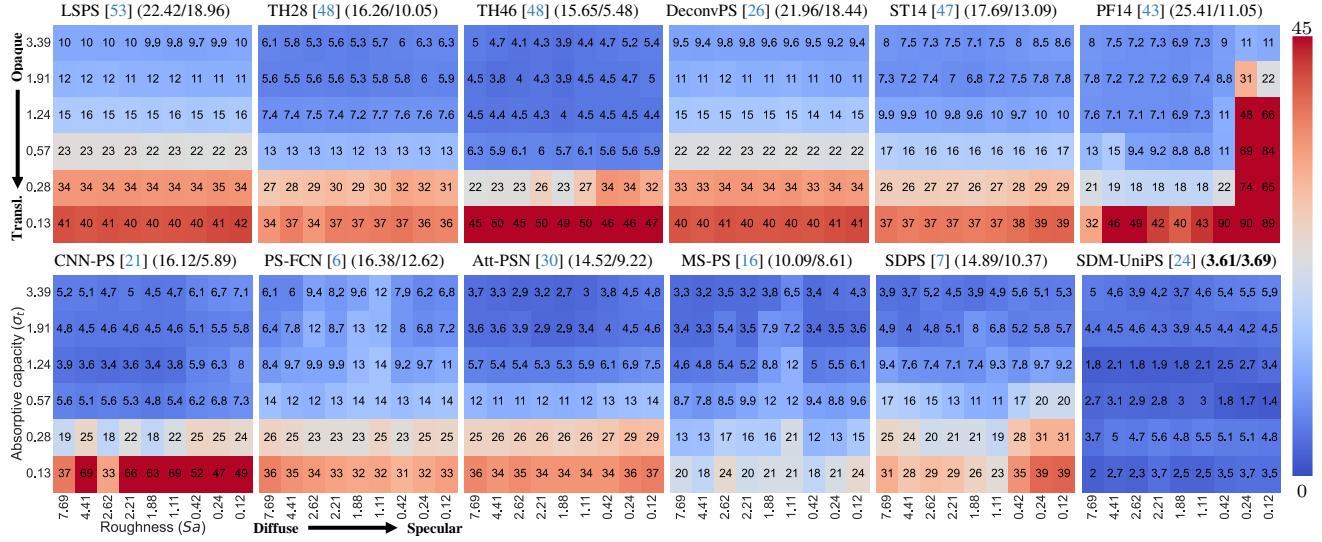


Figure 5. Roughness-translucency MAE matrices for non-learning-based (top) and learning-based (bottom) photometric stereo methods, showing their performance profiles under different levels of reflectance properties. The mean and median of the MAE matrix are presented near the method name. “Att-PSN” and “SDPS” are the abbreviations of NormAttention-PSN [30] and SDPS-Net [7]. The ticks of row and column are  $\sigma_t$  and  $Sa$ . Reducing  $\sigma_t$  corresponds to increasing translucency, while lowering  $Sa$  is associated with decreased roughness.

335 DeconvPS [26] enhances LS [53] by addressing subsurface  
336 scattering and shows smaller MAE values on translucent surfaces.  
337 However, the method is still prone to substantial errors  
338 due to attached shadows and the Lambertian assumption.

339 By using a position thresholding strategy, TH46 [48] out-  
340 performs all other non-learning-based photometric stereo  
341 methods. It indicates that for the mid-range of translucency  
342 levels (*i.e.*,  $\sigma_t \in [0.57, 3.39]$ ), the Lambertian component  
343 is still strongly observed, which can be extracted via the  
344 simple position thresholding strategy. However, a notable  
345 error surge occurs when  $\sigma_t \leq 0.28$ , corresponding to the  
346 solution concentration of 1/32. The Lambertian component  
347 becomes weaker in these materials due to the strong subsur-  
348 face scattering. Under uncalibrated light setting, PF14 [43]  
349 shows a noticeable performance degradation when surface  
350 reflectance crosses certain levels of roughness and trans-  
351 lucency.

352 **Learning-based methods** CNN-PS [21] has a similar per-  
353 formance profile with TH46 [48], which are both per-pixel-  
354 based methods, where the MAE increases significantly in  
355 high translucency levels. It may be due to that these methods  
356 do not consider neighboring pixels, which could be critical  
357 when dealing with significant subsurface scattering. The all-  
358 pixel-based method Att-PSN [30] shows better accuracy than  
359 PS-FCN [6] by addressing blurry normal estimates. This  
360 is particularly beneficial for photometric stereo applied to  
361 translucent surfaces, where subsurface scattering can blur  
362 the surface normal estimates [26].

363 MS-PS [16] uses a multi-scale network structure to learn

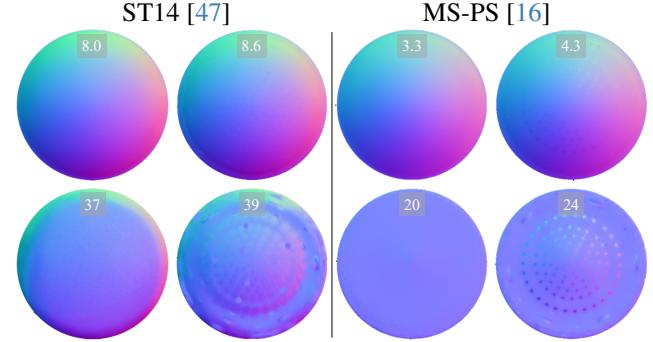


Figure 6. Visualization of estimated surface normals for hemisphere objects at the four corners of the translucency-roughness shown in Fig. 1. The MAEs in degrees are labeled in each sub-figure.

364 the global context of images, which contributes to model-  
365 ing the light transport in subsurface scattering. Using a  
366 large-scale training dataset containing diverse materials, MS-  
367 PS [16] shows robustness across varying translucency levels.  
368 Under the uncalibrated light setting, SDPS [7] shows stabler  
369 accuracy than the non-learning-based method PF14 [43] and  
370 yields smaller average MAEs. SDM-UniPS [24] achieves  
371 the smallest average MAE compared to all other methods  
372 even without knowing the light direction. The average MAE  
373 is threefold smaller than the one achieved by the second-best  
374 method MS-PS [16].

#### 4.2. Analysis of estimated normal map

375 Figure 6 shows estimated surface normal maps using  
376 ST14 [47] and MS-PS [16] for surfaces at the four corners  
377

Table 2. Roughness and translucency profiles of DiLiGenRT-S

| Object                             | PAW   | LEAF  | FLOWER |
|------------------------------------|-------|-------|--------|
| Roughness ( $S_a$ )                | 1.845 | 0.698 | 1.189  |
| Absorptive capacity ( $\sigma_t$ ) | 0.974 | 0.407 | 0.272  |

of the translucency-roughness matrix shown in Fig. 1. It is observed that the estimated shapes become flatter with increasing the translucency level. This is because subsurface scattering makes the observed radiance at different surface regions more uniform. In addition, with a lower roughness level, the estimated shapes exhibit high-frequency spikes due to specular highlights.

The methods ST14 [47] and MS-PS [16] effectively handle specular highlights on opaque surfaces, as shown by the slight increase in MAE between the even and odd columns in the first row of Fig. 6. However, when it comes to more translucent surfaces, the impact of specular highlights grows more significantly. As such, a photometric stereo method’s proficiency in handling specular highlights on opaque surfaces does not necessarily translate to its performance on translucent surfaces.

### 4.3. Evaluation on diverse shapes by DiLiGenRT-S

In addition to sphere shape contained in DiLiGenRT, we also build DiLiGenRT-S to include three different shapes: PAW, LEAF, and FLOWER, as shown in Fig. 1 (right). Their translucency and roughness profiles provided in Table 2 are measured in the same way discussed in Sec. 3.2. The evaluation on DiLiGenRT-S is summarized in Table 3. It shows a consistent trend that the normal estimation error increases with a greater translucency level. The MAEs are generally higher compared to DiLiGenRT, due to cast-shadows and inter-reflections due to the complex shapes. The MAEs of SDM-UniPS [24] have a significant increase compared to other methods. It is perhaps due to that SDM-UniPS [24] is particularly good at sphere-like shapes because the pre-trained model was likely overfitting to such a shape in their synthetic training dataset PS-Wild [23] and PS-Mix [24]. Therefore, we will exclude SDM-UniPS [24] on our subsequent evaluations on DiLiGenRT hereafter due to the possible overfitting.

### 4.4. Roughness-translucency error over #lights

The accuracy of surface normal estimation is influenced by surface reflectance, the selected photometric stereo method, and the number of varying lights. Figure 7 offers evaluation results under sparse (#10) and dense (#100) light sets uniformly distributed over a hemisphere. Each cell within the roughness-translucency error table displays the best-performing photometric stereo method and its MAE for a given material under a specific number of lights.

Table 3. Evaluation on DiLiGenRT-S by MAE in degrees, where the lowest MAE for each object is labeled in bold. In reference to Fig. 7, the MAE of the top-performing method at the corresponding  $S_a$  and  $\sigma_t$  is underlined, while the second-best is denoted with an asterisk (\*).

| Object | LS [53]      | TH28 [48]    | TH46 [48]    | DeconvPS [26]  | ST14 [47] | CNN-PS [21] |
|--------|--------------|--------------|--------------|----------------|-----------|-------------|
| PAW    | 17.9         | 17.2         | <u>16.5*</u> | 17.9           | 17.7      | <u>17.3</u> |
| LEAF   | 21.4         | 23.4         | 22.3         | 21.1           | 21.3      | 19.6        |
| FLOWER | 27           | 29.2         | 29           | 26.6           | 27        | 23.7*       |
| Object | PS-FCN [6]   | Att-PSN [30] | MS-PS [16]   | SDM-UniPS [24] | PF14 [43] | SDPS [7]    |
| PAW    | 16.7         | 17           | 17           | 18.6           | 20.2      | 18.6        |
| LEAF   | <u>17.9*</u> | 20.4         | <u>18.1</u>  | 29.5           | 27.3      | 19.4        |
| FLOWER | 22.6         | 25.2         | <u>20.7</u>  | 23.2           | 29.1      | 22.6        |

As shown in Fig. 7, with 100 lights, per-pixel-based methods like TH46 [48] and CNN-PS [21] yield superior results for semi-translucent materials (*i.e.*,  $\sigma_t$  between 0.57 and 1.24). Conversely, all-pixel-based methods such as Att-PSN [30] and MS-PS [16] excel in handling opaque and highly translucent materials (*i.e.*,  $\sigma_t$  above 1.24 or below 0.57), with MS-PS [16] particularly proficient at handling surfaces with high translucency. This trend remains consistent across varying roughness levels.

With a limited number of lights, *e.g.*, 10, all-pixel-based methods (Att-PSN [30] and MS-PS [16]) show superior accuracy than per-pixel-based ones (CNN-PS [21] and TH46 [48]) for 49 of 54 materials. One potential reason is that per-pixel-based methods predict surface normal only from the image measurements at a single pixel position, ignoring the spatial context provided by neighboring pixels, while all-pixel-based methods take them into account. This difference becomes significant under a sparse light setting, where image observations at a single pixel position are insufficient for per-pixel-based methods to accurately predict surface normals under general reflectance, while the broader observations from neighboring pixels benefit all-pixel-based methods for producing better normal estimates. According to the MAE values summarized in Fig. 7, for opaque and semi-translucent surfaces, 10 light sources suffice for photometric stereo given a maximum MAE tolerance of 10°. For evaluations under 20 and 50 lights, please refer to our supplementary material.

**Method selection based on DiLiGenRT** Figure 7 also presents an overview of the minimum MAE bounds for photometric stereo methods and shows their working ranges. It serves as a guide to selecting an appropriate photometric stereo method based on the reflectance and light setting. For example, the closest cell to the LEAF object’s roughness and absorptive capacity values shown in Table 2 is  $(S_a, \sigma_t) = (0.42, 0.28)$  in DiLiGenRT. The best and second-best methods indexed by these values in Fig. 7 are MS-PS [16] and PS-FCN [6]. The MAEs of these methods shown in Table 3 closely match the smallest MAE observed for the Leaf object. Similar observations are also confirmed

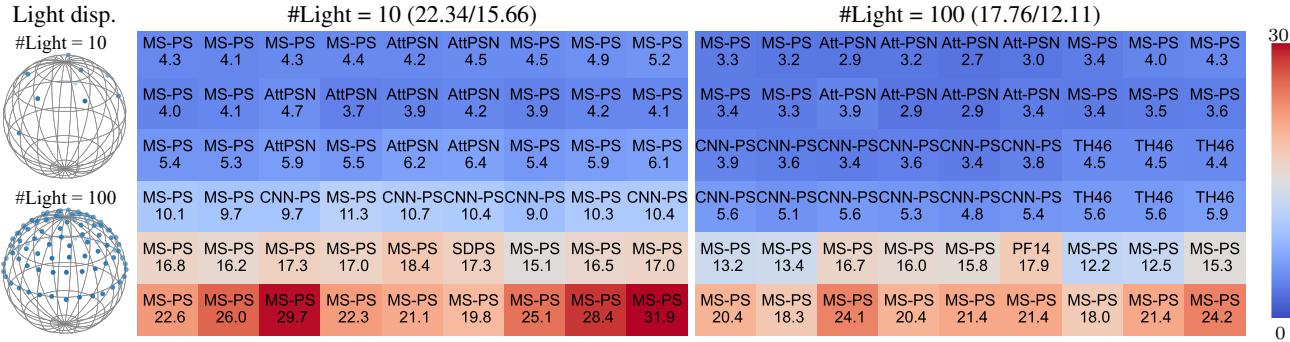


Figure 7. Benchmark evaluation on DiLiGenRT under sparse and dense lights (#10 and #100), summarized by mean/median MAE values. Each cell records the best-performing method for the material along with its MAE.

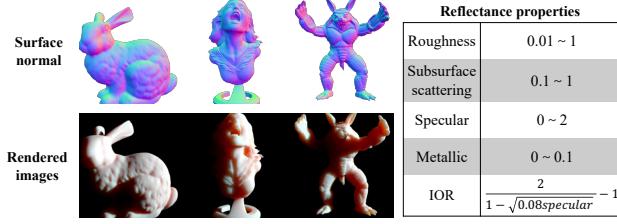


Figure 8. Selected shapes and images of synthetic dataset PS-SSS.

on the other two objects of DiLiGenRT-S. In this way, the quantified reflectances of our DiLiGenRT are useful in recommending the most suitable photometric stereo method given a target object.

#### 4.5. Fine-tuning with a synthetic translucent dataset

The benchmark evaluation illustrates the challenges in accurately recovering surface normal for highly translucent surfaces due to severe subsurface scattering effects. One potential reason is that the training dataset used in existing learning-based methods does not include translucent materials. To address the problem, we propose to boost the performance of existing methods through fine-tuning using a synthetic translucent dataset.

As shown in Fig. 8, we render a synthetic dataset specifically with Sub-Surface Scattering, named PS-SSS. To create PS-SSS, we selected 30 shapes and randomly rotated them 100 times, leading to 3,000 distinct surface normals. Similar to CyclePS [21], we use Blender [10] as our rendering engine and the Disney Principled BSDF [4] as the reflectance model. For each scene, we generate 100 uniformly-distributed light directions [37] and create a unique Principled BSDF whose reflectance properties are randomly assigned within a setting range, as shown in Fig. 8 (right), where the subsurface scattering parameter is adjusted between 0.1 and 1 to control translucency. In total, we generate 3,000 scenes covering distinct translucent surfaces. Please find the complete shape list of PS-SSS in our supplementary material.

We test incorporating data prior from PS-SSS into

existing learning-based methods. As an example, we use PS-FCN [6] as our baseline and fine-tune the normal regression module using our PS-SSS dataset over 9 epochs, where the learning rate is set to 1e-4. As shown in Fig. 9, the mean MAE decreases to 9.92°, which is much smaller compared with 16.38° in the original version presented in Fig. 5. This error reduction is particularly evident in surfaces with high translucency levels ( $\sigma_t \leq 0.57$ ), showing the effectiveness of incorporating data prior for handling subsurface scattering.

## 5. Discussion

**Open problems** Our evaluation on DiLiGenRT shows that existing photometric stereo methods can effectively handle varying levels of roughness. However, highly translucent surfaces remain challenging for most methods, where the MAEs generally exceed 20°. Besides leveraging data prior from the synthetic dataset as discussed in Sec. 4.5, explicitly modeling the subsurface scattering and using differentiable rendering [58] could be a promising approach. Moreover, DiLiGenRT is designed to be a method selector based on measured reflectance properties. Algorithms for selecting suitable photometric stereo methods based on input images remain unexplored.

**Limitations** This paper focuses on assessing photometric stereo under quantified reflectance properties. Besides the reflectance, illumination influences the performance of photometric stereo in real-world scenarios. Therefore, extending the distant directional light setup of DiLiGenRT to diverse illumination conditions including environment lights and near-point lights is our future direction.

527 **References**

- 528 [1] Neil G. Alldrin, Satya P. Mallick, and David J. Kriegman.  
529 Resolving the generalized bas-relief ambiguity by entropy  
530 minimization. In *Proc. of IEEE Conference on Computer  
531 Vision and Pattern Recognition (CVPR)*, 2007. 3
- 532 [2] Neil G. Alldrin, Todd Zickler, and David J. Kriegman. Photometric  
533 stereo with non-parametric and spatially-varying  
534 reflectance. In *Proc. of IEEE Conference on Computer Vision  
535 and Pattern Recognition (CVPR)*, 2008. 3
- 536 [3] Peter N. Belhumeur, David J. Kriegman, and Alan L. Yuille.  
537 The bas-relief ambiguity. *International Journal of Computer  
538 Vision*, 1999. 2
- 539 [4] Brent Burley and Walt Disney Animation Studios. Physically-  
540 based shading at disney. In *Proc. of SIGGRAPH*, pages 1–7.  
541 vol. 2012, 2012. 3, 8
- 542 [5] Mammohan Chandraker, Jiamin Bai, and Ravi Ramamoorthi.  
543 On differential photometric reconstruction for unknown,  
544 isotropic brdfs. *IEEE Transactions on Pattern Analysis and  
545 Machine Intelligence*, 35(12):2941–2955, 2012. 2
- 546 [6] Guanying Chen, Kai Han, and Kwan-Yee K. Wong. PS-FCN:  
547 A flexible learning framework for photometric stereo. In *Proc.  
548 of European Conference on Computer Vision (ECCV)*, 2018.  
549 3, 5, 6, 7, 8
- 550 [7] Guanying Chen, Kai Han, Boxin Shi, Yasuyuki Matsushita,  
551 and Kwan-Yee K. Wong. Self-calibrating deep photometric  
552 stereo networks. In *Proc. of IEEE Conference on Computer  
553 Vision and Pattern Recognition (CVPR)*, 2019. 3, 5, 6, 7
- 554 [8] Guanying Chen, Michael Waechter, Boxin Shi, Kwan-Yee K  
555 Wong, and Yasuyuki Matsushita. What is learned in deep un-  
556 calibrated photometric stereo? In *Proc. of European Confer-  
557 ence on Computer Vision (ECCV)*, pages 745–762. Springer,  
558 2020. 3
- 559 [9] Lixiong Chen, Yinqiang Zheng, Boxin Shi, Art Subpa-Asa,  
560 and Imari Sato. A microfacet-based model for photometric  
561 stereo with general isotropic reflectance. *IEEE Transactions  
562 on Pattern Analysis and Machine Intelligence*, 43(1):48–61,  
563 2019. 2
- 564 [10] Blender Online Community. *Blender - a 3D modelling and  
565 rendering package*. Blender Foundation, Blender Foundation,  
566 Amsterdam, 2021. 3, 8
- 567 [11] Bo Dong, Kathleen D Moore, Weiyi Zhang, and Pieter Peers.  
568 Scattering parameters and surface normals from homoge-  
569 neous translucent materials using photometric stereo. In *Proc.  
570 of IEEE Conference on Computer Vision and Pattern Recog-  
571 nition (CVPR)*, 2014. 3
- 572 [12] Kenji Enomoto, Michael Waechter, Kiriakos N Kutulakos,  
573 and Yasuyuki Matsushita. Photometric stereo via discrete  
574 hypothesis-and-test search. In *Proceedings of the IEEE/CVF  
575 Conference on Computer Vision and Pattern Recognition*,  
576 pages 2311–2319, 2020. 2
- 577 [13] Kenji Enomoto, Michael Waechter, Fumio Okura, Kiriakos N  
578 Kutulakos, and Yasuyuki Matsushita. Discrete search photo-  
579 metric stereo for fast and accurate shape estimation. *IEEE  
580 Transactions on Pattern Analysis and Machine Intelligence*,  
581 2022. 2
- 582 [14] Dan B Goldman, Brian Curless, Aaron Hertzmann, and  
583 Steven M Seitz. Shape and spatially-varying brdfs from pho-  
584 tometric stereo. *IEEE Transactions on Pattern Analysis and  
585 Machine Intelligence*, 2009. 2
- 586 [15] Bjoern Haefner, Songyou Peng, Alok Verma, Yvain Qu  au,  
587 and Daniel Cremers. Photometric depth super-resolution.  
588 *IEEE Transactions on Pattern Analysis and Machine Intelli-  
589 gence*, 2019. 3
- 590 [16] Cl  ment Hardy, Yvain Qu  au, and David Tschumperl  .  
591 MS-PS: A multi-scale network for photometric stereo with  
592 a new comprehensive training dataset. *arXiv preprint  
593 arXiv:2211.14118*, 2022. 3, 5, 6, 7
- 594 [17] Hideki Hayakawa. Photometric stereo under a light source  
595 with arbitrary motion. *JOSA A*, 1994. 2
- 596 [18] Yannick Hold-Geoffroy, Paulo Gotardo, and Jean-Fran  ois  
597 Lalonde. Single day outdoor photometric stereo. *IEEE Trans-  
598 actions on Pattern Analysis and Machine Intelligence*, 2019.  
599 3
- 600 [19] Berthold K.P Horn and Michael J Brooks. The variational  
601 approach to shape from shading. *Computer Vision, Graphics,  
602 and Image Processing*, 33(2):174–208, 1986. 2
- 603 [20] Zhuo Hui and Aswin C Sankaranarayanan. Shape and  
604 spatially-varying reflectance estimation from virtual exem-  
605 plars. *IEEE Transactions on Pattern Analysis and Machine  
606 Intelligence*, 39(10):2060–2073, 2016. 2
- 607 [21] Satoshi Ikehata. CNN-PS: CNN-based photometric stereo for  
608 general non-convex surfaces. In *Proc. of European Confer-  
609 ence on Computer Vision (ECCV)*, 2018. 3, 5, 6, 7, 8
- 610 [22] Satoshi Ikehata. PS-Transformer: Learning sparse photomet-  
611 ric stereo network using self-attention mechanism. *Proc. of  
612 the British Machine Vision Conference (BMVC)*, 2022. 3
- 613 [23] Satoshi Ikehata. Universal photometric stereo network using  
614 global lighting contexts. In *Proc. of IEEE Conference on  
615 Computer Vision and Pattern Recognition (CVPR)*, pages  
616 12591–12600, 2022. 3, 5, 7
- 617 [24] Satoshi Ikehata. Scalable, detailed and mask-free universal  
618 photometric stereo. *arXiv preprint arXiv:2303.15724*, 2023.  
619 3, 5, 6, 7
- 620 [25] Satoshi Ikehata, David Wipf, Yasuyuki Matsushita, and Kiy-  
621 oharu Aizawa. Robust photometric stereo using sparse regres-  
622 sion. In *Proc. of IEEE Conference on Computer Vision and  
623 Pattern Recognition (CVPR)*, 2012. 2
- 624 [26] Chika Inoshita, Yasuhiro Mukaigawa, Yasuyuki Matsushita,  
625 and Yasushi Yagi. Surface normal deconvolution: Photomet-  
626 ric stereo for optically thick translucent objects. In *Proc. of  
627 European Conference on Computer Vision (ECCV)*, pages  
628 346–359, 2014. 3, 5, 6, 7
- 629 [27] Wenzel Jakob. Mitsuba renderer, 2010. 3
- 630 [28] Henrik Wann Jensen, Stephen R Marschner, Marc Levoy,  
631 and Pat Hanrahan. A practical model for subsurface light  
632 transport. In *Proceedings of Annual Conference on Computer  
633 Graphics and Interactive Techniques*, pages 511–518, 2001.  
634 3
- 635 [29] Micah K Johnson and Edward H Adelson. Shape estimation  
636 in natural illumination. In *Proc. of IEEE Conference on  
637 Computer Vision and Pattern Recognition (CVPR)*, 2011. 3
- 638 [30] Yakun Ju, Boxin Shi, Muwei Jian, Lin Qi, Junyu Dong, and  
639 Kin-Man Lam. Normattention-PSN: A high-frequency re-  
640 gion enhanced photometric stereo network with normalized

- 641 attention. *International Journal of Computer Vision*, 130(12):  
642 3014–3034, 2022. 3, 5, 6, 7
- 643 [31] Berk Kaya, Suryansh Kumar, Carlos Oliveira, Vittorio Ferrari,  
644 and Luc Van Gool. Uncalibrated neural inverse rendering for  
645 photometric stereo of general surfaces. In *Proc. of IEEE Conference  
646 on Computer Vision and Pattern Recognition (CVPR)*, 2021. 3
- 647 [32] Junxuan Li and Hongdong Li. Self-calibrating photometric  
648 stereo by neural inverse rendering. In *Proc. of European  
649 Conference on Computer Vision (ECCV)*, pages 166–183.  
650 Springer, 2022. 3
- 651 [33] Junxuan Li, Antonio Robles-Kelly, Shaodi You, and Yasuyuki  
652 Matsushita. Learning to minify photometric stereo. In *Proc.  
653 of IEEE Conference on Computer Vision and Pattern Recog-  
654 nition (CVPR)*, 2019. 3
- 655 [34] Min Li, Zhenglong Zhou, Zhe Wu, Boxin Shi, Changyu Diao,  
656 and Ping Tan. Multi-view photometric stereo: A robust so-  
657 lution and benchmark dataset for spatially varying isotropic  
658 materials. *IEEE Transactions on Image Processing*, 29:4159–  
659 4173, 2020. 1, 3
- 660 [35] Fotios Logothetis, Ignas Budvytis, Roberto Mecca, and  
661 Roberto Cipolla. PX-NET: Simple and efficient pixel-wise  
662 training of photometric stereo networks. In *Proc. of Interna-  
663 tional Conference on Computer Vision (ICCV)*, pages 12757–  
664 12766, 2021. 3, 5
- 665 [36] Fotios Logothetis, Roberto Mecca, Ignas Budvytis, and  
666 Roberto Cipolla. A CNN based approach for the point-light  
667 photometric stereo problem. *International Journal of Com-  
668 puter Vision*, pages 1–20, 2022. 1, 3
- 669 [37] George Marsaglia. Choosing a point from the surface of a  
670 sphere. *The Annals of Mathematical Statistics*, 43(2):645–  
671 646, 1972. 8
- 672 [38] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and  
673 Leonard McMillan. A data-driven reflectance model. In  
674 *Proc. of SIGGRAPH*, 2003. 3
- 675 [39] Thomas G Mayerhöfer, Susanne Pahlow, and Jürgen Popp.  
676 The bouguer-beer-lambert law: Shining light on the obscure.  
677 *ChemPhysChem*, 21(18):2029–2046, 2020. 5
- 678 [40] Roberto Mecca, Fotios Logothetis, Ignas Budvytis, and  
679 Roberto Cipolla. LUCES: A dataset for near-field point light  
680 source photometric stereo. In *Proc. of the British Machine  
681 Vision Conference (BMVC)*, 2021. 1, 3
- 682 [41] Kazuma Minami, Hiroaki Santo, Fumio Okura, and Ya-  
683 suyuki Matsushita. Symmetric-light photometric stereo. In  
684 *IEEE/CVF Winter Conference on Applications of Computer  
685 Vision (WACV)*, pages 2706–2714, 2022. 2
- 686 [42] Thoma Papadimitri and Paolo Favaro. A new perspective on  
687 uncalibrated photometric stereo. In *Proc. of IEEE Conference  
688 on Computer Vision and Pattern Recognition (CVPR)*, pages  
689 1474–1481, 2013. 2
- 690 [43] Thoma Papadimitri and Paolo Favaro. A closed-form, con-  
691 sistent and robust solution to uncalibrated photometric stereo  
692 via local diffuse reflectance maxima. *International Journal  
693 of Computer Vision*, 2014. 2, 5, 6, 7
- 694 [44] Jieji Ren, Feishi Wang, Jiahao Zhang, Qian Zheng, Mingjun  
695 Ren, and Boxin Shi. DiLiGenT10<sup>2</sup>: A photometric stereo  
696 benchmark dataset with controlled shape and material varia-  
697 tion. In *Proc. of IEEE Conference on Computer Vision and  
698 Pattern Recognition (CVPR)*, 2022. 1, 3, 4, 5
- 699 [45] Hiroaki Santo, Masaki Samejima, Yusuke Sugano, Boxin Shi,  
700 and Yasuyuki Matsushita. Deep photometric stereo network.  
701 In *Proc. of International Conference on Computer Vision  
702 Workshops (ICCVW)*, 2017. 2, 3
- 703 [46] Boxin Shi, Yasuyuki Matsushita, Yichen Wei, Chao Xu, and  
704 Ping Tan. Self-calibrating photometric stereo. In *Proc. of  
705 IEEE Conference on Computer Vision and Pattern Recog-  
706 nition (CVPR)*, 2010. 2
- 707 [47] Boxin Shi, Ping Tan, Yasuyuki Matsushita, and Katsushi  
708 Ikeuchi. Bi-polynomial modeling of low-frequency re-  
709 flectances. *IEEE Transactions on Pattern Analysis and Ma-  
710 chine Intelligence*, 2014. 2, 5, 6, 7
- 711 [48] Boxin Shi, Zhipeng Mo, Zhe Wu, Dinglong Duan, Sai-Kit  
712 Yeung, and Ping Tan. A benchmark dataset and evaluation for  
713 non-Lambertian and uncalibrated photometric stereo. *IEEE  
714 Transactions on Pattern Analysis and Machine Intelligence*,  
715 2019. 1, 2, 3, 5, 6, 7
- 716 [49] William M. Silver. *Determining shape and reflectance us-  
717 ing multiple images*. PhD thesis, Massachusetts Institute of  
718 Technology, 1980. 2
- 719 [50] Ashish Tiwari and Shanmuganathan Raman. DeepPS2: Re-  
720 visiting photometric stereo using two differently illuminated  
721 images. In *Proc. of European Conference on Computer Vision  
722 (ECCV)*, pages 129–145. Springer, 2022. 3, 5
- 723 [51] Ashish Tiwari and Shanmuganathan Raman. LERPS: lighting  
724 estimation and relighting for photometric stereo. In *IEEE  
725 International Conference on Acoustics, Speech and Signal  
726 Processing (ICASSP)*, pages 2060–2064. IEEE, 2022. 3
- 727 [52] Feishi Wang, Jieji Ren, Heng Guo, Mingjun Ren, and Boxin  
728 Shi. DiLiGenT-Pi: Photometric stereo for planar surfaces  
729 with rich details-benchmark dataset and beyond. In *Proc. of  
730 International Conference on Computer Vision (ICCV)*, pages  
731 9477–9487, 2023. 1, 3, 4
- 732 [53] Robert J. Woodham. Photometric method for determining sur-  
733 face orientation from multiple images. *Optical engineering*,  
734 1980. 1, 2, 5, 6, 7
- 735 [54] Lun Wu, Arvind Ganesh, Boxin Shi, Yasuyuki Matsushita,  
736 Yongtian Wang, and Yi Ma. Robust photometric stereo via  
737 low-rank matrix completion and recovery. In *Proc. of Asian  
738 Conference on Computer Vision (ACCV)*, 2010. 2
- 739 [55] Ying Xiong, Ayan Chakrabarti, Ronen Basri, Steven J Gortler,  
740 David W Jacobs, and Todd Zickler. From shading to local  
741 shape. *IEEE Transactions on Pattern Analysis and Machine  
742 Intelligence*, 2014. 3
- 743 [56] Ying Xiong, Ayan Chakrabarti, Ronen Basri, Steven J. Gortler,  
744 David W. Jacobs, and Todd Zickler. From shading to local  
745 shape. *IEEE Transactions on Pattern Analysis and Machine  
746 Intelligence*, 2015. 3
- 747 [57] Zhuokun Yao, Kun Li, Ying Fu, Haofeng Hu, and Boxin Shi.  
748 GPS-Net: Graph-based photometric stereo network. *Proc. of  
749 Annual Conference on Neural Information Processing Systems  
750 (NeurIPS)*, 33:10306–10316, 2020. 3, 5
- 751 [58] Hong-Xing Yu, Michelle Guo, Alireza Fathi, Yen-Yu Chang,  
752 Eric Ryan Chan, Ruohan Gao, Thomas Funkhouser, and Jia-  
753 jun Wu. Learning object-centric neural scattering functions  
754

- 755 for free-viewpoint relighting and scene composition. *arXiv*  
756 preprint arXiv:2303.06138, 2023. 8
- 757 [59] Qian Zheng, Yiming Jia, Boxin Shi, Xudong Jiang, Ling-Yu  
758 Duan, and Alex C. Kot. SPLINE-Net: Sparse photometric  
759 stereo through lighting interpolation and normal estimation  
760 networks. In *Proc. of International Conference on Computer  
761 Vision (ICCV)*, 2019. 3, 5
- 762 [60] Zhenglong Zhou and Ping Tan. Ring-light photometric stereo.  
763 In *Proc. of European Conference on Computer Vision (ECCV)*,  
764 2010. 2