Non-Lambertian Multispectral Photometric Stereo via Spectral Reflectance Decomposition

Jipeng Lv[1] , **Heng Guo**[2,3*] , **Guanying Chen**[4] , **Jinxiu Liang**[1] , **Boxin Shi**[1]
[1]Peking University
[2]Beijing University of Posts and Telecommunications
[3]Osaka University
[4]The Chinese University of Hong Kong, Shenzhen
heng.guo@ist.osaka-u.ac.jp, chenguanying@cuhk.edu.cn, {lvjipeng, cssherryliang, shiboxin}@pku.edu.cn

Abstract

Multispectral photometric stereo (MPS) aims at recovering the surface normal of a scene from a single-shot multispectral image captured under multispectral illuminations. Existing MPS methods adopt the Lambertian reflectance model to make the problem tractable, but it greatly limits their application to real-world surfaces. In this paper, we propose a deep neural network named NeuralMPS to solve the MPS problem under non-Lambertian spectral reflectances. Specifically, we present a spectral reflectance decomposition model to disentangle the spectral reflectance into a geometric component and a spectral component. With this decomposition, we show that the MPS problem for surfaces with a uniform material is equivalent to the conventional photometric stereo (CPS) with unknown light intensities. In this way, NeuralMPS reduces the difficulty of the non-Lambertian MPS problem by leveraging the well-studied non-Lambertian CPS methods. Experiments on both synthetic and real-world scenes demonstrate the effectiveness of our method.

1 Introduction

Photometric stereo methods, originally proposed by Woodham [Woodham, 1980] and Silver [Silver, 1980], recover detailed three-dimensional (3D) surfaces from images captured from a fixed camera under varying lighting directions, which are commonly obtained at different timestamps. Since conventional photometric stereo (CPS) methods stack images with *time-multiplexing*, the target surface has to be kept static during the multiple shots. With *spectral-multiplexing*, multispectral photometric stereo (MPS) [Kontsevich *et al.*, 1994] can recover surface normals from a one-shot multispectral image. As shown in Fig. 1, a single multispectral image for MPS encodes observations under varying lighting directions in different spectral bands, conveying the information about surface normals at that timestamp. With input multispectral images recorded by a multispectral camera under spectral light sources, MPS can reconstruct the 3D shapes of dynamic objects[1].

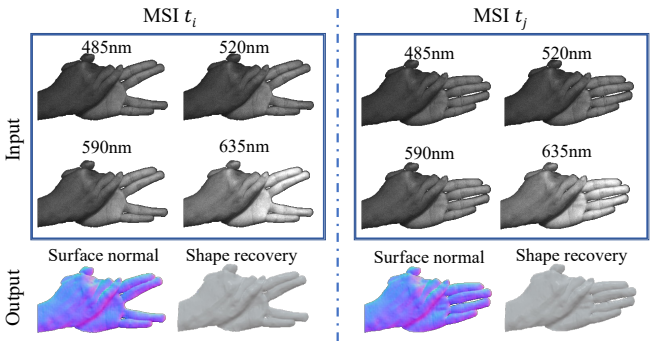

Figure 1: Our NeuralMPS takes multispectral image (MSI) as input, and recovers the dynamic shape (surface normal) of a non-Lambertian surface at different timestamps.

Existing MPS methods [Guo *et al.*, 2021; Chakrabarti and Sunkavalli, 2016; Ozawa *et al.*, 2018; Anderson *et al.*, 2011b] mostly focus on the Lambertian reflectance, which has limited ability to represent the surface appearance in real-world. For non-Lambertian surfaces, the spectral reflectance varies not only with the spectral band, but also the incident-outgoing lighting directions w.r.t. the surface normal at different scene points. There have been a few MPS methods working for non-Lambertian reflectance, which assumed a specific hardware setup [Rahman *et al.*, 2014] with two-shot capturing or an input image containing three spectral channels with fixed spectral bands [Ju *et al.*, 2018; Ju *et al.*, 2020a; Ju *et al.*, 2020b]. Due to the restrictive capture setting and limited surface normal estimation accuracy of existing methods, non-Lambertian MPS remains an open and challenging problem. Alternatively, non-Lambertian reflectance is well-studied in the CPS context and great progress has been achieved with data-driven approaches [Chen *et al.*, 2018; Santo *et al.*, 2017; Yao *et al.*, 2020; Ikehata, 2018]. However, non-Lambertian CPS methods cannot be directly applied to multispectral image observations without considering the wavelength-dependent reflectance responses.

* Corresponding Author

[1]Please check the supplementary video for the full sequence.

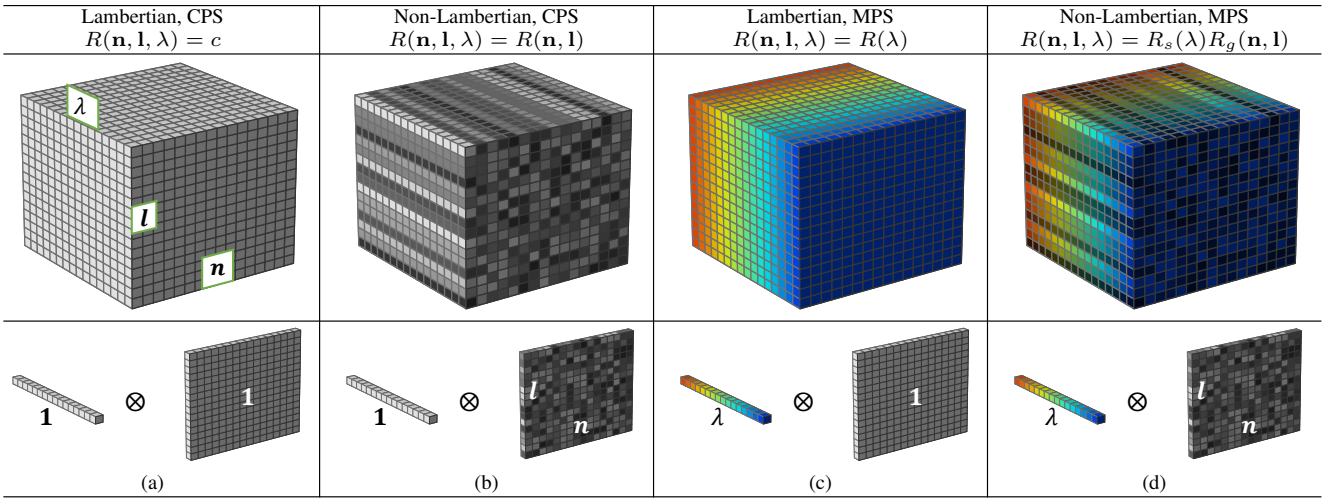

Figure 2: Visualization of the 4 categories of reflectance assumptions in existing photometric stereo methods, where $\otimes$ represents Kronecker product and $\mathbf{1}$ denotes one vector/matrix whose elements are all equal to 1. The reflectance decomposition in different photometric stereo tasks is shown in the third row. The last column illustrates the proposed SRD model.

In this paper, we show that MPS can be reformulated as a CPS problem by introducing a *spectral reflectance decomposition (SRD) model*, which allows us to borrow experiences from well-established CPS theory. Specifically, we assume the target surface has dominant diffuse or speculative reflectance, which is commonly seen in real-world materials (*e.g.* measured material "paper_blue" and "ilm_l3_37_metallic" in Fig. 3). Under this assumption, our SRD model can decompose the non-Lambertian spectral reflectance into two independent components: *geometric* and *spectral*. The geometric component only varies with the incident-outgoing lighting directions, which can be well-modeled by non-Lambertian CPS methods. The spectral component is the response w.r.t. the spectral wavelength. For surfaces covered by a uniform material, we can entangle it with the input light intensity as an *equivalent light intensity*. In this way, the MPS problem can be formulated as an semi-calibrated CPS problem under unknown equivalent light intensities. Following the above insights, we propose a neural network *NeuralMPS* to first predict the equivalent light intensity including the spectral component, and then embed existing CPS methods to recover the surface normal map considering the geometric component only.

To summarize, our contributions are as follows:

- We formulate the MPS problem as a well-studied CPS problem with unknown equivalent light intensity by introducing an SRD model.
- We propose a learning-based MPS network *NeuralMPS* to accurately recover surface normal under diffuse- or specular- dominant non-Lambertian reflectance.

2 Related Works

The surface general reflectance can be modeled by the Bidirectional Reflectance Distribution Function (BRDF) [Szeliski, 2010], which describes how much light at each wavelength arriving at an incident direction is emitted in a reflected direction. With a fixed view direction, the isotropic spectral BRDF is a function $R(\mathbf{n}, \mathbf{l}, \lambda)$ of the surface normal $\mathbf{n}$, the lighting direction $\mathbf{l}$ and the wavelength λ. Therefore, the BRDF response w.r.t. the three variables can be recorded as a cube, as shown in Fig. 2. If we apply the BRDF response to a surface covered by uniform material and ignore the crosstalk, the surface spectral reflectance can be recorded as the same cube. In this section, we review CPS and MPS approaches under Lambertian reflectance and non-Lambertian reflectance, respectively. These four categories are based on different simplifications on the spectral BRDF cube. Among all the categories, we assume distant lights with calibrated directions.

Lambertian, CPS CPS under Lambertian reflectance assumption is the most classic setup. In such case, the spectral reflectance under varying lighting directions and wavelengths is a constant value, *i.e.*, $R(\mathbf{n}, \mathbf{l}, \lambda) = c$ (see Fig. 2 (a)). Without loss of generality, we set $c = 1$. Given image measurements under more than 3 non-coplanar lighting directions, classical photometric stereo works [Woodham, 1980; Silver, 1980] provided a closed-form solution to surface normal estimation.

Non-Lambertian, CPS The spectral reflectance for non-Lambertian CPS is varying with the incident-outgoing lighting directions (geometric component), but omits the variation w.r.t the wavelength (spectral component). Therefore, the spectral BRDF $R(\mathbf{n}, \mathbf{l}, \lambda) = R(\mathbf{n}, \mathbf{l})$ and the spectral cube shown in Fig. 2 (b) is a Kronecker product of an all-in-one spectral response vector and the geometric response matrix. Existing non-Lambertian CPS methods applied parametric-based [Chen *et al.*, 2019b; Shi *et al.*, 2014] or dictionary-based reflectance model [Enomoto *et al.*, 2020; Hui and Sankaranarayanan, 2016] to represent the geometric component $R(\mathbf{n}, \mathbf{l})$. Recently, neural-based photo-

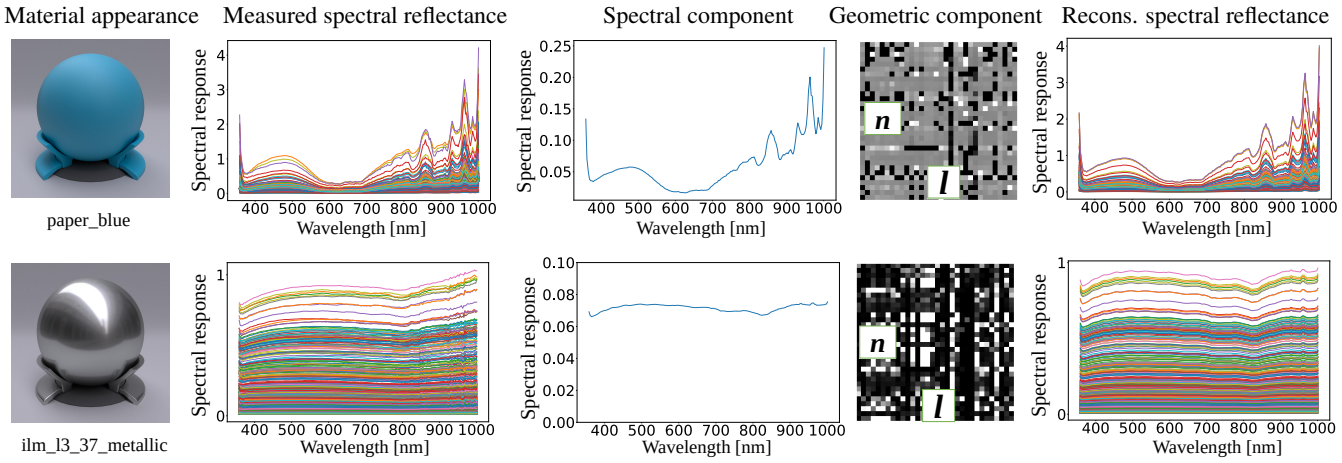

Figure 3: Illustration of SRD on two measured BRDFs "paper_blue" (diffuse dominant) and "ilm_l3_37_metallic" (specular dominant) [Dupuy and Jakob, 2018].

metric stereo methods [Chen *et al.*, 2018; Ikehata, 2018; Yao *et al.*, 2020] achieved high accuracy on surface normal estimation under non-Lambertian reflectance, where the geometric component is learned from data with diverse reflectances. Please refer to the two surveys [Shi *et al.*, 2019; Ju *et al.*, 2022] for comprehensive reviews of non-Lambertian CPS methods.

Lambertian, MPS In MPS, the spectral component has to be taken into account. The Lambertian spectral BRDF keeps constant w.r.t. geometric component but varies over different wavelengths, *i.e.*, $R(\mathbf{n}, \mathbf{l}, \lambda) = R(\lambda)$. Therefore, the spectral cube shown in Fig. 2 (c) is a Kronecker product of a spectral response vector and an all-in-one geometric response matrix. Different from Lambertian CPS, MPS is ill-posed even under Lambertian reflectance. Existing methods require additional shape priors [Anderson *et al.*, 2011b; Anderson *et al.*, 2011a] or reflectance clustering [Ozawa *et al.*, 2018; Chakrabarti and Sunkavalli, 2016] to solve the problem. Recently, Guo *et al.* [Guo *et al.*, 2021] formulated the Lambertian MPS problem into a well-posed one and provided a unique solution for surface normals with image cues only.

Non-Lambertian, MPS Non-Lambertian MPS problem is the most challenging case among the four categories, as the spectral BRDF $R(\mathbf{n}, \mathbf{l}, \lambda)$ is related to both wavelength and incident-outgoing lighting directions. Therefore, the spectral BRDF is recorded as a general cube (see Fig. 2 (d)). Existing non-Lambertian MPS methods assumed analytical BRDF model and two-shot data capturing [Rahman *et al.*, 2014]. However, the analytical BRDF model is limited to a small set of materials, and the two-shot data capturing eliminates the advantage of MPS over CPS on dynamic shape recovery. Ju *et al.* [Ju *et al.*, 2020a; Ju *et al.*, 2020b] propose learning-based methods to solve non-Lambertian reflectance. However, the spectral band of the input multispectral images in the test phase is required to keep fixed in the training stage, which limits their application to real-world scenarios. Our method regards the spectral BRDF cube as a Kronecker product of the spectral response vector (spectral component) and geometric response matrix (geometric component). As we can see in the following sections, such a decomposition enables the application of off-the-shell non-Lambertian CPS methods on the non-Lambertian MPS task.

3 Spectral Reflectance Decomposition Model

In the context of MPS, the material spectral reflectance varies with both the incident-outgoing lighting directions and spectral bands. Under a fixed view direction, it can be formulated as a function of surface normal $\mathbf{n}$, lighting direction $\mathbf{l}$, and wavelength λ. Following the dichromatic model proposed in [Shafer, 1985], we represent the spectral reflectance as

$$R(\mathbf{n}, \mathbf{l}, \lambda) = R_{gd}(\mathbf{n}, \mathbf{l}) R_{sd}(\lambda) + R_{gs}(\mathbf{n}, \mathbf{l}) R_{ss}(\lambda), \quad (1)$$

where R_{g*} and R_{s*} represent the geometric component and the spectral component, the subscript d and s denote diffuse and specular reflectance respectively. We assume the non-Lambertian target surfaces are with dominant diffuse (*i.e.*, $R(\mathbf{n}, \mathbf{l}, \lambda) = R_{gd}(\mathbf{n}, \mathbf{l}) R_{sd}(\lambda)$) or specular (*i.e.*, $R(\mathbf{n}, \mathbf{l}, \lambda) = R_{gs}(\mathbf{n}, \mathbf{l}) R_{ss}(\lambda)$) reflectance. In this way, we introduce our SRD model that disentangles the non-Lambertian spectral reflectance into only two components:

$$R(\mathbf{n}, \mathbf{l}, \lambda) = R_s(\lambda) R_g(\mathbf{n}, \mathbf{l}), \quad (2)$$

where the spectral component $R_s(\lambda)$ is related to the wavelength, and the geometric component $R_g(\mathbf{n}, \mathbf{l})$ is related to the variation w.r.t. surface normal and lighting direction.

We illustrate and verify our SRD model by measured spectral BRDFs [Dupuy and Jakob, 2018]. As shown in Fig. 3, we plot measured spectral reflectances of two materials: "paper blue" and "ilm_l3_37_metallic" in the $[360\text{nm}, 1000\text{nm}]$ wavelength range, where curves in different colors show the spectral reflectance at varying lighting direction and surface normal pairs: $(\mathbf{l}, \mathbf{n})$. If the surface follows Lambertian reflectance [Anderson *et al.*, 2011b; Esteban *et al.*, 2011], the spectral reflectance shown in the figure should contain one single curve, which is not flexible to represent the real-world spectral reflectances. To verify the effectiveness of

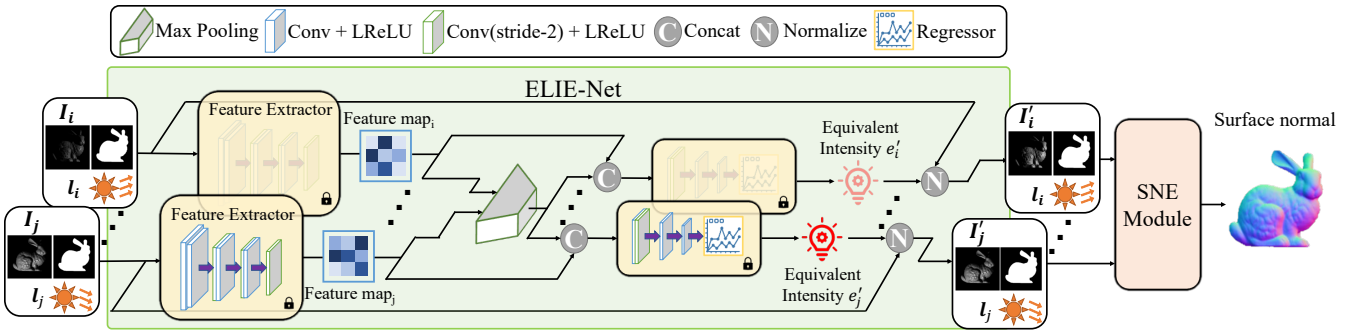

Figure 4: Network structure of our NeuralMPS. The ELIE-Net is fed with image observations under varying calibrated lighting directions and the corresponding masks for estimating equivalent light intensity, from which the images are normalized (denoted as $\mathbf{I}'$) and sent to the SNE Module for estimating the surface normal. The network parameters labeled with the "lock" symbol are shared for different image inputs.

our SRD model, we sample 195 wavelengths and 200×200 $(\mathbf{l}, \mathbf{n})$ pairs to build an spectral reflectance response matrix $\mathbf{R} \in \mathbb{R}^{195,40000}$. We achieve SRD by conducting SVD on $\mathbf{R}$ such that

$$\begin{aligned} \mathbf{R} &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top, \\ \mathbf{r}_s = \mathbf{U}_1, &\quad \mathbf{r}_g = \mathbf{V}_1, \end{aligned} \tag{3}$$

where $\mathbf{r}_s \in \mathbb{R}^{195}$ and $\mathbf{r}_g \in \mathbb{R}^{40000}$ are the decomposed spectral and geometric components in the vector form obtained from the first column of $\mathbf{U}$ and $\mathbf{V}$, respectively. We reshape the $\mathbf{r}_g$ into matrix form and draw the two components in Fig. 3. To measure the SRD accuracy, we define the fitting energy as the ratio of the largest singular value over the sum of all the singular values. The energy over all the 51 materials is 80.1% for the average, 81.6% for the median, and 96.1% for the maximum, which reveals that we can approximate the spectral reflectance in high accuracy with the decomposed $\mathbf{r}_s$ and $\mathbf{r}_g$. We also reconstruct the spectral reflectance response matrix $\mathbf{R}$ from the decomposed spectral and geometric parts. The reconstructed spectral reflectance curves shown in Fig. 3 are shown to be close to the originally measured one.

3.1 SRD-based Image Formation Model

In this section, we show that the MPS problem is equivalent to the CPS under unknown light intensity based on our proposed SRD model.

Given an orthogonal multispectral camera with a linear radiometric response and f spectral directional lights with calibrated lighting directions, we turn on all the spectral light sources and capture a single multispectral image. In this way, the f spectral bands of the multispectral image record measurements under varying lighting directions. Considering a non-Lambertian surface with general isotropic spectral BRDF, the image observation for one scene point under the j-th incoming lighting can be written as follows:

$$m_j = e_j \max(\mathbf{n}^\top \mathbf{l}_j, 0) \int_\lambda R(\mathbf{n}, \mathbf{l}_j, \lambda) C_j(\lambda) E_j(\lambda)\, d\lambda, \tag{4}$$

where $\mathbf{n} \in S^2 \subset \mathbb{R}^3$ represents the unit surface normal vector, $\mathbf{l}_j \in S^2 \subset \mathbb{R}^3$ and $e_j \in \mathbb{R}^+$ are lighting direction and radiance of the j-th light source, respectively, with its spectra defined by $E_j(\lambda) : \mathbb{R}^+ \to \mathbb{R}^+$. The camera spectral sensitivity at the j-th spectral band is represented by $C_j(\lambda)$, and the attach shadow is modeled by $\max(\cdot, 0)$.

Similar to previous MPS works [Guo *et al.*, 2021; Chakrabarti and Sunkavalli, 2016; Ozawa *et al.*, 2018], we assume the crosstalk between spectral bands is negligible, *i.e.*, observations under each spectral light can only be observed in its corresponding camera channel. With negligible crosstalk and the proposed SRD model shown in Eq. (2), the image observation can be re-written as

$$m_j = e'_j R_g(\mathbf{n}, \mathbf{l}_j) \max(\mathbf{n}^\top \mathbf{l}_j, 0), \tag{5}$$

and e'_j is defined by

$$e'_j = e_j \int_{\lambda \in \Omega_j} R_s(\lambda) C(\lambda) E(\lambda)\, d\lambda, \tag{6}$$

where Ω_j is the j-th spectral band of the corresponding spectral light and camera channel. The camera spectral sensitivity, light spectra, material spectral component, and the light intensity are encoded in e'_j. If the material spectral component is uniform for the whole surface (*e.g.* monochromatic material), e'_j keeps constant for all the surface points. Therefore, we name e'_j as the *equivalent light intensity*. According to the image formation model shown in Eq. (5), given the image measurement m and its corresponding lighting direction $\mathbf{l}$, the non-Lambertian MPS task can be reformulated as estimating the surface normal $\mathbf{n}$ under unknown geometric reflectance $R_g(\mathbf{n}, \mathbf{l})$ and the equivalent light intensity e'_j. In this way, *non-Lambertian MPS is equivalent to the non-Lambertian CPS problem with unknown light intensities*.

4 Learning-based Non-Lambertian MPS

Guided by the conclusion in the last section, we present the NeuralMPS to solve MPS under non-Lambertian reflectances. As shown in Fig. 4, our NeuralMPS contains two modules with functionalities indicated by their names: Equivalent Light Intensity Estimation Network (ELIE-Net) and Surface Normal Estimation (SNE) Module.

ELIE-Net ELIE-Net aims at predicting the equivalent light intensities from observed multispectral images and the calibrated lighting directions. As shown in Fig. 4, image observation at each spectral channel, the corresponding calibrated

spectral lighting direction, and the object mask are fed into a shared encoder separately for extracting the local feature maps. The local features of different spectral observations are then pooled into a global one by a max-pooling layer. From a shared decoder fed with concatenated local and global features, the equivalent light intensity for each image observation is regressed. To supervise the ELIE-Net, we minimize the ℓ_2 distance with the ground-truth equivalent lighting intensity extracted from the measured spectral BRDFs:

$$\mathcal{L}_{ins} = \|\mathbf{s} - \hat{\mathbf{s}}\|_2^2, \tag{7}$$

where $\mathbf{s}$ and $\hat{\mathbf{s}}$ correspond to the GT and predicted equivalent lighting intensity.

SNE module With estimated equivalent lighting intensities, we then normalize the input multispectral images at the corresponding spectral bands to remove the influence of spectral components in MPS. Given these normalized image observations ($\mathbf{I}'$ in Fig. 4) as well as the calibrated lighting directions, our SNE module recovers the surface normal map by considering the geometric component only. As discussed in Sec. 3, any existing non-Lambertian CPS methods can be used as our SNE module. In this paper, we choose a learning-based CPS method PS-FCN [Chen *et al.*, 2018] as our SNE module. For better fitting the MPS image observations, the PS-FCN as the SNE module are re-trained on our synthetic multispectral dataset "PS-Spectral", as we will show in the next. Following the training strategy of PS-FCN [Chen *et al.*, 2018], the loss function is defined as the cosine similarity between the estimated and the ground-truth surface normals:

$$\mathcal{L}_n = \frac{1}{p}\sum_{i=1}^{p}(1 - \mathbf{n}_i^\top \hat{\mathbf{n}}_i), \tag{8}$$

where p is the number of valid pixels in the object mask, $\mathbf{n}_i$ and $\hat{\mathbf{n}}_i$ denote the ground-truth and estimated surface normal vectors at the i-th pixel.

Difference from SDPS-Net Our method shares a similar pipeline with the SDPS-Net [Chen *et al.*, 2019a]. Here we elaborate on the differences between our work and SDPS-Net [Chen *et al.*, 2019a]. Firstly, SDPS [Chen *et al.*, 2019a] focuses on the non-Lambertian uncalibrated CPS problem, which has a completely different image formation from the non-Lambertian MPS problem that we are working on. Specifically, the LC-Net of the SDPS-Net [Chen *et al.*, 2019a] learns to estimate the lighting direction and light source irradiance in the uncalibrated CPS, where no assumption is made on the non-Lambertian BRDF. However, the ELIE-Net of our method has calibrated lighting directions as input and focuses on predicting the equivalent light intensity, which encodes not only the light source irradiance but also the integral of camera spectral sensitivity, light spectra, and the material spectral component. In addition, our method for the MPS task relies on the proposed SRD assumption for non-Lambertian spectral BRDF, which ease the difficulties of the non-Lambertian MPS problem by the well-studied CPS problem.

Secondly, LC-Net [Chen *et al.*, 2019a] recasts the regression-based problems of lighting direction and intensity into a classification-based one, where the continuous lighting space is discretized into pre-defined bins and classes. However, we found equivalent light intensity has a high dynamic range (HDR)[2], as it encodes more than the light source irradiance. The discretization in LC-Net [Chen *et al.*, 2019a] is sub-optimal for accurately recovering the equivalent lighting intensity. Therefore, we predict equivalent lighting intensity in our MPS task in a regression manner.

"PS-Spectral" Dataset To train our NeuralMPS, we build a synthetic training dataset with shapes coming from the Blobby dataset [Johnson and Adelson, 2011] as well as the Sculpture dataset [Wiles and Zisserman, 2017]. Each shape in the dataset is rendered with 51 measured isotropic spectral BRDFs [Dupuy and Jakob, 2018].

To build the dataset, we first render a multispectral image tensor and its corresponding spectral reflectance tensor with the size of $[f, p, t]$, where f, p, t denote the number of light directions, pixels in the image, and spectral bands, respectively. In the MPS setting, each lighting direction corresponds to one spectral band. Therefore, we randomly select f out of t spectral bands to produce one multispectral image data with the size of $[f, p, f]$. In our dataset setting, we select $f = 39$ lighting directions distributed uniformly, and $t = 195$ spectral bands following the same wavelength range of [Dupuy and Jakob, 2018]. The spatial resolution of the image is set as 128×128. To generate the ground-truth equivalent lighting intensity, we conduct the SVD decomposition on the spectral reflectance tensor $\mathbf{R}$ with our SRD model following Eq. (3).

We also render a test dataset including two shapes: SPHERE and BUNNY. An example of the rendered multispectral images in our test dataset can be found in the supplementary material.

5 Experiments

In this section, we evaluate our method on both synthetic and real data. In Sec. 5.1, the accuracy of the proposed SRD model is verified on measured spectral BRDF database [Dupuy and Jakob, 2018] firstly. Then, in Sec. 5.2 we compare our NeuralMPS with the state-of-the-art MPS method GO21 [Guo *et al.*, 2021] on surface normal estimation. The ablation study with and without the ELIE-Net is also provided to evaluate the effectiveness of predicting equivalent light intensity. Finally, in Sec. 5.3 we give the qualitative evaluation on real-captured data to verify our application on real scenarios.

5.1 Evaluation of the SRD Model

Our SRD model decomposes the spectral reflectance into the geometric and spectral components via SVD decomposition. We verify the accuracy of the SRD model on diverse measured spectral BRDFs [Dupuy and Jakob, 2018]. Specifically, we choose a sphere shape containing diverse surface normals and render image observations under a set of lighting directions following uniform distribution. We conduct this rendering process on 51 spectral BRDFs [Dupuy and Jakob, 2018].

As mentioned in Sec. 4, the spectral BRDF can be concentrated as a matrix $\mathbf{R}$ with shape of $[f, p, t]$. With our

[2]Please find more details in the supplementary material.

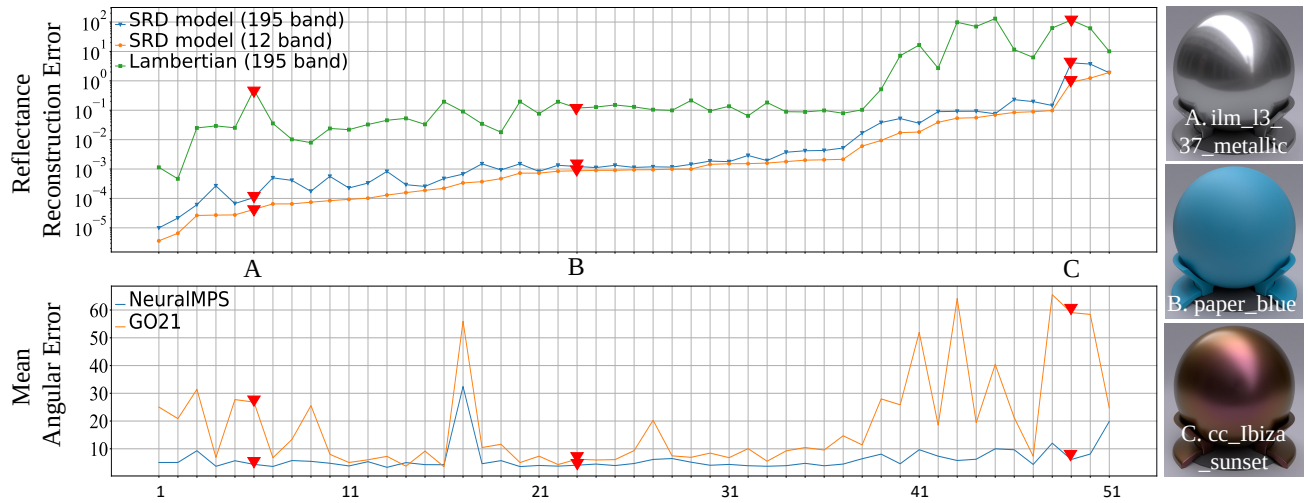

Figure 5: Evaluation of spectral reflectance decomposition (upper part), and surface normal estimation for diverse materials [Dupuy and Jakob, 2018] (lower part). Specular-dominant, diffuse-dominant, and specular-diffuse mixed material examples are labeled with A, B, and C, respectively. The X-axis represents indices of the materials for both parts.

SRD model, $\mathbf{R}$ is decomposed into geometric component $\mathbf{r}_g$ of shape $[f \times p]$ and spectral component $\mathbf{r}_s$ of shape $[t]$ via SVD in Eq. (3). To measure reconstruction error E_r of our SRD model, we reconstruct spectral BRDF $\mathbf{R}'$ from the spectral and geometric components, and calculate the Frobenius norm of the difference between the ground-truth spectral reflectance and the reconstructed ones, *i.e.*,

$$E_r = \frac{1}{51} \sum_i \|\mathbf{R}_i - \mathbf{R}'_i\|_F. \tag{9}$$

To show the effectiveness of our SRD model on non-Lambertian spectral reflectance, we also evaluate the approximation error of Lambertian assumption for the spectral BRDF. Specifically, we set the geometric component as an all-in-one vector, *i.e.*, $\mathbf{r}_g = \mathbf{1}$, and find the spectral component $\mathbf{r}_s$ in the Lambertian case that best fit $\mathbf{r}_g$ and $\mathbf{R}$ via least-square. After that, we also reconstruct the reflectance matrix from the geometric and spectral components for the Lambertian case.

As shown in Fig. 5 (top), the reconstructed spectral reflectance from our SRD model is more accurate than the Lambertian assumption used in GO21 [Guo *et al.*, 2021], revealing the effectiveness of our SRD model for general non-Lambertian spectral reflectance. Besides, we found that the reconstruction errors with $t = 12$ bands as well as $t = 195$ bands have little difference. Therefore, our SRD model can be applied to real-world multispectral cameras with limited spectral bands.

From Fig. 5, we observed that our SRD model works well for most materials. However, there are materials (*e.g.*, "cc_ibiza_sunset") showing mixed specular and diffuse reflectance. The spectral BRDFs for these materials violate our SRD model that assumes the reflectance is either diffuse- or specular- dominant, like materials "ilm_l3_37_metallic" and "paper_blue". During the training of our NeuralMPS, we drop the last 9 materials whose reconstruction errors are larger than 0.05, which are treated as outliers of our SRD model. However, we keep the 9 materials in the test phase for a fair comparison and present the normal estimation results of all the 51 spectral BRDFs in the next section.

ELIE-Net	SNE module	SPHERE	BUNNY
w/o	PS-FCN	7.05	13.12
w/	PS-FCN	6.27	12.89

Table 1: Ablation study of ELIENet on surface normal estimation.

5.2 Evaluation of NeuralMPS

As shown in Fig. 5 (bottom), we evaluate the normal estimation error on the SPHERE test dataset. The state-of-the-art MPS method GO21 [Guo *et al.*, 2021] is based on the assumption of Lambertian reflectance, therefore the surface normal estimation has relatively high accuracy on diffuse materials such as the "paper_blue". However, for materials with specular reflectance, GO21 [Guo *et al.*, 2021] cannot produce reasonable surface normal recovery, such as the case of "ilm_l3_37_metallic" and "cc_ibiza_sunset". Benefited from our SRD model and designed network for non-Lambertian MPS, the surface normal estimation of our method is more accurate compared with the state-of-the-art MPS method GO21 [Guo *et al.*, 2021] on diverse materials, demonstrating the strength of our MPS method on general spectral reflectance. On the other hand, we found the normal estimations for the last 9 materials have relatively higher error, possibly due to the relative lower accuracy of the SRD model in the corresponding spectral BRDFs.

Ablation study of ELIE-Net As shown in Table 1, to make an ablation study of ELIE-Net, we evaluate the surface normal estimation error of our NeuralMPS with and without the module on test dataset SPHERE and BUNNY. Specifically, "w/o" or "w/" in the first column refer to the results recovered from SNE module where the input images are normal-

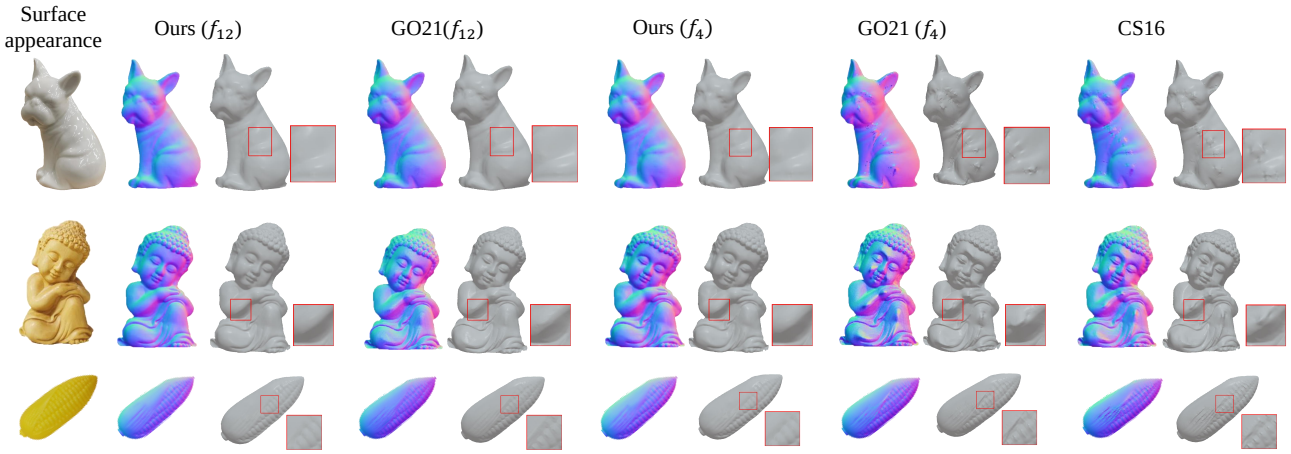

Figure 6: Qualitative comparison with existing MPS methods on real data. Shape distortions caused by non-Lambertian specular highlights are emphasized in the closed-up views.

ized by a all-one vector or the estimated equivalent light intensities from "ELIE-Net". Based on the mean angular error of the surface normal estimates over the test dataset, the absence of equivalent light intensities leads to a performance drop in normal estimation (by 11% and 2% for SPHERE and BUNNY). Therefore, the CPS method used in our SNE module ("w/o ELIE-Net") cannot be directly used for solving the MPS task. It is necessary for non-Lambertian MPS to predict the equivalent lighting intensity with our proposed ELIE-Net.

It's worth noticing that the ELIE-Net and SNE module are trained separately. Therefore the ELIE-Net can be directly applied to SNE modules that adopt other conventional photometric stereo (CPS) methods for solving multispectral photometric stereo (MPS), without the influence of a specific network structure. This setting would be beneficial for non-learning-based CPS works. On the other hand, the proposed NeuralMPS can also be trained end-to-end when the SNE module adopts learning-based CPS methods. We have our NeuralMPS which is trained in this way tested on test dataset SPHERE and BUNNY, and the mean angular error in degree is 6.22 for SPHERE and 9.10 for BUNNY.

5.3 Experiments on Real Data

In Fig. 6, we test our method on real multispectral images of non-Lambertian surfaces released by GO21 [Guo *et al.*, 2021]. Taking the spectral images and calibrated lighting directions as input, we compare our NeuralMPS with existing MPS methods CS16 [Chakrabarti and Sunkavalli, 2016] and GO21 [Guo *et al.*, 2021] on surface normal recovery. As CS16 [Chakrabarti and Sunkavalli, 2016] takes only three-channel RGB images as input, we select the image observations with the wavelength 450, 550, and 650 [nm] to mimic the RGB input. For GO21 [Guo *et al.*, 2021] and our method, we test the methods on multispectral images with 12 and 4 spectral bands (labeled as f_4 and f_{12}) to evaluate the application on real-world multispectral cameras with varying number of spectral channels. Since the ground-truth surface normal and shape are not available in the real-world MPS dataset, we conduct a qualitative evaluation by comparing the integrated shapes from the estimated surface normal maps using the normal integration method [Xie *et al.*, 2014].

As shown in Fig. 6, we present the estimated surface normals from existing methods and ours. Since CS16 [Chakrabarti and Sunkavalli, 2016] mainly focuses on Lambertian reflectance, the recovered shape is distorted due to the specular highlights, as shown in the closed-up views. GO21 [Guo *et al.*, 2021] is adapted to arbitrarily many spectral bands as the input and removes the specular highlights and shadows as outliers with the position thresholding strategy [Shi *et al.*, 2019]. Therefore, the shape recoveries from GO21 (f_{12}) are reasonable taking 12 spectral bands as input. However, the method fails with limited spectral input (GO21 (f_4)), as the position thresholding strategy prefers more data for the outlier removal. Compared with existing methods, our NeuralMPS achieves reasonable results for all the three objects. There are no distortions caused by specular highlights. Since our method learns the non-Lambertian spectral reflectance from the measured spectral BRDF dataset, rather than treating the specular highlights as outliers, the recovered surface normal remains in good quality even with limited spectral image inputs, as shown in Ours (f_4).

6 Conclusion

In this paper, we propose a multispectral photometric stereo method NeuralMPS for surface normal recovery under diffuse- or specular- dominant non-Lambertian spectral reflectance. The spectral reflectance decomposition model is the key of our method, which represents spectral BRDF as a composition of geometry components and spectral components. In this way, the non-Lambertian MPS problem can be recast into the well-studied non-Lambertian CPS problem with unknown equivalent light intensity in which the spectral component is embedded. Our NeuralMPS therefore design the ELIE-Net to estimate the equivalent light intensity and apply the SNE module to further achieve surface normal recovery under non-Lambertian spectral reflectance.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 62088102, 62136001, 62202409), the China Postdoctoral Science Foundation (Grant No. 2022M720236), and the JSPS KAKENHI (Grant No. JP23H05491).

Contribution Statement

Jipeng Lv and Heng Guo contributed equally to this work.

References

[Anderson *et al.*, 2011a] Robert Anderson, Björn Stenger, and Roberto Cipolla. Augmenting depth camera output using photometric stereo. *Machine Vision and Applications*, 2011.

[Anderson *et al.*, 2011b] Robert Anderson, Björn Stenger, and Roberto Cipolla. Color photometric stereo for multicolored surfaces. In *Proc. of International Conference on Computer Vision (ICCV)*, pages 2182–2189, 2011.

[Chakrabarti and Sunkavalli, 2016] Ayan Chakrabarti and Kalyan Sunkavalli. Single-image RGB photometric stereo with spatially-varying albedo. In *International Conference on 3D Vision (3DV)*, pages 258–266, 2016.

[Chen *et al.*, 2018] Guanying Chen, Kai Han, and Kwan-Yee K. Wong. PS-FCN: A flexible learning framework for photometric stereo. In *Proc. of European Conference on Computer Vision (ECCV)*, 2018.

[Chen *et al.*, 2019a] Guanying Chen, Kai Han, Boxin Shi, Yasuyuki Matsushita, and Kwan-Yee K. Wong. Self-calibrating deep photometric stereo networks. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[Chen *et al.*, 2019b] Lixiong Chen, Yinqiang Zheng, Boxin Shi, Art Subpa-Asa, and Imari Sato. A microfacet-based model for photometric stereo with general isotropic reflectance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[Dupuy and Jakob, 2018] Jonathan Dupuy and Wenzel Jakob. An adaptive parameterization for efficient material acquisition and rendering. *ACM Trans. on Graph.*, 37(6):1–14, 2018.

[Enomoto *et al.*, 2020] Kenji Enomoto, Michael Waechter, Kiriakos N Kutulakos, and Yasuyuki Matsushita. Photometric stereo via discrete hypothesis-and-test search. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2311–2319, 2020.

[Esteban *et al.*, 2011] CH Esteban, G Vogiatzis, and R Cipolla. Overcoming shadows in 3-source photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011.

[Guo *et al.*, 2021] Heng Guo, Fumio Okura, Boxin Shi, Takuya Funatomi, Yasuhiro Mukaigawa, and Yasuyuki Matsushita. Multispectral photometric stereo for spatially-varying spectral reflectances: A well posed problem? In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 963–971, 2021.

[Hui and Sankaranarayanan, 2016] Zhuo Hui and Aswin C Sankaranarayanan. Shape and spatially-varying reflectance estimation from virtual exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(10):2060–2073, 2016.

[Ikehata, 2018] Satoshi Ikehata. CNN-PS: CNN-based photometric stereo for general non-convex surfaces. In *Proc. of European Conference on Computer Vision (ECCV)*, 2018.

[Johnson and Adelson, 2011] Micah K Johnson and Edward H Adelson. Shape estimation in natural illumination. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011.

[Ju *et al.*, 2018] Yakun Ju, Lin Qi, Huiyu Zhou, Junyu Dong, and Liang Lu. Demultiplexing colored images for multispectral photometric stereo via deep neural networks. *IEEE Access*, pages 30804–30818, 2018.

[Ju *et al.*, 2020a] Yakun Ju, Xinghui Dong, Yingyu Wang, Lin Qi, and Junyu Dong. A dual-cue network for multispectral photometric stereo. *Pattern Recognition*, page 107162, 2020.

[Ju *et al.*, 2020b] Yakun Ju, Lin Qi, Jichao He, Xinghui Dong, Feng Gao, and Junyu Dong. MPS-Net: Learning to recover surface normal for multispectral photometric stereo. *Neurocomputing*, pages 62–70, 2020.

[Ju *et al.*, 2022] Yakun Ju, Kin-Man Lam, Wuyuan Xie, Huiyu Zhou, Junyu Dong, and Boxin Shi. Deep learning methods for calibrated photometric stereo and beyond: A survey. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[Kontsevich *et al.*, 1994] Leonid L Kontsevich, AP Petrov, and IS Vergelskaya. Reconstruction of shape from shading in color images. *Journal of the Optical Society of America*, pages 1047–1052, 1994.

[Ozawa *et al.*, 2018] Keisuke Ozawa, Imari Sato, and Masahiro Yamaguchi. Single color image photometric stereo for multicolored surfaces. *Computer Vision and Image Understanding*, pages 140–149, 2018.

[Rahman *et al.*, 2014] Sejuti Rahman, Antony Lam, Imari Sato, and Antonio Robles-Kelly. Color photometric stereo using a rainbow light for non-Lambertian multicolored surfaces. In *Proc. of Asian Conference on Computer Vision (ACCV)*, pages 335–350. Springer, 2014.

[Santo *et al.*, 2017] Hiroaki Santo, Masaki Samejima, Yusuke Sugano, Boxin Shi, and Yasuyuki Matsushita. Deep photometric stereo network. In *Proc. of International Conference on Computer Vision Workshops (ICCVW)*, 2017.

[Shafer, 1985] Steven A Shafer. Using color to separate reflection components. *Color Research & Application*, (4):210–218, 1985.

[Shi *et al.*, 2014] Boxin Shi, Ping Tan, Yasuyuki Matsushita, and Katsushi Ikeuchi. Bi-polynomial modeling of low-frequency reflectances. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014.

[Shi *et al.*, 2019] Boxin Shi, Zhipeng Mo, Zhe Wu, Dinglong Duan, Sai-Kit Yeung, and Ping Tan. A benchmark dataset and evaluation for non-Lambertian and uncalibrated photometric stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019.

[Silver, 1980] William M. Silver. *Determining shape and reflectance using multiple images*. PhD thesis, Massachusetts Institute of Technology, 1980.

[Szeliski, 2010] Richard Szeliski. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

[Wiles and Zisserman, 2017] Olivia Wiles and Andrew Zisserman. SilNet: Single-and multi-view reconstruction by learning from silhouettes. In *The British Machine Vision Conference*, 2017.

[Woodham, 1980] Robert J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical engineering*, 1980.

[Xie *et al.*, 2014] Wuyuan Xie, Yunbo Zhang, Charlie CL Wang, and Ronald C-K Chung. Surface-from-gradients: An approach based on discrete geometry processing. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2195–2202, 2014.

[Yao *et al.*, 2020] Zhuokun Yao, Kun Li, Ying Fu, Haofeng Hu, and Boxin Shi. GPS-NET: Graph-based photometric stereo network. *Proc. of Annual Conference on Neural Information Processing Systems (NeurIPS)*, 33, 2020.

Non-Lambertian Multispectral Photometric Stereo via Spectral Reflectance Decomposition (Supplementary Material)

Jipeng Lv[1,*] , **Heng Guo**[2,3*] , **Guanying Chen**[4] , **Jinxiu Liang**[1] , **Boxin Shi**[1]
[1]Peking University
[2]Beijing University of Posts and Telecommunications
[3]Osaka University
[4]The Chinese University of Hong Kong, Shenzhen
heng.guo@ist.osaka-u.ac.jp, chenguanying@cuhk.edu.cn, {lvjipeng, cssherryliang, shiboxin}@pku.edu.cn

1 Dynamic shape recovery

Recall that a single input multispectral image (MSI) for multispectral photometric stereo (MPS) encodes observations under varying lighting directions in different spectral bands. Given MSIs captured by a multispectral camera at different timestamps, the proposed MPS method can recover the surface normal of dynamic objects. Please find the attached video for the dynamic shape recovery results from the proposed method. The input MSI sequence records a moving hand captured with spectral light sources. The spectral reflectance of the human hand roughly follows the assumptions of the proposed SRD model. As shown in the attached video, the 3D shape of the moving hand in the MSI sequence can be stably reconstructed by the proposed method.

2 Code release

We release the implementation, pre-trained model, and test data of our work on GitHub. Please visit https://github.com/Github-Tree-0/NeuralMPS for details.

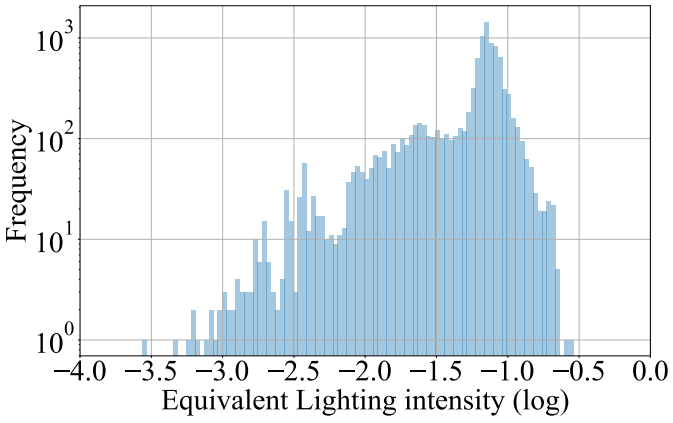

Figure 1: Histogram of equivalent light intensities has a broad range, which shows their property of high dynamic range.

3 Implementation

We train the ELIE-Net and the SNE module of our NeuralMPS separately. Specifically, the GT equivalent light intensity is used as the supervised signal of the ELIE-Net. When training the SNE module, we take the GT equivalent light intensity as input and use the GT surface normal to supervise the SNE module. Similar to SDPS-Net [Chen *et al.*, 2019], we randomly multiply a scale factor in the range of $(0.1, 1)$ on both the GT equivalent light intensity and the rendered image observations during the training to improve the robustness of equivalent lighting intensity prediction.

4 Shape recovery evaluation on synthetic data

As shown in Fig. 4, we compare shape recovery results of the proposed method and GO21 [Guo *et al.*, 2021] on two synthetic objects covered by a specular dominant material "ilm_l3_37_metallic" and a diffuse dominant material "ilm_solo_millennium_falcon", respectively. The estimated normal maps, and the corresponding angular error maps, and the mean angular error (MAE) values are shown in Fig. 4. Compared to GO21 [Guo *et al.*, 2021], our method achieves more accurate surface normal estimation results on both objects, especially on the specular dominant material "ilm_l3_37_metallic", revealing the strength of our NeuralMPS on general non-Lambertian materials.

5 Effectiveness of the proposed ELIE-Net

HDR property of the equivalent light intensity As shown in Fig. 1, we plot the histogram of the equivalent light intensities in log scale from 51 measured materials [Dupuy and Jakob, 2018] following the calculation from Eq. (3) of the main paper. The light intensity or irradiance defined in conventional photometric stereo normally varies in a small range based on the real experimental setup. Therefore, existing uncalibrated photometric stereo method such as SDPS-Net [Chen *et al.*, 2019] divides the range of possible light intensities into a fixed number of classes and predict the light intensities as a classification problem. However, the equivalent light intensity is more complicated as it also entangles the material spectral reflectance component. As shown in Fig. 1, the equivalent light intensity values vary from 10^{-4} to $10^{-0.5}$, which are in high dynamic ranges. Therefore, our ELIE-Net predicts the equivalent light intensity in a regression manner.

Table 1: Quantitative results of the predicted equivalent light intensity, which are measured by the mean ratio over 51 materials of the test dataset. Lower values are better.

Method	SPHERE	BUNNY
Ours	0.24	0.27
GO21 [Guo *et al.*, 2021]	1.23	1.22

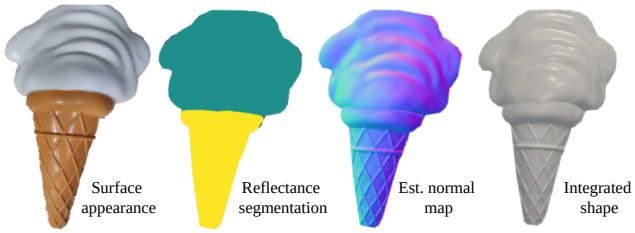

Figure 2: Our method can be applied to spatially-varying general reflectance with given reflectance segmentation. The integrated shape of our normal estimation is consistent with the surface appearance.

Comparison with GO21 on equivalent light intensity GO21 [Guo *et al.*, 2021] is based on a Lambertian setting. The equivalent light intensities of GO21 [Guo *et al.*, 2021] can be obtained from the their estimated chromaticities. We compare the equivalent light intensity estimates of the proposed ELIE-Net and GO21 [Guo *et al.*, 2021] on the SPHERE and BUNNY test dataset. To evaluate the accuracy of the predicted equivalent light intensity, we use a "ratio" metric defined as follows:

$$\text{ratio} = \frac{|e - e_{gt}|}{\max\{e_{gt}, 10^{-3}\}}, \tag{1}$$

where e, e_{gt} denote the predicted and the ground truth equivalent light intensities, respectively. The results in Fig. 5 demonstrate the superior performance of the proposed ELIE-Net over GO21 [Guo *et al.*, 2021] on diverse material. In Table 1, we summarize the mean ratios on equivalent light intensity prediction. The proposed ELIE-Net is shown to be nearly 6 times more accurate than GO21 [Guo *et al.*, 2021].

6 Effectiveness of the proposed SRD model

For general spectral BRDFs, the reflectance varies w.r.t different incident-outgoing directions and wavelengths. The Lambertian model assumes the spectral reflectance is only related to the wavelength. Therefore, the geometric component $\mathbf{r_g}$ shown in Eq. (3) of the main paper can be treated as a all-in-one vector in the Lambertian case. Given the ground-truth spectral BRDF, we can fit the Lambertian spectral component $\mathbf{r_s}$ with $\mathbf{r_g} = \mathbf{1}$ in a least-square manner. On the other hand, our SRD model conducts SVD to extract the spectral and geometry components.

Based on the spectral and geometry components estimated from our SRD model and the Lambertian model, we reconstruct two spectral BRDFs and compute their differences with the ground-truth spectral BRDF. In this way, we can evaluate the effectiveness of our SRD model over the Lambertian model. Denoting the ground-truth and reconstructed spectral BRDF as $\mathbf{R}$ and $\mathbf{R}'$, we define the BRDF reconstruction error ratio as

$$\hat{E}_r = \frac{\|\mathbf{R} - \mathbf{R}'\|_F^2}{\|\mathbf{R}\|_F^2}. \tag{2}$$

As shown in Fig. 6, we plot the reconstruction error ratio over 51 measured spectral BRDFs [Dupuy and Jakob, 2018]. Our SRD model is shown to be more accurate than the Lambertian model, making our NeuralMPS capable to handle non-Lambertian MPS problems.

Table 2: Surface normal estimation errors on our MPS test dataset measured by MAngE in degree by applying different CPS methods: PS-FCN [Chen *et al.*, 2018], ST14 [Shi *et al.*, 2014], and AttentionPSN [Ju *et al.*, 2020] as the SNE module

ELIE-Net	PS-FCN	ST14	AttentionPSN	GO21
w/o	13.12	26.49	16.08	25.44
w/	12.89	23.09	11.78	

7 Different baselines for SNE module

We replace PSFCN [Chen *et al.*, 2018] with other CPS methods as the SNE module, where AttentionPSN [Ju *et al.*, 2020] is one of the latest CPS works, and ST14 [Shi *et al.*, 2014] is a non-learning based CPS method. Here we directly adopted the pre-trained model of AttentionPSN without fine-tuning it on our MPS synthetic training dataset. As shown in Table 2, compared to the stateof-the-art MPS method GO21 [Guo *et al.*, 2021] that is based on Lambertian reflectance, our MPS method with a non-Lambertian SNE module can achieve better results on the synthetic BUNNY test dataset. The estimation accuracy of different CPS baselines is improved by attaching our ELIE-Net, which is the key to transferring CPS methods for solving the non-Lambertian MPS problem. Besides, the estimation error from AttentionPSN [Ju *et al.*, 2020] is lower than PS-FCN [Chen *et al.*, 2018], showing that our MPS method can be enhanced with the latest CPS works. By further conducting fine-tuning on our synthetic MPS training dataset, AttentionPSN [Ju *et al.*, 2020] as our SNE module is expected to achieve more accurate results.

8 Test dataset SPHERE and BUNNY

We render a test dataset including two shapes: SPHERE and BUNNY. In Fig. 3 is an example of the rendered multispectral images.

9 Limitation

We use a uniform material assumption in the main paper so that the spectral reflectance can be decomposed into a global spectral component and geometric component. A possible solution is demonstrated in the next section. Furthermore, our SRD model assumes dominant diffuse- or specular- reflectance, which may comply with some materials in the real world. An intuitive idea is to separate the specular and diffuse parts of the spectral reflectance, which will be considered in our future work.

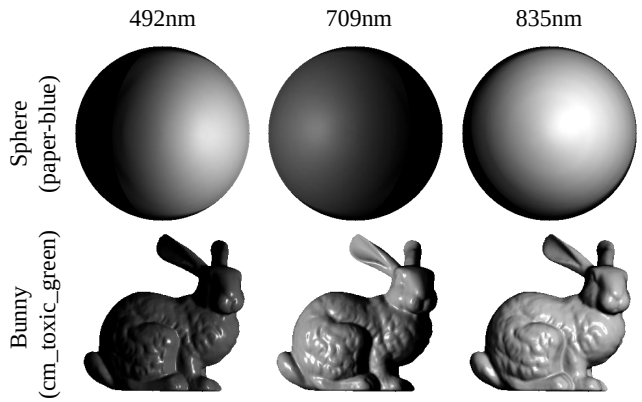

Figure 3: Our synthetic test dataset SPHERE and BUNNY rendered under varying wavelengths, lighting directions, and measured spectral BRDFs [Dupuy and Jakob, 2018].

10 Uniform Material Assumption

We assume the surface is covered by uniform material in the main paper. However, it is desirable to consider the more general spatially-varying non-Lambertian spectral reflectance. Similar to GO21 [Guo *et al.*, 2021], if we have access to reflectance segmentation, we can apply our non-Lambertian MPS method to each region with uniform material and then combine the output surface normals to obtain a complete surface normal map, as shown in Fig. 2.

11 Measured spectral BRDFs for rendering

We select 51 measured spectral BRDFs from [Dupuy and Jakob, 2018] in our synthetic experiment. The material appearances and names are listed in Fig. 7. The material indices shown in the x-axis of Fig. 6 in our main paper are labeled in Fig. 7.

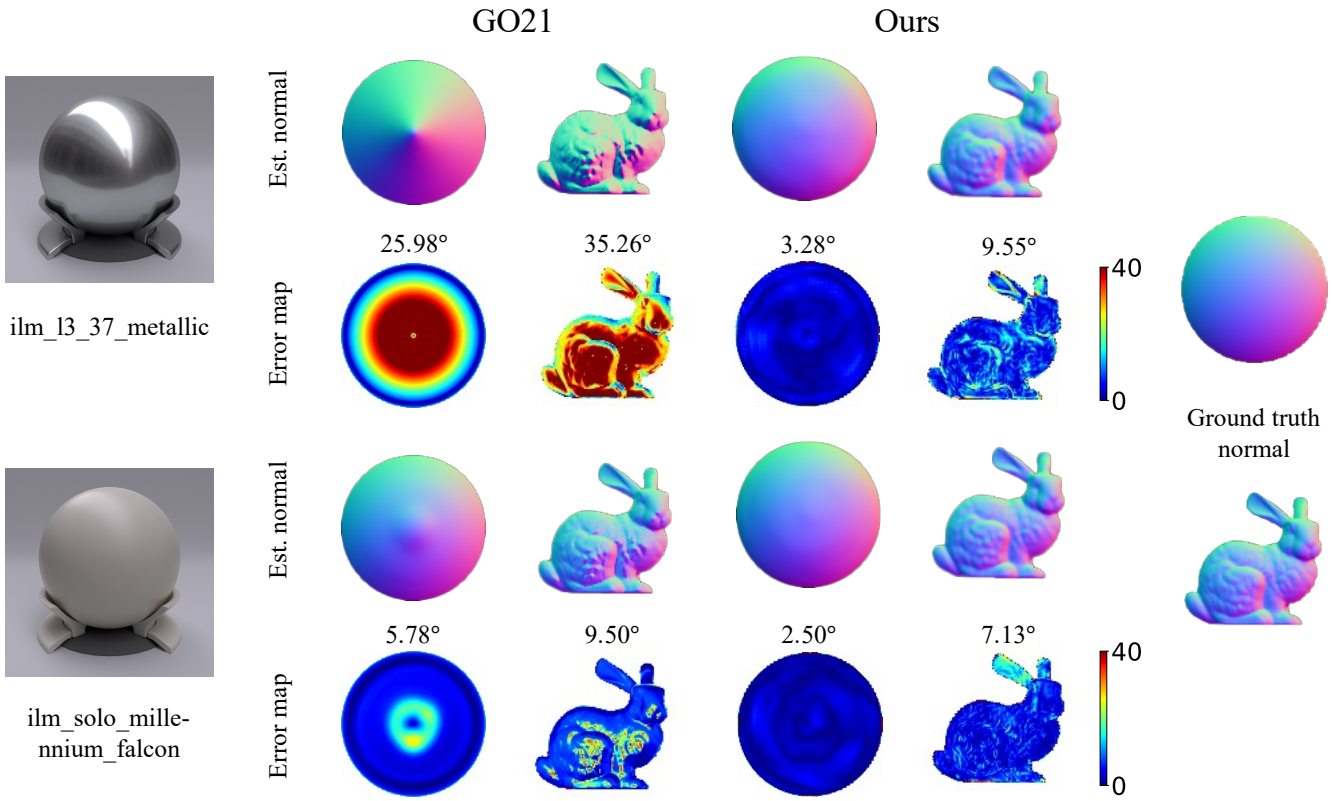

Figure 4: Shape recovery comparison between the proposed method and GO21 [Guo *et al.*, 2021] on synthetic objects SPHERE and BUNNY rendered with "ilm_l3_37_metallic" and "ilm_solo_millennium_falcon" materials. The material appearances under natural illumination are shown in the first column. The estimated normal maps, their corresponding mean angular error (MAE) values, and error maps are shown from the second to the fourth column. The ground-truth surface normal maps are shown in the last column.

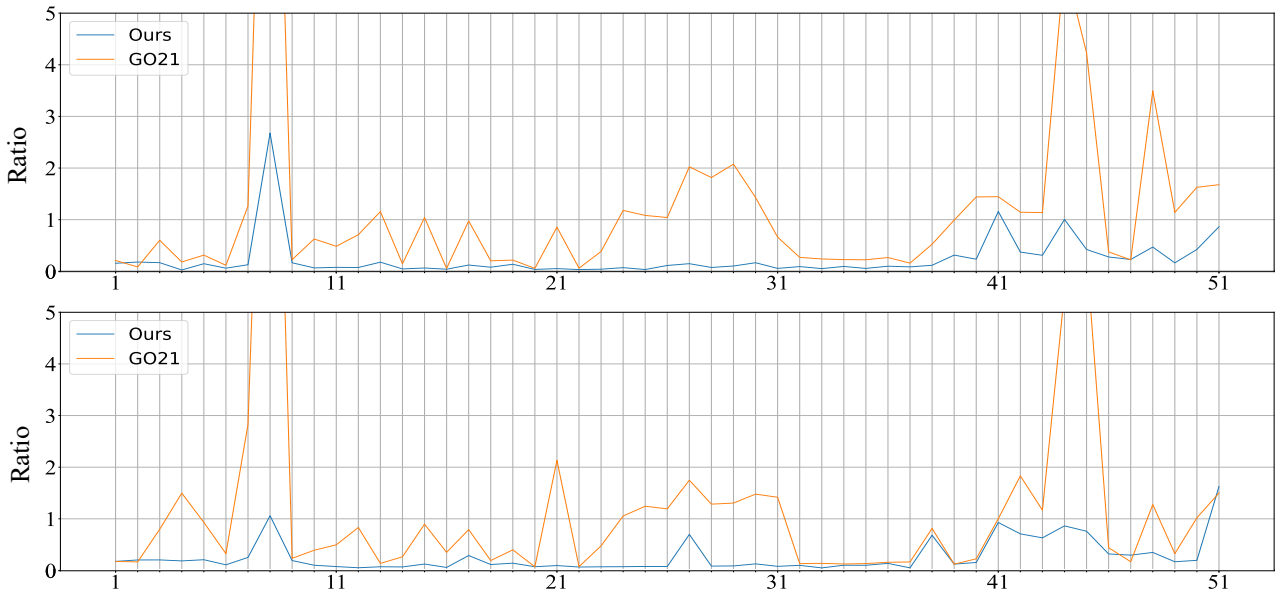

Figure 5: Results of predicted equivalent light intensities for diverse materials [Dupuy and Jakob, 2018] on SPHERE (upper part) and BUNNY (lower part) test dataset. The X-axis indicates indices of the materials. The predicted equivalent light intensities for each material are evaluated using the mean of "ratio" defined in Eq. (1).

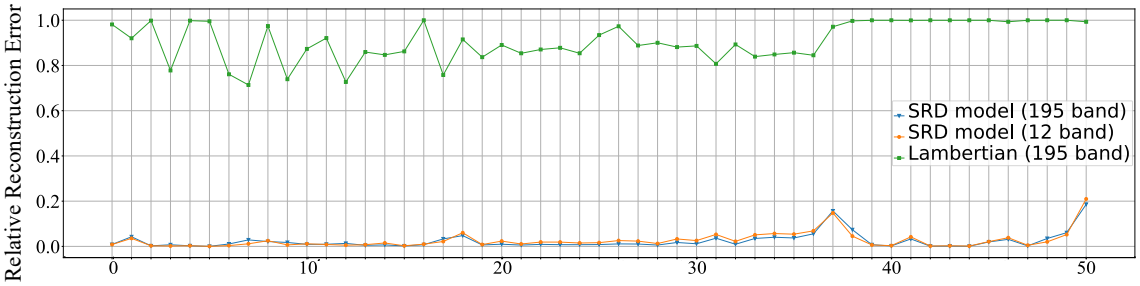

Figure 6: Visualization of the spectral BRDF reconstruction error ratios over 51 measured BRDFs.

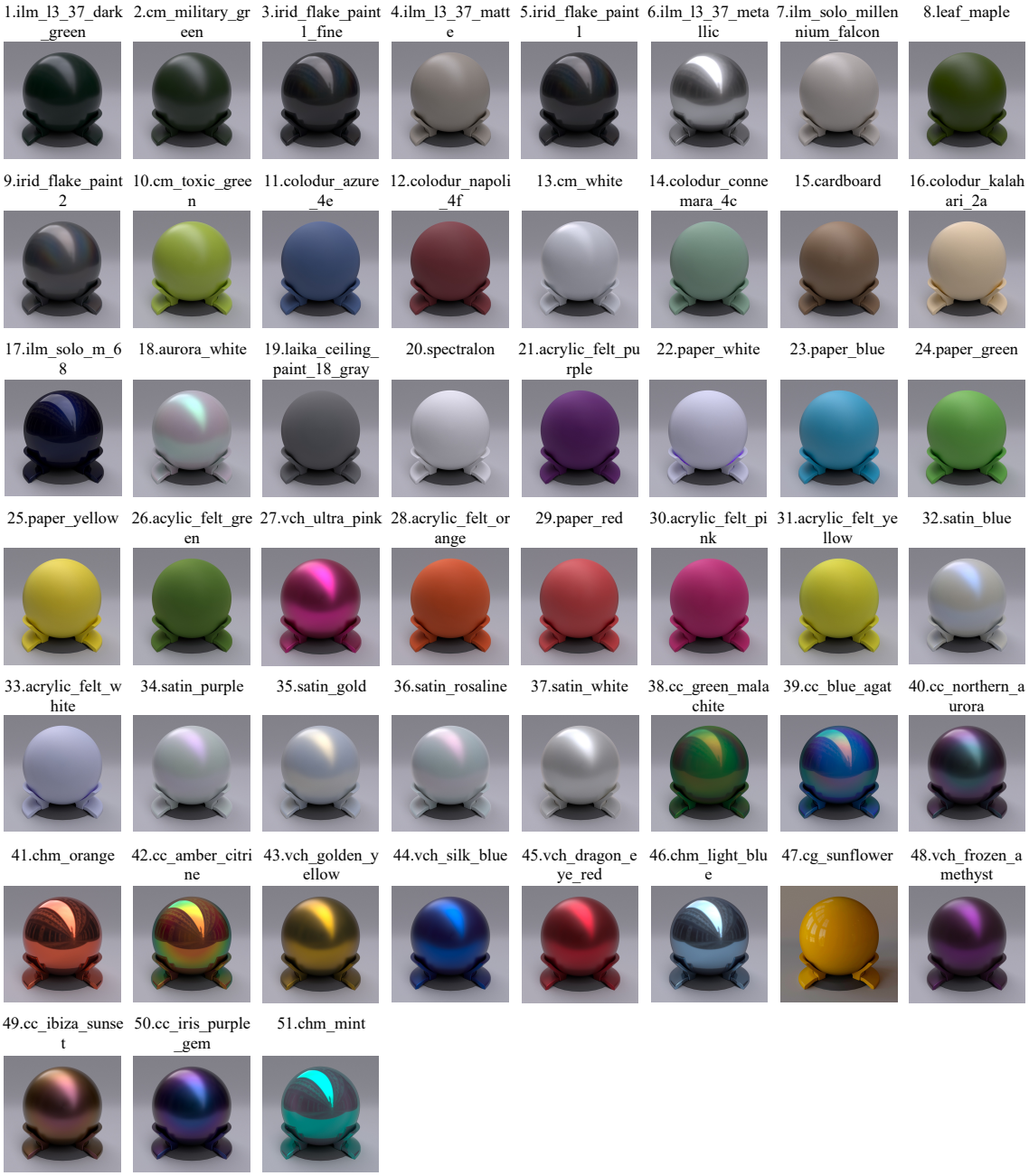

Figure 7: Material IDs and the corresponding appearances [Dupuy and Jakob, 2018] used in this paper.

References

[Chen *et al.*, 2018] Guanying Chen, Kai Han, and Kwan-Yee K. Wong. PS-FCN: A flexible learning framework for photometric stereo. In *Proc. of European Conference on Computer Vision (ECCV)*, 2018.

[Chen *et al.*, 2019] Guanying Chen, Kai Han, Boxin Shi, Yasuyuki Matsushita, and Kwan-Yee K. Wong. Self-calibrating deep photometric stereo networks. In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[Dupuy and Jakob, 2018] Jonathan Dupuy and Wenzel Jakob. An adaptive parameterization for efficient material acquisition and rendering. *ACM Trans. on Graph.*, 37(6):1–14, 2018.

[Guo *et al.*, 2021] Heng Guo, Fumio Okura, Boxin Shi, Takuya Funatomi, Yasuhiro Mukaigawa, and Yasuyuki Matsushita. Multispectral photometric stereo for spatially-varying spectral reflectances: A well posed problem? In *Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 963–971, 2021.

[Ju *et al.*, 2020] Yakun Ju, Kin Man Lam, Yang Chen, Lin Qi, and Junyu Dong. Pay attention to devils: A photometric stereo network for better details. In *Proc. of International Joint Conference on Artificial Intelligence*, pages 964–700, 2020.

[Shi *et al.*, 2014] Boxin Shi, Ping Tan, Yasuyuki Matsushita, and Katsushi Ikeuchi. Bi-polynomial modeling of low-frequency reflectances. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2014.