A Benchmark Dataset and Evaluation for Non-Lambertian and Uncalibrated Photometric Stereo

Boxin Shi, *Member, IEEE,* Zhipeng Mo, Zhe Wu, Dinglong Duan,
Sai-Kit Yeung *Member, IEEE,* and Ping Tan, *Senior Member, IEEE*

Abstract—Classic photometric stereo is often extended to deal with real-world materials and work with unknown lighting conditions for practicability. To quantitatively evaluate non-Lambertian and uncalibrated photometric stereo, a photometric stereo image dataset containing objects of various shapes with complex reflectance properties and high-quality ground truth normals is still missing. In this paper, we introduce the 'DiLiGenT' dataset with calibrated Directional Lightings, objects of General reflectance with different shiness, and 'ground Truth' normals from high-precision laser scanning. We use our dataset to quantitatively evaluate state-of-the-art photometric stereo methods for general materials and unknown lighting conditions, selected from a newly proposed photometric stereo taxonomy emphasizing non-Lambertian and uncalibrated methods. The dataset and evaluation results are made publicly available, and we hope it can serve as a benchmark platform that inspires future research.

Index Terms—Photometric stereo, Benchmark, Dataset, Non-Lambertian, Uncalibrated

✦

1 INTRODUCTION

THE classic photometric stereo problem [1] assumes a Lambertian surface reflectance model, a fixed orthographic camera with linear radiometric response, and known directional illuminations to inversely solve for the radiometric image formation model under different lightings and hence recover a per-pixel surface normal map from at least three images. Different from multi-view stereo, photometric stereo algorithms evolve from diversifying their image formation model assumptions and imaging conditions beyond the classic concept, *e.g.,* from Lambertian to non-Lambertian surface, from known to unknown lighting conditions, from directional to environmental lighting, from orthographic to perspective camera, from single to multiple viewpoints, from static to dynamic scenes, and so on. Modern algorithms are developed to deal with surfaces with general reflectance [2], uncalibrated lighting [3], general environment illumination [4], perspective camera [5], multiple viewpoints [6], moving objects [7], and so on. Such diversification raises the need of an organized categorization or taxonomy for photometric stereo algorithms. It also brings great challenges to quantitatively compare photometric stereo algorithms using a benchmark dataset with ground truth, while successful examples greatly inspiring

future research exist in multi-view stereo [8].

This paper aims to provide a taxonomy, a benchmark dataset and quantitative evaluation for photometric stereo algorithms. To make benchmarking photometric stereo feasible, we narrow our scope to specific categories of algorithms. In particular, we focus on two major difficulties in photometric stereo – non-Lambertian and uncalibrated algorithms, under the classic data capture setup [1] – directional lighting and a fixed orthographic camera. Understanding the performance of photometric stereo under such conditions is basic for many other extensions (*e.g.,* multi-view, general environment lighting, dynamic scenes, *etc.*).

We first provide a taxonomy of photometric stereo methods, with an emphasis on recent non-Lambertian and uncalibrated algorithms, followed by a brief categorization for other extensions beyond the classic data capture setup and different numbers of input images. We then present the 'DiLiGenT' photometric stereo image dataset which is captured under carefully calibrated directional lightings, with reference shapes for all objects provided by a high-end laser scanner. Our dataset contains objects of various materials and shapes, from ideally diffuse to highly reflective, from smooth to bristly. Finally, we quantitatively evaluate recently representative non-Lambertian [2, 9–15] and uncalibrated [3, 16–19] photometric stereo algorithms to understand their pros and cons and motivate future research on unsolved issues. The dataset and evaluation results are available for download from our project website: https://sites.google.com/site/photometricstereodata/.

Our main contributions are threefold:

- An up-to-date survey and taxonomy on recent photometric stereo techniques with a special focus on non-Lambertian and uncalibrated methods;

- *B. Shi (corresponding author) is with the Institute of Digital Media, School of Electronics Engineering and Computer Science, Peking University (PKU), China. Part of this work was finished when B. Shi was working at the Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology (AIST), Japan.*
  E-mail: shiboxin@pku.edu.cn
- *Z. Mo and S.-K. Yeung are with Singapore University of Technology and Design (SUTD); P. Tan is with Simon Fraser University (SFU). Part of this work was finished while Z. Wu and D. Duan were with P. Tan's group at National University of Singapore (NUS).*

- The first photometric stereo dataset with 'ground truth' shapes and objects of complex reflectance for single-view methods under the directional lightings;
- A quantitative evaluation of recent non-Lambertian and uncalibrated photometric stereo algorithms, which serves as a benchmark platform together with the dataset.

This paper extends its preliminary version [20] in all three aspects of contributions above. We provide a more comprehensive taxonomy by categorizing additional relevant and recent works in Section 2. We expand the previous dataset with additional nine sets of data but without releasing the 'ground truth' for better evaluating newly proposed methods and we show details in creating the 'DiLiGenT' dataset including the lighting calibration, 'ground truth' shape scanning, and shape-normal alignment in Section 3. Finally, we add analysis to the performance of state-of-the-art non-Lambertian and uncalibrated photometric stereo algorithms under varying lighting distributions and with various error inducing factors such as attached/cast shadow, interreflection, elevation angle of normal, and half angle (specularity) in Section 4.

2 A PHOTOMETRIC STEREO TAXONOMY

We start categorizing photometric stereo algorithms under the classic data capture setup, *i.e.*, the scene is illuminated by directional lightings and recorded with a fixed orthographic camera. With such a setup, the main challenges for photometric stereo are how to deal with non-Lambertian materials (Section 2.1) and unknown lighting conditions (Section 2.2). To provide a complete taxonomy, we briefly review other photometric stereo algorithms beyond the classic data capture setup in Section 2.3 and number of images required for different algorithms in Section 2.4. There are many different ways to categorize photometric stereo algorithms; we refer the reader to [21] for discussing the relationship between photometric and other 3D modeling approaches and [22] for detailed review of representative methods.

2.1 Non-Lambertian photometric stereo

Lambert's reflectance model is widely adopted in photometric stereo for its simplicity. However, the real-world is full of non-Lambertian objects. Many photometric stereo algorithms have been developed to deal with non-Lambertian materials.

Non-Lambertian photometric stereo problems could be summarized as in the top row of Table 1[1]. Here, $\mathbf{I}_{P \times Q}$ is the observation matrix of P points under Q different lighting conditions; $\mathbf{N}_{P \times 3}^{\top}$ and $\mathbf{L}_{3 \times Q}$ are stacks of normal vector $\mathbf{n} \in \mathbb{R}^{3 \times 1}$ and lighting vector $\mathbf{l} \in \mathbb{R}^{3 \times 1}$ respectively; $\boldsymbol{\rho}(\mathbf{n}, \mathbf{l})_{P \times Q}$ are the Bidirectional Reflectance Distribution Function (BRDF) values for each observation, which is a function of normal and lighting directions given viewing direction fixed at $\mathbf{v} = (0, 0, 1)^{\top}$; and '∘' denotes the element-wise multiplication.

The Lambertian photometric stereo [1] (BASELINE) simply assumes $\boldsymbol{\rho}(\mathbf{n}, \mathbf{l})$ is a constant function (unknown scaling)

1. Please refer to the supplementary material for an illustration of the coordinate system and notations.

of $\mathbf{n}$ and $\mathbf{l}$, and attached shadow ($\max\{\cdot, 0\}$) can also be ignored. Given calibrated $\mathbf{L}$, the surface normal matrix $\mathbf{N}$ is solved by linear Least Squares, and the reflectance values are approximated by normalizing each row of the estimated $\mathbf{N}$. The goal of non-Lambertian photometric stereo methods is to estimate $\mathbf{N}$ when $\boldsymbol{\rho}(\mathbf{n}, \mathbf{l})$ is a general unknown reflectance function.

Outlier rejection based methods assume non-Lambertian phenomena (such as shadow and specular highlight) are local and sparse, such that they can be detected and discarded as outliers. Earlier methods [23–25] select three optimal lights out of four where the surface appears mostly Lambertian to estimate normal. With more input images, this subset of Lambertian images can be extracted in a more robust manner by Markov random field [26], graph cuts [27], RANSAC [28, 29], maximum-likelihood estimation [30], maximum feasible subsystem [31], taking the median values [32], using robust SVD [33], or expectation maximization [34]. Instead of discarding non-Lambertian reflection as outliers, Zickler *et al.* [35] analyze in a subspace that is free from highlights.

Recent methods apply robust statistical techniques to reject outliers, by assuming a low-rank (Lambertian) observation matrix plus a sparse outlier matrix (shadow and specularity) [9, 10] (WG10, IW12), as shown in Table 1. The difference is that WG10 [9] aims at minimizing the rank of matrix, while IW12 [10] employs sparse Bayesian regression to explicitly enforce the rank-3 constraint. Such rank minimization approaches have the advantages of robustness, since the only assumption to the outlier matrix is its sparsity regardless of the error magnitude and distribution.

Outlier rejection based methods generally cannot deal with materials with broad and soft specular reflection, where the non-Lambertian outliers are dense. They often require more input images to make statistical analysis.

Analytic reflectance models are employed in some photometric stereo algorithms to model non-Lambertian reflectance. Instead of discarding specular reflection as outliers, they fit a nonlinear analytic BRDF to interpret all the observed data. Therefore, they have the advantage of exploiting all available data. Along this direction, various analytic BRDF models have been incorporated, including the Blinn-Phong model [36], the Torrance-Sparrow model [37], the Ward model [38], a mixture of multiple Ward models [2, 39], specular spike model [40, 41], and a microfacet BRDF with ellipsoid normal distribution function [42].

Take the method in [2] (GC10) as an example. GC10 [2] adopts the Ward model to explicitly represent $\boldsymbol{\rho}(\mathbf{n}, \mathbf{l})$. Each Ward lobe involves several parameters – diffuse strength d, specular strength s, and surface roughness α – to describe the shininess of the surface, as summarized in Table 1. In order to handle spatially-varying BRDFs, GC10 [2] linearly combines several $\boldsymbol{\rho}_i(d_i, s_i, \alpha_i)$ (i-th basis materials) multiplied by a spatially-varying weight $\mathbf{w}_i$. GC10 [2] needs to solve a nonlinear system for $\{d_i, s_i, \alpha_i, \mathbf{w}_i\}$ besides estimating surface normal $\mathbf{n}$, which is computationally expensive.

The drawback of such approaches is that the analytic models vary significantly from material to material, and each of them is limited to a narrow class of materials. Furthermore, these models are highly nonlinear. So such approaches in principal require complicated case by case

TABLE 1
Summary of non-Lambertian photometric stereo assumptions and formulations.

	Solve $\mathbf{N}$ from $\mathbf{I} = \max\{\boldsymbol{\rho}(\mathbf{n},\mathbf{l}) \circ (\mathbf{N}^\top \mathbf{L}), 0\}$ by using different assumptions and constraints on $\boldsymbol{\rho}(\mathbf{n},\mathbf{l})$
	Notations: $\mathbf{h} = (\mathbf{l}+\mathbf{v})/\|\mathbf{l}+\mathbf{v}\|$, $\theta_h = \langle \mathbf{n}, \mathbf{h}\rangle = \arccos(\mathbf{n}^\top \mathbf{h})$, $\theta_d = \langle \mathbf{l}, \mathbf{h}\rangle = \arccos(\mathbf{l}^\top \mathbf{h})$
BASELINE [1]	$\boldsymbol{\rho}(\mathbf{n},\mathbf{l}) \approx \mathbf{D}$, where each row of $\mathbf{D}$ is a constant representing the albedo of a Lambertian surface
WG10 [9]	$\boldsymbol{\rho}(\mathbf{n},\mathbf{l}) \approx \mathbf{D}+\mathbf{E}$, where $\mathbf{E}$ is sparse and $\mathrm{rank}(\mathbf{I})$ is minimized
IW12 [10]	$\boldsymbol{\rho}(\mathbf{n},\mathbf{l}) \approx \mathbf{D}+\mathbf{E}$, where $\mathbf{E}$ is sparse and $\mathrm{rank}(\mathbf{I}) = 3$
GC10 [2]	$\boldsymbol{\rho}(\mathbf{n},\mathbf{l}) \approx \sum_i \mathbf{w}_i \circ \boldsymbol{\rho}_i(d_i, s_i, \alpha_i)$, where $\boldsymbol{\rho}_i(d_i, s_i, \alpha_i) = \frac{d_i}{\pi} + \frac{s_i}{4\pi\alpha_i^2 \sqrt{(\mathbf{n}^\top \mathbf{l})(\mathbf{n}^\top \mathbf{v})}} \exp\left(\frac{(1-1/\mathbf{n}^\top \mathbf{h})}{\alpha_i^2}\right)$
AZ08 [11]	$\boldsymbol{\rho}(\mathbf{n},\mathbf{l})$ is isotropic and depends only on (θ_h, θ_d)
ST12 [12]	$\boldsymbol{\rho}(\mathbf{n},\mathbf{l})$ is isotropic, depends only on θ_h, and is monotonic about $\mathbf{n}^\top \mathbf{h}$
HM10 [13]	$\boldsymbol{\rho}(\mathbf{n},\mathbf{l})$ is isotropic, monotonic about $\mathbf{n}^\top \mathbf{l}$, and $\boldsymbol{\rho}(\mathbf{n},\mathbf{l}) = 0$ for $\mathbf{n}^\top \mathbf{l} \leq 0$
ST14 [14]	The low-frequency part of $\boldsymbol{\rho}(\mathbf{n},\mathbf{l})$ is a bi-polynomial $A(\cos(\theta_h))B(\cos(\theta_d))$, where A and B are polynomials
IA14 [15]	$\boldsymbol{\rho}(\mathbf{n},\mathbf{l}) \approx \sum_i \boldsymbol{\rho}_i(\mathbf{n}^\top \alpha_i)$, where $\alpha_i = (p_i\mathbf{l}+q_i\mathbf{v})/\|p_i\mathbf{l}+q_i\mathbf{v}\|$, p_i, q_i are nonnegative unknown values

analysis for different material classes.

General properties of a BRDF such as isotropy, reciprocity, and monotonicity are valid for a broader class of materials. Photometric stereo algorithms capitalizing on these general properties have the potential to deal with much broader types of materials.

Many real world materials are isotropic, which show structured BRDF values. Isotropy BRDFs have the property that equal reflectance values are observed if $\mathbf{n}$ is symmetric about the plane spanned by $\mathbf{v}$ and $\mathbf{l}$. By exploring isotropy, the azimuth angle of a surface normal can be computed [43], and 'iso-depth' and 'iso-slope' contours [43–45] can be derived. It has been proved in [44, 45] that these two sets of contours determine the surface shape. The isotropic BRDFs can also be represented using a dictionary built from MERL database [46] to regularize the surface normal and solve the spatially-varying reflectance simultaneously [47].

Many isotropic BRDFs can be well approximated by bivariate functions [48], which allows iterative estimation of surface normals and BRDF represented using a 2D discrete table [11] (AZ08). By further assuming this bivariate function is monotonic in one dimension, Shi *et al.* [12] (ST12) estimate the elevation angle of a surface normal by searching for the correct candidate showing monotonic BRDF values. AZ08 [11], ST12 [12], and more recently in [49] are 'two-step' approaches that first estimate the azimuth angle using isotropic constraint [43] before the elevation angle estimation. Isotropy and monotonicity are further combined with visibility constraint in [13] (HM10) to deal with general materials whose BRDF consists of a single lobe. Some recent methods [14, 15] (ST14, IA14) develop novel bivariate BRDF representations to facilitate normal estimation. They use two thresholds, one to exclude shadows and the other one to exclude specular reflections, so the remaining reflectance values vary slowly. The specular threshold is necessary for ST14 [14], but optional for IA14 [15]. ST14 [14] models such low-frequency reflectance using a bi-polynomial representation, while IA14 [15] models it as a sum of lobes with unknown center directions. Please refer to Table 1 for their analytic formulae and constraints on general BRDFs.

Handling anisotropic BRDF is a well-known difficult problem for photometric stereo. Holroyd *et al.* [50] utilize a symmetric elliptical microfacet normal distribution function to solve this issue. This can also be solved by example-based methods like [51–53], which capture example (known shape) and target (unknown shape) objects of the same material under varying lighting conditions. The example-based approach has been adopted to build a high-resolution

microgeometry sensor [54].

The photometric stereo problem can also be solved by embedding all pixels to a hemisphere (*i.e.*, a 2D manifold representing all possible normal directions) while keeping appropriate distances between different pixels. Such distances are measured before embedding by similarity of pixel intensity profiles derived from analytic BRDFs [55], general isotropic BRDFs [19, 56], or attached shadow codes [57] which in principal can also handle anisotropic BRDFs.

2.2 Uncalibrated photometric stereo

Many photometric stereo algorithms assume known lighting conditions. Typically, this requires inserting additional calibration objects, such as a mirror sphere, into the scene during data capture. This process is tedious and the calibration sphere often causes interreflection on other scene objects. So uncalibrated methods are developed to automatically calibrate lighting conditions.

Most uncalibrated photometric stereo methods are Lambertian and based on the factorization technique proposed in [58]. The Lambertian assumption simplifies $\boldsymbol{\rho}(\mathbf{n},\mathbf{l})$ as $\mathbf{D}$ with each row of $\mathbf{D}$ being a constant, so that $\mathbf{I} = \mathbf{S}^\top \mathbf{L}$ and $\mathbf{S}^\top$ encodes the albedo-scaled normals. By applying SVD (or matrix factorization considering missing elements [59]), the pseudo-normal $\tilde{\mathbf{S}}$, which is different from the true albedo-scaled normal $\mathbf{S}$ by a 3×3 ambiguity matrix $\mathbf{A}$, can be decomposed, as shown in the top row of Table 2. Resolving this intrinsic shape-lighting ambiguity requires additional information. It is solved up to a rotation ambiguity, if six points with the same albedo or six lights of the same intensity can be identified [58]. If the surface is integrable [60] or the shadow boundary is observed [61], this ambiguity is reduced to the General Bas-Relief (GBR) ambiguity [62], which is denoted as the matrix $\mathbf{G}$ with only three unknowns. Since then, many algorithms have been developed to further resolve this GBR ambiguity. There are methods relying on analyzing the specular spikes [63, 64], parametric specular reflection [37], isotropic specular reflection [18, 65, 66], interreflection [67], a ring of light sources [68], and a perspective camera [69, 70].

We summarize some recent methods and their constraints for solving the ambiguity matrix $\mathbf{G}$ in Table 2. Alldrin *et al.* [3] (AM07) assume the distribution of albedos in a natural image contains a few dominant values (*i.e.*, a few dominant colors of a scene), so that the GBR transform can be found by minimizing the entropy of this distribution. Shi *et al.* [16] (SM10) automatically identify pixels with

TABLE 2
Summary of uncalibrated photometric stereo constraints and solutions.

Solve $\mathbf{N}$ from $\mathbf{I} = \max\{\boldsymbol{\rho}(\mathbf{n},\mathbf{l}) \circ (\mathbf{N}^\top \mathbf{L}), 0\}$ when $\mathbf{L}$ is unknown	
For Lambertian objects, $\mathbf{I} = \max\{\mathbf{D} \circ (\mathbf{N}^\top \mathbf{L}), 0\} = \mathbf{S}^\top \mathbf{L} = \tilde{\mathbf{S}}^\top \mathbf{A}^\top \mathbf{A}^{-1} \tilde{\mathbf{L}} = \hat{\mathbf{S}}^\top \mathbf{G}^\top \mathbf{G}^{-1} \hat{\mathbf{L}}$	
AM07 [3]	$\mathbf{D}$ has only a few different albedos, *i.e.*, the rows of $\mathbf{S}$ have only a few different lengths
SM10 [16]	Several surface points have equal albedo, *i.e.*, several rows of $\mathbf{S}$ having equal length are identified
PF14 [17]	Several points with locally maximum intensity on a Lambertian surface, *i.e.*, points with $\mathbf{n} = \mathbf{l}$ are identified
WT13 [18]	$\boldsymbol{\rho}(\mathbf{n},\mathbf{l}) \approx \mathbf{D} + \boldsymbol{\rho}_s(\theta_h, \theta_d)$, *i.e.*, the specular reflection depends only on $\{\theta_h, \theta_d\}$

the same albedo by chromaticity clustering to solve $\mathbf{G}$. Papadhimitri and Favaro [17] (PF14) propose to identify local diffuse reflectance (LDR) maxima, where the normal vector is coincident with the lighting direction. Wu and Tan [18] (WT13) assume an additive bivariate specular reflection to solve the GBR ambiguity. SM10 [16] and PF14 [17] need to select desired pixels (or lighting directions), but their solutions for $\mathbf{G}$ are quite simple, *i.e.*, a linear problem for SM10 [16] and a close-form solution for PF14 [17]. AM07 [3] and WT13 [18] seek the optimal three unknowns in $\mathbf{G}$ by a brute-force search.

Manifold embedding based methods [19, 55–57] can also work with uncalibrated lighting conditions. They first recover normals up to a global rotation ambiguity, which is removed by boundary or integrability constraint. In particular, the method in [19] (LM13) directly handles general isotropic BRDFs by mapping the distances of intensity profiles to angular differences of normals, and its extension [56] improves the accuracy of elevation angle estimation by BRDF symmetry based on the azimuth angle estimation from [19]; both methods cannot be described using the framework in Table 2.

2.3 Beyond the classic data capture setup

Point lighting is a more precise model and its lighting direction varies spatially. Such a model is often used for a nearby light source, with the lighting intensity fall-off being considered [71, 72]. This model is adopted in [73] to build a handheld photometric stereo device. The LCD monitor can be used as the light source for photometric stereo by assuming an extended near point lighting model [74, 75]. Involving the near point lighting model will make the problem becomes nonlinear, but a recent work formulates the problem in a mesh deformation framework to avoid conventional integration and nonlinear optimization [76].

General lighting (or environment lighting) based methods allow photometric stereo to work in less constrained lighting conditions. Environment lighting could be represented as a spherical function. Basri *et al.* [4] take spherical harmonics representation and extend the factorization based photometric stereo method [58] to arbitrary lighting and solve the problem up to four or nine dimensional linear ambiguity. Such a spherical representation is adopted by many methods [77–80] that study photometric stereo in general lighting. Environment lighting can also be represented as a sum of many directional lights and calibrated by a mirror sphere [81, 82]. General lighting model is often employed by outdoor photometric stereo problems. Jung *et al.* [83] use quadratic skylight plus a Gaussian sunlight, which is more general than considering only the sunlight as a directional lighting [39, 84]. Hold-Geoffroy *et al.* [85, 86] analyze the condition to apply outdoor photometric stereo

with a lighting model considering illumination from sun, sky, and clouds.

Perspective camera model is a more precise camera model. A perspective camera model will improve photometric stereo at the cost of nonlinear optimization [5] or solving differential equation for Lambertian [87] and non-Lambertian materials [88]. Perspective model is also useful in building a handheld photometric stereo camera [73], binocular photometric stereo [89], and auto-calibrating the lighting conditions [69, 70].

Cameras with non-linear response distort the scene radiance values. While most of the methods apply a separate radiometric calibration to correct this distortion, some methods [16, 39, 80, 84] include automatic radiometric calibration. Reflectance monotonicity [13], symmetry [43], manifold embedding [55, 57], and example-based methods [51–53] are naturally unaffected by the nonlinear camera response. Thus, radiometric calibration is not required for these methods.

Color lighting based methods simplify data capture by treating different color channels as independent images. Under separate red, green, and blue lighting, a single image allows Lambertian photometric stereo to be applied for dynamic shape reconstruction [90–93].

Depth prior from other sensors (*e.g.*, the Kinect sensor) can correct the low-frequency shape distortion [94] often associated with the results from photometric stereo. Fusing depth and normal produces high quality surface reconstructions [95–98], even if the reflectance and illumination information are not very accurate [99].

Multiple viewpoints provide useful constraints in various aspects. Depth estimates from binocular stereo can be fused with photometric stereo [100]. While single-view methods only recover a normal map, multi-view photometric stereo methods reconstruct a complete 3D model [6, 101–103]. The rough shape from multi-view stereo can calibrate directional lighting [6, 7, 103–105], near point lighting [73], or general lighting [80] for Lambertian surfaces. For arbitrary isotropic surfaces, Zhou *et al.* [102] fuse the iso-depth contours reconstructed by [43] from multiple viewpoints. Simultaneous shape and reflectance recovery is more tractable given multi-view data for directional lighting [102, 106], general lighting [107], or polarized spherical gradient illumination [108–110].

Object motion provides stronger constraint than multi-view data when illumination condition is fixed, since the angle between lighting and surface normal varies under object motion. The 3D shape can be recovered from object motion with as few as two frames [111]. With more frames, it can calibrate the directional lighting [7, 112] and general lighting [77] for Lambertian surfaces, and directional lighting for general BRDFs [113].

Special applications of photometric stereo which are beyond all above categories are summarized here. Photometric stereo can be extended to multi-spectral domain [114] by exploring the wavelength dependence of material reflectance. Photometric stereo can also be extended to deal with more complicated materials such as translucent objects by combining with deconvolution technique [115, 116] or fluorescent materials for either calibrated [117] or uncalibrated [118] lighting condition. By compensating backward scattering of murky media [119], photometric stereo works for objects immersed in murky water. By using a high-speed video camera and synchronized lighting [120], photometric stereo works in real time. Photometric stereo can also be combined with shape from polarization to complement the weakness of each other in either calibrated or uncalibrated setup [121].

2.4 Number of input images

We further categorize photometric stereo methods according to their number of input images. The basic trend is the more general assumption and complicated problem require more images as input.

A small number of images (at least three) are required for classic photometric stereo [1]. By regarding the RGB channels as three independent images, a single color image under independent red, green, and blue lights [90, 91, 93] or natural illumination [122][2] allows photometric stereo. More input images help to enhance the system robustness, such as four-light [23–25, 35, 71] or five-light configuration [32]. Under general lighting, the first order and second order spherical harmonics approximations require at least four [78] or nine images [4, 77]. Many photometric stereo algorithms take more images than the theoretical minimum number for better robustness. Typically 10 - 20 images are used for calibrated methods with shadow analysis [27], analytic BRDF models [2, 37, 38, 123], near point lighting [71, 72, 74, 76], and general lighting [81, 82]. Most uncalibrated methods under directional lighting [3, 16–18, 29, 44, 63, 64, 67, 69] or near point lighting [61, 70] also take 10 - 20 input images.

A medium number of images (about 50 - 100) are required for many outlier rejection based methods [9, 10, 28, 30, 31, 33]. Most methods handling general BRDFs also take about 50 - 100 images [11–15, 19, 43, 47]. Multi-view methods naturally take more images since they need to capture images from different viewpoints. Lambertian multi-view photometric stereo methods [6, 73] often take 50 - 100 images to cover more than 36 viewpoints.

A large number of images (about 500 - 1000) are required for outdoor photometric stereo [39, 84], manifold embedding based methods [55, 57], handling anisotropic BRDFs [50], transparent object by observing specular spike [41], or outliers using statistical methods [26, 34]. Some multi-view photometric stereo methods also take such a number of images to deal with dynamic shapes [101], intricate geometries [103], or non-Lambertian materials [53, 102]. If considering all frames taken by a video camera as input images, some methods relying on object motion [7, 112] also take about 1000 images.

2. This work is claimed as a shape from shading method since it only needs one image, but its formulation is indeed photometric stereo by treating three channels as three images.

3 PHOTOMETRIC STEREO DATASET

In this section, we introduce the 'DiLiGenT' dataset to benchmark photometric stereo algorithms under classic data capture setup (directional lighting, fixed and orthographic camera) for non-Lambertian and uncalibrated methods. There are a few publicly available dataset for Lambertian[3,4] and non-Lambertian photometric stereo[5] without ground truth shapes; the point feature dataset in [124] and multi-view stereo dataset in [125, 126] also contain images under varying lighting conditions and they have scanned shapes provided. But none of existing datasets are suitable for benchmarking non-Lambertian photometric stereo, because most of their objects are simple in reflectance and the number of different lighting conditions is small. This motivates us to create a new dataset with objects of various shapes and BRDFs, carefully calibrated lighting conditions, and 'ground truth' 3D shapes (normals).

3.1 Overview of the dataset

The 'DiLiGenT' main dataset contains ten objects, as shown in the top two rows of Figure 1. In terms of surface BRDFs, it covers materials that are mostly diffuse (CAT) or with a rough surface (POT1), with strong and sparse specular spikes (BALL, READING), with broad and soft specular lobes on uniform (BEAR, BUDDHA) and spatially-varying materials (POT2, GOBLET[6]), and with metallic paint on mostly uniform (COW) and spatially-varying (HARVEST) surfaces. In terms of surface shapes, we have the simple sphere (BALL), smoothly curved surfaces (BEAR, CAT, GOBLET, COW), smooth surfaces with local details (POT1, POT2), surfaces with complicated geometry (BUDDHA, READING), and delicate shapes with concave parts (HARVEST).

Data capture is performed using a Point Grey Grasshopper (GRAS-50S5C-C) camera that has a linear radiometric response function with a $50mm$ lens. All images have a resolution of 2448×2048. The objects are about $20cm$ in diameter and placed about $1.5m$ from the camera to approximate orthographic projection. To avoid interreflection, all images are recorded in a dark room with everything covered by black cloth except the target object. For each lighting condition, we take four images under different exposure times and combine them to obtain a single input HDR image, so that the highly specular observations can also be precisely stored. A typical set of exposure time settings for our data capture is $\{20, 40, 70, 100\}ms$.

According to Section 2.4, most non-Lambertian methods require about 100 differently illuminated images, so we design our light sources as 96 white LED bulbs fixed on a rectangular metal frame. The frame is divided into 12×8 regular samples spanning $74.6° \times 51.4°$, with one LED fixed at each of the grid points. We design the light frame in this way for its portability and efficiency in cost and space.

6. This object is made from one material, but the rust on surface makes its BRDF spatially-varying. We exclude the concave interior of this object for evaluation due to its strong interreflection, which is not modeled in almost all existing photometric stereo algorithms.
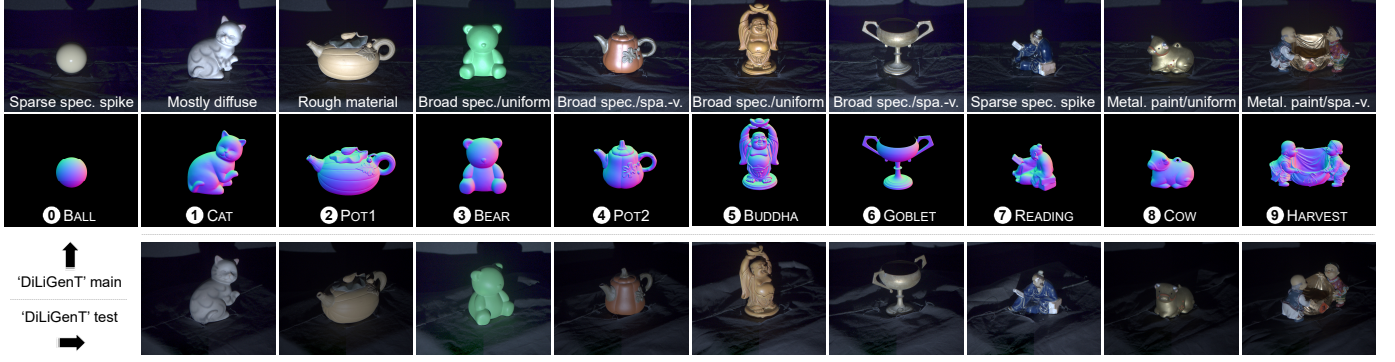
Fig. 1. Samples of photometric stereo images (tone-mapped HDR images) and 'ground truth' normals for ten objects in the 'DiLiGenT' main (top two rows) and sample images in the 'DiLiGenT' test (bottom row) datasets. The texts in brackets indicate the dominant reflectance properties of the objects. Please zoom in the electronic version for better details.

During data capture, the light frame is fixed with the camera at its center.

Test dataset for online benchmarking: We further capture these objects (except for BALL) from a different viewpoint using the same lighting setup, as shown in the bottom row of Figure 1. The images and lighting data are made publicly available for both datasets, while the 'ground truth' normals are not released with the test dataset for the purpose of evaluating newly proposed algorithms. Please refer to our website for procedures of submitting and evaluating new results.

3.2 Lighting condition calibration

Lighting condition calibration includes calibrating both light source intensities and directions. The lighting intensity is calibrated using a Macbeth white balance chart, which is a good approximation of a uniform Lambertian surface. All our HDR images are pre-normalized by the calibrated lighting intensities when evaluating a photometric stereo method.

Conventional lighting direction calibration for photometric stereo assumes directional lighting and an orthographic camera. By identifying the specular point in the image of a mirror sphere and calculating its normal, it is easy to infer the lighting direction with the assumption that the normal at the specular point bisects the angle between viewing direction and lighting direction according to the law of reflection: $\mathbf{l} = 2(\mathbf{n}^\top \mathbf{v})\mathbf{n} - \mathbf{v}$.

The simple approach above cannot meet the high demand of accuracy as a benchmark dataset. By still using a mirror sphere and considering the LED bulbs on a fixed grid as a constraint, we could calibrate their 3D positions simultaneously by assuming the point light source and perspective camera models. With the lighting positions available, the lighting directions for directional lighting assumption are simply calculated by subtracting the sphere center position from the positions of LED bulbs. We use the lighting intensities and directions for the evaluation in Section 4, but we release the point light positions and camera calibration information on our website so that people can also use our dataset for evaluating methods that assume point light source and/or perspective camera models.

The first step for calibration is to estimate the 3D position of the sphere center $\mathbf{C}$[7]. Estimating the sphere center starts from acquiring an image taken under natural illumination. We then manually mark out an ellipse as boundary of the sphere. Pixels on the ellipse correspond to 3D points where the viewing direction is tangent to the sphere. Assuming a pixel on the ellipse has a homogeneous coordinate $\mathbf{x}$, then its corresponding viewing direction is $\mathbf{v} = \mathbf{K}^{-1}\mathbf{x}$, where $\mathbf{K}$ is the camera intrinsic matrix. Given the known radius r of the sphere, the distance from sphere center $\mathbf{C}$ to a line is $\|\mathbf{C} - (\mathbf{v}\mathbf{v}^\top)\mathbf{C}\| = r$. Then $\mathbf{C}$ can be estimated by solving an optimization problem with all pixels on the ellipse considered. The 3D position of specular point on the mirror sphere is denoted as $\mathbf{P}_j$. We then denote the specular point in the captured image as $\mathbf{p}_j$, thus the ray starting from $\mathbf{p}_j$ is $\mathbf{K}^{-1}\mathbf{p}_j$. $\mathbf{P}_j$ is found by intersecting $\mathbf{K}^{-1}\mathbf{p}_j$ with the sphere, whose 3D points are determined by the sphere center $\mathbf{C}$ and radius r.

Solving positions of all LEDs are performed by finding an optimal rigid transformation that maps all LED bulbs from the world coordinate to the camera coordinate. Denoting the positions of all LEDs are $\{\hat{\mathbf{Q}}_1, \hat{\mathbf{Q}}_2, \cdots, \hat{\mathbf{Q}}_{96}\}$ in the world coordinate system, the only set of parameters to be estimated will be a rigid transformation $\{\mathbf{R}, \mathbf{T}\}$ which transforms $\hat{\mathbf{Q}}_j$ to its position $\mathbf{Q}_j$ in the camera coordinate system. We then define the following cost function:

$$f(\mathbf{R}, \mathbf{T}) = \sum_{j=1}^{96} \text{dist}(\mathbf{R}\hat{\mathbf{Q}}_j + \mathbf{T}, \mathbf{l}_j, \mathbf{P}_j). \qquad (1)$$

Here the direction $\mathbf{l}_j$ is determined as the reflection of $\mathbf{K}^{-1}\mathbf{p}_j$ about $\mathbf{n}_{P_j}$, which is the surface normal direction at $\mathbf{P}_j$, and $\text{dist}(\mathbf{Q}, \mathbf{l}, \mathbf{P})$ is the squared distance from the light source $\mathbf{Q}$ to a line passing the point $\mathbf{P}$ with direction $\mathbf{l}$:

$$\text{dist}(\mathbf{Q}, \mathbf{l}, \mathbf{P}) = \|\mathbf{Q} - \mathbf{P} - (\mathbf{l}\mathbf{l}^\top)(\mathbf{Q} - \mathbf{P})\|. \qquad (2)$$

This is a non-linear Least Squares problem that can be solved by Matlab function 'lsqnonlin'. Such a simultaneous estimation to all LED positions is more robust than calibrating each LED one by one, due to the hard constraint from accurately measured positions of all LED bulbs $\{\hat{\mathbf{Q}}_1, \hat{\mathbf{Q}}_2, \cdots, \hat{\mathbf{Q}}_{96}\}$ when designing the light frame.

7. Please refer to the supplementary material for an illustration of the coordinate system and notations.

3.3 Obtaining 'ground truth' shapes

We use the Rexcan CS scanner synchronized with the TA-300 turn table (such a combination is denoted as Rexcan CS+). The scanner is based on structured light technology, which is claimed to produce 3D point cloud with point spacing in the interval of $[0.035, 0.2]mm$ and the surface noise lower than $0.01mm$. We equip the scanner with two $200mm$ lens whose FOV is $150° \times 105° \times 185°$ (width, height, diagonal); such a set of lens allows scanning with the point spacing as $0.097mm$. The turn table has a diameter of $30cm$; the diameters of our objects are around $10cm$ so that they could be placed at the center of the turn table and completely captured within the FOV of the scanner's cameras.

The turn table can perform two-axis movement, which allows it to rotate for $\pm180°$ and swing for $\pm40°$. We program the turn table to move at a step of $20°$ for rotation and a step of $10°$ for swing, so in each round of scanning task the object is scanned for $18 \times 9 = 162$ times automatically. The eZScan 7 software affiliated with the scanner performs an automatic registration to all scans and merge them as a complete mesh. We check histograms of registration error from the scanner's output, and observe that the average distance is usually at the order of $0.01mm$[8]. Considering the sizes of our objects are at the order of $10cm$, the registration error is low enough to provide us a sufficiently accurate shape. We perform spray coating to highly specular objects like (READING) and (HARVEST) before scanning. Finally, we have obtained 3D meshes in which no visible holes or bumps are observed from the frontal view for surface normal calculation.

3.4 Shape-to-image alignment

After the 3D shapes are scanned and repaired, we need to find the shape-to-image alignment that transforms the 3D shape from the scanner coordinate system to the photometric stereo image coordinate system. Such registration is not an easy task, since it requires establishing correspondence between a textureless 3D shape and a 2D image. We cannot access the raw data of the commercial scanner we used for calibration, so that registering the shape and image coordinate system from camera calibration like [124–126] is not feasible. As an alternative, we use the mutual information method [127] in Meshlab, which has been adopted in [128] to perform the similar alignment task. We find such a registration task could be performed in a much more stable way if we replace the 2D image with its normal map, because surface normal consistency is a key feature used by such an alignment method [127]. We provide the normal map estimated by a state-of-the-art method that deals with general BRDFs [14] as the reference image. The camera used to capture photometric stereo images is geometrically calibrated, so its intrinsic parameters are fixed input for the alignment algorithm and only the extrinsic parameters (rotation, translation, and a uniform scaling) are required to be estimated, which makes the automatic alignment much more reliable. Meshlab[9] provides an easy-to-use UI

to operate the 3D shape, so that we can interactively rotate, translate, and scale the 3D shape to a roughly consistent position with the reference normal map. Then we apply the automatic alignment algorithm in [127] until it converges to a stable solution.

We use synthetic experiments to verify the accuracy of such alignment procedure, and find that the automatic alignment converges quite stably in synthetic case. By testing several shapes from the Stanford 3D Scanning Repository, we observe an average angular error between the aligned and true normal maps about $1°$ for a smooth surface like the BUNNY and more than $3°$ for a complicated surface like the HAPPY BUDDHA[10]. However, real data contain noise from camera calibration, lens distortion, and other unexpected sources, we find for objects with large depth variation (in particular POT1, POT2, and BUDDHA), the automatic alignment results are obviously sub-optimal even if we tried our best to provide good initialization. Therefore, we have to manually adjust the 3D shape by carefully rotating and translating it in minimum steps allowed by Meshlab based on the automatic alignment result. Upon finishing a single step of manual operation, we compare the normal map rendered from the aligned shape and the reference normal map by closely checking every small feature on both normal maps, until all parts are aligned with subpixel precision[11]. The code and instruction of our alignment procedure are available on our project website.

4 QUANTITATIVE BENCHMARK RESULTS

Our evaluation criteria is based on the statistics of angular error. For each pixel, the angular error is calculated as $\arccos(\mathbf{n}_0^\top \mathbf{n})$ in degrees, where $\mathbf{n}_0$ and $\mathbf{n}$ are 'ground truth' and estimated normals respectively. In addition to the mean angular error, which is a commonly adopted metric in papers surveyed in Section 2, we also calculate the minimum, maximum, median, the first quartile, and the third quartile of angular errors for each estimated normal map[12].

4.1 Evaluation for non-Lambertian methods

We evaluate and compare non-Lambertian photometric stereo methods summarized in Table 1, *i.e.*, BASELINE [1], WG10 [9], IW12 [10], GC10 [2], AZ08 [11], ST12 [12], HM10 [13], ST14 [14], and IA14 [15]. We choose these methods because they cover most categories of non-Lambertian methods and achieve state-of-the-art performance. For all evaluated methods, we use the parameters provided in the original codes or suggested by the original papers. The evaluation results[13] are summarized in the first and second rows of Figure 2.

The BASELINE [1] method has the largest errors in most data. When sparse outliers could be efficiently removed,

---

8. Please refer to the supplementary material for a picture of the scanning setup and an example of the scanner's output.

9. http://meshlab.sourceforge.net/

10. Please refer to the supplementary materials for complete details of synthetic evaluation.

11. An example of alignment with and without manual adjustment is provided in the supplementary material

12. All evaluations in this section are conducted using the 'DiLiGenT' main dataset, and the evaluation results using the test dataset are in the supplementary material.

13. Estimated normal maps and difference maps w.r.t.'ground truth' for all objects and methods are in the supplementary material.
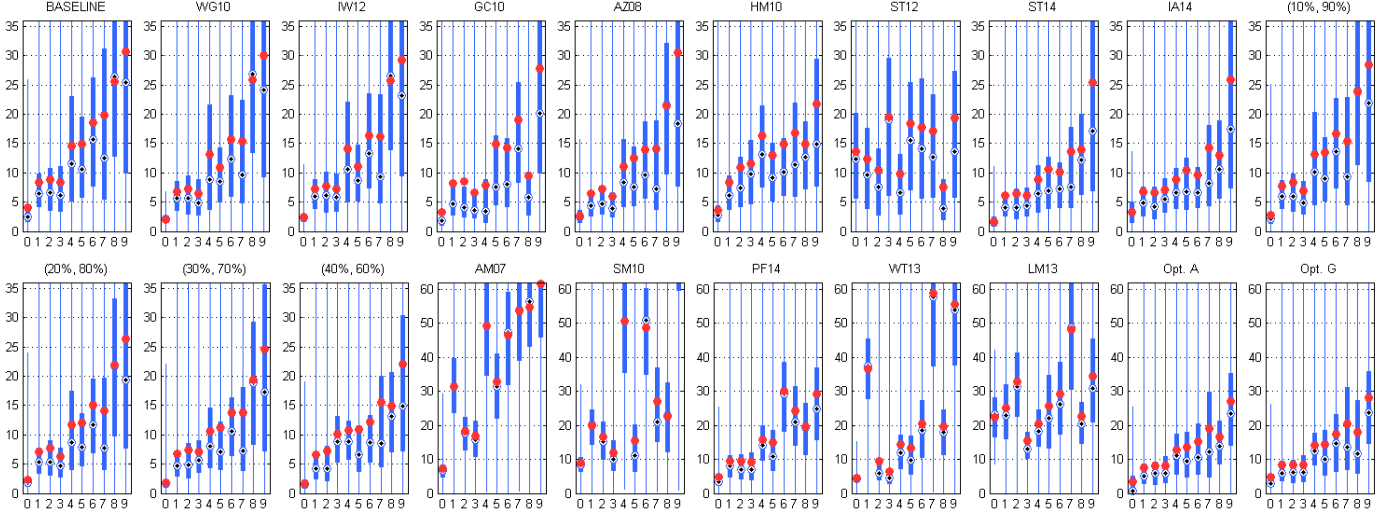
Fig. 2. Benchmark results for calibrated non-Lambertian (from BASELINE to (40%, 60%)) and uncalibrated (from AM07 to Opt. G) photometric stereo: Each subplot shows the results by one evaluated method for each data; the X-axis is the ID of data w.r.t. the main dataset in Figure 1, and the Y-axis is the angular error in degrees; the statistics of angular errors for all pixels per normal map are displayed using the box-and-whisker plot: The red dot indicates the mean value, the black dot is the median, the top and bottom bounds of the blue box indicate the first and third quartile values, and the top and bottom ends of the vertical blue line indicate the minimum and maximum errors.

WG10 [9] and IW12 [10] show improvement on objects including a dominant Lambertian component (BALL, CAT, POT1). But they cannot handle broad and soft specularity well (POT2, COW). Adopting a mixture of the Ward model, GC10 [2] produces superior results on objects with broad and soft specularity (POT2, COW). Its performance drops on objects with complicated spatially-varying BRDFs (GOBLET). Further, the mixture model fitting might suffer from some local minima as indicated by the relatively large errors in some objects (BUDDHA, READING). The bivariate BRDF constraint from AZ08 [11] consistently achieves moderate accuracy on all data, and performs better when the highlight is sparser or weaker (BALL, BEAR). ST12 [12] suffers from large error when the BRDF is not 1D monotonic (BALL, BEAR), but achieves best results among all evaluated methods on one challenging example (COW), where its assumption is satisfied. HM10 [13] works reasonably well for monotonic reflectance described by a single lobe BRDF (GOBLET), but not for the mixture of diffuse and specular components (BEAR, POT2). In general, the two most recent methods ST14 [14] and IA14 [15] show best performance for most data. All evaluated methods show large errors in concave areas (HARVEST), due to frequent cast shadow and interreflection.

A position threshold method: Non-Lambertian phenomena such as shadows and specular highlights are the main problems for calibrated photometric stereo. Since shadows and highlights are often associated with dark and bright pixels, we could simply sort pixels by their intensities and discard shadows and highlights as bottom and top ranked observations respectively. The classic photometric stereo (BASELINE [1]) can be safely applied once these non-Lambertian observations are largely discarded. Such a strategy has been used in many non-Lambertian methods (*e.g.*, WG10 [9], GC10 [2], ST14 [14], IA14 [15]) as initialization.

We experiment this simple position threshold method by setting the position thresholds for shadow and high-light among $T_{low} = \{0\%, 10\%, \cdots, 40\%\}$ and $T_{high} = \{100\%, 90\%, \cdots, 60\%\}$ respectively. The benchmark results are summarized in the second row of Figure 2. It is interesting to note that a narrower interval produces more accurate normal estimates on most of the data. In particular, the tightest thresholding ($T_{low} = 40\%, T_{high} = 60\%$) produces comparable results to state-of-the-art methods ST14 [14] and IA14 [15].

Observations: High-quality photometric stereo for complicated BRDFs (COW), spatially-varying materials (GOBLET), and concave shapes (READING, HARVEST) is still challenging. Methods exploiting general reflectance properties such as symmetry or monotonicity often require far more input images than the BASELINE [1]. Non-Lambertian photometric stereo also involves extensive computation, and all evaluated methods here take much more memory and computation resources than the BASELINE [1]. Computationally efficient non-Lambertian photometric stereo methods using a small number of images are still missing.

4.2 Evaluation for uncalibrated methods

We evaluate and compare recent uncalibrated methods summarized in Table 2, *i.e.*, AM07 [3], SM10 [16], PF14 [17], and WT13 [18], plus a manifold embedding based method LM13 [19]. These methods so far are evaluated by their consistency with calibrated methods. Using our dataset with 'ground truth' shapes, we can directly evaluate the accuracy in surface normal directions. To fairly evaluate and compare AM07 [3], SM10 [16], and PF14 [17], we start from the same GBR-distorted normal map that is generated from the code of PF14 [17]. Unlike these three methods, WT13 [18] applies diffuse-specular separation [129] beforehand and LM13 [19] is not specific to GBR distortion, so we simply take their complete system for evaluation. We summarize the evaluation results in the bottom row of Figure 2.

The uncalibrated methods produce much larger errors than the calibrated ones especially for non-Lambertian ma-

terials. AM07 [3] and SM10 [16] only produce reasonable results on near Lambertian objects (BALL, POT1, BEAR). AM07 [3] assumes piecewise constant albedo, and is sensitive to objects with smoothly varying texture (CAT). SM10 [16] requires albedo based segmentation and hence fails on objects with high-frequency material changes (GOBLET). Though PF14 [17] is designed for Lambertian surfaces, it shows excellent robustness and tolerance to non-Lambertian objects (POT2, BUDDHA, COW). Its average error for all test data is the smallest. The specular reflection based method WT13 [18] outperforms other methods on materials with broad and soft highlights (BEAR, POT2, BUDDHA, COW). Its performance deteriorates when the specular reflection is weak (CAT) or spatially-varying (GOBLET, HARVEST) and it suffers from catastrophic failure when diffuse-specular separation is poor (READING). The performance of LM13 [19] is poor in our evaluation, partially due to its requirement on uniformly distributed directional lighting.

We further evaluate the linearly-distorted and GBR-distorted normals, which are inputs to auto-calibration. We solve the optimal linear transformation $\mathbf{A}$ and GBR transformation $\mathbf{G}$ by registering the distorted normals to the 'ground truth' normals. The mean angular errors between the registered and 'ground truth' normals are reported as Opt. A and Opt. G for linearly- and GBR-distorted normals respectively. Both results show surprisingly large residual errors, which suggests a more robust method is needed for recovering the pseudo-normal $\tilde{\mathbf{S}}$ (up to $\mathbf{A}$) and $\hat{\mathbf{S}}$ (up to $\mathbf{G}$).

Observations: Uncalibrated photometric stereo is still a very difficult problem for real-world objects and a major bottleneck is from the widely adopted factorization step [58] which is designed for Lambertain materials. Pre-processing the data by using outlier rejection methods [9, 10] can only solve the problem in a limited scale, since they need to assume the existence of a dominant Lambertian component. LM13 [19] presents a successful trial to avoid this factorization for general BRDFs, but it is largely limited by the requirement of uniformly distributed lighting. Exploring general BRDF properties and how to accurately estimate the pseudo-normal for non-Lambertian objects could be interesting future topics. In terms of resolving the GBR ambiguity, it might be interesting to combine the strength of diffuse reflection based method (*e.g.*, PF14 [17]) and specular reflection based method (*e.g.*, WT13 [18]).

4.3 Evaluation for varying lighting distribution

We select nine subsets from our complete set of 96 different lightings, as shown on the left of Figure 3, to evaluate how non-Lambertian and uncalibrated photometric stereo perform under varying lighting distribution. Lighting distributions 'A', 'B', and 'C' contain 12 lights with increasing angular variations; small angular variation is desired when designing a compact device, but it also brings larger errors from various noise origins [130]. Lighting distributions 'D', 'E', and 'F' use 36 lights; they are all biased distributions with all lights concentrated in part of the hemisphere. Lighting distributions 'G', 'H', and 'I' use randomly distributed lights with increasing numbers as 12, 36, and 72. We choose to evaluate non-Lambertian methods IW12 [10] (outlier rejection), GC10 [2] (fitting an analytic model), position

thresholdings of $(20\%, 80\%)$ and $(40\%, 60\%)$ (initial results for ST14 [14] and IA14 [15], which show similar performance to their final results) in addition to the BASELINE [1]; for uncalibrated methods we show results of PF14 [17] (best performance) and the ideal case of Opt. A.

We calculate the mean angular error for each object first, and then plot the statistics of these ten mean values per lighting distribution in Figure 3. From the results of calibrated methods, we verify that too small angular variation ('A') and biased lighting distributions ('D', 'E', and 'F') should be avoided for all evaluated photometric stereo methods here. If the number of lights is small, a too narrow threshold ('G') causes obviously large errors, but such a narrow threshold shows best performance when enough lights are given ('I'). To ensure reliable initialization for methods using low-frequency BRDF properties like ST14 [14] and IA14 [15], capturing data with nearly 100 different lightings is recommended, but the requirement on lighting distribution is less important, since 72 random lights show similar performance to 96 lights on a regular grid. If only a dozen of lights are used, fitting an analytic BRDF as GC10 [2] would be the optimal choice, since it takes all measured data into computation. Uncalibrated methods are generally much more sensitive to lighting numbers and distributions. Though for the ideal case of Opt. A only one type of biased distribution causes problems, the state-of-the-art uncalibrated method like PF14 [17] significantly drops in performance when fewer lights are used, especially when the distribution is biased. Increasing the number of lights and maintaining a relatively uniform distribution are important factors to run uncalibrated methods reliably.

4.4 Regional error analysis

We further investigate how normal estimates distribute in different regions that might cause higher angular errors, in particular attached shadow, cast shadow, and interreflection. To identify these regions, we render images with the same camera setup of our data capture using the Mitsuba renderer[14] given our 'ground truth' normal map. The lighting distribution is fixed the same as the 96 directional lights in the dataset throughout all experiments here. For all renderings, we fix the BRDF as the 'roughplastic' material provided in Mitsuba, with 'alpha' (roughness) fixed as 0.2. The attached and cast shadow regions could be easily identified from the rendered direct-only images. We then enable the global illumination by setting the 'maxDepth' as 5 in the path tracer, and the pixels whose differences are greater than 10% of the corresponding direct-only intensities are indicated as interreflection regions.

Given the regional labels (a binary number for each pixel), we introduce the *error inducing rate* caused by attached shadow, cast shadow, and interreflection as the summation of labels across all lighting directions and then normalize them to $[0, 1]$ for each pixel. With a bit abuse of concept, we further introduce the error inducing rate caused by elevation angle of surface normal[15] and minimum half angle (θ_h, see definition in Table 1) across all lightings to further evaluate the influence of pixels near the occluding

---

14. https://www.mitsuba-renderer.org/
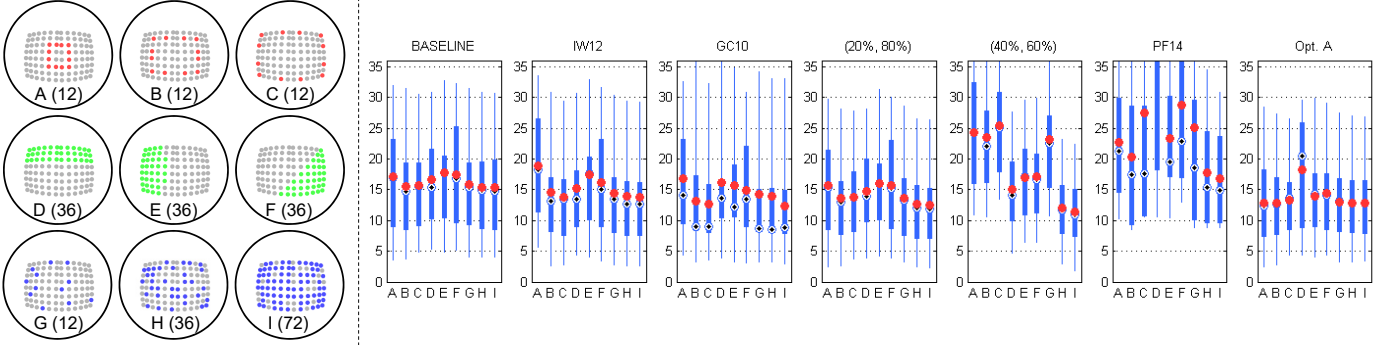15. Varying in the range of $[0, \frac{\pi}{2}]$ from the equator to the north pole.

Fig. 3. Evaluation results for varying lighting distribution: Each subplot (right) shows the statistics of mean angular errors for all ten objects per lighting distribution; the X-axis is the ID of lighting distribution visualized on the left (the camera is at the center of the grid and numbers in parenthesis are the numbers of lights), and the Y-axis is the angular error in degrees.
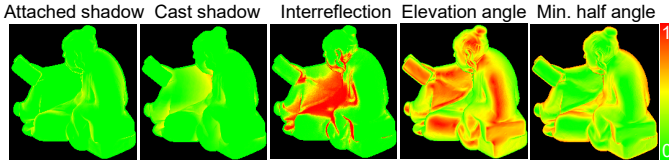

Fig. 4. Visualization of five types of error inducing rates for READING.

boundary and degree of specularity. Both types of angular values in $[0, \frac{\pi}{2}]$ are normalized to $[0, 1]$ for consistency. An example[16] visualization of error inducing rate is provided in Figure 4. Now we can plot the angular errors varying with these five types of error inducing rates for each object (except for BALL); for all the same rate value observed at different pixels, we take the median of their angular errors as the current rate. We also provide the average angular errors of all pixels as a reference. The median curves[17] of all objects from BASELINE [1], WG10 [9], GC10 [2], ST14 [14], position thresholdings of $(40\%, 60\%)$, and PF14 [17] are shown in Figure 5.

Observations: The errors from both shadow and interreflection show a clear increasing trend w.r.t. the corresponding error inducing rates. The difference is that shadows 'drastically' increase the errors, while interreflections show a 'calmly' negative effect, which is a bit against intuition that expects interreflection might be the main cause of larger errors. Elevation angles closing to zero and half angles closing to $\frac{\pi}{2}$[18] include many pixels near the occluding boundary, and they show quite large errors. Mirror-like strong specularity is mainly observed when elevation angles approach top of the sphere (viewpoint) or half angles approach zero (normal bisects the lighting and viewing direction); in these regions, methods assuming dominant Lambertian reflectance like WG10 [9] and PF14 [17] show larger errors, while recent non-Lambertian methods like GC10 [2] and ST14 [14] efficiently deal with such errors susceptible to specularity. It is interesting to note that ro-
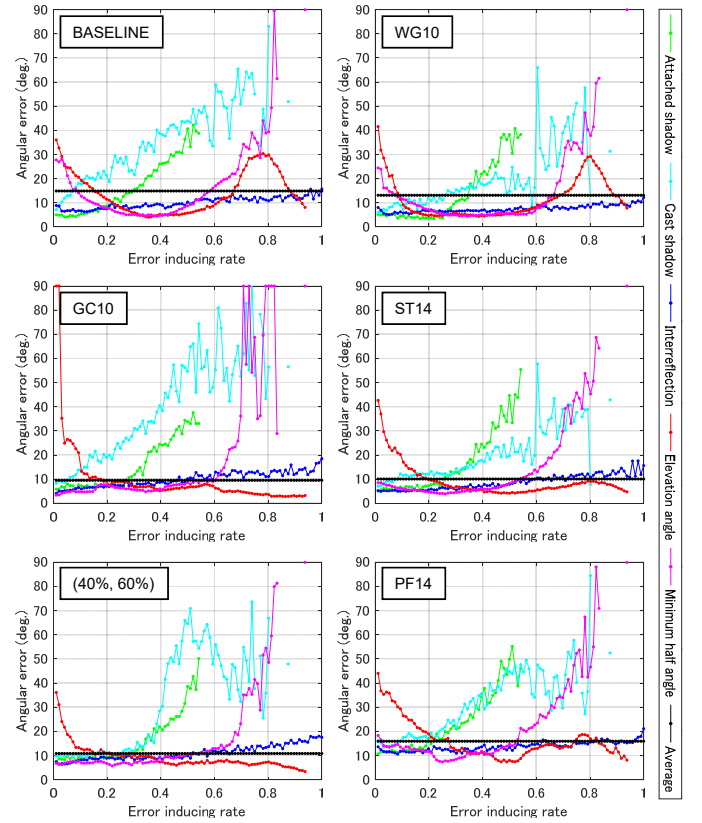
16. Please refer to the supplementary material for the complete results of all objects.

17. The complete results for all objects and methods evaluated in Figure 2 are in the supplementary material.

18. Our light distribution concentrates around the north pole of the sphere (please refer to the left part of Figure 3), so the half vectors **h** also concentrate around top of the sphere.

Fig. 5. Angular error in degrees (Y-axis, median values of all objects) varying with five types of error inducing rates (X-axis); the black horizontal line is the average value of all pixels.

bust methods like WG10 [9] shows slight advantage in reducing the errors near occluding boundary and caused by interreflection. The analysis here reveals that shadows and pixels near occluding boundary worth special treatment to avoid large angular errors, however it might not worth making the BRDF model over-complicated to explicitly deal with interreflection. Existing non-Lambertian methods are quite effective in reducing the impact from specularity, and outlier rejection has unique contribution to lowering errors near boundary or caused by interreflection. Future methods should consider their joint force to achieve even higher performance.

5 CONCLUSION

We surveyed recent photometric stereo methods according to our taxonomy focusing on non-Lambertian and uncalibrated photometric stereo algorithms and presented photometric stereo images dataset with carefully calibrated lighting conditions as well as scanned and registered 'ground truth' shapes (normals). Eight recent non-Lambertian photometric stereo methods and five uncalibrated methods have been quantitatively evaluated and compared using our data.

Challenges: Our evaluation demonstrates that although existing methods can efficiently deal with non-Lambertian materials with sparse or broad speculiarty, metal-like materials are still challenging, which demand more specific reflectance modeling. The pixels frequently falling into shadows and near occluding boundaries often induce larger angular errors than specularity and interreflection, so pertinent solutions to them may benefit in lowering the average error.

Limitations: We have tried to include surfaces with more delicate structures, but the scanned geometry looks more blurred than photometric stereo results[19]. Such a scanned shape cannot be used to evaluate photometric stereo. Though we have tried our best in achieving shape-to-image alignment, our ground truth normal is still not the real ground truth of surface normal measurements.

REFERENCES[20]

[1] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering* 19(1):139–144, 1980, [LCDS]. 1, 2, 3, 5, 7, 8, 9, 10

[2] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and spatially-varying BRDFs from photometric stereo. *IEEE TPAMI* 32(6):1060–1071, 2010, [ACDS]. 1, 2, 3, 5, 7, 8, 9, 10, 11

[3] N. G. Alldrin, S. P. Mallick, and D. J. Kriegman. Resolving the generalized bas-relief ambiguity by entropy minimization. In *Proc. CVPR*, 2007, [LUDS]. 1, 3, 4, 5, 8, 9

[4] R. Basri, D. Jacobs, and I. Kemelmacher. Photometric stereo with general, unknown lighting. *IJCV* 72(3):239–257, 2007, [LUGS]. 1, 4, 5

[5] A. Tankus and N. Kiryati. Photometric stereo under perspective projection. In *Proc. ICCV*, 2005, [LCDS-PC]. 1, 4

[6] C. Hernández, G. Vogiatzis, and R. Cipolla. Multiview photometric stereo. *IEEE TPAMI* 30(3):548–554, 2008, [LUDM-MV]. 1, 4, 5

[7] L. Zhang, B. Curless, A. Hertzmann, and S. M. Seitz. Shape and motion under varying illumination: Unifying structure from motion, photometric stereo, and multiview stereo. In *Proc. ICCV*, 2003, [LUDL-MV-OM]. 1, 4, 5

[8] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Proc. CVPR*, 2006. 1

[9] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *Proc. ACCV*, 2010, [RCDM]. 1, 2, 3, 5, 7, 8, 9, 10

[10] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In *Proc. CVPR*, 2012, [RCDM]. 2, 3, 5, 7, 8, 9

[11] N. G. Alldrin, T. Zickler, and D. J. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *Proc. CVPR*, 2008, [BCDM]. 3, 5, 7, 8

[12] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi. Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances. In *Proc. ECCV*, 2012, [BCDM]. 3, 7, 8

[13] T. Higo, Y. Matsushita, and K. Ikeuchi. Consensus photometric stereo. In *Proc. CVPR*, 2010, [BCDM-NL]. 3, 4, 7, 8, 11

[14] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi. Bi-polynomial modeling of low-frequency reflectances. *IEEE TPAMI* 36(6):1078–1091, 2014, [BCDM]. 3, 7, 8, 9, 10

[15] S. Ikehata and K. Aizawa. Photometric stereo using constrained bivariate regression for general isotropic surfaces. In *Proc. CVPR*, 2014, [BCDM]. 1, 3, 5, 7, 8, 9

[16] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan. Self-calibrating photometric stereo. In *Proc. CVPR*, 2010, [LUDS-NL]. 1, 3, 4, 5, 8, 9

[17] T. Papadhimitri and P. Favaro. A closed-form, consistent and robust solution to uncalibrated photometric stereo via local diffuse reflectance maxima. *IJCV* 107(2):139–154, 2014, [LUDS]. 4, 8, 9, 10

[18] Z. Wu and P. Tan. Calibrating photometric stereo by holistic reflectance symmetry analysis. In *Proc. CVPR*, 2013, [BUDS]. 3, 4, 5, 8, 9

[19] F. Lu, Y. Matsushita, I. Sato, T. Okabe, and Y. Sato. Uncalibrated photometric stereo for unknown isotropic reflectances. In *Proc. CVPR*, 2013, [BUDM]. 1, 3, 4, 5, 8, 9

[20] B. Shi, Z. Wu, Z. Mo, D. Duan, S.-K. Yeung, and P. Tan. A benchmark dataset and evaluation for non-Lambertian and uncalibrated photometric stereo. In *Proc. CVPR*, 2016. 2

[21] S. Herbort and C. Wöhler. An introduction to image-based 3D surface reconstruction and a survey of photometric stereo methods. *3D Research* 2(3):1–17, 2011. 2

[22] J. Ackermann and M. Goesele. A survey of photometric stereo techniques. *Foundations and Trends in Computer Graphics and Vision* 9(3-4):149–254, 2015. 2

[23] E. N. Coleman and R. Jain. Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry. *CGIP* 18(4):309–328, 1982, [RCDS]. 2, 5

[24] F. Solomon and K. Ikeuchi. Extracting the shape and

19. An example is provided in the supplementary material

20. The captical letters in brackets are photometric stereo category labels defined as: The first letter indicates reflectance model (Section 2.1): **L** – Lambert's model, **R** – Lambert's model + outliers (solved by Robust methods), **A** – Analytic non-Lambertian model, and **B** – General properties of BRDF; the second letter is lighting calibration requirement (Section 2.2): **C** – Calibrated and **U** – Uncalibrated; the third letter is lighting model (Section 2.3): **D** – Directional lighting, **P** – Point lighting, and **G** – General (environment) lighting; and the fourth letter is number of input images (Section 2.4): **S** – Small, **M** – Medium, and **L** – Large. Other properties (Section 2.3) are labeled after the hyphen as: **PC**: Perspective Camera, **NL**: Non-Linear camera response, **CL**: Color Lighting, **DP**: Depth Prior, **MV**: Multi-View setup, **OM**: Object Motion, and **SP**: SPecial applications.

roughness of specular lobe objects using four light photometric stereo. *IEEE TPAMI* 18(4):449–454, 1996, [RCDS].

[25] S. Barsky and M. Petrou. The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows. *IEEE TPAMI* 25(10):1239–1252, 2003, [RCDS]. 2, 5

[26] T.-P. Wu, K.-L. Tang, C.-K. Tang, and T.-T. Wong. Dense photometric stereo: A Markov random field approach. *IEEE TPAMI* 28(11):1830–1846, 2006, [RCDL]. 2, 5

[27] M. Chandraker, S. Agarwal, and D. J. Kriegman. Shadow-Cuts: Photometric stereo with shadows. In *Proc. CVPR*, 2007, [RCDS]. 2, 5

[28] Y. Mukaigawa, Y. Ishii, and T. Shakunaga. Analysis of photometric factors based on photometric linearization. *JOSA A* 24(10):3326–3334, 2007, [RCDM]. 2, 5

[29] K. Sunkavalli, T. Zickler, and H. Pfister. Visibility subspaces: Uncalibrated photometric stereo with shadows. In *Proc. ECCV*, 2010, [RUDS]. 2, 5

[30] F. Verbiest and L. Van Gool. Photometric stereo with coherent outlier handling and confidence estimation. In *Proc. CVPR*, 2008, [RCDM]. 2, 5

[31] C. Yu, Y. Seo, and S. W. Lee. Photometric stereo from maximum feasible lambertian reflections. In *Proc. ECCV*, 2010, [RCDM]. 2, 5

[32] D. Miyazaki, K. Hara, and K. Ikeuchi. Median photometric stereo as applied to the segonko tumulus and museum objects. *IJCV* 86(2):229–242, 2010, [RCDS]. 2, 5

[33] D. Miyazaki and K. Ikeuchi. Photometric stereo under unknown light sources using robust SVD with missing data. In *Proc. ICIP*, 2010, [LUDM]. 2, 5

[34] T.-P. Wu and C.-K. Tang. Photometric stereo via expectation maximization. *IEEE TPAMI* 32(3):546–560, 2010, [RCDL]. 2, 5

[35] T. Zickler, S. P. Mallick, D. J. Kriegman, and P. N. Belhumeur. Color subspaces as photometric invariants. *IJCV* 79(1):13–30, 2008, [RCDS]. 2, 5

[36] S. Tozza, R. Mecca, M. Duocastella, and A. D. Bue. Direct differential photometric stereo shape recovery of diffuse and specular surfaces. *Journal of Mathematical Imaging and Vision* 56:57–76, 2016, [ACDS]. 2

[37] A. S. Georghiades. Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo. In *Proc. ICCV*, 2003, [AUDS]. 2, 3, 5

[38] H. S. Chung and J. Jia. Efficient photometric stereo on glossy surfaces with wide specular lobes. In *Proc. CVPR*, 2008, [ACDS]. 2, 5

[39] J. Ackermann, F. Langguth, S. Fuhrmann, and M. Goesele. Photometric stereo for outdoor webcams. In *Proc. CVPR*, 2012, [ACDM-NL]. 2, 4, 5

[40] T. Chen, M. Goesele, and H.-P. Seidel. Mesostructure from specularity. In *Proc. CVPR*, 2006, [ACDM]. 2

[41] S.-K. Yeung, T.-P. Wu, C.-K. Tang, T. F. Chan, and S. Osher. Normal estimation of a transparent object using a video. *IEEE TPAMI* 37(4):890–897, 2015, [ACDL]. 2, 5

[42] L. Chen, Y. Zheng, B. Shi, A. Subpa-Asa, and I. Sato. A microfacet-based refectance model for photometric stereo with highly specular surfaces. In *Proc. ICCV*, 2017, [ACDM]. 2

[43] N. G. Alldrin and D. J. Kriegman. Toward reconstructing surfaces with arbitrary isotropic reflectance: A stratified photometric stereo approach. In *Proc. ICCV*, 2007, [BCDM-NL]. 3, 4, 5

[44] P. Tan, L. Quan, and T. Zickler. The geometry of reflectance symmetries. *IEEE TPAMI* 33(12):2506–2520, 2011, [BUDS]. 3, 5

[45] M. Chandraker, J. Bai, and R. Ramamoorthi. On differential photometric reconstruction for unknown, isotropic BRDFs. *IEEE TPAMI* 35(12):2941–2955, 2013, [BUDS]. 3

[46] W. Matusik and M. Brand. A data-driven reflectance model. *ACM TOG (Proc. SIGGRAPH)* 22(3):759–769, 2003. 3

[47] Z. Hui and A. C. Sankaranarayanan. A dictionary-based approach for estimating shape and spatially-varying reflectance. In *Proc. ICCP*, 2015, [BCDM]. 3, 5

[48] F. Romeiro, Y. Vasilyev, and T. Zickler. Passive reflectometry. In *Proc. ECCV*, 2008. 3

[49] S. Li and B. Shi. Photometric stereo for general isotropic reflectances by spherical linear interpolation. *Optical Engineering* 54(8):083104, 2015, [BCDM]. 3

[50] M. Holroyd, J. Lawrence, G. Humphreys, and T. Zickler. A photometric approach for estimating normals and tangents. *ACM TOG (Proc. SIGGRAPH Asia)* 27(5):133, 2008, [BCDL]. 3, 5

[51] A. Hertzmann and S. M. Seitz. Example-based photometric stereo: Shape reconstruction with general, varying BRDFs. *IEEE TPAMI* 27(8):1254–1264, 2005, [BUDS-NL]. 3, 4

[52] J. Ackermann, M. Ritz, A. Stork, and M. Goesele. Removing the example from example-based photometric stereo. In *Proc. ECCV Workshop*, 2010, [BUDM-NL-MV].

[53] J. Ackermann, F. Langguth, S. Fuhrmann, A. Kuijper, and M. Goesele. Multi-view photometric stereo by example. In *Proc. 3DV*, 2014, [BUDL-NL-MV]. 3, 4, 5

[54] M. K. Johnson, F. Cole, A. Raj, and E. H. Adelson. Microgeometry capture using an elastomeric sensor. *ACM TOG (Proc. SIGGRAPH)* 30(4):46:1–46:8, July 2011, [BCPS]. 3

[55] I. Sato, T. Okabe, Q. Yu, and Y. Sato. Shape reconstruction based on similarity in radiance changes under varying illumination. In *Proc. ICCV*, 2007, [AUDL-NL]. 3, 4, 5

[56] F. Lu, I. Sato, and Y. Sato. Uncalibrated photometric stereo based on elevation angle recovery from BRDF symmetry of isotropic materials. In *Proc. CVPR*, 2015, [BUDM]. 3, 4

[57] T. Okabe, I. Sato, and Y. Sato. Attached shadow coding: Estimating surface normals and lighting conditions. In *Proc. ICCV*, 2009, [BUDL-NL]. 3, 4, 5

[58] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *JOSA A* 11(11):3079, 1994, [LUDS]. 3, 4, 9

[59] A. D. Bue, J. Xavier, L. Agapito, and M. Paladini. Bilinear factorization via augmented Lagrange multipliers. In *Proc. ECCV*, 2010, [LUGS]. 3

[60] A. L. Yuille and D. Snow. Shape and albedo from multiple images using integrability. In *Proc. CVPR*, 1997, [LUDS]. 3

[61] D. J. Kriegman and P. N. Belhumeur. What shadows reveal about object structure. *JOSA A* 18(8), 2001, [LUPS]. 3, 5

[62] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille. The bas-relief ambiguity. *IJCV* 35(1):33–44, 1999, [LUDS]. 3

[63] O. Drbohlav and R. Sara. Specularities reduce ambiguity of uncalibrated photometric stereo. In *Proc. ECCV*, 2002, [AUDS]. 3, 5

[64] O. Drbohlav and M. Chantier. Can two specular pixels calibrate photometric stereo? In *Proc. ICCV*, 2005, [AUDS]. 3, 5

[65] P. Tan, S. Mallick, L. Quan, D. Kriegman, and T. Zickler. Isotropy, reciprocity and the generalized bas-relief ambiguity. In *Proc. CVPR*, 2007, [BUDM]. 3

[66] P. Tan and T. Zickler. A projective framework for radiometric image analysis. In *Proc. CVPR*, 2009, [BUDM]. 3

[67] M. Chandraker, F. Kahl, and D. J. Kriegman. Reflections on the generalized bas-relief ambiguity. In *Proc. CVPR*, 2005, [LUDS]. 3, 5

[68] Z. Zhou and P. Tan. Ring-light photometric stereo. In *Proc. ECCV*, 2010, [LUDS]. 3

[69] T. Papadhimitri and P. Favaro. A new perspective on uncalibrated photometric stereo. In *Proc. CVPR*, 2013, [LUDS-PC]. 3, 4, 5

[70] T. Papadhimitri and P. Favaro. Uncalibrated near-light photometric stereo. In *Proc. BMVC*, 2014, [LUPS-PC]. 3, 4, 5

[71] Y. Iwahori, H. Sugie, and N. Ishii. Reconstructing shape from shading images under point light source illumina-

tion. In *Proc. ICPR*, 1990, [LCPS]. 4, 5

[72] J. J. Clark. Active photometric stereo. In *Proc. CVPR*, 1992, [LCPS]. 4, 5

[73] T. Higo, Y. Matsushita, N. Joshi, and K. Ikeuchi. A hand-held photometric stereo camera for 3-D modeling. In *Proc. ICCV*, 2009, [LUPM-PC-MV]. 4, 5

[74] G. Schindler. Photometric stereo via computer screen lighting for real-time surface reconstruction. In *Proc. 3DPVT*, 2008, [LUPS]. 4, 5

[75] J. J. Clark. Photometric stereo using LCD displays. *IVC* 28(4):704–714, 2010, [LCPS]. 4

[76] W. Xie, C. Dai, and C. C. L. Wang. Photometric stereo with near point lighting: A solution by mesh deformation. In *Proc. CVPR*, 2015, [LCPS]. 4, 5

[77] D. Simakov, D. Frolova, and R. Basri. Dense shape reconstruction of a moving object under arbitrary, unknown lighting. In *Proc. ICCV*, 2003, [LUGS-OM]. 4, 5

[78] C. Chen and C. Chen. The 4-source photometric stereo under general unknown lighting. In *Proc. ECCV*, 2006, [LUGS]. 5

[79] L. Shen and P. Tan. Photometric stereo and weather estimation using Internet images. In *Proc. CVPR*, 2009, [LUGM].

[80] B. Shi, K. Inose, Y. Matsushita, P. Tan, S.-K. Yeung, and K. Ikeuchi. Photometric stereo using Internet images. In *Proc. 3DV*, 2014, [LUGM-NL]. 4

[81] L.-F. Yu, S.-K. Yeung, Y.-W. Tai, D. Terzopoulos, and T. F. Chan. Outdoor photometric stereo. In *Proc. ICCP*, 2013, [LCGS]. 4, 5

[82] C.-H. Hung, T.-P. Wu, Y. Matsushita, L. Xu, J. Jia, and C.-K. Tang. Photometric stereo in the wild. In *Proc. WACV*, 2015, [LCGS-SP]. 4, 5

[83] J. Jung, J.-Y. Lee, and I. S. Kweon. One-day outdoor photometric stereo via skylight estimation. In *Proc. CVPR*, 2015, [LUGS]. 4

[84] A. Abrams, C. Hawley, and R. Pless. Heliometric stereo: Shape from sun position. In *Proc. ECCV*, 2012, [LCDL-NL]. 4, 5

[85] Y. Hold-Geoffroy, J. Zhang, P. F. U. Gotardo, and J.-F. Lalonde. What is a good day for outdoor photometric stereo? In *Proc. ICCP*, 2015, [LCGM]. 4

[86] Y. Hold-Geoffroy, J. Zhang, P. F. U. Gotardo, and J.-F. Lalonde. x-hour outdoor photometric stereo. In *Proc. 3DV*, 2015, [LCGM]. 4

[87] R. Mecca, A. Tankus, A. Wetzler, and A. M. Bruckstein. A direct differential approach to photometric stereo with perspective viewing. *SIAM Journal on Imaging Sciences* 7(2):579–612, 2014, [LCDS-PC]. 4

[88] R. Mecca, Y. Quéau, F. Logothetis, and R. Cipolla. A single lobe photometric stereo approach for heterogeneous material. *SIAM Journal on Imaging Sciences* 9(4):1858–1888, 2016, [ACPS-PC]. 4

[89] H. Du, D. B. Goldman, and S. M. Seitz. Binocular photometric stereo. In *Proc. BMVC*, 2011, [LCDS-PC-MV]. 4

[90] C. Hernández, G. Vogiatzis, G. J. Brostow, B. Stenger, and R. Cipolla. Non-rigid photometric stereo with colored light. In *Proc. ICCV*, 2007, [LCDS-CL]. 4, 5

[91] H. Kim, B. Wilburn, and M. Ben-Ezra. Photometric stereo for dynamic surface orientations. In *Proc. ECCV*, 2010, [LCDS-CL]. 5

[92] G. Fyffe, X. Yu, and P. Debevec. Single-shot photometric stereo by spectral multiplexing. In *Proc. ICCP*, 2011, [LCDS-CL].

[93] R. Anderson, B. Stenger, and R. Cipolla. Color photometric stereo for multicolored surfaces. In *Proc. ICCV*, 2011, [LCDS-CL]. 4, 5

[94] D. Nehab, S. Rusinkiewicz, J. Davis, and R. Ramamoorthi. Efficiently combining positions and normals for precise 3D geometry. *ACM TOG (Proc. SIGGRAPH)* 24(3):536–543, July 2005, [LCDS]. 4

[95] Z. Lu, Y.-W. Tai, F. Deng, M. Ben-Ezra, and M. S. Brown.

[96] A 3D imaging framework based on high-resolution photometric-stereo and low-resolution depth. *IJCV* 102(1-3), 2013, [LCDS-DP]. 4

[96] Q. Zhang, M. Ye, R. Yang, Y. Matsushita, B. Wilburn, and H. Yu. Edge-preserving photometric stereo via depth fusion. In *Proc. CVPR*, 2012, [LCDS-DP].

[97] S. M. Haque, A. Chatterjee, and V. M. Govindu. High quality photometric reconstruction using a depth camera. In *Proc. CVPR*, 2014, [LUDS-DP-MV].

[98] A. Chatterjee and V. M. Govindu. Photometric refinement of depth maps for multi-albedo objects. In *Proc. CVPR*, 2015, [LUDS-DP]. 4

[99] T. Okatani and K. Deguchi. Optimal integration of photometric and geometric surface measurements using inaccurate reflectance/illumination knowledge. In *Proc. CVPR*, 2012, [BCDS-DP]. 4

[100] K. Ikeuchi. Determining a depth map using a dual photometric stereo. *IJRR* 6(1):15–31, 1987, [LCDS-MV]. 4

[101] D. Vlasic, P. Peers, I. Baran, P. Debevec, J. Popović, S. Rusinkiewicz, and W. Matusik. Dynamic shape capture using multi-view photometric stereo. *ACM TOG (Proc. SIGGRAPH Asia)* 28(5):174, 2009, [LCDL-MV]. 4, 5

[102] Z. Zhou, Z. Wu, and P. Tan. Multi-view photometric stereo with spatially varying isotropic materials. In *Proc. CVPR*, 2013, [BCDL-MV]. 4, 5

[103] J. Park, S. N. Sinha, Y. Matsushita, Y.-W. Tai, and I. S. Kweon. Multiview photometric stereo using planar mesh parameterization. In *Proc. ICCV*, 2013, [LUDL-MV]. 4, 5

[104] N. Joshi and D. J. Kriegman. Shape from varying illumination and viewpoint. In *Proc. ICCV*, 2007, [LUDS-MV].

[105] R. Sabzevari, A. Del Bue, and V. Murino. Multi-view photometric stereo using semi-isometric mappings. In *Proc. 3DIMPVT*, 2012, [LUDL-MV]. 4

[106] M. Chandraker. What camera motion reveals about shape with unknown BRDF. In *Proc. CVPR*, 2014, [BUDS-PC-MV]. 4

[107] G. Oxholm and K. Nishino. Multiview shape and reflectance from natural illumination. In *Proc. CVPR*, 2014, [BCGS-MV]. 4

[108] W.-C. Ma, T. Hawkins, P. Peers, C.-F. Chabert, M. Weiss, and P. Debevec. Rapid acquisition of specular and diffuse normal maps from polarized spherical gradient illumination. In *Proc. EGSR*, 2007, [BCGS-MV]. 4

[109] A. Ghosh, G. Fyffe, B. Tunwattanapong, J. Busch, X. Yu, and P. Debevec. Multiview face capture using polarized spherical gradient illumination. *ACM TOG (Proc. SIGGRAPH Asia)* 30(6):129, 2011, [BCGL-MV].

[110] B. Tunwattanapong, G. Fyffe, P. Graham, J. Busch, X. Yu, A. Ghosh, and P. Debevec. Acquiring reflectance and shape from continuous spherical harmonic illumination. *ACM TOG* 32(4):109, 2013, [BCGL-MV]. 4

[111] R. Basri and D. Frolova. A two-frame theory of motion, lighting and shape. In *Proc. CVPR*, 2008, [LCDS-OM]. 4

[112] J. Lim, J. Ho, M.-H. Yang, and D. J. Kriegman. Passive photometric stereo from motion. In *Proc. ICCV*, 2005, [LUDL-OM]. 4, 5

[113] M. Chandraker, D. Reddy, Y. Wang, and R. Ramamoorthi. What object motion reveals about shape with unknown BRDF and lighting. In *Proc. CVPR*, 2013, [BUDS-PC-OM]. 4

[114] T. Takatani, Y. Matsushita, S. Lin, Y. Mukaigawa, and Y. Yagi. Enhanced photometric stereo with multispectral images. In *Proc. MVA*, 2013, [LCDM-SP]. 5

[115] B. Dong, K. D. Moore, and C. William. Scattering parameters and surface normals from homogeneous translucent materials using photometric stereo. In *Proc. CVPR*, 2014, [BCDS-SP]. 5

[116] C. Inoshita, Y. Mukaigawa, Y. Matsushita, and Y. Yagi. Surface normal deconvolution: Photometric stereo for optically thick translucent objects. In *Proc. ECCV*, 2014, [BCDS-SP]. 5

[117] I. Sato, T. Okabe, and Y. Sato. Bispectral photometric stereo based on fluorescence. In *Proc. CVPR*, 2012, [BCDS-SP]. 5

[118] T. Treibitz, Z. Murez, B. G. Mitchell, and D. J. Kriegman. Shape from fluorescence. In *Proc. ECCV*, 2012, [BUDS-SP]. 5

[119] C. Tsiotsios, M. E. Angelopoulou, T.-K. Kim, and A. J. Davison. Backscatter compensated photometric stereo with 3 Sources. In *Proc. CVPR*, 2014, [LCPS-SP]. 5

[120] T. Malzbender, B. Wilburn, D. Gelb, and B. Ambrisco. Surface enhancement using real-time photometric stereo and reflectance transformation. In *Proc. EGSR*, 2006, [LCDS-SP]. 5

[121] T. T. Ngo, H. Nagahara, and R. Taniguchi. Shape and light directions from shading and polarization. In *Proc. CVPR*, 2015, [LGUS-SP]. 5

[122] M. K. Johnson and E. H. Adelson. Shape estimation in natural illumination. In *Proc. CVPR*, 2011, [LUGS]. 5

[123] S. K. Nayar, K. Ikeuchi, and T. Kanade. Determining shape and reflectance of hybrid surfaces by photometric sampling. *IEEE TRA* 6(4):418–431, 1990, [ACPS]. 5

[124] H. Aanæs, A. L. Dahl, and K. S. Pedersen. Interesting interest points: A comparative study of interest point performance on a unique data set. *IJCV* 97:18–35, 2012. 5, 7

[125] R. Jensen, A. Dahl, G. Vogiatzis, E. Tola, and H. Aanæs. Large scale multi-view stereopsis evaluation. In *Proc. CVPR*, 2014. 5

[126] H. Aanæs, R. R. Jensen, G. Vogiatzis, E. Tola, and A. B. Dahl. Large-scale data for multiple-view stereopsis. *IJCV* 120(2):153–168, 2016. 5, 7

[127] M. Corsini, M. Dellepiane, F. Ponchio, and R. Scopigno. Image-to-geometry registration: A mutual information method exploiting illumination-related geometric properties. *CGF* 28(7):1755–1764, 2009. 7

[128] S. R. Richter and S. Roth. Discriminative shape from shading in uncalibrated illumination. In *Proc. CVPR*, 2015. 7

[129] Y. Sato and K. Ikeuchi. Temporal-color space analysis of reflection. *JOSA A* 11(11):2990–3002, 1994. 8

[130] J. Wang, Y. Matsushita, B. Shi, and A. C. Sankaranarayanan. Photometric stereo with small angular variations. In *Proc. ICCV*, 2015, [LCDS]. 9

Zhipeng Mo received the B.E. degree from Tianjin University in 2014. He is currently pursuing his Ph.D. degree at Singapore University of Technology and Design, under the supervision of Prof. Sai-Kit Yeung and Prof. Boxin Shi. His research interest is photometric methods in computer vision.

Zhe Wu received the B.E. degree from Tsinghua University in China in 2010 and the Ph.D. degree in computer vision from National University of Singapore in 2015. He is currently a vision engineer at DJI Innovations developing autonomous navigation systems for drones.

Dinglong Duan received the B.S. degree in computer science and technology from Zhejiang University in 2014, supervised by Prof. Hujun Bao and Prof. Guofeng Zhang. In 2013, he was with Prof. Ping Tan and Dr. Zhe Wu at National University of Singapore, as a research assistant. He is currently a research and development engineer at Tencent Audio Video Laboratory.

Sai-Kit Yeung is currently an Assistant Professor at the Singapore University of Technology and Design (SUTD), where he leads the Vision, Graphics and Computational Design (VGD) Group. He was also a Visiting Assistant Professor at Stanford University and MIT. Before joining SUTD, he had been a Postdoctoral Scholar in the Department of Mathematics, University of California, Los Angeles (UCLA). He was also a visiting student at the Image Processing Research Group at UCLA in 2008 and at the Image Sciences Institute, University Medical Center Utrecht, the Netherlands in 2007. He received his PhD in Electronic and Computer Engineering from the Hong Kong University of Science and Technology (HKUST) in 2009. He also received a BEng degree (First Class Honors) in Computer Engineering in 2003 and a MPhil degree in Bioengineering in 2005 from HKUST. His research interests include computer vision, computer graphics and computational design.

Boxin Shi is currently an Assistant Professor (1000 Youth Talents Professorship) at Peking University, where he leads the Camera Intelligence Group. Before joining PKU, he did postdoctoral research at MIT Media Lab, Singapore University of Technology and Design, Nanyang Technological University from 2013 to 2016, and worked as a Researcher at the National Institute of Advanced Industrial Science and Technology from 2016 to 2017. He received the B.E. degree from Beijing University of Posts and Telecommunications in 2007, M.E. degree from Peking University in 2010, and Ph.D. degree from the University of Tokyo in 2013. He won the Best Paper Runner-up award at International Conference on Computational Photography 2015. He has served as Area Chairs for ACCV 2018, MVA 2017.

Ping Tan received his Ph.D. degree from the Hong Kong University of Science and Technology in 2007, and his Master and Bachelor degrees from Shanghai Jiao Tong University, China, in 2003 and 2000 respectively. He is currently an Associate Professor at the School of Computing Science at Simon Fraser University (SFU). Before that, he was an Associate Professor at the Department of Electrical and Computer Engineering at the National University of Singapore. He has served as an editorial member of the International Journal of Computer Vision, Computer Graphics Forum, and the Machine Vision and Applications. He has served in the program committees of SIGGRAPH, SIGGRAPH Asia.