From Intensity Profile to Surface Normal: Photometric Stereo for Unknown Light Sources and Isotropic Reflectances

Feng Lu, Yasuyuki Matsushita, Imari Sato, Takahiro Okabe, and Yoichi Sato

Abstract—We propose an uncalibrated photometric stereo method that works with general and unknown isotropic reflectances. Our method uses a pixel intensity profile, which is a sequence of radiance intensities recorded at a pixel under unknown varying directional illumination. We show that for general isotropic materials and uniformly distributed light directions, the geodesic distance between intensity profiles is linearly related to the angular difference of their corresponding surface normals, and that the intensity distribution of the intensity profile reveals reflectance properties. Based on these observations, we develop two methods for surface normal estimation; one for a general setting that uses only the recorded intensity profiles, the other for the case where a BRDF database is available while the exact BRDF of the target scene is still unknown. Quantitative and qualitative evaluations are conducted using both synthetic and real-world scenes, which show the state-of-the-art accuracy of smaller than $10°$ without using reference data and $5°$ with reference data for all 100 materials in MERL database.

Index Terms—Uncalibrated photometric stereo, general reflectance, BRDF, intensity profile

✦

1 INTRODUCTION

P HOTOMETRIC stereo recovers surface normals of a scene from a set of images recorded under varying lighting conditions. A variety of methods have been proposed in the literature under the assumptions of Lambertian reflectance and known lighting directions. These assumptions, however, greatly limit the applicability of photometric stereo, because real-world scenes exhibit more general reflectances and precise calibration of lighting directions is laborious in practice. Therefore, it is desired to develop a technique that works with general surface reflectance and unknown lighting directions simultaneously with good accuracy.

In most photometric stereo approaches, observed intensities are treated as independent measurements rather than *ordered* measurements. The ordered measurements offer additional information that cannot be found in the un-ordered measurements as demonstrated by clustering of surface normals [1]. The strength of the ordered measurements in photometric stereo has been shown by some previous approaches [2], [3], [4], [5]. These previous methods make use of an *intensity profile*, which is a sequence of radiance intensities recorded at a pixel under varying lightings. These methods are shown effective

- F. Lu and Y. Sato are with the Institute of Industrial Science, the University of Tokyo, Tokyo 153-8505, Japan. E-mail: {lufeng,ysato}@iis.u-tokyo.ac.jp
- Y. Matsushita is with Microsoft Research Asia, Beijing 100080, China. E-mail: yasumat@microsoft.com
- I. Sato is with National Institute of Informatics, Tokyo 101-0003, Japan. E-mail: imarik@nii.ac.jp
- T. Okabe is with Kyushu Institute of Technology, Fukuoka 820-8502, Japan. E-mail: okabe@ai.kyutech.ac.jp

Manuscript received xxx.

under individual assumptions: calibrated light directions and Lambertian reflectance assumption for [4], known occluding boundaries [3], and calibrated reference object with the same reflectance property [2].

In this paper, we further exploit the observations on intensity profiles to develop two methods for accurate surface normal estimation from a scene with general and unknown isotropic reflectances observed under unknown lighting directions. Our methods are built upon the notion of *conditional linearity*. In particular, we show that there exist two strong relations: one between the geodesic distance of intensity profiles and the angular difference of surface normals, and the other between the intensity distribution of intensity profiles and surface reflectance property. We investigate these relations quantitatively and develop a technique that recovers surface normals purely from intensity profiles (Sec. 3). We also develop the second method in the case where a database of bi-directional reflectance distribution functions (BRDFs) is available. For this setting, we present a technique that recovers not only surface normals but also unknown light sources with improved accuracy (Sec. 4).

Our methods make the following assumptions: (1) directional light sources of the same intensity are uniformly positioned around the target surface, (2) the test surface is made of a uniform reflectance up to scaling. We begin with these assumptions for development of the method; however, later in this paper, we show that they can be relaxed to some extent. In addition to these assumptions, we assume a common orthographic camera and exclude extreme cases where the test surface contains only a few planar regions. With these assumptions, our method is able to recover surface normals up to a binary convex/concave ambiguity.

Our methods have the following advantages over the previous approaches. First, our methods handle both uncalibrated light sources and general isotropic reflectances simultaneously; therefore, both of the two proposed methods remove the conventional assumptions about known lighting directions and reflectances. Second, by using a MERL BRDF database [6], our second method is able to estimate lighting directions as well as material and surface normals, which has been difficult with previous methods. Third, derived surface normals have only a binary convex/concave ambiguity, while general uncalibrated photometric stereo methods suffer from generalized bas-relief (GBR) ambiguity [7]. Our average normal estimation error evaluated for 100 real-world materials is smaller than $10°$ without BRDF database (the first method) and smaller than $5°$ with BRDF database (the second method).

1.1 Previous work

A variety of previous studies have been conducted to relax the assumptions of the Lambertian reflectance model and known lighting directions in photometric stereo. Non-Lambertian reflectances have been handled either by regarding non-Lambertian components as outliers, or using more general reflectance models. The former class of methods finds non-Lambertian observations in a robust estimation framework using color-cues [8], median filter [9], rank-minimization [10], sparse regression [11], or Markov random field [12]. The latter class of methods studies reflectance properties, such as bilateral symmetry [13], reflective symmetry about the halfway vector [14], isotropy and monotonicity [15], [16], and other reflectance symmetries [17]. There are methods that use complex parametric models, such as ones that use the Ward model [18], [19], Lambertian+specular models [20] and bi-variate BRDF representation [21]. These methods assume that the lighting directions are known.

There are uncalibrated photometric stereo techniques assuming unknown lighting directions. Most of the existing methods attempt to resolve the Generalized Bas-Relief (GBR) ambiguity [7] with the Lambertian model. Various surface properties are used such as diffuse maxima [22], specularity [23], [24], low-dimensional space [25], minimum entropy [26], interreflections [27], color profiles [28], reflectance symmetry [17], holistic reflectance symmetry [29], and certain configurations of the light sources [30]. If considering the perspective projection case, camera calibration can solve the ambiguity [31]. These methods rely on the assumption that the diffuse reflectance component obeys the Lambertian law. Therefore, the applicability to real-world surfaces that have non-Lambertian diffuse components is restricted.

Handling both non-Lambertian reflectances and uncalibrated light sources simultaneously is far more challenging and has been less studied. Silver [32] and Hertzmann and Seitz [2] capture reference objects with known and proper reflectances together with the target
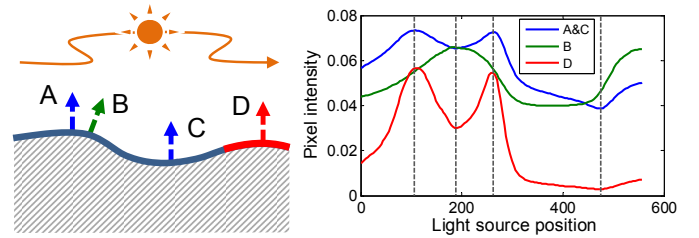

Fig. 1. Illustration of intensity profiles. Surface points A, B, and C have the same reflectance, but D is different. A, C, and D have the same surface normal, while B has a different normal.

scene, while Ren *et al.* [33] captures a well-designed calibration board with multiple reference materials and shapes. These methods use intensity values of reference objects to determine surface normals of the test scene. Because reference objects are captured with the test scene, the effect of uncalibrated light sources is eliminated. Georghiades [34] uses the Torrance-Sparrow model and proposes to optimize over a large set of variables. Chandraker *et al.* [35] recover surface iso-contours from images taken under light sources positioned in a ring. An initial normal needs to be assigned to determine surface normals. Sato *et al.* [3] use the similarity of intensity profiles to recover the normal difference with the assumptions about uniform light sources and Lambertian or Torrance-Sparrow reflectance models. Okabe *et al.* [36] use attached shadow codes to deal with general BRDFs under dense and uniform light sources. Both of the two methods require visible occluding contours for resolving the ambiguity. Unlike these methods, our method does not require any reference objects being captured under the same lights with the test scene, and assumes none of initial normals, certain reflectance models, visible occluding contours, or strictly uniform light sources.

2 INTENSITY PROFILES FOR PHOTOMETRIC STEREO

An intensity profile is a sequence of radiance intensities recorded at a pixel across varying lighting conditions. It has been used in various contexts due to its useful properties when assuming distant lightings, no cast shadow, and no inter-reflections.

Orientation-consistency: Intensity profiles become exactly the same, if and only if they correspond to the same surface normal and material (A and C in Fig. 1). Using this simple observation, surface normals can be determined by looking up a pre-stored table indexed with surface brightness values [32], or match the intensity profiles to those from a reference object that is made of the same material [2].

Geometry-extrema: For isotropic BRDFs, intensity profiles reach the extremas synchronously, if and only if they correspond to the same surface normal (A and D in Fig. 1) even for scene points with different materials. This

observation is used for clustering surface orientations [1] without determining the exact orientations.

Similarity: Similarity of two intensity profiles is strongly related to the angular difference between two surface normals for the same material (A and B in Fig. 1). Sato *et al.* [3] analyze this relationship and recover surface normals in the cases of Lambertian and Torrance-Sparrow reflectances illuminated under evenly distributed light directions with an assumption of known occluding boundaries.

This paper makes a further observation about intensity profiles and introduces the property of *conditional linearity*. Unlike [3], we do not restrict our analysis to certain reflectance models. Instead, we take into account more general isotropic reflectances such as materials in the MERL BRDF database [6].

Conditional linearity: For various real-world isotropic reflectances, we observe a strong *linear* relation between the distance among intensity profiles seen under evenly distributed lightings and the angular difference of surface normals, up to a certain normal angular difference. We also observe that the linear coefficient is material-dependent and closely related to the intensity distribution of the intensity profiles. These observations allow us to develop uncalibrated photometric stereo methods that work with unknown isotropic reflectances.

2.1 Angular difference of Surface normals

Let us now assume evenly distributed but unknown light directions and a scene with a uniform material; we show later that these assumptions can be relaxed to some extent. Let $\mathbf{n}_p, \mathbf{n}_q \in \mathbb{R}^{3 \times 1}$ be a surface normal pair at pixel locations p and q, and $\{\mathbf{I}_p, \mathbf{I}_q\}$ be the corresponding pixel intensity profiles in a normalized form as:

$$\mathbf{I}_p = \left[I_p^1, \ldots, I_p^L\right]^{\mathrm{T}} = \left[\tilde{I}_p^1, \ldots, \tilde{I}_p^L\right]^{\mathrm{T}} \Big/ \sqrt{\Sigma_l(\tilde{I}_p^l)^2}, \quad (1)$$

where $\tilde{I}_p^l$ is the recorded intensity at the p-th pixel for a scene point ($p = 1, \ldots, P$), under the l-th lighting direction ($l = 1, \ldots, L$), and I_p^l is the normalized intensity.

Previous methods have shown that the similarity of intensity profiles and surface normals are strongly correlated [1], [3]. Similarity can be defined using the Euclidean distance of two intensity profiles as $\|\mathbf{I}_p - \mathbf{I}_q\|_2$ in a straightforward manner. It indeed correlates with the angular difference $\cos^{-1}(\mathbf{n}_p^{\mathrm{T}}\mathbf{n}_q)$ of surface normals $\mathbf{n}_p$ and $\mathbf{n}_q$ at scene points p and q; however, this linear relationship only holds in a limited range as depicted in Fig. 2 (left). To extend the range, our method uses a geodesic distance $d_G(\mathbf{I}_p, \mathbf{I}_q)$ instead of $\|\mathbf{I}_p - \mathbf{I}_q\|_2$ as it is used in [3] to measure the similarity of more distinct normals (Fig. 2 (right)).

The geodesic distance corresponds to the distance of the shortest path between two nodes in a graph, which is computed by adding up the Euclidean distances of neighboring nodes along the path [37]. Here we refer to "neighbors" in the space of intensity profiles rather
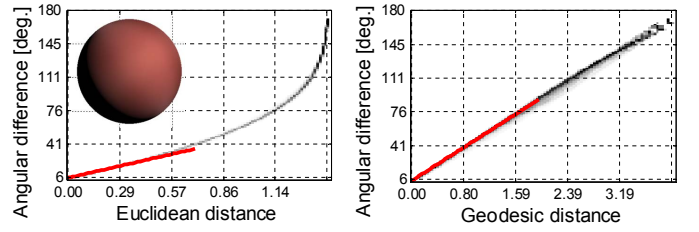

Fig. 2. Using geodesic distance (right) preserves a linear relationship over a greater range of angular differences in comparison with using Euclidean distance (left).
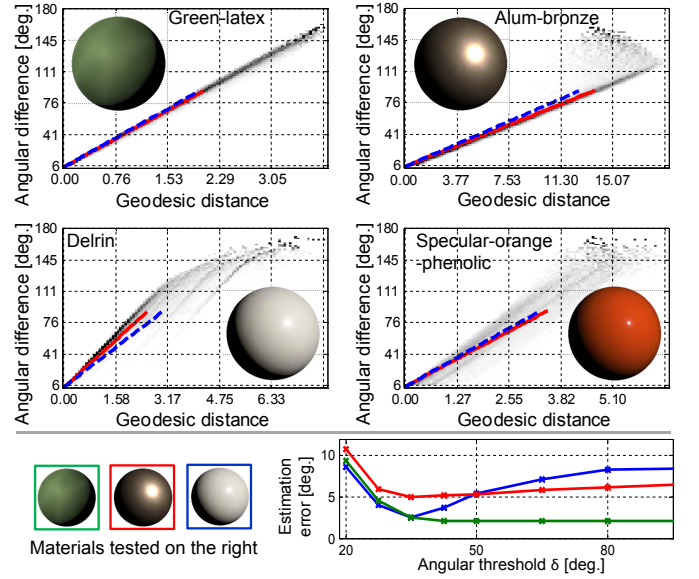

Fig. 3. Examples of the linear relationship. Top: four typical distributions and their partial linear fittings (solid lines) are shown under uniform lightings. Dotted lines show non-uniform light cases. Bottom: setting different angular threshold δ varies the estimation errors.

than pixel coordinates. In our case, we first compute Euclidean distances between all possible combinations of intensity profiles $\mathbf{I}_p$ and $\mathbf{I}_q$ as

$$d(\mathbf{I}_p, \mathbf{I}_q) = \begin{cases} \|\mathbf{I}_p - \mathbf{I}_q\|_2 & \text{if } \|\mathbf{I}_p - \mathbf{I}_q\|_2 < \varepsilon_p \\ +\infty & \text{otherwise,} \end{cases} \quad (2)$$

where each $\mathbf{I}_p$ determines a threshold ε_p; it is empirically set to the 10-th shortest distance from this $\mathbf{I}_p$ to all other $\mathbf{I}_q$. Then geodesic distance between $\mathbf{I}_u$ and $\mathbf{I}_v$ is computed using the pre-computed Euclidean distances $\{d(\mathbf{I}_p, \mathbf{I}_q)\}$ as

$$d_G(\mathbf{I}_u, \mathbf{I}_v) = D_{sp}(\{d(\mathbf{I}_p, \mathbf{I}_q)\}), \quad (3)$$

where function $D_{sp}(\cdot)$ applies a Dijkstra's algorithm to find the shortest path that connects $\mathbf{I}_u$ and $\mathbf{I}_v$, and adds up multiple $d(\mathbf{I}_p, \mathbf{I}_q)$ along the path. In this manner, in the geodesic distance, linearity is locally well preserved and extended to a greater range of angular differences as shown in Fig. 2 (right).

We examine this linearity for all 100 materials in the MERL BRDF database [6] by plotting the angular differences $\cos^{-1}(\mathbf{n}_u^{\mathrm{T}}\mathbf{n}_v)$ and geodesic distances $d_G(\mathbf{I}_u, \mathbf{I}_v)$. Fig. 3 shows four typical plots, from which we can make the following observations. First, the use of geodesic distance generally shows the greater linear relationship to the angular difference of normals in a larger range. Second, for most materials, even those having complex reflectances, the linearity still largely holds in the range of $[0°, 70°]$ of the angular difference, while it does not span the entire range for many materials. Finally, the linear coefficient, or slope, varies with the material (solid lines). The slope is insensitive to random noise in lighting directions as shown by the dotted lines in Fig. 3, which are the line fittings to the plots produced with fluctuating light directions by $7°$ standard deviations. A more thorough analysis on the linearity under non-uniform light distributions is given in Sec. 5.1.

Interestingly, the first case in Fig. 3 is actually the problem solved in [3], which corresponds to an ideal subset of the more general reflectances. Based on the above observations, we define a partial linear conversion from the geodesic distance $d_G(\mathbf{I}_u, \mathbf{I}_v)$ to the angular difference of surface normals $\cos^{-1}(\mathbf{n}_u^{\mathrm{T}}\mathbf{n}_v)$ in a certain range bounded by a threshold δ as

$$\cos^{-1}(\mathbf{n}_u^{\mathrm{T}}\mathbf{n}_v) \approx \alpha_m d_G(\mathbf{I}_u, \mathbf{I}_v) \text{ if } \cos^{-1}(\mathbf{n}_u^{\mathrm{T}}\mathbf{n}_v) < \delta. \quad (4)$$

While the result is not sensitive to the choice of threshold δ, extreme values should be avoided. We examine the error across varying δ in the bottom of Fig. 3. It shows that for non-diffused materials, δ should be set no larger than $\pi/4$ to guarantee a good estimation. As for the material-dependent slope α_m, the inference of α_m enables us to eliminate the necessity of the known occluding boundary assumption that is used in [3] as we will see later in this paper.

2.2 Material reflectance property

The linear coefficient α_m described in the previous section is material-dependent, i.e., it is related to the surface reflectance property of a material. To characterize such a reflectance property, we begin with showing that the *intensity distribution* observed in an intensity profile conveys information about the reflectance property. Fig. 4 shows four intensity profiles, where each row corresponds to the same material but different surface normals. These figures indicate that an intensity profile's shape depends on both reflectance and surface normal. On the other hand, its intensity distribution appears stable against surface normal changes while it varies with materials as shown in Fig. 4.

Based upon this observation, we develop a new measure for characterizing the relationship between the intensity distribution and the corresponding material. The intuitive examples in Fig. 4 suggest that a shinny material shows a biased distribution due to sporadic specular observations compared to a more symmetric
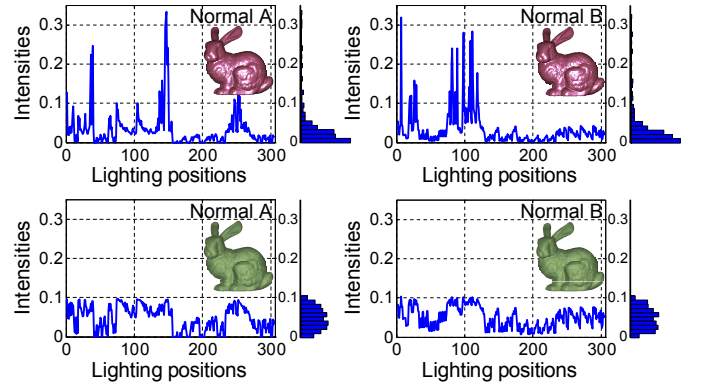

Fig. 4. Intensity profiles w.r.t. material and surface normal. The top row shows intensity profiles of a specular surface recorded at two distinct surface normals. The bottom row shows those of a diffuse surface. The intensity distributions are plotted on their right-hand side.

distribution from a diffused material. We therefore use *skewness* for characterizing the relationship. The skewness γ of an intensity distribution, which is invariant to the intensity/lighting order, is calculated as

$$\gamma(\mathbf{I}) = L^{\frac{1}{2}} \sum_l (I^l)^3 \bigg/ \left(\sum_l (I^l)^2 \right)^{\frac{3}{2}}, \quad (5)$$

where $\mathbf{I}$ is an intensity profile, and I^l is its l-th element that corresponds to the l-th observation under the L lighting direction. If we look at examples in Fig. 4 again, the skewness value is large for the shiny material and is small for the matte one. Therefore, we quantitatively examine the relationship between the skewness of the intensity distributions and the linear coefficient α_m^{-1}.

To verify the effectiveness of the skewness measure, we render a spherical surface under varying numbers of light directions (162 and 42) using all the 100 materials in MERL BRDF database. The skewness values γ computed from all pixels and inverse slopes α_m^{-1} are plotted in Fig. 5. As we see in the figure, they have a strong linear relationship, which in fact yields the correlation coefficient of 0.99. Our experiment shows that varying numbers of light directions show nearly identical linear relations. From this result, we use the skewness of the intensity distribution for estimating α_m of the unknown material using the linear relationship in Fig. 5, and it results in constraining Eq. (4).

2.3 Photometric stereo using intensity profiles

The two linear relationships described in the previous sections 2.1 and 2.2 are our key findings in this paper. Based on them, we develop two uncalibrated photometric stereo methods that work for unknown isotropic BRDFs. We begin with the case where only intensity profiles are available in Sec. 3, and extend the method for the case where a database of BRDFs is available in Sec. 4. We denote the former method as a method *without* reference data, and the latter as a method *with*
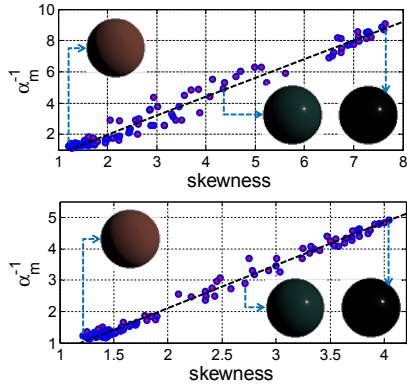
Fig. 5. Relationship between skewness values and α_m^{-1} values of 100 materials. The correlation coefficients are 0.99 (Top/bottom: $162/42$ light directions)

reference data. The method without reference data uses the two linear relationships to quantitatively convert the observed intensity profiles into surface normals, while the method with reference data casts the original problem into a matching of intensity profiles, where its search space is bounded by the two linear relationships.

In practice, they can be applied in accordance with the scenarios: quick implementation without a pre-stored database, or simultaneous light source recovery and a generally better guaranteed accuracy (Sec. 5).

In what follows, we begin with describing the method without reference data to explain how surface normals can be determined only from intensity profiles. Then we describe the method with reference data, in which the problem can be more constrained.

3 SURFACE NORMAL ESTIMATION WITHOUT REFERENCE DATA

We describe our first method for recovering surface normals only from the intensity profiles without reference data. From the observed intensity profiles, we compute geodesic distance $d_G(\mathbf{I}_u, \mathbf{I}_v)$ using Eqs. (2) and (3). We then convert $d_G(\mathbf{I}_u, \mathbf{I}_v)$ into a normal angular difference $\cos^{-1}(\mathbf{n}_u^{\mathrm{T}}\mathbf{n}_v)$ using Eq. (4). As discussed in the previous section, since the conversion is valid only when $\cos^{-1}(\mathbf{n}_u^{\mathrm{T}}\mathbf{n}_v) < \delta$ holds, we rewrite Eq. (4) as the following:

$$\cos^{-1}(\mathbf{n}_u^{\mathrm{T}}\mathbf{n}_v) = \begin{cases} \alpha_m d_G(\mathbf{I}_u, \mathbf{I}_v) & \text{if } \alpha_m d_G(\mathbf{I}_u, \mathbf{I}_v) \leq \delta \\ \text{Undefined} & \text{otherwise,} \end{cases}$$

(6)

where α_m is obtained using the skewness measure, and threshold δ is set to $\pi/4$ to ensure good linear regions for diverse classes of materials as discussed in Sec. 2.1.

3.1 Formulation

We wish to recover surface normals of scene points that correspond to P pixels in the observed image from a set of images taken under varying unknown lightings. Let the surface normal matrix be

$N = [\mathbf{n}_1 \mid \mathbf{n}_2 \mid \ldots \mid \mathbf{n}_P] \in \mathbb{R}^{3 \times P}$ that we solve for. We define the observation matrix O as

$$O = [o_{u,v} = \mathbf{n}_u^{\mathrm{T}}\mathbf{n}_v]_{P \times P} \in \mathbb{R}^{P \times P},$$

(7)

whose elements $o_{u,v}(= \mathbf{n}_u^{\mathrm{T}}\mathbf{n}_v)$ are readily obtained from Eq. (6). The diagonal elements $o_{u,u}$ in O are all ones, which ensures the unit normal length for $\{\mathbf{n}_p\}$. Equation (6) has undefined cases for some u and v; therefore, not all O's elements have well-defined values. In other words, the observation matrix O has missing elements. With a matrix E that accounts for the errors due to the missing entries, the relationship between the observation matrix O and surface normal N can be written as

$$O = N^{\mathrm{T}}N + E.$$

(8)

We wish to solve for surface normal N by using the incomplete observation matrix O and unknown but sparse error matrix E.

3.2 Matrix decomposition with missing data

Solving Eq. (8) for N involves recovering and decomposing the incomplete observation matrix O. We use a matrix A for $A = N^{\mathrm{T}}N$, where we know that $\mathrm{rank}(A) = 3$ since $\mathrm{rank}(N) = 3$. Let Ω be a set of indices where $o_{u,v}$ has a defined value in O, and let its complement set be Ω^c. By restricting the error matrix E only to account for the missing entries Ω^c, the original problem of Eq. (8) can be written as

$$\underset{A}{\operatorname{argmin}} \|A - O + E\|_F^2 \quad \text{s.t.} \quad \mathrm{rank}(A) = 3, \Phi_\Omega(E) = \mathbf{0},$$

(9)

where $\Phi_\Omega(E)$ is an operator that only keeps E's entries in Ω unchanged and sets others zero. We solve the problem of Eq. (9) by alternately estimating A and E. The optimization begins by initializing $E = \mathbf{0}$ and setting the missing entries of O to zeros so that $\Phi_{\Omega^c}(O) = \mathbf{0}$.

At the k-th iteration, we update A_{k+1} by singular value decomposition (SVD) as

$$\begin{cases} USV^{\mathrm{T}} \leftarrow \mathrm{SVD}(O - E_k), \\ A_{k+1} \leftarrow U\left(S\begin{bmatrix} I_{(3)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\right)V^{\mathrm{T}} = US_{(3)}V^{\mathrm{T}}, \end{cases}$$

(10)

where A_{k+1} is reconstructed using only the first three singular values of S to ensure $\mathrm{rank}(A_{k+1}) = 3$. $I_{(3)}$ is a 3×3 identity matrix. We then update E_{k+1} by

$$E_{k+1} \leftarrow \Phi_{\Omega^c}(O - A_{k+1}).$$

(11)

Equation (11) assigns values to E_{k+1} only for those entries in Ω^c. This ensures the second constraint of Eq. (9) to hold.

The iteration stops when it converges using the criterion $\|E_k - E_{k+1}\|_F < \xi\|O\|_F$, where ξ is a small value (set to 10^{-4}). Finally, the method obtains the surface normal estimates $\hat{N}$ as

$$\hat{N} = S_{(3)}^{\frac{1}{2}}U^{\mathrm{T}} = S_{(3)}^{\frac{1}{2}}V^{\mathrm{T}}.$$

(12)

3.3 Convex/Concave ambiguity

Like any other uncalibrated approach, our solution contains ambiguity. In our case, any matrix $Q \in \mathbb{R}^{3 \times 3}$ that satisfies $Q^{\mathrm{T}}Q = I$ can be multiplied with the solution of Eq. (8) without violating the equality [38]:

$$N^{\mathrm{T}}Q^{\mathrm{T}}QN = (QN)^{\mathrm{T}}(QN) = O - E. \qquad (13)$$

Therefore, $QN \in \mathbb{R}^{3 \times P}$ is also a solution to the problem. In a previous study, Belhumeur *et al.* [7] show that if $\mathrm{rank}(N) = 3$, the ambiguity due to a general invertible 3×3 matrix can be reduced to the generalized bas-relief (GBR) ambiguity using the integrability constraint [39]. This is also true in our case since $Q \in \mathbb{R}^{3 \times 3}$. Therefore, by enforcing integrability, we have a simpler ambiguity described by a matrix $\tilde{Q}$. As a result, in our case, $\tilde{Q}$ should take a form of a GBR transformation, and at the same time, satisfy $\tilde{Q}^{\mathrm{T}}\tilde{Q} = I$. Therefore, it becomes

$$\tilde{Q} \Rightarrow \underbrace{\frac{1}{\lambda}\begin{bmatrix} 1 & 0 & \mu \\ 0 & 1 & \nu \\ 0 & 0 & \lambda \end{bmatrix}}_{\text{GBR transformation}} \Rightarrow \underbrace{\frac{1}{\lambda}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \lambda \end{bmatrix}}_{\text{Satisfy } \tilde{Q}^{\mathrm{T}}\tilde{Q}=I}, \quad \lambda = \pm 1. \quad (14)$$

This is a binary ambiguity where $\lambda = \pm 1$ corresponds to convex/concave surfaces that are not distinguishable without light calibration. Therefore, by using the integrability constraint, our method recovers surface normals up to only a binary convex/concave ambiguity. To apply the integrability constraint, our method uses the approach of [39].

4 NORMAL ESTIMATION BY USING REFERENCE DATA

Section 3 recovers surface normals for general isotropic materials using only intensity profiles without any reference data. In this section, we describe an extended method by assuming that a database of BRDFs is additionally available.

A BRDF database includes dense measurements of real material reflectances, and diverse appearances can be synthesized using varying surface normals and lightings. If such a synthetic data is dense enough, any test data can be considered as its subset, and the problem can be made more bounded and tractable by using the known synthetic data as a reference. In this extended method, we first generate the reference dataset using the BRDF database, and then use the reference data for deriving the estimates of surface normals and lighting directions.

4.1 Reference data

We produce a reference dataset containing intensity values that are sampled from a large variety of surface normals, light directions, and materials. Let $\mathbf{s}_k^r \in \mathbb{R}^3$ $(k = 1, \ldots, S)$ and $\mathbf{n}_i^r$ $(i = 1, \ldots, N)$ denote the sampled light directions and surface normals, respectively, and let $m(= 1, \ldots, M)$ be the index for materials included in the reference dataset. Then $I^r(k, i, m)$ is the
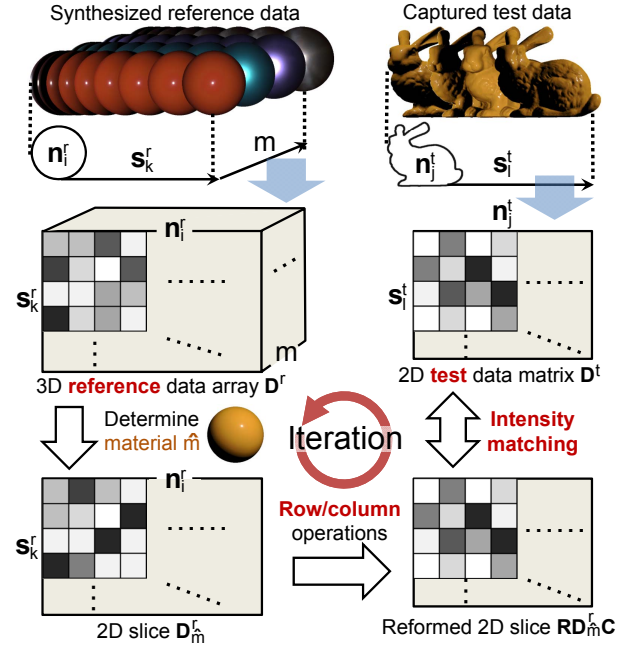

Fig. 6. Matching-based photometric stereo: determining surface normals, light directions, and material of the test scene by matching intensity values between test data and reference data.

corresponding pixel intensity value for k-th light, i-th surface normal, and m-th material. In other words, the reference dataset $\{I^r(k, i, m)\}$ defines a mapping $(\mathbf{s}_k^r, \mathbf{n}_i^r, \mathrm{BRDF}(m)) \mapsto I^r(k, i, m)$. In the rest of the paper, we use superscripts r and t for indicating reference and test data, respectively.

We assume that all light sources have an identical intensity; therefore, both light source $\mathbf{s}_k^r$ and surface normal $\mathbf{n}_i^r$ are represented as unit vectors. To densely sample the reference dataset I^r, we use 1249 frontal vertices on a uniformly subdivided icosahedron to define light directions $\mathbf{s}_k^r$ and surface normals $\mathbf{n}_i^r$ visible to the camera. Pixel intensity values are then rendered for each combination of light source $\mathbf{s}_k^r$, surface normal $\mathbf{n}_i^r$, and material m selected from the MERL BRDF dataset [6] that contains 100 isotropic reflectances. In this manner, we obtain the reference data $\{I^r(k, i, m)\}$ for all combinations of sampled light directions, surface normals, and 100 different surface reflectances.

4.2 Formulation

We formulate the problem as intensity matching between test data and reference dataset. For a test scene, we obtain its test data by recording a sequence of images from a fixed viewpoint with varying directional illumination. Let $\mathbf{s}_l^t$ $(l = 1, \ldots, S^t)$ denote the unknown light vectors and $\mathbf{n}_j^t$ $(j = 1, \ldots, N^t)$ be the unknown surface normals for the recorded pixels. Their corresponding pixel intensity values $\{I^t(l, j)\}$ are stored as the test data, which form a test matrix $D^t = [I^t(l, j)]_{S^t \times N^t}$.

Let $\boldsymbol{D}^r = [I^r(k, i, m)]_{S^r \times N^r \times M^r}$ denote the three dimensional tensor containing all the intensity values from the reference dataset. Then, our task is to match intensity values between $\boldsymbol{D}^r$ and $\boldsymbol{D}^t$. To reduce the computational complexity, our method first obtains the estimate of material $\hat{m}$ using the skewness measure described in Sec. 2.2. Then, its corresponding 2D matrix $\boldsymbol{D}^r_{\hat{m}} = [I^r(k, i, m = \hat{m})]_{S^r \times N^r}$, which is a slice of $\boldsymbol{D}^r$, is used for matching with $\boldsymbol{D}^t$ to solve for surface normals and light directions. An overview of this approach is illustrated in Fig. 6.

To match pixel intensities between $\boldsymbol{D}^r_{\hat{m}}$ and $\boldsymbol{D}^t$, let us define two operators as two matrices $\boldsymbol{R} \in \mathbb{R}^{S^t \times S^r}$ and $\boldsymbol{C} \in \mathbb{R}^{N^r \times N^t}$. When the reference matrix is left-multiplied by $\boldsymbol{R}$, its q-th row is selected and placed in the p-th row of the resulting matrix. Similar to a permutation matrix, this is achieved by setting $R_{p,q} = 1$ to be the only non-zero element in the p-th row of $\boldsymbol{R}$. Similarly, matrix $\boldsymbol{C}$ applies column selection. When the reference matrix is right-multiplied by $\boldsymbol{C}$, its p-th column is selected and placed in the q-th column of the resulting matrix. This is performed by setting the element $C_{p,q} = 1$ to be the only non-zero element in the q-th column of $\boldsymbol{C}$.

Using these operators $\boldsymbol{R}$ and $\boldsymbol{C}$ applied to the reference matrix $\boldsymbol{D}^r$, we can use pixel intensities of the selected rows (light directions) and columns (surface normals) of the reference matrix $\boldsymbol{D}^r$ to compare with the test matrix $\boldsymbol{D}^t$. If m, $\boldsymbol{R}$, and $\boldsymbol{C}$ are properly chosen, the following objective function can be minimized:

$$\{\hat{\boldsymbol{R}}, \hat{\boldsymbol{C}}, \hat{m}\} = \underset{\boldsymbol{R}, \boldsymbol{C}, m, \alpha}{\operatorname{argmin}} \|\boldsymbol{R} \boldsymbol{D}^r_m \boldsymbol{C} - \alpha \boldsymbol{D}^t\|_F, \qquad (15)$$

where α is a scalar value to account for exposure differences. This matching procedure is illustrated in Fig. 6 (right). In the optimal solution set $\{\hat{\boldsymbol{R}}, \hat{\boldsymbol{C}}, \hat{m}\}$, $\hat{m}$ is considered the index of the reference material that best approximates the unknown test material m^t. Besides, positions of '1's in $\hat{\boldsymbol{R}}$ relate each unknown light source directions $\mathbf{s}^t_l$ to their closest reference light source directions in $\{\mathbf{s}^r_k\}$. Similarly, $\hat{\boldsymbol{C}}$ maps each unknown surface normals $\mathbf{n}^t_j$ to their closest reference surface normals in $\{\mathbf{n}^r_i\}$. In this manner, the problem can be formulated as finding the best matching between the test data and reference data.

Due to the discontinuous nature of the operators $\boldsymbol{R}$ and $\boldsymbol{C}$, where only binary values $\{0, 1\}$ are contained, directly finding the global optima via optimization is computationally difficult. We solve this problem in the subsequent sections in an iterative refinement framework. Namely, our strategy is to first determine the material $\hat{m}$, then infer light directions, and finally estimate surface normals, and iterate these steps for reaching to a better solution.

4.3 Material estimation

To solve the problem of Eq. (15), we first obtain the initial solution by inferring the material. By assuming

a uniform surface reflectance in the test scene, or a surface patch with a unique material after segmentation, we infer the surface reflectance from the observed pixel intensity values by computing the skewness measure using Eq. (5). Let the skewness for the test material and those of 100 reference materials be γ^t and $\{\gamma^r_m\}$, respectively. Then a best matching material $\hat{m}$ is found as:

$$\hat{m} = \underset{m}{\operatorname{argmin}} \|\gamma^t - \gamma^r_m\|_2. \qquad (16)$$

However, for highly specular materials, if the test surface is lit by much fewer light directions than those in the reference data, γ^t computed for a single pixel may be unstable due to the limited angular resolution of lightings. To avoid such a situation, we take multiple pixels with similar surface normals (whose intensity profiles have positive correlation coefficients), and use all of them to compute γ^t. We also retain a list (empirically 20) matching candidates $\{\hat{m}\}$ with their ascending matching costs, among which the final solution will be obtained via optimization described later in this paper.

4.4 Initialization of light directions

In this section, we describe a method for initializing light directions, i.e., corresponding to the estimation of $\boldsymbol{R}$, by assuming a roughly uniform light direction distribution. In estimating the light directions, our method takes a two-step approach, i.e., it first determines the azimuth angles, then estimates the elevation angles.

To determine the azimuth angles of the light directions in the test data, we use the angular difference calculation described in Sec. 2.1. We begin with approximately determining two special elements in the test data: the light direction $\mathbf{s}^t_{/\!/}$ and surface normal $\mathbf{n}^t_{/\!/}$, which are parallel to the viewing direction. To identify $\mathbf{s}^t_{/\!/}$, we take a simple strategy by selecting the lighting $\mathbf{s}^t$ that illuminates the greatest number of surface points. Specifically, we examine every image captured under a different light $\mathbf{s}^t_l$, and select the one that contains the largest number of pixels whose intensity values are greater than a threshold; the threshold for detecting shadows is determined based on the first quartile of all measured intensities, following [36]. The frontal surface normal $\mathbf{n}^t_{/\!/}$ is then determined by selecting the surface normal that produces the brightest pixel intensity under the frontal lighting direction $\mathbf{s}^t_{/\!/}$.

At the same time, we find one set of surface normals $\{\mathbf{n}^t_e\}$ with similar elevation angles as illustrated in Fig. 7 after grouping similar intensity values under the frontal light source $\mathbf{s}^r_{/\!/}$. From this set, our method estimates the azimuth angles of $\{\mathbf{n}^t_e\}$ using their intensity profiles $\{\mathbf{I}^t_e\}$; specifically, the linear relationship between the angular difference of surface normals and geodesic distance of their intensity profiles described in Sec. 2.1. For obtaining angular differences among $\{\mathbf{n}^t_e\}$, we apply Multi Dimensional Scaling (MDS) [40], which is a common technique for distance-preserving dimensionality reduction using the geodesic distances among $\{\mathbf{I}^t_e\}$ as input. The output
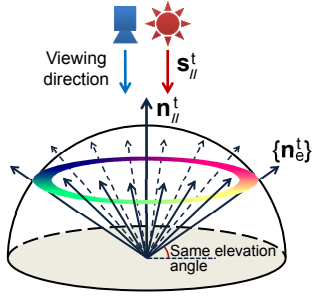
Fig. 7. For light source estimation, we find light source $\mathbf{s}_\parallel^t$ and surface normal $\mathbf{n}_\parallel^t$ parallel to the viewing direction, also a set of normals $\{\mathbf{n}_e^t\}$ with similar elevation angles.
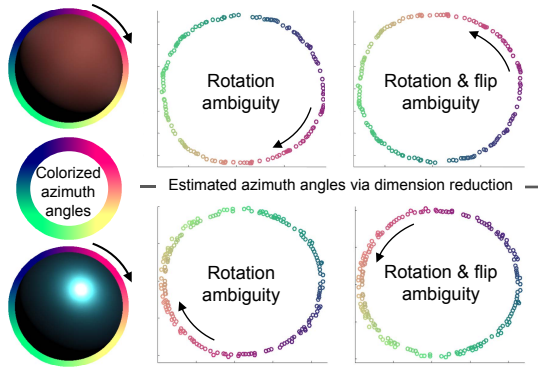

Fig. 8. The estimated azimuth angles of $\{\mathbf{n}_e^t\}$ for two materials are shown in two rows. The spots' positions in the plots are the dimensionality reduction results of $\{\mathbf{I}_e^t\}$. They correctly convey the relative azimuth angles of $\{\mathbf{n}_e^t\}$ and the color shows the ground truth azimuth angles. Note the 2D rotation and flip ambiguity.

are a set of 2D coordinates that preserve the distances among $\{\mathbf{I}_e^t\}$. The geodesic distances among these 2D coordinates are linearly related to the angular difference of $\{\mathbf{n}_e^t\}$; in other words, they show a similar structure with that of $\{\mathbf{n}_e^t\}$ in the 2D space up to a 2D rotation and flip ambiguity as shown in the example in Fig. 8. Since $\{\mathbf{n}_e^t\}$ have similar elevation angles, the 2D coordinates represent the distribution of their azimuth angles.

For each light source $\mathbf{s}_l^t$, we examine its corresponding pixel intensity values at surface normals $\{\mathbf{n}_e^t\}$. The brightest intensity observation should correspond to one of $\{\mathbf{n}_e^t\}$ that has the closest azimuth angle to $\mathbf{s}_l^t$ assuming that the reflectance is isotropic. Based on this, we assign an azimuth angle obtained from $\{\mathbf{n}_e^t\}$ to each light source $\{\mathbf{s}_l^t\}$. Note that, at this point, the 2D rotation and flip ambiguity in $\{\mathbf{n}_e^t\}$ is still embedded in $\mathbf{s}_l^t$.

Once the azimuth angles of $\{\mathbf{s}_l^t\}$ are determined, their elevation angles are ordered in accordance with the brightness of the intensity observations at the frontal surface normal $\mathbf{n}_\parallel^t$, based on the fact that when the light direction's elevation angle gets closer to the frontal surface normal $\mathbf{n}_\parallel^t$, it yields a brighter intensity observation. In the meantime, we remove light sources from $\{\mathbf{s}_l^t\}$ which cannot illuminate $\mathbf{n}_\parallel^t$ to only keep frontal lights,

while notation $\{\mathbf{s}_l^t\}$ does not change for convenience. By assuming a uniform distribution of $\{\mathbf{s}_l^t\}$ in the frontal hemisphere, we order uniform reference lights $\{\mathbf{s}_k^r\}$ in accordance with their elevation angles, sub-sample their elevation angles to a total number equal to the number of $\{\mathbf{s}_l^t\}$ and assign them to $\{\mathbf{s}_l^t\}$.

In this procedure, we have assumed a uniform light source distribution for the purpose of initialization. However, iteration in the next section removes this assumption and handles non-uniform lights for deriving the final solution. The 2D rotation and flip ambiguity in this section will also be addressed later.

4.5　Surface normal estimation

Previous sections estimate a list of candidate materials $\hat{m}$ and light selection matrix $\hat{R}$. What remains unknown in Eq. (15) is $\hat{C}$ that encodes the surface normals of the test scene.

Let $\boldsymbol{A} = \hat{\boldsymbol{R}}\boldsymbol{D}_{\hat{m}}^r$, then Eq. (15) can be rewritten as

$$\hat{\boldsymbol{C}} = \underset{\boldsymbol{C},\alpha}{\operatorname{argmin}} \|\boldsymbol{A}\boldsymbol{C} - \alpha\boldsymbol{D}^t\|_F, \qquad (17)$$

where $\boldsymbol{C}$ contains a single '1' in each column while all the other elements are zero. Because of this discontinuous and non-differentiable constraint, Eq. (17) is not computationally tractable by efficient convex/non-convex optimization methods. Instead of directly optimizing Eq. (17), we use a nearest-neighbor search in the normalized column vector domain. Namely, we sequentially select a column of $\boldsymbol{D}^t$ and find the most similar column in $\boldsymbol{A}$. The similarity is measured using the normalized column vectors to eliminate the effect of α, and then computing the cosine of their inner product. For instance, for the q-th column vector of $\boldsymbol{D}^t$, if its most similar column in $\boldsymbol{A}$ is found to be the p-th one, $C_{p,q} = 1$ is set to be the only non-zero element in $\boldsymbol{C}$'s q-th column.

By sequentially determining each column of $\boldsymbol{C}$, correspondences between test surface normals $\{\mathbf{n}_j^t\}$ and reference normals $\{\mathbf{n}_i^r\}$ are obtained to give us an initial estimate of the test surface geometry. At this point, a 2D rotation and flip ambiguity of light directions (Sec. 4.4) are still embedded in the recovered surface normal $\{\mathbf{n}_j^t\}$.

4.6　Iterative refinement

To further refine the estimates of light and surface normal selection matrices $\hat{R}$ and $\hat{C}$, we take an alternating update procedure. For updating $\hat{R}$, we fix the previous estimate of $\hat{C}$ and compute $\hat{\alpha}$ using Eq. (17). Then $\hat{R}$ is updated by

$$\hat{\boldsymbol{R}} \leftarrow \underset{\boldsymbol{R}}{\operatorname{argmin}} \|\boldsymbol{R}\boldsymbol{D}_{\hat{m}}^r\hat{\boldsymbol{C}} - \hat{\alpha}\boldsymbol{D}^t\|_F, \qquad (18)$$

which is solved in a similar manner as Eq. (17). We then fix $\hat{R}$ and update $\hat{C}$ using Eq. (17) as described in Sec. 4.5. This procedure continues iteratively until the cost decrement of Eq. (17) becomes smaller than 0.1%.

We also define an outer loop aiming at finding the best matching material. As described in Sec. 4.3, we have a

---

Algorithm 1 Reference-based method algorithm

Find candidate materials $\{\hat{m}\}$ via Eq. (16).

// Outer loop: test each candidate material

for Select a new candidate material $\hat{m}$ **do**

    Initial estimation of light directions $\hat{R}$ (Sec. 4.4)

    Initial estimation of surface normals $\hat{C}$ (Sec. 4.5)

    // Inner loop: update light directions and surface normals

    while Not converge **do**

        Update $\hat{R}$ by using Eq. (18)

        Update $\hat{C}$ by using Eq. (17)

        Update the minimum cost value.

    end while

    Update the solution with minimum cost value

end for

Return the final solution $\{\hat{m}, \hat{R}, \hat{C}\}$

---

list of candidate materials obtained by evaluating the skewness. For each of them, we estimate the light and surface normal selection matrices $\hat{R}$ and $\hat{C}$ using the procedure described above. Among those material-wise estimates, the minimizer $\{\hat{R}, \hat{C}, \hat{m}\}$ of the cost of Eq. (15) is taken as the final solution. The entire procedure is summarized in Algorithm 1.

4.7 Convex/Concave ambiguity

As discussed in the end of Sec. 4.5, the solution is not yet correctly determined since it contains ambiguity in the surface normal estimates. Namely, the derived surface normals can rotate and flip about the viewing direction together with light sources by still satisfying Eq. (15).

Similar to Sec. 3.3, the integrability constraint can be used to reduce this ambiguity to a binary one. The 2D rotation and flipping ambiguity B for a surface normal $\mathbf{n}$ is written as

$$B\mathbf{n} = \begin{bmatrix} \pm\cos\phi & -\sin\phi & 0 \\ \pm\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{n}, \qquad (19)$$

where ϕ is the 2D rotation angle around the viewing direction. Using the integrability constraint, the general invertible 3×3 ambiguity can be reduced to a GBR ambiguity. Therefore, in our case, the GBR transformation that satisfies our particular ambiguity B becomes

$$\hat{B} = \begin{bmatrix} \kappa & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \kappa = \pm 1, \qquad (20)$$

which correspond to two surfaces whose surface normals differ by a π rotation. Therefore, the ambiguity corresponds to a binary convex/concave ambiguity. Applying integrability constraint in our case is simple, mainly because the original ambiguity in Eq. (19) only contains an angle ϕ and sign '$\pm$' as unknowns. Compared to general invertible 3×3 matrices solved by conventional methods, our problem is much easier to solve numerically.
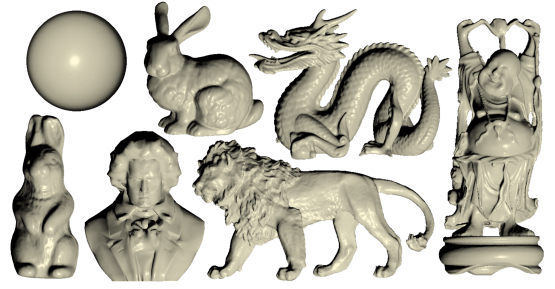

Fig. 9. Synthetic surfaces used in the experiments: hemisphere, Bunny, Dragon, Rabbit, Beethoven, Lion, and Happy Buddha.

TABLE 1
Average normal / light errors of 100 materials. With reference data, only frontal lights are used (89 and 25).

Mat– erials	162 (89*) lights		42 (25*) lights	
	No ref.	With ref.*	No ref.	With ref.*
All 100	$9.65°/-$	$4.36°/5.45°$	$9.30°/-$	$4.88°/5.84°$
Best 67	$6.77°/-$	$2.58°/2.98°$	$7.02°/-$	$3.13°/3.64°$
Best 33	$3.65°/-$	$1.94°/2.03°$	$4.72°/-$	$2.18°/2.32°$

Since our iterative update scheme does not explicitly enforce the ambiguity to be a rotation and flipping ambiguity, for a Lambertian surface, the ambiguity may become the GBR ambiguity during the iterative update. This is due to the fact that Eq. (15) has minimizers up to the GBR ambiguity in the Lambertian case. For general non-Lambertian reflectances, the ambiguity will not be generalized as such because of the information provided by specular components. This is consistent with previous studies that resolve the GBR ambiguity using specular observations [23], [24], [34].

5 EXPERIMENTAL RESULTS

In this section, we evaluate our methods, one without reference data (Sec. 3) and the other without reference data (Sec. 4), using synthetic and real-world data for both quantitative and qualitative assessments.

5.1 Synthetic scenes

We quantitatively assess the proposed methods using synthetic scenes. We use several 3D models include a hemisphere, Bunny, Dragon, and Happy Buddha from Stanford 3D model dataset [41], and also Rabbit, Beethoven and Lion as shown in Fig. 9. These 3D models are rendered using BRDFs of the MERL BRDF database under varying lighting directions. The rendering settings are detailed in the each test described below.

Estimation accuracy We first assess the estimation accuracy of surface normal and light directions using a hemisphere that contains all visible surface normals. We render images using 100 different materials of the MERL BRDF database illuminated under multiple light
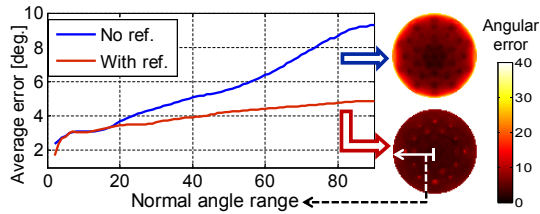
Fig. 10. Average surface normal error w.r.t.normal range. Spherical plots show the average error maps.

TABLE 2
Comparison of average angular errors of surface normal estimates for 100 materials and $162/42$ light directions.

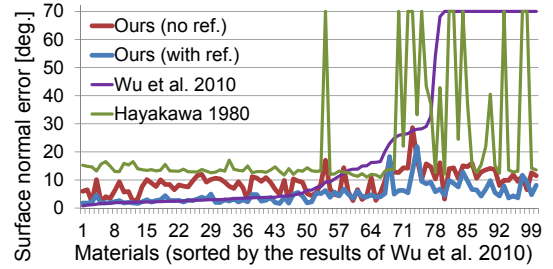| Lights	Ours (no ref.)	Ours* (with ref.)	Wu *et al.* [10]	Hayakawa [38]
162 ($89^\star$)	$9.60°$	$\mathbf{4.36°}$	$18.52°$	$22.09°$
42 ($25^\star$)	$9.30°$	$\mathbf{4.88°}$	$20.05°$	$26.95°$

Fig. 11. Comparison of average normal errors for 100 materials with 42 lights. Large errors are truncated at $70°$.

directions uniformly distributed over a sphere. In particular, we first test 162 uniform light directions, and then 42. For the method with reference data (Sec. 4), we only use the frontal light directions that span in the upper hemisphere, resulting in 89 and 25 light directions.

Table 1 shows the average accuracy for 100 materials. Without reference data, our method in Sec. 3 achieves average errors smaller than $10°$ for all 100 materials, while for the best 33 materials the average accuracy is around $4°$. On the other hand, if using reference data, our method in Sec. 4 has better accuracies than $5°$, while the number for the best 33 materials is around $2°$. In addition, it also allows for light sources estimation, whose accuracy is similar to those for surface normals.

We study the distribution of surface normal errors. Fig. 10 shows the relationship between normal angle range and average error. Without reference data, our method in Sec. 3 accurately recovers surface normals when they are more close to the viewing direction; while errors become bigger for those normals in the boundary. On the other hand, our method in Sec. 4 using reference data performs more stably for variant surface normals. Note that when using reference data, the discrete sampling of $\mathbf{s}_k^r$ and $\mathbf{n}_i^r$ in Sec. 4.1 is one source of the error. In particular, the average sampling interval is about $4.1°$ in our setting, and thus it will cause a $1.6°$ average error for arbitrary surface normal or light source. This error exists in the final results inevitably.

Comparison with previous methods In this section, we further compare our results with those of previous methods. The same dataset as above, i.e., hemispherical surfaces rendered with 100 materials and both 162 and 42 light directions, are used by all methods. Strictly speaking, it is not easy to find prior methods that can completely handle unknown reflectances and uncalibrated illuminations without additional assumptions to ours. In our experiments, we choose the following ones that at least separate the reflectance and illumination factors without requiring known light directions:

1) *Hayakawa* [38]: a baseline method that assumes Lambertian reflectance and no shadows.
2) *Wu et al.* [10]: a state-of-the-art method that robustly handles non-Lambertian components and also cast shadows.

During experiments, we find these methods difficult to solve the GBR ambiguity for very specular materials. Therefore we provide them with the ground truth light

sources. In this sense, they actually work in a calibrated mode and thus give their theoretically best results.

Average estimation errors of all methods are shown and compared in Table 2. Our method achieves significantly better estimation accuracy than others for both 162 and 42 light sources. Note that errors for methods [42] and [10] are around $20°$ in average, but it does not mean these methods cannot handle all materials. In fact, as per-material results shown in Fig. 11 for 42 lights, our method clearly achieves good accuracies for all materials while the other two methods only fail for those difficult ones. This shows the significant merit of our method to handle general and unknown reflectances.

Another significant merit of our method is that, if reference data is available, we can also recover light source directions with high precision. This cannot be done by conventional methods without accurately knowing both reflectance and shape.

Inexact reference material We test the method with a reference database where the exact reference material does not exist. We perform the experiments for 100 materials by forcing the method not to choose the exact reference material. The resulting accuracy does not reduce significantly, as shown in Table 3. It shows that most common reflectances can be approximated by similar ones in the database when only considering their grayscale values. In addition, we also try interpolating the unknown material by using three best matched inexact materials during optimization, where the weights are set inversely proportional to their individual matching costs. The accuracy in Table 3 shows an even more stable result with the interpolation method than when given the exact reference material, since the non-interpolation method may not always select that best material.

Convergence The problem in Algorithm 1 is clearly non-convex, and we test its convergence experimentally.
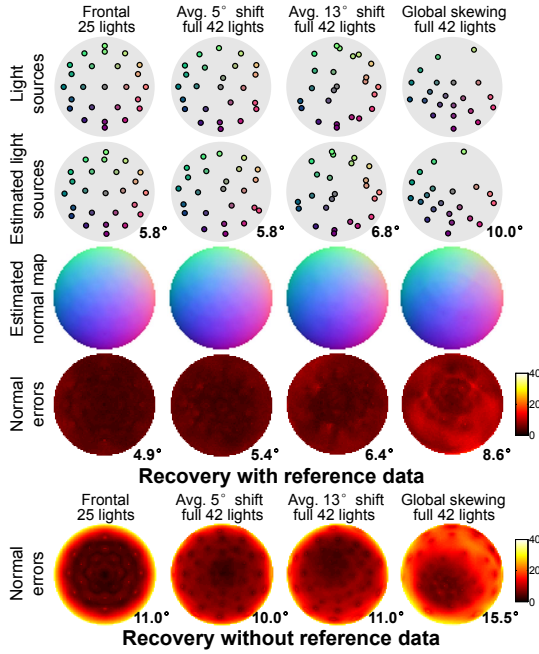
Fig. 12. Average surface normal estimation results for 100 materials under non-uniform distributions of light directions. Values represent the average angular errors.

TABLE 3
Results with/without exact reference material.

Lights	Exact mat.	Inexact mat.	Interpolated mat.
89	$4.36°$	$4.65°$	$\mathbf{4.25°}$
25	$4.88°$	$5.31°$	$\mathbf{4.86°}$

We select eight materials which represent four typical reflectance types: exceptional, specular, glossy and diffuse. We test the convergence with different initial states by adding Gaussian noise to the ground truth light directions to cause average biases from $0°$ to $50°$. Results in Fig. 13 show that: 1) convergence becomes bad for glossy materials with sharp specular lobes when the initial light bias $> 20°$; 2) some exceptional materials have globally large errors since they are not ideally isotropic; 3) other types of materials show good convergence even with large initialization biases. For reference, the average bias of our light direction initialization is below $10°$.

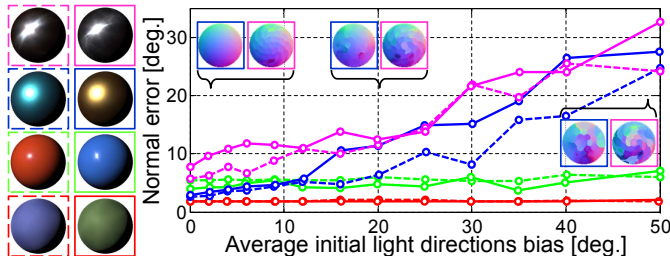**Non-uniform distribution of light directions** To as-

Fig. 13. Results by distorting the initial light directions. Each material on the left corresponds to an error curve.

TABLE 4
Normal error increments due to non-uniformity of lights.

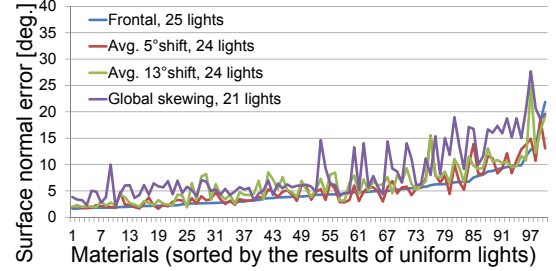|	Uniform	Frontal	5° shift	13° shift	Skewing
No ref.	$9.30°$	$+1.67°$	$+0.67°$	$+1.74°$	$+6.17°$
W. ref.	$4.88°$	$+0°$	$+0.51°$	$+1.48°$	$+3.70°$

Fig. 14. Per-material surface normal estimation errors under different non-uniform light sources.

sess the effects of non-uniform distributions of light directions, we test our method with four representative non-uniform light patterns. These light source directions are generated by 1) only keeping the frontal ones of the original uniform lights; 2) randomly shifting the uniform lights with an average bias of $5°$; 3) randomly shifting the uniform lights with an average bias of $13°$; and 4) globally skewing the uniform lights. These light patterns and the estimation results are summarized in Fig. 12. Note that with reference data, our method in Sec. 4 by default only uses frontal lights.

Intuitively, both surface normal map and light source pattern can be correctly obtained. Randomly shifted light sources do not affect the accuracy too much, while global skewing causes relatively larger errors. This intuition can be supported by the error increments given in Table 4, where the errors increase by $10\% \sim 20\%$ due to random shifts, and the number is around 70% in the cases of global skewing. These results indicate that uniform light sources are preferred by our method, and non-uniformity due to slight random shifts is also well acceptable. Good accuracy is achievable by just assigning light sources to be roughly uniform in practice.

Per-material results are further given in Fig. 14 for easy comparison. The results obtained by using reference data are shown. The uniform-light case achieves best accuracies, while other light sources produce bigger errors but maintain the same error-material correspondences.

Results for synthetic 3D surfaces We conduct experiments using synthetic 3D surfaces shown in Fig. 9. They are rendered with six typical *non-Lambertian* materials under 42 light sources with $5°$ random shifts. Their normal maps are recovered by using our methods both with/without reference data. Representative results are given in Fig. 15. The recovered normal maps by two methods look highly similar, although error maps show that using reference data improves accuracy especially for boundary areas. Surface normal errors and light
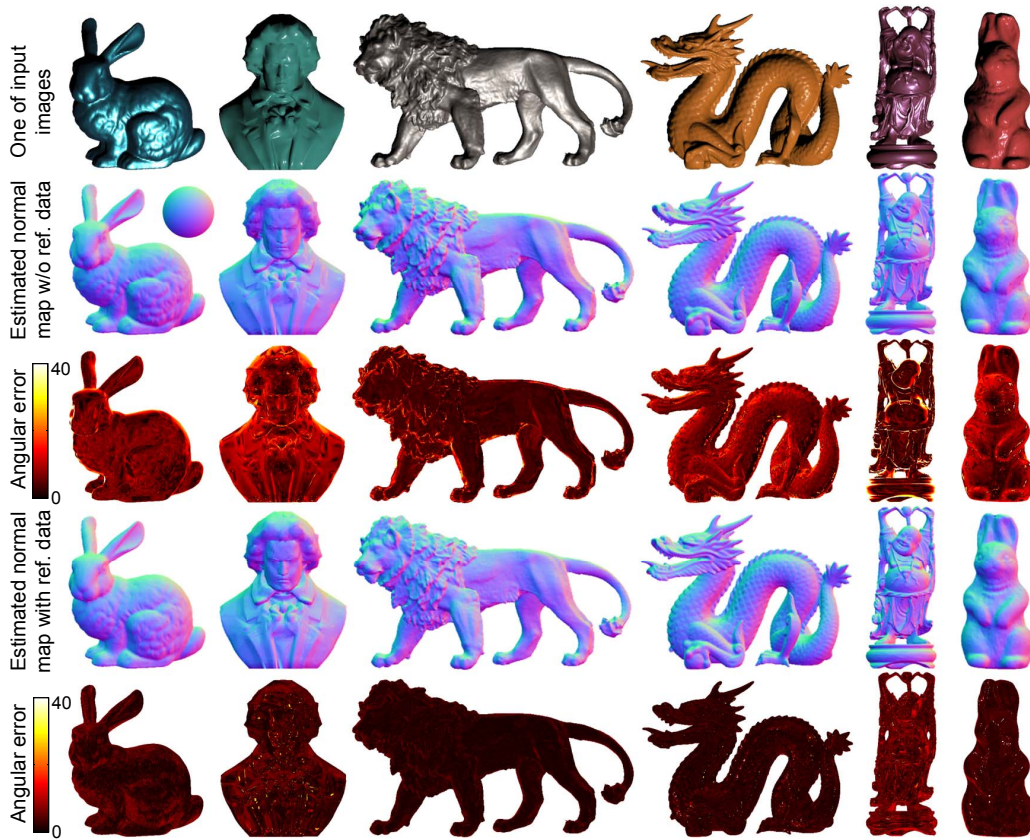
Fig. 15. Results for synthetic surfaces with isotropic reflectances. Top row shows one of the images for each scene. Second and third rows respectively show the recovered surface normal and its error map of our method without reference data (Sec. 3). Fourth and fifth rows show those of our method with reference data (Sec. 4).

TABLE 5
Results for 3D models with six non-Lambertian materials.

3D Model	No ref. normal err.	With ref. normal err.	With ref. lights err.
Bunny	8.91°	4.41°	6.00°
Dragon	8.06°	3.72°	5.12°
Rabbit	8.30°	4.39°	5.84°
Beethoven	6.65°	4.46°	5.96°
Lion	7.12°	3.72°	5.33°
Happy Buddha	6.23°	4.22°	5.44°
Average	7.55°	4.15°	5.62°

source errors are further provided in Table 5. The average normal errors are $7.55°$ and $4.15°$ for all six non-Lambertian reflectances with uncalibrated light sources.

5.2 Real-world scenes

We conduct experiments for real world scenes. Images for real-world objects are captured by rotating the light source around the objects. Some objects with large concavity suffer from severe cast shadows produced by sidelights and back-lights. To avoid this, we use lights mainly from the frontal hemisphere for these objects. In most cases around 50 light sources are finally used.

Experimental results are summarized in Fig. 16 where estimated surface normal maps are presented in color. Results in the second row are obtained without using reference data, and results in the last row are obtained by using reference data. Both methods give nearly identical outputs. Note that different colors on the same object are well handled, because our methods use gray-scale pixel intensities that are normalized. Besides, in all results, there exists a convex/concave ambiguity as discussed. The convex/concave surfaces are theoretically undistinguishable by our method. However, if use additional information such as simple user interaction or boundary assumption, it is feasible to resolve the ambiguity because our ambiguity is only a binary one. For example, the rightmost object "red-leaf" in Fig. 16 has a concave shape that is correctly and automatically recovered by using the outward-looking boundary normals.

We make quantitative evaluations for these results. First, the average angular error of "metallic hemisphere" (the leftmost case in Fig. 16), which has a known shape, is $3.45°$ in the second row, and is $9.99°$ in the third row. The latter one has a larger error partially because intensity matching is more sensitive to intensity loss in saturated areas, as we did not purposely use a high bit-depth camera. As for other objects, we do not have ground truth of their surface normals. However, we can
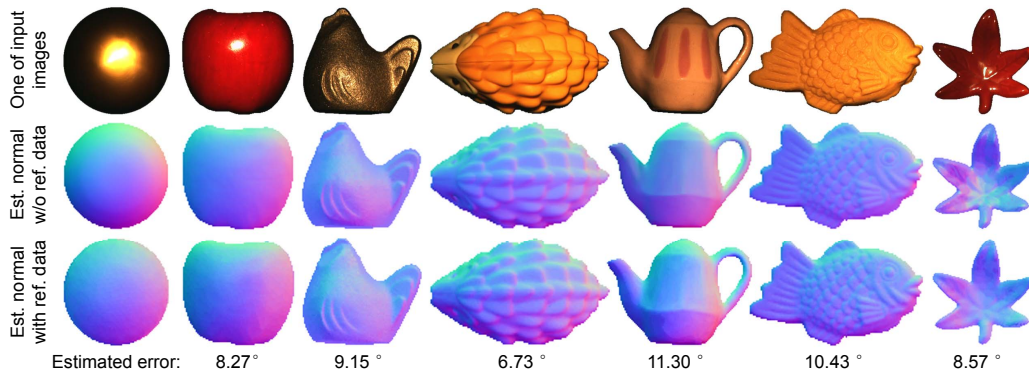
Fig. 16. Result of real objects using our methods without (middle row) and with (bottom row) reference data.
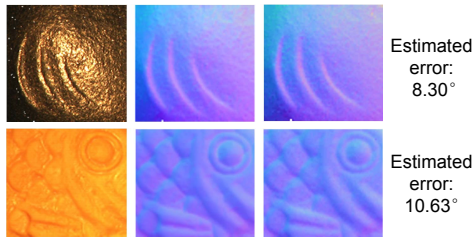

Fig. 17. Real scenes without occluding boundaries.

assess their light source estimation accuracy since we know their ground truth. Because the average error of light sources are similar to (usually larger than) that of surface normals as shown in Table 1 and Table 5, they can be used to estimate the surface normal accuracy. These estimated errors are shown in the bottom of Fig. 16.

We also test a few scenes that have no occluding boundaries, as shown in Fig. 17. As the results show, our methods, no matter with or without reference data, can estimate surface normals purely from the images of the surface patches. This shows the advantage of the proposed methods over the prior methods [3], [36].

6 DISCUSSION AND CONCLUSION

In this paper, we propose photometric stereos methods that handle both uncalibrated light sources and general unknown reflectances. Our approach uses intensity profiles, and it is built upon our findings on the relations between the geodesic distance of intensity profiles and the angular distance of surface normals, and also between the intensity distribution of intensity profiles and surface reflectance property. We investigate these relations and develop methods that recover surface normals under two conditions: purely from intensity profiles and by using reference data. Experimental results show impressive results in surface normal estimation with unknown light sources and general reflectances.

The assumptions about uniform lights and reflectance are somehow restrictive. Although we have shown that they can be relaxed to some degree without significantly affecting the results, further investigation on how to remove them can be a good future research direction.

REFERENCES

[1] S. Koppal and S. Narasimhan, "Clustering appearance for scene analysis," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2006, pp. 1323–1330.

[2] A. Hertzmann and S. Seitz, "Example-based photometric stereo: shape reconstruction with general, varying BRDFs," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1254–1264, 2005.

[3] I. Sato, T. Okabe, Q. Yu, and Y. Sato, "Shape reconstruction based on similarity in radiance changes under varying illumination," in *Proc. of Int'l Conf. on Computer Vision (ICCV)*, 2007, pp. 1–8.

[4] G. Oxholm, P. Bariya, and K. Nishino, "The scale of geometric texture," in *Proc. of European Conf. on Computer Vision (ECCV)*, 2012.

[5] F. Lu, Y. Matsushita, I. Sato, T. Okabe, and Y. Sato, "Uncalibrated photometric stereo for unknown isotropic reflectances," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2013.

[6] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," in *Proc. of ACM SIGGRAPH*, 2003, pp. 27–31.

[7] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille, "The bas-relief ambiguity," *Int'l Journal of Computer Vision*, vol. 35, no. 1, pp. 33–44, 1999.

[8] S. Barsky and M. Petrou, "The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, pp. 1239–1252, 2003.

[9] D. Miyazaki, K. Hara, and K. Ikeuchi, "Median photometric stereo as applied to the segonko tumulus and museum objects," *Int'l Journal of Computer Vision*, vol. 86, no. 2, pp. 229–242, 2010.

[10] 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. of Asian Conf. on Computer Vision (ACCV)*, 2010, pp. 703–717.

[11] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Robust photometric stereo using sparse regression," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2012, pp. 318–325.

[12] T. Wu, K. Tang, C. Tang, and T. Wong, "Dense photometric stereo: A markov random field approach," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 11, pp. 1830–1846, 2006.

[13] N. Alldrin and D. Kriegman, "Toward reconstructing surfaces with arbitrary isotropic reflectance : A stratified photometric stereo approach," in *Proc. of Int'l Conf. on Computer Vision (ICCV)*, 2007, pp. 1–8.

[14] M. Holroyd, J. Lawrence, G. Humphreys, and T. Zickler, "A photometric approach for estimating normals and tangents," *ACM Transactions on Graphics*, vol. 27, p. 133, 2008.

[15] T. Higo, Y. Matsushita, and K. Ikeuchi, "Consensus photometric stereo," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 1157–1164.

[16] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances," in *Proc. of European Conf. on Computer Vision (ECCV)*, 2012.

[17] P. Tan, L. Quan, and T. Zickler, "The geometry of reflectance symmetries," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 12, pp. 2506–2520, 2011.

[18] D. Goldman, B. Curless, A. Hertzmann, and S. Seitz, "Shape and spatially-varying brdfs from photometric stereo," in *Proc. of Int'l Conf. on Computer Vision (ICCV)*, 2005, pp. 341–348.

[19] H. Chung and J. Jia, "Efficient photometric stereo on glossy surfaces with wide specular lobes," in *Proc. of Int'l Conf. on Computer Vision (ICCV)*, 2008, pp. 1–8.

[20] S. Nayar, K. Ikeuchi, and T. Kanade, "Determining shape and reflectance of hybrid surfaces by photometric sampling," *IEEE Transactions on Robotics and Automation*, vol. 6, no. 4, pp. 418–431, 1990.

[21] N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.

[22] P. Favaro and T. Papadhimitri, "A closed-form solution to uncalibrated photometric stereo via diffuse maxima," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2012, pp. 821–828.

[23] O. Drbohlav and R. Šára, "Specularities reduce ambiguity of uncalibrated photometric stereo," in *Proc. of European Conf. on Computer Vision (ECCV)*, 2002, pp. 644–645.

[24] O. Drbohlav and M. Chaniler, "Can two specular pixels calibrate photometric stereo?" in *Proc. of Int'l Conf. on Computer Vision (ICCV)*, 2005, pp. 1850–1857.

[25] R. Basri, D. Jacobs, and I. Kemelmacher, "Photometric stereo with general, unknown lighting," *Int'l Journal of Computer Vision*, vol. 72, no. 3, pp. 239–257, 2007.

[26] N. Alldrin, S. Mallick, and D. Kriegman, "Resolving the generalized bas-relief ambiguity by entropy minimization," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2007.

[27] M. Chandraker, F. Kahl, and D. Kriegman, "Reflections on the generalized bas-relief ambiguity," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 788–795.

[28] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan, "Self-calibrating photometric stereo," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 1118–1125.

[29] Z. Wu and P. Tan, "Calibrating photometric stereo by holistic reflectance symmetry analysis," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2013.

[30] Z. Zhou and P. Tan, "Ring-light photometric stereo," in *Proc. of European Conf. on Computer Vision (ECCV)*, 2010, pp. 265–279.

[31] T. Papadhimitri and P. Favaro, "A new perspective on uncalibrated photometric stereo," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2013.

[32] W. Silver, "Determining shape and reflectance using multiple images," Master's thesis, M.I.T., Dept. of Electrical Engineering and Computer Science, 1980.

[33] P. Ren, J. Wang, J. Snyder, X. Tong, and B. Guo, "Pocket reflectometry," in *Proc. of ACM SIGGRAPH*, 2011, pp. 45:1–45:10.

[34] A. Georghiades, "Incorporating the torrance and sparrow model of reflectance in uncalibrated photometric stereo," in *Proc. of Int'l Conf. on Computer Vision (ICCV)*, 2003, pp. 816–823.

[35] M. Chandraker, J. Bai, and R. Ramamoorthi, "A theory of differential photometric stereo for unknown isotropic brdfs," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 2505–2512.

[36] T. Okabe, I. Sato, and Y. Sato, "Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting conditions," in *Proc. of Int'l Conf. on Computer Vision (ICCV)*, 2009, pp. 1693–1700.

[37] J. Tenenbaum, V. De Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.

[38] H. Hayakawa, "Photometric stereo under a light source with arbitrary motion," *JOSA A*, vol. 11, no. 11, pp. 3079–3089, 1994.

[39] A. Yuille, D. Snow, R. Epstein, and P. Belhumeur, "Determining generative models of objects under varying illumination: Shape and albedo from multiple images using svd and integrability," *Int'l Journal of Computer Vision*, vol. 35, no. 3, pp. 203–222, 1999.

[40] J. B. Kruskal and M. Wish, *Multidimensional scaling*. Sage, 1978, vol. 11.

[41] "The stanford 3d scanning repository," http://graphics.stanford.edu/data/3Dscanrep/.

[42] R. Woodham, "photometric method for determining surface orientation from multiple images," *Optical Engineering*, vol. 1, no. 7, pp. 139–144, 1980.

Feng Lu received the BS and MS degrees in Automation from Tsinghua University in 2007 and 2010, and the PhD degree in information science and technology from the University of Tokyo in 2013 respectively. He is currently a project research associate in the Institute of Industrial Science, the University of Tokyo. His research interests include human gaze analysis, photometric stereo and reflectance analysis.

Yasuyuki Matsushita received his B.S., M.S. and Ph.D. degrees in EECS from the University of Tokyo in 1998, 2000, and 2003, respectively. He joined Microsoft Research Asia in April 2003. He is a Senior Researcher in Visual Computing Group. His areas of research are computer vision (photometric techniques, such as radiometric calibration, photometric stereo, shape-from-shading), computer graphics (image relighting, video analysis and synthesis). He is on the editorial board of IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), International Journal of Computer Vision (IJCV), IPSJ Journal of Computer Vision and Applications (CVA), The Visual Computer Journal, and Encyclopedia of Computer Vision. He served/is serving as a Program Co-Chair of PSIVT 2010, 3DIMPVT 2011, ACCV 2012, and ICCV 2017, and a General Co-Chair for ACCV 2014. He is a senior member of IEEE.

Imari Sato received her B.S. degree in policy management from Keio University in 1994 and M.S. and Ph.D. degrees in interdisciplinary information studies from the University of Tokyo in 2002 and 2005. She is an Associate Professor at the National Institute of Informatics. She serves concurrently as Associate Professor at the Graduate university for Advanced Studies and as Visiting Associate Professor at Department of Information Processing, Tokyo Institute of Technology. Her research interests include physics-based vision and image-based modeling.

Takahiro Okabe received the B.S. and M.S. degrees in physics, and the Ph.D. degree in information science and technology from the University of Tokyo, Japan, in 1997, 1999 and 2011 respectively. After working at the Institute of Industrial Science, the University of Tokyo, he joined Kyushu Institute of Technology, Japan, as an associate professor in 2013. His research interests include computer vision, image processing, pattern recognition, and computer graphics, in particular their physical and mathematical aspects.

Yoichi Sato received the BS degree from the University of Tokyo in 1990, and the MS and PhD degrees in robotics from the School of Computer Science, Carnegie Mellon University, in 1993 and 1997, respectively. He is a professor at the Institute of Industrial Science, the University of Tokyo. His research interests include physics-based vision, reflectance analysis, and gaze and gesture analysis. He served/is serving in several conference organization and journal editorial roles including IEEE Transactions on Pattern Analysis and Machine Intelligence, International Journal of Computer Vision, IPSJ Journal of Computer Vision and Applications, ECCV2012 Program Co-Chair and MVA2013 General Chair.