Photometric Stereo for General BRDFs via Reflection Sparsity Modeling

Tian-Qi Han and Hui-Liang Shen, *Member, IEEE*

***Abstract*— This paper proposes a pixelwise photometric stereo method for object surfaces with general bidirectional reflectance distribution functions (BRDFs) via appropriate reflection modeling. The modeling is based on three general characteristics of reflection components, i.e., the smooth variation of diffuse reflection, the concentration of specular reflection, and the low-intensity nature of shadow. A graph, whose nodes are light directions, is introduced to model these characteristics. In the graph, the neighborhood of nodes is determined by finding the light sources with close directions. The smoothness of the diffuse component is termed as the summation of local variations under all light sources. The specular reflection is modeled by group sparsity, and the shadow is determined via weighted ℓ_1-norm modeling. The optimization problem, which incorporates these three modeling terms, is cast as a second-order cone programming problem. The proposed method is evaluated on both synthetic and real-world scenes with both isotropic and anisotropic materials. The experimental results show that the method is effective for object surfaces with general BRDFs and outperforms the state-of-the-arts.**

***Index Terms*— Photometric stereo, non-Lambertian, general BRDF, surface normal, surface reconstruction, reflectance model, diffuse reflection, specular reflection, shadow, graph, group sparse, sparsity modeling.**

I. Introduction

PHOTOMETRIC stereo estimates the normals of surface points from multiple images, which are captured under different light directions using a fixed camera. Under the strict Lambertian assumption, surface normals can be estimated by least squares (LS) with at least 3 images [1]. However, in practice, the images may be corrupted by non-Lambertian effects such as specular highlight and shadow. In addition, real diffuse reflection also violates the ideal Lambertian assumption. These reflections of object surfaces are generally depicted by the bidirectional reflectance distribution functions (BRDFs) [2]. In the literature, a number of photometric stereo works have been proposed to deal with the non-Lambertian issue, which can be approximately categorized as the Lambertian-enforced methods, BRDF-based methods, and shadow-involved methods.

Manuscript received January 31, 2015; revised June 17, 2015 and July 26, 2015; accepted August 18, 2015. Date of publication August 20, 2015; date of current version September 18, 2015. This work was supported by the National Natural Science Foundation of China under Grant 61371160. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Guoliang Fan. *(Corresponding author: Hui-Liang Shen.)*

The authors are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: htq2002@163.com; shenhl@zju.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2015.2471081

The Lambertian-enforced methods assume a basic Lambertian reflectance model and treat highlight and shadow as outliers [3]–[8]. These methods detect highlight and shadow simultaneously and perform well in case of ideal Lambertian reflection with sparsely distributed outliers. However, for materials with non-Lambertian diffuse reflection or broad specular reflection, the performance of these methods may degrade due to the strict assumption on reflections.

The BRDF-based methods estimate the parameters of BRDF models [9]–[12] or directly use the general properties of the BRDFs such as isotropy, reciprocity symmetry, and monotonicity [13]–[22]. To achieve close conformity to real materials, the BRDF is generally of high dimensionality with respect to surface normal, light direction, and viewing direction. Hence it is challenging to estimate the complete BRDF in a photometric stereo system as the number of light directions is usually not sufficiently large and the normals are not known *a priori*. In addition, many BRDF-based methods [11]–[13], [15], [16], [18] assume shadow is discarded in advance.

The shadow-involved methods employ shadow cues in surface reconstruction [10], [23], [24]. However, it is nontrivial to detect shadow in a real scene. In addition, as the information conveyed in shadow is relatively limited, most existing works choose to eliminate shadow before normal estimation.

This paper proposes a pixel-wise photometric stereo method by estimating the reflectance variation with respect to light directions. Thanks to the pixel-wise nature, the normal estimation does not require the local interaction among neighboring surface points. The reflection is represented by a combination of three components, i.e., the diffuse, specular highlight and shadow components.[1] We model these three reflection components according to their general characteristics. First, the diffuse reflection always varies smoothly with respect to light directions. Second, the specular reflection generally concentrates to highlight regions with high image intensities. Third, the shadow is of low intensity but possibly nonzero. An undirected graph, whose nodes are the light directions, is introduced to model these reflections. The diffuse component is modeled by the summation of local variation of diffuse reflectance. The highlight regions are determined according to its group sparse nature [25]–[27], and the shadow is depicted by a weighted ℓ_1-norm term. The optimization problem is cast as a second order cone programming (SOCP) problem [28], which avoids the difficulty in non-convex optimization.

[1] Strictly speaking, shadow is not a type of reflection. Here we abuse the concept of reflection a little for description simplicity.

The effectiveness of the proposed method is extensively evaluated using both synthetic and real scenes.

A. Related Work

Photometric stereo [1] was introduced by Woodham for surface reconstruction under the strict assumption of Lambertian reflection. Early works extend this method by introducing a forth light source or using color information to deal with highlight and shadow [29]–[31]. These methods can keep the number of photometric images quite low, but the estimation accuracy degrades in case of heavy highlight or non-Lambertain diffuse reflection. To deal with this problem, methods that use more images have been presented for robustness and versatility.

1) Lambertian-Enforced Methods: The methods in this category assume the diffuse component is Lambertian and treat the highlight and shadow as outliers. A Markov network for dense photometric stereo was formulated by using graph cuts and tensor belief propagation with numerous input images [3]. To reduce the number of images, a random sample consensus (RANSAC) method was introduced to classify diffuse reflection, specular reflection, attached shadow, and cast shadow [4]. In [5], median-based photometric stereo was further presented to eliminate the effect of outliers. Note that these mentioned methods optimize the solutions in an iterative way. Another class of Lambertian-enforced methods formulate an optimization problem directly by minimizing the number of the outliers, under the assumption that the outliers are sparsely distributed. For example, the Big-M method [6] was developed to obtain the maximum subset of Lambertian images by solving an integer programming problem. In the low-rank photometric stereo scheme, the integer programming is avoided by using the ℓ_1-norm to approximate the ℓ_0 norm [7]. Motivated by this, a recent work [8] explicitly used the rank-3 Lambertian constraint and cast the problem as a sparse Bayesian learning (SBL) framework. As the SBL framework approximates the ℓ_0 norm better than the ℓ_1-norm, the method [8] could produce improved accuracy in surface normal estimation.

2) BRDF-Based Methods: It is noted that the assumption of Lambertian diffuse reflectance and sparsely distributed specular component do not hold for many real materials. Hence a number of photometric stereo methods, which either employ parametric reflectance models or use reflection properties, have been proposed. In earlier works, the reflectance model like Torrance-Sparrow model [32] or Ward model [33] were used to fit the specular reflection [9]–[11]. In additional to these photorealistic rendering models, a biquadratic reflectance model was recently introduced to represent the low-frequency component of reflectance for photometric stereo [12]. Compared with the classical reflectance models, this model is concise and accurate in representing the low-frequency components, but the high-frequency components (e.g., highlight) should be discarded beforehand. The other BRDF-based methods employed general reflectance properties for normal estimation or surface reconstruction [13]–[20]. For example, the properties of isotropy and reciprocity symmetry was explored in [14], [19] and the photometric invariants were used in [20]. The monotonicity property is employed in [15] to estimate the elevation angles, provided that the azimuth angles are known. The general properties of reflectance were also used together to extend the applications. For example, the work in [16] investigated the monotonicity, visibility, and isotropy simultaneously to obtain inequity constraints of the normal and solve the normal by taking the intersection of the solution spaces. This method could work with a camera with nonlinear radiometric response, but was limited to materials with either diffuse or specular reflections. The photometric technique in [13] exploited the BRDF symmetry and recovered the surface normals of isotropic and anisotropic materials using a large number of images. A recent method [17] incorporated the monotonicity of reflectance by using a piecewise linear approximation to an inverse diffuse model. To further enforce the reflectance properties, a bivariate monotonic smooth function was used to approximate the observed intensity of general isotropic materials [18]. This method is capable of handling both diffuse and specular reflections using convex optimization. However, the orders of the Bernstein polynomials should be increased for versatility, which in turn increases the required number of images (e.g. 300 images for the MERL BRDF database [2], [34]).

Besides, example-based methods [21], [22] were proposed by finding corresponding points with similar appearance between the target and reference objects which are of same materials. These methods are data-driven, and do not assume a specific BRDF model. Thanks to point correspondence, these methods can also deal with attached shadow.

3) Shadow-Involved Methods: In additional to specular highlight, shadow is another influential factor in photometric stereo. Some works [10], [23] employed the shadow information for normal estimation based on simple shadow detection. In practice, the intensity of shadow is affected by complex factors such as environment illumination, interreflection, noise, etc. The misclassification of shadow may considerably affect the estimation of surface normal. In [24], a shadow was detected using graph cut and the obtained shadow map could be used for surface reconstruction. In case of Lambertian reflectance, this method could deal with albedo variation and complex geometry. Nevertheless, the works in this category are relatively limited.

B. Contributions

The contributions of this work are twofold. First, by modeling the general characteristics of different reflection components, a novel photometric stereo method is proposed. Second, the reflection modeling terms are formulated on a graph of light directions, and the unified optimization problem is solved via second order cone programming (SOCP).

The proposed method combines the advantages of the Lambertian-enforced methods and the BRDF-based methods. Similar to the former, the reflection is separated into diffuse, specular, and shadow components. Different from the latter, the reflection is explicitly represented as a function with the light direction as its only variable. In this way, the proposed method reduces the dimensionality in BRDF modeling, and it does not require a large number of photometric images.

TABLE I
MAIN NOTATIONS USED IN THIS WORK

Notations	Meanings
$\mathbf{l}$, $\mathbf{L}$	Light direction vector and matrix
$\mathbf{n}$, $\mathbf{v}$	Surface normal and viewing direction
o, $\mathbf{O}$	Observed intensity in scalar and diagonal matrix form
ρ_d, ρ_s	Diffuse and specular reflectances
s, $\mathbf{s}$	Diffuse component in scalar and vector form
e, $\mathbf{e}$	Specular component in scalar and vector form
τ, $\boldsymbol{\tau}$	Shadow component in scalar and vector form
J_{d},J_{s},J_{w}	Diffuse, specular and shadow modeling terms
λ_{s}, λ_{w}	Weights of the specular and shadow modeling terms
$\mathcal{G}$	Graph of light directions
(V, E)	Node set and edge set of graph $\mathcal{G}$
d_{ij}	Distance between light directions $\mathbf{l}_i$ and $\mathbf{l}_j$
$\mathbf{D}$	Difference matrix for diffuse component
μ_{ij}^-, μ_{ij}^+	Low and high ratio bounds
g_k, $\mathbf{u}_k$	kth group and latent vectors in group sparse modeling
$\tilde{\mathbf{u}}_k$	Latent vector composed of non-zero covariants in $\mathbf{u}_k$
$\tilde{\mathbf{u}}$	Vector formed by stacking all latent vectors $\tilde{\mathbf{u}}_{\mathbf{k}}$
$\mathbf{G}$	Transform matrix for the specular modeling term
w, $\mathbf{w}$	Weight of the shadow term in scalar and vector form
η	Parameter determining μ_{ij}^- and μ_{ij}^+
ξ	Weight factor determining w
$\mathcal{L}$	Label of highlight levels
M	Number of light directions
m	Parameter determining the neighborhood threshold

The proposed method is also different from existing works [3], [20] that use neighborhood or differential relationships. Those methods use the neighborhood information of pixels and thus require the local interactions among neighboring pixels. In contrast, the proposed method actually works in a pixel-wise manner by using the neighborhood information of light directions but not surface points.

C. Notations

Table I lists the main notations and their corresponding meanings used in this work.

II. PHOTOMETRIC STEREO VIA SPARSITY MODELING

In this section we propose an optimization framework by modeling the general characteristics of diffuse reflection, specular reflection, and shadow. It is assumed that the diffuse reflection varies smoothly with respect to light directions; the smoothness is measured by the summation of local variations. The highlight reflection always concentrates to regions and is hence modeled by group sparsity. Considering its low-intensity nature, the shadow is depicted by an intensity-adaptive ℓ_1-norm term. An optimization framework, which incorporates all these reflection modeling terms, is finally proposed.

A. Analysis of Reflectance

Suppose that a surface point is illuminated from a light direction $\mathbf{l}$, the total reflectance ρ can be represented as the linear combination of diffuse reflectance ρ_d and specular reflectance ρ_s. Accordingly, the image intensity o can be formulated as

$$o(\mathbf{n}, \mathbf{l}, \mathbf{v}) = \rho_d(\mathbf{n}, \mathbf{l}, \mathbf{v})\mathbf{n}^{\mathsf{T}}\mathbf{l} + \rho_s(\mathbf{n}, \mathbf{l}, \mathbf{v})\mathbf{n}^{\mathsf{T}}\mathbf{l}, \tag{1}$$

where $\mathbf{n} = (n_x, n_y, n_z)^{\mathsf{T}}$ denotes the surface normal and $\mathbf{v}$ denotes the viewing direction. Without loss of generality, the reflectances are modeled as the functions of $\mathbf{n}$, $\mathbf{l}$, and $\mathbf{v}$. The aim of this work is to estimate the surface normal $\mathbf{n}$, as well as the reflection components, using a number of observed intensities o under various light directions $\mathbf{l}$.

As mentioned above, we do not attempt to solve the high-dimensional parametric BRDF model, but instead model the general characteristics of reflection components. In photometric stereo, the surface normal and viewing direction are both fixed for a specific surface point, and the light direction is the only variable that affects the diffuse and specular reflections. Thus (1) can be simplified as

$$o(\mathbf{l}) = \rho_d(\mathbf{l})\mathbf{n}^{\mathsf{T}}\mathbf{l} + \rho_s(\mathbf{l})\mathbf{n}^{\mathsf{T}}\mathbf{l}. \tag{2}$$

In this way, the reflection components can be regarded as the functions of light direction $\mathbf{l}$, which is known *a prior* in calibrated photometric stereo.

In this work we assume that these two reflections have the following general characteristics. First, the diffuse reflection varies smoothly with respect to light directions, and can be non-Lambertian. Actually, the diffuse components of real surfaces can be depicted by low-frequency reflectance [12], which exhibits smooth characteristic. Second, the specular pixels concentrate to regions in the light direction domain. For many materials, highlight only appears at the surface points whose normals are close to the bisector of light and viewing directions, which is in accordance with many BRDF models [33], [35]–[37]. Equivalently, for light sources that generate highlight at a surface point, their light directions should also be close.

The mentioned two assumptions are appropriate for materials with general BRDFs. Figure 1 illustrates the reflection characteristics of the *yellow-phenolic* material in the MERL database [2]. To illustrate the characteristics, we densely sample the lighting directions from a hemisphere and render the *Bunny* surface under these directions. A rendered image is shown in Fig. 1(a). We select one surface point to discuss the distribution of intensity, reflectance, and norm of

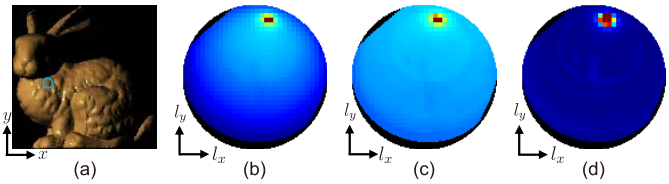

Fig. 1. Illustration of reflection characteristics. (a) A *Bunny* image rendered using the *yellow-phenolic* material under a certain light direction. The interested surface point (circled) is illuminated in turn by a dense light configuration covering the whole hemisphere. (b) Distribution of image intensity for the interested point. (c) Distribution of reflectance. (d) Distribution of the norm of reflectance gradient. The image (a) is shown in the pixel domain (x, y), and the figures (b)-(d) are plotted in the light direction domain (l_x, l_y). In (b)-(d), the pixels with higher values are encoded by warmer colors, and the shadow pixels are in black.

the gradients of reflectances, all with respect to light directions. It is observed in Fig. 1(b) that the highly-valued specular pixels concentrate to a region. As shown in Fig. 1(c), the reflectances vary smoothly in the diffuse region of the light direction domain. It is also clear in Fig. 1(d) that the norms of the gradients of reflectances in the diffuse region are low but not zero. This indicates that the Lambertian assumption for diffuse reflectance is too strict for real materials.

Note that the coincidence of the unknown reflectance and normal in (2) results in a complicated optimization problem. We separate the variables by dividing ρ_d in the two sides, which yields the image model

$$\frac{1}{\rho_d} o = \mathbf{n}^\mathsf{T}\mathbf{l} + \frac{\rho_s}{\rho_d}\mathbf{n}^\mathsf{T}\mathbf{l}. \tag{3}$$

To simplify notations, we introduce two new variables, $s = \frac{1}{\rho_d}$ and $e = \frac{\rho_s}{\rho_d}\mathbf{n}^\mathsf{T}\mathbf{l}$. Then (3) becomes

$$so = \mathbf{n}^\mathsf{T}\mathbf{l} + e. \tag{4}$$

As s is the reciprocal of ρ_d, it also has the smooth property in the light direction domain. Note that e is actually the specular reflection component. Without causing confusion, s and e are, respectively, also referred as diffuse and specular reflection components hereafter.

Note that the above discussion overlooks the shadow which is inevitable in real world. To take the shadows into account, we introduce an extra compensation term τ, and the image model then becomes

$$so = \mathbf{n}^\mathsf{T}\mathbf{l} + e + \tau. \tag{5}$$

When a surface point is not directly illuminated, the observed intensity deviates from the general reflectance model. In this case $\tau \neq 0$ compensates the deviation. On the other hand, for an illuminated point, we have $\tau = 0$.

In a real photometric system, the scene is illuminated in turn under M different light directions. Then (5) can be written in its matrix form as

$$\mathbf{Os} = \mathbf{L}^\mathsf{T}\mathbf{n} + \mathbf{e} + \boldsymbol{\tau}, \tag{6}$$

where $\mathbf{O} = diag(o_1, \cdots, o_M) \in \mathbb{R}^{M\times M}$ denotes the observation matrix and $\mathbf{L} = (\mathbf{l}_1, \cdots, \mathbf{l}_M) \in \mathbb{R}^{3\times M}$ denotes the light direction matrix. The vectors $\mathbf{s} = (s_1, \cdots, s_M)^\mathsf{T} \in \mathbb{R}^M$, $\mathbf{e} = (e_1, \cdots, e_M)^\mathsf{T} \in \mathbb{R}^M$, and $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_M)^\mathsf{T} \in \mathbb{R}^M$ represent the diffuse, specular, and shadow components.

B. Optimization Framework

The optimization problem is formulated by minimizing the combination of the diffuse modeling term J_d, specular modeling term J_s, and shadow modeling term J_w,

$$\min\ J = J_\mathrm{d} + \lambda_\mathrm{s} J_\mathrm{s} + \lambda_\mathrm{w} J_\mathrm{w}, \tag{7}$$

where the weight parameters λ_s and λ_w balance the relative contributions of the specular and shadow modeling terms, respectively.

Note that the modeling terms J_d, J_s, and J_w are computed in the light direction domain. As illustrated in Fig. 2(a), the light directions in a real system are usually irregularly distributed.

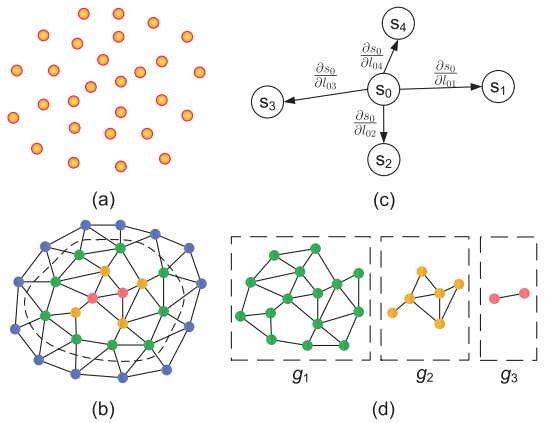

Fig. 2. Illustration of the graph of light directions. (a) The distribution of light source directions. (b) The graph constructed from light directions. The higher pixel intensities are coded with warmer colors. (c) For diffuse reflection, the gradient is a vector that includes all the directional derivative along the connected edges. (d) The specular reflections concentrate into several possible highlight groups.

We construct an undirected graph $\mathcal{G} \in (V, E)$ as shown in Fig. 2(b), in which the elements of node set V are light directions and the edge set E is determined by finding the light source pairs with close directions. The characteristics of the diffuse and specular reflections can then be explored using the neighborhood of light directions, as shown in Fig. 2(c) and (d).

We begin the discussion from the diffuse component, which changes smoothly with respect to light direction. As illustrated in Fig. 2(c), the nodes of graph $\mathcal{G}$ are assigned with the diffuse reflection s. The local variation of a node is measured as the norm of its gradient of diffuse reflection (or more precisely, diffuse reflectance). The diffuse modeling term is computed as the summation of these local variations,

$$J_\mathrm{d} = \frac{1}{2}\sum_{i\in V} \|\nabla_i s\|_2^2, \tag{8}$$

where ∇_i is the gradient operator performed on the diffuse reflection at the ith light direction. The detailed computation of J_d will be elaborated in Section III-A.

It is known that a specular pixel is generally of high intensity. However, it is nontrivial to accurately locate the specular region as the width and shape of the specular lobe is unavailable. In other words, it is not straightforward to determine where the specular region cuts off. Nevertheless, we can still use the intensity information to approximately determine some possible highlight regions with different lobe widths; see Fig. 2(d). In this work, each possible highlight region is treated as a *group* so that we can employ the *group sparse* scheme to determine the most possible highlight regions. In a photometric stereo system with M light sources, we define a group of covariates as a subset $g_k \subset \{1, 2, \cdots, M\}$, then $G = \{g_1, g_2, \cdots, g_K\}$ denotes a set of *overlapped* groups. Note that two groups overlap if they have at least one covariate in common. Then the specular modeling term is defined as

$$J_\mathrm{s} = \min \sum_{k=1}^{K} \beta_k \|\mathbf{u}_k\|_2, \quad \text{s.t.} \sum_{k=1}^{K} \mathbf{u}_k = \mathbf{e}, \tag{9}$$

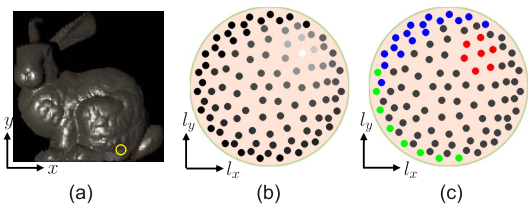

Fig. 3. The resultant reflection labels of a surface point. (a) A *Bunny* image rendered using the *two-layer-gold* material under a certain light direction. The interested surface point is circled. (b) Image intensity of the interested point under 100 light directions. The image intensities are indicated by the brightness levels of the dots. (c) The pixels labeled as highlight (in red), diffuse (in gray), attached shadow (in green), and cast shadow (in blue). In (a) and (b), the image intensities are nonlinearly adjusted for visualization. The image (a) is shown in the pixel domain (x, y), and the figures (b) and (c) are plotted in the light direction domain (l_x, l_y).

where β_k is the weight of the kth group and $\mathbf{u}_k$ is the latent vector of the kth group that has the same length of $\mathbf{e}$. The ith element of $\mathbf{u}_k$ is nonzero only when the covariate $i \in g_k$. The minimization of $\sum_{k=1}^{K} \beta_k \|\mathbf{u}_k\|_2$ with respect to $\mathbf{u}_k$ is a mixed (ℓ_1, ℓ_2)-norm optimization problem [26]. At the group level, it behaves like an ℓ_1-norm and induces group sparsity. On the other hand, within the groups the ℓ_2-norm does not promote sparsity. Therefore the vectors $\mathbf{u}_k$ will be shrunk or non-shrunk to zero as a whole. Then, for the specular component, e_i can be nonzero as long as i belongs to at least one non-shrunk group $\mathbf{u}_k$ due to $\mathbf{e} = \sum_{k=1}^{K} \mathbf{u}_k$. Therefore the nonzero covariates of $\mathbf{e}$ is generally the union of the groups. Especially, when the groups satisfy $g_K \subset g_{K-1} \subset \cdots \subset g_1$, the nonzero covariates of $\mathbf{e}$ are inclined to be in the subset of $G = \{g_1, g_2, \cdots, g_K\}$ [26], [27]. This leads to a group sparse solution from all the possible highlight regions. The details for determining the groups will be described in Section III-B.

There are two types of shadow: attached shadow and cast shadow. The attached shadow occurs when the surface does not face the light direction, while the cast shadow is caused by occlusion which depends on the overall 3D shape of the object. Ideally, a shadow pixel is of zero intensity and thus can be detected by simple thresholding. However, some influential factors, such as environment illumination, inter-reflection, and noises, always lead to nonzero-valued shadow. A high threshold may exclude diffuse pixels that are important to photometric stereo, while a low threshold cannot locate all shadow pixels. In this work we use the weighted ℓ_1-norm to model the shadow term

$$J_{\mathrm{w}} = \|\mathbf{w}^{\mathsf{T}} \boldsymbol{\tau}\|_1, \tag{10}$$

where $\mathbf{w}$ is the weight vector. The adaptive weight is based on the fact that shadow is of low intensity. Compared with the ℓ_2-norm, the ℓ_1-norm has a tendency to yield more zero-valued entries in $\boldsymbol{\tau}$. This is in accordance to the property of shadow distribution, as only a small proportion of light sources will produce attached or cast shadow. The details for (10) will be discussed in Section III-C.

With the above mentioned modeling terms, the image pixels can be determined as highlight, diffuse, or shadow components. Figure 3 illustrates the resultant reflection labels of a *Bunny* surface rendered under 100 light directions. It is observed that the pixels with highest intensities are judged as highlight ones, and the pixels with medium intensities and smooth variations are judged as diffuse ones. The pixels with low intensities are classified as either attached shadow or cast shadow.

III. Algorithm Implementation

In this section, we elaborate the computation of the above mentioned modeling terms in the optimization framework (7). We first present the concrete expression of the diffuse modeling term, and then introduce the way to determine the highlight groups. Next, we determine the weight in the shadow modeling term. The algorithm is finally summarized.

A. Modeling Diffuse Reflection

In the above section, we have constructed an undirected graph without self loops, $\mathcal{G} \in (V, E)$, to incorporate the neighborhood information of light directions. Here V is a finite set of nodes and E is a connected subset, $E \subseteq V \times V$. Two nodes i and j are connected if the unordered pair $(i, j) \in E$. For two light directions $\mathbf{l}_i$ and $\mathbf{l}_j$, we use the distance $d_{ij} = \|\mathbf{l}_i - \mathbf{l}_j\|_2$ to determine whether they are connected,

$$\begin{cases} (i, j) \in E, & d_{ij} < T_d \quad \text{and} \quad i \neq j \\ (i, j) \notin E, & \text{otherwise,} \end{cases} \tag{11}$$

where T_d is a threshold. A larger value of T_d will result in more connected light sources and vice versa. In this work we use a parameter $m \in \mathcal{Z}^+$ to determine the value of T_d. For each light direction, we compute its distances to the other ones, and find the distance d_m which is among the mth minimum value. Taking all the light directions into consideration, we construct a vector $\mathbf{d}_m$ by stacking these d_m values. Then the threshold T_d is set as $T_d = \mathtt{mean}(\mathbf{d}_m) + 3 \cdot \mathtt{std}(\mathbf{d}_m)$. The set of connected pairs, E, can then be determined according to (11) based on the computed T_d. In this way, the number of connected pairs for each node is usually larger than m.

For two connected nodes i and j, the directional derivative of the diffuse component corresponding to node i is computed as

$$\frac{\partial s_i}{\partial l_{ij}} = \frac{s_i - s_j}{d_{ij}}, \tag{12}$$

where l_{ij} is the direction from $\mathbf{l}_i$ to $\mathbf{l}_j$. Let $\Omega_i = \{j | (i, j) \in E\}$ be the set of nodes that connect to node i, the gradient vector of the diffuse reflection at light source i is defined as

$$\nabla_i s = \left(\cdots \frac{\partial s_i}{\partial l_{ij}} \cdots \right)^{\mathsf{T}}, \quad j \in \Omega_i, \tag{13}$$

where $\nabla_i : V \to \mathbb{R}^{|\Omega_i|}$ is a vector gradient operator and $|\Omega_i|$ denotes the number of elements in the set Ω_i. Based on the gradient vector, the local variation of the diffuse reflection at light source i is measured as

$$\|\nabla_i s\|_2 = \sqrt{\sum_{j \in \Omega_i} \left(\frac{\partial s_i}{\partial l_{ij}} \right)^2} = \sqrt{\sum_{j \in \Omega_i} \left(\frac{s_i - s_j}{d_{ij}} \right)^2}. \tag{14}$$

By summing up the local variations at all light sources, the smoothness of the diffuse reflection is computed as

$$\frac{1}{2}\sum_{i\in V}\|\nabla_i s\|_2^2 = \frac{1}{2}\sum_{i\in V}\sum_{j\in\Omega_i}\left(\frac{s_i - s_j}{d_{ij}}\right)^2 = \sum_{(i,j)\in E}\left(\frac{s_i - s_j}{d_{ij}}\right)^2. \tag{15}$$

Note that the factor $\frac{1}{2}$ is introduced to ensure that the directional derivative is computed once for each node. From (15), the smoothness can be regarded as the weighted quadratic sum of all the differences along the connected edges.

For a graph with M nodes and n edges, (15) can be written in its matrix form by defining a difference matrix $\mathbf{D}\in\mathbb{R}^{n\times M}$ with each row expressing one edge difference,

$$\begin{cases} D_{ki} = -D_{kj} = \frac{1}{d_{ij}}, & \{i,j\}\in E_k \\ D_{kl} = 0, & \{l\}\notin E_k, \end{cases} \tag{16}$$

where E_k $(1\le k\le n)$ denotes the kth pair in E. Note that in (16) we abuse the notation a little, by indexing the nodes according to E_k. Then the diffuse modeling term can be expressed as

$$J_{\mathrm{d}} = \frac{1}{2}\sum_{i\in V}\|\nabla_i s\|_2^2 = \|\mathbf{Ds}\|_2^2. \tag{17}$$

B. Modeling Specular Reflection

In the following, we elaborate how to determine the possible highlight regions according to the observed intensities. For two light directions i and j, we compute the corresponding intensity ratio as

$$\frac{o_i}{o_j} = \frac{\rho_i}{\rho_j}\frac{\mathbf{l}_i^{\mathsf{T}}\mathbf{n}_0}{\mathbf{l}_j^{\mathsf{T}}\mathbf{n}_0} = \frac{\rho_i}{\rho_j}\mu_{ij}(\mathbf{n}_0), \tag{18}$$

where ρ_i and ρ_j are the total reflectances under light directions i and j, respectively. For simplicity, we denote $\mu_{ij}(\mathbf{n}_0)\triangleq(\mathbf{l}_i^{\mathsf{T}}\mathbf{n}_0)/(\mathbf{l}_j^{\mathsf{T}}\mathbf{n}_0)$. In the following, we show how to infer the relationship between ρ_i and ρ_j from intensities o_i and o_j, when $\mu_{ij}(\mathbf{n}_0)$ is actually unknown.

Intuitively, for two close light directions $\mathbf{l}_i$ and $\mathbf{l}_j$, the ratio $\mu_{ij}(\mathbf{n}_0)$ should vary in a small range for the majority of normals. Thus we densely sample the normals on a hemisphere and compute the ratios μ_{ij} for each normal under light directions $\mathbf{l}_i$ and $\mathbf{l}_j$. Figure 4(b) illustrates the spatial distribution of μ_{ij} for the given light directions and Fig. 4(c) shows the corresponding probability density function $p_{ij}(\mu)$. We introduce a low bound μ_{ij}^- and a high bound μ_{ij}^+ such that, with a relatively large probability, $\mu_{ij}(\mathbf{n}_0)\in[\mu_{ij}^-,\mu_{ij}^+]$. This implies that, if the intensity ratio o_i/o_j is outside of the range $[\mu_{ij}^-,\mu_{ij}^+]$, the change of intensity value is more likely caused by reflectances ρ_i or ρ_j, but not $\mu_{ij}(\mathbf{n}_0)$. We introduce the label $\mathcal{L}\in\mathcal{Z}_0^+$ to represent the highlight level of each intensity. An image intensity with a larger label value is more likely to be highlight. Then, for two light sources i and j, their labels $\mathcal{L}_i$ and $\mathcal{L}_j$ are related as

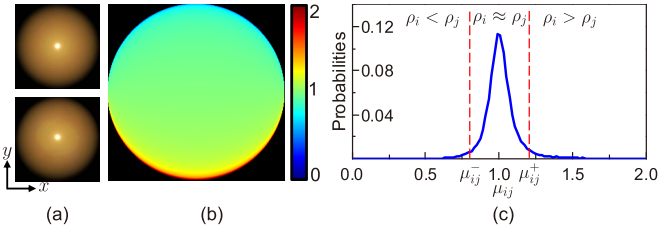

Fig. 4. Illustration of the ratio μ_{ij} under two close light directions $\mathbf{l}_i = (0,0,1)^{\mathsf{T}}$ and $\mathbf{l}_j = (0,\sin\frac{5}{180}\pi,\cos\frac{5}{180}\pi)^{\mathsf{T}}$. (a) Two ball images rendered using the material *yellow-phenolic*. The images are shown in the pixel domain (x,y). (b) Spatial distribution of μ_{ij}, i.e., the value of each pixel is computed as the ratio of the corresponding intensities from the two images in (a). (c) Probability distribution function of μ_{ij}. The relationship between ρ_i and ρ_j are determined by the low bound μ_{ij}^- and high bound μ_{ij}^+.

$$\mathcal{L}_i - \mathcal{L}_j = \Delta_{ij} = \begin{cases} 1, & o_i/o_j > \mu_{ij}^+ \\ -1, & o_i/o_j < \mu_{ij}^- \\ 0, & \text{otherwise.} \end{cases} \tag{19}$$

In this work, we use a parameter $\eta\in(0.5,1)$ to determine the bounds μ_{ij}^+ and μ_{ij}^-, by solving $\int_0^{\mu_{ij}^-} p_{ij}(\mu)d\mu = \eta$ and $\int_{\mu_{ij}^+}^{+\infty} p_{ij}(\mu)d\mu = \eta$, respectively. It will be seen in the experiments that our method is insensitive to the value of parameter η.

We then use (19) to determine the possible highlight regions by assigning the labels to the connected nodes in the graph $\mathcal{G}$. Considering that highlight is generally of high intensity, we assume the intensity below the median of $\mathbf{o}$ is non-specular, and correspondingly, the labels of these nodes are set as 0.

To compute the labels of the q possible specular nodes, we extract the subset $\widetilde{E}\subseteq E$ whose edges are connected by these nodes. According to (19), each edge in $\widetilde{E}$ determines a linear equation. To write these equations in the matrix form, we construct a difference matrix $\mathbf{D}_{\mathcal{L}}\in\mathbb{R}^{\widetilde{n}\times q}$, where $\widetilde{n}=|\widetilde{E}|$ denotes the number of edges in $\widetilde{E}$. As the kth row of $\mathbf{D}_{\mathcal{L}}$ corresponds to the kth edge with connected nodes i and j, we set $D_{ki}=1$, $D_{kj}=-1$ and set 0s to all other elements. Then the label vector $\boldsymbol{\mathcal{L}} = (\mathcal{L}_1,\mathcal{L}_2,\cdots,\mathcal{L}_q)^{\mathsf{T}}$ satisfies

$$\mathbf{D}_{\mathcal{L}}\boldsymbol{\mathcal{L}} = \boldsymbol{\Delta}, \tag{20}$$

where $\boldsymbol{\Delta} = (\Delta_1,\Delta_2,\cdots,\Delta_{\widetilde{n}})^{\mathsf{T}}$ denotes the label difference vector. The element Δ_k is determined by the kth edge with connected nodes i and j, as shown in (19).

The problem (20) is ill-posed as it has only the difference relationships between node pairs. Note that if an unknown label is connected with several zero-valued labels, itself also tends to be non-specular. With this constraints, (20) is reformulated as

$$\min_{\boldsymbol{\mathcal{L}}}\|\mathbf{D}_{\mathcal{L}}\boldsymbol{\mathcal{L}} - \boldsymbol{\Delta}\|_2^2 + \sum_{i=1}^{q}\gamma_i^2\mathcal{L}_i^2, \tag{21}$$

where γ_i is the number of zero-valued labels connecting to node i. As labels are of discrete values, (21) is an integer programming problem which is NP-hard. We relax this problem by using the least squares solution as

$$\widetilde{\boldsymbol{\mathcal{L}}} = (\mathbf{D}_{\mathcal{L}}^{\mathsf{T}}\mathbf{D}_{\mathcal{L}} + \boldsymbol{\Gamma}^{\mathsf{T}}\boldsymbol{\Gamma})^{-1}\mathbf{D}_{\mathcal{L}}^{\mathsf{T}}\boldsymbol{\Delta}, \tag{22}$$

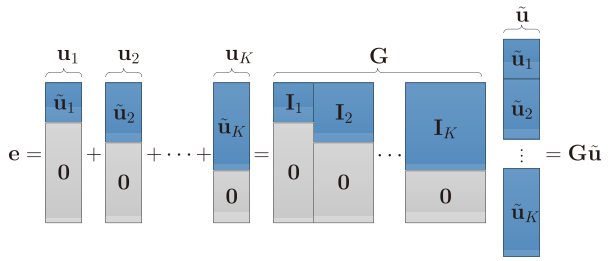

Fig. 5. Illustration of the covariant duplication trick. The sum of vectors $\mathbf{u}$ is converted to matrix multiplication by constructing the matrix $\mathbf{G}$.

where $\mathbf{\Gamma} = \text{diag}(\gamma_1, \gamma_2, \cdots, \gamma_q)$. The rounded label $\boldsymbol{\mathcal{L}} = \lceil \widetilde{\boldsymbol{\mathcal{L}}} \rceil$ is regarded as the final solution. Then we obtain groups $G = \{g_1, g_2, \cdots, g_K\}$, where $g_k = \{\bigcup i | \mathcal{L}_i \geq k\}$ that satisfies $g_K \subset g_{K-1} \subset \cdots \subset g_1$.

After obtaining the groups G, (9) can be expressed in the matrix form by covariant duplication. We first extract the non-zero covariants of each latent vector $\mathbf{u}_k$ to obtain a new vector $\widetilde{\mathbf{u}}_k \in \mathbb{R}^{|g_k|}$, where $|g_k|$ denotes the number of elements in g_k. We then stack all these vectors to form a new vector $\widetilde{\mathbf{u}} \in \mathbb{R}^p$, where p denotes the total number of non-zero latent variables. Finally a transform matrix $\mathbf{G} \in \mathbb{R}^{M \times p}$ is defined as

$$\mathbf{G} = (\widetilde{\mathbf{I}}_1, \widetilde{\mathbf{I}}_2, \cdots, \widetilde{\mathbf{I}}_K), \tag{23}$$

where $\widetilde{\mathbf{I}}_k = \begin{pmatrix} \mathbf{I}_k \\ \mathbf{0} \end{pmatrix} \in \mathbb{R}^{M \times |g_k|}$ is the augmented matrix of identity matrices $\mathbf{I}_k \in \mathbb{R}^{|g_k| \times |g_k|}$. This formulation is illustrated in Fig. 5. Then the specular modeling term J_s in (9) is expressed as

$$J_\text{s} = \min \sum_{k=1}^{K} \beta_k \|\widetilde{\mathbf{u}}_k\|_2. \quad \text{s.t. } \mathbf{e} = \mathbf{G}\widetilde{\mathbf{u}}. \tag{24}$$

Note that $\|\widetilde{\mathbf{u}}_k\|_2 = \|\mathbf{u}_k\|_2$. We set the weights $\beta_k = \sqrt{|g_k| + \kappa\sqrt{|g_k|}}$ with $\kappa = 1$ to avoid redundant groups [26].

C. Modeling Shadow

In the following, we discuss how to determine the adaptive weight in the shadow modeling term $J_\text{w} = \|\mathbf{w}^\mathsf{T}\boldsymbol{\tau}\|_1$. Recall that the whole optimization framework for photometric stereo is

$$\min \; J = J_\text{d} + \lambda_\text{s} J_\text{s} + \lambda_\text{w} \|\mathbf{w}^\mathsf{T}\boldsymbol{\tau}\|_1, \tag{25}$$

where $\mathbf{w} = (w_1, \cdots, w_M)^\mathsf{T}$, with M being the number of light directions. In case of a large weight w_i, the shadow component τ_i will shrink to zero. On the other hand, when w_i approaches to zero, τ_i is more inclined to be non-zero under the constraint of the image model (5). Hence an adaptive weight w_i, which increases monotonically with intensity o_i, can be adopted based on the low-intensity nature of the shadow. In this work we set w_i as

$$w_i = (\xi o_i)^2, \tag{26}$$

where ξ is the weight factor. Note that ξ and λ_w function similarly and thus we fixed $\lambda_\text{w} = 1$ in the optimization model (25).

With the estimated shadow component, we can further classify shadows according to the sign of τ. The attached shadow satisfies $\mathbf{l}^\mathsf{T}\mathbf{n} < 0$ and $e = 0$, thus $\tau > 0$. The cast shadow, which is caused by occlusion, satisfies $\mathbf{l}^\mathsf{T}\mathbf{n} > 0$ and so that $\tau < 0$.

Algorithm 1 Normal Estimation

Input : Light directions $\{\mathbf{l}_i\}_{i=1}^M$, image intensities $\{\mathbf{o}_k\}_{k=1}^P$, parameters $\lambda_\text{s} = 1$ or 0.1, $\lambda_\text{w} = 1, \eta, \xi$.
Output: Surface normal $\{\mathbf{n}_k\}_{k=1}^P$

/* Initialization Stage */
1. Determine the set $\mathcal{C}(d_{ij}) = \{d_{ij}\}_{i=1,j=1}^{M,M}$ and T_d from $\{\mathbf{l}_i\}_{i=1}^M$;
2. Construct the graph $\mathcal{G} \in (V, E)$ from $\{\mathcal{C}(d_{ij}), T_d\}$ using (11);
3. Sample normals on a *Ball*;
for $k = 1$ **to** n **do**
  4. Determine $(\mu_{ij}^-, \mu_{ij}^+), (i,j) \in E_k$ from $\{\eta, \mathcal{G}\}$;
end
5. Obtain the set $\mathcal{C}(\mu_{ij}^-, \mu_{ij}^+) = \{(\mu_{ij}^-, \mu_{ij}^+) | (i,j) \in E_k\}$;
6. Determine $\mathbf{D}$ and J_d from $\{\mathcal{C}(d_{ij}), \mathcal{G}\}$ using (16) and (17);

/* Normal Computation Stage */
for $k = 1$ **to** P **do**
  7. Determine $\boldsymbol{\mathcal{L}}$ and $G = \{g_1, g_2, \cdots, g_K\}$ from $\{\mathcal{C}(\mu_{ij}^-, \mu_{ij}^+), \mathbf{o}_k, \mathcal{G}\}$ using (19) and (22);
  8. Determine J_s from G using (24);
  9. Determine J_w from $\mathbf{o}_k$ as in (25);
  10. Solve $\mathbf{n}_k$ from $J_\text{d}, J_\text{s}, J_\text{w}$ using (27).
end

D. Algorithm

After obtaining the expressions of the three terms J_d, J_s, and J_w, the optimization problem can be finally formulated as

$$\begin{aligned} \mathbf{n} = \arg \min_{\mathbf{n},\mathbf{s},\mathbf{e},\tilde{\mathbf{u}},\tau} \; & \|\mathbf{D}\mathbf{s}\|_2^2 + \lambda_\text{s} \sum_{k=1}^{K} \beta_k \|\tilde{\mathbf{u}}_k\|_2 + \lambda_\text{w} \|\mathbf{w}^\mathsf{T}\boldsymbol{\tau}\|_1 \\ \text{s.t. } & \mathbf{O}\mathbf{s} = \mathbf{L}^\mathsf{T}\mathbf{n} + \mathbf{e} + \boldsymbol{\tau} \\ & \mathbf{e} = \mathbf{G}\tilde{\mathbf{u}}, \quad \mathbf{s} \geq \mathbf{0}. \end{aligned} \tag{27}$$

The constraint $\mathbf{s} \geq \mathbf{0}$ is due to the non-negative nature of reflectance. To avoid the degenerate solution, we solve the normals by setting $n_z = 1$ instead of the scaled normal. The optimization problem can be further formulated as a second order cone programming (SOCP) problem [28]. In this work the SOCP problem is solved using the MOSEK package [38], which uses the primal or dual interior point method [39].

The algorithm for estimating the normal map of a scene with P surface points is summarized in Algorithm 1. In the Algorithm, the graph is constructed at the *Initialization Stage*, and the surface normal is estimated in a pixel-wise manner at the *Normal Computation Stage*. The computational cost is mainly determined by solving the SOCP problem (27). In general, the SOCP problem is of polynomial complexity depending on the number of both variables and cones. Denote M the number of nodes in the graph, n the number of connected edges, and K the number of groups determined

by pixel intensities. Note that the n is proportional to M, and K is much smaller than M. Then the diffuse modeling term consists of about n variables and one cone. According to (24), the specular modeling term has less than $K \times M$ variables and K cones. The shadow modeling term requires $2M$ variables to formulate the ℓ_1-norm penalty. Hence the computational time mainly depends on the number of light directions M. Our investigation indicates that image noise and shadow almost do not affect the computational efficiency of the algorithm.

IV. Experiments

We evaluate the proposed method on both synthetic and real scenes, and compare its performance with the baseline least squares (LS) method [1] and three state-of-the-art methods, i.e., sparse Bayesian learning (SBL) [8], biquadratic reflectance model (BQ) [12], and constrained bivariate regression (CBR) [18]. The SBL is a Lambertian-enforced method that detects the sparsely distributed outliers. The BQ and CBR are both BRDF-based methods. The BQ method models the low-frequency reflection component, while the CBR method deals with both diffuse and specular reflections. In accordance with [18], the CBR method includes the retro-reflection detection step in the experiment.

We rendered 32-bit high dynamic range (HDR) synthetic image scenes using the BRDF models and MERL material database [2], [34] mainly on two surfaces, *Ball* and *Bunny*. In the rendered images, both attached and cast shadows are of zero intensity values. As the reflectance in the MERL database is of RGB format, the rendered color images were transformed to the grayscale ones. The image intensities were scaled to the range $[0, 1]$ and then uniformly quantized to 2^{32} levels. This data representation is similar to that in [8]. Our investigation indicates that the quantization error is negligible to normal estimation. The parameters of these methods were selected as follows. In the proposed method, the weight parameters $\lambda_s = 1$ and $\lambda_w = 1$. As shadow is ideal in the synthetic scenes, we set the weight factor ξ with a large value 10^7. With the incorporated shadow modeling term in the proposed method, shadow thresholding is no longer needed. The selection of the parameters η and m will be discussed in the following experiment. For the other methods, we adopted the parameter values as suggested in the corresponding works. The shadow pixels were first detected by a threshold 10^{-6}. The weight parameter of the SBL method is set as $\lambda = 10^{-6}$ and the low frequency threshold of the BQ method is $T_{\text{low}} = 25\%$. The order of the polynomial in the CBR method is $(3, 5)$, which is in accordance with [18].

In additional to the synthetic data, four real surfaces, i.e., *Lady*, *Horse*, *Boy*, and *Satin Cylinder* were used in the experiment. The *Lady* surface is of relatively ideal Lambertian diffuse reflection but spatially-varying albedos, the *Horse* and *Boy* surfaces are of metal-like materials and the *Satin Cylinder* surface is formed by wrapping an anisotropic reflectance *Satin* fabric. The images of these surfaces (see Fig. 15, Fig. 16, Fig. 17, and Fig. 19) were acquired using a single-view based imaging system illustrated in Fig. 6. In the system, the camera is fixed and its viewing direction is orthographic to the scene. The light sources are installed on a rotatable arc.

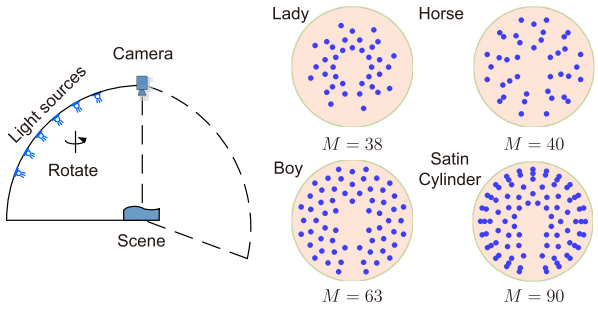

Fig. 6. (Left) Illustration of the imaging system for real data acquisition. The camera is orthographic to the scene, and light sources are installed on the arc which can be rotated around the viewing direction. (Right) The light configurations for the real surfaces. Note that, for the *Satin Cylinder* surface, three different light configurations are used, while only the one with $M = 90$ light directions are shown here.

The elevation angles of these light sources range from 20° to 70°, with an interval step 10°. The distance from light sources to scene are about 10 times larger than the object size. Hence it is reasonable to assume that the light sources are directional to object surface. The intensity variations of these light sources were corrected by imaging a white diffuse board. In the imaging process, the arc rotated around the viewing direction and the light sources illuminated the object in turn. The images were acquired under a set of multiple exposure times due to the severe intensity variation of real objects. The image intensities of some pixels may still be saturated even at the lowest exposure time. Based on these images, a 32-bit HDR image was generated for each object. For the real scene data, the weight parameter λ_s was set as 0.1. As shadow is complicated in the real scene case, the weight factor was set as $\xi = 10/\texttt{median}(\mathbf{o})$. By adopting the median operator, the effect of albedo variation can be eliminated to some extent. For the competing methods, we manually chose an appropriate threshold for shadow removal. The other parameters were set as follows. In the SBL method, $\lambda = 10^{-2}$. In the BQ method $T_{\text{low}} = 25\%$, and in the CBR method the order of polynomial is $(1, 5)$.

A. Influence of Factors

As will be shown in Section IV-B, the accuracy of normal estimation is generally higher when more light sources are employed. However, the improvement will be insignificant when the number of light sources M is quite large (100+). In this section we mainly investigate the influences of other factors on the proposed method, by setting the number of light directions $M = 50$ or $M = 100$.

We first explore the effects of parameter m in diffuse modeling and parameter η in specular modeling. The analysis is based on a system with 100 uniformly distributed light directions (see Fig. 7 (a)) and the 100 MERL materials. We compute the angular error of each material and evaluate the accuracy of the method in terms of the average angular error on all materials. As observed in Table II, the angular accuracy is very stable when the value of m varies from 1 to 4. The reason is that, in a system with uniformly distributed light directions, the distance d_m changes slightly.

Table III lists the angular errors with respect to the parameter η, which determines the bounds μ_{ij}^- and μ_{ij}^+ in

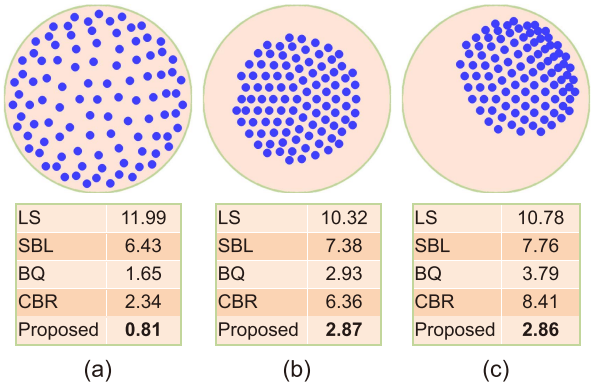

Fig. 7. Three different configurations of 100 light directions and the corresponding average angular errors (in degrees) produced by five methods. The errors are computed on the 100 MERL materials. (a) Uniformly distributed light directions covering the whole hemisphere. (b) Concentrated light directions. (c) Biased and concentrated light directions.

TABLE II

AVERAGE ANGULAR ERRORS (IN DEGREES) OF THE *Ball* SURFACE WITH RESPECT TO THE PARAMETER m WHEN FIXING $\eta = 80\%$. THE ERRORS ARE COMPUTED ON THE 100 MERL MATERIALS

m	1	2	3	4
Error	0.811	0.808	0.804	0.807

TABLE III

AVERAGE ANGULAR ERRORS (IN DEGREES) OF THE *Ball* SURFACE WITH RESPECT TO η VALUES WHEN FIXING $m = 4$. THE ERRORS ARE COMPUTED ON THE 100 MERL MATERIALS

η	70%	75%	80%	85%	90%	95%	99%
Error	0.80	0.80	0.81	0.81	0.82	0.83	0.84

specular modeling. It is observed the angular accuracy is also very stable. This is mainly because that the specular reflection changes abruptly under neighboring light directions. Hence the variation of $(\mu_{ij}^{-}, \mu_{ij}^{+})$ causes very limited influence on specular detection. Based on the above analysis, we fix $\eta = 80\%$ and $m = 4$ hereafter.

We then discuss the influence of light direction configuration on surface normal estimation. Figure 7 shows three different configurations and the corresponding angular errors of the estimated normals. In the first configuration (a), the 100 light directions cover the whole hemisphere and are uniformly distributed. In the second configuration (b), the light directions concentrate to some extent and cover partial hemisphere. In the third configuration (c), the concentrated light directions are further biased from the viewing direction. It is observed from Fig. 7 that, except for the LS method, almost all the methods perform better under the first configuration than under the second and third configurations. It is observed from (b) and (c) that the accuracy of the BQ and CBR methods degrades obviously when the light directions are biased. Interestingly, the accuracy of the proposed method almost keeps the same in the second and third configurations. In summary, the accuracy of the proposed method is affected by the concentration, but not the bias, of light directions. Hereafter we adopt the first configuration in the experiments.

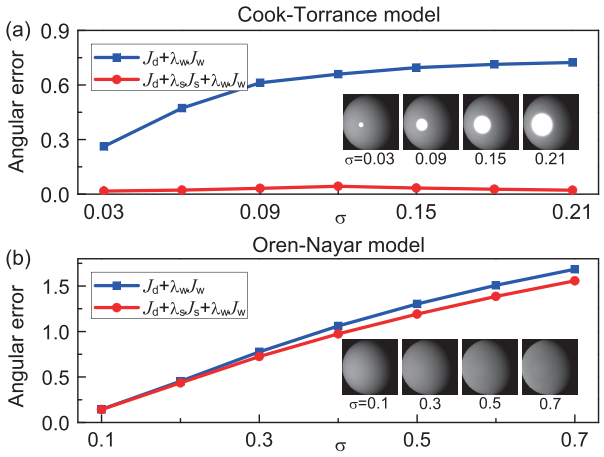

Fig. 8. Angular errors (in degrees) of the *Ball* surfaces rendered using two BRDF models under 50 light sources. (a) Cook-Torrance model. (b) Oren-Nayar model. Please refer to Appendix for the roughness parameter σ in the two models.

In the following, we verify the assumptions of the diffuse and specular modeling terms using images rendered by two BRDF models, i.e., the Cook-Torrance model [35] and Oren-Nayar model [40]. The expressions and parameter settings of the BRDF models are given in Appendix. The Cook-Torrance model consists of an ideal Lambertian diffuse term and a specular term. The Oren-Nayar model is a non-Lambertian diffuse reflection model without explicit specular term. Figure 8 illustrates the accuracy of the proposed method in case of with or without the specular modeling term J_s. For the Cook-Torrance model, the angular errors become quite lower under various roughness σ when the specular modeling term is employed. This verifies that the specular modeling term is effective in handling highlights. For the Oren-Nayar model, the additional employment of specular modeling term does not affect much the estimation accuracy. This validates that, for objects without specular highlight, the normal estimation is actually dominated by the diffuse modeling term.

We also investigate the effect of the shadow modeling term J_w in the proposed method. Our experimental results show that, when the shadow modeling term is not incorporated, the angular errors of the rendered objects will be above 5°. The reason is that the existence of shadow pixels breaks down the smoothness assumption of diffuse reflection, and consequently, the estimated normal is considerably biased. Therefore it is necessary to employ the shadow modeling term in the proposed method. Actually, as shadow has obvious influence on normal estimation, some state-of-the-art methods [12], [18] explicitly remove the shadow pixels in advance.

We further evaluate the robustness of the proposed method to image noise using the MERL materials. We adopt a commonly used noise model which includes both signal-dependent and signal-independent noises [41]. Before the quantization process, noise is first added to the observed intensity o as $\hat{o} = o + \sqrt{\sigma_a^2 o + \sigma_b^2} X$, where $X \sim \mathcal{N}(0, 1)$ is a random variable following Gaussian distribution with zero mean and standard deviation 1.0. Here σ_a and σ_b denote the standard deviations of the signal-dependent and independent

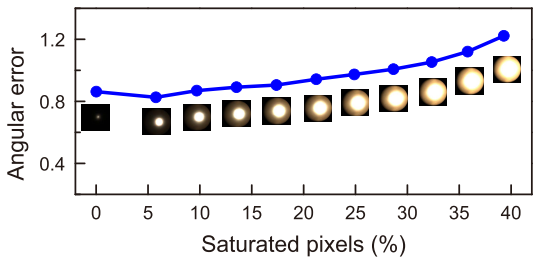

Fig. 9. Angular errors (in degrees) of the *two-layer-gold* material with respect to the percent of saturated pixels under 50 light sources.

TABLE IV
AVERAGE ANGULAR ERRORS (IN DEGREES) OF THE *Ball* SURFACE COMPUTED ON THE 100 MERL MATERIALS UNDER VARIOUS NOISE LEVELS. THE PARAMETERS $\sigma_a(\times 10^{-3})$ AND $\sigma_b(\times 10^{-7})$ ARE THE STANDARD DEVIATION OF THE SIGNAL-DEPENDENT AND SIGNAL-INDEPENDENT NOISE RESPECTIVELY

(σ_a, σ_b)	BQ	CBR	Proposed
$(0, 0)$	1.65	2.34	**0.81**
$(1, 0)$	1.91	2.37	**0.93**
$(0, 2)$	1.66	2.34	**0.94**
$(1, 2)$	1.92	2.37	**1.06**
$(3, 0)$	3.11	2.46	**1.83**
$(0, 10)$	18.78	26.20	**3.78**
$(3, 10)$	19.27	26.22	**4.39**

noises, respectively. Then we apply the 32-bit quantization process as mentioned above. Table IV lists the average angular errors of the *Ball* surface under various noise levels. The errors of all the methods increase a little in case of small noise levels. However, when the signal-independent noise reaches the shadow detection threshold, the accuracy of the BQ and CBR methods degrades abruptly. This is because that the shadow pixels are misclassified as the illuminated ones, which greatly biases the normal estimation. In comparison, the angular errors of the proposed method are relatively low, thanks to its reasonable modeling of the shadow component.

Besides noise, the images of some materials may also suffer from intensity saturation due to the extreme strong specular highlight. In the following we evaluate the robustness of the proposed method to intensity saturation. In the experiment, intensity saturation is simulated by varying exposure time. Figure 9 illustrates the angular errors of the *two-layer-gold* material with respect to the percentage of saturated pixels. It is observed that the angular errors do not change largely when 20% or less pixels are saturated. The angular error is just around 1.2° even at a very large percentage (e.g., 40%) of saturation. This validates that the proposed method is robust to intensity saturation.

As known, the proposed method runs in a pixel-wise manner and hence the normal estimation should not be affected by spatially-varying albedo. To illustrate this, in Fig. 10 we rendered the *Mozart* surface using either four different materials or a single material. It is observed that the estimated normal maps in both cases are quite close to the ground truth. The average difference between the two estimated normal maps is below 1.2°.

TABLE V
AVERAGE ANGULAR ERRORS (IN DEGREES) OF THE *Ball* AND *Bunny* SURFACES PRODUCED BY DIFFERENT METHODS WHEN USING VARIOUS NUMBERS OF LIGHT DIRECTIONS.[2] THE ERRORS ARE COMPUTED ON THE 100 MERL MATERIALS

Ball					
M	LS	SBL	BQ	CBR	Proposed
20	10.56	6.57	-	5.01	**3.22**
50	12.30	6.68	1.90	3.55	**1.56**
100	11.99	6.43	1.65	2.34	**0.81**
150	11.89	6.36	1.17	1.89	**0.55**
200	11.27	6.15	1.07	1.90	**0.49**
Bunny					
M	LS	SBL	BQ	CBR	Proposed
20	12.19	7.63	-	7.27	**4.22**
50	13.51	7.60	2.93	6.05	**2.81**
100	13.33	7.32	2.51	4.59	**2.09**
150	13.27	7.23	2.16	3.98	**1.77**
200	12.82	7.06	2.04	3.81	**1.69**

B. Results on Synthetic Data

In this section we evaluate the accuracy of different methods using synthetic data. Table V lists the average angular errors of the *Ball* and *Bunny* surfaces rendered with 100 MERL materials under various numbers of light directions. For all methods, the errors of the *Bunny* surface are larger than those of the *Ball* surface in case of identical number of light directions. This is because that the extra cast shadow of *Bunny* decreases the valid number of data. It is observed that the angular errors decrease when more light directions are employed. In all cases, the proposed method performs the best, followed by the BQ and CBR methods. The SBL method performs better than the LS method, but its angular errors are still quite large. It is worth noting that the accuracy of the proposed method under 100 light directions is better than, or comparable to, that of the BQ and CBR methods under 200 light directions.

Figure 11 illustrates the angular errors for the *Ball* surface rendered using 100 MERL materials. Overall, the proposed method performs the best, followed by the BQ and CBR methods. Figure 12 further illustrates the spatial angular errors of the *Bunny* surface rendered with the seven typical materials chosen in Fig. 11. Among these seven materials, three ones (*blue-rubber*, *beige-fabric*, and *pink-plastic*) contain mainly diffuse reflection, while the remaining ones have obvious specular highlight. It is observed from Fig. 12 that, for the diffuse materials, the proposed method performs better than other methods on *blue-rubber* and *beige-fabric*. However, its angular error is larger than those of the competitors on *pink-plastic*. For the specular materials, the proposed method performs the best on *two-layer-gold* and *red-metallic-paint*,

[2]The angular errors of the BQ method with 20 light directions is omitted because the valid number of data is not adequate for fitting.

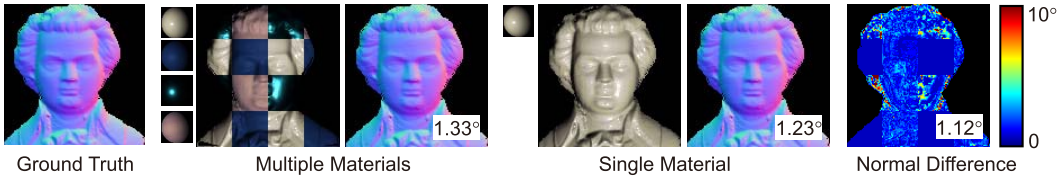

Fig. 10. Influence of spatially-varying BRDFs on normal estimation when using 50 light directions. In the multiple materials case, the *Mozart* surface is rendered using four materials (*white-acrylic*, *blue-rubber*, *green-metallic-paint*, and *pink-fabric*). In the single material case, the surface is rendered using the *white-acrylic* material. The rightmost figure illustrates the difference of the two estimated normal maps.

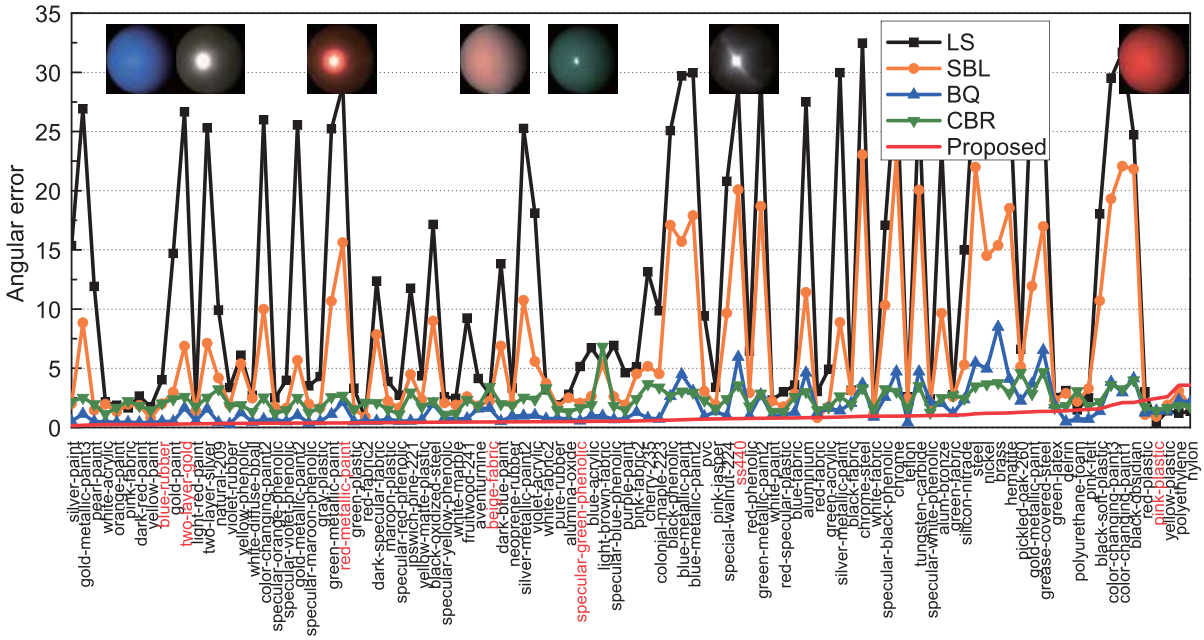

Fig. 11. Angular errors (in degrees) of the *Ball* surface on the 100 MERL materials in case of 100 light sources. Seven typical materials are chosen for further discussion. Their material names are in red and the corresponding images are shown on the top.

but performs slightly worse than the BQ method on *specular-green-phenolic*. For the metal material *ss440*, the angular error produced by the competing methods are around 6.5° or much larger. In comparison, the angular error produced by the proposed method is around 3.5°. This validates the capability of the proposed method in dealing with metal-like materials.

Figure 13 further illustrates the angular errors of the seven MERL materials under various numbers of light directions. As expected, the angular errors decrease when adopting more light directions. For most materials, the angular errors converge when the number of light directions M is larger than 100. The angular accuracy of the *blue-rubber* material is satisfactory when using only 50 light directions. For many materials (*ss440*, *beige-fabric*, *specular-green-phenolic*, *red-metallic-paint*, and *two-layer-gold*), the adoption of 100 light directions is a good choice in balancing the trade-off between system complexity and estimation accuracy. For the *pink-plastic* material, however, the accuracy is not so satisfactory even when 200 light directions are employed. The above findings indicate that, for materials with different reflection properties, the necessary numbers of light directions could be quite different.

In the following, we evaluate the methods using anisotropic materials. Figure 14 shows the example images and the normal errors when using $M = 50$, 100, and 150 light directions. The images are rendered using two anisotropic BRDF models, i.e., the Kurt model [36] and the Ashikhmin-Shirley (AS) model [37]. The expressions and parameter settings of both models are given in Appendix. The images rendered by the Kurt model do not contain strong specular reflections. The proposed method produces quite low angular errors and significantly outperforms the other methods. On the other hand, the images rendered by the AS model have heavy specular highlights. Although the proposed method performs best in most circumstances, its advantage over the BQ method is not obvious. These findings indicate that the proposed method can handle anisotropic materials when an appropriate number of light directions is employed.

C. Results on Real Scenes

We evaluate the methods on four real surfaces, namely, *Lady*, *Horse*, *Boy*, and *Satin Cylinder*. In accordance to the complexity of the reflectances, the numbers of light directions are different for these surfaces (see Fig. 6). Based on the estimated surface normals, the 3D surfaces are constructed using an ℓ_1-norm based integration method which is robust to noise and outlier corruptions [42]. For the *Lady*, *Horse* and *Boy* surfaces, only visual comparisons are conducted. For the *Satin Cylinder* surface, quantitative comparison is also presented.

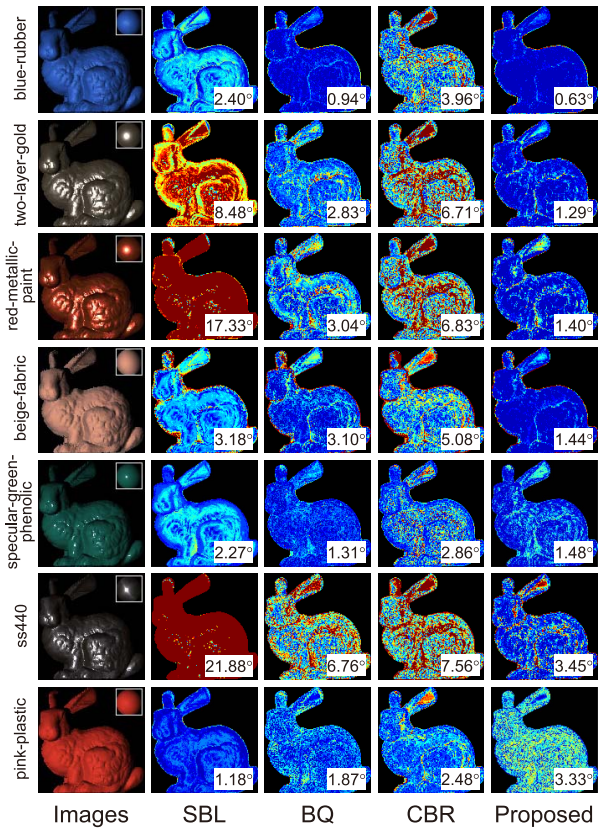

Fig. 12. Angular errors (in degrees) of the *Bunny* surface under 100 light directions rendered using seven MERL materials. The angular errors range from 0° to 10°. The higher errors are encoded by warmer colors. The images of materials *two-layer-gold*, *red-metallic-paint*, and *ss440* are nonlinearly adjusted for better visualization.

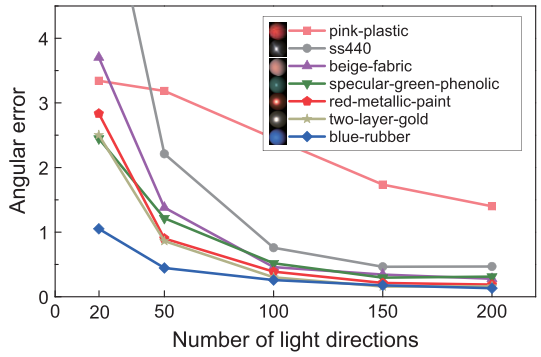

Fig. 13. Angular errors (in degrees) of the *Ball* surface on seven MERL materials when using various numbers of light directions.

The *Lady* surface shown in Fig. 15 is of relatively ideal Lambertian diffuse reflection, strong specular reflection, and spatially-varying albedos. It is observed that the normal maps produced by the CBR method is biased by specular highlight in the second close-up view. In comparison, the proposed method performs the best in the two close-up views, followed by the SBL and BQ methods.

The *Horse* surface in Fig. 16 is of metal-like material and complex geometry. There are either shadow or highlight in the image areas corresponding to the two close-up views. As observed, the SBL method produces noisy normals and the CBR method results in artifacts. The 3D surface produced by the BQ method appears flat due to the misclassification of low frequency reflection component. In comparison, the proposed method produces both good normal map and 3D shape for this complex object.

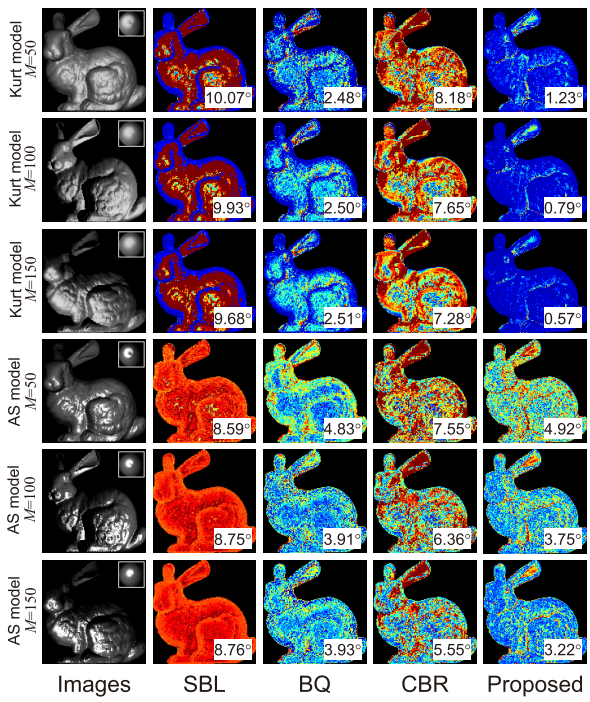

Fig. 14. Angular errors (in degrees) of the *Bunny* surfaces rendered using the anisotropic Kurt model [36] and AS model [37]. The results under various numbers ($M = 50$, 100, 150) of light directions are illustrated. The errors range from 0° to 10°. The higher errors are encoded by warmer colors.

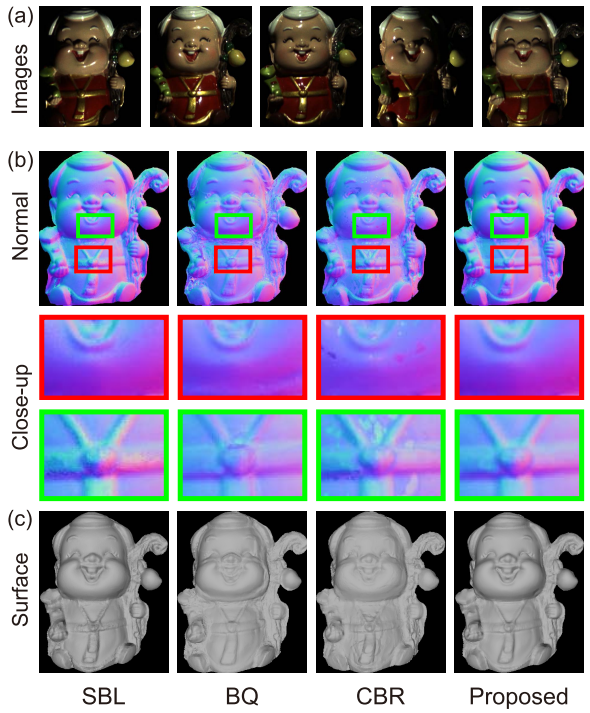

Fig. 15. Results of the real *Lady* surface. (a) Example images. (b) Estimated normal maps. (c) Reconstructed 3D surfaces.

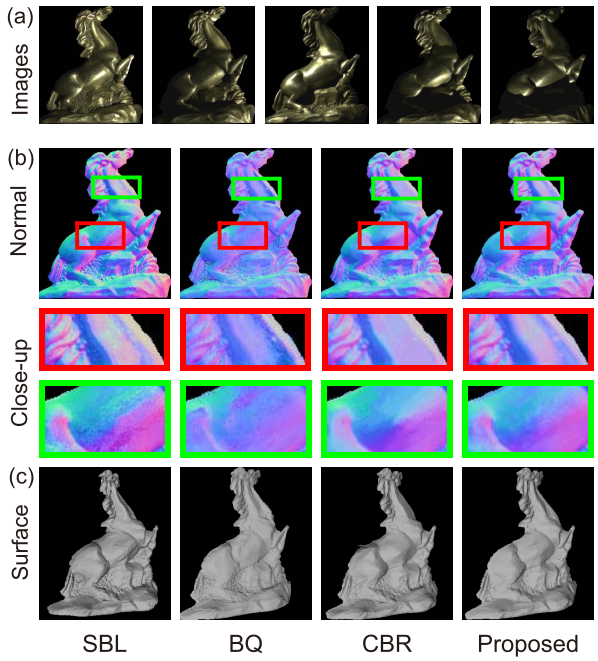

Fig. 16. Results of the real *Horse* surface. (a) Example images. (b) Estimated normal maps. (c) Reconstructed 3D surfaces. The images in (a) are nonlinearly adjusted and the surfaces in (c) are rotated for better visualization.

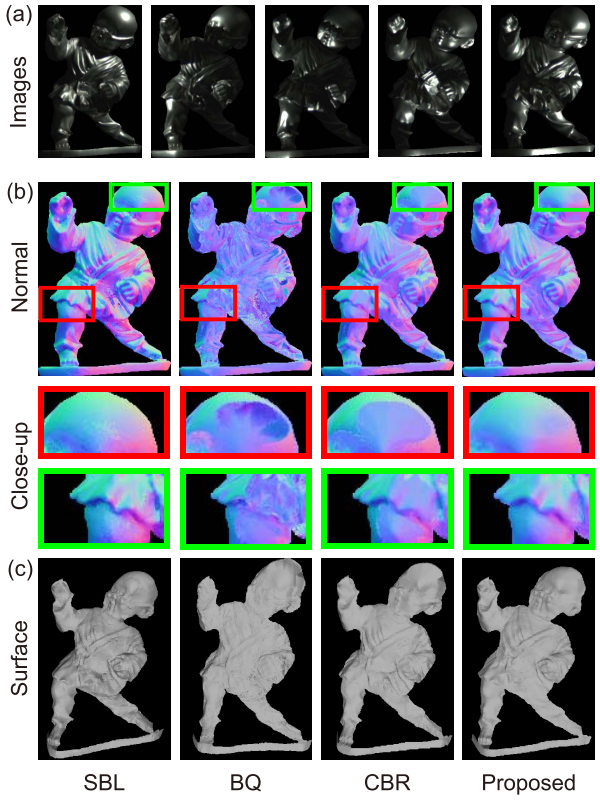

Fig. 17. Results of the real *Boy* surface. (a) Example images. (b) Estimated normal maps. (c) Reconstructed 3D surfaces. The images in (a) are nonlinearly adjusted and the surfaces in (c) are rotated for better visualization.

The *Boy* surface is also of metal-like material with not only complicated reflection but also a large amount of shadow, as shown in Fig. 17. It is observed that the results produced by the BQ and CBR methods are not acceptable. The normal maps produced by the SBL method contain noticeable noise and artifacts. In contrast, the proposed method performs better regarding both the normal map and 3D reconstruction.

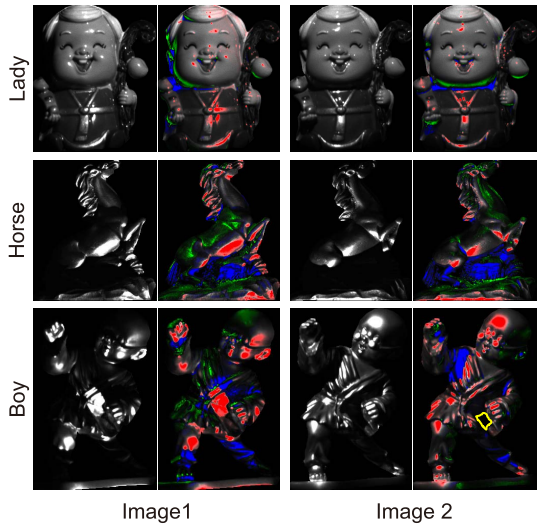

Fig. 18. The example images of the real surfaces and corresponding reflection labeling results. The specular highlight, attached shadow, and cast shadow are labeled in red, green, and blue. The overall reflection labeling is satisfactory, but it is incorrect In the marked area in the right-bottom image.

Figure 18 illustrates the reflection labeling results of the real surfaces *Lady*, *Horse*, and *Boy*. Note that the shadow is further classified as the attached one and cast one according to the principle stated in Section III-C. It is observed that the reflection labeling results agree well with visual judgement. However, as the reflection characteristic of real surface is rather complex, the labeling for some pixels may be erroneous. For example, in the right-bottom image, the shadow pixels in the marked area is incorrectly labeled as diffuse ones. Nevertheless, the overall reliability of reflection labeling is good, and consequently the estimated normal maps of these three real surfaces are satisfactory.

The *Satin Cylinder* surface was formed by wrapping a *Satin* fabric, which is of anisotropic reflectance, onto a cylinder surface. The camera was geometrically calibrated such that the central axis of the cylinder aligns with the column direction of the image. The diameter of the cylinder was determined according to the object boundary. The ground truth surface normal and height were computed using the obtained cylinder diameter. In additional to the mean angular error of surface normals, we also compute the mean height error $\epsilon = \frac{1}{P}\sum_{k=1}^{P}|z_k - \tilde{z}_k|$, where z_k and $\tilde{z}_k$ are, respectively, the heights of ground truth and reconstructed surfaces, and P is the number of pixels. Note that, due to the integration method, lower angular errors do not guarantee lower height errors. Table VI illustrates the mean angular errors and mean height errors of different photometric stereo methods when using $M = 50, 90, 130$ light directions. It is observed that, in all cases, the angular and height errors of the proposed method are lower than those of the competitors. The improvement of normal accuracy is not significant when the number of light directions increases from $M = 90$ to $M = 130$. This finding agrees with the trends in the synthetic data experiment, as shown in Fig. 13.

TABLE VI
MEAN ANGULAR ERRORS (IN DEGREES) AND MEAN HEIGHT ERRORS (IN PIXELS) OF THE *Satin Cylinder* SURFACE WHEN USING VARIOUS NUMBERS OF LIGHT DIRECTIONS

Mean angular errors				
M	SBL	BQ	CBR	Proposed
50	16.27	9.14	10.35	**6.55**
90	14.07	8.87	7.63	**4.85**
130	15.94	8.80	7.40	**4.58**
Mean height errors				
M	SBL	BQ	CBR	Proposed
50	12.71	4.43	6.23	**3.57**
90	10.70	4.62	3.20	**2.18**
130	11.88	4.43	3.25	**2.15**

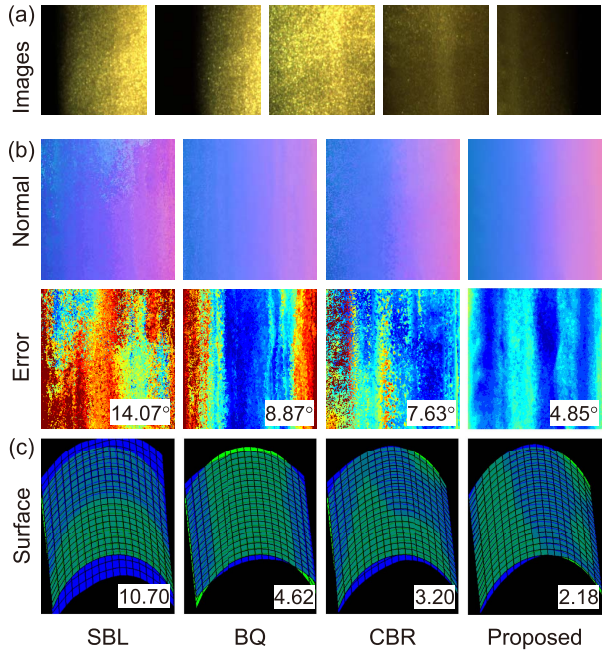

Fig. 19. Results of the real *Satin Cylinder* surface under 90 light directions. (a) Example images. (b) Estimated normal maps and their angular errors with respect to the ground truth. The angular errors range from 0° to 20°. The higher errors are encoded by warmer color. (c) Reconstructed surfaces (in blue) and ground truth surfaces (in green). The mean height errors are shown on the right-bottom of the reconstructed surfaces.

Figure 19 illustrates the estimated normal maps and reconstructed surfaces of the *Satin Cylinder* surface when using $M = 90$ light directions. As shown, the spatial distribution of normal errors produced by the proposed method is much lower than those by the other methods. In addition, the 3D surface reconstructed by the proposed method is quite close to the ground truth surface. This verifies the capability of the proposed method in handling real anisotropic materials.

D. Computational Time

Figure 20 illustrates the mean computational time on the 100 MERL materials for each pixel. All methods were implemented using MATLAB on a personal computer with 2.20 GHz GPU and 8 GB RAM. As shown in Fig. 20, the LS method runs very fast due to its simplicity in computation.

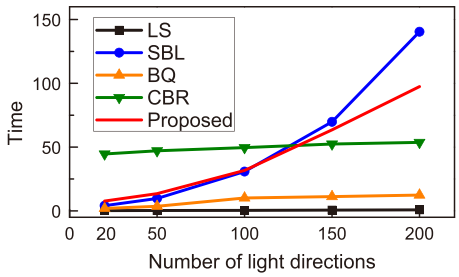

Fig. 20. Average computational times (unit: millisecond) for each pixel of different methods in case of different number of light sources with surface *Ball*.

The computational time of the BQ and CBR method is nearly constant with respect to the number of light directions M. The computational time of the proposed time is approximately proportional to number M. Overall, the proposed method runs faster than the SBL method. Compared with the CBR method, the proposed method runs faster when using 100 or less light directions, but runs slower when using 150 or more light directions. It will be desired if the computational efficiency of the proposed method can be further improved.

V. CONCLUSIONS

This paper has proposed a novel photometric stereo method for object surfaces with various unknown BRDFs by modeling the general characteristics of reflection components. Based on a graph of light directions, the diffuse reflection is modeled as the summation of local variation, while the specular and shadow components are modeled using group sparse and weighted ℓ_1-norm optimization, respectively. The whole optimization problem of photometric stereo is formulated as a second order cone programming (SOCP) problem. The proposed method is extensively evaluated on both synthetic data and real scenes. Experimental results validate that the proposed method is robust to non-Lambertian reflections and is capable of dealing with complex materials (e.g., metal-like, anisotropic) with good accuracy.

There are three limitations in the proposed photometric stereo method. The first limitation is that the method uses directional lighting and cannot acquire images in real time. The second limitation is the high computational complexity when using a general software package. The third limitation is that, for object surfaces with complex BRDFs, the estimation accuracy is not satisfactory when using an extremely small number of light directions.

APPENDIX
BRDF MODELS

The expressions and parameter settings for the BRDF models used in this work are given below. Some common notations and their corresponding meanings are listed in Table VII.

A. Cook-Torrance Model [35]

The expression of Cook-Torrance model is

$$\rho = k_d + k_s \frac{DGF}{(\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v})},$$

TABLE VII
NOTATIONS USED IN THE BRDF MODELS

Notations	Meanings
$\mathbf{h}$	Unit-length halfway vector $(\mathbf{l}+\mathbf{v})/\|(\mathbf{l}+\mathbf{v})\|_2$
θ_i,θ_o,θ_h	Angles between $\mathbf{n}$ and $\mathbf{l}$, $\mathbf{n}$ and $\mathbf{v}$, $\mathbf{n}$ and $\mathbf{h}$
ϕ_i, ϕ_o, ϕ_h	Azimuthal angles of $\mathbf{l}$, $\mathbf{v}$, and $\mathbf{h}$
F	Fresnel reflectance $f_0+(1-f_0)(1-(\mathbf{l}\cdot\mathbf{h})^5)$
k_d,k_s	Diffuse albedo and specular reflectivity

where

$$D=\frac{1}{\sigma^2\cos^4\theta_h}\exp\left(-\frac{\tan^2\theta_h}{\sigma^2}\right),$$

is the fraction of facets, and

$$G=\min\left\{1,\frac{2(\mathbf{n}\cdot\mathbf{h})(\mathbf{n}\cdot\mathbf{v})}{(\mathbf{v}\cdot\mathbf{h})},\frac{2(\mathbf{n}\cdot\mathbf{h})(\mathbf{n}\cdot\mathbf{l})}{(\mathbf{v}\cdot\mathbf{h})}\right\},$$

is the geometrical attenuation. In this work, the parameters are set as $k_d=0.5$, $k_s=0.5$, and $f_0=0.8$. The roughness σ ranges from 0.03 to 0.21.

B. Oren-Nayar Model [40]

The Oren-Nayar model is expressed as

$$\rho=A+B\max(0,\cos(\phi_i-\phi_o))\sin\alpha\tan\beta,$$

where

$$\alpha=\max(\theta_i,\theta_o),\quad \beta=\min(\theta_i,\theta_o),$$

and

$$A=1-0.5\frac{\sigma^2}{\sigma^2+0.33},\quad B=0.45\frac{\sigma^2}{\sigma^2+0.09}.$$

In this work the roughness σ ranges from 0.1 to 0.7.

C. Kurt Model [36]

The Kurt model is expressed as

$$\rho=\frac{k_d(1-F)}{\pi}+\frac{k_sFD}{4(\mathbf{v}\cdot\mathbf{h})((\mathbf{n}\cdot\mathbf{l})(\mathbf{n}\cdot\mathbf{v}))^{\alpha}},$$

where the fraction of facets

$$D=\frac{1}{\pi m_x m_y\cos^4\theta_h}\exp\left(-\tan^2\theta_h\left(\frac{\cos^2\phi_h}{m_x^2}+\frac{\sin^2\phi_h}{m_y^2}\right)\right).$$

The parameters are set as $k_d=0.75$, $k_s=0.25$, $m_x=0.3$, $m_y=1.5$, $\alpha=0$, and $f_0=0.8$. The tangent is set 45° off the plane determined by the viewing and normal directions.

D. Ashikhmin-Shirley (AS) Model [37]

The AS model includes a non-Lambertian diffuse term and an anisotropic specular term. The model is expressed as

$$\rho=\rho_d+\rho_s,$$

where diffuse reflectance

$$\rho_d=\frac{28k_d}{23\pi}(1-k_s)\left(1-\left(1-\frac{\mathbf{n}\cdot\mathbf{l}}{2}\right)^5\right)\left(1-\left(1-\frac{\mathbf{n}\cdot\mathbf{v}}{2}\right)^5\right),$$

and the specular reflectance

$$\rho_s=\frac{\sqrt{(n_u+1)(n_v+1)}}{8\pi}\frac{(\mathbf{n}\cdot\mathbf{h})^{n_u\cos^2\phi_h+n_v\sin^2\phi_h}}{(\mathbf{v}\cdot\mathbf{h})\max(\mathbf{n}\cdot\mathbf{l},\mathbf{n}\cdot\mathbf{v})}F.$$

In this work the parameters are set as $k_d=0.75$, $k_s=0.25$, $n_u=100$, $n_v=1$, and $f_0=0.8$. The tangent is set 45° off the plane determined by the viewing and normal directions.

ACKNOWLEDGMENT

The authors sincerely thank the anonymous reviewers for their valuable comments that have improved the quality of this paper.

REFERENCES

[1] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Opt. Eng.*, vol. 19, no. 1, p. 191139, 1980.
[2] A. Ngan, F. Durand, and W. Matusik, "Experimental analysis of BRDF models," in *Proc. 16th Eurograph. Conf. Rendering Techn.*, 2005, pp. 117–126.
[3] T.-P. Wu, K.-L. Tang, C.-K. Tang, and T.-T. Wong, "Dense photometric stereo: A Markov random field approach," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1830–1846, Nov. 2006.
[4] Y. Mukaigawa, Y. Ishii, and T. Shakunaga, "Analysis of photometric factors based on photometric linearization," *J. Opt. Soc. Amer. A*, vol. 24, no. 10, pp. 3326–3334, 2007.
[5] D. Miyazaki, K. Hara, and K. Ikeuchi, "Median photometric stereo as applied to the Segonko Tumulus and Museum objects," *Int. J. Comput. Vis.*, vol. 86, nos. 2–3, pp. 229–242, 2010.
[6] C. Yu, Y. Seo, and S. W. Lee, "Photometric stereo from maximum feasible Lambertian reflections," in *Proc. 11th Eur. Conf. Comput. Vis.*, 2010, pp. 115–126.
[7] 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. Asian Conf. Comput. Vis.*, 2010, pp. 703–717.
[8] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Robust photometric stereo using sparse regression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 318–325.
[9] A. S. Georghiades, "Incorporating the Torrance and Sparrow model of reflectance in uncalibrated photometric stereo," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, Oct. 2003, pp. 816–823.
[10] H.-S. Chung and J. Jia, "Efficient photometric stereo on glossy surfaces with wide specular lobes," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.
[11] D. B. Goldman, B. Curless, A. Hertzmann, and S. M. Seitz, "Shape and spatially-varying BRDFs from photometric stereo," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 6, pp. 1060–1071, Jun. 2010.
[12] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "Bi-polynomial modeling of low-frequency reflectances," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 6, pp. 1078–1091, Oct. 2013.
[13] M. Holroyd, J. Lawrence, G. Humphreys, and T. Zickler, "A photometric approach for estimating normals and tangents," *ACM Trans. Graph.*, vol. 27, no. 5, p. 133, 2008.
[14] P. Tan, L. Quan, and T. Zickler, "The geometry of reflectance symmetries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2506–2520, Dec. 2011.
[15] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "Elevation angle from reflectance monotonicity: Photometric stereo for general isotropic reflectances," in *Proc. 12th Eur. Conf. Comput. Vis.*, 2012, pp. 455–468.
[16] T. Higo, Y. Matsushita, and K. Ikeuchi, "Consensus photometric stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1157–1164.
[17] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Photometric stereo using sparse Bayesian regression for general diffuse surfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 9, pp. 1816–1831, Sep. 2014.
[18] S. Ikehata and K. Aizawa, "Photometric stereo using constrained bivariate regression for general isotropic surfaces," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2187–2194.
[19] N. G. Alldrin and D. J. Kriegman, "Toward reconstructing surfaces with arbitrary isotropic reflectance: A stratified photometric stereo approach," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.

[20] M. Chandraker, J. Bai, and R. Ramamoorthi, "On differential photometric reconstruction for unknown, isotropic BRDFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 12, pp. 2941–2955, Dec. 2013.
[21] A. Hertzmann and S. M. Seitz, "Example-based photometric stereo: Shape reconstruction with general, varying BRDFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1254–1264, Aug. 2005.
[22] R. F. V. Saracchini, J. Stolfi, and H. C. da Gama Leitão, "A uniform grid structure to speed up example-based photometric stereo," *IEEE Trans. Image Process.*, vol. 20, no. 12, pp. 3495–3507, Dec. 2011.
[23] T. Okabe, I. Sato, and Y. Sato, "Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting conditions," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 1693–1700.
[24] M. K. Chandraker, S. Agarwal, and D. J. Kriegman, "ShadowCuts: Photometric stereo with shadows," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
[25] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. Roy. Statist. Soc. B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.
[26] G. Obozinski, L. Jacob, and J.-P. Vert. (2011). "Group lasso with overlaps: The latent group lasso approach." [Online]. Available: http://arxiv.org/abs/1110.0413
[27] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Optimization with sparsity-inducing penalties," *Found. Trends Mach. Learn.*, vol. 4, no. 1, pp. 1–106, 2012.
[28] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[29] E. N. Coleman, Jr., and R. Jain, "Obtaining 3-dimensional shape of textured and specular surfaces using four-source photometry," *Comput. Graph. Image Process.*, vol. 18, no. 4, pp. 309–328, 1982.
[30] S. Barsky and M. Petrou, "The 4-source photometric stereo technique for three-dimensional surfaces in the presence of highlights and shadows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1239–1252, Oct. 2003.
[31] S. P. Mallick, T. E. Zickler, D. J. Kriegman, and P. N. Belhumeur, "Beyond lambert: Reconstructing specular surfaces using color," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 2. Jun. 2005, pp. 619–626.
[32] K. E. Torrance and E. M. Sparrow, "Theory for off-specular reflection from roughened surfaces," *J. Opt. Soc. Amer.*, vol. 57, no. 9, pp. 1105–1112, 1967.
[33] G. J. Ward, "Measuring and modeling anisotropic reflection," *ACM SIGGRAPH Comput. Graph.*, vol. 26, no. 2, pp. 265–272, 1992.
[34] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A data-driven reflectance model," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 759–769, 2003.
[35] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Trans. Graph.*, vol. 1, no. 1, pp. 7–24, 1982.
[36] M. Kurt, L. Szirmay-Kalos, and J. Křivánek, "An anisotropic BRDF model for fitting and Monte Carlo rendering," *ACM SIGGRAPH Comput. Graph.*, vol. 44, no. 1, 2010, Art. ID 3.
[37] M. Ashikhmin and P. Shirley, "An anisotropic Phong BRDF model," *J. Graph. Tools*, vol. 5, no. 2, pp. 25–32, 2000.
[38] MOSEK ApS. *The MOSEK Optimization Software*. [Online]. Available: http://www.mosek.com/, accessed Jul. 20, 2015.
[39] E. D. Andersen, C. Roos, and T. Terlaky, "On implementing a primal-dual interior-point method for conic quadratic optimization," *Math. Program.*, vol. 95, no. 2, pp. 249–277, 2003.
[40] M. Oren and S. K. Nayar, "Generalization of Lambert's reflectance model," in *Proc. 21st Annu. Conf. Comput. Graph. Interact. Techn.*, 1994, pp. 239–246.
[41] S. Pyatykh and J. Hesser, "Image sensor noise parameter estimation by variance stabilization and normality assessment," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 3990–3998, Sep. 2014.
[42] Y. Cheng, H.-L. Shen, and X. Du, "Robust surface reconstruction from gradient fields," *Electron. Lett.*, vol. 48, no. 7, pp. 375–376, 2012.

Tian-Qi Han received the B.Eng. degree from Tianjin University, China, in 2011. He is currently pursuing the Ph.D. degree with the College of Information Science and Electronic Engineering, Zhejiang University, China. His research interest is computer vision, especially in photometric stereo.

Hui-Liang Shen received the B.Eng. and Ph.D. degrees from Zhejiang University, China, in 1996 and 2002, respectively. He was a Research Associate and Research Fellow with The Hong Kong Polytechnic University, from 2001 to 2005. He is currently a Full Professor with the College of Information Science and Electronic Engineering, Zhejiang University. His research interests include color imaging, image processing, and 3D computer vision.