Efficient Photometric Stereo Using Kernel Regression

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

***Abstract*— Photometric stereo estimates surface normals from multiple images captured under different light directions using a fixed camera. To deal with non-Lambertian reflections, the recent photometric stereo methods employ iterative or optimization frameworks that are computationally expensive. This paper proposes an efficient photometric stereo method using kernel regression, which can be transformed to an eigendecomposition problem. The kernel parameter is variable for each surface point so that it can cope with the variety of general reflectances. The best kernel parameter is automatically determined by leave-one-out cross validation. To improve computational efficiency, the leave-one-out process is accelerated by fast matrix computation and proper normal initialization. The proposed photometric stereo method is extensively evaluated on synthetic and real surfaces with various reflectances. Experimental results validate that the method is computationally efficient and achieves the state-of-the-art accuracy in surface normal estimation.**

***Index Terms*— Photometric stereo, kernel regression, eigen decomposition, surface normal, surface reconstruction, non-Lambertian, general reflectance, BRDF, variable kernel, leave-one-out, cross-validation, acceleration, computational efficiency.**

I. Introduction

PHOTOMETRIC stereo has been widely employed in surface reconstruction since the pioneer work of Woodham [1]. In photometric stereo, an object surface is illuminated by multiple directional light sources and its images are captured by a camera at a fixed viewpoint [2]. Under the assumption of ideal Lambertian reflection, surface normal can be computed from only three images. For object surface with non-Lambertian reflection, a larger number of images are always necessary. By treating highlight and shadow pixels as outliers, surface normal can be computed from the remaining pixels under the assumption of Lambertian reflection [3]–[10].

For many real object surfaces, however, highlight detection is not an easy task due to rough specular highlight or even non-Lambertian diffuse reflection. In recent years, some works handle various bidirectional reflectance distribution functions (BRDFs) [11], [12] by either employing parametric reflectance models [13]–[17] or exploiting reflectance properties [18]–[26]. A limitation is that, as most methods employ iterative scheme or optimization framework, normal estimation is computationally expensive. This limitation becomes a practical issue when a large number of light sources are used in photometric stereo.

This work proposes an efficient photometric stereo method for surfaces with general reflectances, under the condition of known light directions. In the method, normal estimation is formulated as kernel regression, which is further formulated as eigen decomposition. To our knowledge, despite its application in image restoration [27], kernel regression has not been introduced in photometric stereo. To appropriately describe the reflectances of various materials, the best kernel parameters are determined by leave-one-out (LOO) cross-validation. In the LOO process, the computation of inverse Gram matrix is accelerated by fast matrix manipulation, and eigen decomposition is speeded up by proper normal initialization. By this mean, the computational efficiency of the proposed photometric stereo method is much improved. Experiments have been conducted to evaluate the proposed method on both synthetic and real-world scenes.

A. Related Work

Photometric stereo was introduced based on the strict assumption of Lambertian reflection and known directional illumination condition [1]. It has been extended in early works by introducing extra light sources to deal with highlight and shadow [3], [4]. To cope with heavy highlight or shadow, some works [5]–[10] employ more images and treat highlight or shadow as outliers, while keeping the underlying Lambertian diffuse reflection assumption. The outliers have been detected using iterative schemes such as random sample consensus (RANSAC) [5], graph cuts [6], and median filtering (MH10, [7]). Alternatively, outlier detection has been cast as optimization problems that can be solved by integer programming [8], low-rank matrix recovery (WG10, [9]), or sparse Bayesian learning (IW12, [10]). With well formulated optimization frameworks, these methods are capable of producing improved accuracy on materials with Lambertian diffuse reflection. The computational burden of these iterative or optimization methods can be very heavy when employing a large number of light sources.

The outlier-detection based methods cannot work well on object surfaces with rough specular highlight or

Manuscript received April 7, 2016; revised August 26, 2016, September 27, 2016, and November 4, 2016; accepted November 4, 2016. Date of publication November 10, 2016; date of current version November 29, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 61371160 and Grant 61571392, and in part by the National Program for Special Support of Eminent Professionals. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Xiaochun Cao. *(Corresponding author: Hui-Liang Shen.)*

The authors are with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: shenhl@zju.edu.cn; htq2002@163.com; cgli@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.2016.2627805

1057-7149 © 2016 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

non-Lambertian diffuse reflection. To deal with complex reflectances, some works use typical BRDF models [13], [14] or a set of non-parametric basis BRDFs [15], [16]. Others explore the general reflectance properties such as isotropy, reciprocity symmetry, monotonicity, and reflection sparsity [20]–[26]. Instead of being regarded as outlier, in recent works shadow has also been employed for normal estimation based on simple shadow detection [13], [28] or Lambertian reflection assumption [29].

It is noted that general real-world reflectances are complex [11], [12] and cannot be described by classical parametric BRDF models [30], [31] with high accuracy. Recently a novel biquadratic reflectance model has been introduced for normal estimation in photometric stereo (ST14, [17]). Compared with the classical ones, the model is concise and accurate in representing the low-frequency components. A limitation is that the high-frequency components should be discarded beforehand. In [19] a bivariate monotonic smooth function is used to approximate the observed intensity in formulating a constrained bivariate regression problem. A most recent work (HS15, [26]) models the specular highlight, diffuse and shadow components based on their respective reflection characteristics, and formulates normal estimation as a second order cone programming problem. This method works well on surfaces with general reflectances at the cost of high computational complexity.

Uncalibrated photometric stereo, which does not require known illumination condition, has also been developed. By assuming Lambertian reflection and integrable surface, light direction ambiguity is reduced to the General Bas-Relief (GBR) ambiguity [32]. The GBR ambiguity can be further resolved by normal and albedo grouping [33] or entropy distribution minimization [34]. In manifold embedding based methods [28], [35], surface normals are recovered up to a global rotation ambiguity. To deal with general isotropic BRDFs, the work [36] maps the distances of intensity profiles to angular differences of normals without strict reflection assumption. By exploiting the BRDF symmetry of general isotropic reflectances, the elevation angles of surface normals have been reliably recovered [37]. In [38] the complete 3D model is obtained under multiview images. Compared with calibrated photometric stereo methods, the uncalibrated methods are more desired in practical data acquisition but less accurate in normal estimation [39].

B. Contributions

As mentioned, this paper proposes an efficient photometric stereo method based on kernel regression with known illumination condition. The difference between the proposed method and related work is as follows. In [15] and [16], the reflectance of each surface point is represented by a few basis BRDFs. By imposing necessary constraints, the basis BRDFs and surface normal are iteratively solved via alternating constrained least squares. In this work, the reflectance is represented by a large number of fixed (known) nonlinear basis functions, which are determined by light directions. By formulating the photometric stereo problem as kernel regression, surface normal can be efficiently solved in closed form, without alternating optimization. Different from [19] that uses the Bernstein polynomials, the proposed method employs the general kernel function that maps light direction to the Hilbert feature space. The space can have high or even infinite dimensionality, and hence is more appropriate in representing the variety of reflectances.

In summary, the main contributions of this work are twofold:

- A novel photometric stereo method is proposed for general isotropic reflectances using kernel regression, which is further transformed to an eigen decomposition problem.
- The best parameter of each variable kernel is determined by accelerated cross-validation with fast matrix computation and proper normal initialization.

The accuracy and efficiency of the proposed photometric stereo method have been validated by extensive experiments on both synthetic and real-world data.

II. Photometric Stereo Using Kernel Regression

A. Formulation

Suppose that a surface point with normal $\mathbf{n}$ is illuminated by a light source from the direction $\mathbf{l}$, then the observed intensity $o(\mathbf{n}, \mathbf{l})$ of the surface point is modeled as

$$o(\mathbf{n}, \mathbf{l}) = \rho(\mathbf{n}, \mathbf{l})\mathbf{l}^{\mathsf{T}}\mathbf{n}, \tag{1}$$

where $\rho(\mathbf{n}, \mathbf{l})$ denotes the reflectance relating to surface normal and light direction. Note that shadow pixels are not considered here.[1] As in [26], we introduce a new reflectance $s(\mathbf{n}, \mathbf{l}) := 1/\rho(\mathbf{n}, \mathbf{l})$, then (1) becomes

$$s(\mathbf{n}, \mathbf{l}) = \frac{\mathbf{l}^{\mathsf{T}}\mathbf{n}}{o}. \tag{2}$$

Note that the variables $\mathbf{n}$ and $\mathbf{l}$ of intensity o are omitted to simplify notation. As the normal of a given surface point is fixed, reflectance can be written as $s(\mathbf{l})$ because light direction is the only variable. Hence (2) can be rewritten as

$$s(\mathbf{l}) = \frac{\mathbf{l}^{\mathsf{T}}\mathbf{n}}{o}. \tag{3}$$

By disregarding normal $\mathbf{n}$, the reflectance $s(\mathbf{l})$ does not contain any unknown variables. However, this treatment may result in a highly nonlinear reflectance function due to the complex nature of general BRDFs. To deal with the nonlinearity of $s(\mathbf{l})$, we resort to the "kernel trick" by introducing a general nonlinear basis function $\boldsymbol{\phi}(\cdot) : \mathbb{R}^3 \rightarrow \mathcal{H}$ that maps $\mathbf{l}$ to a reproducing kernel Hilbert space $\mathcal{H}$. Then the reflectance is represented as $s(\mathbf{l}) = \boldsymbol{\phi}(\mathbf{l})^{\mathsf{T}}\mathbf{w}$, with $\mathbf{w}$ being the weight vector to be estimated. Consequently (3) becomes

$$\boldsymbol{\phi}(\mathbf{l})^{\mathsf{T}}\mathbf{w} = \frac{\mathbf{l}^{\mathsf{T}}\mathbf{n}}{o}. \tag{4}$$

For a photometric stereo system with M light directions, (4) can be represented in matrix form

$$\boldsymbol{\Phi}\mathbf{w} = \widetilde{\mathbf{L}}^{\mathsf{T}}\mathbf{n}, \tag{5}$$

[1] This work supposes that shadow pixels have been detected by a threshold beforehand.

where $\mathbf{\Phi} = (\boldsymbol{\phi}(\mathbf{l}_1), \boldsymbol{\phi}(\mathbf{l}_2), \cdots, \boldsymbol{\phi}(\mathbf{l}_M))^\mathsf{T}$ is the basis function matrix, and $\widetilde{\mathbf{L}} = (\frac{\mathbf{l}_1}{o_1}, \frac{\mathbf{l}_2}{o_2}, \cdots, \frac{\mathbf{l}_M}{o_M})$ is referred as light-intensity matrix.

The normal $\mathbf{n}$ and weight $\mathbf{w}$ can be robustly estimated using ridge regression,

$$\{\mathbf{n}, \mathbf{w}\} = \arg\min_{\mathbf{n},\mathbf{w}} \|\widetilde{\mathbf{L}}^\mathsf{T}\mathbf{n} - \mathbf{\Phi}\mathbf{w}\|_2^2 + \mu\mathbf{w}^\mathsf{T}\mathbf{w}, \quad \text{s.t. } \mathbf{n}^\mathsf{T}\mathbf{n} = 1, \tag{6}$$

where μ is the regularization parameter. Note that the constraint of normal, $\mathbf{n}^\mathsf{T}\mathbf{n} = 1$, is also incorporated in (6). Using kernel trick, $\mathbf{w}$ is represented by the linear combination of $\mathbf{\Phi}$, i.e., $\mathbf{w} = \mathbf{\Phi}^\mathsf{T}\mathbf{a}$, and then (6) can be reformulated as

$$\min_{\mathbf{n},\mathbf{a}} \|\widetilde{\mathbf{L}}^\mathsf{T}\mathbf{n} - \mathbf{K}\mathbf{a}\|_2^2 + \mu\mathbf{a}^\mathsf{T}\mathbf{K}\mathbf{a} \quad \text{s.t.} \quad \mathbf{n}^\mathsf{T}\mathbf{n} = 1, \tag{7}$$

which is a constrained kernel regression problem. Here $\mathbf{K} = \mathbf{\Phi}\mathbf{\Phi}^\mathsf{T}$ is the Gram matrix with entries

$$K_{ij} = \langle \boldsymbol{\phi}(\mathbf{l}_i), \boldsymbol{\phi}(\mathbf{l}_j) \rangle = k(\mathbf{l}_i, \mathbf{l}_j), \tag{8}$$

where $k(\mathbf{l}_i, \mathbf{l}_j)$ is the kernel function. This work adopts the Gaussian kernel with the form

$$k(\mathbf{l}_i, \mathbf{l}_j) = \exp(-\beta\|\mathbf{l}_i - \mathbf{l}_j\|_2^2), \tag{9}$$

where β is the kernel parameter. Note that the value of β determines the complexity of the kernel function. A smaller β generally results in a simpler reflectance model and *vice versa*. The detailed interpretation is given in Appendix A.

Note that although the expression of basis function ϕ is not explicitly speficied, the basis function matrix $\mathbf{\Phi}$ is known as both the light directions ($\mathbf{l}_i$ and $\mathbf{l}_j$) and kernel function $k(\cdot, \cdot)$ are known. The photometric stereo problem (7) is to solve the weight coefficients $\mathbf{a}$ and normal $\mathbf{n}$ with the known $\mathbf{K}$ and $\widetilde{\mathbf{L}}$.

B. Solution

The optimization problem (7) can be transformed to eigen decomposition. The derivation is based on the following two Lemmas.

Lemma 1: For a definite matrix $\mathbf{A}$ *and any vector* $\mathbf{b}$*, the quadratic function*

$$f(\mathbf{x}) = \frac{\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}}{2} - \mathbf{x}^\mathsf{T}\mathbf{b}, \tag{10}$$

has a global minimum $f(\mathbf{x}_0) = -\frac{1}{2}\mathbf{b}^\mathsf{T}\mathbf{A}^{-1}\mathbf{b}$ *at* $\mathbf{x}_0 = \mathbf{A}^{-1}\mathbf{b}$.

Proof: See Appendix B. ∎

Lemma 2: For a symmetric matrix $\mathbf{K}$*, if* $(\mathbf{K} + \mu\mathbf{I})$ *is invertible, we have* $\mathbf{I} - (\mathbf{K} + \mu\mathbf{I})^{-1}\mathbf{K} = \mu(\mathbf{K} + \mu\mathbf{I})^{-1}$.

Proof: See Appendix C. ∎

By expanding the L_2 norm squared and collecting the terms relating to $\mathbf{a}$ together, problem (7) can be written as

$$\min_{\mathbf{n},\mathbf{a}} \ \mathbf{n}^\mathsf{T}\widetilde{\mathbf{L}}\widetilde{\mathbf{L}}^\mathsf{T}\mathbf{n} + \mathbf{a}^\mathsf{T}(\mathbf{K}^2 + \mu\mathbf{K})\mathbf{a} - 2\mathbf{n}^\mathsf{T}\widetilde{\mathbf{L}}\mathbf{K}\mathbf{a} \quad \text{s.t. } \mathbf{n}^\mathsf{T}\mathbf{n} = 1. \tag{11}$$

Note that coefficient $\mathbf{a}$ relates only to the second and third terms, we define a subproblem

$$\Omega(\mathbf{n}) = \min_{\mathbf{a}} \mathbf{a}^\mathsf{T}(\mathbf{K}^2 + \mu\mathbf{K})\mathbf{a} - 2\mathbf{n}^\mathsf{T}\widetilde{\mathbf{L}}\mathbf{K}\mathbf{a}. \tag{12}$$

Algorithm 1 Estimating Surface Normal

Input : Light directions $\{\mathbf{l}_k\}_{k=1}^M$, observed intensity $\mathbf{o}$.
Output: Surface normal $\mathbf{n}$.
1) Compute light-intensity matrix $\widetilde{\mathbf{L}}$ and Gram matrix $\mathbf{K}$;
2) Compute $\mathbf{P} = \widetilde{\mathbf{L}}(\mathbf{K} + \mu\mathbf{I})^{-1}\widetilde{\mathbf{L}}^\mathsf{T}$;
3) Compute the eigenvectors corresponding to the minimum eigenvalue of $\mathbf{P}$ and obtain the normal $\mathbf{n}$.

Then the original problem (11) can be written as

$$\min_{\mathbf{n}} \ \mathbf{n}^\mathsf{T}\widetilde{\mathbf{L}}\widetilde{\mathbf{L}}^\mathsf{T}\mathbf{n} + \Omega(\mathbf{n}) \quad \text{s.t. } \mathbf{n}^\mathsf{T}\mathbf{n} = 1. \tag{13}$$

According to **Lemma 1**,

$$\Omega(\mathbf{n}) = -(\mathbf{n}^\mathsf{T}\widetilde{\mathbf{L}})(\mathbf{K} + \mu\mathbf{I})^{-1}\mathbf{K}(\widetilde{\mathbf{L}}^\mathsf{T}\mathbf{n}). \tag{14}$$

Substituting (14) into (13) yields

$$\min_{\mathbf{n}} \ \mathbf{n}^\mathsf{T}\widetilde{\mathbf{L}}\widetilde{\mathbf{L}}^\mathsf{T}\mathbf{n} - (\mathbf{n}^\mathsf{T}\widetilde{\mathbf{L}})(\mathbf{K} + \mu\mathbf{I})^{-1}\mathbf{K}(\widetilde{\mathbf{L}}^\mathsf{T}\mathbf{n}), \quad \text{s.t. } \mathbf{n}^\mathsf{T}\mathbf{n} = 1. \tag{15}$$

It can be further transformed to

$$\min_{\mathbf{n}} (\mathbf{n}^\mathsf{T}\widetilde{\mathbf{L}})(\mathbf{I} - (\mathbf{K} + \mu\mathbf{I})^{-1}\mathbf{K})(\widetilde{\mathbf{L}}^\mathsf{T}\mathbf{n}) \quad \text{s.t.} \quad \mathbf{n}^\mathsf{T}\mathbf{n} = 1. \tag{16}$$

According to **Lemma 2**, (16) can be simplified to

$$\min_{\mathbf{n}} \ \mu\mathbf{n}^\mathsf{T}\mathbf{P}\mathbf{n}, \quad \text{s.t. } \mathbf{n}^\mathsf{T}\mathbf{n} = 1, \tag{17}$$

where

$$\mathbf{P} = \widetilde{\mathbf{L}}(\mathbf{K} + \mu\mathbf{I})^{-1}\widetilde{\mathbf{L}}^\mathsf{T} = \widetilde{\mathbf{L}}\mathbf{X}\widetilde{\mathbf{L}}^\mathsf{T}, \tag{18}$$

with $\mathbf{X} = (\mathbf{K} + \mu\mathbf{I})^{-1}$. In (17), the minimum is attained at the eigenvector corresponding to the minimal eigenvalue of $\mathbf{P}$, and the surface normal is the corresponding eigenvector $\mathbf{n}$ [40]. Hence (17) is actually an eigen decomposition problem.

To summarize, Algorithm 1 outlines the computation of surface normal. The computational cost is analyzed as follows:

1) Computation of the Gram matrix $\mathbf{K}$. As the shaded light direction set is the subset of the whole light direction set, we can compute a complete Gram matrix $\widehat{\mathbf{K}}$ of the whole light direction set beforehand. The actual Gram matrix $\mathbf{K}$ can be directly obtained by extracting the subset of $\widehat{\mathbf{K}}$. Hence the computational cost of this step is negligible.
2) Computation of the $M \times M$ matrix $\mathbf{X} = (\mathbf{K} + \mu\mathbf{I})^{-1}$. The computational complexity of this matrix is generally $O(M^3)$. Though its computation speed can be improved by some efficient solvers, the complexity is generally higher than $O(M^2)$.
3) Computation of the eigenvectors corresponding to the minimum eigenvalue of $\mathbf{P}$. The computational complexity is almost irrelevant to M.

It is clear from the above analysis that the computational cost of Algorithm 1 is mainly determined by the computation of $(\mathbf{K} + \mu\mathbf{I})^{-1}$ and eigenvector $\mathbf{n}$.

Algorithm 2 Determining Kernel Parameter

Input : Light directions $\mathbf{L}$, intensity $\mathbf{o}$, parameter set $\mathcal{C}(\beta_l) = \{\beta_l\}_{l=1}^{L}$.
Output: Kernel parameter $\hat{\beta}$.
/* Cross validation */
Compute light-intensity matrix $\widetilde{\mathbf{L}}$;
for $l = 1$ **to** L **do**
  /* Compute baseline normal */
  Compute Gram matrix $\mathbf{K}$ corresponding to β_l;
  Solve $\mathbf{n}_{\beta_l}$ from $\widetilde{\mathbf{L}}$ and $\mathbf{K}$ according to (19);
  /* LOO computation process */
  for $i = 1$ **to** M **do**
    Compute $\widetilde{\mathbf{L}}^{(i)}$ and $\mathbf{K}^{(i)}$ by dropping the ith light source;
    Solve $\mathbf{n}_{\beta_l}^{(i)}$ from $\widetilde{\mathbf{L}}^{(i)}$ and $\mathbf{K}^{(i)}$ according to (19);
  end
  /* Compute total angular error */
  Compute $E(\beta_l)$ according to (20);
end
Determine the best $\hat{\beta}$ according to (21).

C. Determining Kernel Parameter

As noted above, surface normal is computed based on the Gaussian kernel with parameter β. Or equivalently, given the kernel parameter β, the corresponding normal $\mathbf{n}_\beta$ is solved as

$$\mathbf{n}_\beta = \underset{\mathbf{n}}{\operatorname{argmin}} \|\widetilde{\mathbf{L}}^{\mathsf{T}}\mathbf{n} - \mathbf{K}\mathbf{a}\|_2^2 + \mu\mathbf{a}^{\mathsf{T}}\mathbf{K}\mathbf{a}, \quad \text{s.t. } \mathbf{n}^{\mathsf{T}}\mathbf{n} = 1. \tag{19}$$

The computation steps are outlined in Algorithm 1.

It has been observed in experiments that, due to the variety of general reflectances, the kernel parameter β should be optimally determined for the material at hand. Accordingly, the kernel is referred as variable kernel. In this work, the LOO cross-validation scheme is employed to determine the best parameter. Suppose that the parameter set consists of L parameters, i.e., $\mathcal{C}(\beta_l) = \{\beta_l\}_{l=1}^{L}$. For each candidate β_l, we consequently drop the ith light source, where $i \in \{1, 2, \cdots, M\}$. Denoting $\widetilde{\mathbf{L}}^{(i)}$ and $\mathbf{K}^{(i)}$ as, respectively, the light-intensity and Gram matrices after dropping the ith light source, the corresponding trial normal $\mathbf{n}_{\beta_l}^{(i)}$ is solved according to (19). It is reasonable to assume that, for a proper kernel parameter β_l, the computed trial $\mathbf{n}_{\beta_l}^{(i)}$ should be close to the baseline normal $\mathbf{n}_\beta$. The average LOO angular error for parameter β_l is defined as

$$E(\beta_l) = \frac{1}{M}\sum_{i=1}^{M} \arccos(\mathbf{n}_{\beta_l}^{\mathsf{T}}\mathbf{n}_{\beta_l}^{(i)}), \tag{20}$$

where M is the number of light sources. Then, the best β value is determined by

$$\hat{\beta} = \underset{\beta_l}{\operatorname{argmin}} \quad E(\beta_l), \quad 1 \le l \le L. \tag{21}$$

Algorithm 2 outlines the cross-validation scheme determining the best kernel parameter. The algorithm includes two loops. The inner loop estimates normal $\mathbf{n}_{\beta_l}^{(i)}$ after dropping one light source in the LOO process, and the outer loop computes the LOO angular error of each candidate parameter.

D. Practical Considerations

The above algorithms performs quite well in ideal scenarios. The acquired images of real surfaces, however, may be corrupted by noise, attach shadow, cast shadow, interreflection, *etc*. In the photometric stereo field, shadows are always simply detected by thresholding. Due to the existence of noise and interreflection, shadow pixels may be misclassified, which can severely bias normal estimation. To avoid this problem, we also compute normal using the image intensities in the tightest threshold range ($T_{\text{low}} = 40\%, T_{\text{high}} = 60\%$) as suggested in [39]. Our investigation shows that when the normal is largely biased, the LOO error is usually high. Hence we adopt the normal corresponding to the relatively lower LOO angular error as the desired solution.

III. Algorithm Acceleration

Figure 1(a) illustrates the flowchart of cross-validation in Algorithm 2. The LOO computation should be implemented L times when L candidate parameters are used. In basic LOO computation [see Fig. 1(b)], Algorithm 1 should be called M times for each candidate parameter to solve trial normals in a photometric stereo system with M light sources. Hence, basically, Algorithm 1 should be called LM times in determining the best kernel parameter. Clearly the computational burden will be quite heavy when the system employs a large number of light sources. Therefore it is necessary to accelerate the LOO computation which is dependent on Algorithm 1.

A. Acceleration Scheme

As discussed in Section II-C, the main computation of Algorithm 1 is involved in solving $\mathbf{X} = (\mathbf{K} + \mu\mathbf{I})^{-1}$ and solving the eigenvectors corresponding to the minimum eigenvalue of $\mathbf{P} = \widetilde{\mathbf{L}}(\mathbf{K}+\mu\mathbf{I})^{-1}\widetilde{\mathbf{L}}^{\mathsf{T}} = \widetilde{\mathbf{L}}\mathbf{X}\widetilde{\mathbf{L}}^{\mathsf{T}}$. As illustrated in Fig. 1(c), this work employs two strategies to accelerate LOO computation:

1) Fast computation of $\mathbf{X}^{(i)} = (\mathbf{K}^{(i)} + \mu\mathbf{I})^{-1}$ based on the result of $\mathbf{X} = (\mathbf{K} + \mu\mathbf{I})^{-1}$.
2) Setting $\mathbf{n}_{\beta_l}$ as the initial point of $\mathbf{n}_{\beta_l}^{(i)}$ in the conjugate gradient algorithm [41] when solving eigenvector.

The second strategy is straightforward. Generally the trial normal $\mathbf{n}_\beta^{(i)}$, which is obtained after dropping one light source, is close to the baseline normal $\mathbf{n}_\beta$. By setting $\mathbf{n}_\beta$ as the initial point, the conjugate gradient algorithm will converge rapidly. The fast computation of matrix $\mathbf{X}^{(i)} = (\mathbf{K}^{(i)} + \mu\mathbf{I})^{-1}$ is elaborated in the following.

B. Implementation

Lemma 3: For a symmetric matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ *in the form* $\begin{pmatrix}\mathbf{A}_{n-1} & \mathbf{a}_n \\ \mathbf{a}_n^{\mathsf{T}} & a_{nn}\end{pmatrix}$, *the inverse matrix* $\mathbf{B}$ *has the form* $\begin{pmatrix}\mathbf{B}_{n-1} & \mathbf{b}_n \\ \mathbf{b}_n^{\mathsf{T}} & b_{nn}\end{pmatrix}$. *By removing the nth column and nth row of matrix* $\mathbf{A}$, *the inverse matrix of the submatrix* $\mathbf{A}_{n-1}$ *satisfies*

$$\mathbf{B}^{(n)} = \mathbf{B}_{n-1} - \mathbf{b}_n\mathbf{b}_n^{\mathsf{T}}/b_{nn}. \tag{22}$$

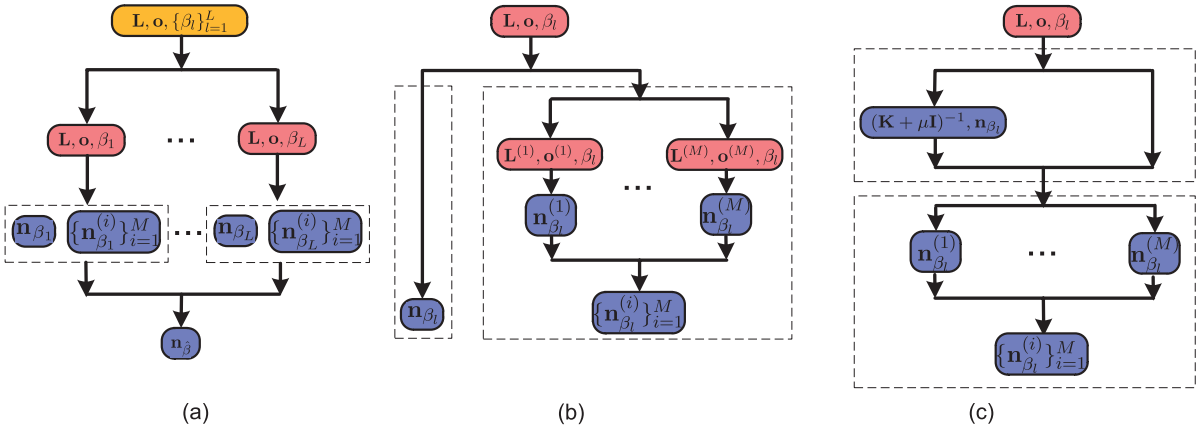

Fig. 1. Computation involved in cross-validation. (a) In the cross-validation of Algorithm 2, the baseline normal $\mathbf{n}_{\beta_l}$ and the trial normal set $\{\mathbf{n}_{\beta_l}^{(i)}\}_{i=1}^{M}$ are computed for each candidate kernel parameter β_l, $l \in \{1, 2, \cdots, L\}$. The best parameter $\hat{\beta}$ is determined in terms of average LOO angular error. (b) In basic LOO computation, all trial normals are individually computed using Algorithm 1. The computational burden is heavy. (c) The LOO computation is accelerated by fast matrix computation and proper normal initialization.

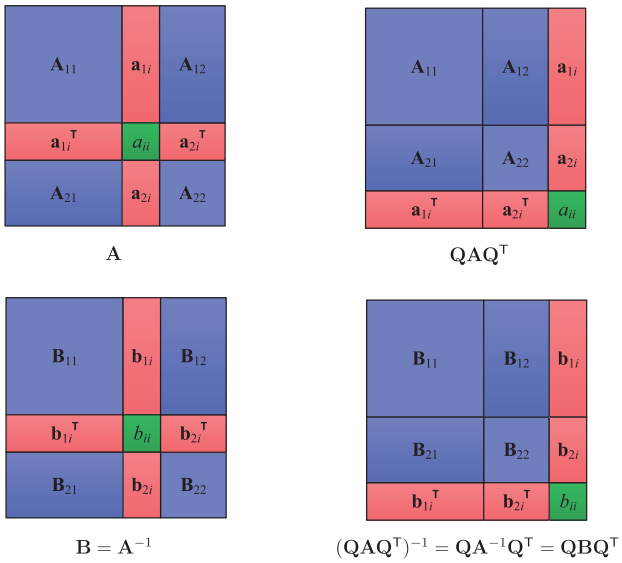

Fig. 2. Extension of **Lemma 3** to the general case of removing the ith row and ith column by introducing the permutation matrix $\mathbf{Q}$.

Proof: See Appendix D. ■

Note that **Lemma 3** can be extended to the ith column and ith row for any $i \in [1, n]$ by introducing a permutation matrix $\mathbf{Q}$ [40]. As illustrated in Fig. 2, with the transform $\mathbf{QAQ}^\mathsf{T}$ the ith row and ith column of $\mathbf{A}$ can be moved to nth row and nth column; likewise, with the transform $\mathbf{QBQ}^\mathsf{T}$ the ith row and ith column of $\mathbf{B}$ can be moved to the nth row and nth column. As $\mathbf{QBQ}^\mathsf{T} = (\mathbf{QAQ}^\mathsf{T})^{-1}$, the inverse of the moved matrix $\mathbf{A}$ corresponds to the moved matrix $\mathbf{B}$. Hence if $\mathbf{X} = (\mathbf{K} + \mu\mathbf{I})^{-1}$ is known, the inverse of $(\mathbf{K} + \mu\mathbf{I})^{(i)}$, i.e., $\mathbf{X}^{(i)}$, can be obtained by

$$\mathbf{X}^{(i)} = \mathbf{X}_{-i} - \mathbf{x}_i \mathbf{x}_i^\mathsf{T} / x_{ii}, \tag{23}$$

where $\mathbf{X}_{-i}$ is the submatrix of $\mathbf{X}$ after removing the ith row and ith column, x_{ii} is the (i, i)th entry of matrix $\mathbf{X}$, and $\mathbf{x}_i$ is the ith column of matrix $\mathbf{X}$ after removing x_{ii}. With (23), $\mathbf{X}^{(i)}$ can be obtained by simple matrix operation.

The fast computation of $\mathbf{P}^{(i)}$ is introduced in the following. To simplify presentation, we discuss the special case of dropping the Mth light source in the LOO computation process. Its extension to the general case of dropping the ith light source is straightforward via matrix permutation. Accordingly, (18) can be written as

$$\mathbf{P} = \left(\widetilde{\mathbf{L}}_{M-1}\ \widetilde{\mathbf{l}}_M\right) \begin{pmatrix} \mathbf{X}_{M-1} & \mathbf{x}_M \\ \mathbf{x}_M^\mathsf{T} & x_{MM} \end{pmatrix} \begin{pmatrix} \widetilde{\mathbf{L}}_{M-1}^\mathsf{T} \\ \widetilde{\mathbf{l}}_M^\mathsf{T} \end{pmatrix}, \tag{24}$$

where $\widetilde{\mathbf{L}}_{M-1}$ is the submatrix of $\widetilde{\mathbf{L}}$ by removing its last column vector $\widetilde{\mathbf{l}}_M$. After dropping the Mth light source in the LOO process, $\mathbf{P}^{(M)}$ becomes

$$\begin{aligned} \mathbf{P}^{(M)} &= \widetilde{\mathbf{L}}^{(M)} \mathbf{X}^{(M)} (\widetilde{\mathbf{L}}^{(M)})^\mathsf{T} \\ &= \widetilde{\mathbf{L}}_{M-1} (\mathbf{X}_{M-1} - \mathbf{x}_M \mathbf{x}_M^\mathsf{T} / x_{MM}) \widetilde{\mathbf{L}}_{M-1}^\mathsf{T}. \end{aligned} \tag{25}$$

Note that the second equality is due to (23) and the fact $\widetilde{\mathbf{L}}^{(M)} = \widetilde{\mathbf{L}}_{M-1}$. Based on (24) and (25), we have

$$\begin{aligned} \mathbf{P}^{(M)} = \mathbf{P} &- (\widetilde{\mathbf{L}}_{M-1}\mathbf{x}_M + x_{MM}\widetilde{\mathbf{l}}_M) \\ &\times (\widetilde{\mathbf{L}}_{M-1}\mathbf{x}_M + x_{MM}\widetilde{\mathbf{l}}_M)^\mathsf{T} / x_{MM} \end{aligned} \tag{26}$$

after simple matrix manipulation. It is noticed that the term $\widetilde{\mathbf{L}}_{M-1}\mathbf{x}_M + x_{MM}\widetilde{\mathbf{l}}_M$ has been obtained when computing $\mathbf{P}$. More explicitly, the second matrix multiplication in (24) is

$$\begin{aligned} &\begin{pmatrix} \mathbf{X}_{M-1} & \mathbf{x}_M \\ \mathbf{x}_M^\mathsf{T} & x_{MM} \end{pmatrix} \begin{pmatrix} \widetilde{\mathbf{L}}_{M-1}^\mathsf{T} \\ \widetilde{\mathbf{l}}_M^\mathsf{T} \end{pmatrix} \\ &= \begin{pmatrix} (\widetilde{\mathbf{L}}_{M-1}\mathbf{X}_{M-1} + \widetilde{\mathbf{l}}_M\mathbf{x}_M^\mathsf{T})^\mathsf{T} \\ (\widetilde{\mathbf{L}}_{M-1}\mathbf{x}_M + x_{MM}\widetilde{\mathbf{l}}_M)^\mathsf{T} \end{pmatrix} \overset{\text{def}}{=} \begin{pmatrix} * \\ \mathbf{y}^\mathsf{T} \end{pmatrix}. \end{aligned} \tag{27}$$

Hence (26) can be simplified using the already computed vector $\mathbf{y}$,

$$\mathbf{P}^{(M)} = \mathbf{P} - \mathbf{y}\mathbf{y}^\mathsf{T} / x_{MM}. \tag{28}$$

Note that (28) only involves 3×3 matrix operation, which is quite efficient and irrelevant to the number of light sources M.

TABLE I

COMPUTATION TIME (UNIT: millisecond) PER PIXEL OF THE ALGORITHM WITH AND WITHOUT ACCELERATION. THE NUMBER OF LIGHT SOURCES IS $M = 100$

	Normal comput.	LOO comput.	Total
Without acc.	2.71	267.60	270.31
With acc.	3.24	0.80	4.04

C. Improvement of Computational Efficiency

Table I lists the computation time of the algorithm with and without acceleration when the number of light sources $M = 100$. The algorithm runs on a personal computer with 2.20 GHz GPU and 8 GB RAM. In the algorithm without acceleration, we directly compute $\widetilde{\mathbf{L}}(\mathbf{K} + \mu\mathbf{I})^{-1}$ instead of $(\mathbf{K} + \mu\mathbf{I})^{-1}$ as the former is more efficient. As expected, in the case of without acceleration, the running time of cross-validation is almost M times of normal computation. With acceleration, the running time of cross-validation is irrelevant to M. As a consequence, the algorithm runs 65× faster after employing the acceleration scheme.

IV. EXPERIMENTS

The performance of the proposed method is mainly compared with those of the baseline least squares (LS, [1]) and five state-of-the-art ones, including sparse Bayesian learning (IA12, [10]), [2] median photometric (MH10, [7]), biquadratic reflectance model (ST14, [17]), constrained bivariate regression (IA14, [19]), [3] and sparsity reflection modeling (HS15, [26]). [4] These methods are evaluated on both synthetic and real scenes. The synthetic data are rendered using MERL material database [12] on two surfaces, *Ball* and *Bunny*. The real data include those captured by ourselves [26] and the most recent DiLiGenT dataset [39]. In the DiLiGenT dataset, the recovered surface normals of many photometric stereo methods, as well as the ground truths, are available. In the experiments we conduct quantitative comparisons on the MERL database and DiLiGenT dataset, and qualitative comparisons on other real surfaces.

The parameters of the competing methods are set as follows. The weight of IW12 [10] is $\lambda = 10^{-6}$ for synthetic data and $\lambda = 10^{-2}$ for real data. The parameters in MH10 [7] are set as $\lambda_{\rho,\text{med}} = 0$ and $\lambda_{\rho,\text{avg}} = 0$ so that the method runs in a pointwise manner. The low-frequency threshold of ST14 [17] is $T_{\text{low}} = 25\%$. The algorithm is conducted with retro-reflection detection and the order of the polynomial is $(3, 5)$ for synthetic data and $(1, 5)$ for real data. In HS15 [26], the parameters are $\xi = 10^7$ and $\xi = 10^{1.5}/\texttt{median}(\mathbf{o})$ for synthetic and real data, respectively, and $\lambda_{\text{w}} = \lambda_{\text{s}} = 1$ for both. For all those methods, the shadow in synthetic scenes is detected by a threshold $T_{\text{shadow}} = 10^{-6}$. For the proposed method, the regularization parameter in (6) is set as $\mu = 0.01$.

A. Effects of Kernels

We first evaluate the effect of fixed kernels whose parameters β are manually set. The findings could assist the practical setting of variable kernels, including the parameter range and number of parameter candidates in cross-validation. Figure 3 illustrates the mean angular errors of the estimated surface normals for the *Ball* surface. The surfaces are rendered using all the 100 MERL materials [12] under various number of light sources. The light directions are uniformly distributed on the hemisphere covering the object surface. As expected, the angular errors decrease when adopting more light sources, and are considerably low when the number of light sources $M = 100$. There is no additional accuracy improvement when increasing M from 150 to 200. It is observed that the proposed method is less sensitive to the kernel parameter when employing a larger number of light sources, and performs quite stable in a wide range $[10^{-3}, 10^{0.6}]$ when using $M = 100$ or more light sources.

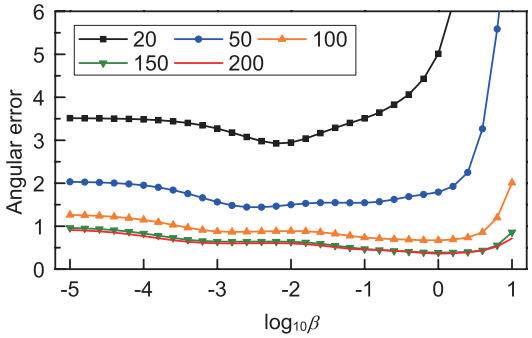

Fig. 3. Mean angular errors (in degrees) of the *Ball* surface with respect to different β values when using $M = 20, 50, 100, 150$, and 200 light sources. The mean angular errors are computed from all the 100 MERL materials.

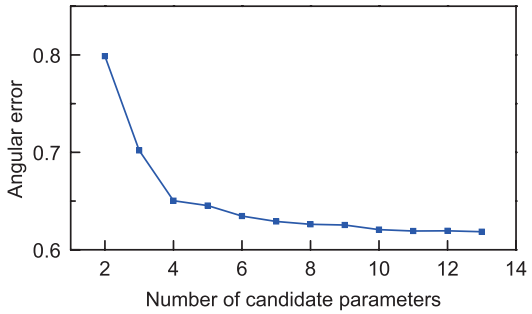

Fig. 4. Mean angular errors (in degrees) of the *Ball* surface on the 100 MERL materials with different number of candidate parameters in the range $[10^{-3}, 10^{0.6}]$. The number of light sources is $M = 100$.

The trends of angular errors in Fig. 3 also reveal the role of kernel parameter β in modeling reflectance. The optimal value of β is relevant to the number of light directions M. Take the case of $M = 50$ as an instance, the optimal value of β should be around $10^{-2.5}$. As a simpler model (with small β value) is unable to represent the variety of reflectances, the angular error is consequently large. On the other hand, overfitting occurs when the reflectance model becomes much complex. For example, the angular error becomes very large when $\beta = 10$. The above findings indicate that the automatic determination of parameter β is important to the proposed method.

[2] Source code available at https://www.hal.t.u-tokyo.ac.jp/~ikehata/
[3] Source code available at https://www.hal.t.u-tokyo.ac.jp/~ikehata/
[4] Source code available at http://www.ivlab.org/publications.html

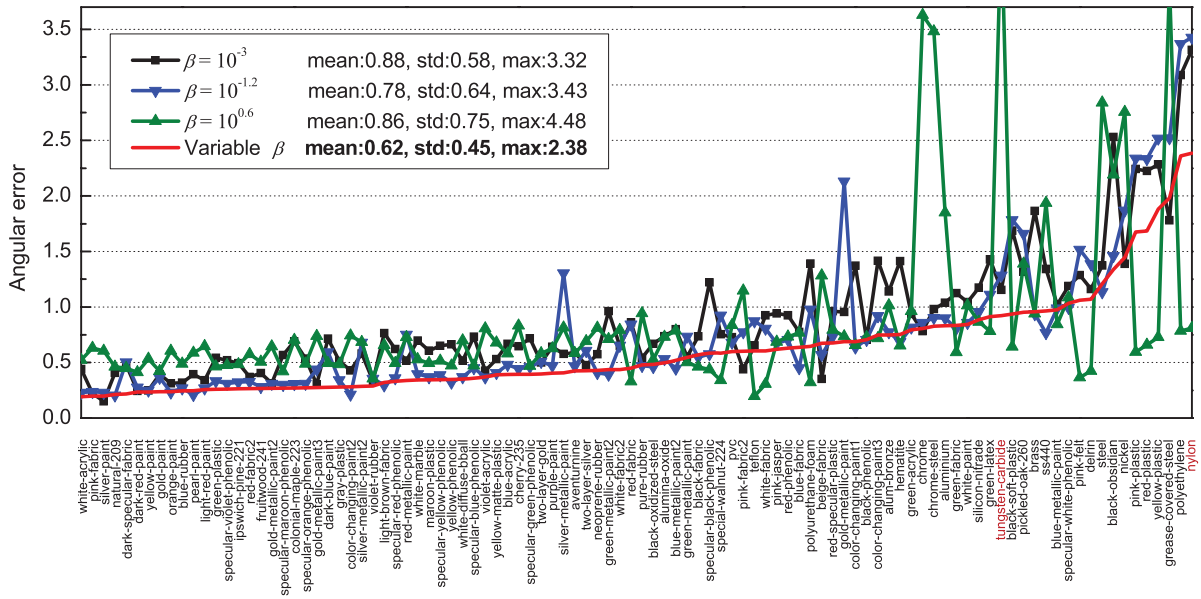

Fig. 5. Angular errors (in degrees) of the *Ball* surface on the 100 MERL materials when employing $M = 100$ light sources.

TABLE II

MEAN ANGULAR ERRORS (IN degrees) OF THE *Ball* SURFACE UNDER DIFFERENT NUMBERS OF LIGHT DIRECTIONS WHEN USING FIXED KERNEL AND VARIABLE KERNEL. EACH LISTED ERROR IS THE AVERAGE OF THE 100 ERRORS FROM THE 100 MERL MATERIALS

M	20	50	100	150	200
$\beta = 10^{-3}$	**3.27**	1.56	0.88	0.64	0.60
$\beta = 10^{-1.2}$	3.40	1.4	0.78	0.50	0.47
$\beta = 10^{0.6}$	13.94	3.27	0.86	0.44	0.43
Variable β	3.46	**1.36**	**0.62**	**0.34**	**0.31**

We set an appropriate number of candidate parameters for the cross-validation of variable kernel. Based on the observation that the method is not quite sensitive to parameter β, we equally sample the logarithm range of $[10^{-3}, 10^{0.6}]$ to get various numbers of candidate parameters L. Figure 4 shows that the mean angular error becomes lower when increasing L from 2 to 13. To keep a tradeoff between normal accuracy and computational efficiency, the number of candidate samples is set as $L = 10$ in the following experiments.

Table II lists the mean angular errors produced by fixed kernel (with $\beta = 10^{-3}, 10^{-1.2}$, and $10^{0.6}$) and variable kernel under various numbers of light sources. Due to the variety of general reflectances, it is almost impossible for a given fixed kernel to produce satisfactory accuracy in all cases. For example, the fixed kernel with $\beta = 10^{0.6}$ performs better than the other two fixed kernels when $M = 150$ and 200 but performs worst when $M = 20$. Note that in a real system the ground truth is unknown, thus it is extremely difficult to choose a good fixed kernel. In contrast, the best parameters of variable kernels are automatically determined from known data via cross-validation. Consequently, variable kernel performs better than fixed kernel in all cases except the one when $M = 20$. Note that even in this exception case, the mean angular error produced by variable kernel is just a little higher than the best one.

Figure 5 plots the angular errors of the *Ball* surface for all the 100 MERL materials in case of $M = 100$ light sources. The materials are sorted with respect to the angular errors produced by the proposed method with variable kernel. It is observed that variable kernel performs better than fixed kernel on most materials. Although the fixed kernel with $\beta = 10^{0.6}$ performs best on some materials (e.g., *nylon*), it also produces the largest errors on other ones (e.g., *tungsten-carbide*).

B. Results on Synthetic Data

The normal accuracy of the proposed method, as well as LS [1], IW12 [10], MH10 [7], ST14 [17], IA14 [19], and HS15 [26], are quantitatively evaluated on synthetic data. Table III lists the mean angular errors of the *Ball* and *Bunny* surfaces under different number of light sources. It is not surprise that LS [1] always performs worst because of its strict assumption of Lambertian reflection. The angular errors produced by IA14 [10] and MH10 [7] are also quite large although highlight pixels have been treated as outliers. The reason is that these two methods assume Lambertian diffuse reflection after outlier removal, but general diffuse reflection is rather complex. In most cases the proposed method performs better than the competitors. In occasions that the number of light sources M is 50 or less, its accuracy is a little worse than that of HS15 [26]. Nevertheless, it will be seen in Section IV-E that the proposed method runs much faster than HS15 [26].

Figure 6 illustrates the angular errors of the *Ball* surface produced by ST14 [17], IA14 [19], HS15 [26] and the proposed methods. It is observed that overall the proposed method performs better than the other methods on the 100 MERL materials. Figure 7 shows the spatial angular

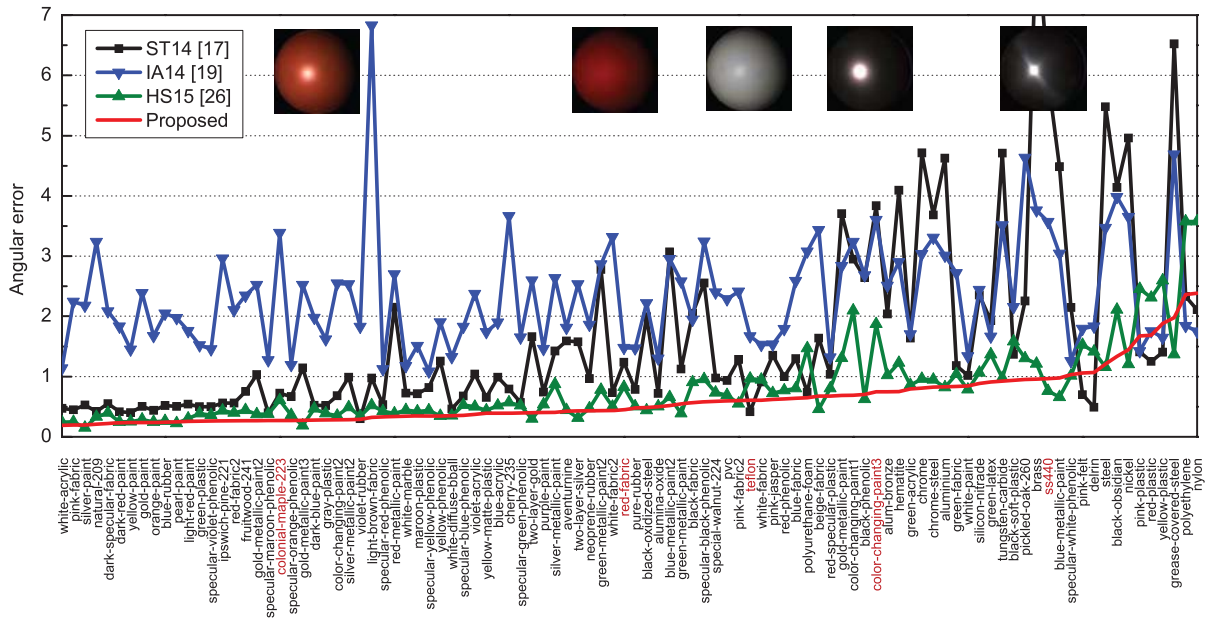

Fig. 6. Angular errors (in degrees) of the *Ball* surface on the 100 MERL materials when using $M = 100$ light sources. Five typical materials, with names in red and images on top, are chosen for further discussion in Fig. 7.

TABLE III

MEAN ANGULAR ERRORS (IN degrees) OF THE *Ball* AND *Bunny* SURFACES WHEN USING VARIOUS NUMBER (M) OF LIGHT SOURCES. EACH LISTED ERROR IS THE AVERAGE OF THE 100 ERRORS FROM THE 100 MERL MATERIALS [5]

Ball							
M	LS [1]	IW12 [10]	MH10 [7]	ST14 [17]	IA14 [19]	HS15 [26]	Proposed
20	10.56	6.57	7.10	-	5.01	**3.22**	3.46
50	12.30	6.68	6.06	1.90	3.55	1.56	**1.36**
100	11.99	6.43	6.20	1.65	2.34	0.81	**0.62**
150	11.89	6.36	6.36	1.17	1.89	0.55	**0.34**
200	11.27	6.15	6.75	1.07	1.90	0.49	**0.31**
Bunny							
M	LS [1]	IW12 [10]	MH10 [7]	ST14 [17]	IA14 [19]	HS15 [26]	Proposed
20	12.19	7.63	7.86	-	7.27	**4.22**	4.66
50	13.51	7.60	7.24	2.93	6.05	**2.81**	2.82
100	13.33	7.32	7.22	2.51	4.59	2.09	**1.99**
150	13.27	7.23	7.36	2.16	3.98	1.77	**1.70**
200	12.82	7.06	7.52	2.04	3.81	1.69	**1.68**

errors of the *Bunny* surface rendered by the five highlighted materials in Fig. 6. The proposed method performs better than all other methods on the materials *colonial-maple-223* and *red-fabric*, but performs worse than ST14 [17] on the material *teflon*. For the other two metal-like materials *color-changing-paint3* and *ss440*, the proposed method performs better than ST14 [17] and IA14 [19], but its angular errors are also not low. These observations indicate that, although the proposed method exhibits superiority over the competitors, its capability of reflectance modeling is still limited for certain materials with complex reflectances.

[5]The angular errors of the ST14 method with $M = 20$ light directions is unavailable because the valid number of data is inadequate for fitting.

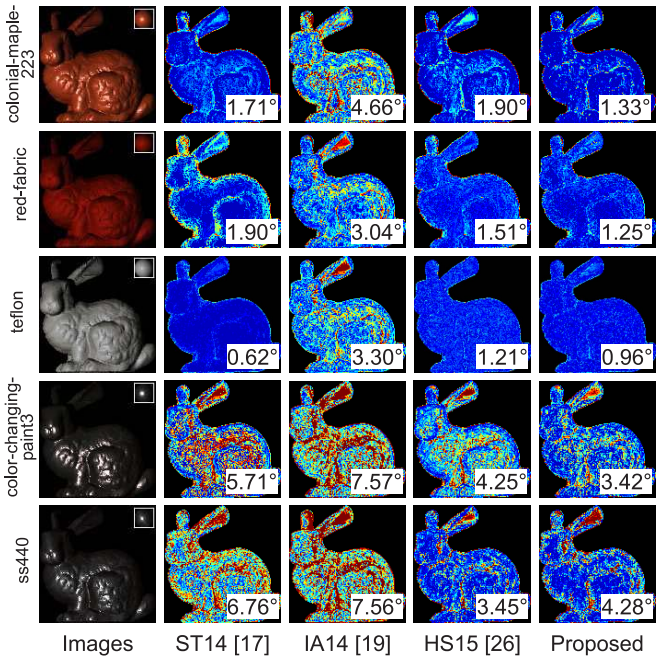

Fig. 7. Spatial Angular errors of the *Bunny* surface rendered using 5 MERL materials. The angular errors are displayed in the range $[0°, 10°]$. Higher errors are encoded by warmer colors. The images of materials *color-changing-paint3* and *ss440* are nonlinearly adjusted for visualization.

C. Results on DiLiGenT Dataset

The DiLiGenT dataset [39] contains the photometric stereo images of 10 real object surfaces with a variety of materials;

Fig. 8. Sample images and corresponding ground truth normals (shown at bottom right) of the ten objects in the DiLiGenT dataset.

see Fig. 8. In the dataset, the intensities and directions of the 96 light source are carefully calibrated. The shapes are obtained using a 3D scanner, based on which the ground truth surface normals have been obtained. In additional to LS [1], IW12 [10], ST14 [17], and IA14 [19], the normals recovered by WG10 [9], GC10 [14], AZ08 [16], ST12 [25], and HM10 [20] are also available in the dataset. Table IV lists the mean angular errors of the twelve photometric stereo methods on the 10 object surfaces. As observed, the proposed method is among the best in the twelve methods. It produces the lowest errors on 3 objects, namely CAT, BUDDHA, and READING. For the first six objects, the mean errors produced by the proposed method are all less than 10°. These objects are mostly of diffuse reflection (BALL, POT1), smoothly varying texture (CAT), or broad highlights (BUDDHA, POT2, BEAR). This validates that the proposed method can handle objects with a variety of reflectances. However, it does not performs well on the last four surfaces (GOBLET, READING, COW, and HARVEST). Figure 9 shows the corresponding spatial angular errors produced by ST12 [25], ST14 [17], and the proposed method. As observed, the proposed method does not produce satisfactory results on surfaces with concave shapes (READING, HARVEST) or of metal-like materials (GOBLET, COW, HARVEST). ST12 [25] performs comparatively well on surfaces COW and HARVEST, but results in large errors on GOBLET and READING. This example demonstrates the common limitation of these methods in dealing with complex reflectance and shape.

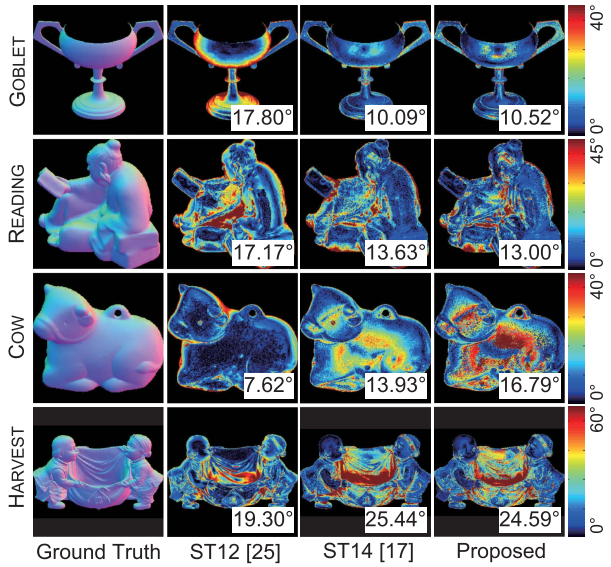

Fig. 9. Ground truth normals and spatial angular errors of the four real surfaces (GOBLET, READING, COW, and HARVEST) in the DiLiGenT dataset.

Fig. 10. Example images of three real object surfaces, namely, *Lady*, *Sheep*, and *Boy*.

D. Results on Other Real Surfaces

The other real surfaces include *Lady*, *Sheep*, and *Boy*; see Fig. 10. The results of the *Lady* surface are illustrated in Fig. 11. The normal maps estimated by MH10 [7], HS15 [26] and the proposed methods are visually satisfactory. ST14 [17] introduces erroneous discontinuity in the close-up normal map. The normal map produced by IW12 [10] is a little noisy. The normal estimated by IA14 [19] is obviously biased by specular highlight. In Fig. 12, the normal maps recovered by the proposed method and MH10 [7] are visually pleasing. In the results produced by IW12 [10] and ST14 [17], artifacts are observed in the close-ups. Figure 13 illustrates the results of the *Boy* surface. Due to its metal-like material and complex shape, this surface contains strong highlight and a large amount of shadow. As can be observed from the close-ups, the estimated normal maps are greatly biased in IW12 [10], ST14 [17], and IA14 [19]. The normal map produced by the proposed method is slightly better than that by HS15 [26].

E. Computational Time

In addition to estimation accuracy, computational efficiency is also an important issue in practical photometric stereo.

TABLE IV
MEAN ANGULAR ERRORS (IN degrees) ON THE TEN OBJECTS IN THE DiLiGenT DATASET

Objects	LS [1]	WG10 [9]	IW12 [10]	MH10 [7]	GC10 [14]	AZ08 [16]	ST12 [25]	HM10 [20]	ST14 [17]	IA14 [18]	HS15 [26]	Proposed
⓪ BALL	4.10	2.06	2.54	2.28	3.21	2.71	13.58	3.55	**1.74**	3.34	2.73	2.17
① CAT	8.41	6.73	7.21	7.08	8.22	6.53	12.34	8.40	6.12	6.74	6.66	**5.64**
② POT1	8.89	7.18	7.74	7.73	8.53	7.23	10.37	10.85	**6.51**	6.64	6.89	7.28
③ BEAR	8.39	6.50	7.32	6.61	6.62	5.96	19.44	11.48	6.12	7.11	**5.12**	5.31
④ POT2	14.65	13.12	14.09	13.26	**7.90**	11.03	9.84	16.37	8.78	8.77	9.80	8.43
⑤ BUDDHA	14.92	10.91	11.11	11.88	14.85	12.54	18.37	13.05	10.60	10.47	12.29	**9.30**
⑥ GOBLET	18.50	15.70	16.25	16.20	14.22	13.93	17.80	14.89	10.09	**9.71**	11.71	10.52
⑦ READING	19.80	15.39	16.17	14.03	19.07	14.17	17.17	16.82	13.63	14.19	14.56	**13.00**
⑧ COW	25.60	25.89	25.70	23.63	9.55	21.48	**7.62**	14.95	13.93	13.05	17.20	16.79
⑨ HARVEST	30.62	30.01	29.26	27.78	27.84	30.50	**19.30**	21.79	25.44	25.95	25.25	24.59
Average	15.39	13.35	13.74	13.05	12.00	12.61	14.58	13.22	**10.30**	10.60	11.22	**10.30**

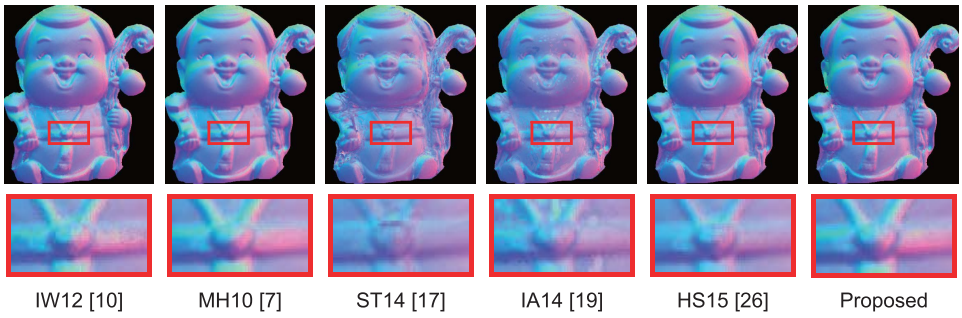

Fig. 11. Normal maps of the real surface *Lady* produced by different methods.

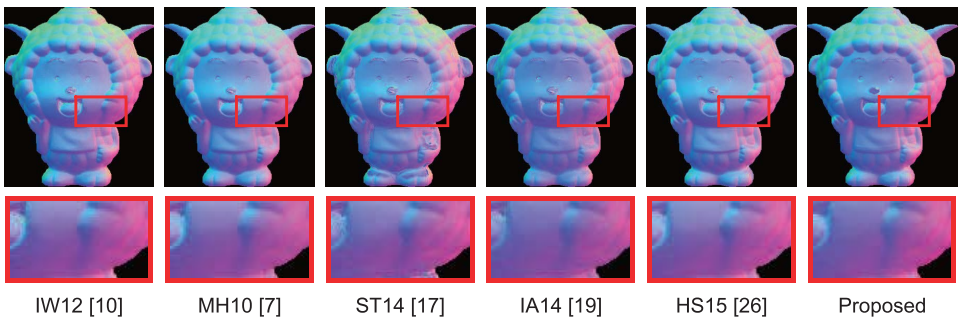

Fig. 12. Normal maps of the real surface *Sheep* produced by different methods.

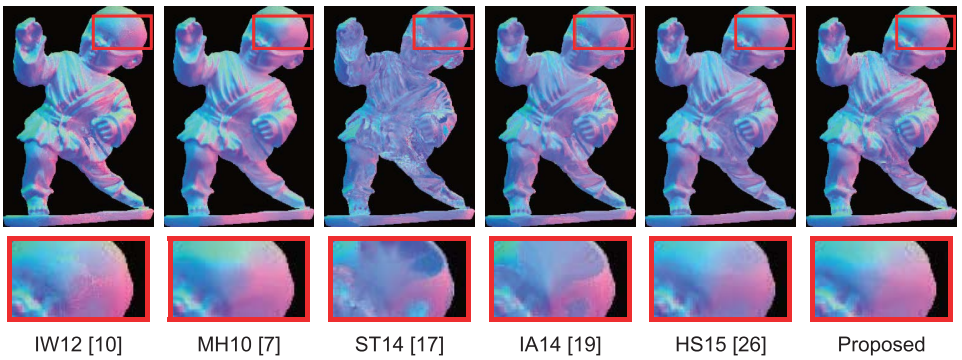

Fig. 13. Normal maps of the real surface *Boy* produced by different methods.

Figure 14 illustrates the average computational time per pixel computed from all the 100 MERL materials. As MH10 [7] is extremely slow in case of large number of light sources, its computational time is not shown. All methods are implemented using MATLAB on a personal computer with 2.20 GHz GPU and 8 GB RAM. In the proposed method, the

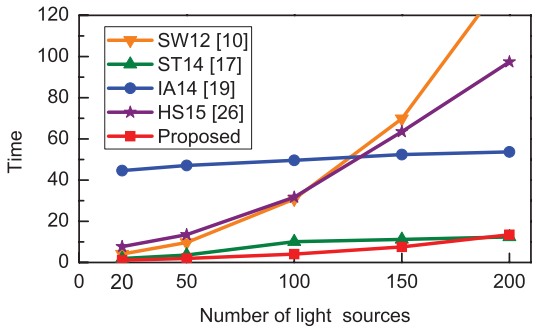

Fig. 14. Average computational times (unit: millisecond) per pixel of IW12 [10], ST14 [17], IA14 [19], HS15 [26], and the proposed methods when using various numbers of light sources.

number of candidate parameters is $L = 10$. It is observed that the proposed method runs slightly faster than ST14 [17] and much faster than the other three methods.

V. Conclusions

This paper has proposed a new point-wise photometric stereo method for general isotropic reflectances using kernel regression, under the assumption of known light directions condition. The best kernel parameter is determined by leave-one-out cross-validation such that the kernel can well characterize a variety of reflectances. To keep computational efficiency, the leave-one-out computation is speeded up using an acceleration scheme. Experimental results validate that, on both synthetic and real surfaces, the method generally performs better than, or comparable to, the state-of-the-art photometric stereo methods.

Appendix A
Interpretation of Kernel Function

Given two light directions ($\mathbf{l}_i$ and $\mathbf{l}_j$) and parameter β, the kernel function is

$$\begin{aligned} k(\mathbf{l}_i, \mathbf{l}_j) &= \exp(-\beta \|\mathbf{l}_i - \mathbf{l}_j\|_2^2) \\ &= \exp(-\beta \mathbf{l}_i^\mathsf{T}\mathbf{l}_i) \cdot \exp(2\beta \mathbf{l}_i^\mathsf{T}\mathbf{l}_j) \cdot \exp(-\beta \mathbf{l}_j^\mathsf{T}\mathbf{l}_j) \\ &= \exp(-2\beta) \cdot \exp(2\beta \mathbf{l}_i^\mathsf{T}\mathbf{l}_j). \end{aligned} \tag{29}$$

The second equality is due to the fact $\mathbf{l}^\mathsf{T}\mathbf{l} = 1$. Using power series expansion, the term $\exp(2\beta \mathbf{l}_i^\mathsf{T}\mathbf{l}_j)$ becomes

$$\exp(2\beta \mathbf{l}_i^\mathsf{T}\mathbf{l}_j) = \sum_{k=0}^{\infty} \frac{(2\beta \mathbf{l}_i^\mathsf{T}\mathbf{l}_j)^k}{k!} = \sum_{k=0}^{\infty} \frac{(2\beta)^k}{k!} (\mathbf{l}_i^\mathsf{T}\mathbf{l}_j)^k. \tag{30}$$

Hence the corresponding feature space contains all the monomials of order k [42]. Theoretically, the feature space is of infinite dimension when $k \to \infty$. However, given a constant β, the high-order monomials are negligible because the factor $\frac{(2\beta)^k}{k!}$ decreases significantly when k increases. Specially, the basis function $\phi(\mathbf{x}) \equiv 1$ when $\beta = 0$, indicating a constant reflectance. On the other hand, a larger β value can keep higher-order monomials, thus corresponds to a more complex reflectance model.

Appendix B
Proof of Lemma 1

Proof: Define a function $g(\mathbf{x}) = -\frac{1}{2}\mathbf{b}^\mathsf{T}\mathbf{A}^{-1}\mathbf{b} - f(\mathbf{x})$. For any $\mathbf{x}$, it is computed as

$$\begin{aligned} g(\mathbf{x}) &= -\frac{1}{2}\mathbf{b}^\mathsf{T}\mathbf{A}^{-1}\mathbf{b} - \frac{\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x}}{2} + \mathbf{x}^\mathsf{T}\mathbf{b} \\ &= -\frac{1}{2}(\mathbf{b}^\mathsf{T}\mathbf{A}^{-1} - \mathbf{x}^\mathsf{T})\mathbf{b} - \frac{1}{2}\mathbf{x}^\mathsf{T}\mathbf{A}(\mathbf{x} - \mathbf{A}^{-1}\mathbf{b}) \\ &= -\frac{1}{2}(\mathbf{b}^\mathsf{T}\mathbf{A}^{-1} - \mathbf{x}^\mathsf{T})\mathbf{A}(\mathbf{A}^{-1}\mathbf{b} - \mathbf{x}). \end{aligned} \tag{31}$$

It is clear that $g(\mathbf{x}) \le 0$ when $\mathbf{A}$ is definite. The minimum value $f(\mathbf{x}_0) = -\frac{1}{2}\mathbf{b}^\mathsf{T}\mathbf{A}^{-1}\mathbf{b}$ is reached at $\mathbf{x_0} = \mathbf{A}^{-1}\mathbf{b}$. ■

Appendix C
Proof of Lemma 2

Proof: Define a matrix $\widetilde{\mathbf{X}}$ in the form

$$\widetilde{\mathbf{X}} = \mathbf{I} - (\mathbf{K} + \mu\mathbf{I})^{-1}\mathbf{K}. \tag{32}$$

By simple arrangement,

$$(\mathbf{K} + \mu\mathbf{I})(\mathbf{I} - \widetilde{\mathbf{X}}) = \mathbf{K}, \tag{33}$$

hence

$$\widetilde{\mathbf{X}} = \mu(\mathbf{K} + \mu\mathbf{I})^{-1}. \tag{34}$$

Combining (32) and (34) yields $\mathbf{I} - (\mathbf{K} + \mu\mathbf{I})^{-1}\mathbf{K} = \mu(\mathbf{K} + \mu\mathbf{I})^{-1}$, which proves the Lemma. ■

Appendix D
Proof of Lemma 3

Proof: Since matrix $\mathbf{B}$ is the inverse of matrix $\mathbf{A}$, we have $\mathbf{AB} = \mathbf{I}$, or

$$\begin{pmatrix} \mathbf{A}_{n-1} & \mathbf{a}_n \\ \mathbf{a}_n^\mathsf{T} & a_{nn} \end{pmatrix} \begin{pmatrix} \mathbf{B}_{n-1} & \mathbf{b}_n \\ \mathbf{b}_n^\mathsf{T} & b_{nn} \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix}. \tag{35}$$

Expanding (35) yields two equalities relating to $\mathbf{A}^{(n)}$,

$$\begin{cases} \mathbf{A}_{n-1}\mathbf{B}_{n-1} + \mathbf{a}_n\mathbf{b}_n^\mathsf{T} = \mathbf{I} \\ \mathbf{A}_{n-1}\mathbf{b}_n + \mathbf{a}_n b_{nn} = \mathbf{0}. \end{cases} \tag{36}$$

From the second equality, the vector $\mathbf{a}_n$ is computed as

$$\mathbf{a}_n = -\mathbf{A}_{n-1}\mathbf{b}_n / b_{nn}. \tag{37}$$

Substituting (37) into the first equality of (36) yields

$$\mathbf{A}_{n-1}\mathbf{B}_{n-1} - \mathbf{A}_{n-1}\mathbf{b}_n\mathbf{b}_n^\mathsf{T}/b_{nn} = \mathbf{I}, \tag{38}$$

or equivalently

$$\mathbf{A}_{n-1}(\mathbf{B}_{n-1} - \mathbf{b}_n\mathbf{b}_n^\mathsf{T}/b_{nn}) = \mathbf{I}, \tag{39}$$

which proves that the inverse of the submatrix $\mathbf{A}_{n-1}$ is

$$\mathbf{B}^{(n)} = \mathbf{B}_{n-1} - \mathbf{b}_n\mathbf{b}_n^\mathsf{T}/b_{nn}. \tag{40}$$

■

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, pp. 139–144, 1980.

[2] J. Ackermann and M. Goesele, *Survey of Photometric Stereo Techniques*. Delft, The Netherlands: Now Publishers Incorporated, 2015.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] C. Yu, Y. Seo, and S. W. Lee, "Photometric stereo from maximum feasible Lambertian reflections," in *Proc. Eur. Conf. Comput. Vis.*, 2010, pp. 115–126.

[9] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma, "Robust photometric stereo via low-rank matrix completion and recovery," in *Proc. Asian Conf. Comput. Vis.*, 2010, pp. 703–717.

[10] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa, "Robust photometric stereo using sparse regression," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, May 2012, pp. 318–325.

[11] 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.

[12] A. Ngan, F. Durand, and W. Matusik, "Experimental analysis of BRDF models," in *Proc. 16th Eurogr. Conf. Rendering Techn.*, 2005, pp. 117–126.

[13] 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.

[14] 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.

[15] J. Lawrence *et al.*, "Inverse shade trees for non-parametric material representation and editing," *ACM Trans. Graph.*, vol. 25, no. 3, pp. 735–745, 2006.

[16] N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[17] 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.

[18] 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.

[19] 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.

[20] T. Higo, Y. Matsushita, and K. Ikeuchi, "Consensus photometric stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1157–1164.

[21] 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.

[22] N. G. Alldrin and D. J. Kriegman, "Toward reconstructing surfaces with arbitrary isotropic reflectance : A stratified photometric stereo approach," in *Proc. 11th IEEE Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.

[23] 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.

[24] 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.

[25] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi, "A biquadratic reflectance model for radiometric image analysis," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 230–237.

[26] T.-Q. Han and H.-L. Shen, "Photometric stereo for general BRDFs via reflection sparsity modeling," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 4888–4903, Dec. 2015.

[27] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process.*, vol. 16, no. 2, pp. 349–366, Feb. 2007.

[28] T. Okabe, I. Sato, and Y. Sato, "Attached shadow coding: Estimating surface normals from shadows under unknown reflectance and lighting conditions," in *Proc. 12th IEEE Int. Conf. Comput. Vis.*, Sep. 2009, pp. 1693–1700.

[29] M. Chandraker, S. Agarwal, and K. David, "Shadowcuts: Photometric stereo with shadows," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.

[30] 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.

[31] G. J. Ward, "Measuring and modeling anisotropic reflection," *ACM SIGGRAPH Comput. Graph.*, vol. 26, no. 2, pp. 265–272, 1992.

[32] P. N. Belhumeur, D. J. Kriegman, and A. L. Yuille, "The bas-relief ambiguity," *Int. J. Comput. Vis.*, vol. 35, no. 1, pp. 1060–1066, Jan. 1997.

[33] B. Shi, Y. Matsushita, Y. Wei, C. Xu, and P. Tan, "Self-calibrating photometric stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1118–1125.

[34] N. G. Alldrin, S. P. Mallick, and D. J. Kriegman, "Resolving the generalized bas-relief ambiguity by entropy minimization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–7.

[35] I. Sato, T. Okabe, Q. Yu, and Y. Sato, "Shape reconstruction based on similarity in radiance changes under varying illumination," in *Proc. 11th IEEE Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.

[36] F. Lu, Y. Matsushita, I. Sato, T. Okabe, and Y. Sato, "From intensity profile to surface normal: Photometric stereo for unknown light sources and isotropic reflectances," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 10, pp. 1999–2012, Oct. 2015.

[37] F. Lu, I. Sato, and Y. Sato, "Uncalibrated photometric stereo based on elevation angle recovery from BRDF symmetry of isotropic materials," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 168–176.

[38] C. Wu, Y. Liu, Q. Dai, and B. Wilburn, "Fusing multiview and photometric stereo for 3D reconstruction under uncalibrated illumination," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 8, pp. 1082–1095, Aug. 2011.

[39] B. Shi, Z. Wu, Z. Mo, D. Duan, S.-K. Yeung, and P. Tan, "A benchmark dataset and evaluation for non-lambertian and uncalibrated photometric stereo," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 3707–3716.

[40] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD, USA: Johns Hopkins Univ. Press, 1989.

[41] X. Yang, T. K. Sarkar, and E. Arvas, "A survey of conjugate gradient algorithms for solution of extreme eigen-problems of a symmetric matrix," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 37, no. 10, pp. 1550–1556, Oct. 1989.

[42] C. M. Bishop and N. M. Nasrabadi, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.

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

Tian-Qi Han received the B.Eng. degree from Tianjin University, China, in 2011, and the Ph.D. degree from Zhejiang University, China, in 2016. He is currently with Huawei Technologies Co. Ltd., China. His research interest is computer vision, especially photometric stereo.

Chunguang Li (M'14–SM'14) received the M.S. degree in pattern recognition and intelligent systems and the Ph.D. degree in circuits and systems from the University of Electronic Science and Technology of China, Chengdu, China, in 2002 and 2004, respectively.

He is currently a Professor with the College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China. His current research interests include statistical signal processing, machine learning, wireless sensor network, and image processing.