Variational estimation of inhomogeneous specular reflectance and illumination from a single view

Kenji Hara[1,*] and Ko Nishino[2]

[1]*Department of Visual Communication Design, Faculty of Design, Kyushu University, 4–9–1 Shiobaru, Minami-ku, Fukuoka-shi, 815–8540 Japan*
[2]*Department of Computer Science, College of Engineering, Drexel University, 3141 Chestnut Street, Philadelphia, Pennsylvania 19104, USA*
**Corresponding author: hara@design.kyushu-u.ac.jp*

Received July 16, 2010; revised November 24, 2010; accepted November 25, 2010; posted November 29, 2010 (Doc. ID 131830); published January 17, 2011

Estimating the illumination and the reflectance properties of an object surface from a few images is an important but challenging problem. The problem becomes even more challenging if we wish to deal with real-world objects that naturally have spatially inhomogeneous reflectance. In this paper, we derive a novel method for estimating the spatially varying specular reflectance properties of a surface of known geometry as well as the illumination distribution of a scene from a specular-only image, for instance, recovered from two images captured with a polarizer to separate reflection components. Unlike previous work, we do not assume the illumination to be a single point light source. We model specular reflection with a spherical statistical distribution and encode its spatial variation with a radial basis function (RBF) network of their parameter values, which allows us to formulate the simultaneous estimation of spatially varying specular reflectance and illumination as a constrained optimization based on the I-divergence measure. To solve it, we derive a variational algorithm based on the expectation maximization principle. At the same time, we estimate optimal encoding of the specular reflectance properties by learning the number, centers, and widths of the RBF hidden units. We demonstrate the effectiveness of the method on images of synthetic and real-world objects.

OCIS codes: 100.2960, 100.3190, 150.2950.

1. INTRODUCTION

Real-world objects and scenes exhibit various visual intricacies resulting from the complex interaction of light with their surfaces. This interaction can be described by that of the incident illumination with the surface geometry modulated by the reflectance properties of the object surface. Inverting this complex process to estimate their sources, namely, the illumination, geometry, and reflectance from images, has important applications in a wide range of areas. For instance, estimated illumination can be used to "unshade" the scene, recovered object/scene geometry provides a direct 3D cue, and inferred reflectance properties can offer unique material classifications, all of which become invaluable in practical applications, such as object/scene recognition and tracking. These estimates of the scene information can be directly used for image synthesis as well—to achieve photorealistic renderings of the object/scene under novel lighting and from novel viewpoints.

Sifting out all three constituents from a single observation, namely, recovering illumination, geometry, and reflectance properties from a single image, is, however, devastatingly ill posed. A more realistic and practical problem is to recover as much information from as few images as possible, making fewer assumptions as needed. Among all the possible problem settings, estimating illumination and reflectance properties from images assuming known geometry has drawn large interest due to its direct relevance to photometric scene understanding (estimating radiometric properties of a scene from images). Past methods recover the illumination [1] or the reflectance properties [2,3] separately, assuming the other to be known, or recover both [4] from a large set of images of object surfaces with known geometry. To make the problem tractable, these approaches, however, assume that the reflectance properties of the object surface are more or less uniform. For instance, Sato *et al.* [2] assume that the specular reflectance properties of the object surface change very smoothly on the surface such that they can be represented with a small set of sparsely sampled points on the object surface. If we further assume that the specular reflectance properties of the object surface are completely homogeneous across the surface, we may drastically reduce the number of required input images and still simultaneously estimate those parameter values as well as the illumination [5–8]. This inverse problem is, however, inherently ill posed and becomes even harder if we wish to account for the spatial variation of material properties on the surface.

Real-world surfaces usually have reflectance properties that vary across their surfaces, and their variations are not necessarily always smooth. Devising a computational method for estimating both the illumination and reflectance properties of such general object surfaces from a few images would be a vital step to make photometric scene understanding methods work in the real world. Zickler *et al.* [9] proposed a framework for recovering spatially varying reflectance properties from a sparse set of images based on the spatial coherence between spatial and angular data. They formulated reflectance estimation as the problem of interpolating scattered data in a mixed

spatial and angular domain and stably recovering the spatially varying reflectance by considering the trade-off relation between spatial and angular reflectance resolution. In their work, however, the illumination was assumed to be known in advance. To our knowledge, the problem of recovering spatially varying specular reflectance and scene illumination from a sparse set of images has not been tackled before.

In this paper, we derive a novel method for estimating the spatially varying specular reflectance properties of a surface of known geometry as well as the illumination distribution from a specular-only image (one that only contains the specular reflection of the surface). Such an image can be captured, for instance, using polarization to separate reflection components. The remaining reflectance property, i.e., the diffuse reflection, can be estimated by unshading the preseparated diffuse-only image using the estimated illumination, for instance, by assuming Lambertian reflection. Unlike previous methods that achieve the same goal, but from a large number of images, we do not assume the illumination to be a single point light source [10]. In contrast to the method by Zickler *et al.* [9], we analytically model the surface reflectance, in particular its specular component, to derive a sound variational method that can jointly estimate the illumination and reflectance from a sparse angular and spatial sampling.

We make minimal assumptions about the scene illumination and specular reflectance property. We assume that the illumination environment is composed of a set of discrete distant point light sources (directional illumination) with the same color (thus, finite area light sources are not included), and the specular peak corresponding to all sources can be observed in the input specular image. The number and direction of distant point light sources can be estimated by using conventional methods, as in [5,6]. For simplicity, we use a semiautomatic method for achieving this estimation. The light source directions can be easily obtained by calculating the perfect mirror direction at each specular peak pixel and, thus, the unknown illumination parameters to be estimated are the light source intensities.

The key contribution of the proposed method is its formulation of the simultaneous estimation of spatially varying specular reflectance and illumination as a variational probabilistic inference problem. This particular formulation enables us to devise canonical statistical error measures and derive a robust joint estimation algorithm for minimizing it, which are vital to solving the otherwise ill-conditioned problem. To this end, we model specular reflection with a spherical statistical distribution and encode the spatial variation with radial basis functions (RBFs) of its parameters. We then formulate the joint estimation as a constrained optimization based on Csiszár's I-divergence measure [11] (hereinafter referred to as the "I-divergence") and derive an iterative algorithm similar to expectation maximization (EM) [12] to solve it. We demonstrate the effectiveness of the method on synthetic and real-world scenes. The results show that we can successfully estimate both the scene illumination and spatially varying specular reflectance properties from a single image and, as a result, the full reflectance properties may be estimated from only two images from the same view. To our knowledge, this paper is the first to demonstrate such capability.

2. RELATED WORK

Computationally inverting the radiometric image formation to estimate various combinations of the three key ingredients of scene appearance, namely, the geometry, reflectance, and illumination, has been an important area of research in computer vision and optics communities. Among all the possible problem settings, joint estimation of reflectance and illumination, as well as joint estimation of reflectance and geometry, have been of particular relevance in these disciplines. Here, we review some of the representative works that tackle these two joint estimations that are relevant to our work. Other joint estimations are excluded from this section, as they are either less relevant for practical scenarios (joint estimation of illumination and geometry) or remain open to canonical approaches (joint estimation of all three factors).

Joint estimation of reflectance and geometry has drawn wide attention, especially in the computer vision community. This joint estimation is a key hurdle to extend conventional three-dimensional (3D) reconstruction methods to handle real-world surfaces that inevitably have more complex reflectance properties than a constant Lambertian albedo. Nayar *et al.* [13] proposed to determine the shape and reflectance parameters of Lambertian, specular, and hybrid surfaces using a photometric sampling method. The shape information is estimated without prior knowledge of the relative strengths of the Lambertian and specular reflection components. Sato and Ikeuchi [14] analyzed a sequence of RGB color images in a four-dimensional space, referred to as the temporal-color space, to recover albedo values of specular and diffuse reflection components and the shape of the object. Yuile *et al.* [15] generalized the singular value decomposition (SVD) approach for estimating shape and Lambertian albedo from multiple images to include ambient background illumination. They analyzed the linear ambiguity remaining after SVD and developed an iterative algorithm to improve the accuracy of the estimates by detecting shadows as outliers and removing them. Zhao and Chellappa [16] presented a symmetric shape-from-shading (SFS) approach to recover shape and spatially varying albedo from a single image of symmetric Lambertian objects. They related the self-ratio image irradiance equation to the standard image irradiance equation used in SFS. Prior to that, Zheng and Chellappa [17] introduced a method that estimates all three factors in a restricting setting, namely, a single light source with Lambertian albedo surface. They estimated the direction of a single light source and Lambertian albedo of the surface, taking into consideration the effect of self-shadowing, and then computed the surface shape by imposing a smoothness constraint. Smith and Hancock [18] embedded a statistical model of facial shape within a SFS algorithm. They presented two SFS-based methods for fitting the model to image brightness data. Birkbeck *et al.* [19] presented a method that alternatively reconstructs shape and general reflectance of both textured and nontextured surfaces from images. Their algorithm recovers the geometry of objects with large concavities and specular reflections.

A few methods that handle inhomogeneous reflectance and go beyond previous assumptions of simplistic textured Lambertian plus homogeneous specular reflection have been introduced. Alldrin *et al.*[20] introduced a method for simultaneously estimating nonparametric inhomogeneous bidirectional reflectance distribution function (BRDF) of a surface

as well as the 3D shape from images of an object taken under varying points of source lighting directions. The method requires many images, on the order of at least tens, as it uses a data-driven reflectance model based on a bivariate approximations of isotropic reflectance functions. Biswas *et al.* [21] solved the general SFS problem of estimating the shape and albedo of an object with a varying albedo map and unknown illuminant direction from a single image. Their method effectively utilized the statistics of errors in illumination and normal information for robust estimation of albedo for images illuminated by single and multiple light sources. These past works largely focus on dealing with spatially varying Lambertian reflectance for 3D reconstruction. We believe that a variational approach, as we propose in this paper, for joint estimation of reflectance and illumination can be generalized to further extend such work to deal with more general spatially inhomogeneous reflectance in the context of geometry recovery.

Joint estimation of reflectance and illumination is another important, challenging problem that finds many practical applications, especially in computer vision and graphics ranging from robust tracking to photorealistic image synthesis. Ikeuchi and Sato [22] used a simplified Torrance–Sparrow reflection model to estimate its parameters together with the Lambertian diffuse albedo and the light source direction using iterative least squares optimization from a single image of an object with known geometry. Sato *et al.* [23] analyzed the radiance distribution inside shadows cast by an object in the scene to recover the illumination distribution and the surface reflectance properties. In particular, they introduced an adaptive sampling framework for efficient estimation. Tominaga and Tanaka [24] analyzed the color histogram of a single color image of an object to estimate the parameters of the Phong reflection model as well as the direction and the color of a directional light source. Ramamoorthi and Hanrahan [7] derived a signal-processing framework for inverse rendering. They represented the reflected light field with a convolution of the lighting and BRDF in the angular space and demonstrated a method for factoring the light field to simultaneously estimate the lighting and BRDF. Nishino *et al.* [6] presented a method to estimate a diffuse texture map, the specular reflectance parameter, and the illumination distribution from a few images of an object taken from different viewpoints under static illumination. They achieved the simultaneous estimation by formulating the specular reflection mechanism as a two-dimensional blind convolution on the surface of a hemisphere, and then running an iterative algorithm to deconvolve it. Hara *et al.* [5] proposed to simultaneously estimate the illumination of a scene and the specular reflectance property of a surface from single-view images. They modeled the specular reflection as a finite mixture of a directional statistics distribution and estimated the mixture parameters corresponding to the illumination and reflectance parameters.

These past methods fundamentally assume homogeneous reflectance, except for Lambertian albedo variation expressed as texture. Recently, Yu *et al.* [25] recovered spatially varying diffuse/specular albedo maps together with an illumination map from multiview images by using a tensor factorization framework. This method fundamentally relies on a data-driven representation and inevitably requires tens of images captured from different viewpoints. In order to handle the appearance of real-world surfaces and estimate their reflectance properties as well as the scene illumination from as few images as possible, we need to be able to model and jointly estimate spatially varying, inhomogeneous surface reflectance in a compact analytical form. We focus on achieving this from as few as two input images taken from a single view. To our knowledge, joint estimation of inhomogeneous reflectance and illumination from such few images has not been studied in the past.

3. SPHERICAL TORRANCE–SPARROW MODEL

The key idea underlying the proposed method is to cast the joint estimation of illumination and specular reflectance properties as a probabilistic inference problem so that a joint estimation algorithm based on statistical principals can be derived. A crucial step to derive this formulation is to represent the illumination and the reflectance properties with the same parametric form that, at the same time, enables statistical interpretation. To this end, we consider encoding the specular light interaction with a spherical distribution, i.e., a directional statistics distribution. In particular, we adopt the spherical Torrance–Sparrow reflection model introduced by Hara *et al.* [5]. It is worth noting that our proposed method can be adapted to use a wide variety of other BRDF models, including the Blinn–Phong model, and microfacet models, such as the original Torrance–Sparrow. The spherical Torrance–Sparrow reflection model, however, has the strong advantage of being expressed in a spherical domain that allows us to encode both the reflectance and illumination in a unified analytic expression.

The spherical Torrance–Sparrow reflection model is based on and approximates the Torrance–Sparrow reflection model. The original Torrance–Sparrow reflection model [26] represents specular reflection with the aggregated light reflection from a collection of microfacets lying within an infinitesimally small surface region corresponding to a pixel, each having perfect mirror reflection but whose orientations are statistically distributed. Assuming that the distribution of the orientations of microfacets can be modeled with the von Mises–Fisher distribution [27,28]—a Gaussian distribution on the unit sphere (see Fig. 1)—the spherical Torrance–Sparrow reflection model [5] represents specular reflection as

$$\mathbf{I}_S = \int_{-\pi}^{\pi} \int_{0}^{\frac{\pi}{2}} \frac{\mathbf{K}_S F G}{\cos\theta_r} L_i(\theta_i, \phi_i) \exp[-2\kappa \sin^2 \alpha] \sin\theta_i \mathrm{d}\theta_i \mathrm{d}\phi_i, \quad (1)$$

where $\mathbf{I}_S$ denotes a three-band color vector of the specular reflection radiance. As in the original Torrance–Sparrow

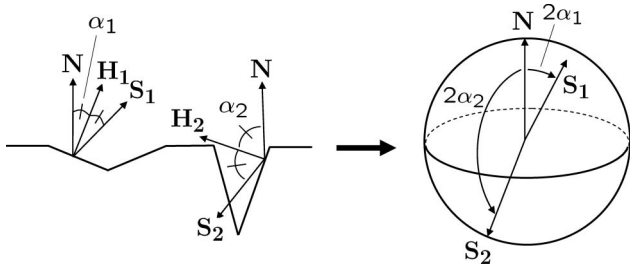

Fig. 1. Spherical Torrance–Sparrow reflection model [5] represents the microfacet orientation distribution of the Torrance–Sparrow reflection model [26] with a directional statistics distribution, namely, the von Mises–Fisher distribution [27].

reflection model, $\mathbf{K}_S$ is the color vector of the specular reflection, F is the Fresnel reflectance coefficient, G is the geometrical attenuation factor, θ_r is the angle between the viewing direction and the surface normal, and θ_i and ϕ_i are the polar and azimuth angles, respectively. $L_i(\theta_i, \phi_i)$ is the illumination radiance per unit solid angle coming from the direction (θ_i, ϕ_i), $\sin\theta_i d\theta_i d\phi_i$ is the infinitesimal solid angle, and α is the angle between the surface normal and the bisector of the viewing direction and the light source direction [26]. Similar to the spherical Torrance–Sparrow reflection model, κ encodes the surface roughness, which is related to the surface roughness σ of the original Torrance–Sparrow reflection model [26] by $\kappa \simeq 1/4\sigma^2$. Note that the value of this vector subsumes the normalization factor of the exponential function, the reflectivity of the surface, and the scaling factor between scene radiance and a pixel value.

As in the case of the original Torrance–Sparrow reflection model [26], we may simplify Eq. (1) by assuming that the Fresnel factor and the geometric attenuation factor remain constant, which is valid for nongrazing angle incident light [29]. Then the specular image irradiance $\mathbf{I}_S(\mathbf{x})$ at image coordinate $\mathbf{x} = (x, y)^T$ of surface point $P(\mathbf{x})$ can be written as

$$\mathbf{I}_S(\mathbf{x}) = \int_{-\pi}^{\pi}\int_{0}^{\frac{\pi}{2}} \frac{\mathbf{K}_S(\mathbf{x})}{\cos\theta_r(\mathbf{x})} L_i(\theta_i, \phi_i) \times \exp[-2\kappa(\mathbf{x})\sin^2\alpha(\mathbf{x})]\sin\theta_i \mathrm{d}\theta_i \mathrm{d}\phi_i. \quad (2)$$

We assume that the illumination environment can be decomposed into a set of distant point light sources of uniform color $\boldsymbol{L}$, which is valid as long as we have an environmental illumination, e.g., the object size is much smaller compared to the distance to the closest light source, of roughly the same color. Under this assumption, Eq. (2) can be further reduced and discretized

$$\mathbf{I}_S(\mathbf{x}) \simeq I_S(\mathbf{x})\mathbf{L}, \quad (3)$$

$$I_S(\mathbf{x}) = \frac{2\pi}{N_L}\frac{\mathbf{K}_S(\mathbf{x})}{\cos\theta_r(\mathbf{x})}\sum_{l=1}^{L} L_l \exp[-2\kappa(\mathbf{x})\sin^2\alpha_l(\mathbf{x})], \quad (4)$$

where all the parameters are scalars: L_l is the magnitude of the color vector of the lth point light source and N_L is the number of nodes in the geodesic hemisphere [6].

4. PROBABILISTIC FORMULATION

The spherical Torrance–Sparrow reflection model provides a clean directional statistics interpretation of the specular reflection as described in Eq. (4), which encodes the interaction of illumination, i.e., the distribution of L_l on a unit hemisphere and the reflectance parameters K_s and κ in a unified representation. We can now formulate the joint estimation of illumination and reflectance properties as simultaneously estimating all parameters in this equation. To proceed, we first rewrite Eq. (4) as

$$\tilde{I}(\mathbf{x}) = \sum_{l=1}^{L} \mathcal{I}_l(\mathbf{x}), \quad (5)$$

$$\mathcal{I}_l(\mathbf{x}) = \tilde{L}_l K_s(\mathbf{x}) \exp[-\kappa(\mathbf{x})\omega_l(\mathbf{x})], \quad (6)$$

where $\tilde{I}(\mathbf{x}) = I_S(\mathbf{x})\cos\theta_r(\mathbf{x})$, $\tilde{L}_l = 2\pi L_l/N_L$, and $\omega_l(\mathbf{x}) = 2\sin^2\alpha_l(\mathbf{x})$. Here, we assume that the extrinsic camera parameters as well as the shape and position of the object are known. Thus, $\theta_r(\mathbf{x})$ is known. Note, however, that the two unknown parameters $\tilde{L}_l$ and $K_s(\mathbf{x})$ cannot be uniquely determined from Eq. (6): the equation is bilinear in these two parameters. Hence, we impose an additional constraint without loss of generality:

$$\sum_{l=1}^{L} \tilde{L}_l = 1, \quad (7)$$

which corresponds to redefining $\tilde{L}_l$ as the relative radiance of the lth light source.

In order to represent inhomogeneous specular reflectance properties, we assume that the variation of the values of the two parameters $K_s(\mathbf{x})$ and $\kappa(\mathbf{x})$ can be encoded with a set of RBFs defined on a uniform grid in the image plane

$$K_s(\mathbf{x}) \simeq K_s(\mathbf{x};\boldsymbol{\Theta}) = \exp\left[\sum_{i=1}^{I} K_i \Phi_i(\mathbf{x})\right], \quad (8)$$

$$\kappa(\mathbf{x}) \simeq \kappa(\mathbf{x};\boldsymbol{\theta}) = \sum_{j=1}^{J} \kappa_j \phi_j(\mathbf{x}), \quad (9)$$

where the logarithm of the approximating function $K_s(\mathbf{x};\boldsymbol{\Theta})$ for $K_s(\mathbf{x})$ is represented as a weighted sum of I RBFs $\{\Phi_i(\mathbf{x})\}_{i=1}^{I}$ with unknown weight coefficients $\boldsymbol{\Theta} = \{K_i\}_{i=1}^{I}$, while the approximating function $\kappa(\mathbf{x};\boldsymbol{\theta})$ for $\kappa(\mathbf{x})$ is directly represented as a weighted sum of J RBFs $\{\phi_j(\mathbf{x})\}_{j=1}^{J}$ with unknown weight coefficients $\boldsymbol{\theta} = \{\kappa_j\}_{j=1}^{J}$. Also, we use the RBF network for function approximation in the physical parameter space, while most conventional methods directly approximate the observed data (e.g., [9]). Such approaches that essentially just interpolate the observed data would not provide a physical interpretation of the surface reflectance. In contrast, representing the parameters of the reflection model with an RBF network directly provides a physical interpretation of the surface variation, e.g., its glossiness or roughness. Here, we use a Gaussian distribution for each RBF

$$\Phi_i(\mathbf{x}) = \exp\left[-\frac{\|\mathbf{x} - \boldsymbol{C}_i\|^2}{2R_i^2}\right] \quad (i = 1, 2, \cdots, I), \quad (10)$$

$$\phi_j(\mathbf{x}) = \exp\left[-\frac{\|\mathbf{x} - \boldsymbol{c}_j\|^2}{2r_j^2}\right] \quad (j = 1, 2, \cdots, J), \quad (11)$$

where $\boldsymbol{C}_i$ and R_i are the center and width (variance) for $\Phi_i(\mathbf{x})$, respectively, and $\boldsymbol{c}_j$ and r_j are the center and width (variance) for $\phi_i(\mathbf{x})$, respectively. Typically, the RBFs for function approximation are chosen such that the number of basis functions is less than the number of data samples, the centers are spaced at regular intervals, and the widths are double the center intervals. Based on this heuristic, the initial values of $\{I, J, \{\boldsymbol{C}_i, R_i\}_{i=1}^{I}, \{\boldsymbol{c}_j, r_j\}_{j=1}^{J}\}$ were set to constant values (Section 5) and then properly optimized (Section 6). We empirically found the optimization results to be robust to changes in these constant initial values.

By substituting Eqs. (8) and (9) into Eqs. (5) and (6), we obtain

$$\tilde{I}(\mathbf{x}) \simeq \tilde{I}(\mathbf{x};\boldsymbol{\Omega}) = \sum_{l=1}^{L} \mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l), \tag{12}$$

$$\begin{aligned}\mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l) &= \tilde{L}_l \exp\left[\sum_{i=1}^{I} K_i \Phi_i(\mathbf{x}) - \omega_l(\mathbf{x}) \sum_{j=1}^{J} \kappa_j \phi_j(\mathbf{x})\right] \\ &= \exp\left[\ln \tilde{L}_l + \sum_{i=1}^{I} K_i \Phi_i(\mathbf{x}) - \omega_l(\mathbf{x}) \sum_{j=1}^{J} \kappa_j \phi_j(\mathbf{x})\right], \end{aligned} \tag{13}$$

where $\boldsymbol{\Omega} = \{\boldsymbol{\Theta}, \boldsymbol{\theta}, \mathcal{L}\} = \{\{K_i\}_{i=1}^{I}, \{\kappa_j\}_{j=1}^{J}, \{\tilde{L}_l\}_{l=1}^{L}\}$ and $\boldsymbol{\Omega}_l = \{\boldsymbol{\Theta}, \boldsymbol{\theta}, \tilde{L}_l\} = \{\{K_i\}_{i=1}^{I}, \{\kappa_j\}_{j=1}^{J}, \tilde{L}_l\}$.

The problem of simultaneously estimating the specular reflection parameters $\{K_s(\mathbf{x}), \kappa(\mathbf{x})\}$ and illumination $\mathcal{L} = \{\tilde{L}_l\}_{l=1}^{L}$ has thus been formulated as determining the unknown parameters $\boldsymbol{\Omega} = \{\{K_i\}_{i=1}^{I}, \{\kappa_j\}_{j=1}^{J}, \{\tilde{L}_l\}_{l=1}^{L}\}$ in Eqs. (12) and (13). From Eq. (13), the joint estimation of the spatially varying inhomogeneous reflectance and illumination leads to optimizing the weight coefficients $\boldsymbol{\Omega}_l$ of each lth RBF network for approximating its corresponding continuous function $\ln(\mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l))$ whose hidden units are $\{\{\Phi_i(\mathbf{x})\}_{i=1}^{I}, \{\omega_l(\mathbf{x})\phi_j(\mathbf{x})\}_{j=1}^{J}, 1\}$ and whose second-layer weights are $\{\{K_i\}_{i=1}^{I}, \{-\kappa_j\}_{j=1}^{J}, \ln \tilde{L}_l\}$ (Fig. 2). Although $\{\omega_l(\mathbf{x})\phi_j(\mathbf{x})\}_{j=1}^{J}$ are not, strictly speaking, radially symmetric (we refer to them as "pseudo-RBFs"), just like the common RBFs, these basis functions yield coefficient matrices whose nonzero matrix elements are concentrated around the diagonal due to their locality and smoothness and, hence, numerically stable learning can be expected to be achieved even if the number of unknown weight coefficients is large.

5. PARAMETER ESTIMATION AS PROBABILISTIC INFERENCE

We formulate parameter estimation, $\boldsymbol{\Omega}$, as a probabilistic inference problem, in which the predicted specular intensities should become as close as possible to the observed intensities. Since we have expressed the predicted values as a directional statistics distribution $\tilde{I}(\mathbf{x})$, we may naturally measure the discrepancy between this distribution and the observed distribution, which we define as $D(\mathbf{x})$ that encodes the input specular image irradiance at $\mathbf{x}$ multiplied by $\cos\theta_r(\mathbf{x})$, using a statistical measure. In particular, we use the I-divergence [11], which is a generalization of the well-known Kullback–Leibler divergence and is a natural discrepancy measure for nonnegative data:

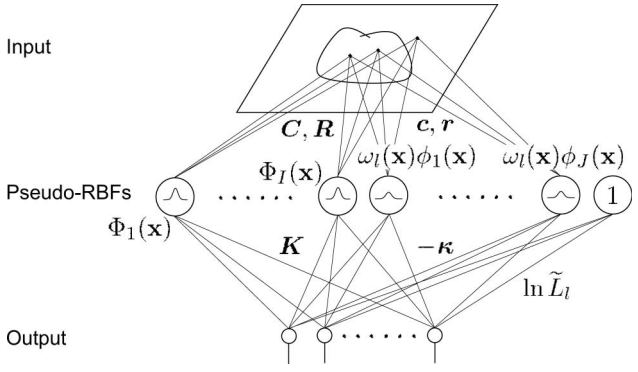

Fig. 2. We formulate the joint estimation of spatially varying, inhomogeneous reflectance and illumination as learning the hidden units of a pseudo-RBF network.

$$E(\boldsymbol{\Omega}) = \iint_S \left\{ D(\mathbf{x}) \log \frac{D(\mathbf{x})}{\tilde{I}(\mathbf{x};\Omega)} - [D(\mathbf{x}) - \tilde{I}(\mathbf{x};\boldsymbol{\Omega})] \right\} \mathrm{d}\mathbf{x}, \tag{14}$$

where S is the image region corresponding to the target surface area. The goal is then to minimize this I-divergence with respect to $\boldsymbol{\Omega} = \{\{K_i\}_{i=1}^{I}, \{\kappa_j\}_{j=1}^{J}, \{\tilde{L}_l\}_{l=1}^{L}\}$, subject to the constraint given by Eq. (7). The numbers ($\Lambda = \{I, J\}$) and the centers and widths ($\boldsymbol{\Gamma} = \{\{\boldsymbol{C}_i, R_i\}_{i=1}^{I}, \{\boldsymbol{c}_j, r_j\}_{j=1}^{J}\}$) of the RBF hidden units in Eqs. (10) and (11) were fixed constant in this section.

The advantage of this statistical formulation is that we can devise sound probabilistic inference algorithms to solve the otherwise tangled joint estimation problem. In other words, we can leverage established alternating minimization algorithms that do not just empirically iterate between minimizing the error function with respect to each parameter set, but are constructed based on the principals of EM with sound probabilistic interpretations. The problem is well studied in the signal processing and machine learning communities. Kameoka *et al.* [30] derived an iterative algorithm, which is formally similar to EM, to minimize the I-divergence. In this section, we derive a variational estimation algorithm for minimizing the statistical discrepancy measure in Eq. (14) based on the fundamentals of their iterative algorithm.

The derived EM algorithm is not guaranteed to converge to the global optimum as other EM-based algorithms. We may consider a more greedy approach, like the annealing-based EM algorithms [31,32], for guaranteed global optimum, but this will incur significant computational overhead. An efficient computation of such an approach is left for future work.

First, let us introduce a set of functions, $\boldsymbol{m} = \{m_l(\mathbf{x})\}_{l=1}^{L}$, that satisfies both $m_l(\mathbf{x}) \in (0,1)$ and $\sum_{l=1}^{L} m_l(\mathbf{x}) = 1$ for $\forall l \in \{1, 2, \cdots, L\}$ and $\forall \mathbf{x} \in S$ and partitions the values of $D(\mathbf{x})$ into L light sources. We refer to these functions as the masking functions. Note that the masking functions take on fractional values that each encode the spatial contribution of the corresponding light source. Also, the number, L, of distant point light sources can be estimated by using methods as in [5]. In this paper, we focus on the problem of simultaneous estimation of inhomogeneous specular reflectance and point light source intensities, and assume L is assumed to be known. Using these functions, we obtain a useful variational upper bound

$$\begin{aligned} E(\boldsymbol{\Omega}) &\le E^{+}(\boldsymbol{\Omega}, \boldsymbol{m}) \\ &= \iint_S \left\{ \sum_{l=1}^{L} m_l(\mathbf{x}) D(\mathbf{x}) \log \frac{m_l(\mathbf{x}) D(\mathbf{x})}{\mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)} \right. \\ &\quad \left. - \sum_{l=1}^{L} [m_l(\mathbf{x}) D(\mathbf{x}) - \mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)] \right\} \mathrm{d}\mathbf{x} \\ &= \sum_{l=1}^{L} \iint_S \left\{ \mathcal{D}_l(\mathbf{x}) \log \frac{\mathcal{D}_l(\mathbf{x})}{\mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)} - [\mathcal{D}_l(\mathbf{x}) - \mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)] \right\} \mathrm{d}\mathbf{x}, \end{aligned} \tag{15}$$

where $\mathcal{D}_l(\mathbf{x}) = m_l(\mathbf{x}) D(\mathbf{x}) (l = 1, 2, \cdots, L)$.

We may then first minimize $E^{+}(\boldsymbol{\Omega}, \boldsymbol{m})$ with respect to $\boldsymbol{m}(\mathbf{x})$ while keeping $\boldsymbol{\Omega}$ fixed, which leads to $E(\boldsymbol{\Omega}) = E^{+}(\boldsymbol{\Omega}, \boldsymbol{m})$, and preserving the inequality in Eq. (15). The masking functions $\boldsymbol{m}(\mathbf{x})$ that minimize $E^{+}(\boldsymbol{\Omega}, \boldsymbol{m})$ can be obtained analytically as

$$m_l(\mathbf{x}) = \frac{\mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)}{\sum_{l=1}^{L}\mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)} \qquad (l = 1, 2, \cdots, L). \tag{16}$$

Then, we may minimize $E^+(\boldsymbol{\Omega}, \boldsymbol{m})$ with respect to $\boldsymbol{\Omega}$ while keeping $\boldsymbol{m}$ fixed. These two computations are iterated until convergence.

We now derive the update equations for $\boldsymbol{\Omega}$. First, the partial derivatives of $E^+(\boldsymbol{\Theta}, \boldsymbol{m})$ [Eq. (15)] with respect to an I-dimensional vector $\boldsymbol{K} = [K_1, \cdots, K_I]^T$ and a J-dimensional vector $\boldsymbol{\kappa} = [\kappa_1, \cdots, \kappa_J]^T$ are

$$\frac{\partial E^+}{\partial \boldsymbol{K}} = -\sum_{l=1}^{L}\iint_S [\mathcal{D}_l(\mathbf{x}) - \mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)]\boldsymbol{\Phi}(\mathbf{x})\mathrm{d}\mathbf{x}, \tag{17}$$

$$\frac{\partial E^+}{\partial \boldsymbol{\kappa}} = \sum_{l=1}^{L}\iint_S [\mathcal{D}_l(\mathbf{x}) - \mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)]\omega_l(\mathbf{x})\boldsymbol{\phi}(\mathbf{x})\mathrm{d}\mathbf{x}, \tag{18}$$

where $\boldsymbol{\Phi}(\mathbf{x}) = [\Phi_1(\mathbf{x}), \cdots, \Phi_I(\mathbf{x})]^T$ and $\boldsymbol{\phi}(\mathbf{x}) = [\phi_1(\mathbf{x}), \cdots, \phi_J(\mathbf{x})]^T$. Note that, in this paper, the partial derivatives of $E^+(\cdot)$ with respect to column vectors $\boldsymbol{K} = [K_1, \cdots, K_I]^T$ and $\boldsymbol{\kappa} = [\kappa_1, \cdots, \kappa_J]^T$ are defined as column vectors $\frac{\partial E^+}{\partial \boldsymbol{K}} = [\frac{\partial E^+}{\partial K_1}, \cdots, \frac{\partial E^+}{\partial K_I}]^T$ and $\frac{\partial E^+}{\partial \boldsymbol{\kappa}} = [\frac{\partial E^+}{\partial \kappa_1}, \cdots, \frac{\partial E^+}{\partial \kappa_J}]^T$, respectively.

Using these partial derivatives, $\{K_i^{(t-1)}\}_{i=1}^I$ and $\{\kappa_j^{(t-1)}\}_{j=1}^J$ can be updated using gradient-based local minimization techniques, such as steepest descent and conjugate gradient, where $\{K_i^{(t-1)}\}_{i=1}^I$ and $\{\kappa_j^{(t-1)}\}_{j=1}^J$ are the values assumed by $\{K_i\}_{i=1}^I$ and $\{\kappa_j\}_{j=1}^J$ at the $t-1$th step of Algorithm 1 described later in this section, respectively.

To derive the update rule for $\{\tilde{L}_l\}_{l=1}^L$, we must account for the constraint given by (7). Here, we minimize the Lagrange function

$$J(\boldsymbol{\Omega}, \boldsymbol{m}, \lambda) = E^+(\boldsymbol{\Omega}, \boldsymbol{m}) - \lambda\left(\sum_{l=1}^{L}\tilde{L}_l - 1\right), \tag{19}$$

where λ is a Lagrange multiplier.

The partial derivatives of $J(\boldsymbol{\Omega}, \boldsymbol{m}, \lambda)$ with respect to $\tilde{L}_l$ and λ are

$$\frac{\partial J}{\partial \tilde{L}_l} = \frac{1}{\tilde{L}_l}\iint_S [-\mathcal{D}_l(\mathbf{x}) + \mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)]\mathrm{d}\mathbf{x} - \lambda \qquad (l = 1, \cdots, L), \tag{20}$$

$$\frac{\partial J}{\partial \lambda} = -\sum_{l=1}^{L}\tilde{L}_l + 1. \tag{21}$$

Setting Eqs. (20) and (21) to zero, we obtain

$$\tilde{L}_l = -\frac{q_l^{(t+1)}}{\lambda - p_l^{(t+1)}} \qquad (l = 1, \cdots, L), \tag{22}$$

$$-\sum_{l=1}^{L}\tilde{L}_l + 1 = 0, \tag{23}$$

where $\{p_l^{(t+1)}\}_{l=1}^L$ and $\{q_l^{(t+1)}\}_{l=1}^L$ are the values obtained from the values of $\{K_i^{(t+1)}\}_{i=1}^I$ and $\{\kappa_j^{(t+1)}\}_{j=1}^J$ as

$$p_l^{(t)} = \iint_S K_s^{(t)}(\mathbf{x})\exp[-\kappa^{(t)}(\mathbf{x})\omega_l(\mathbf{x})]\mathrm{d}\mathbf{x} \qquad (l = 1, \cdots, L), \tag{24}$$

$$q_l^{(t)} = \iint_S \mathcal{D}_l^{(t)}(\mathbf{x})\mathrm{d}\mathbf{x} \qquad (l = 1, \cdots, L), \tag{25}$$

where $K_s^{(t)}(\mathbf{x}) = \exp[\sum_{i=1}^I K_i^{(t)}\Phi_i(\mathbf{x})]$, $\kappa^{(t)}(\mathbf{x}) = \sum_{j=1}^J \kappa_j^{(t)}\phi_j(\mathbf{x})$, and $\mathcal{D}_l^{(t)}(\mathbf{x}) = m_l^{(t)}(\mathbf{x})D(\mathbf{x})$.

To update $\tilde{L}_l$, we solve $L+1$ simultaneous Eqs. (22) and (23) with $L+1$ unknowns $\{\{\tilde{L}_l\}_{l=1}^L, \lambda\}$. By substituting Eq. (22) into Eq. (23), we get a fractional equation with an unknown variable λ:

$$g^{(t+1)}(\lambda) = \sum_{l=1}^{L}\frac{q_l^{(t+1)}}{\lambda - p_l^{(t+1)}} + 1 = 0. \tag{26}$$

Hence, we numerically solve this equation using the Newton–Raphson method, which iterates the following procedure until convergence:

$$\lambda^{(\tau_\lambda+1)} = \lambda^{(\tau_\lambda)} - \frac{g^{(t+1)}(\lambda^{(\tau_\lambda)})}{g^{(t+1)}(\lambda^{(\tau_\lambda)})}. \tag{27}$$

Finally, by substituting those values into each of Eq. (22), $\{\tilde{L}_l\}_{l=1}^L$ can be updated as follows:

$$\tilde{L}_l^{(t+1)} = -\frac{q_l^{(t+1)}}{\lambda^{(\tau_\lambda)} - p_l^{(t+1)}} \qquad (l = 1, \cdots, L), \tag{28}$$

where $\{\tilde{L}_l^{(t+1)}\}_{l=1}^L$ are the values assumed by $\{\tilde{L}_l\}_{l=1}^L$ at the $t+1$th step of Algorithm 1. The iterative algorithm in this section can be summarized as follows.

Algorithm 1

Step 1. Initialize:
Set $\boldsymbol{\Omega}^{(0)} \Leftarrow \{\{K_i^{(0)}\}_{i=1}^I, \{\kappa_j^{(0)}\}_{j=1}^J, \{\tilde{L}_l^{(0)}\}_{l=1}^L\}$. Set $t \Leftarrow 0$.
Step 2. Repeat the steps below until convergence.
E-step: set $\boldsymbol{m}^{(t+1)} \Leftarrow \underset{\boldsymbol{m}}{\mathrm{argmin}} E^+(\boldsymbol{\Omega}^{(t)}, \boldsymbol{m})$.
M-step: set $\boldsymbol{\Omega}^{(t+1)} \Leftarrow \underset{\boldsymbol{\Omega}}{\mathrm{argmin}} E^+(\boldsymbol{\Omega}, \boldsymbol{m}^{(t+1)})$. Set $t \Leftarrow t+1$.
Step 3. Output:
Set $K_s(\mathbf{x}) \Leftarrow \exp[\sum_{i=1}^I K_i^{(t)}\Phi_i(\mathbf{x})]$. Set $\kappa(\mathbf{x}) \Leftarrow \sum_{j=1}^J \kappa_j^{(t)}\phi_j(\mathbf{x})$.
Set $\{\tilde{L}_l\}_{l=1}^L \Leftarrow \{\tilde{L}_l^{(t)}\}_{l=1}^L$.

6. REFINEMENT OF THE RBF NETWORK STRUCTURE

To obtain a minimal RBF network that explains the input data faithfully, we remove the unnecessary hidden units while optimizing the network parameters $\boldsymbol{\Omega}$ and $\boldsymbol{\Gamma}$. $\mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l)$ in Eq. (13) and $E^+(\boldsymbol{\Omega}, \boldsymbol{m})$ in Eq. (15) are hereinafter also denoted as $\mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l, \boldsymbol{\Gamma})$ and $E^+(\boldsymbol{\Omega}, \boldsymbol{\Gamma}, \boldsymbol{m})$, respectively.

A. Updating C_i and c_j

The partial derivatives of $E^+(\boldsymbol{\Theta}, \boldsymbol{m})$ with respect to a $I \times 2$ matrix $\boldsymbol{C} = [\boldsymbol{C}_1, \cdots, \boldsymbol{C}_I]^T$ and a $J \times 2$ matrix $\boldsymbol{c} = [\boldsymbol{c}_1, \cdots, \boldsymbol{c}_J]^T$ are

$$\frac{\partial E^+}{\partial \boldsymbol{C}} = -\sum_{l=1}^{L}\iint_S [\mathcal{D}_l(\mathbf{x}) - \mathcal{I}_l(\mathbf{x};\boldsymbol{\Omega}_l, \boldsymbol{\Gamma})][\boldsymbol{\Psi}_1(\mathbf{x})\mathbf{x}^T - \boldsymbol{\Psi}_2(\mathbf{x})]\mathrm{d}\mathbf{x}, \tag{29}$$

$$\frac{\partial E^+}{\partial \boldsymbol{c}} = \sum_{l=1}^{L} \iint_S [\mathcal{D}_l(\mathbf{x}) - \mathcal{I}_l(\mathbf{x}; \boldsymbol{\Omega}_l, \boldsymbol{\Gamma})] \omega_l(\mathbf{x}) [\boldsymbol{\psi}_1(\mathbf{x})\mathbf{x}^T - \boldsymbol{\psi}_2(\mathbf{x})] \mathrm{d}\mathbf{x}, \tag{30}$$

where $\boldsymbol{\Psi}_1(\mathbf{x})$ ($\boldsymbol{\psi}_1(\mathbf{x})$) is an I (J) dimensional vector $[\frac{K_1\Phi_1(\mathbf{x})}{R_1^2} \cdots \frac{K_I\Phi_I(\mathbf{x})}{R_I^2}]^T$ ($[\frac{\kappa_1\phi_1(\mathbf{x})}{r_1^2} \cdots \frac{\kappa_J\phi_J(\mathbf{x})}{r_J^2}]^T$), and $\boldsymbol{\Psi}_2(\mathbf{x})$ ($\boldsymbol{\psi}_2(\mathbf{x})$) is a $I \times 2$ ($J \times 2$) matrix $[\frac{K_1\Phi_1(\mathbf{x})}{R_1^2}\boldsymbol{C}_1 \cdots \frac{K_I\Phi_I(\mathbf{x})}{R_I^2}\boldsymbol{C}_I]^T$ ($[\frac{\kappa_1\phi_1(\mathbf{x})}{r_1^2}\boldsymbol{c}_1 \cdots \frac{\kappa_J\phi_J(\mathbf{x})}{r_J^2}\boldsymbol{c}_J]^T$). Note that derivation of a scalar E^+ with respect to a matrix $\boldsymbol{C} = [\boldsymbol{C}_1 \cdots \boldsymbol{C}_I]^T$ is defined as $\frac{\partial E^+}{\partial \boldsymbol{C}} = [\frac{\partial E^+}{\partial \boldsymbol{C}_1} \cdots \frac{\partial E^+}{\partial \boldsymbol{C}_I}]^T$.

Using these partial derivatives, $\{\boldsymbol{C}_i\}_{i=1}^I$ and $\{\boldsymbol{c}_j\}_{j=1}^J$ can be updated using gradient-based local minimization techniques, such as steepest descent and conjugate gradient.

B. Updating R_i and r_j

R_i and r_j are updated using the $\{\boldsymbol{C}_i\}_{i=1}^I$ and $\{\boldsymbol{c}_j\}_{j=1}^J$ values, respectively, at each iteration of the above update process as follows:

$$R_i = \frac{1}{M} \sum_{i' \in \mathcal{M}i} \|\mathbf{C}_{i'} - \mathbf{C}_i\|, \qquad r_j = \frac{1}{N} \sum_{j' \in \mathcal{N}j} \|\mathbf{c}_{j'} - \mathbf{c}_j\|, \tag{31}$$

where $\mathcal{M}_i$ ($\mathcal{N}_j$) is the set of M (N) nearest neighbors of $\{\boldsymbol{C}_i\}_{i=1}^I$ ($\{\boldsymbol{c}_j\}_{j=1}^J$) from $\boldsymbol{C}_i$ ($\boldsymbol{c}_j$).

C. Pruning the Hidden Units

The problem of approximating functions using a minimal RBF network is well studied in the neural network community. Yingwei *et al.* [33] proposed a pruning strategy to remove those hidden units that make little contribution to the output of the network.

1. For every observation (x_n, y_n), compute the outputs of all hidden units $o_k^n (k = 1, \cdots, K)$ and find the largest absolute hidden unit output value $\|o_{\max}^n\|$.
2. Remove the hidden units for which the normalized output $r_k^n = \|\frac{o_k^n}{o_{\max}^n}\|$ is less than a threshold δ for M consecutive observations.

As a postprocessing step of our method, we adopt this pruning strategy, which determines the numbers of hidden units (Λ). Our preliminary experiments, however, showed that the use of $r_k^n = \|\frac{o_k^n}{\sum_{k=1}^K o_k^n}\|$, instead of $r_k^n = \|\frac{o_k^n}{o_{\max}^n}\|$, in the definition of the normalized output in step 2 of Algorithm 1 yielded better results. Thus, based on the definition of $r_k^n = \|\frac{o_k^n}{\sum_{k=1}^K o_k^n}\|$, we select the following pruning criterion:

$$\|\frac{K_i^{(t)}\Phi_i(\mathbf{x})}{\sum_{i=1}^I K_i^{(t)}\Phi(\mathbf{x})}\| \le \delta, \qquad \|\frac{\omega_l(\mathbf{x})\kappa_j^{(t)}\phi_j(\mathbf{x})}{\sum_{j=1}^J \omega_l(\mathbf{x})\kappa_j^{(t)}\phi_j(\mathbf{x})}\| \le \delta'. \tag{32}$$

The iterative algorithm presented in this section can be summarized as follows. Note that, since the process of updating $\boldsymbol{R_i}$ and $\boldsymbol{r_j}$ (Subsection 6.B) does not necessarily minimize Eq. (15), step 2c of Algorithm 2 is computed outside the M-step.

Algorithm 2

Step 1. Initialize:
- Set $\boldsymbol{\Gamma} \Leftarrow \{\{\boldsymbol{C}_i, R_i\}_{i=1}^I, \{\boldsymbol{c}_j, r_j\}_{j=1}^J\}$.
- Set $\boldsymbol{\Omega}^{(0)} \Leftarrow \{\{K_i^{(0)}\}_{i=1}^I, \{\kappa_j^{(0)}\}_{j=1}^J, \{\tilde{L}_l^{(0)}\}_{l=1}^L\}$.
- Set $t \Leftarrow 0$.

Step 2. Repeat the steps 2a through 2c until convergence.
- Step 2a. Set $\tilde{\boldsymbol{\Gamma}} \Leftarrow \boldsymbol{\Gamma}$.
- Step 2b. Repeat the following EM steps until convergence.
  - E-step: set $\boldsymbol{m}^{(t+1)} \Leftarrow \underset{\boldsymbol{m}}{\operatorname{argmin}} E^+(\boldsymbol{\Omega}^{(t)}, \tilde{\boldsymbol{\Gamma}}, \boldsymbol{m})$
  - M-step: set $\boldsymbol{\Omega}^{(t+1)} \Leftarrow \underset{\boldsymbol{\Omega}}{\operatorname{argmin}} E^+(\boldsymbol{\Omega}, \tilde{\boldsymbol{\Gamma}}, \boldsymbol{m}^{(t+1)})$. Set $t \Leftarrow t+1$.
- Step 2c. Update $\boldsymbol{\Gamma}$ (see Subsections 6.A and 6.B).

Step 3. Remove the unnecessary hidden units (see Subsection 6.C).

Step 4. Output:
- Set
- $K_s(\mathbf{x}) \Leftarrow \exp[\sum_{i=1}^I K_i^{(t)}\Phi_i(\mathbf{x})] = \exp[\sum_{i=1}^I K_i^{(t)}\rho(\mathbf{x}; \boldsymbol{C}_i, R_i)]$.
- Set $\kappa(\mathbf{x}) \Leftarrow \sum_{j=1}^J \kappa_j^{(t)}\phi_j(\mathbf{x}) = \sum_{j=1}^J \kappa_j^{(t)}\rho(\mathbf{x}; \boldsymbol{c}_j, r_j)$.
- Set $\{\tilde{L}_l\}_{l=1}^L \Leftarrow \{\tilde{L}_l^{(t)}\}_{l=1}^L$.

7. EXPERIMENTAL RESULTS

We evaluated the effectiveness of the proposed method on a synthetic scene and a real scene. The experimental conditions are summarized in Table 1.

A. Synthetic Scene

We first evaluate the accuracy of the method using a synthetic image of a specular object. The simulated setup allows us to quantitatively evaluate the accuracy of the specular reflectance parameter and illumination estimates. We rendered an image of a 3D model [Fig. 3(a)] with spatially varying inhomogeneous specular reflectance whose parameter values are shown in Figs. 4(a) and 4(b). The illumination consisted of five point light sources. This synthetic image [Fig. 3(b)] was then used as the input to the proposed algorithm. From the input image, we semiautomatically detected a set of locally brightest points by manually drawing a set of their candidate regions (their number is fixed), and then, automatically finding the brightest point in each candidate region, used the viewing direction reflected about the surface normals at those surface points as the point source directions.

We located RBF centers $\boldsymbol{C}_i$ and $\boldsymbol{c}_j$ at regular intervals in the input image and fixed the values of RBF widths, R_i and r_j, at the interval between RBF centers ($\{\boldsymbol{C}_i, R_i\}_{i=1}^I$ and $\{\boldsymbol{c}_j, r_j\}_{j=1}^J$ were equally set). The initial states of $\boldsymbol{c}_j$ and r_j are illustrated in Fig. 3(c). The center and radius of each circle in Fig. 3(c) indicate $\boldsymbol{c}_j$ and one quarter r_j, respectively. The final states of $\boldsymbol{c}_j$ and r_j are shown in Fig. 3(d). Figures 4(a) and

Table 1. Experimental Conditions

Initial values for unknown parameters	$K_1, \cdots, K_I$	0.9
	$\kappa_1, \cdots, \kappa_J$	8.0
	$\tilde{L}_1, \cdots, \tilde{L}_L$	$1/L$
	$\boldsymbol{C}_1, \cdots, \boldsymbol{C}_I$	15 (located every 15 pixels)
	$R_1, \cdots, R_I$	15
	$\boldsymbol{c}_1, \cdots, \boldsymbol{c}_J$	15 (located every 15 pixels)
	$r_1, \cdots, r_J$	15
Values for the constant parameters	μ_K	1.0×10^{-4}
	μ_κ	1.0×10^{-3}
	μ_C	1.0×10^{-2}
	μ_c	1.0×10^{-1}
	M, N	8
	δ	0.005

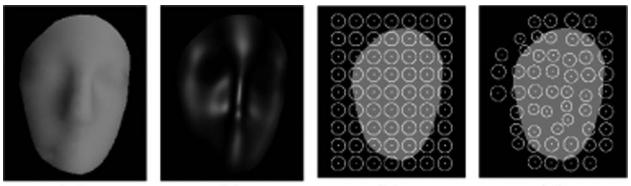

Fig. 3. (a) Surface geometry used to render the synthetic specular image. (b) Synthetic specular image. (c) Initial values of $\boldsymbol{c}_j$ and r_j. (d) Final values of $\boldsymbol{c}_j$ and r_j.

4(b) [respectively, Figs. 4(c) and 4(d)] show the ground truth and estimated $K_s(\mathbf{x})$ (also the ground truth and estimated $\kappa(\mathbf{x})$, respectively) encoded in RGB.

Figure 5 shows the results of relighting: rendering the object under different novel lighting conditions. Specular images were rendered using the estimated parameter values with the spherical Torrance–Sparrow reflection model, which were then composed together with the relit diffuse images. The diffuse images were relit by simply taking the ratio of the incident irradiance between the original and new illumination environment, assuming Lambertian reflection. Figures 5(a) and 5(b) [also Figs. 5(c) and 5(d), respectively] show the synthesized images using the ground truth and estimated values of $K_s(\mathbf{x})$, $\kappa(\mathbf{x})$, and $\{\tilde{L}_l\}_{l=1}^{L}$, respectively. In Figs. 5(a) and 5(b), we turned off three light sources in the original illumination environment of five light sources. In Fig. 5(c) and 5(d), we set point light sources uniformly distributed over all the directions using a geodesic dome.

The ground truth and estimated point source intensities are tabulated in Table 2. The results clearly show that the proposed method successfully estimates the illumination and inhomogeneous specular reflectance properties from a single specular reflection image with high accuracy.

B. Real Scene

We also evaluated the effectiveness of the proposed method on real scenes. We set up a scene consisting of an assortment of fruits and vegetables: a watermelon, an orange pepper, a mango, and a squash [Fig. 6(a)], and a scene of a mask made in Nepal [Fig. 6(e)]. The illumination (three point light sources) and viewpoint conditions of both is the same. We used a digital single-lens reflex (SLR) camera (Canon 30D) to capture the image. In order to obtain a specular-only image, we used a polarization filter [34,35] in front of the light sources and camera lens, whose orientations were manually set to be perpendicular to each other to capture a diffuse-only image and then captured another image where the polarization filters were set to be parallel. Figures 6(a) and 6(e) show the images taken with parallel polarization (diffuse + specular), Figs. 6(b) and 6(f) show the diffuse-only images, and Figs. 6(c) and 6(g) show the specular-only images computed by subtracting the diffuse-only image from the other input image.

We also obtained a 3D geometric model of the objects using a laser range finder (Minolta Vivid 910) whose external

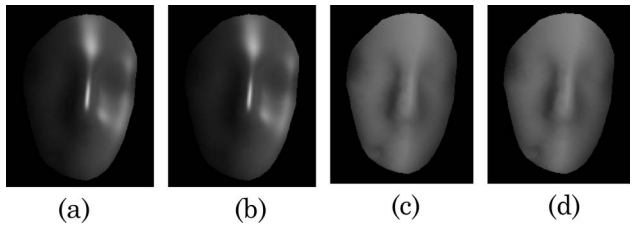

Fig. 5. Synthesized image under novel lighting conditions.

Table 2. Estimated Light Intensities

	Light 1	Light 2	Light 3	Light 4	Light 5
Estimated light intensity	0.1084	0.2010	0.2525	0.2410	0.1971
Ground truth	0.10	0.20	0.25	0.25	0.20

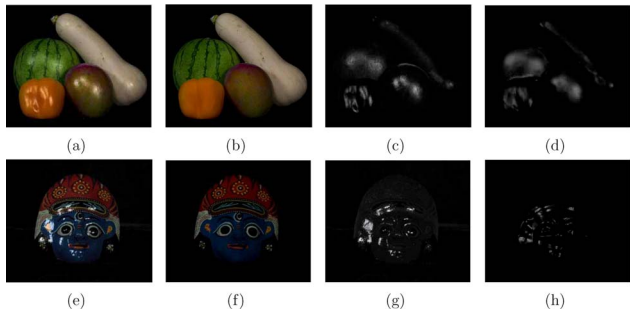

Fig. 6. (Color online) (a), (e) One out the input images. (b), (f) Other input image taken with orthogonal polarization filters to observe diffuse reflection only. (c), (g) Specular image computed by subtracting (b) from (a). (d), (h) Synthesized specular image using the estimated illumination and specular reflectance properties.

Table 3. Estimated Light Intensities

	Light 1	Light 2	Light 3
Estimated light intensity (fruits)	0.2826	0.3357	0.3817
Estimated light intensity (mask)	0.2831	0.3470	0.3699
Ground truth	0.2937	0.3371	0.3692

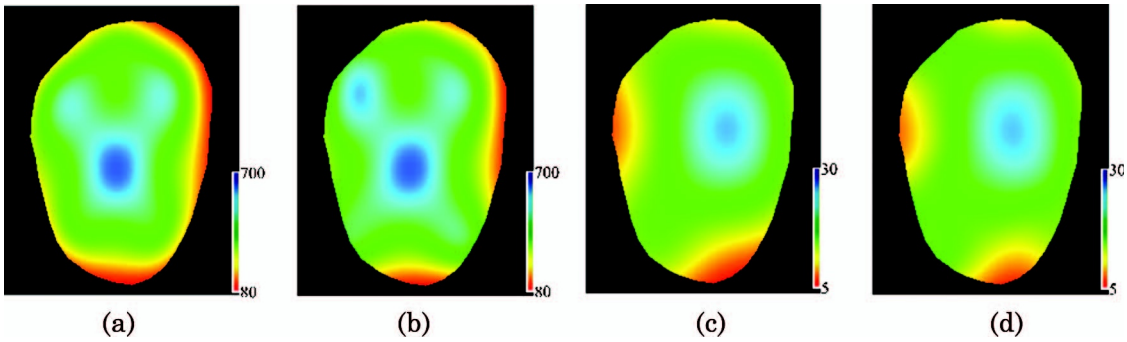

Fig. 4. (a) $K_s(\mathbf{x})$ (ground truth values). (b) $K_s(\mathbf{x})$ (estimated values). (c) $\kappa(\mathbf{x})$ (ground truth values). (d) $\kappa(\mathbf{x})$ (estimated values).

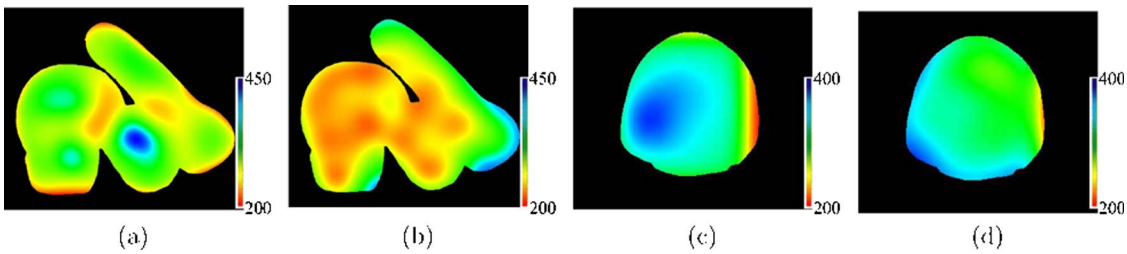

Fig. 7. (Color online) Estimated spatially varying specular reflectance properties: (a), (c) $K_s(\mathbf{x})$, (b), (d) $\kappa(\mathbf{x})$.

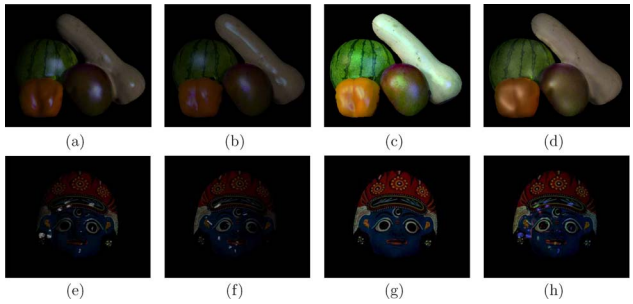

Fig. 8. (Color online) Synthesized images of the scene under novel lighting conditions.

orientation was precalibrated with the digital SLR. As in the previous example, from the specular image, we detected a set of locally brightest points and used them to estimate the point source directions. Figures 6(d) and 6(h) show rendered specular images using the estimated values of $K_s(\mathbf{x})$, $\kappa(\mathbf{x})$, and $\{\tilde{L}_l\}_{l=1}^{L}$. The estimated and ground truth point source intensities are tabulated in Table 3. We measured the ground truth intensities by capturing an image of a black shiny sphere placed in the scene. The synthesized specular image clearly shows that the proposed method successfully estimates the illumination and inhomogeneous specular reflectance properties accurately. Note that the errors at the object boundaries, e.g., between the orange pepper and watermelon, are mainly caused by inaccuracies in the scanned geometry.

Figure 7 shows the estimated values of $K_s(\mathbf{x})$ and $\kappa(\mathbf{x})$ encoded in RGB, respectively. One can see that the proposed method estimates the spatial variation of the specular reflection properties. For instance, in Figs. 7(a) and 7(b), $K_s(\mathbf{x})$ of the mango and $\kappa(\mathbf{x})$ of the squash are relatively large compared to the other three objects, and even within each object the values smoothly vary. Such information can be useful to classify the objects based on material properties. Note that, as shown in Fig. 7, we cannot estimate the specular reflection properties where specular reflection is not observed. If necessary, we may directly extrapolate the estimated values using the RBFs or by assuming uniformity across local regions.

Figure 8 shows the results of relighting. In Figs. 8(a) and 8(e), we turned off two light sources in the original illumination environment, and in Figs. 8(b) and 8(f), we moved that light source to the left. In Figs. 8(c) and 8(g), we changed each color of the three light sources in the original illumination environment to red, green, and blue. Note that, although the original illumination needs to consist of light sources of the same color, this assumption is not required in relighting. In Figs. 8(d) and 8(h), we set light sources uniformly distributed over all the directions using a geodesic dome. The results clearly show that we can achieve photorealistic relighting of the scene with the estimated illumination and specular reflectance properties, which again strongly suggests that the proposed method accurately estimates these variables from a single specular image.

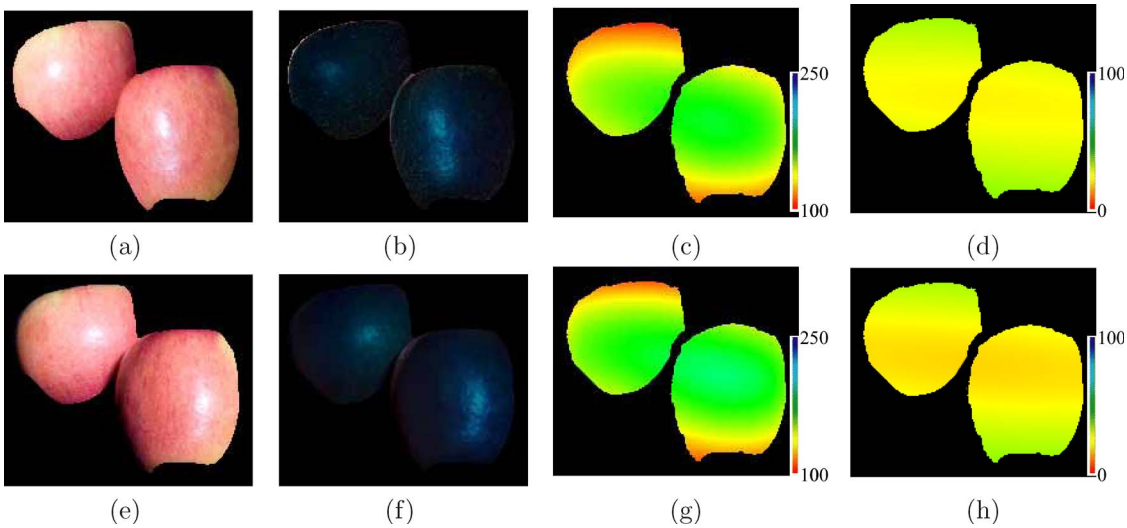

Fig. 9. (Color online) (a) Image captured under a particular lighting condition. (b) Specular image separated from (a). (c) $K_s(\mathbf{x})$ estimated (b). (d) $\kappa(\mathbf{x})$ estimated (b). (e) Image captured under a lighting condition different from (a). (f) Specular image separated from (e). (g) $K_s(\mathbf{x})$ estimated (f). (h) $\kappa(\mathbf{x})$ estimated (f).

Finally, to verify the correctness of the estimated reflectance properties, we compared the estimated values under two different illumination conditions. Figures 9(a) and 9(e) show the captured images under a single light source placed at two different positions, and Figs. 9(b) and 9(f) show the corresponding specular images. Figures 9(c) and 9(d) [also 9(g) and 9(h)] show the estimated $K_s(\mathbf{x})$ and $\kappa(\mathbf{x})$ encoded in RGB from the image shown in Fig. 9(b) [the image shown in Fig. 9(f)]. Since Figs. 9(c) and 9(d) are similar to Figs. 9(g) and 9(h), respectively, one can see that each estimation was successfully accomplished.

8. CONCLUSION

We introduced a novel method for simultaneously estimating the illumination and spatially varying inhomogeneous specular reflectance properties of an object, given a few images taken from the same viewpoint (e.g., in our setup two images) of an object of known geometry. The proposed method represents specular reflection with a directional distribution and the spatial variation of its parameter values with a set of RBFs. We showed that this method enables the formulation of simultaneous estimation as a sound probabilistic inference based on the I-divergence measure, which we solved with an EM-like algorithm. We showed its effectiveness through experiments with synthetic and real images. In future work, we would like to investigate the use of multiple images taken under varying illumination or viewpoints to densely estimate the specular reflectance properties across the entire object surface.

ACKNOWLEDGMENTS

This work was supported in part by the Japanese Ministry of Education, Science, and Culture under grant-in-aid for scientific research 20500155 to K. Hara, the National Science Foundation (NSF) CAREER awards IIS-0746717 and IIS-0964420, and the Office of Naval Research (ONR) grant N00014-11-1-0099 to K. Nishino.

REFERENCES

1. G. Kay and T. Caelli, "Inverting an illumination model from range and intensity maps," CVGIP, Image Underst. **59**, 183–201 (1994).
2. Y. Sato, M. Wheeler, and K. Ikeuchi, "Object shape and reflectance modeling from observation," in *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques* (Association for Computing Machinery, 1997), pp. 379–387.
3. P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (Association for Computing Machinery, 2000), pp. 145–156.
4. S. Marschner and D. Greenberg, "Inverse lighting for photography," In *Proceedings of IS&T/SID Fifth Color Imaging Conference* (The Society for Imaging Science and Technology, 1997), pp. 262–265.
5. K. Hara, K. Nishino, and K. Ikeuchi, "Mixture of spherical distributions for single-view relighting," IEEE Trans. Patt. Anal. Mach. Intell. **30**, 25–35 (2008).
6. K. Nishino, K. Ikeuchi, and Z. Zhang, "Re-rendering from a sparse set of images," Tech. Rep. DU-CS-05-12 (Drexel University, 2005).
7. R. Ramamoorthi and P. Hanrahan, "A signal processing framework for inverse rendering," in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (Association for Computing Machinery, 2001), pp. 117–128.
8. H. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H. Seidel, "Image-based reconstruction of spatially varying materials," in *Proceedings of the 12th Eurographics Workshop on Rendering Techniques*, S. J. Gortler and K. Myszkowski, eds. (Springer, 2001), pp. 104–115.
9. T. Zickler, R. Ramamoorthi, S. Enrique, and P. Belhumeur, "Reflectance sharing: predicting appearance from a sparse set of images of a known shape," IEEE Trans. Patt. Anal. Mach. Intell. **28**, 1287–1302 (2006).
10. D. Goldman, B. Curless, A. Hertzmann, and S. Seitz, "Shape and spatially-varying BRDFs from photometric stereo," IEEE Trans. Patt. Anal. Mach. Intell. **32**, 1060–1071 (2010).
11. I. Csiszár, "Why least squares and maximum entropy? An axiomatic approach to inverse problems," Ann. Stat. **19**, 2032–2066 (1991).
12. A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," J. R. Stat. Soc. Ser. B. Methodol. **39**, 1–38 (1977).
13. S. Nayar, K. Ikeuchi, and T. Kanade, "Determining shape and reflectance of Lambertian, specular, and hybrid surfaces using extended sources," in *Proceedings of the International Workshop on Industrial Applications of Machine Intelligence and Vision* (IEEE, 1989), pp. 169–175.
14. Y. Sato and K. Ikeuchi, "Temporal-color space analysis of reflection," J. Opt. Soc. Am. A **11**, 2990–3002 (1994).
15. A. L. Yuille, D. Snow, R. Epstein, and P. N. Belhumeur, "Determining generative models of objects under varying illumination: shape and albedo from multiple images using SVD and integrability," Int. J. Comput. Vis. **35**, 203–222 (1999).
16. W. Y. Zhao and R. Chellappa, "Symmetric shape-from-shading using self-ratio image," Int. J. Comput. Vis. **45**, 55–75 (2001).
17. Q. Zheng and R. Chellappa, "Estimation of illuminant direction, albedo and shape from shading," IEEE Trans. Patt. Anal. Mach. Intell. **13**, 680–702 (1991).
18. W. A. P. Smith and E. R. Hancock, "Recovering facial shape using a statistical model of surface normal direction," IEEE Trans. Patt. Anal. Mach. Intell. **28**, 1914–1930 (2006).
19. N. Birkbeck, D. Cobzas, P. F. Sturm, and M. Jagersand, "Variational shape and reflectance estimation under changing light and viewpoints," in *Proceedings of the 9th European Conference on Computer Vision* (Springer, 2006), pp. 536–549.
20. N. Alldrin, T. Zickler, and D. Kriegman, "Photometric stereo with non-parametric and spatially-varying reflectance," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2008), pp. 1–8.
21. S. Biswas, G. Aggarwal, and R. Chellappa, "Robust estimation of albedo for illumination-invariant matching and shape recovery," IEEE Trans. Patt. Anal. Mach. Intell. **31**, 884–899 (2009).
22. K. Ikeuchi and K. Sato, "Determining reflectance properties of an object using range and brightness images," IEEE Trans. Patt. Anal. Mach. Intell. **13**, 1139–1153 (1991).
23. I. Sato, Y. Sato, and K. Ikeuchi, "Illumination from shadows," IEEE Trans. Patt. Anal. Mach. Intell. **25**, 290–300 (2003).
24. S. Tominaga and N. Tanaka, "Estimating reflection parameters from a single color image," IEEE Comput. Graph. Appl. **20**, 58–66 (2000).
25. T. Yu, H. Wang, N.Ahuja, and W.-C. Chen, "Sparse lumigraph relighting by illumination and reflectance estimation from multi-view images," in *Rendering Techniques 2006: Eurographics Symposium on Rendering*, T. Akenine-Möller and W. Heidrich, eds. (Eurographics Association, 2006), pp. 41–50.
26. K. E.Torrance and E. M.Sparrow, "Theory for off-specular reflection from roughened surfaces," J. Opt. Soc. Am. **57**, 1105–1112 (1967).
27. R. Fisher, "Dispersion on a sphere," Proc. R. Soc. Lond., Ser. A. **217**, 295–305 (1953).
28. K. Mardia and P. Jupp, *Directional Statistics* (Wiley, 2000).
29. F. Solomon and K. Ikeuchi, "Extracting the shape and roughness of specular lobe objects using four light photometric stereo," IEEE Trans. Patt. Anal. Mach. Intell. **18**, 449–454 (1996).
30. H. Kameoka, T. Nishimoto, and S. Sagayama, "A multipitch analyzer based on harmonic temporal structured clustering,"

IEEE Trans. Audio Speech Lang. Process. **15**, 982–994 (2007).
31. N. Ueda and R. Nakano, "Deterministic annealing EM algorithm," Neural Netw. **11**, 271–282 (1998).
32. M. Lavielle and E. Moulines, "A simulated annealing version of the EM algorithm for non-Gaussian deconvolution," Stat. Comput. **7**, 229–236 (1997).
33. L. Yingwei, N. Sundararajan, and P. Saratchandran, "A sequential learning scheme for function approximation using minimal radial basis function neural networks," Neural Comput. **9**, 461–478 (1997).
34. S. Nayar, X. Fang, and T. Boult, "Removal of specularities using color and polarization," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 1993), pp. 583–590.
35. L. Wolff and T. Boult, "Constraining object features using a polarization reflectance model," IEEE Trans. Patt. Anal. Mach. Intell. **13**, 635–657 (1991).