

Single Image Multimaterial Estimation

Stephen Lombardi Ko Nishino
Department of Computer Science
Drexel University, Philadelphia, PA
{sal64, kon}@drexel.edu

Abstract

Estimating the reflectance and illumination from a single image becomes particularly challenging when the object surface consists of multiple materials. The key difficulty lies in recovering the reflectance from sparse angular samples while correctly assigning them to different materials. We tackle this problem by extracting and fully leveraging reflectance priors. The idea is to strongly constrain the possible solutions so that the recovered reflectance conform with those of real-world materials. We achieve this by modeling the parameter space of a directional statistics BRDF model and by extracting an analytical distribution of the subspace that real-world materials span. This is used, with other priors, in a layered MRF-based formulation that models material regions and their spatially varying reflectance with continuous latent layers. The material regions and their reflectance, and the direction and strength of a single point source are jointly estimated. We demonstrate the effectiveness of the method on real and synthetic images.

1. Introduction

Identifying materials from their images is a long quested ability of computer vision systems to better interact with the real world. Estimating the radiometric properties of materials from their images (i.e., recovering the reflectance of object surfaces) is a fundamental step towards achieving this goal. If successful, the estimated reflectance can give robust cues to other descriptive attributes, such as its tactile and geometric properties (e.g., surface roughness and rigidity), and also greatly narrow down the search space for recognizing materials from their raw image cues.

In this paper, we will focus on the challenging but practical task of estimating the reflectance of object surfaces made of multiple materials from as few as a single image. We assume that the reflectance of the surface can be approximated well with pointwise reflectance, namely the Bidirectional Reflectance Distribution Function (BRDF), and its geometry can be estimated beforehand, in our case with a

laser-stripe range sensor. We also assume that the illumination consists of an unknown single directional light source.

Multimaterial estimation under these assumptions is still a very challenging task mainly for two reasons: material assignment and limited angular samples. Consider estimating the reflectance of an object surface that occupies N pixels in its image. All N combinations between the two extrema, the surface consisting of a single material and the image providing N angular samples of its reflectance on one hand and the surface consisting of N materials and the image providing on average only one angular sample of each material’s reflectance on another hand, are plausible solutions. The main challenge, however, lies in the assignment of the pixels to distinct materials as each k -th combination can have k^N valid solutions each of which corresponds to a different spatial segmentation of the object surface with up to k materials.

Most important, as the number of materials increases, the angular sampling of each material’s reflectance decreases. The fact we have only a single point source compounds this problem as each angular sample would only capture a single combination of incident and exiting directions. As such, even if we could manually give a fairly reliable estimate of the number of materials (e.g., with the aid of color segmentation) segmenting the surface into individual materials while estimating the reflectance of each material poses significant challenges that requires interpretation of sparse and under-sampled angular observations of each reflectance. Surface normals of real object surfaces rarely cover the full frontal half of a sphere as often assumed in other works, which further exacerbates this problem.

We address these challenges by deriving reflectance priors that accurately encode the space of real-world materials and construct a novel probabilistic formulation of multimaterial estimation that fully leverages those reflectance priors. The key idea is to constrain each material’s reflectance to lie within the distribution of real-world materials so that the estimation can reliably find the most “realistic” combination of material and surface segmentation.

We build upon the work of Nishino and Lombardi [8]

to derive statistical priors on the parameter values of the isotropic Directional Statistics BRDF (DSBRDF) model that capture the gamut of real-world materials. This isotropic DSBRDF prior provides concise yet accurate representation of what a “realistic” material reflectance should be by limiting the variation across color channels and reflectance lobes based on functional bases extracted from measured data of real-world materials.

We derive a layered Markov random field formulation of multimaterial estimation, inspired by the work of Sun et al. [12], to fully leverage these reflectance priors. Each material is represented with an MRF in this formulation. The spatial extent of each material is modeled with a continuous latent layer that encodes soft assignments of pixels to that material. The reflectance of each material is then modeled using a set of DSBRDF parameter values. This formulation nicely captures the spatial segmentation of the multiple materials while allowing us to place constraints on the solution. We jointly estimate the material segmentation and each material reflectance, together with the strength and direction of a single point source.

We experimentally evaluate the effectiveness of the method on a number of synthetic and real object surfaces that consists of different numbers of materials. The results clearly demonstrate that the method can achieve accurate segmentation and recovery of the reflectance. They also show that the method can handle real-world materials that exhibit reflectance that cannot be modeled with conventional parametric models. We believe the method yields an important step towards realizing radiometric material estimation in general images.

2. Related Work

Various methods have been proposed for single material estimation from a single image. Ramamoorthi and Hanrahan [9] used spherical harmonics to represent BRDFs so that both the reflectance and illumination can be estimated with frequency decomposition, but with manual specification of the directional light source. Hara et al. [4] derived a spherical Torrance-Sparrow reflection model to jointly estimate multiple point sources and the reflectance parameters through mixture modeling on a unit sphere. Chandraker and Ramamoorthi [2] study the conditions under which a 1D slice of a BRDF is uniquely estimable. By contrast, we estimate full BRDFs by utilizing strong priors to constrain the solution space. These past methods also fundamentally rely on the assumption that the object surface consists of a single material whose reflectance can be modeled with a diffuse plus specular reflection. Our method is not limited to this assumption.

Lensch et al. [5] estimate reflectance and refine shape of multi-material objects from multiple views by clustering pixels into a set of BRDFs using a coarse-to-fine approach.

We derive a novel layered MRF formulation to segment and estimate the reflectance of multi-material objects from a single image. Romeiro and Zickler [11] achieve material estimation under natural lighting but of single material surfaces by representing the reflectance with data-driven bases computed from measured data using a bivariate nonparametric reflectance model [10]. The reflectance prior we introduce is built upon a parametric reflectance model that has comparable accuracy to the bivariate nonparametric model [7] and let’s us obtain tight encoding of the space of real-world materials by modeling its parameter values for measured data.

Zickler et al. [13] “share” angular samples of reflectance across different spatial locations to estimate a nonparametric representation of each material. Hara and Nishino [3] model the spatial variation of just the specular reflection by superimposing a radial basis function network and by estimating the parameters of a spherical Torrance-Sparrow model at each node using variational inference. These methods essentially estimate the materials and their reflectances through interpolation either in the angular domain or spatial domain by trading off each other. Such approaches are suitable for object surfaces that are almost homogeneous but has slight spatial variations like human skin, or object surfaces with materials that smoothly blend into one another. Object surfaces with distinct materials would pose significant challenges as the effective angular samples for each material is strictly bounded by the region the material occupies and at the same time do not continuously blend together. In such cases, the ability to reliably extrapolate possible reflectance of real-world material from possible material segments becomes crucial.

3. Characterizing Real-World BRDFs

We aim to extract strong but analytically tractable priors on the material reflectance. For this, we need a reflectance model that can accurately encode a wide-range of real-world materials which would, at the same time, make explicit the variability of them in the space it spans. For this, we build upon the directional statistics BRDF model first introduced by Nishino [7].

In this paper, we specifically focus on real-world materials each of whose reflectance can be expressed with an isotropic Bidirectional Reflectance Distribution Function (BRDF). We employ the isotropic Directional Statistics BRDF (DSBRDF) model [7] that encodes each slice of a BRDF for a fixed θ_d with a mixture of unnormalized hemispherical exponential power distributions

$$\varrho^{(\lambda)}(\theta_h, \theta_h; \kappa^{(\lambda)}, \gamma^{(\lambda)}) = \sum_r \exp \left[\kappa^{(r, \lambda)}(\theta_d) \cos^{\gamma^{(r, \lambda)}(\theta_d)}(\theta_h) \right] - 1, \quad (1)$$

where (θ_h, ϕ_h) and (θ_d, ϕ_d) are the halfway direction and the light source direction expressed with the “difference” to

the halfway vector, respectively. Each mixture component encodes a single reflectance “lobe” indexed by r , λ denotes the color channel, and κ and γ are functions on the domain $[0, \frac{\pi}{2}]$. In general, κ controls the brightness of the BRDF while γ controls the sharpness of the specular highlights. These two parameters are different for each θ_d slice and thus are functions of θ_d .

With the DSBDRF model, each isotropic BRDF can be encoded with a set of the two parameters for discretely sampled θ_d values $\{(\kappa^{(r,\lambda)}(\theta_d), \gamma^{(r,\lambda)}(\theta_d)) \mid \theta_d = \frac{s\pi}{2S}, s = 0, \dots, S, r = 0, \dots, R, \lambda = R, G, B\}$, where S is the number of BRDF slices and R is the number of lobes used. Nishino and Lombardi [8] showed that these parameter series can be expressed as continuous functions of θ_d using functional bases computed from real-world isotropic BRDFs. This in turn means that we may succinctly characterize the space of real-world reflectance by modeling the distribution of isotropic BRDFs in the space spanned by those functional bases.

Following the analysis by Nishino and Lombardi [8], we first estimate the DSBDRF parameters $\{(\kappa^{(r,\lambda)}(\theta_d), \gamma^{(r,\lambda)}(\theta_d))\}$ of 100 isotropic real-world BRDFs [6] using $R = 3$ lobes and $S = 18$ θ_d slices. We then fit a second degree B-spline with nine knots to the logarithm of the series of parameter values across θ_d variation for each BRDF, which results in 6 variables per curve. Through multivariate functional principal component analysis on all the resulting continuous parameter curves (3 colors per BRDF, 3 reflectance lobes per color, 2 functions— $\kappa(\theta_d)$ and $\gamma(\theta_d)$ —per lobe, and 6 parameters per function for a total of 108 parameters), we obtain a set of basis functions (eigenfunctions) $b_i(\theta_d)$ for $i = 1, \dots, 108$, and mean basis functions $b_\mu(\theta_d)$. Note that the mean function and each basis function are 108-dimensional vectors that contain the B-spline values for all color channels, reflectance lobes, and κ and γ . If we denote the operation of extracting the B-spline coefficients for color channel λ , reflectance lobe r , and function κ with $b(\theta_d; \kappa, r, \lambda)$ (similarly for γ), we can express the DSBDRF parameters κ and γ as a linear combination of these functions

$$\kappa^{(r,\lambda)}(\theta_d) = \exp \left[b_\mu(\theta_d; \kappa, r, \lambda) + \sum_i \psi_i b_i(\theta_d; \kappa, r, \lambda) \right], \quad (2)$$

$$\gamma^{(r,\lambda)}(\theta_d) = \exp \left[b_\mu(\theta_d; \gamma, r, \lambda) + \sum_i \psi_i b_i(\theta_d; \gamma, r, \lambda) \right]. \quad (3)$$

Each coefficient ψ_i for the i -th basis function is a scalar that controls the influence of the m -th eigenfunction to the DSBDRF parameters across all colors and reflectance lobes.

The basis functions span the subspace in which BRDFs of real-world materials lie. As in principal component analysis, the basis functions are ordered by importance (the functional variance of the data) and we may use far fewer basis function than the full dimension and still retain accuracy. As shown by Nishino and Lombardi [8] we found that

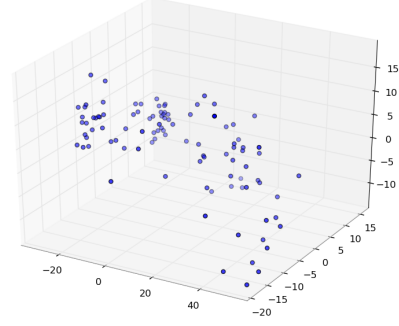


Figure 1: The distribution of 100 real-world isotropic BRDFs [6] in the subspace spanned by the first three basis functions of DSBDRF parameters (Eq. 2 and 3). The distribution is roughly elliptical and can be modeled with a multivariate Gaussian distribution.

about 15 basis functions give an accurate low-dimensional representation. This in turn means that we may now characterize the variability of isotropic real-world BRDFs in this low-dimensional space by modeling their projections, i.e., the coefficients. Each BRDF can now be written with a 15-dimensional vector $\Psi = [\psi_1, \dots, \psi_{15}]$ which we refer to as the DSBDRF coefficients.

Figure 1 shows the distribution of the first three DSBDRF coefficients of the 100 BRDFs from the MERL database [6]. We model this subspace distribution with a multivariate Gaussian $\mathcal{N}(\Psi \mid 0, \Sigma_\Psi)$. It is important to note that the basis functions not only encode the principal modes of variation of the DSBDRF parameters κ and γ independently but also their joint variation for each BRDF, each color channel, and each reflectance lobe. This means that the coefficients also naturally capture these covariances in their values and distribution. As a result, despite its mathematical simplicity, the Gaussian coefficient distribution encodes the variability of real-world reflectance across different color channels and reflectance lobes in its covariance matrix. The model and the priors can accurately encode the characteristics and distribution of a wide range of isotropic BRDFs—well beyond conventional diffuse plus specular models.

4. Estimating A Single Material

The analytical distribution of the DSBDRF parameters expressed in the subspace spanned by FPCA basis functions (i.e., DSBDRF coefficients) provides a faithful representation of the possible reflectance of real-world materials. This lends powerful means to constrain the solution space when estimating the materials in an image. In this section, we demonstrate this by considering the simpler problem of single material estimation.

4.1. Inferring Material Reflectance

We may now formulate single material estimation as estimating the DSBDRF coefficients Ψ from the image. Note that, for single material estimation, the entire object surface shares the same coefficient set Ψ . We achieve this with a Bayesian formulation that lets us fully leverage the distribution of real-world materials derived in Section 3. Given an image $\mathbf{I} = \{\mathbf{I}_{\mathbf{x}} = (I_{\mathbf{x}}^R, I_{\mathbf{x}}^G, I_{\mathbf{x}}^B) | \mathbf{x} = (x, y) \in \Omega\}$, our goal is to find the maximum a posteriori estimate of the DSBDRF coefficients Ψ and the point source $\mathbf{L} = (L_x, L_y, L_z)$ including its direction and intensity

$$p(\Psi, \mathbf{L} | \mathbf{I}) \propto \prod_{\mathbf{x}} p(\mathbf{I}_{\mathbf{x}} | \Psi, \mathbf{L}) p(\Psi).$$

We assume a uniform distribution for the normalization and also for the prior on the light source.

We model the image formation noise with a Gaussian distribution and thus the likelihood as a Gaussian centered on the predicted irradiance (in RGB) $\mathbf{E}_{\mathbf{x}}(\Psi)$ at pixel $\mathbf{x}$ with variance σ^2 :

$$p(\mathbf{I}_{\mathbf{x}} | \Psi, \mathbf{L}) \sim \mathcal{N}(\mathbf{I}_{\mathbf{x}} | \mathbf{E}_{\mathbf{x}}(\Psi, \mathbf{L}), \sigma^2). \quad (4)$$

The image irradiance at each pixel $\mathbf{E}_{\mathbf{x}}$ can be computed using the DSBDRF model (Eq. 1, 2 and 3) using its coefficient estimates Ψ and the light source estimate $\mathbf{L}$

$$E_{\mathbf{x}}(\Psi, \mathbf{L}) = \varrho(\theta_d(\mathbf{x}, \mathbf{L}), \theta_h(\mathbf{x}, \mathbf{L}); \Psi) |\mathbf{L}| \cos \theta_i(\mathbf{x}, \mathbf{L}), \quad (5)$$

where the polar angles θ_h and θ_d , and the incident angle θ_i are computed using the known surface normal at point $\mathbf{x}$ and the light source direction $\frac{\mathbf{L}}{|\mathbf{L}|}$. Note that we are dropping the color channel index for brevity.

As we discussed in Section 3, the prior distribution of the DSBDRF coefficients Ψ is modeled with a Gaussian distribution

$$p(\Psi) \sim \mathcal{N}(\Psi | 0, \Sigma_{\Psi}).$$

We compute the MAP estimate by minimizing the negative log posterior

$$\arg \min_{\Psi, \mathbf{L}} \sum_{\mathbf{x}} \frac{1}{\sigma^2} (\mathbf{I}_{\mathbf{x}} - E_{\mathbf{x}}(\Psi, \mathbf{L}))^2 + \Psi^T \Sigma_{\Psi}^{-1} \Psi. \quad (6)$$

We alternate between minimizing Eq. 6 with respect to each latent variable, Ψ and $\mathbf{L}$, separately. When minimizing with respect to Ψ using the current light source estimate $\mathbf{L}$, our method successively increases the number of coefficients being estimated after each iteration. That is, we begin by estimating only ψ_1 , followed by $\{\psi_1, \psi_2\}$, and so on, up to the complete set Ψ . This allows the estimation to build a solution from the more important basis functions first, which can be considered as a coarse-to-fine estimation

in the reflectance space. This contributes considerably to the reliable estimation of reflectance, and is a strong advantage of the reflectance representation and prior we use.

We minimize with respect to $\mathbf{L}$ by evaluating the error function with the lighting direction set as a node on a unit radius geodesic dome (of 2562 nodes). For each potential point source direction, we evaluate the intensity $|\mathbf{L}|$ as the minimum of the least squares error between the input image to that of the rendered image given the current reflectance estimate Ψ , and find the direction that minimizes Eq. 6. For speed, we only test potential light source directions within a certain angle (e.g., 20°) of the current estimate.

4.2. Surface Normal Refinement

We note that, in practice, “ground truth” surface geometry is often not accurate enough for reflectance estimation. In particular, sharp specular highlight regions require very accurate surface normals to properly predict the outgoing radiance at a given pixel. To remedy this, we use an additional model that makes small adjustments to the surface normals when they make large improvements in the probability of a solution. We do this by modeling the true latent surface normals as an MRF that is observed through both the input image and the observed geometry.

Let $\mathbf{N}$ be a Markov random field that represents the true, latent surface normals over the image pixel grid. We will first make the assumption that the surface normals vary smoothly across the surface by placing a prior distribution on $\mathbf{N}$. We use the von Mises-Fisher distribution for the prior on $\mathbf{N}$,

$$p(\mathbf{N}) \propto \prod_{\mathbf{x}} \prod_{\mathbf{x}' \in \Omega(\mathbf{x})} \exp [\lambda_s \mathbf{N}_{\mathbf{x}}^T \mathbf{N}_{\mathbf{x}'}],$$

where $\Omega(\mathbf{x})$ is the set of neighbors of pixel $\mathbf{x}$ and λ_s is the concentration parameter of the distribution which enforces a smoother surface as its value increases.

We assume Gaussian noise in $\mathbb{S}^2$ for the observed surface normals and model the likelihood with a von Mises-Fisher distribution centered on the latent surface normals

$$p(\hat{\mathbf{N}}_{\mathbf{x}} | \mathbf{N}_{\mathbf{x}}) \propto \exp [\lambda_o \mathbf{N}_{\mathbf{x}}^T \hat{\mathbf{N}}_{\mathbf{x}}],$$

where λ_o is the concentration parameter for the distribution and $\hat{\mathbf{N}}$ are the observed surface normals.

Finally, we modify the image formation likelihood (Eq. 4) to explicitly include the latent surface normals rather than being implicitly computed with the observed surface normals. The entire posterior then becomes

$$p(\Psi, \mathbf{L}, \mathbf{N} | \mathbf{I}, \hat{\mathbf{N}}) \propto \prod_{\mathbf{x}} p(\mathbf{I}_{\mathbf{x}} | \Psi, \mathbf{L}, \mathbf{N}) p(\hat{\mathbf{N}}_{\mathbf{x}} | \mathbf{N}_{\mathbf{x}}) p(\Psi) p(\mathbf{N}).$$

We estimate Ψ and $\mathbf{L}$ as before but now have added the additional step of estimating $\mathbf{N}$ by optimizing the new posterior while keeping the other variables fixed. By tuning

λ_o to a large value, we prevent the surface normals from moving too much while the smoothness constraint resolves ambiguity among good surface normal choices.

5. Estimating Multiple Materials

The reflectance representation and priors give us means to reliably estimate “realistic” reflectance by extrapolating the reflectance from observed angular samples in the subspace of real-world materials. When estimating the reflectance of multiple materials from a single image, we have the added challenge of assigning each pixel to the appropriate material. We tackle this by deriving a novel layered MRF-based formulation that fully leverages the reflectance and other priors to reliably confine the solution space.

5.1. A Layered MRF Formulation

We solve single image multimaterial estimation by fully leveraging additional characteristics of real-world surfaces. We observe that adjacent pixels tend to belong to different materials when their intensities differ greatly. Although adjacent pixels in an image will hardly ever be the exact same color, they will usually differ smoothly due to lighting if they belong to the same material.

We derive a probabilistic formulation to exploit this observation together with the reflectance model and priors. The formulation, inspired by the work of Sun et al. [12], uses a layered graphical model where each k -th layer is responsible for both the spatial extent and surface geometry of the k -th material out of K distinct materials. Each layer models the spatial extent and surface geometry with two Markov random fields (MRFs), $\mathbf{m}_k$ and $\mathbf{N}_k$, respectively. The MRF $\mathbf{m}_k$ specifies the spatial region of the k -th material with continuous latent variables and the MRF $\mathbf{N}_k$ specifies the surface normals at each surface point $\mathbf{N}_{k,\mathbf{x}}$. We denote the collection of these MRFs for all K materials with $\mathbf{m}$ and $\mathbf{N}$.

We say that a pixel $\mathbf{x}$ belongs to the k -th material if and only if $m_{k,\mathbf{x}}(\in \mathbf{m}_k)$ is the maximum value among all other layer values $\{m_{k',\mathbf{x}} \mid k' = 1, \dots, K\}$. For this, we define an indicator random field $\mathbf{s}$ that, for each pixel $\mathbf{x}$, encodes the material of the surface in a 1-of- K coding

$$s_{k,\mathbf{x}} = \mathbf{1}\{m_{k,\mathbf{x}} = \max_{k'} m_{k',\mathbf{x}}\},$$

where $\mathbf{1}$ is an indicator function returning one when the expression inside the braces are true, zero otherwise.

The assumptions that the illumination will cause a material to vary smoothly across an image, whereas changes in material are more likely to happen between pixels with large intensity differences, can be modeled with a spatial prior on each material MRF

$$p(\mathbf{m}) = \prod_k \prod_{\mathbf{x}} \prod_{\mathbf{x}' \in N(\mathbf{x})} \exp \left[-w(\mathbf{x}, \mathbf{x}') (m_{k,\mathbf{x}} - m_{k,\mathbf{x}'})^2 \right],$$

where $N(\mathbf{x})$ denotes the 4-neighbors of pixel $\mathbf{x}$ and the weights reflect the intensity differences

$$w(\mathbf{x}, \mathbf{x}') = \max \left\{ \exp \left[-\frac{1}{2\sigma_c^2} \|\mathbf{I}_{\mathbf{x}} - \mathbf{I}_{\mathbf{x}'}\|^2 \right], \delta \right\},$$

where the difference is the l_2 norm of RGB vectors and the parameters σ_c and δ control the sharpness and penalty of material boundary edges.

Finally, we assume Gaussian noise, so that the likelihood can be expressed with an exponential function

$$p(\mathbf{I} \mid \Psi, \mathbf{m}, \mathbf{L}, \mathbf{N}) = \prod_k \prod_{\mathbf{x}} \exp \left[-s_{k,\mathbf{x}} \left(I_{\mathbf{x}} - \mathbf{E}_{\mathbf{x}}(\Psi_{k,\mathbf{x}}, \mathbf{L}, \mathbf{N}_{k,\mathbf{x}}) \right)^2 \right],$$

where the image irradiance is computed with Eq. 5.

5.2. Inferring Material Reflectance and Region

We estimate the reflectance of each surface point from a single image by computing the maximum a posteriori estimates of the reflectance Ψ_k , the spatial extent $\mathbf{m}_k$, and the surface normals $\mathbf{N}_k$ of each of the k -th material

$$p(\Psi, \mathbf{m}, \mathbf{L}, \mathbf{N} \mid \mathbf{I}, \hat{\mathbf{N}}) \propto p(\mathbf{I} \mid \Psi, \mathbf{m}, \mathbf{L}, \mathbf{N}) p(\Psi \mid \mathbf{m}) p(\mathbf{m}) p(\hat{\mathbf{N}} \mid \mathbf{N}) p(\mathbf{N}). \quad (7)$$

We assume a uniform distribution for the light source prior $p(\mathbf{L})$. In practice, we compute the MAP estimate via energy minimization which corresponds to minimizing the negative log posterior while weighting the prior terms.

We alternate between estimating the light source $\mathbf{L}$ and the material $\Psi, \mathbf{m}$ and $\mathbf{N}$. We employ the same method described in Section 4 to estimate the light source using a geodesic dome. In the material estimation step, we compute the reflectance coefficients for each material Ψ_k given the current material segmentation $\mathbf{m}$ to compute the unary prior $p(\Psi \mid \mathbf{m})$. For each k -th material, this is equivalent to the single material estimation discussed in Section 4. The error term in an energy potential form corresponding to the unary prior $p(\Psi \mid \mathbf{x})$ is also weighted proportionally to the number of pixels assigned to material k

$$E(\Psi \mid \mathbf{m}) = \sum_k \sum_{\mathbf{x}} -s_{k,\mathbf{x}} \log p(\Psi_k)$$

to avoid penalizing small material regions more heavily. Once we compute the reflectance for each material, given the material segmentation from the last iteration, we can estimate the spatial extent $\mathbf{m}_k$ of each material by minimizing the negative log posterior using the L-BFGS algorithm [1].

We compute an initial estimate of the light source direction by thresholding the image to roughly obtain the specular highlights and then by averaging all the reflected directions of the viewing direction mapped on the geodesic

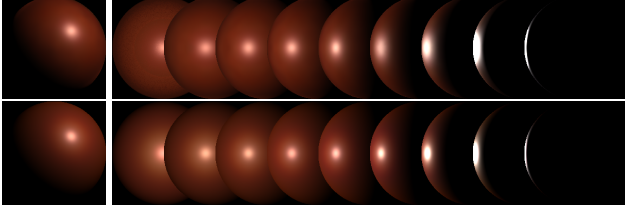


Figure 2: Single material estimation results for the colonial-maple-223 material [6]. Top row: the input image (left) and ground truth renderings of varying incident light source directions (right). Bottom row: synthesized images using estimated reflectance (left) and under varying incident light source directions (right). The results show that the method successfully extrapolates the reflectance from its limited angular samples with good accuracy.

dome using the known surface normals of these highlight pixels. Although this only gives a rough estimate as highlights can never be exactly localized with simple thresholding, we found it to be good enough to start the estimation.

We obtain an initial estimate of the material segmentation $\mathbf{m}$ through k -means clustering on the chromaticity of the single input image. The clustering gives us a set of centroids $\{\mathbf{c}_k | k = 1, \dots, K\}$, where each centroid $\mathbf{c}_k$ is a three-dimensional chromaticity vector $\tilde{\mathbf{I}} = \frac{1}{I^R + I^G + I^B} (I^R, I^G, I^B)$. We set the initial estimates $\mathbf{m}^0$ to be inversely proportional to the distance of each pixel in the chromaticity image to the centroids

$$m_{k,\mathbf{x}}^0 = \exp \left[-(\tilde{\mathbf{I}}_{\mathbf{x}} - \mathbf{c}_k)^2 \right].$$

Finally, we approximate the max operator to compute the segmentation map $\mathbf{s}$ from the material MRFs $\mathbf{m} = \{\mathbf{m}_k | k = 1, \dots, K\}$ with a differentiable function so that the gradient can be computed when optimizing for Ψ and $\mathbf{m}$

$$s_{k,\mathbf{x}} \approx \frac{\exp \gamma m_{k,\mathbf{x}}}{\sum_{k'} \exp \gamma m_{k',\mathbf{x}}}.$$

In this approximation, γ controls its accuracy: the approximation becomes exact as $\gamma \rightarrow \infty$. We, however, keep γ to a moderately large number to avoid arithmetic overflow. It is worth pointing out that $\sum_k s_{k,\mathbf{x}} = 1$, indicating that our selection of γ will never over- or under-penalize the pixel $\mathbf{x}$. A consequence of this is that we may set γ low to allow materials to blend together, which is useful for objects that do not have sharp material boundaries.

6. Experimental Results

We conducted a number of experiments on real and synthetic data to evaluate the effectiveness of our method. For the synthetic data, we rendered spheres with various materials. Using the synthetic data we can evaluate the accuracy of the light source and reflectance estimates. We also captured high dynamic range images of several scenes with objects

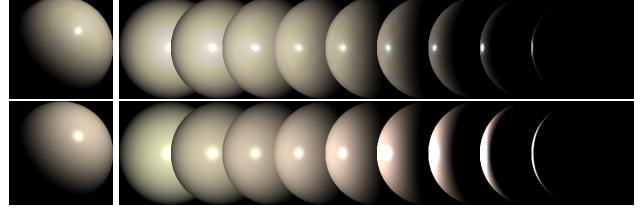


Figure 3: Single material estimation results for the white-acrylic material [6]. The input, ground truth, and results are shown in the same layout as Figure 2. The method recovers the full BRDF from a single input image with reasonable accuracy even for this challenging material.

consisting of multiple materials. For all these real images, we used a laser-stripe range finder (Minolta VIVID 910) to acquire the scene geometry. We also took an image of a small black sphere to accurately capture the ground truth light source direction for each example.

6.1. Single Material Estimation

Figures 2 and 3 show the input image and the reflectance estimates together with the ground truth of single image single material estimation for two materials from rendered input images. The materials used to test the estimation were not used to compute the reflectance priors. The method successfully extrapolates the reflectance from a limited set of angular samples, even for fairly complex materials like the white-acrylic shown in Figure 3. Notice that the behavior of the reflectance as the light source approaches grazing angles is not captured in the input image, but still successfully recovered with high accuracy for the colonial-maple-223 and reasonable accuracy for the white-acrylic. The reflectance behavior of the white-acrylic material is unique and may require a more complex distribution model for the reflectance prior, whereas in this paper we adopt a Gaussian model for computational simplicity. We continued the process of estimating synthetic materials for all 100 BRDFs of the MERL database. We found that the mean and standard deviation of the relative RMS errors of the estimated BRDFs was 0.617 and 0.165, respectively. The error is obviously higher than if the BRDFs were fit with measured data [7] but they nonetheless give a good qualitative result. Our successful recovery of full BRDFs is possible as the reflectance prior strongly constrains the estimate to lie within the distribution of real-world materials in the reflectance space represented with DSBRDF coefficients.

6.2. Multimaterial Estimation

Figure 4 shows the results of multimaterial estimation for several real-world scenes. The results show that the method successfully segments the object surface into distinct materials and accurately estimates the reflectance at each surface point. The painted mask object has a very smoothly varying



Figure 4: Multimaterial estimation results for three real scenes. Each row, from left to right, shows the input image, a synthesized image of the scene using estimated reflectance and light source, the material segmentation result, a relit image of the object, and a ground truth image of the relighting result for each scene. The results demonstrate that the method successfully recovers the reflectance for complex scenes, for instance the gold paint on the mask, except when interreflections and shadowing are prevalent in the scene.

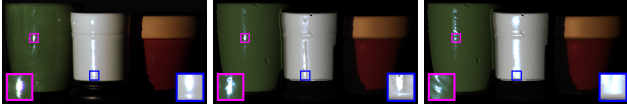


Figure 5: Comparison between normals estimated using the normal refinement technique and the normals measured with a range scanner. The leftmost image is the observed image, the middle image is a synthesized image of the estimated reflectance and segmentation with refined normals, and the right image is the same estimated reflectance and segmentation rendered with the input normals. The input normals are so erroneous that refining them is necessary in this case.

surface; because of this, normal refinement was not necessary. The cup scene, on the other hand, features particularly sharp specular highlights. For this reason, normal refinement was necessary to correctly estimate the reflectance of each material. Figure 5 compares images of the cup scene rendered with the observed normals and with the refined normals.

Some of these materials have reflectance that cannot be captured with conventional diffuse plus specular reflectance models. For instance, the gold paint on the forehead of the mask does not have any diffuse component and its color is actually in its specular reflection. The method successfully recovers the reflectance of such challenging real-world materials from the limited information that can be extracted from a single image, which truly demonstrates the power of the reflectance priors and the probabilistic formulation. For the scenes in figure 4 the estimated light source direction was within 13.5 degrees of the ground truth direction. We

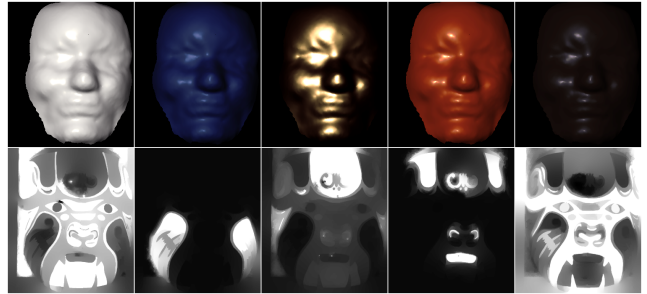


Figure 6: The reflectance and spatial extents of the five materials in the mask scene. Each column represents a different material; the top row shows the geometry of the mask rendered with that material and the bottom row shows the material assignment $m_{k,x}$.

believe these errors were mainly caused by the imperfect surface normal measurements.

Figure 6 gives some insight to how the method operates by showing the geometry of the mask rendered with each material, and the values of material assignment $m_{k,x}$. The method attempts to minimize the difference between adjacent values of $m_{k,x}$ while simultaneously minimizing the difference between the predicted and the observed data. It allows material regions to be discontinuous (jump in the values of $m_{k,x}$) when the difference between adjacent pixel intensities are high, and this effect is visible in the figure.

Our method does not model global illumination effects like interreflection and cast shadows which can contribute greatly to the scene irradiance of more complex objects. This effect is noticeable in the cups scene: the center cup is reflected in the green cup and a heavy shadow cast on the

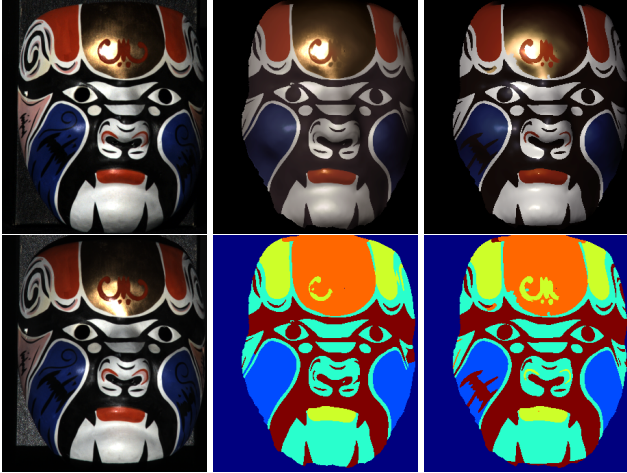


Figure 7: Comparison of using a single input image versus two input images. The left column shows the input images, the middle column shows the results when using a single input image, and the right column shows the results when using the two input images. We examine the results by comparing the predicted images in the top row to the input image in the top-left. Our method can leverage the additional observations to estimate reflectance and segmentation more accurately.

red cup causes that area to be erroneously assigned to a dark material. In the future, we plan to explore methods to predict these global effects by modeling the complete geometry of the scene rather than just the surface normals.

Our method is also capable of using multiple input images, taken from the same camera point of view, to improve the result. We may simply treat the additional image as additional observations of each material and allow the method to make separate light estimates for each input image. Figure 7 compares the results of using multiple input images to using a single image. As evidenced by the improved segmentation of the mask, the additional observations increase the accuracy of the results because of the added information. One of the greatest strengths of our method is its ability to incorporate more information if it becomes available.

We note some limitations of our method as well. Our formulation assumes that a scene is made up of only a handful of materials. As a result, the method is not suited to scenes that feature smoothly varying BRDFs along the scene surface. As we mentioned, it is also unable to handle scenes with significant global illumination effects. Finally, a scene including materials with a limited number of angular samples (e.g., a scene composed of a set of planes) can cause the method to fail to accurately capture the BRDFs. Using multiple input images can alleviate this problem.

7. Conclusion

In this paper, we introduced a novel method for jointly estimating the light source and reflectance of object surfaces

consisting of multiple materials from a single image. We showed how the space of real-world materials can be characterized as statistical priors, together with other priors on the spatial extents of materials and their reflectances’ spatial variations, in a probabilistic formulation. The experimental results demonstrate that this successfully confines the solution space to reliably estimate the reflectance of real-world materials. We believe capitalizing on what we can learn and characterize about the reflectance space of real-world materials is crucial for estimating radiometric properties from images, and plan to further investigate its applications.

Acknowledgments This work was supported in part by the Office of Naval Research grant N00014-11-1-0099, and the National Science Foundation (NSF) awards IIS-0746717 and IIS-0964420.

References

- [1] R. Byrd, P. Lu, J. Nocedal, and C. Zhu. A Limited Memory Algorithm for Bound Constrained Optimization. *SIAM JSC*, 16:1190–1208, Sept. 1995. 5
- [2] M. Chandraker and R. Ramamoorthi. What an image reveals about material reflectance. In *IEEE ICCV*, pages 1–8, Nov. 2011. 2
- [3] K. Hara and K. Nishino. Variational Estimation of Inhomogeneous Specular Reflectance and Illumination from a Single View. *JOSA A*, 28(2):136–146, Feb. 2011. 2
- [4] K. Hara, K. Nishino, and K. Ikeuchi. Mixture of Spherical Distributions for Single-View Relighting. *IEEE TPAMI*, 30(1):25–35, Jan. 2008. 2
- [5] H. P. A. Lensch, J. Kautz, M. Goesele, W. Heidrich, and H.-P. Seidel. Image-based Reconstruction of Spatial Appearance and Geometric Detail. *ACM ToG*, 22(2):234–257, Apr. 2003. 2
- [6] W. Matusik, H. Pfister, M. Brand, and L. McMillan. A Data-Driven Reflectance Model. *ACM ToG*, 22(3):759–769, July 2003. 3, 6
- [7] K. Nishino. Directional Statistics BRDF Model. In *IEEE ICCV*, pages 476–483, Sept. 2009. 2, 6
- [8] K. Nishino and S. Lombardi. Directional Statistics-based Reflectance Model for Isotropic Bidirectional Reflectance Distribution Functions. *JOSA A*, 28(1):8–18, Jan. 2011. 1, 3
- [9] R. Ramamoorthi and P. Hanrahan. A Signal-Processing Framework for Reflection. *ACM ToG*, 23(4):1004–1042, Oct. 2004. 2
- [10] F. Romeiro, Y. Vasilyev, and T. Zickler. Passive Reflectometry. In *ECCV*, pages 859–872, 2008. 2
- [11] F. Romeiro and T. Zickler. Blind Reflectometry. In *ECCV*, pages 45–58, Berlin, Heidelberg, 2010. 2
- [12] D. Sun, E. Sudderth, and M. Black. Layered Image Motion with Explicit Occlusions, Temporal Consistency, and Depth Ordering. *NIPS*, Dec. 2010. 2, 5
- [13] T. Zickler, R. Ramamoorthi, S. Enrique, and P. N. Belhumeur. Reflectance Sharing: Predicting Appearance from A Sparse Set of Images of a Known Shape. *IEEE TPAMI*, 28(8):1287–1302, Aug. 2006. 2