Blind Reflectometry

Fabiano Romeiro and Todd Zickler

Harvard University
33 Oxford St., Cambridge, MA, USA, 02138
{romeiro,zickler}@seas.harvard.edu

Abstract. We address the problem of inferring homogeneous reflectance (BRDF) from a single image of a known shape in an unknown real-world lighting environment. With appropriate representations of lighting and reflectance, the image provides bilinear constraints on the two signals, and our task is to blindly isolate the latter. We achieve this by leveraging the statistics of real-world illumination and estimating the reflectance that is most likely under a distribution of probable illumination environments. Experimental results with a variety of real and synthetic images suggest that useable reflectance information can be inferred in many cases, and that these estimates are stable under changes in lighting.

1 Introduction

The optical properties of a material often provide a clue for how it will behave when acted upon. They help inform us, for example, if the material is hard, soft, hot, cold, rigid, pliable, brittle, heavy, or lightweight. It makes sense, then, that people can infer materials' optical properties from their images; and building similar functionality into computer vision systems seems worthwhile.

The optical properties of many materials are adequately summarized by the bidirectional reflectance distribution function (BRDF), which describes how flux at a surface patch is absorbed and reflected over the output hemisphere. The BRDF provides a complete description of lightness, gloss, sheen, and so on; and in this paper, we explore when and how it can be recovered from an image. This task is made difficult by the fact that reflectance is confounded with shape, lighting, and viewpoint, all of which may be unknown. Even when the shape and relative viewpoint are provided (say, by contours, shadows, or other cues), the blind separation of BRDF from lighting is something that computer vision systems cannot yet do well.

This paper considers the following problem, depicted in Fig. 1. We are given a single high dynamic range (HDR) image of a known shape under an unknown, real-world lighting environment, and our task is to infer the material's BRDF. Our approach is to compute the BRDF that is most likely under a distribution of probable lighting environments—a strategy that is motivated by previous successes for other ill-posed vision problems, including color constancy and blind image deblurring. In our case, we show that by choosing an appropriate BRDF representation, we can leverage the statistics of real-world lighting to accurately infer materials' optical properties in a variety of lighting environments.

K. Daniilidis, P. Maragos, N. Paragios (Eds.): ECCV 2010, Part I, LNCS 6311, pp. 45–58, 2010.

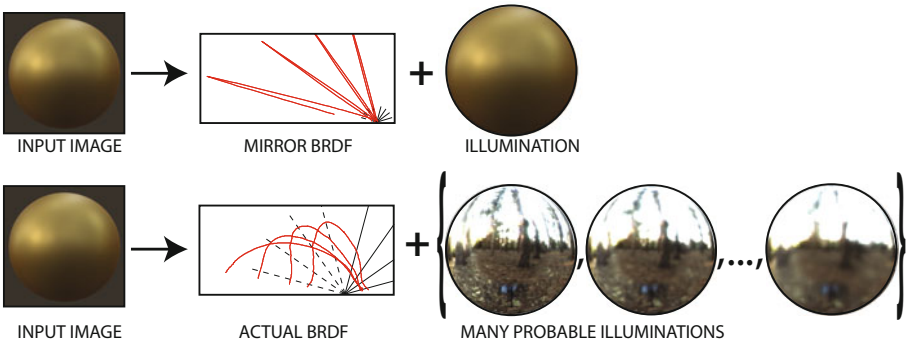

Fig. 1. Our goal is estimating the BRDF from an image of a known shape in unknown real-world lighting. *Top*: The trivial solution is a mirror-like BRDF, which exactly predicts the input for a carefully-crafted "blurry" environment. *Bottom*: To avoid this, we choose a BRDF that predicts the input for a distribution of probable environments.

2 Background and Related Work

People are quite adept at inferring reflectance information from images, and there have been a number of psychophysical studies that explore the underlying mechanisms [1,22,7,33,31,35]. Results suggest that people do not require contextual knowledge of the environment to infer reflectance [7], but that performance decreases when the directional statistics of the environment deviate significantly from those found in nature [7,4,5]. These findings provide motivation for our work, which also leverages the directional statistics of natural environments.

When it comes to computational approaches for recovering reflectance from images, most have been developed for controlled or known lighting (e.g., [30,14,9]). Fewer methods have been designed for cases where the lighting is not known, and of these, most assume reflectance to be well-represented by a pre-chosen "parametric" BRDF model, such as the Phong, Ward, or Lafortune models (e.g., [21,36,11]). Parametric BRDF models place considerable restrictions on reflectance, and as a result, they allow inferring quite a bit about a scene. For example, the method of Georghiades [8] can simultaneously infer everything—shape, lighting and reflectance—provided that the material is well-represented by a simplified Torrance-Sparrow model (and that lighting consists of a moving point light source). While parametric models continue to improve (e.g., [20]), their use typically has two significant limitations. First, it severely restricts the space of materials (see [19,32]); and second, because these models are non-linear in their parameters, the required computation ends up being model-specific and cannot easily be transferred from one material class to another.

An attractive alternative to parametric BRDF models is using a linear combination of reflectance basis functions. This way, the representation can be grown to include the entire world of BRDFs, at least in theory. Moreover, when object shape is known, it leads to a simple bilinear relationship between the unknown reflectance parameters (i.e., the coefficients in the basis) and the lighting

parameters. This bilinearity has already been exploited in both vision [13,10] and graphics (e.g., [25,18]), and it is the key to making our approach tractable.

The choice of bases for reflectance and lighting are important, and we discuss them in detail in subsequent sections. But once these choices are made, we obtain a bilinear inferrence problem that resembles others in vision: Given an image of a known shape, we must find probable lighting and reflectance parameters that could have created it. Color constancy and blind image deblurring can be formulated analogously [2,6,12], and our work leverages insight gained from their analyses. Specifically, instead of simultaneously estimating the BRDF and environment that best explain a given image, we obtain better results by estimating the BRDF that is most likely under a *distribution* of lighting environments (Fig. 1). This process is termed "MAP$_k$ estimation" in the context of blind deblurring [12], and the same basic idea forms the core of our approach.

A natural comparison for our approach is the framework of Ramamoorthi and Hanrahan [24,26], which uses spherical harmonics to represent lighting and reflectance and expresses their interaction as a convolution. Since spherical harmonics are eigenfunctions of the convolution operator, this leads to elegant closed-form expressions for the lighting and reflectance coefficients. But this representation cannot easily incorporate a meaningful prior probability distribution for natural lighting environments (see [5]), and it either requires that the entire 4D light field is available as input or that the BRDF can be restricted to being a "radially-symmetric" function over a one-dimensional domain.

Another natural comparison is the method of Haber et al. [10], which also represents lighting and reflectance using linear bases. Their approach differs in terms of its input and output (multiple images instead of one; spatially-varying BRDFs instead of uniform) and has two technical distinctions. It does not explicitly model the statistics of natural lighting, and it jointly estimates lighting and reflectance instead of marginalizing over a distribution of environments.

3 Approach

We assume all sources and reflecting surfaces in the environment to be far from the object in question so that the angular distribution of incident lighting does not vary over the object's surface. This allows the unknown lighting L to be represented as an "environment map"—a positive-valued function on the sphere of directions; $L\colon \mathbb{S}^2 \to \mathbb{R}^+$. We also assume that the camera and object geometry are known, that mutual illumination is negligible, and that the unknown BRDF (F) is isotropic. Then, a linear measurement made at pixel i can be written

$$I_i = \int_\Omega L_i(\omega) V_i(\omega) F_i(\omega)(n_i \cdot \omega) d\omega, \tag{1}$$

where n_i is the surface normal at the back-projection of pixel i, $L_i(\omega)$ is the hemisphere of the unknown lighting centered at direction n_i, and $V_i(\omega)$ is a binary-valued hemispherical "visibility" function that encodes the object's self-shadowing at the back-projection of i (e.g., [18,10]). Finally, F_i is a 2D slice of the unknown BRDF determined by the normal n_i and the local view direction.

Since everything in Eq. 1 is known except the lighting and BRDF, an image $I = \{I_i\}$ imposes a set of constraints upon them. One approach to estimating the BRDF, then, is to define prior probability distributions for the unknown lighting $p(L)$ and BRDF $p(F)$ and find the functions that maximize the posterior

$$p(L, F|I) \propto p(I|L, F)p(L)p(F), \tag{2}$$

using a likelihood $p(I|L, F)$ based on Eq. 1. This is closely related to the approach of Haber et al. [10], and it suffers from a preference for the trivial mirror-like solution. Any image can be perfectly explained by a mirror-like BRDF and a carefully crafted "blurry" environment that exactly matches the image [7], so the likelihood (and usually the posterior) are maximal for these functions.

In this paper we avoid this problem in the following manner. Instead of selecting the single BRDF/lighting pair that best explain an input image, we select the BRDF that is most likely under a *distribution* of lighting environments. We do this by computing the mean of the marginalized posterior:

$$F_{opt} \triangleq \int Fp(F|I)dF = \int F\left(\int p(F, L|I)dL\right)dF. \tag{3}$$

The intuition here—adapted directly from the related problem of blind image deblurring [6,12]—is that instead of selecting a BRDF that perfectly explains the image for a *single* lighting environment (the trivial solution), we select one that reasonably explains the image for many probable lighting environments.

Evaluating the expression on the right of Eq. 3 requires prohibitive computation, and to make it feasible, we employ a variational Bayesian technique. Following [16,17] we approximate the posterior using a separable function,

$$p(L, F|I) \approx q(L, F) = q(L)q(F), \tag{4}$$

with components having convenient parametric forms. Given an input image, we compute the parameters of this approximate posterior using fixed point iteration, and then we trivially approximate the solution (Eq. 3) as the mean of $q(F)$.

Pursuing this approach requires suitable representations for lighting and reflectance. In particular, we require each to be a linear combination of basis functions, and we require the prior probability distributions of their coefficients to be well-approximated by exponential forms. We describe our choices next.

3.1 Representing Illumination

We represent spherical lighting using a wavelet basis. As depicted in Fig. 2 and following [34], we do this by mapping the sphere to a plane with an octahedral map [23] and using a Haar wavelet basis in this plane. Notationally, we write $L = \sum_{m=1}^{M} \ell_m \psi_m$ with ψ_m the basis functions and ℓ_m the corresponding coefficients. This choice of basis is motivated by the fact that statistics of band-pass filter coefficients of real-world lighting display significant regularity. Much like real-world images, the distributions of these coefficients are highly kurtotic, having

heavy tails [4,5]. Our choice is also motivated by the apparent utility of the related image statistics for tasks like compression, denoising, and deblurring.

To develop prior distributions for the coefficients $\ell \triangleq \{\ell_m\}$, we collected 72 environments (nine from the ICT Graphics Lab[1] and the remainder from the sIBL Archive[2]), normalized each so that it integrates to one, and studied the coefficient distributions at different scales. Like Dror et al. [4,5], we found these statistics to be notably non-stationary, especially at coarser scales. Figure 2 shows empirical distributions and parametric fits for a variety of scales using a 32×32 discretization of the sphere. At the finest scales (scales 4 and 5), the distributions are quite stationary, and we employ a single zero-mean Gaussian mixture for all of the coefficients at each scale (with 4 and 5 components, respectively). At the middle scales (scales 2 and 3), the statistics change significantly depending on elevation angle and basis type (vertical, diagonal, horizontal), and accordingly we use distinct distributions for each basis type both above and below the horizon. Each distribution is a zero-mean Gaussian mixture, and we use three components for groups in scale 3 and two components for groups in scale 2. Finally, at the coarsest scale (scale 1) we use zero-mean, two-component Gaussian mixtures for the diagonal and horizontal basis types, and a Gaussian rectified at a negative value for the vertical basis type to capture the fact that lighting is dominant from above. Note that the DC value ℓ_1 is the same in all cases since the illuminations are normalized. Additional details are in [28].

With these definitions we can write our illumination prior as

$$p(\ell) = \prod_{m=2}^{M} \sum_{n=1}^{N_m} \pi_{nm}\ p_{nm}(\ell_m), \tag{5}$$

with N_m the number of mixture components for coefficient m and π_{nm} the mixing weights. The group structure described above is implicit in this notation: All coefficients in any one group share the same N_m, π_{nm} and p_{nm}.

3.2 Representing Reflectance

We represent BRDFs as a linear combination of non-negative basis functions learned through non-negative matrix factorization (NMF) of all 100 materials in the MERL/MIT database [15]. This produces the linear representation $F = \sum_{k=1}^{K} f_k \phi_k$ and has the advantage of allowing non-negativity constraints on the recovered BRDF F to be naturally enforced through non-negativity constraints on the coefficients $f \triangleq \{f_k\}$. Also, we find that the empirical distributions of the resulting coefficients f_k can be well-approximated by exponentials (see Fig. 3, right), making them well-suited for inference using variational Bayes.

Each BRDF in the database is represented using a $90 \times 90 \times 180$ discretization of the 3D isotropic BRDF domain, parameterized in terms of the half-vector and difference-vector [29]. For computational convenience, we reduce each material to

[1] `http://www.debevec.org/probes`; `http://gl.ict.usc.edu/Data/HighResProbes`
[2] `http://www.hdrlabs.com/sibl/archive.html`

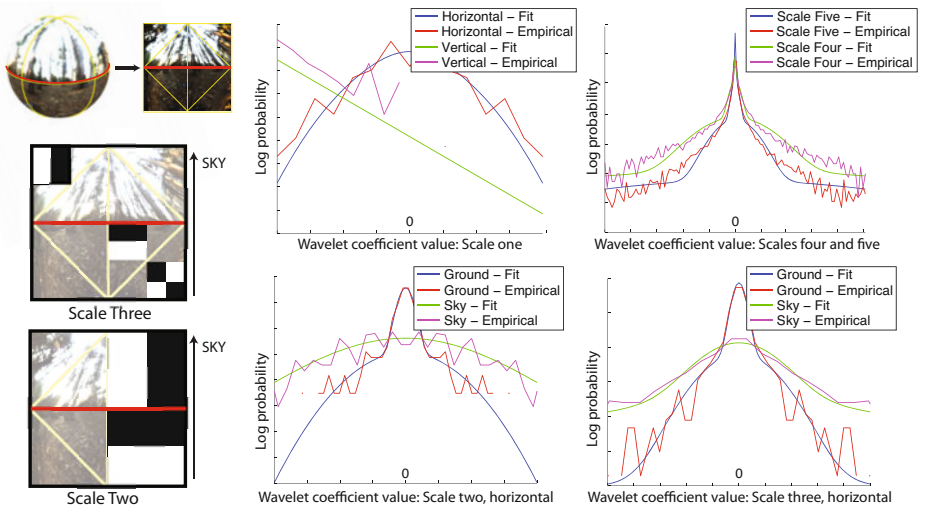

Fig. 2. *Left*: We represent lighting using a Haar wavelet basis on the octahedral domain [23] discretized to 32×32. Statistics of wavelet coefficients are non-stationary, so we fit distinct distributions for coefficients above and below the horizon. *Right*: Empirical distributions and their parametric fits for a variety of wavelet coefficient groups.

the 90×90 bivariate domain of Romeiro et al. [27] and scale it to have a maximum value of one before computing the NMF. The bivariate reduction allows handling the entire database (and potentially much more) without resulting to out-of-core methods, and as shown in [27], it has a limited effect on accuracy. The resulting basis functions are defined on the two dimensional domain $(\theta_h, \theta_d) \in [0, \pi/2) \times [0, \pi/2)$, where θ_h and θ_d are the halfway angle and difference angle, respectively (see [29]). We can visualize the basis functions as images, and two of them are in the right of Fig. 3. In this visualization, specular reflection is at small halfway angles (top edge), grazing effects are at large difference angles (top-right corner) and retro-reflection is at small difference angles (left edge).

The left of Fig. 3 qualitatively evaluates the NMF model's fit to the original BRDF data for different numbers (K) of basis functions. We also compare to a parametric BRDF model (Cook-Torrance) as fit by Ngan et al. [19]. While it remains perceptually distinguishable from ground truth for some materials, we find that the NMF model's fit with ten basis functions ($K = 10$) provides a good balance between complexity and accuracy. It provides fits that are comparable to the Cook-Torrance model, but it is linear, which is important to our approach.

We can also evaluate the fit quantitatively by computing RMS error in the BRDF domain. According to this metric, the NMF approach significantly outperforms the parametric model. The mean and median RMS error computed using the green channels of all materials are 1.58 and 0.45 for the NMF model, and 46.91 and 17.11 for the Cook-Torrance fit. Some of this significant difference is due to the fact that we have chosen to perform NMF based on L_2 cost

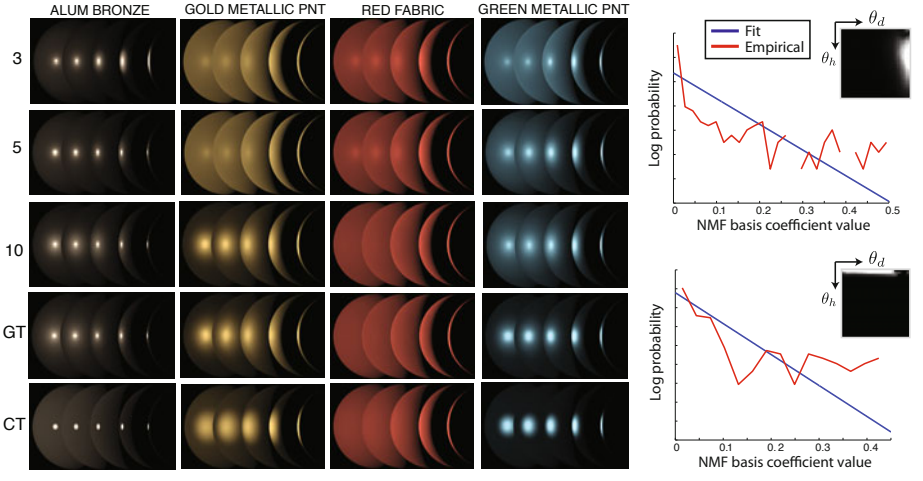

Fig. 3. *Left*: Qualitative evaluation of the NMF BRDF model. Top to bottom: NMF model with 3, 5, and 10 basis functions; ground truth; and Cook-Torrance fit from [19]. *Right*: Empirical distributions and parametric fits for NMF coefficients corresponding to basis elements that roughly account for grazing (top) and specular (bottom) effects.

in the BRDF domain, whereas the parametric Cook-Torrance fit is performed with an approximate perceptual metric [19]. In fact, one can view the metric used in NMF as a choice that can be tuned for each application. If one ultimately seeks to infer BRDFs for the purposes of material recognition, then the L_2 cost used here may be preferred. If the inferred BRDF is to be used for image synthesis, however, it may be more desirable to use a perceptual metric (e.g., [35]) within a kernel-NMF framework (e.g., [3]).

Having computed basis functions ϕ_k and the parameters λ_k of the coefficient distributions, we obtain the following prior distribution for reflectance:

$$p(f) = \prod_{k=1}^{K} \lambda_k \exp(-\lambda_k f_k), \qquad f_k \geq 0. \tag{6}$$

3.3 A Bilinear Likelihood

Having defined linear representations of the lighting and reflectance, we can write an expression for the likelihood of their coefficients given a particular image. We begin by updating the imaging model to include a camera exposure parameter (γ) and a crude model for noise:

$$I_i = \gamma \int_{\Omega} L_i(\omega) V_i(\omega) F_i(\omega)(n_i \cdot \omega) d\omega + \epsilon, \tag{7}$$

with $\epsilon \sim N(0, \sigma^2)$. The exposure parameter compensates for the difference between the absolute scale of the intensity measurements and the combined scale of

the illumination and reflectance functions. This is important because the prior distributions for lighting and reflectance are estimated from normalized data while the intensity measurements may be at an arbitrary scale.

Substituting $L = \sum \ell_m \psi_m$ and $F = \sum f_k \phi_k$ into this expression, one can re-write this as (see details in [28]):

$$I_i = \gamma \ell^T M_i f + \epsilon, \tag{8}$$

where the per-pixel matrices M_i are determined by the shape (V_i, n_i), view direction, and the lighting and reflectance basis functions $\{\psi_m\}$ and $\{\phi_k\}$. For an input image of a known shape, these matrices can be pre-computed, and we assume them to be constant and known.

By treating the pixels of an input image as independent samples, this measurement model leads directly to our desired expression for the likelihood of a set of model parameters given image I:

$$p(I|\ell, f, \sigma, \gamma) = \prod_{i=1}^{N} \frac{\sigma^{-1}}{\sqrt{2\pi}} \exp\left(-\frac{\sigma^{-2}}{2}(I_i - \gamma \ell^T M_i f)^2\right). \tag{9}$$

We treat the exposure and noise variance (γ, σ^2) as model parameters to be estimated along with illumination and reflectance, and for these we define prior distributions $p(\sigma^{-2}) \sim \Gamma(a, b)$ and $p(\gamma) \sim Exp(\lambda_\gamma)$.

3.4 Inferrence

The definitions of the previous sections (Eqs. 5, 6, 9, and the noise and exposure priors) provide everything we need to write the posterior

$$p(\ell, f, \gamma, \sigma^{-2}|I) \propto p(I|\ell, f, \gamma, \sigma^{-2}) p(\ell) p(f) p(\gamma) p(\sigma^{-2}). \tag{10}$$

As described in Sect. 3, we wish to marginalize over lighting (as well as noise, and exposure) and compute the mean of the marginalized posterior. Following Miskin and MacKay [16,17], we do this by approximating the posterior with a separable function, $p(\theta|I) \approx q(\theta) = q(\ell)q(f)q(\sigma^{-2})q(\gamma)$, with $\theta \triangleq (\ell, f, \sigma^{-2}, \gamma)$. The function $q(\theta)$ is computed by minimizing a cost based on the Kullback-Leibler divergence between it and the posterior [16,17]:

$$\int q(\theta) \left(\log \frac{q(\ell)}{p(\ell)} + \log \frac{q(f)}{p(f)} + \log \frac{q(\sigma^{-2})}{p(\sigma^{-2})} + \log \frac{q(\gamma)}{p(\gamma)} - \log p(I|\theta)\right) d\theta \tag{11}$$

We provide an overview of the optimization here, and details can be found in [28]. The basic idea is to use coordinate descent, with each distribution $q(\cdot)$ being updated using the current estimates of the others. The update equations are derived by integrating all of the terms but one in Eq. 11 (the one containing $q(f)$, say), taking the derivative with respect to the remaining distribution ($q(f)$

in this example) and equating the result to zero. In our case, this procedure reveals that the approximating distributions $q(\cdot)$ are of the following forms

$$q(f) = \prod q_k(f_k), \text{ with } q_k \sim N_R(u_k, w_k), \tag{12}$$

$$q(\ell) = \prod q_m(\ell_m), \text{ with } q_m \sim \begin{cases} N(u_m, w_m) & \text{if } m \neq 3 \\ N_{RC}(u_m, w_m, T) & \text{otherwise,} \end{cases} \tag{13}$$

$$q(\gamma) \sim N_R(u_\gamma; w_\gamma) \text{ and } q(\sigma^{-2}) \sim \Gamma(\sigma^{-2}; a_p, b_p), \tag{14}$$

where N_R is a Gaussian distribution rectified at 0, and N_{RC} is a Gaussian distribution rectified at T. The same procedure also provides closed form expressions for the updated parameters of each distribution $q(\cdot)$ in terms of the current parameters of the others (see [28]). One strategy, then, is to cycle through these distributions, updating each in turn; but as described in [16,17], convergence can be accelerated by updating all parameters in parallel and then performing a line search between the current and updated parameter-sets.

Specifically, we define intermediate variables that are sufficient to determine all of the distribution parameters: $\Phi = (\Phi_1, \Phi_2, \Phi_3, \Phi_4, \Phi_5, \Phi_6, \Phi_7)$ where Φ_1 and Φ_2 are K-vectors such that $\Phi_1(k) = \frac{u_k}{w_k}$ and $\Phi_2(k) = \log(\frac{1}{w_k})$; Φ_3 and Φ_4 are M-vectors such that $\Phi_3(m) = \frac{u_m}{w_m}$ and $\Phi_4(m) = \log(\frac{1}{w_m})$; $\Phi_5 = \log(\frac{b_p}{a_p})$; $\Phi_6 = \frac{u_\gamma}{w_\gamma}$; and $\Phi_7 = \log(\frac{1}{w_\gamma})$. These intermediate variables are iteratively updated according to Algorithm 1, and once they converge, they determine the distribution $q(f)$, whose mean is the BRDF we seek. The noise variable Φ_5 is not updated at every iteration, but only when the other variables have converged at the current noise level. This is a strategy borrowed from Miskin's implementation[3].

We initialize the algorithm with the posterior means $\{u_k^{(0)}\}$ and $\{u_m^{(0)}\}$ corresponding to a random BRDF and lighting environment, respectively. The initial posterior variances $\{w_m^{(0)}\}$ and $\{w_k^{(0)}\}$ are set to relatively large values (10^{-1}) to account for the uncertainty in our initial estimates. Exposure parameters u_γ and w_γ are initialized to 1 and 10 respectively. Finally, parameter $\frac{b_p}{a_p}$ is initialized to 1 so that we have a broad initial posterior on the inverse noise variance.

4 Evaluation and Results

We begin our evaluation using images synthesized[4] with the MERL/MIT BRDF data and our collection of measured illumination environments. Using these tools, we can render HDR images for input to our algorithm as well as images with the recovered BRDFs for comparison to ground truth.

There is a scale ambiguity for each image because we can always increase the overall brightness of the illumination by making a corresponding decrease in the BRDF. Accordingly, we only seek to estimate the BRDF up to scale. We

[3] `http://www.inference.phy.cam.ac.uk/jwm1003/train_ensemble.tar.gz`.
[4] PBRT: `http://www.pbrt.org/`

Algorithm 1. Fit ensemble of approximating distributions

$\phi_1^{(0)}(k) \leftarrow \frac{u_k^{(0)}}{w_k^{(0)}}, \phi_2^{(0)} \leftarrow \log\left(\frac{1}{w_k(0)}\right), \phi_3^{(0)}(m) \leftarrow \frac{u_m^{(0)}}{w_m^{(0)}}, \phi_4^{(0)} \leftarrow \log\left(\frac{1}{w_m(0)}\right)$

$\phi_5^{(0)} \leftarrow \log(1), \phi_6^{(0)} \leftarrow \frac{u_\gamma^{(0)}}{w_\gamma^{(0)}}, \phi_7^{(0)} \leftarrow \log\left(\frac{1}{w_\gamma(0)}\right), i = 0$

repeat

  repeat

    $\boldsymbol{\Phi}^* = \text{Update}(\boldsymbol{\Phi}^{(i)})$ (see [28] for update equations), $\Delta\boldsymbol{\Phi} = \boldsymbol{\Phi}^* - \boldsymbol{\Phi}^{(i)}$

    $\alpha^* = \arg\min_\alpha C_{KL}(\boldsymbol{\Phi}^{(i)} + \alpha\Delta\boldsymbol{\Phi})$ (see [28] for cost expression C_{KL})

    $\boldsymbol{\Phi}^{(i+1)} = \boldsymbol{\Phi}^{(i)} + \alpha^*\Delta\boldsymbol{\Phi}$, $\boldsymbol{\Phi}_5^{(i+1)} = \boldsymbol{\Phi}_5^{(i)}$, $i = i + 1$

  until $|C_{KL}^{(i+1)} - C_{KL}^{(i)}| < 10^{-4}$

  $\boldsymbol{\Phi}_5^{(i)} = \boldsymbol{\Phi}_5^{(i-1)} + \alpha^*\Delta\boldsymbol{\Phi}_5$

until $||\boldsymbol{\Phi}_5^{(i)} - \boldsymbol{\Phi}_5^{(i-1)}|| < 10^{-4}$

also ignore wavelength-dependent (color) effects by performing inference on the luminance channel and recovering a monochrome BRDF as output. Inferring wavelength-dependent reflectance effects (i.e., a spectral BRDF) would require solving the color constancy problem in conjunction with the reflectometry problem, and we leave this problem for future research.

While we operate in grayscale, we display the input and output using color in this paper. The displayed input is the color image prior to extracting luminance, and the displayed output is the outer product of the recovered monochromatic BRDF and the RGB vector that provides the best fit to the ground truth. This visualization strategy produces artifacts in some cases. For example, the color-visualization of the recovered red material in Fig. 6 does not (and cannot) match the highlight colors of the reference image.

Given a rendered input image of a defined shape (we use a sphere for simplicity), we collect observations from 12,000 normals uniformly sampled on the visible hemisphere. We discard normals that are at an angle of more than 80 degrees from the viewing direction (since the signal to noise ratio is very low at these points) as well as normals that are close to the poles of our parametrization of the sphere (as Eq. 8 is not a good approximation in these regions). This results in an observation vector I of length 8,600.

Each column of Fig. 4 shows a BRDF recovered from a single input image synthesized with either the St. Peter's Basilica or Grace Cathedral environment. Following [27], the recovered BRDFs are compared to ground truth by synthesizing images in a novel environment, and close inspection shows them to be visually quite accurate. Figure 5 further explores stability under changes in lighting. For this, we run our algorithm twice for each material using two different environments and compare the recovered BRDFs. We visualize these BRDFs along with ground truth by using them to synthesize a "spheres image" inspired by [32]. The BRDF estimates are quite consistent across the two environments, and they provide imperfect but reasonable approximations to ground truth.

The same procedure was applied to the captured data from [27]. As above we operate on the luminance channel of HDR images and estimate a monochrome

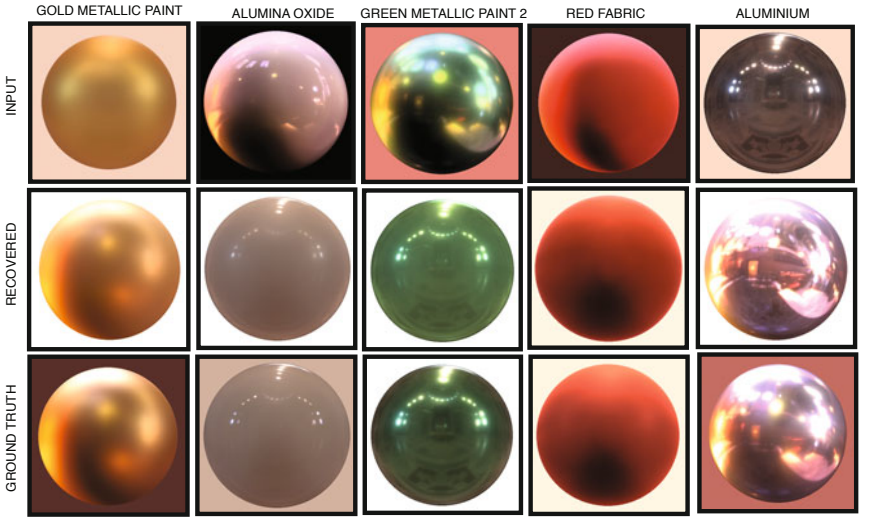

Fig. 4. Evaluation with synthetic input. *Top*: Single image used as input. *Middle*: Appearance predicted in a novel environment using the recovered BRDF. *Bottom*: Ground truth image in the same novel environment.

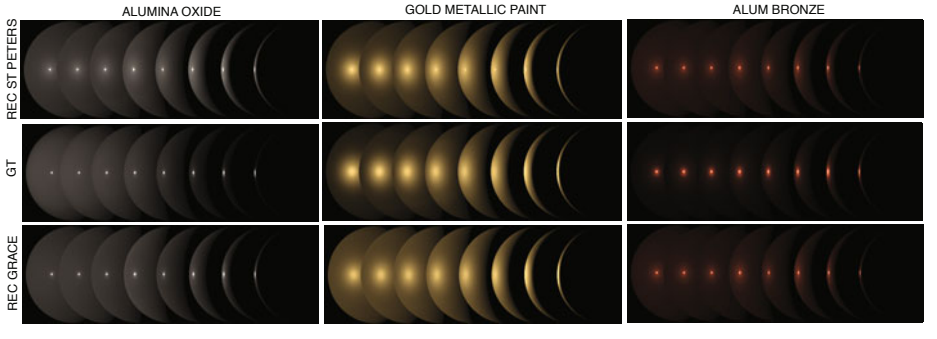

Fig. 5. Stability under changes in lighting: BRDFs recovered when the same material is seen in different environments. *Top to bottom*: BRDF recovered in the St. Peter's environment; ground truth; and BRDF recovered in the Grace Cathedral environment.

BRDF, but now we visualize the output in color by taking the outer product of the monochrome BRDF and the median RGB color of the input image. Figure 6 shows results with the BRDFs recovered from single input images (top row) being used to render synthetic images of the same material under novel environments (more precisely, the same environment from a different viewpoint). Accuracy is assessed by comparing these synthetic images to real images captured in the same novel environments. While the recovered reflectance is clearly distinguishable from ground truth, we see that useful qualitative reflectance information is still

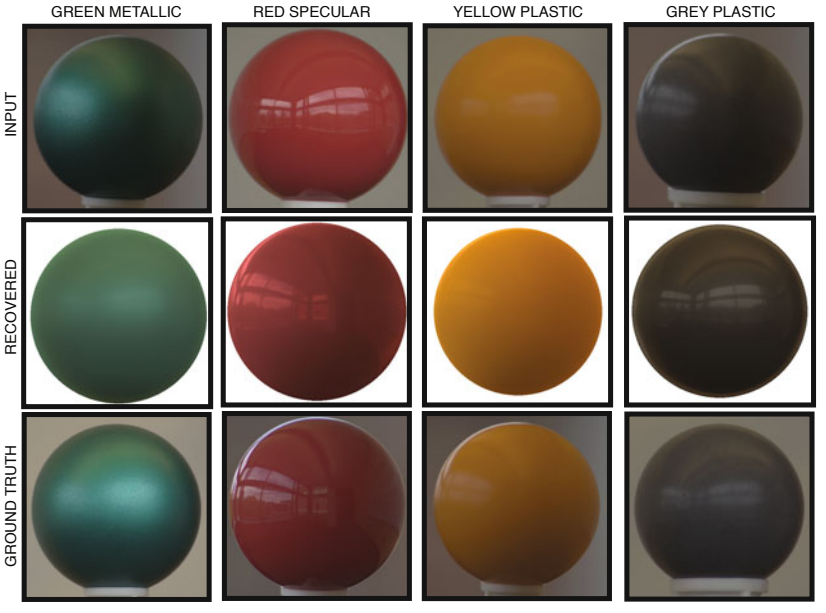

Fig. 6. Evaluation with captured input. *Top*: Image used as input. *Middle*: Appearance predicted in a novel environment using the recovered BRDF. *Bottom*: Ground truth images captured in the same novel environments. (Image data provided by [27].)

obtained. Based on the inferred BRDFs, for example, it would be straightforward to create an ordering of the four materials based on gloss.

These results reveal two limitations of the approach. First, one should expect less accuracy when the input image contains significant mesostructure (e.g., `green metallic`) or texture because these small variations effectively increase noise. Second, performance will be diminished when the illumination is "inadequate", meaning that it does not induce significant specular, grazing, and/or retro reflections, and does not sufficiently constraint the BRDF (e.g., `red specular` and `yellow plastic`). This latter limitation is consistent with perceptual findings [7] and frequency-domain arguments [26], and it has been documented for cases in which the environment is known [27]. Romeiro et al. [27] also describe why quantitative analysis of the conditions for adequate illumination are difficult: Unlike the spherical harmonic approach [26], lighting and reflectance are not related by a convolution operator in the present case. Perhaps a quantitative description of the conditions for "adequate illumination" in material recognition will be a fruitful direction for future work.

5 Discussion

Our results suggest that for a range of homogeneous diffuse and glossy materials, reflectance information can be inferred from unknown real-world illumination when the object shape is known. They also suggest that these estimates are fairly stable when the illumination undergoes a change.

The approach has at least two features worth highlighting. First, it uses a linear basis for reflectance. This allows a seamless trade between complexity and accuracy, and it is very different from "parametric" BRDF models (Phong, etc.) that are non-linear in their variables and are only suitable for a particular material class. Second, it is an inference system built upon a probabilistic generative image model, and this makes it amenable to combination with other contextual cues and vision subsystems. In particular, we might explore combinations with shape-from-X techniques (shading, contours, shadows, etc.) to assess how well reflectance can be recovered when shape is not known *a priori*.

Acknowledgments. The authors thank Bill Freeman and Yair Weiss for helpful discussions regarding bilinear inference problems. This work was supported by the Office of Naval Research through award N000140911022, NSF Career Award IIS-0546408, and a fellowship from the Alfred P. Sloan Foundation.

References

1. Beck, J., Prazdny, S.: Highlights and the perception of glossiness. Perception & Psychophysics 30(4) (1981)
2. Brainard, D., Freeman, W.: Bayesian color constancy. J. Opt. Soc. Am. A 14(7) (1997)
3. Ding, C., Li, T., Jordan, M.: Convex and semi-nonnegative matrix factorizations. IEEE T. Pattern Anal. 32(1) (2010)
4. Dror, R.: Surface reflectance recognition and real-world illumination statistics. Ph.D. thesis, Massachusetts Institute of Technology (2002)
5. Dror, R., Willsky, A., Adelson, E.: Statistical characterization of real-world illumination. J. Vision 4 (2004)
6. Fergus, R., Singh, B., Hertzmann, A., Roweis, S.T., Freeman, W.T.: Removing camera shake from a single photograph. ACM T. Graphics (Proc. ACM SIGGRAPH) 25(3) (2006)
7. Fleming, R., Dror, R.O., Adelson, E.H.: Real-world illumination and the perception of surface reflectance properties. J. Vision 3(5) (2003)
8. Georghiades, A.: Recovering 3-D shape and reflectance from a small number of photographs. In: Proc. Eurographics Workshop on Rendering (2003)
9. Ghosh, A., Achutha, S., Heidrich, W., O'Toole, M.: BRDF acquisition with basis illumination. In: Proc. IEEE Int. Conf. Computer Vision (2007)
10. Haber, T., Fuchs, C., Bekaer, P., Seidel, H., Goesele, M., Lensch, H.: Relighting objects from image collections. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition (2009)
11. Hara, K., Nishino, K., Ikeuchi, K.: Mixture of spherical distributions for single-view relighting. IEEE T. Pattern Anal. 30(1) (2008)
12. Levin, A., Weiss, Y., Durand, F., Freeman, W.: Understanding and evaluating blind deconvolution algorithms. In: Proc. IEEE Conf. Computer Vision and Pattern Recognition (2009)
13. Mahajan, D., Ramamoorthi, R., Curless, B.: A theory of spherical harmonic identities for BRDF/lighting transfer and image consistency. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3954, pp. 41–55. Springer, Heidelberg (2006)

14. Marschner, S., Westin, S., Lafortune, E., Torrance, K., Greenberg, D.: Image-based BRDF measurement including human skin. In: Proc. Eurographics Symposium on Rendering (1999)
15. Matusik, W., Pfister, H., Brand, M., McMillan, L.: A data-driven reflectance model. ACM T. Graphics (Proc. ACM SIGGRAPH) 22(3) (2003)
16. Miskin, J.: Ensemble Learning for Independent Component Analysis. Ph.D. thesis, University of Cambridge (2000)
17. Miskin, J., MacKay, D.: Ensemble learning for blind source separation. Independent Component Analysis: Principles and Practice (2001)
18. Ng, R., Ramamoorthi, R., Hanrahan, P.: Triple product wavelet integrals for all-frequency relighting. ACM T. Graphics (Proc. ACM SIGGRAPH) 23(3) (2004)
19. Ngan, A., Durand, F., Matusik, W.: Experimental analysis of BRDF models. In: Eurographics Symposium on Rendering (2005)
20. Nishino, K.: Directional statistics BRDF model. In: Proc. IEEE Int. Conf. Computer Vision (2009)
21. Nishino, K., Zhang, Z., Ikeuchi, K.: Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis. In: Proc. IEEE Int. Conf. Computer Vision (2001)
22. Pellacini, F., Ferwerda, J., Greenberg, D.: Toward a psychophysically-based light reflection model for image synthesis. In: Proc. ACM SIGGRAPH (2000)
23. Praun, E., Hoppe, H.: Spherical parametrization and remeshing. ACM T. Graphics 22(3) (2003)
24. Ramamoorthi, R., Hanrahan, P.: A signal-processing framework for inverse rendering. In: Proceedings of ACM SIGGRAPH, pp. 117–128 (2001)
25. Ramamoorthi, R., Hanrahan, P.: Frequency space environment map rendering. ACM Transactions on Graphics (TOG) 21(3), 517–526 (2002)
26. Ramamoorthi, R., Hanrahan, P.: A signal-processing framework for reflection. ACM T. Graphics 23(4) (2004)
27. Romeiro, F., Vasilyev, Y., Zickler, T.: Passive reflectometry. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 859–872. Springer, Heidelberg (2008)
28. Romeiro, F., Zickler, T.: Ensemble learning for reflectometry. Tech. Rep. TR-06-10, Harvard School of Engineering and Applied Sciences (2010), `ftp://ftp.deas.harvard.edu/techreports/tr-06-10.pdf`
29. Rusinkiewicz, S.: A new change of variables for efficient BRDF representation. In: Eurographics Rendering Workshop '98 (1998)
30. Sato, Y., Wheeler, M., Ikeuchi, K.: Object shape and reflectance modeling from observation. In: Proceedings of ACM SIGGRAPH (1997)
31. Sharan, L., Li, Y., Motoyoshi, I., Nishida, S., Adelson, E.: Image statistics for surface reflectance perception. J. Opt. Soc. Am. A 25(4) (2008)
32. Stark, M., Arvo, J., Smits, B.: Barycentric parameterizations for isotropic BRDFs. IEEE T. Vis. Computer Graphics 11(2) (2005)
33. Todd, J., Farley norman, J., Mingolla, E.: Lightness constancy in the presence of specular highlights. Psychological Science 15(1), 33–39 (2004)
34. Wang, R., Ng, R., Luebke, D., Humphreys, G.: Efficient wavelet rotation for environment map rendering. In: Eurographics Symposium on Rendering (2006)
35. Wills, J., Agarwal, S., Kriegman, D., Belongie, S.: Toward a perceptual space for gloss. ACM T. Graphics 28(4) (2009)
36. Yu, T., Wang, H., Ahuja, N., Chen, W.: Sparse lumigraph relighting by illumination and reflectance estimation from multi-view images. In: Eurographics Symposium on Rendering (2006)