Inferring Reflectance under Real-world Illumination

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:24947961">http://nrs.harvard.edu/urn-3:HUL.InstRepos:24947961</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Inferring Reflectance under Real-world Illumination

Fabiano Romeiro
and
Todd Zickler

TR-10-10

Computer Science Group
Harvard University
Cambridge, Massachusetts
Inferring Reflectance under Real-world Illumination

Fabiano Romeiro · Todd Zickler

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 suggest that usable reflectance information can be often be inferred, and that these estimates are stable under changes in lighting.

Keywords Reflectance · Bi-directional reflectance distribution function · Natural image statistics · Material recognition · Blind source separation

1 Introduction

The optical properties of a material often provide information regarding 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 complicated 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, which is depicted in the first row of 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. We assume that all light sources and objects in the environment are far from the shape of interest (distant lighting), and we assume that the material is well-described by an isotropic BRDF (i.e., it has no “grain”).

This problem is difficult for two reasons. First, the set of possible outputs is large relative to our measurement space. Our desired output—an isotropic BRDF—is a function on a three-dimensional domain (Nicodemus et al, 1977), while our input image is function of only two-dimensions. This means that our problem is under-determined, or at least it appears to be. The second difficulty arises from the fact that there are many lighting/reflectance-pairs that can explain any given image. As depicted in the middle of Fig. 1, in addition to the actual lighting and reflectance, an image can always be trivially explained by a mirror-like BRDF and a carefully-crafted “blurry” lighting environment.

Fabiano Romeiro
Harvard University
E-mail: romeiro@fas.harvard.edu

Todd Zickler
Harvard University
33 Oxford St., Cambridge MA 02138
Phone: (617) 495-4390
E-mail: zickler@seas.harvard.edu
Of course, there are many other solutions as well, so in addition to being under-determined, our problem is ill-posed.

Our attack on this problem is two-pronged. First, we restrict the output space by imposing a non-traditional symmetry condition on the BRDF. This symmetry, which we call half-vector symmetry, is empirically-motivated, and it reduces the isotropic BRDF domain from three dimensions to two, thereby matching the domain of the output to that of the input. By expressing reflectance as a linear combination of basis functions in this reduced domain, we arrive at a representation that provides a fine balance between accuracy and tractability. It is capable of representing important phenomena, such as specular and off-specular peaks, retro-reflection, and grazing angle effects; and at the same time, it is simple enough to be used for inference. This is in contrast to more commonly-used non-linear BRDF models (Phong, Ward, Lafortune, Cook-Torrance, Oren-Nayar, etc.), which have limited accuracy or are tied to particular material classes.

The second prong of our attack deals with the ambiguity between reflectance and lighting, and it does so by exploiting the statistics of real-world lighting environments in a Bayesian framework. In this we are motivated by perceptual studies suggesting that humans are capable of inferring reflectance information when the directional statistics of the environment do not deviate significantly from those found in nature (Fleming et al, 2003). Our approach is to compute the BRDF that is most likely under a distribution of probable lighting environments—a Bayesian computational strategy adopted from solutions to other ill-posed vision and imaging problems, including color constancy (Brainard and Freeman, 1997), image super-resolution (Tipping and Bishop, 2002), and blind image deblurring (Fergus et al, 2006).

We evaluate our approach using both real and synthetic images, and we consider cases in which the lighting is both known and unknown a priori. Our results suggest that, much like the human visual system, we can infer usable reflectance information whenever the lighting environment is “adequately complex”. Preliminary versions of this work appeared in (Romeiro et al, 2008; Romeiro and Zickler, 2010a).

2 Background and Related Work

The BRDF describes the manner in which incident radiant flux is modulated by a surface patch. It is a positive function of four angular dimensions and can be written $F(\omega_i, \omega_o)$, where $\omega_i$ and $\omega_o$ are unit vectors in the hemisphere centered about the patch normal. These are the directions of incident and reflected flux, respectively, and they are often expressed in spherical coordinates: $(\theta_i, \phi_i)$ and $(\theta_o, \phi_o)$. The BRDF provides an accurate and complete description of appearance for optically-thick materials that are observed at an appropriate scale, meaning one for which mesostructure effects (i.e., mutual illumination and shadowing between nearby pixels) and sub-surface scattering effects are negligible (Nicodemus et al, 1977). We assume this scale condition to be satisfied throughout this document.

One can measure the BRDF of a planar material by sampling the double-hemisphere of input and output directions with a gonioreflectometer. Since this is extremely slow, and since a slight loss of accuracy is often acceptable for vision and graphics applications, a number of camera-based alternatives have been proposed. When a camera is used with a curved mirror (Ward, 1992) or a curved material sample (Marschner et al, 1999), one image provides a dense sampling of a two-dimensional slice of the BRDF. To sample the complete BRDF domain, these can be combined with a moving light source (Ward, 1992; Marschner et al, 1999) or a projector (Ghosh et al, 2007). These camera-based systems significantly reduce measurement time, but they also require special-purpose hardware and active control of lighting.

Reflectometry in more natural conditions requires a passive approach. When we do not have the luxury of manipulating incident lighting, we must exploit instead
the lighting that naturally exists. Previous computational methods for passive reflectometry fall into two categories—those that assume the lighting to be known a priori and those that do not.

2.1 Passive methods: known lighting

In the computer graphics community, the inference of reflectance from natural images has been studied under the banner of “inverse rendering”. Many approaches exist, and almost all of them make the problem tractable by restricting the space of materials through use of “parametric” BRDF models—such as the Phong, Ward, Cook-Torrance, and Lafortune models—that are non-linear in a small number of parameters. These parametric BRDF models impose strong constraints on reflectance, and as a result, one can exploit them to recover substantial information about a scene. For example, there are methods for recovering the spatially-varying BRDF of a complex scene exhibiting mutual illumination from one (Boivin and Gagalowicz, 2001) or more (Yu et al., 1999; Debevec et al., 2004) images of known geometry under known real-world lighting.

The use of parametric BRDF models typically has two significant limitations. First, it restricts the space of materials because most models are only accurate for a particular material class (Ngan et al., 2005; Stark et al., 2005). Second, because these models are non-linear in their parameters, the required optimization ends up being model-specific, and it 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 (Mahajan et al., 2006; Haber et al., 2009) and graphics (e.g., Ramamoorthi and Hanrahan, 2001; Ng et al., 2004), and it is the key to making our approach tractable.

The most straight-forward linear representation is a tabulated BRDF (Matusik et al., 2003), which implies an interpolation basis. Equally fundamental is the spherical harmonic basis, which allows one to exploit the fact that image data can be expressed as a convolution between reflectance and lighting (Ramamoorthi and Hanrahan, 2001). Spherical harmonics enable a frequency domain analysis that, among other things, provides BRDF coefficients in closed form when lighting and shape are known. The drawback of this approach is that it can only yield general isotropic BRDFs if the full four-dimensional light field is observed as input. When only a single input image is available, it can only yield a “radially-symmetric” BRDF—a function of one angle and of the form $F(\omega_i, \omega_o) = F(\langle g(\omega_i), \omega_o \rangle)$ for some map $g$ from the hemisphere to itself—that is incapable of representing grazing-angle effects (Ramamoorthi and Hanrahan, 2004).

2.2 Passive methods: unknown lighting

All of the techniques described above assume that lighting (and shape) are known, and comparatively few methods have been developed for the case in which lighting is not pre-determined. In this case one must deal with the lighting/reflectance ambiguity depicted in Fig. 1 and somehow avoid the trivial solution of a mirror-like BRDF.

Existing approaches to reducing the ambiguity include restricting the space of allowable reflectances and/or lighting environments so that the trivial solution lies outside of the pre-determined feasible set. This approach is taken by Hara et al. (2008), who recover both lighting and reflectance from a single image by assuming that lighting consists of a few point sources and reflectance is well-represented by a modified Torrance-Sparrow model. Another approach is to capture multiple images of a material from distinct view directions. In the limit of increasing view directions, one samples the entire four-dimensional light field, and the ambiguity goes away (Ramamoorthi and Hanrahan, 2001). A method that seems to combine both of these insights is that of Haber et al. (2009). It employs a linear BRDF basis that may not include the mirror BRDF, and at the same time, it benefits from having multiple images of the same material (under distinct and unknown illuminations).

2.3 Human inspiration

Perceptual studies suggest that humans can also infer reflectance information from a single image. The mechanisms that underly human material recognition are still being studied (Beck and Prazdny, 1981; Pellacini et al., 2000; Fleming et al., 2003; Todd et al., 2004; Sharan
et al, 2008; Wills et al, 2009), but there is compelling evidence that people do not require contextual knowledge of the environment to infer reflectance, and that they can do so reliably when the directional statistics of the environment do deviate significantly from those found in nature (Fleming et al, 2003). These findings provide motivation for our work, which also leverages the directional statistics of natural environments.

Much like real-world images, the statistics of band-pass filter coefficients of real-world lighting display significant regularity: their distributions are highly kurtotic, having heavy tails (Dror, 2002; Dror et al, 2004). We exploit this insight computationally by using it to define a prior probability distribution of lighting environments, and we augment this with a computational process inspired by other bilinear inference problems in vision: color constancy, super-resolution, and blind image deblurring (Brainard and Freeman, 1997; Tipping and Bishop, 2002; Fergus et al, 2006; Levin et al, 2009). 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 (Levin et al, 2009), and the same basic idea forms the core of our approach.

3 A Bivariate BRDF for Reflectometry

For many materials, the dimension of the BRDF domain can be reduced without incurring a significant loss of detail. The domain can be folded in half, for example, because reciprocity ensures that BRDFs are symmetric about the directions of incidence and reflection: \( F(\omega_i, \omega_o) = F(\omega_o, \omega_i) \). For isotropic materials, meaning those without a preferred direction for local scattering, the BRDF is unchanged as the input and output directions are rotated about the surface normal as a fixed pair, so the domain \((\theta_i, \phi_i, \theta_o, \phi_o)\) can be reduced to the three-dimensional domain \((\theta_i, \theta_o, \phi_i - \phi_o)\). The BRDF of an isotropic material is also unchanged when the output direction is reflected across the incidence plane—Marschner (1998) refers to this property separately as bilateral symmetry—and this induces an additional folding of the domain onto \((\theta_i, \theta_o, |\phi_i - \phi_o|)\).

It is convenient to parameterize the BRDF domain in terms of halfway and difference angles (Rusinkiewicz, 1998). Accordingly, the complete 4D domain is written in terms of the spherical coordinates of the halfway vector, \( h = (\omega_i + \omega_o)/|\omega_i - \omega_o| \), and those of the input direction with respect to the halfway vector: \((\theta_h, \phi_h, \theta_d, \phi_d)\) (see Fig. 2). In this parameterization, Rusinkiewicz (1998) notes that the folding due to reciprocity corresponds to \( \phi_d \rightarrow \phi_d + \pi \), and the projection due to isotropy (without bilateral symmetry) is one onto \((\theta_h, \theta_d, \phi_d)\). In Appendix A, we show that bilateral symmetry induces the additional folding \( \phi_d \rightarrow -\pi - \phi_d \). This seems to have been previously neglected in the literature (cf. (Rusinkiewicz, 1998; Matusik et al, 2003)), and when it is taken into account, we obtain the domain \((\theta_h, \theta_d, \phi_d) \subset [0, \pi/2]^3\) for isotropic BRDFs.

In this paper, we consider an additional projection of the BRDF domain, one that reduces it from three dimensions down to two. We project \((\theta_h, \theta_d, \phi_d) \subset [0, \pi/2]^3\) to \((\theta_h, \theta_d) \subset [0, \pi/2]^2\), as depicted in the left of Fig. 2. The projection is acceptable whenever a BRDF exhibits little change for rotations of the input and output directions (as a fixed pair) about the halfway vector. This is a direct generalization of isotropy, bilateral symmetry and reciprocity, which already restrict the BRDF to be completely determined over these same rotations by the sub-interval \([0, \frac{\pi}{2}],\). We refer to materials that satisfy this requirement (for some definition of “little change”) as being bivariate or half-vector symmetric. As shown in the right of Fig. 2, the bivariate representation leads to an intuitive RGB visualization of reflectance. Specular (and off-specular) peaks appear along the left edge, grazing-angle effects appear in the bottom-left corner, and retro-reflections appear along the top edge.

The accuracy of such bivariate representations of the materials in the MERL/MIT BRDF database (Matusik et al, 2003) is shown in Fig. 3, where materials are...
sorted by relative RMS BRDF error:

$$E_{\text{rms}} = \left( \sum_{\theta_h, \theta_d, \phi_d} \frac{(F(\theta_h, \theta_d, \phi_d) - F(\theta_h, \theta_d))^2}{(F(\theta_h, \theta_d))^2} \right)^{\frac{1}{2}},$$

with

$$F(\theta_h, \theta_d) = \frac{1}{\phi(\theta_h, \theta_d)} \sum_{\phi(\theta_h, \theta_d)} F(\theta_h, \theta_d, \phi_d).$$

Here, $\phi(\theta_h, \theta_d)$ is the set of valid $\phi_d$ values given fixed values of $\theta_h$ and $\theta_d$. These relative RMS errors are shown in red in this figure, and they can be decomposed into two separate sources. Some of this “error”, shown in blue, comes from the fact that the MERL data ignores bilateral symmetry and tabulates the isotropic domain using $(\theta_h, \theta_d, \phi_d) \subset [0, \pi/2] \times [0, \pi/2] \times [0, \pi)$ instead of $(\theta_h, \theta_d, \phi_d) \subset [0, \pi/2]^3$. This over-parameterization permits anisotropic highlights in the grease-covered-steel and green-acrylic materials, for example, and it accounts for a significant portion of the total difference between the original data and our bivariate reduction of it.

More important than RMS error in the BRDF domain is the difference we observe in a typical image, so Fig. 3 also includes synthetic images of materials that are more and less well-represented by a bivariate BRDF. Overall, this qualitative test suggests that the overwhelming majority of the materials in the database are reasonably well-represented by bivariate functions, and that the bivariate reduction may even have a positive effect in some cases. For example, the star-shaped highlights of the original green-acrylic material are caused by lens flare artifacts embedded in the measurements\(^1\), and these are removed by the bivariate reduction.

Motivation for a bivariate representation is provided by Rusinkiewicz (1998), who notes that a microfacet BRDF will be dominated by a separable bivariate representation $F(\theta_h)F(\theta_d)$ away from grazing angles. Additional motivation comes from the work of Stark et al (2005), who show empirically that a carefully-selected two-dimensional domain is often sufficient for capturing (off-)specular reflections, retro-reflections, and important Fresnel effects. The two-dimensional domain $(\theta_h, \theta_d)$ that we propose here is homeomorphic to that of Stark et al, which is why it possesses these same properties. Stark et al propose the ‘$\alpha$-$\sigma$-parameterization’ for two-dimensional BRDFs, and this is related to $(\theta_h, \theta_d)$

\(^1\) W. Matusik, personal communication.
Fig. 4 Qualitative evaluation of the linear BRDF model learned via non-negative matrix factorization (NMF) of the MERL/MIT database. Left, top to bottom: NMF model with 3, 5, and 10 basis functions; ground truth; and Cook-Torrance fit by Ngan et al (2005). Right: Ten non-negative basis functions, depicted as images in the bivariate domain.

For this reason, Figs. 2 and 3 can be seen as providing a new interpretation and validation for their model. (The original paper examined Cornell BRDF data (Westin, 2003), which is arguably more accurate but also sparse.)

One important advantage of the \((\theta_h, \theta_d)\) parameterization is that it provides an intuitive means for choosing bases for the bivariate domain. In this paper, we employ two approaches that differ in their trade-off between complexity and generality. When inferring reflectance under known lighting (Sect. 4), we can tolerate a more general representation, and we use an interpolating basis, representing each material as a tabulated set of samples. We employ non-uniform sampling in \((\theta_h, \theta_d)\) to better represent specular highlights, retro-reflections and Fresnel effects, and we implement this by defining continuous functions \(\zeta(\theta_h, \theta_d)\) and \(\xi(\theta_h, \theta_d)\) before sampling uniformly in \((\zeta, \xi)\). Inspired by Matusik et al (2003), we use \(\zeta = 2\theta_d/\pi, \xi = \sqrt{2\theta_h}/\pi\) in this paper, since this increases the sampling density near specular reflections \((\theta_h \approx 0)\). Accordingly, each reflectance function is defined on the square, \((\zeta, \xi) \in [0, 1] \times [0, 1]\), and by creating a regular grid of samples on this square, \(\Xi = \{[\zeta_s, \xi_t]\}\), each material is represented by a vector of coefficients \(f = \{f_{st}\}\). In Sect. 4 we find it sufficient to use a 32 \times 32 discretization, giving a total of \(|\Xi| = 1024\) coefficients.

When lighting is not known (Sect. 5), we find that accurate inference requires a more restricted representation of reflectance. For this, we reduce the number of basis functions by applying non-negative matrix factorization (NMF) to all of the 100 materials in the MERL/MIT database. The bivariate reduction is useful for this factorization because it significantly reduces the computational burden: By first applying the bivariate reduction, we shrink each material to a 90 \times 90 array, and we can easily handle the entire database (and potentially much more) without requiring out-of-core factorization methods. The left of Fig. 4 evaluates the NMF representation qualitatively as we increase the number of non-negative bivariate basis functions. The BRDF visualization in this figure is inspired by Stark et al (2005) and consists of a collage of sphere-images, each rendered under a distinct directional illuminant. For additional comparison, the left of the figure also includes a non-linear BRDF representation (Cook-Torrance) as fit to the same data by Ngan et al (2005). In Sect. 5, we choose 10 NMF basis functions as a good balance between complexity and accuracy for inference using current commodity computers, and these ten basis functions are shown in the right of Fig. 4. A discussion of alternative bases can be found in Sect. 6.

4 Reflectance from known lighting

We now turn to the task of inferring a BRDF from a single HDR image, and we begin by assuming the light-
ing environment is known, such as that measured by an HDR illumination probe. In this case, each pixel in the images provides a linear constraint on the BRDF parameters, and our goal is to infer the reflectance function from these constraints. Previous analysis tells us that the constraints from a single image are not sufficient to recover a general three-dimensional isotropic BRDF (Ramamoorthi and Hanrahan, 2001), but the results in this section suggest that they often are sufficient to recover plausible bivariate reflectance.

We assume all sources and reflecting surfaces in the environment to be far from the object in question so that the (known) 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; \( L : S^2 \rightarrow \mathbb{R}^+ \). We also assume that the camera and object geometry are known as stated earlier, and that mutual illumination is negligible. Then, a linear measurement made at pixel \( j \) can be written (e.g., Ng et al (2004); Haber et al (2009))

\[
I_j = \int_{\Omega} L(R_j^{-1} \omega_i) V_j(\omega_i) \Phi(\omega_i, R_n v) \cos \theta_i d\omega_i, \tag{3}
\]

where \( v \in S^2 \) is the view direction in a global coordinate system, \( R_j \in SO(3) \) is any rotation that simultaneously maps the surface normal at the back-projection of pixel \( j \) to the north pole and the view direction to the equator (i.e., it converts global spherical coordinates to coordinates of the local hemisphere \( \Omega \)), and \( V_j(\omega) \) is a binary hemispherical “visibility” function that encodes the object’s self-shadowing at the back-projection of the pixel.

Equation 3 imposes a linear constraint on our BRDF coefficients \( f = \{ f_{\sigma} \} \), and to obtain an expression for this constraint, we approximate the rendering equation by a sum over a discrete set \( \Omega_d \) of lighting directions on the hemisphere:

\[
I_j \approx \frac{2\pi}{|\Omega_d|} \sum_{\omega_i \in \Omega_d} \left( \sum_{\zeta, \xi \in N_k} \alpha_{jk}^L \Phi(R_n^{-1} \omega^k_i) f_{\sigma} \right) \cos \theta_i^k, \tag{4}
\]

where \( N_k \) is the set of four bivariate grid points that are closest to \( \zeta(\omega_i^k, R_j v) \), \( \xi(\omega_i^k, R_j v) \)—this is overloaded notation for \( \zeta \) and \( \xi \), which depend on the incident and reflected directions through \( (\theta_h, \theta_r) \)—and \( \alpha_{jk}^L \) is the coefficient of the bilinear interpolation associated with these grid points. (We find a piecewise linear approximation of the BRDF to be adequate.) This equation can be rewritten as

\[
I_j \approx \frac{2\pi}{|\Omega_d|} \sum_{(\zeta, \xi) \in \Xi} \sum_{\omega_i^k \in \text{bin}_{\zeta, \xi}} \alpha_{jk}^L \Phi(R_n^{-1} \omega_i^k) \cos \theta_i^k, \tag{5}
\]

to emphasize its interpretation as an inner product involving vector \( f \).

Pixels corresponding to surface points with distinct normals and/or visibilities provide constraints that lead to a system of equations

\[
I = Lf \tag{6}
\]

where \( I = \{ I_j \} \), and \( L \) is a lighting matrix whose rows are given by the non-BRDF terms in Eq. 5. These constraints are visualized in Fig. 5 for the special case of a convex surface for which \( V_j(\omega) = 1 \forall j \). At each pixel, the integral in Eq. 3 is computed over the visible hemisphere of light directions \( \Omega \). Our use of a bivariate BRDF induces a “folding” of this hemisphere across the view/normal plane because, due to reciprocity and bilateral symmetry, light directions \( \omega \) and \( \omega' \) that are symmetric about this plane always correspond to the same BRDF value. The constraint from each pixel can therefore be interpreted as an inner product between the unknown BRDF coefficients \( f \) and a hemisphere of illumination that is weighted, folded across the local view/normal plane, and warped onto the \( \zeta, \xi \)-plane.

To infer a BRDF, we seek a solution \( f \) that satisfies the constraints of Eq. 6. A simple least squares estimate might work well in the noiseless case, but in practice we require regularization to handle noise caused by the sensor, the bivariate approximation, the discretization of the rendering equation, errors in the assumed surface shape, and perhaps most importantly, a lighting matrix \( L \) that is not full rank.
As with general 4D BRDFs, bivariate BRDFs vary slowly over much of their domain. Regularization can therefore be implemented in the form of a smoothness constraint in the $\zeta\xi$-plane. There are many choices here, and we have found spatially-varying Tikhonov-like regularization to be effective. According to this design choice, the optimization becomes

$$\arg\min_f \| I - \mathcal{L}f \|^2 + \alpha \left( \| \Lambda_{\zeta}^{-1} D_{\zeta} f \|^2 + \| \Lambda_{\xi}^{-1} D_{\xi} f \|^2 \right)$$

subject to $f \geq 0,$

(7)

where $D_{\zeta}$ and $D_{\xi}$ are $|\Xi| \times |\Xi|$ finite-difference matrices, and $\alpha$ is a tunable scalar regularization parameter. The matrices $\Lambda_{\zeta}$ and $\Lambda_{\xi}$ are diagonal $|\Xi| \times |\Xi|$ matrices that affect non-uniform regularization in the bivariate BRDF domain. Their diagonal entries are learned from the MERL/MIT database by setting each to the variance of the partial derivative at the corresponding $\zeta\xi$-domain point, where the variance is computed across all materials in the database. Probabilistically, this approach can be interpreted as seeking the MAP estimate with independent, zero-mean Gaussian priors on the bivariate BRDF’s partial derivatives.

4.2 Evaluation and Results

We begin with an evaluation that uses images synthesized with tabulated BRDF data from the MERL/MIT database, measured illumination\(^2\), and a physically based renderer\(^3\). Using these tools, we can render images for input to our algorithm as well as images with the recovered BRDFs for direct comparison to ground truth. In all cases, we use complete 3D isotropic BRDF data to create the images for input and ground-truth comparison, since this is closest to a real-world setting. Also, we focus our attention on the minimal case of a single input image; with additional images, the performance can only improve. It is worth emphasizing that this data is not free of noise. Sources of error include the fact that the input image is rendered with a 3D BRDF as opposed to a bivariate one, that normals are computed from a mesh and are stored at single precision, and that a discrete approximation to the rendering equation is used.

Given a rendered input image of a defined shape (we use a sphere for simplicity), we harvest observations from 8,000 normals uniformly sampled on the visible hemisphere to create an observation vector $I$ of length 8,000. We discard normals that are at an angle of more than 80° from the viewing direction, since the signal to noise ratio is very low at these points. The bivariate BRDF domain is represented using a regular $32 \times 32$ grid on the $\zeta\xi$-plane, and our observation matrix $L$ is therefore $M \times 1024$, where $M$ is the number of useable normals. The entries in $L$ are computed using Eq. 5 with 32,000 points uniformly distributed on the illumination hemisphere. With $I$ and $L$ determined, we can solve for the unknown BRDF as described previously.

We find it beneficial to use a small variant of the optimization in Eq. 7: We solve the problem twice using two separate pairs of diagonal weight matrices ($\Lambda_{\zeta}, \Lambda_{\xi}$). One pair gives preference to diffuse reflectance, while the other gives preference to gloss. This provides two solutions, and we choose the one with lowest residual.

\(^2\) [http://www.debevec.org/Probes/](http://www.debevec.org/Probes/)

\(^3\) [http://www.pbrt.org/](http://www.pbrt.org/)
Fig. 6 Visual evaluation with MERL BRDF data. A bivariate BRDF is estimated from a single input image (top), and this estimate is used to render a new image under novel lighting (second row). Ground truth images for the novel environments are shown for comparison, along with difference images scaled by 100. Few noticeable differences exist. Far right: Environment maps used in the paper, top to bottom: St. Peter’s Basilica, Grace Cathedral, Uffizi Gallery, Cafe and Corner Office.

Using this procedure, we were able to use the same weight matrices and regularization parameter ($\alpha$) for all results in this section. In every case, the optimizations were initialized with a Lambertian BRDF, and color is handled by running the algorithm independently in each of the RGB color channels.

Results are shown in Fig. 6. Each row shows an input (HDR) image and compares the recovered bivariate BRDF to the (3D) ground truth by synthesizing images in an environment that is different from the input environment. Close inspection reveals very little noticeable difference between the two images, and the recovered BRDF is visually quite accurate. There are numerical differences, however, and these have been scaled by 100 for visualization. Note that some of this error is simply due to the bivariate approximation, and this is demonstrated in the corresponding scatter plots. These are incident-plane plots for several incident directions, with the incident and mirror-reflection angles shown in solid and dashed black, respectively, and they show the original 3D MERL/MIT data (“ground truth”), the best-possible bivariate representation as computed in Sect. 3, and the BRDF recovered from the single input image. While the scatter plots reveal clear deviations from ground truth, they suggest that the approach provides reasonable approximations for a variety of materials. This is true even though just a single image is used as input—many fewer than the 300 images that were used to collect the original data (Matusik et al, 2003).

Figure 7 explores the notion of adequate illumination. This figure compares estimates of yellow-matte-plastic obtained using two different input images. The Uffizi Gallery environment (top left) does not provide strong observations of grazing angle effects, so this portion of the BRDF is not accurately estimated. This leads to noticeable artifacts near grazing angles when the recovered BRDF is used for rendering, and it is clearly visible in a scatter plot. When a different environment is used as input, however, more accurate behavior near grazing angles is obtained.

In addition to this evaluation using synthetic data, the same experimental procedure was applied to captured data. Figure 8 shows the results for a number of materials. As before, each BRDF is recovered from a single input image (left), and the recovered BRDFs are used to render synthetic images of the same object from a novel viewpoint (a reasonable proxy for a novel environment). The synthetic images are directly compared to real images captured in the same novel positions. Again, the visual accuracy is reasonable, and while there is no ground truth available for comparison in the scatterplots, the recovered BRDFs seem plausible.

Captured data contains at least three significant sources of noise in addition to what exists in the rendered data above: 1) errors in the assumed surface geometry; 2) surface mesostructure (e.g., the green sphere); and 3) spatial reflectance variations (e.g., the grey sphere). Presently, surface shape is computed by assuming the camera to be orthographic and estimating the center.
and radius of the sphere in the camera’s coordinate system. Errors in this process, coupled with errors in the alignment with the illumination probe, lead to structured measurement noise. Despite this, our results suggest that plausible BRDFs can be recovered for a diversity of materials. Note that additional results can be found in an earlier version of this work (Romeiro et al, 2008).

5 Blind Reflectometry

Let us return to the rendering equation (Eq. 3), and consider the case in which the lighting $L$ is not known. Since everything in the equation is known except the lighting and BRDF, an image $I = \{I_j\}$ 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

$$p(L,F|I) \propto p(I|L,F)p(L)p(F), \quad (8)$$

using a likelihood $p(I|L,F)$ based on Eq. 3. This is closely related to the approach of Haber et al (2009), and when only a single input image is available, 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 (Fleming et al, 2003), 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:

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

The intuition here—adapted directly from the related problem of blind image deblurring (Fergus et al, 2006; Levin et al, 2009)—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. 9 requires prohibitive computation, and to make it feasible, we employ a variational Bayesian technique. Following (Miskin, 2000; Miskin and MacKay, 2001) we approximate the posterior using a separable function,

$$p(L,F|I) \approx q(L,F) = q(L)q(F), \quad (10)$$

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. 9) 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.

5.1 Representing illumination

We represent spherical lighting using a wavelet basis. As depicted in Fig. 9 and following (Wang et al, 2006), we do this by mapping the sphere to a plane with an octahedral map (Praun and Hoppe, 2003) and using a Haar wavelet basis in this plane. Notationally, we write $L = \sum_{m=1}^{M} \ell_m \psi_m$ with $\psi_m$ the basis functions and $\ell_m$ 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 have heavy tails (Dror, 2002; Dror et al, 2004). 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
Fig. 8 Results using captured data. A BRDF is estimated from a single input image (top) under a known environment. This recovered BRDF is used to render a synthetic image for novel view within the same environment (middle). An actual image for the same novel position is shown for comparison (bottom). Despite the existence of non-idealities such as surface mesostructure and spatial inhomogeneity, plausible BRDFs are recovered.

Graphics Lab\(^4\) and others from the sIBL Archive\(^5\), normalized each so that it integrates to one, and studied the coefficient distributions at different scales. Like Dror et al (2004), we find these statistics to be notably non-stationary, especially at coarser scales. Figure 9 shows empirical distributions and parametric fits for a variety of scales using a $32 \times 32$ 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 $\ell_1$ is the same in all cases since the illuminations are normalized.

With these definitions we can write our illumination prior as

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

with $N_m$ the number of mixture components for coefficient $m$, and $\pi_{nm}$ the mixing weights. The group structure described above is implicit in this notation: All coefficients in any one group share the same $N_m$, $\pi_{nm}$ and $p_{nm}$. The specific forms for each group can be found in Appendix B.

5.2 Representing reflectance

As described in Sect. 3, we require a more restrictive representation of reflectance for accurate inference when the lighting is unknown, and we represent BRDFs as a linear combination of positive basis functions learned through non-negative matrix factorization (NMF). This produces the linear representation, $F = \sum_{k=1}^{K} f_k \phi_k$, and the advantage of the non-negative basis functions $\phi_k \geq 0$ is that non-negativity constraints on the recovered BRDF can be naturally enforced through non-negativity constraints on the coefficients $f \triangleq \{ f_k \}$. Moreover, we find that the empirical distributions of the resulting coefficients $f_k$ can be reasonably approximated by exponential distributions (see Fig. 10), making them well-suited for inference using our chosen flavor of variational Bayes.

Having computed basis functions $\phi_k$ and the parameters $\lambda_k$ of the coefficient distributions, we obtain the following prior distribution for reflectance:

$$p(f) = \prod_{k=1}^{K} \lambda_k \exp(-\lambda_k f_k), \quad f_k \geq 0.$$
5.3 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 ($\gamma$) and a crude model for noise:

$$I_j = \gamma \int_{\Omega} L_j(\omega) V_j(\omega) F_j(\omega) \cos \theta_i d\omega + \epsilon,$$

(13)

with $\epsilon \sim N(0, \sigma^2)$. The role of the exposure parameter is to compensate 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 both the lighting and reflectance have been learned from normalized data, while the intensity measurements may be at an arbitrary scale.

Substituting $L = \sum m \psi_m$ and $F = \sum k \phi_k$ into this expression, one can re-write this as (see Appendix C):

$$I_j = \gamma \ell^T M_j f + \epsilon,$$

(14)

where the per-pixel matrices $M_j$ depend on the surface shape (visibility and surface normal), the known 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 from this point on.

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 $I$:

$$p(I|\ell, f, \sigma, \gamma) = \prod_j \frac{1}{\sqrt{2\pi} \sigma} \exp \left( -\frac{(I_j - \gamma \ell^T M_j f)^2}{2} \right).$$

(15)

Note that the exposure and noise variance ($\gamma, \sigma^2$) are treated as model parameters to be estimated along with illumination and reflectance. For these, we define prior distributions $p(\sigma^{-2}) \sim \Gamma(\alpha, \beta)$ and $p(\gamma) \sim \text{Exp}(\lambda)$. 

5.4 Inference

The definitions of the previous sections (Eqs. 11, 12, 15, and the noise and exposure priors) provide everything...
we need to write the posterior
\[ p(\ell, f, \gamma, \sigma^{-2}|I) \propto p(I|\ell, f, \gamma, \sigma^{-2})p(\ell)p(f)p(\gamma)p(\sigma^{-2}). \]  
(16)

As described above, we wish to marginalize over lighting (as well as noise, and exposure) and compute the mean of the marginalized posterior. Following Miskin (2000) and Miskin and MacKay (2001), we do this by approximating the posterior with a separable function, \( p(\theta|I) \approx q(\theta) = q(\ell)q(f)q(\sigma^{-2})q(\gamma) \), where \( \theta \triangleq (\ell, f, \sigma^{-2}, \gamma) \). The function \( q(\theta) \) is estimated by minimizing the following cost based on the Kullback-Leibler divergence between it and the posterior (Miskin, 2000; Miskin and MacKay, 2001):
\[ C_{KL} = \int q(\theta)(\log q(\ell) + \log q(f) + \log q(\sigma^{-2}) + \log q(\gamma))d\theta \]
\[ + \log \frac{q(\gamma)}{p(\gamma)} - \log p(I|\theta)d\theta \]  
(17)

We provide an overview of the optimization strategy here, and details can be found in the appendices. The basic idea is to use coordinate descent, whereby each distribution \( q(\cdot) \) in the ensemble is updated using the current estimates of the others. The update equations are derived by carrying out all of the integrations but one in Eq. (17) (the one containing \( q(\ell) \), 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 forms
\[ q(f) = \prod q_k(f_k), \text{ with } q_k \sim N_R(u_k, w_k), \]  
(18)
\[ q(\ell) = \prod q_m(\ell_m), \]  
(19)
\[ \text{with } q_m \sim \begin{cases} N(u_m, w_m) & \text{if } m \neq 3 \\ N_{RC}(u_m, w_m, T) & \text{otherwise}, \end{cases} \]  
(20)
\[ q(\gamma) \sim N_R(u_\gamma, w_\gamma), \]  
(21)
\[ q(\sigma^{-2}) \sim \Gamma(\sigma^{-2}; a_p, b_p), \]

and it provides closed form expressions for the updated parameters of each approximating distribution in terms of the current parameters of the others (see Appendix E). \( N_R \) corresponds to a Gaussian distribution rectified at 0 and \( N_{RC} \) to a Gaussian distribution rectified at \( T \) (see Appendix B). One strategy, then, is to cycle through these distributions, updating each in turn. But as described in (Miskin, 2000; Miskin and MacKay, 2001), convergence can be accelerated by updating all parameters in parallel and then performing a line search between the current and updated parameter-sets.

### Algorithm 1 Fit ensemble

\[ \phi_1^{(0)}(k) \leftarrow \frac{u_k^{(0)}}{w_k^{(0)}}, \phi_2^{(0)} \leftarrow \log \left( \frac{w_k^{(0)}}{w_k^{(0)}} \right), \phi_3^{(0)}(m) \leftarrow \frac{w_m^{(0)}}{w_m^{(0)}}, \phi_4^{(0)} \leftarrow \log \left( \frac{1}{w_m^{(0)}} \right), \phi_5^{(0)} \leftarrow \log(1), \phi_6^{(0)} \leftarrow \frac{w_k^{(0)}}{w_k^{(0)}}, \phi_7^{(0)} \leftarrow \log \left( \frac{1}{w_k^{(0)}} \right), i = 0 \]

repeat
  repeat
    \( \Phi^* = \text{Update}(\Phi^{(i)}) \) (see Appendix E)
    \[ \Delta \Phi = \Phi^* - \Phi^{(i)} \]
    \( \alpha^* = \arg\min_\alpha C_{KL}(\Phi^{(i)} + \alpha \Delta \Phi) \) (see Appendix D)
    \[ \Phi^{(i+1)} = \Phi^{(i)} + \alpha^* \Delta \Phi, \Phi_k^{(i+1)} = \Phi_k^{(i)}, i = i + 1 \]
  until \( |C_{KL}^{(i+1)} - C_{KL}^{(i)}| < 10^{-4} \)
  \[ \Phi_k^{(i)} = \Phi_k^{(i+1)} + \alpha^* \Delta \Phi_k \]
until \( ||\Phi_k^{(i)} - \Phi_k^{(i-1)}|| < 10^{-4} \)

Specifically, we define the parameters \( \{u_k^{(0)}\} \) and \( \{w_k^{(0)}\} \) corresponding to a random BRDF and lighting environment, respectively. The initial posterior variances \( \{w_m^{(0)}\} \) and \( \{u_k^{(0)}\} \) are set to relatively large values (\( 10^{-1} \)) to account for the uncertainty in our initial estimates. Exposure parameters \( \gamma \) and \( \sigma \) 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.

#### 5.5 Evaluation and Results

As before, we begin our evaluation using images synthesized with the MERL/MIT BRDF data and our collection of measured illumination environments so that 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. This means that we must ignore color (wavelength dependent) reflectance effects in this section as well, because running the algorithm separately in each color channel would produce three independent scale ambiguities that cannot be resolved without solving the color constancy problem. We infer the inference of spectral BRDFs for future research, and instead, we only perform inference on the luminance channel to recover monochrome BRDFs in this section.

Note that while we operate in grayscale, we continue to display the input and output using color in this section. 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. 13 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. 14 is not a good approximation in these regions). This results in an observation vector $I$ of length 8,600.

Each column of Fig. 11 shows a BRDF recovered from a single input image synthesized with either the St. Peter’s Basilica or Grace Cathedral environment. As before, 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 12 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 using the spheres scene. 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. As above we operate on the luminance channel of HDR images and estimate a monochrome 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 13 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). As before, 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 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 the same two limitations seen previously, and as might be expected they become more important when the lighting is unknown. First, we expect less accuracy when the input image contains significant mesostructure (e.g., green metallic) or texture because these small variations effectively increase noise. Second, as described before, 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).

6 Discussion

This paper addresses the problem of recovering homogeneous reflectance when surface shape is known and
when the illumination is natural. In cases where the illumination is known or can be measured with a light probe, our approach can be seen as providing “lightweight” reflectometry that eliminates the need for active illumination and requires minimal acquisition infrastructure. In cases where the illumination is not known, we show that we can recover reflectance information by leveraging natural lighting statistics.

One notable feature of our approach is its use of 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. It is also very different from the convolution framework (Ramamoorthi and Hanrahan, 2001), which is restricted to one-dimensional (radially-symmetric) BRDFs. In comparison to these, our approach can be applied to a much broader class of surfaces. However, one of the important things we give up in exchange for generality is the intuition provided by the convolution framework. It becomes difficult to characterize the necessary conditions for adequate illumination, and this suggests a direction for future work.

Our linear representation is based on a bivariate parameterization of BRDFs, which reduces the domain of isotropic reflectance functions from three dimensions to two. We have found this representation to provide a good balance between complexity and accuracy for the reflectometry problem, and it is likely to be useful for other applications as well. Examples that already exist include interactive rendering (Sitthi-Amorn et al, 2010) and photometric stereo (Alldrin et al, 2008).

A second notable feature of our approach is that it is based on a probabilistic generative image model. This makes it amenable to combination with other contextual cues and vision subsystems. In particular, we should 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.

While we have proposed one way to learn a linear material basis—NMF with the Euclidean metric—there is almost surely a better alternative. For example, one might consider using a perceptual metric instead (Ngan et al (e.g. 2006); Wills et al (e.g. 2009)). This would optimize representation accuracy in the image plane instead of the BRDF domain, and it may enable more effective inference. There is also a question of whether the dictionary of basis materials is best made under-complete (with dense representations) or over-complete (with sparse ones). Our choice of an under-complete representation, as learned via NMF, is driven by the currently limited availability of empirical BRDF data.
But as additional datasets become available, a sparse representation (see Aharon et al., 2006) may be worth considering.

Finally, this work suggests an approach to material recognition by which we infer explicit BRDF parameters before classification. This stands in contrast to systems that have sought material classification directly from image statistics (e.g., Dror, 2002). The latter seems to have been motivated, at least in part, by the belief that explicit reflectometry “in the wild” is simply too hard. The results in this paper suggest that this may not be the case.

Appendix

A Folding due to bilateral symmetry

For an isotropic (and reciprocal) BRDF, bilateral symmetry induces a folding of the domain, $\phi_d \rightarrow \pi - \phi_d$. To see this, recall that bilateral symmetry means that the BRDF is unchanged when the output direction $\omega_{out}$ is reflected across the incident plane (Marschner, 1998, p.10). In the figure below, the output direction $\omega_{out}$ is reflected across the incident plane to give $\omega_{out}'$. The half-vector and azimuthal difference angle prior to the reflection is $h$ and $\phi_d$, respectively, and those corresponding to the reflected output direction are $h'$ and $\phi_d'$. The vectors $h$ and $h'$ are themselves reflections about the incident plane, so the spherical triangles $\Delta nh\omega_{in}$ and $\Delta nh\omega_{in}'$ are similar, and $\phi_d + \phi_d' = \pi$. Finally, since $(n,h) = (n',h') = \cos\theta_h$ and $(\omega_{in},h) = (\omega_{in},h') = \cos\theta_d$, the angles $(\theta_h,\theta_d)$ are unchanged by the reflection, and we have $f(\theta_h,\theta_d,\phi_d) = f(\theta_h,\theta_d,\phi_d') = f(\theta_h,\theta_d,\pi - \phi_d) \forall \phi_d$.

B Probability distributions for lighting

The mixture components in the probability distributions for lighting (Eq. 11) are as follows.

\begin{align}
p_{nm} & \sim \begin{cases} 
N(0,v_{nm}), & \text{if } m \neq 3 \\
N_{RC}(\ell_m,v_{nm},T), & \text{if } m = 3
\end{cases} 
\tag{22}
\end{align}

\begin{align}
v_{nm} & = \begin{cases} 
\ell_{nr}(m), & \text{if } r(m) \in \{2,3\} \\
v_{nr}(m), & \text{if } r(m) \in \{4,5\} \\
v_{nm}, & \text{otherwise}
\end{cases} 
\tag{23}
\end{align}

\begin{align}
N_m & = \begin{cases} 
r(m), & \text{if } r(m) \in \{3,4,5\} \\
2, & \text{if } r(m) \in \{1,2\}, m \neq 3 \\
1, & \text{if } m = 3
\end{cases} 
\tag{24}
\end{align}

where $m = 3$ is the wavelet coefficient corresponding to the vertical basis element at the coarsest scale, $r(m)$ indicates the scale of coefficient $m$, and $g(m)$ indicates to which grouping (spatial location and basis element type) coefficient $m$ belongs.

In this expression, the notation $N_{RC}$ corresponds to a Gaussian distribution rectified at a given point. Letting the rectification point be $T$, say, then if $x \sim N_{RC}(u,w,T)$ we can write

\begin{align}
p(x) & = \begin{cases} 
\frac{\sqrt{2}}{\sqrt{\pi}\erfcx\left(-\frac{x-u}{2w}\right)} \exp\left(-\frac{(x-u)^2}{4w^2}\right), & \text{if } x \geq T \\
0, & \text{otherwise}
\end{cases} 
\tag{25}
\end{align}

It is related to the “standard” rectified Gaussian distribution by $x \sim N_{RC}(u,w) \leftrightarrow x \sim N_{RC}(u,w,T)$.) Furthermore, if $x \sim N_{RC}(u,w,T)$, then

\begin{align}
\langle x \rangle_{p(x)} & = u + \frac{\sqrt{2w}}{\sqrt{\pi}\erfcx\left(-\frac{u}{2w}\right)} \\
\langle x^2 \rangle_{p(x)} & = u(x) + w + T(x - u),
\end{align}

where $\langle \cdot \rangle_{p(\cdot)}$ represents the expectation under distribution $p(\cdot)$.

C Discrete rendering equation

Given linear representations, $L = \sum \ell \psi(\ell)$ and $F = \sum \ell \phi(\ell)$, we can write $L(\omega) \approx (\ell) r(\ell), \phi(\ell)$ and $F(\omega) \approx (\ell) r(\ell), \phi(\ell)$, where $\rho(\ell)$ is the delta basis in a discretized local hemisphere, $R$ is the normal-dependent linear map from the wavelet basis in domain of the octahedral map into this delta basis (see Wang et al., 2006), and $W$ is the view-dependent linear map from the NMF basis to this delta basis. We also represent the visibility in the same basis as $\langle V(\omega) \approx \sum (v) r(\ell), \phi(\ell) \rangle$, and substitute these into the rendering equation:

\begin{align}
\int_{\Omega} L(\omega) V(\omega) F(\omega) \cos \theta_d d\omega & \approx \sum_{r_1,r_2,r_3} (R\ell) r(\ell), (W)f(\ell), (v) r(\ell) C_{r_1,r_2,r_3},
\end{align}

with $C_{r_1,r_2,r_3} = \int_{\Omega} \rho(\ell) \rho(\ell), \phi(\ell) \phi(\ell) \cos \theta_d d\omega$.

\begin{align}
\int_{\Omega} L(\omega) V(\omega) F(\omega) \cos \theta_d d\omega & \approx \sum_{(\ell)} (R\ell)r(\ell), (W)f(\ell), (v) C_{\ell}, 
\end{align}

which can be written in matrix form as $\ell^T M f$, where the pixel-dependent matrix $M$ is given by $M = R^T \text{diag}(C \circ v) W$, with $\circ$ the Hadamard (or entrywise) product and $\text{diag}(\cdot)$ a square matrix with its argument along the diagonal.

D Cost function $C_{KL}$

Following Miskin (2000) and Miskin and MacKay (2001), we approximate some of the terms in Eq. 17 with their upper bounds and re-write this expression as

\begin{align}
C_{KL} = C_{KL}^{(f)} + C_{KL}^{(g^2)} + C_{KL}^{(g)} - \log p(D|\theta) > q(\theta), \tag{27}
\end{align}

where $C_{KL}^{(f)}$, $C_{KL}^{(g^2)}$, and $C_{KL}^{(g)}$ are the KL divergence terms for the functions $f$, $g^2$, and $g$, respectively.
with each term as given below.

\[
C^{(f)}_{KL} = \sum_{m=2}^{M} \left( \frac{1}{2} \log 2\pi \sigma_m - \frac{1}{2} \frac{(\ell_m - u_m)^2}{\sigma_m} \right) q_{\ell_m}(\ell_m) - \sum_{m=2}^{M} \sum_{n=1}^{N_m} \delta_{nm} \left( \log \pi_{nm} \sigma_m + \log p_{nm}(\ell_m) \right) q_{\ell_m}(\ell_m)
\]

\[
C^{(f)}_{K} = \sum_{k=1}^{K} \left( \frac{1}{2} \log 2\pi \sigma_k - \frac{1}{2} \frac{(f_k - u_k)^2}{\sigma_k} \right) q_{f_k}(f_k)
\]

\[
C^{(\sigma-2)}_{KL} = \langle \log \sigma_{\ell_m}(\ell_m) \rangle - \langle \log \sigma_{f_k}(f_k) \rangle
\]

\[
C^{(\gamma)}_{KL} = \langle \log \gamma_{\ell_m}(\ell_m) \rangle - \langle \log \gamma_{f_k}(f_k) \rangle
\]

\[
\langle \log p_D(\theta) \rangle_{q(\theta)} = -\frac{N}{2} \log 2\pi + \frac{N}{2} \langle \log \sigma^{-2} \rangle_{q(\sigma^{-2})} - \frac{1}{2} \langle \sigma^{-2} \rangle_{q(\sigma^{-2})} \sum_{i=1}^{N} \langle (\ell_i M f - I_\ell)^2 \rangle_{q(\ell,f)}
\]

\[
\frac{1}{w_k} = \frac{(\sigma^{-2})_{q(\sigma^{-2})}}{2} \sum_{j=1}^{N} C_{kj} - \lambda_k,
\]

\[
\frac{1}{w_k} = \frac{(\sigma^{-2})_{q(\sigma^{-2})}}{2} \sum_{j=1}^{N} \left( \sum_{m} \ell_m M_{j,mk} \right)^2 \langle \gamma^{-2} \rangle_{q(\gamma)} \Rightarrow \gamma,
\]

\[
\frac{1}{w_\gamma} = \frac{(\sigma^{-2})_{q(\sigma^{-2})}}{2} \sum_{j=1}^{N} \langle (\ell_i M f - I_\ell)^2 \rangle_{q(\ell,f)} - \lambda_\gamma,
\]

\[
\frac{1}{w_m} = \frac{(\sigma^{-2})_{q(\sigma^{-2})}}{2} \sum_{j=1}^{N} \langle (\ell_i M f - I_\ell)^2 \rangle_{q(\ell,f)},
\]

\[
\frac{1}{w_m} = \frac{(\sigma^{-2})_{q(\sigma^{-2})}}{2} \sum_{j=1}^{N} E_{mj} + \sum_{n=1}^{N_m} \delta_{nm} \bar{\ell}_{nm},
\]

\[
\frac{1}{w_m} = \frac{(\sigma^{-2})_{q(\sigma^{-2})}}{2} \sum_{j=1}^{N} \left( \sum_{k} M_{m,k} f_k \right)^2 \langle \gamma^{-2} \rangle_{q(\gamma)}
\]

\[
\frac{1}{v_{nm}} = \frac{(\sigma^{-2})_{q(\sigma^{-2})}}{2} \sum_{j=1}^{N} \delta_{nm} \bar{v}_{nm},
\]

with,

\[
a_p = a + \sum_{j=1}^{N} \langle \gamma M f - I_\ell \rangle^2 \rangle_{q(\ell,f)} / 2
\]

\[
b_p = b + N / 2
\]

\[
C_{kj} = 2 \left( \sum_{m} \gamma_{\ell_m} M_{j,mk} f - I_\ell \right) \left( \sum_{m} \gamma_{\ell_m} M_{j,mk} f \right) \langle \gamma_{f,k} \rangle_{q(\gamma_{f,k})}
\]

\[
E_{mj} = 2 \left( \sum_{j} \gamma_{\ell_m} M_{j,mk} f - I_\ell \right) \left( \sum_{k} \gamma_{f,k} \right) \langle \gamma_{f,m} \rangle_{q(\gamma_{f,m})}
\]

\[
\delta_{nm} = \alpha_m \exp \left( \langle \log (\pi_{nm} p_{nm}(\ell_m)) \rangle_{q(\ell_m)} \right),
\]

where \( \alpha_m \) are such that \( \sum_{m=1}^{N_m} \delta_{nm} = 1 \).

Simplified expressions for some of these expectations, as well as sample derivations of the parametric forms for the ensemble distributions \( q(\cdot) \) can be found in an associated technical report (Romeiro and Zickler, 2010b).

**Acknowledgements** The authors thank Wojciech Matusik for helpful discussions regarding the MERL/MIT BRDF database, as well as Bill Freeman and Yair Weiss for sharing their insights into genericity and bilinear inference problems. This work was supported by the Office of Naval Research through award N000140911022, by National Science Foundation Career Award IIS-0546408, and by a fellowship from the Alfred P. Sloan Foundation.

**References**


