

Reflectance and Illumination Recovery in the Wild

Stephen Lombardi, *Student Member, IEEE*, and Ko Nishino, *Member, IEEE*

Abstract—The appearance of an object in an image encodes invaluable information about that object and the surrounding scene. Inferring object reflectance and scene illumination from an image would help us decode this information: reflectance can reveal important properties about the materials composing an object; the illumination can tell us, for instance, whether the scene is indoors or outdoors. Recovering reflectance and illumination from a single image in the real world, however, is a difficult task. Real scenes illuminate objects from every visible direction and real objects vary greatly in reflectance behavior. In addition, the image formation process introduces ambiguities, like color constancy, that make reversing the process ill-posed. To address this problem, we propose a Bayesian framework for joint reflectance and illumination inference in the real world. We develop a reflectance model and priors that precisely capture the space of real-world object reflectance and a flexible illumination model that can represent real-world illumination with priors that combat the deleterious effects of image formation. We analyze the performance of our approach on a set of synthetic data and demonstrate results on real-world scenes. These contributions enable reliable reflectance and illumination inference in the real world.

Index Terms—Reflectance estimation, natural illumination estimation, real-world reflectance, DSRDF



1 INTRODUCTION

AN image is a function of several scene components: object geometry, object reflectance, and scene illumination. Inferring one or more of these components from an image can, therefore, inform us about the world. The reflectance properties of an object, for instance, can provide valuable information about the materials that make up the object (e.g., [1]). The illumination environment can tell us about the scene itself—for example, it can indicate whether a scene is indoors or outdoors. Acquiring the reflectance of an object would enable the prediction of object appearance in a novel scene, helping support appearance-based object tracking and recognition. The most critical factor in solving this problem successfully is that we must design an approach with real-world input in mind. Past methods have mostly relied on limiting assumptions that confine them to laboratory settings with specific requirements (e.g., point light sources and Lambertian reflection).

In this work, we investigate the problem of estimating object reflectance and scene illumination from a single image given the geometry of that object. We would like to solve this problem for images taken in the wild—images of real-world objects taken under natural, complex illumination. For this, we make as few limiting assumptions about reflectance and illumination as possible: we don't assume that reflectance can be accurately represented by simple models like Lambert's law [2] or Torrance-Sparrow [3] and we don't assume that illumination can be modeled by a small set of point lights or a

small set of generic linear bases. Rather, we will leverage highly expressive reflectance and illumination models that do not constrain us to simple synthetic scenes. This added flexibility significantly complicates the problem by introducing many additional variables. To address this difficulty, we must analyze the natural variation of real-world reflectance and the effect of reflectance on the illumination environment.

Reflectance and illumination separately contribute to the difficulty of the joint inference problem. Fitting a reflectance function from a single image is difficult even when the illumination is known because a single image only reveals a small fraction of the full reflectance function [4]. An inference approach must therefore be able to sensibly extrapolate this unseen information. The addition of unknown natural illumination further complicates the problem because of the sheer number of variables introduced. Many approaches assume that natural illumination can be modeled with a low-order parametric model like spherical harmonics or a small number of point light sources. These models fail to accurately predict scene irradiance when the reflectance model is not Lambertian. A full-color non-parametric illumination model representing the sphere of incident illumination is necessary to model real-world scenes.

Reflectance and illumination themselves are not the only problem—their interaction compounds the difficulty. Several ambiguities exist between reflectance and illumination caused by the image formation process that thwart inference algorithms. One is a bilinear ambiguity between the magnitude of the reflectance and the illumination. The consequence of the ambiguity is that multiplying the reflectance by a scale factor and dividing the illumination by that same scale factor will

• The authors are with the Department of Computing, Drexel University, Philadelphia, PA, 19104.
E-mail: {sal64, kon}@drexel.edu

produce the same image as the unscaled reflectance and illumination. The problem is compounded by the fact that the bilinear ambiguity exists among each color channel independently and therefore gives rise to the color constancy problem. If we observe an object in an image, we cannot be sure whether the color of that object is due to the reflectance function, the incident illumination, or some combination of both. An ambiguity also exists between the specularity of the reflectance function and the sharpness of the illumination map, as first noted by Ramamoorthi and Hanrahan [5]. This means that if we increase the specularity of the reflectance function and blur the illumination environment, we would produce the same image as the original reflectance and illumination. Consequently, a trivial solution to the reflectance and illumination inference problem is a perfect mirror reflectance function. Careful use of prior knowledge is required to resolve these ambiguities.

In this paper, we develop a framework for reflectance and illumination inference from a single image. Our approach incorporates an expressive yet low-dimensional reflectance model. This is achieved through the use of a flexible bidirectional reflectance distribution function (BRDF) model with a data-driven component that permits compactness. We augment the reflectance model with a data-driven prior that enables intelligent extrapolation of unobserved portions of the reflectance function. In addition, we use a flexible illumination representation with simple but powerful priors that enable tractable illumination inference. We combine these elements using a Bayesian framework that allows for canonical inference by means of maximum a posteriori (MAP) estimation.

The reflectance model is based on the Directional Statistics BRDF (DSBRDF) model originally introduced by Nishino [6] and later extended by Nishino and Lombardi [7]. The DSBRDF represents reflectance as a sum of “lobes” that are each written as a directional statistics distribution in the half vector BRDF parameterization [8]. We further enhance the model by separating color and intensity from each reflectance lobe. This separation allows us to develop an additional reflectance prior that captures the joint variation of lobe color. We also extend the reflectance prior introduced by Lombardi and Nishino [9] by modeling the distribution of DSBRDF coefficients as a mixture of Gaussians. These extensions allow for a more expressive yet compact modeling of real-world reflectance.

Our illumination model is a non-parametric representation of the incident illumination field. A non-parametric model allows for a great deal of expressibility that must be constrained in order to reduce the solution space. We use several priors to do this. First, we adopt the entropy prior introduced by Lombardi and Nishino [10] that models the entropy loss of the illumination due to its interaction with the object reflectance. We also utilize a natural image statistic prior that encourages a plausible natural illumination environment to be recovered.

We thoroughly evaluate the effectiveness of our Bayesian joint estimation with synthetic and real images. First, we show how our reflectance model can accurately express a wide variety of real, measured reflectance functions from the MERL BRDF database [11]. Next, we demonstrate the ability of our model to successfully infer reflectance and illumination in a number of synthetic scenes. We quantitatively evaluate results on synthetic scenes by computing the log-space root-mean-square error between the ground truth and inferred reflectance and illumination. We qualitatively discuss results on real scenes and compare inferred illumination to the ground truth. In the end, we show that the key features of the model greatly ease this difficult inference problem.

2 RELATED WORK

The problem of inferring reflectance and other scene properties from an image has received considerable attention in computer vision. Photometric stereo [12] and shape-from-shading [13] are two early algorithms that attempt to solve an instance of this problem. Many early methods, however, impose simplistic assumptions about objects (e.g., that they exhibit ideal Lambertian reflectance) that inherently limit their applicability. Although many of these restrictions have been relaxed as the body of work has grown, there is still a need for methods that function effectively in the wild.

It is commonly assumed in the literature that materials exhibit simple reflectance behavior. Ikeuchi and Sato [14] propose a method to recover diffuse (Lambertian) and specular (Torran-Sparrow) reflectance from known geometry under a single unknown point light source. Although this is done from a single view, the reflectance model is unable to capture the visual appearance of many real-world materials. Sato et al. [15] recover spatially-varying reflectance properties from the full geometry and image intensity from multiple views. This work also uses a Lambertian plus Torran-Sparrow reflectance and a point-light illumination model that limits its applicability.

Advanced models of reflectance have been developed that can capture real-world behavior more accurately. Properly measuring a real-world reflectance function requires taking many different angular samples of both the incident and exitant directions of light. Numerous methods have been developed to do this. Ward [16] develops a curved mirror apparatus to easily capture multiple incident light directions at once. This apparatus is used to measure real-world reflectance functions and fit them to a parametric anisotropic BRDF model. Marschner et al. [17] move the camera around a curved object illuminated by a point light to measure the reflectance function. This method is primarily designed for BRDF measurement in laboratory settings. Dana et al. [18] extend the concept of a BRDF by including the spatial domain, giving rise to the bidirectional texture function (BTF). They use a robotic arm to automatically rotate material samples to

sufficiently cover the angular domain of the reflectance function. Matusik et al. [11] introduce an image-based BRDF capture apparatus that relies on curved surfaces to provide many surface normals so that only a light source must be moved. They construct a novel BRDF model that leverages a large set of measured reflectance functions to create non-parametric data-driven reflectance bases. These developments have made it possible to capture a large number of real-world reflectance functions from an image easily. Unfortunately, these methods require a specialized apparatus in a controlled laboratory environment to acquire reflectance data and many observations to properly acquire a BRDF.

Many authors have adopted sophisticated reflectance models to recover object reflectance outside the laboratory. Zickler et al. [19] develop a method to measure spatially-varying reflectance from a small set of images. This work uses a non-parametric BRDF model but overcomes the need for many reflectance samples by “sharing” observations between the spatial and angular domain. Lombardi and Nishino [9] estimate spatially-varying reflectance using a parametric but flexible reflectance model. The reflectance model used is a key contribution, as it provides a low-dimensional representation that is still able to capture real-world BRDF behavior, but the illumination is assumed to be a point light. Although these methods use more complex models for reflectance, expressive illumination models are also necessary for the real world.

Some past works have explored sophisticated illumination models for real-world radiometric scene property recovery. Marschner and Greenberg [20] estimate a lighting distribution using a sum of basis functions from a single image, but assume Lambertian reflectance. Nishino et al. [21] recover a Lambertian plus Torrance-Sparrow reflectance model under an unknown lighting distribution from a small set of images. Hara et al. derive a spherical Torrance-Sparrow reflection model to jointly estimate multiple point sources and the reflectance parameters through mixture modeling on a unit sphere [22]. The main limitation of this work is the use of polarization filters to manually separate specular highlights. Barron and Malik [23] construct a complete framework for joint spatially-varying reflectance, spatially-varying illumination, and geometry estimation. In this work, the illumination is modeled using spherical harmonics, but the reflectance model is simply Lambertian. For recovering reflectance and illumination in the wild, we need to represent both reflectance and illumination in a general way.

Ramamoorthi and Hanrahan [5] introduced a signal-processing framework for unconstrained reflectance and illumination environments. This work analyzed the theoretical ambiguities that exist between the reflectance and illumination by representing both with spherical harmonics and enumerated the situations under which they cannot be separated. A major practical concern of this work, however, is the large number of input images

required and the limited expressiveness of generic bases like spherical harmonics that require excessively many coefficients to express high-frequency illumination and reflectance (e.g., strong directional light in a scene and specularities, respectively).

Romeiro and Zickler [24] introduced a practical method to estimate real-world reflectance functions under known natural illumination from a single image. This method is able to effectively tackle real-world scenes by using a non-parametric data-driven BRDF model. Later, Romeiro and Zickler extended this work to *unknown* natural illumination [25]. The main drawbacks of their method are that the illumination environment is not inferred but instead marginalized out and that the reflectance and illumination are actually monochrome (color is added later as a post-processing step). The illumination environment itself conveys a wealth of information as it is the scene itself surrounding the object; inferring it jointly with reflectance is therefore useful.

Preliminary work of our method appeared in [10] which introduced a method for solving this problem fully (i.e., a method for recovering real-world reflectance functions and natural illumination from a single image with known geometry). The development underpinning this method is the DSBDRF [7], a low-order parametric BRDF model capable of representing a great deal of real-world reflectance functions. The DSBDRF is highly expressive because it represents the BRDF as a sum of reflectance “lobes” that are written as a probability distribution using the half vector parameterization of the BRDF [8]. It achieves compactness due to a set of data-driven basis functions that capture common modes of variation among real-world BRDFs. Combined with a novel entropy prior on the natural illumination, this method can enable the estimation of real-world reflectance and illumination. In this paper, we expand this work by enhancing the reflectance model and priors and by more deeply analyzing the model.

3 MODELING MATERIAL REFLECTANCE

There is often a trade off between the number of parameters of a reflectance model and its expressiveness (i.e., the variety of reflectance it can model). We want a model that is both extremely expressive (i.e., one that can model most real-world objects) and low-dimensional (so that it can be reliably estimated). We also want the model to be amenable to constraints that help to encourage reflectance estimation to take on only plausible values.

3.1 Directional Statistics BRDF

We adopt the Directional Statistics Bidirectional Reflectance Distribution Function (DSBRDF) model [6], [7] that has these desirable properties. The DSBDRF model is able to express an extremely wide variety of measured material reflectance functions while being compact due to a set of data-driven reflectance bases.

The DSBDRF model uses the halfway vector parameterization introduced by Rusinkiewicz [8]. This parameterization is centered around the half vector (i.e., the vector between the viewing direction and the lighting direction) where θ_d is the angle between the viewing direction and the half vector and θ_h is the angle between the half vector and the surface normal. We write the DSBDRF as a sum of reflectance “lobes,” each a statistical distribution over the hemispherical domain of θ_h ,

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

Here κ controls the overall brightness of the lobes and γ controls the specularity. In practice, we represent κ and γ with degree two B-splines with nine knots [7]. Note that there is a unique κ and γ for each color channel λ and lobe r . Most important, κ and γ are functions of θ_d —this gives the model its expressiveness [6] [7].

A BRDF represented with the DSBDRF model can now be thought of as a set of (κ, γ) -curves, one for each lobe. The (κ, γ) -curves control the behavior of the reflectance function across different slices of the BRDF. Capturing the variation of these curves is the key to achieving a compact but expressive model.

Because (κ, γ) -curves give a complete description of the BRDF, we can fit them to a set of measured reflectance data (such as the MERL BRDF database [11]). After acquiring fits for a set of measured BRDFs, we use functional principal component analysis (FPCA) on the (κ, γ) -curves to compute a set of reflectance bases b_i [7]. We then represent the (κ, γ) -curves using a log-linear combination of the bases:

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

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

where b_μ is the mean function, b_i are the learned basis functions, and ψ_i are the DSBDRF coefficients. By using a subset of the basis functions, we can trade off accuracy and compactness.

3.2 Modeling Color Explicitly

The reflectance parameter Ψ controls the intensity of each reflectance lobe and color channel jointly. While a single set of parameters is an elegant solution for modeling reflectance, we can improve the representation by modeling the color of each lobe explicitly. Doing this allows us to place separate, specialized priors on the color and luminance of the BRDF.

We represent the BRDF color by introducing a chromaticity vector for each lobe. The chromaticity, c , modulates the color of each lobe,

$$\varrho_\lambda(\theta_d, \theta_h; \kappa, \gamma, c) = \sum_r c_{r,\lambda} \left(\exp \left[\kappa_{r,\lambda}(\theta_d) \cos^{\gamma_{r,\lambda}(\theta_d)} \theta_h \right] - 1 \right), \quad (4)$$

where λ is the color channel, and with the constraint that $\sum_\lambda c = 1$, which ensures that the chromaticity vector does not influence the overall luminance of the lobe. This removes the dependency of κ and γ on the color channel λ . The separation is key to more tightly modeling real-world reflectance.

Figure 1 compares the fits of several variants of the DSBDRF model with measured reflectance functions. The variants of the DSBDRF shown are the (κ, γ) -curves modeled as B-splines, (κ, γ) -curves modeled as a combination of the learned bases, and the color-separated model presented in this work. The figure illustrates how the color-separated model maintains expressiveness while drastically increasing compactness. These careful steps to lower the dimensionality of the model but preserve the quality are essential to the framework.

Figure 2 shows a comparison of the modeling error of BRDFs fit to measured data from the MERL database [11]. We compare the DSBDRF model (with color separation) to the bivariate model proposed by Romeiro and Zickler [24] and a simple Lambertian plus Cook-Torrance BRDF [26]. Here we measure accuracy using root-mean-square error in log-space,

$$E_{\ell\text{-RMS}} = \sqrt{\sum_{\theta_h, \theta_d, \phi_d} \frac{(\log f(\theta_h, \theta_d, \phi_d) - \log \varrho(\theta_h, \theta_d))^2}{N}}, \quad (5)$$

between the measured BRDF $f(\theta_h, \theta_d, \phi_d)$ and the fit BRDF $\varrho(\theta_h, \theta_d)$. As noted by [11], when comparing two BRDFs it is critical to take the log because the specular peaks of the BRDFs typically have values orders of magnitude greater than the diffuse portions. The bivariate BRDF model is a non-parametric representation that depends only on θ_d and θ_h and is computed by averaging over ϕ_d . The figure shows that the DSBDRF model can accurately model real-world BRDFs with less free parameters than a combination of Lambert’s Law and Cook-Torrance.

3.3 Reflectance Priors

With color separated from the rest of the model, we can develop more specialized priors. We first place a prior on the (κ, γ) -curves. In the earlier DSBDRF work [9], [10], a single Gaussian distribution was used to model the distribution of basis coefficients. We extend this prior by utilizing a mixture of Gaussians,

$$p(\Psi) = \sum_i \pi_i N(\Psi | \mu_i, \Sigma_i). \quad (6)$$

This extension provides a tighter fit to the data without sacrificing generality.

Figure 3 shows a visualization of the prior overlaid onto the DSBDRF space. The DSBDRF space is visualized by plotting the projections of each BRDF onto the first two basis functions b_0, b_1 . We overlay the ellipses of the Gaussian mixture to observe how it models the space of BRDFs. We can see that the mixture of Gaussians is

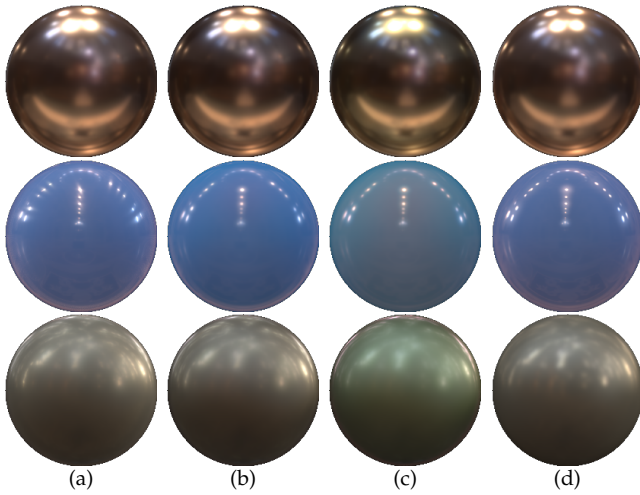


Fig. 1. Fitting variants of the DSBDRF model to measured BRDF data. This figure illustrates how we arrive at a compact analytical reflectance model while retaining expressiveness. Column (a) is a ground-truth rendering of the MERL BRDF in three different illumination environments. Column (b) shows renderings of the DSBDRF model with (κ, γ) -curves represented as B-splines with color integrated into each lobe representation [7]. This model has three colors, three lobes, and six parameters per B-spline for a total of 108 free variables. Column (c) shows renderings of the DSBDRF model with (κ, γ) -curves represented with the learned bases b_i truncated at 16 parameters with color integrated into each lobe representation. This model has 16 free variables. Column (d) shows renderings of the DSBDRF model with color represented explicitly for each lobe (see main text). This model uses 10 basis coefficients and 2 variables per lobe for chromaticity for a total of 16 free variables. From the figure we can see qualitatively that the DSBDRF model with color represented separately has expressiveness comparable to the “full” DSBDRF model (Column (b)) but with only 16 parameters.

able to naturally identify different types of reflectance functions. For example, it captures primarily diffuse reflectance functions in one cluster, many shiny metals and plastics in another cluster, and those in-between in the third cluster. This observation supports the use of a mixture model as a distribution on likely reflectance parameters.

We can now place a novel prior on the chromaticity vectors. The most important variation we should capture is the variation in color among the lobes—that is, the joint variation of lobe chromaticity. We observe that the chromaticity vectors in a BRDF will often have similar hue but will not necessarily have similar saturation. We use this observation to construct a simple prior distribution that quantifies the relationship of lobe chromaticity. This is difficult to do, however, as there is no obvious reparameterization of chromaticity into “hue” and “saturation”. We must therefore derive a hue representation.

Hue is typically encoded as a rotation around the center of a color triangle. Chromaticity is represented with Barycentric coordinates so it would be unnatural to try to compute this rotation as if it were a Euclidean space. Instead, we will use a function that gives a vector that is invariant to saturation. First, consider the function,

$$g_i(\mathbf{x}, \alpha) = \frac{x_i^\alpha}{\sum_{i'} x_{i'}^\alpha}, \quad (7)$$

where $\mathbf{x}$ is a Barycentric coordinate. As α grows, the point $\mathbf{x}$ will move from the center of the color triangle toward the nearest corner. The derivative of g with respect to α at $\alpha = 0$,

$$\mathbf{g}'(\mathbf{x}, 0) \propto \mathcal{L} \log \mathbf{x}, \quad (8)$$

where

$$\mathcal{L} = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}, \quad (9)$$

will give a vector whose direction can be interpreted as “hue” and whose magnitude can be interpreted as “saturation”. Normalizing it will then give us “hue”:

$$\mathbf{h}(\mathbf{x}) = \frac{\mathcal{L} \log \mathbf{x}}{\|\mathcal{L} \log \mathbf{x}\|}. \quad (10)$$

Given that the hue vector lies on the unit circle, it is a natural choice to place a von Mises distribution [27] on it. In order to address the relationship of lobe chromaticity vectors in a BRDF, we will construct a prior distribution for pairs of lobe chromaticity vectors. The prior distribution for the chromaticity of lobe r given the chromaticity of lobe r' is,

$$p(\mathbf{h}(c_r) | \mathbf{h}(c_{r'})) \propto \exp \left[\kappa_h \mathbf{h}^\top(c_{r'}) \mathbf{h}(c_r) \right], \quad (11)$$

where κ_h is the concentration parameter. Assuming there are three conditionally independent lobes and assuming a uniform marginal distribution, we can write the joint distribution for $\mathbf{c}$,

$$p(\mathbf{c}) = p(\mathbf{h}(c_2) | \mathbf{h}(c_1)) p(\mathbf{h}(c_3) | \mathbf{h}(c_1)). \quad (12)$$

Intuitively, this prior encourages lobes to have similar hues regardless of saturation.

Figure 4 visualizes the distribution of chromaticities among the lobes in several BRDFs and gives empirical evidence for our prior. For these example BRDFs, we also illustrate the path of the Eq. 7 as α is varied from zero to infinity. Any two lobes having the same dashed path will have the same chromaticity hue vector. Also shown is the distribution of the angles between chromaticity hue vectors for each unique pair of lobes for all MERL BRDFs. This data is used to fit the parameter κ_h for the von Mises prior distribution. We can see that the distribution is peaked where the hue vectors are pointing in the same direction, indicating a strong correlation of hue among lobes.

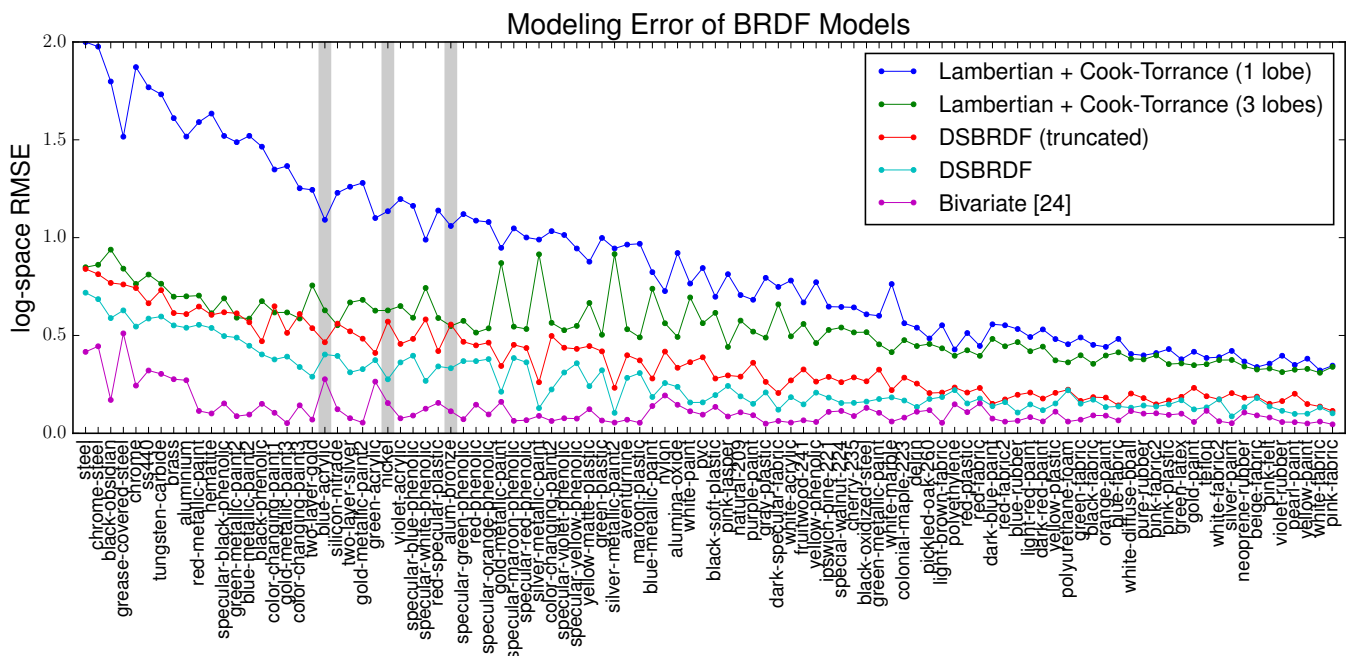


Fig. 2. Comparison of the DSBRDF model to the non-parametric bivariate model [24] and Cook-Torrance [26]. This figure shows the log-space RMSE of fitting Lambertian and 1 lobe of Cook-Torrance (10 free parameters), Lambertian and 3 lobes of Cook-Torrance (24 free parameters), the DSBRDF with color separation modeled with a small number of learned bases (13 free parameters), the full DSBRDF model with color separation (42 free parameters), and the non-parametric bivariate BRDF (24,300 parameters). The vertical grey bars highlight the BRDFs used in Figure 1. The figure demonstrates that the DSBRDF model accurately captures real-world reflectance functions with a low-order parameterization.

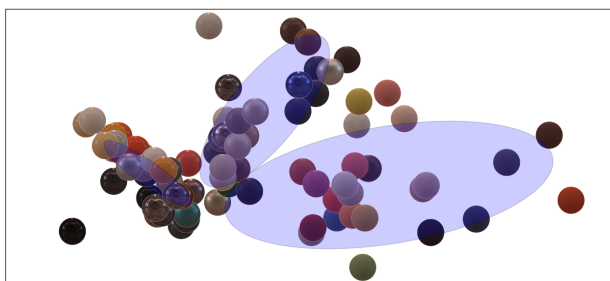


Fig. 3. Reflectance prior overlaid onto a 2D slice of the DSBRDF space. The BRDFs of the MERL database are visualized by the first two eigenfunction coefficients in the DSBRDF model. The ellipses of the Gaussian mixture are shown. We can see that the mixture of Gaussians is a good model for this distribution.

4 MODELING NATURAL ILLUMINATION

Handling natural illumination is essential to enable radiometric scene property inference from real-world images. Real-world scenes cannot be adequately represented by simple models like a small set of point lights or spherical harmonics. We must consider that a scene is illuminated by every visible direction in full color. Although this ostensibly burdens the inference problem, there is a great deal of latent structure in the natural illumination environment that we can exploit.

4.1 A Non-parametric Illumination Model

In the past there have been numerous approaches for modeling natural illumination environments. Barron and Malik use spherical harmonics as a linear basis [23]. Romeiro and Zickler model natural illumination as a linear combination of data-derived basis functions [25]. Rather than attempt to reduce the variability of the representation, we use a very expressive representation but control variability with carefully constructed priors. The priors will help overcome the ambiguities of the problem and encourage the recovery of accurate illumination maps. We represent natural illumination non-parametrically as a dense grid on the sphere parameterized by latitude and longitude which we call the illumination map $\mathbf{L} = \{L_{\theta, \phi, \lambda}\}$ for incident angle $\{\theta, \phi\}$ and color channel λ . We can think of this representation as a wide-angle image of the surrounding environment. Because of the great expressiveness afforded by this model, we must apply intelligent priors to constrain the solution space and recover sensible estimates.

4.2 Illumination Priors

As shown in previous work, reflectance acts as a band-pass filter on the illumination environment causing an ambiguity between reflectance and illumination. The BRDF acting as a bandpass filter causes a blurring of the illumination and thus a spreading of the histogram. As

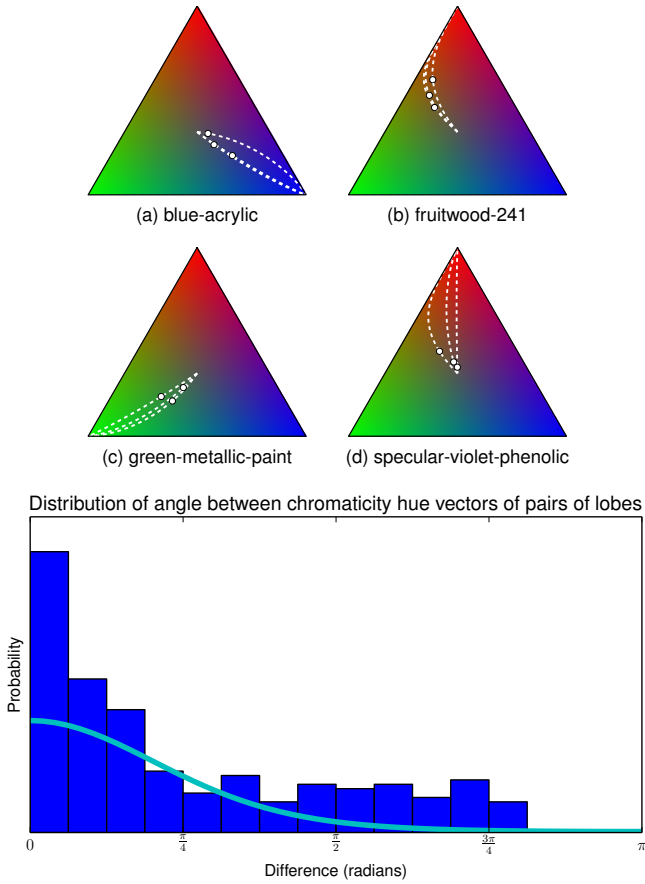


Fig. 4. Empirical support for the novel BRDF chromaticity prior. The top two rows show the lobe chromaticity values fit to four example BRDFs from the MERL database. Solid dots represent the chromaticity values. The dashed line is Eq. 7 as α is varied. The bottom plot shows the distribution of the angles between chromaticity hue vectors for each unique pair of lobes among all MERL BRDFs. We then fit a von Mises distribution to this (cyan curve). The strong correlation of lobe chromaticity hue vectors shows that the novel BRDF chromaticity prior effectively constrains the BRDF color to reflect that of real-world BRDFs.

shown by Lombardi and Nishino [10], this increases the entropy of the reflected radiance. We'd like to recover the true illumination environment and to do this we assume that the entropy increase in the observed image is due entirely to the BRDF. To this end, we constrain the illumination to have minimum entropy, so that the BRDF will be responsible for causing the increase in entropy of the outgoing radiance.

We minimize the entropy of the illumination map with an exponential distribution,

$$p_1(\mathbf{L}) = \lambda_E \exp \left[-\lambda_E E(\mathbf{L}) \right]. \quad (13)$$

We use the continuous form of entropy,

$$E(\mathbf{L}) = - \int H(x) \log H(x) dx, \quad (14)$$

so that we can use gradient-based optimization methods. We also use kernel density estimation to model the histogram of illumination map pixel intensities H for the same reason,

$$H(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{\theta, \phi, \lambda} \exp \left[-\frac{(x - L_{\theta, \phi, \lambda})^2}{2\sigma^2} \right]. \quad (15)$$

We know that the natural illumination environment can be interpreted as a wide-angle image of the scene surrounding the object. It follows, then, that we can place natural image statistics priors on the illumination environment to properly constrain it (e.g., [28]). We therefore place a simple heavy-tailed prior distribution on the gradient magnitudes of the illumination map using a hyper-Laplacian distribution,

$$p_2(\mathbf{L}) \propto \prod_{\theta, \phi} \exp \left[-b^{-1} \left(\sqrt{\sum_{\lambda} \frac{\partial L_{\theta, \phi, \lambda}}{\partial \theta}^2 + \frac{\partial L_{\theta, \phi, \lambda}}{\partial \phi}^2} \right)^\alpha \right], \quad (16)$$

with $\alpha < 1$. Note this formulation of the image gradient magnitude takes the color of the gradient into account.

This prior is important for two reasons. First, it helps promote natural image statistics in recovered illumination maps. It will also help conquer the ambiguity between the specularity of the reflectance and sharpness of the illumination map by deterring the optimization from choosing the trivial solution—a mirror-like reflectance function and a blurred illumination map. If the current estimate is the trivial solution, the soft gradients of the illumination map will cause the image-gradient prior to have very low probability.

Multiplying the two illumination priors together, we obtain the complete illumination prior,

$$p(\mathbf{L}) \propto p_1(\mathbf{L})p_2(\mathbf{L}). \quad (17)$$

Finally, we must address the color constancy problem. As mentioned, there is a bilinear ambiguity between the reflectance and illumination that exists independently in each color channel. This gives rise to the color constancy problem: we cannot know whether an observed pixel color is due to the object reflectance or the color of the illumination. To address this problem, we use the same approach taken by Lombardi and Nishino [10] by adopting the grey world assumption. In effect, the reflectance function will explain most of the color variation of the scene, leaving the illumination map to fill the gaps. To implement this in our framework, we will estimate reflectance and illumination in two passes. In the first pass, we constrain the illumination to be greyscale, forcing the reflectance function to explain the colors observed of the image. In the second pass, we use the previous results to initialize but allow illumination to estimate in full color.

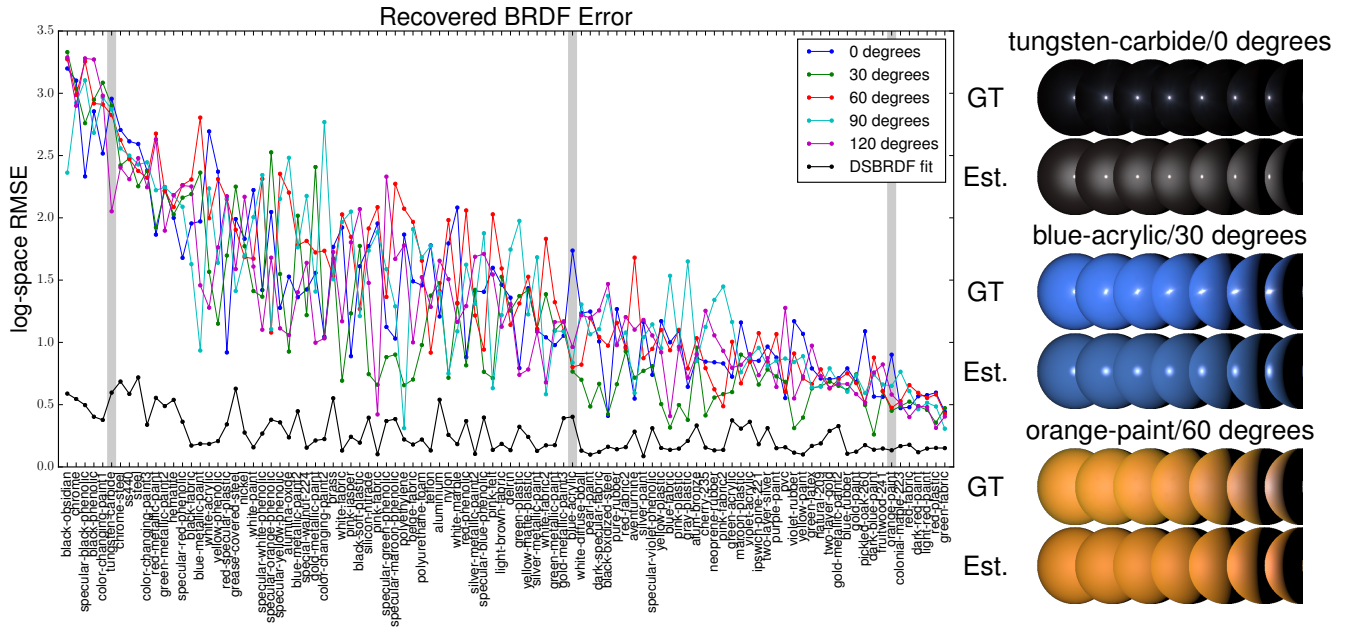


Fig. 5. Quantitative evaluation of single point light experiments and select results. On the left, this figure shows the log-space RMSE (Eq. 5) of recovered BRDFs from 500 single point light experiments. Each BRDF from the MERL database was illuminated by a point light at 0, 30, 60, 90, and 120 degrees from the viewing direction. On the right, we highlight several example results to illustrate how certain log-space RMSE values correspond to perceptual accuracy. The BRDFs on the right are visualized by rendering a sphere under several point light directions.

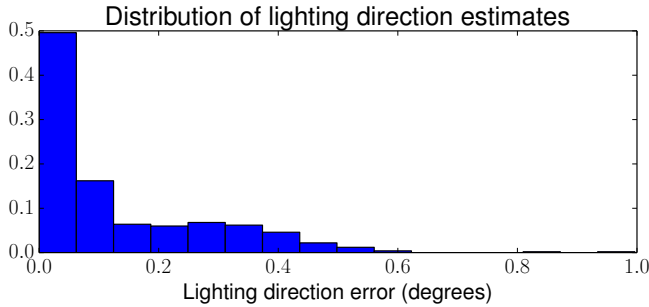


Fig. 6. Quantitative evaluation of point light direction estimation. This figure shows the distribution of point light direction error in degrees. As shown, all estimates are within a single degree of the ground-truth lighting direction.

5 INFERRING REFLECTANCE AND ILLUMINATION

At the heart of our framework is a Bayesian formulation that combines the priors and a likelihood distribution and provides an inference approach. In this work, we will perform inference by finding the maximum a posteriori (MAP) estimate of the posterior distribution,

$$p(\Psi, \mathbf{c}, \mathbf{L} | \mathbf{I}) = p(\mathbf{I} | \Psi, \mathbf{c}, \mathbf{L}) p(\Psi) p(\mathbf{c}) p(\mathbf{L}), \quad (18)$$

where Ψ are the reflectance parameters, $p(\Psi)$ is given in Eq. 6, $\mathbf{c}$ are the reflectance lobe chromaticities, $p(\mathbf{c})$ is given in Eq. 12, $\mathbf{L}$ is the illumination map, $p(\mathbf{L})$ is given in Eq. 17, and $\mathbf{I}$ is the observed image.

We write the likelihood distribution using a Laplacian distribution on the logarithm of the irradiance values,

$$p(\mathbf{I} | \Psi, \mathbf{c}, \mathbf{L}) \propto \prod_{\mathbf{x}, \lambda} \exp \left[-\beta_I \left| \log I_{\mathbf{x}, \lambda} - \log E_{\mathbf{x}, \lambda}(\Psi, \mathbf{c}, \mathbf{L}) \right| \right], \quad (19)$$

where $E_{\mathbf{x}, \lambda}(\Psi, \mathbf{c}, \mathbf{L})$ is the mean of the distribution and represents a function that computes the irradiance at pixel $\{\mathbf{x}, \lambda\}$ given parameters for reflectance $\{\Psi, \mathbf{c}\}$ and illumination $\mathbf{L}$ and β_I is a parameter controlling the scale of the distribution. We use a Laplacian distribution to model the image formation process because it allows for robust estimation [29]. It's important to place the distribution in log-space so that the algorithm is invariant to the scale of the input image.

With the complete posterior distribution specified, we can now perform inference by computing a maximum a posteriori estimate. In practice, we will minimize the negative log of the posterior with an alternating minimization approach. After initialization, we fix the reflectance and chromaticity estimates and estimate the illumination. Then, we fix the illumination estimate and estimate the reflectance and chromaticity. We alternate estimating reflectance and illumination in this fashion until convergence (determined by the difference in the negative log posterior between successive iterations), which typically takes 5 to 10 alternations. Although our priors help to keep the optimization on the right track, there are cases where the optimization gets stuck in a local minima.

We initialize our algorithm with a very straightforward

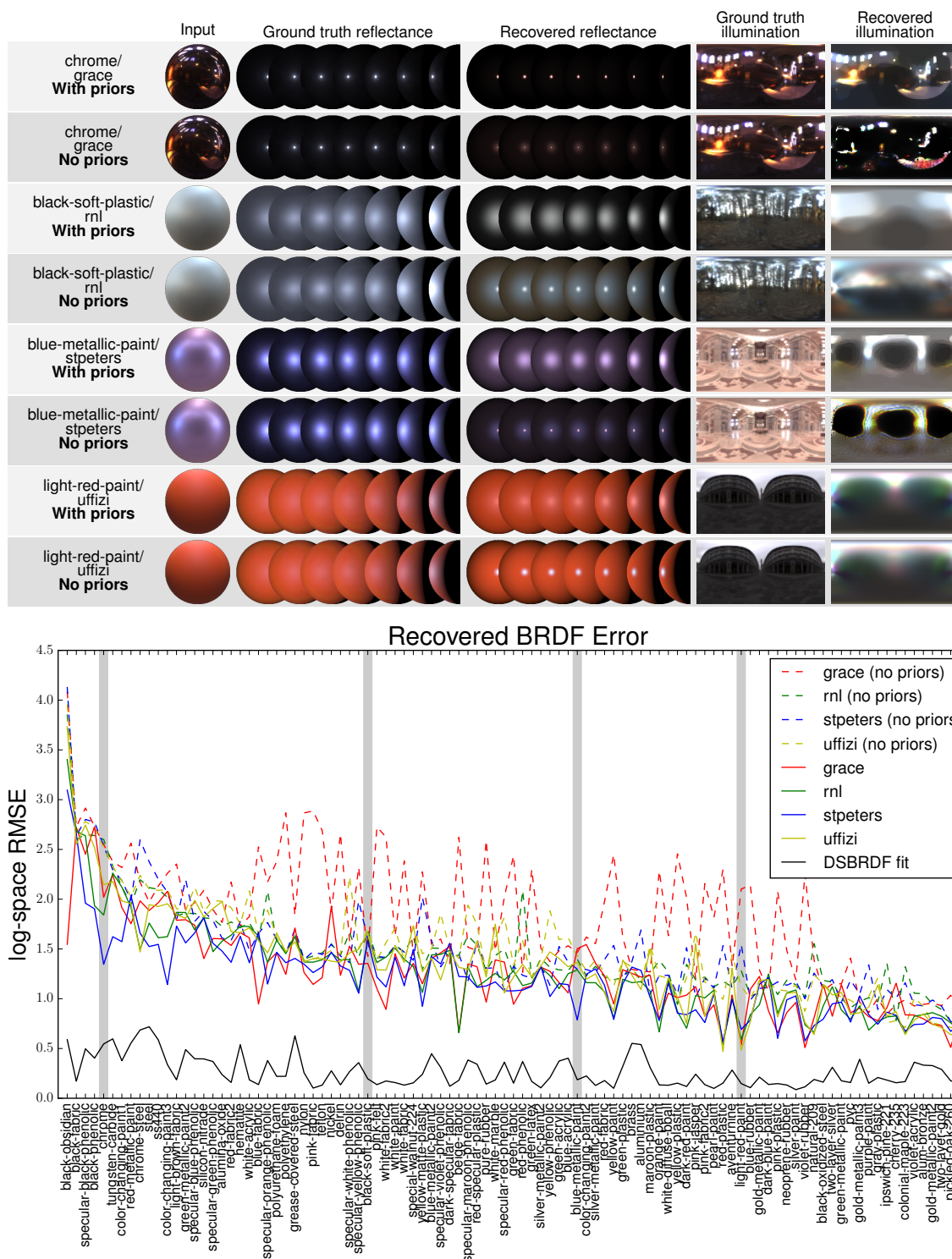


Fig. 7. Evaluation of synthetic experiments under natural illumination. The bottom plot shows the log-space RMSE for 100 MERL BRDFs under four natural illumination environments. Highlighted are eight results from that plot shown at the top. This plot demonstrates the ability of the model to successfully infer reflectance and illumination in a variety of illumination environments for many different materials.

ward approach. The reflectance, Ψ , is initialized to all zeros. This can be interpreted as the “mean BRDF” because of the data-driven bases derived from the MERL database. The chromaticity, c , is initialized to random

values. The illumination map, L , is also initialized to all zeros. We don’t rely on a complex initialization—instead, the framework is able to find good solutions even with this simple initialization.

Relationship of natural illumination accuracy versus BRDF specularity

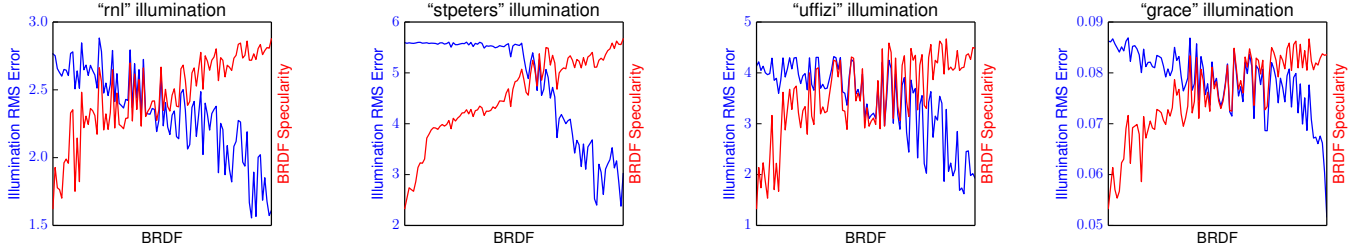


Fig. 8. Quantitative evaluation of natural illumination estimation versus BRDF specularity. This figure shows the RMS error of the natural illumination estimates in blue for 100 MERL BRDFs in four illumination environments. In red is the “specularity” of the BRDF, as measured by the γ value of the most specular lobe. This figure illustrates the general trend that highly glossy BRDFs enable more accurate illumination estimation.

6 EXPERIMENTAL RESULTS

We thoroughly evaluate the effectiveness of our method with a number of synthetic and real-world images. First, we investigate inference under a simple point light illumination model. This scenario will allow us to observe how our reflectance model and priors are explaining unseen slices of the reflectance function. Next, we explore full reflectance and natural illumination inference in synthetic scenes. Dealing with synthetic scenes allows us to ignore some confounding variables and gives us the opportunity to better analyze the results. Finally, we use our framework to estimate reflectance and natural illumination in the real world.

6.1 Synthetic images; single point light source

A simple way to evaluate our reflectance model is to set up a series of synthetic experiments under an unknown point light source. Although an unrealistic assumption in general, it will allow us to understand how well the model is able to extrapolate unseen slices of the reflectance function. This is because an object imaged under a single point light will only express a small portion of its reflectance function (i.e., θ_d is fixed). Any inference procedure that attempts to recover a full BRDF from an object illuminated by a single point light must therefore make assumptions about the other parts of the reflectance function that are not observed. In this section, we show that our model is making intelligent assumptions about the behavior of unseen BRDF slices.

To perform this evaluation, we set up 500 single point light experiments. For each of the 100 BRDFs in the MERL database [11], we render a sphere with that BRDF under five point light directions: 0, 30, 60, 90, and 120 degrees from the viewer. We then run our inference algorithm for each rendered sphere and quantitatively evaluate the estimated BRDF. Note that we exclude the BRDF being estimated in the data-driven basis and prior computations.

We use a simple procedure for estimating the point light parameters. We model the light as a point on a sphere with a non-negative intensity value. During the

inference algorithm, when the reflectance estimate is fixed, we simply use Powell’s method [30] to update the lighting parameters.

To evaluate BRDF estimates, we again use the log-space RMSE (Eq. 5). As previously discussed, there is an ambiguity in scale between the light and reflectance function that cannot be known without additional information. Therefore, there is an unknown scale factor between the recovered BRDF and the ground-truth BRDF. Because we know the ground-truth lighting intensity, we can simply invert the estimated lighting intensity to arrive at the correct scale factor. Note that this procedure is not as simple for the case of natural illumination.

Figure 5 shows quantitative results for all 500 experiments. To give these quantitative results meaning, we compare several BRDFs to ground truth as a cascaded rendering of spheres from different point light directions. This helps visualize how log-space RMSE values correspond to perceptual BRDF accuracy. From the plot we can see that specular BRDFs are more difficult to estimate. This is likely because the specular BRDFs have a greater degree of variation, especially as θ_d varies.

Figure 6 shows a quantitative evaluation of the point light estimates. As the figure shows, all estimates of the illumination direction are within one degree of the ground truth lighting.

6.2 Synthetic images; unknown, unconstrained natural illumination

While point light experiments can tell us a great deal about the models strengths, our ultimate goal is to recover reflectance and natural illumination. To do this, we will use our full framework, including the natural illumination representation and priors. In this section, we show experimental results from synthetic scenes.

To evaluate our algorithm synthetically, we created 400 experiments using 100 MERL BRDFs in four different illumination environments. Each material comes from the MERL BRDF database [11]. Illumination environments are from Paul Debevec’s Light Probe Gallery [31].

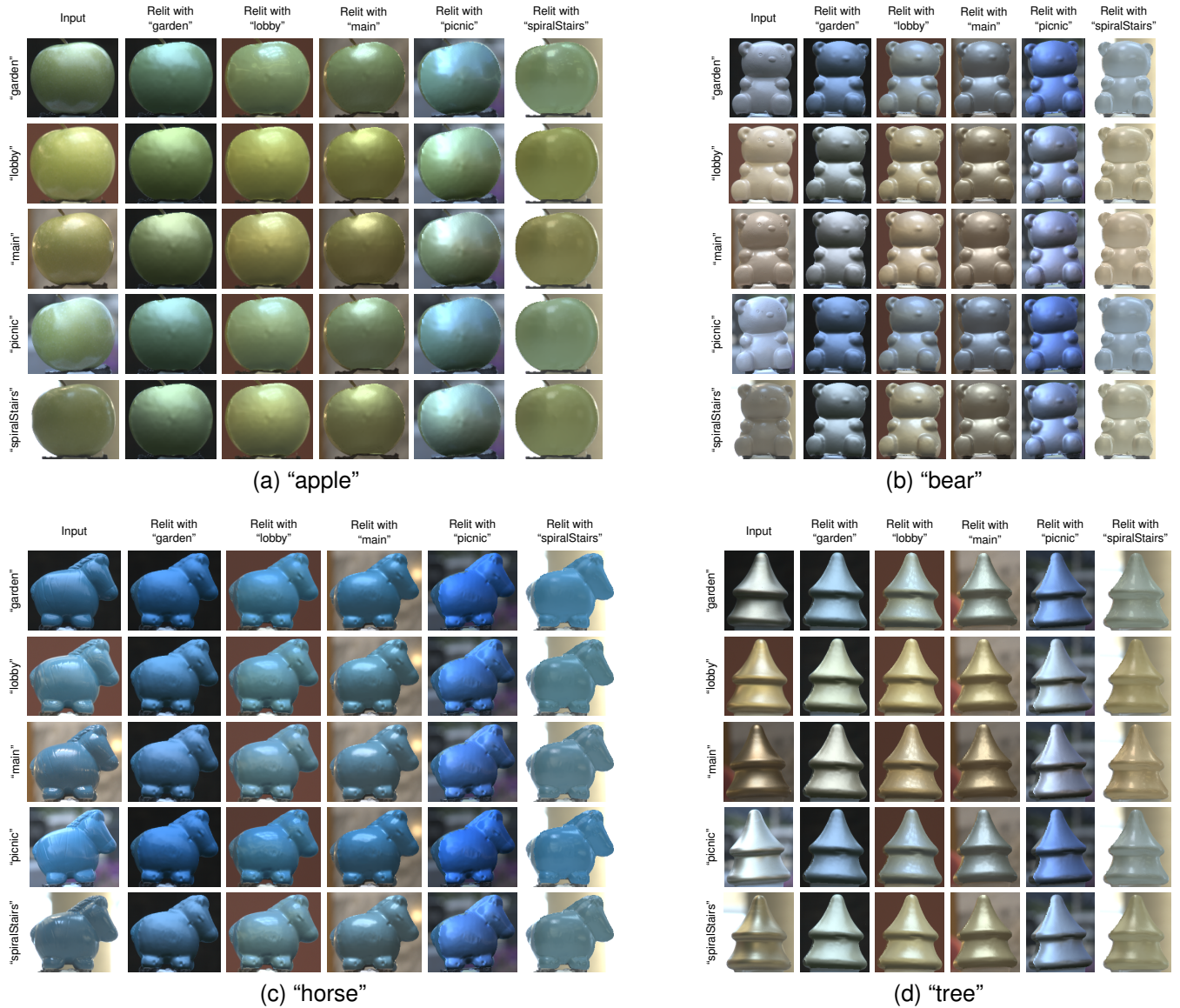


Fig. 9. Relighting comparisons. The first column in each subfigure show a photograph of an object under several different illumination environments. We recover the reflectance and illumination of the object and then use the recovered reflectance to relight the object with each illumination environment which is shown in the subsequent columns. The visual similarity of the relighting to the ground truth demonstrates the accuracy of our reflectance estimates.

We compare each recovered material and illumination estimate to the ground-truth values.

Like the single point light experiments, we use log-space RMSE to quantitatively evaluate the results. However, determining the scale difference between the recovered and ground-truth BRDF is not straightforward now that the illumination is no longer a single point. Therefore, we will use the scale factor that minimizes the log-space RMSE,

$$\hat{E}_{\ell\text{-RMS}} = \min_{\alpha} \sqrt{\sum_{\theta_h, \theta_d, \phi_d} \frac{(\log f(\theta_h, \theta_d, \phi_d) - \log \alpha \varrho(\theta_h, \theta_d))^2}{N}}, \quad (20)$$

Figure 7 shows the accuracy of the reflectance estimates compared to a baseline method. For the baseline method we simply run our method without priors. In general, the proposed priors provide a large qualitative improvement over the baseline method, even when the quantitative improvement is modest. For example, “black-soft-plastic/rnl” “with priors” slightly outperforms “no priors” quantitatively yet the qualitative results are compelling: the specular lobes have the right size and shape, whereas with “no priors” the reflectance is overly-specular. The overly-specular results when priors are disabled can also be seen in the “blue-metallic-paint/stpeters” and “light-red-paint/uffizi” examples, which have a larger difference in quantitative error, and

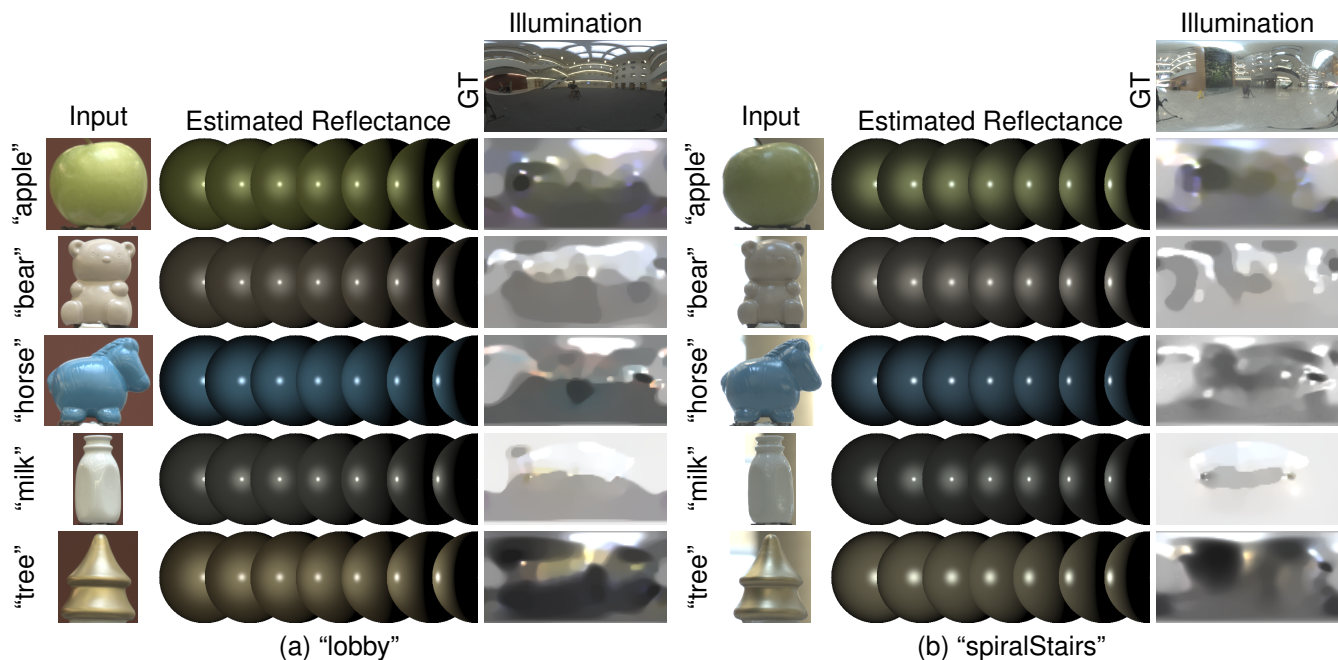


Fig. 10. Results on “lobby” and “spiralStairs” illumination environments. Each row shows the recovered reflectance and illumination of a different object with the ground-truth illumination for comparison. Note how features of the reflectance function are accurately captured: the “apple”, “bear”, “horse”, and “milk” objects are shiny and the recovered reflectance function has sharp specular highlights; the “tree” object has a softer glossy reflection that is captured in the reflectance estimate.

show that our priors are working as intended.

Figure 8 examines the relationship between the specular of BRDFs and the accuracy of the estimated illumination. We can see that, in general, more specular BRDFs allow for more accurate natural illumination estimates. This makes sense, because less information is destroyed during the image formation process for specular BRDFs than diffuse. This relationship is especially prevalent in the “rnl” and “uffizi” lighting environments. Figure 7 also illustrates this effect.

6.3 Real scenes

The most challenging scenes—and the ultimate goal of this paper—are ones in the real world. They may present significant global illumination effects, camera noise, and other unique problems. Despite these great difficulties, we show that our model is able to accurately capture reflectance and illumination from a single image.

We use the Drexel Objects under Natural Illumination database [10] to evaluate our model. This database includes up to 6 objects in 5 environments, each captured with high dynamic range, and includes ground-truth geometry and illumination.

Figure 9 examines the accuracy of the reflectance estimates through a series of relightings. For each image in the data set, we use the reflectance recovered by our method to relight the object using the ground truth illumination of each environment. In this way, we can visually compare reflectance results to the ground-truth

images. This figure shows many compelling results, including those of the “bear” object and the “tree” object. It can be seen that the method is able to recover reflectance and illumination in a wide variety of scenes and it can be effectively used to predict object appearance in novel environments.

Figures 10 and 11 show real-world results for each illumination environment. In each environment, we can see that our algorithm recovers plausible reflectance estimates through a cascaded rendering of spheres. For example, in Figure 10, note the diffuse highlight of the “tree” object. In addition, many important features of the natural illumination environments are recovered. For example, in Figure 11 (b), many illumination estimates (especially from the horse object) capture the position and detail of the skylight that is casting much of the light in the scene. Also note the “milk” object in each figure, which does not contain a full hemisphere of surface normals. In these cases, unseen portions of the lighting environment are assumed to be black. Overall, the objects that allow the best recovery of the illumination environment seem to be the “apple”, “bear”, and “horse”. This is likely because the surface normals of these objects span the entire hemisphere facing the camera, and the objects are mostly concave so they do not cause significant inter-reflection.

Figure 12 shows a direct comparison to the method of Romeiro and Zickler [25] on the real-world dataset captured by the authors. As shown our method can more

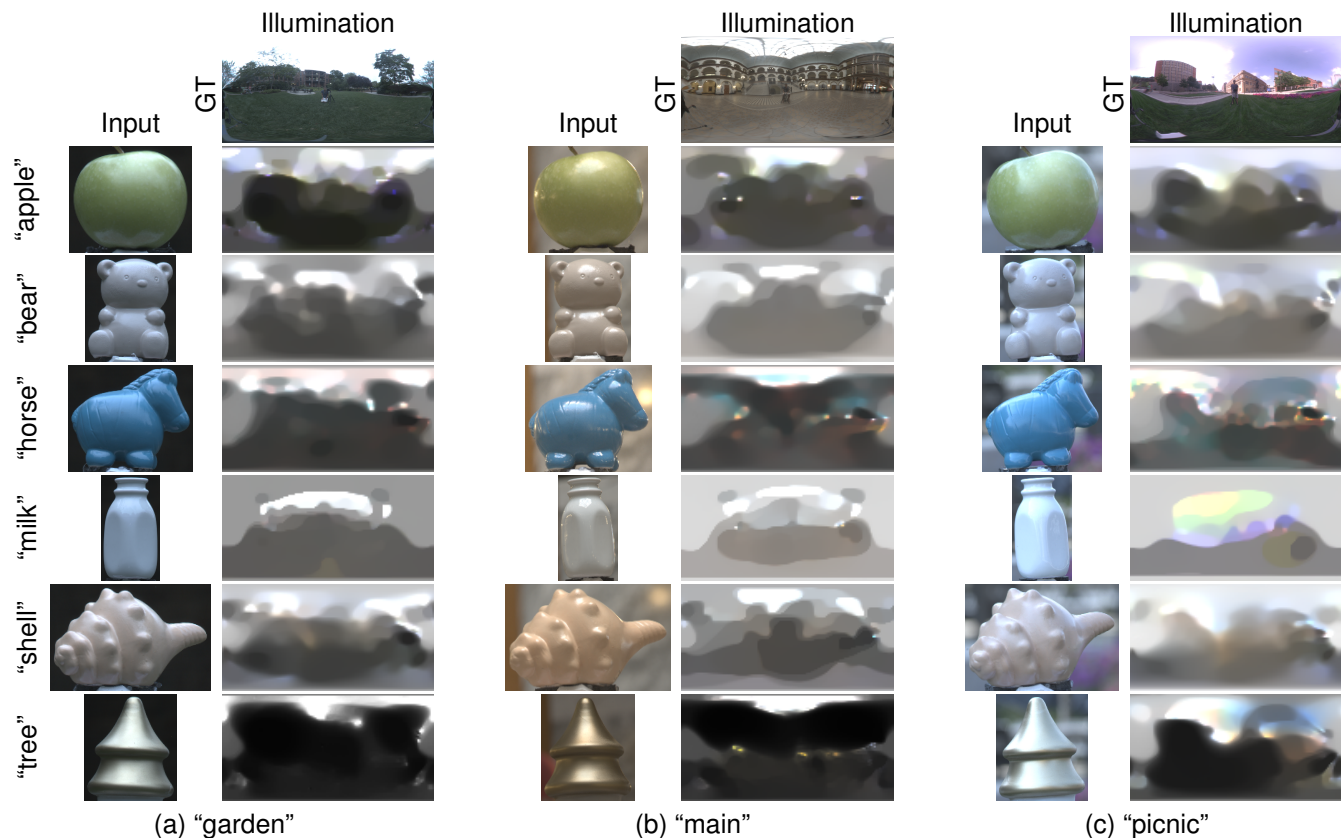


Fig. 11. Results on “garden”, “main”, and “picnic” illumination environments. Each row shows the recovered illumination of a different object with the ground-truth illumination for comparison. This illustrates that our algorithm is able to capture important features of the illumination environment. For example, in the “picnic” scene, the method captures the sun behind a building from the “apple” object.

accurately capture the frequency characteristics of the reflectance function. For instance, we correctly estimate the soft specular highlight of the gray plastic sphere in the fourth column whereas the method of [25] recovers a more specular BRDF. We believe this is because the joint inference of reflectance and illumination has a better opportunity to reconstruct the true reflectance than marginalization as the number of possible illumination environments is vast. Representing the distribution and marginalizing over all these possibilities accurately is very difficult—directly representing and solving for the illumination allows us to avoid this problem. Instead, our method exploits specific illumination characteristics that would lead to the most plausible illumination environment.

7 CONCLUSION

Reflectance and natural illumination recovery in the real world is an important yet challenging problem in computer vision. In this paper, we presented a novel method that uses flexible reflectance and illumination models in order to handle real-world scenes. We introduced strong priors that keep inference tractable and enable us to overcome the ambiguities due to the image formation

process. We used synthetic examples to further analyze the performance of our model under a variety of scenarios. Finally, we demonstrated the performance of our algorithm on real-world scenes. We believe this work is an important step towards making physics-based computer vision methods fully appreciate the world.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research grants N00014-11-1-0099 and N00014-14-1-0316, and the National Science Foundation awards IIS-0746717, IIS-0964420 and IIS-1353235.

REFERENCES

- [1] J. Gu and C. Liu, “Discriminative illumination: Per-pixel classification of raw materials based on optimal projections of spectral brdf,” in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, 2012, pp. 797–804.
- [2] J. Lambert, *Photometria*, 1760.
- [3] K. E. Torrance and E. M. Sparrow, “Theory for off-specular reflection from roughened surfaces,” *J. Opt. Soc. Am.*, vol. 57, no. 9, pp. 1105–1112, Sep 1967.
- [4] M. Chandraker and R. Ramamoorthi, “What an image reveals about material reflectance,” in *ICCV*, 2011, pp. 1–8.
- [5] R. Ramamoorthi and P. Hanrahan, “A Signal-Processing Framework for Reflection,” *ACM Trans. on Graphics*, vol. 23, no. 4, pp. 1004–1042, Oct. 2004.

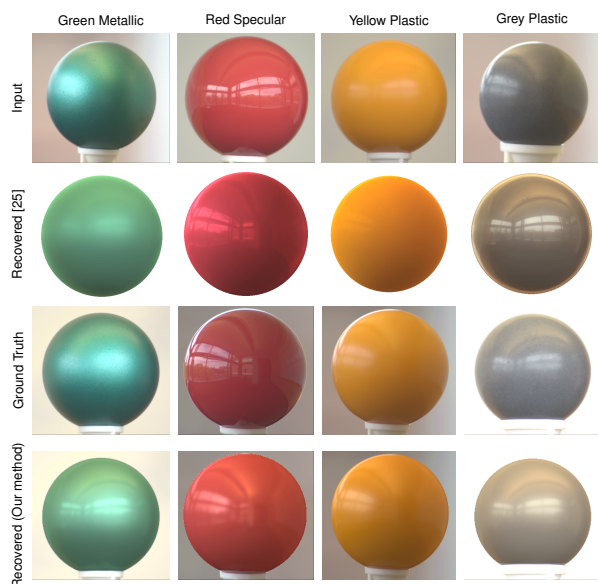


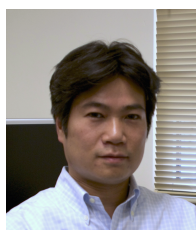
Fig. 12. Comparison to Romeiro and Zickler [25]. We compare our results qualitatively on a data set of four real spheres created by [25]. The top row shows the input image, the second row shows the predicted appearance from a novel viewpoint using the BRDF recovered using [25], the third row shows the ground truth appearance of the object from the novel viewpoint, the fourth row shows the predicted image from the novel viewpoint using the recovered reflectance function using our method and the ground truth illumination map. As shown, our method in general recovers a more accurate BRDF.

- [6] K. Nishino, "Directional Statistics BRDF Model," in *ICCV*, 2009, pp. 476–483.
- [7] K. Nishino and S. Lombardi, "Directional Statistics-based Reflectance Model for Isotropic Bidirectional Reflectance Distribution Functions," *Journal of Optical Society America, A*, vol. 28, no. 1, pp. 8–18, Jan. 2011.
- [8] S. Rusinkiewicz, "A New Change of Variables for Efficient BRDF Representation," in *Eurographics Workshop on Rendering*, 1998, pp. 11–22.
- [9] S. Lombardi and K. Nishino, "Single image multimaterial estimation," in *CVPR*, 2012, pp. 238–245.
- [10] —, "Reflectance and natural illumination from a single image," in *ECCV*, 2012.
- [11] W. Matusik, H. Pfister, M. Brand, and L. McMillan, "A Data-Driven Reflectance Model," *ACM Trans. on Graphics*, vol. 22, no. 3, pp. 759–769, Jul. 2003.
- [12] R. J. Woodham, "Photometric method for determining surface orientation from multiple images," *Optical Engineering*, vol. 19, no. 1, pp. 191–199, 1980.
- [13] B. K. Horn, "Shape from shading: A method for obtaining the shape of a smooth opaque object from one view," Cambridge, MA, USA, Tech. Rep., 1970.
- [14] K. Ikeuchi and K. Sato, "Determining reflectance properties of an object using range and brightness images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 13, no. 11, pp. 1139–1153, Nov 1991.
- [15] Y. Sato, M. D. Wheeler, and K. Ikeuchi, "Object Shape and Reflectance Modeling from Observation," in *Proc. of ACM SIGGRAPH*, 1997, pp. 379–387.
- [16] G. J. Ward, "Measuring and modeling anisotropic reflection," in *Proceedings of the 19th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '92. New York, NY, USA: ACM, 1992, pp. 265–272.
- [17] S. R. Marschner, S. H. Westin, E. P. F. Lafortune, K. E. Torrance, and D. P. Greenberg, "Image-based brdf measurement including

- human skin," in *Proceedings of the 10th Eurographics Conference on Rendering*, ser. EGWR'99. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 1999, pp. 131–144.
- [18] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Trans. Graph.*, vol. 18, no. 1, pp. 1–34, Jan. 1999.
- [19] T. Zickler, R. Ramamoorthi, S. Enrique, and P. N. Belhumeur, "Reflectance Sharing: Predicting Appearance from A Sparse Set of Images of a Known Shape," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1287–1302, Aug. 2006.
- [20] S. Marschner and D. Greenberg, "Inverse lighting for photography," in *IS&T/SID Fifth Color Imaging Conference*. The Society for Imaging Science and Technology, 1997, pp. 262–265.
- [21] K. Nishino, Z. Zhang, and K. Ikeuchi, "Determining reflectance parameters and illumination distribution from a sparse set of images for view-dependent image synthesis," in *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, vol. 1. IEEE, 2001, pp. 599–606.
- [22] K. Hara, K. Nishino, and K. Ikeuchi, "Mixture of Spherical Distributions for Single-View Relighting," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, no. 1, pp. 25–35, Jan. 2008.
- [23] J. T. Barron and J. Malik, "Intrinsic scene properties from a single rgb-d image," *CVPR*, 2013.
- [24] F. Romeiro, Y. Vasilyev, and T. Zickler, "Passive Reflectometry," in *ECCV*, 2008, pp. 859–872.
- [25] F. Romeiro and T. Zickler, "Blind Reflectometry," in *ECCV*, 2010, pp. 45–58.
- [26] R. L. Cook and K. E. Torrance, "A reflectance model for computer graphics," *ACM Trans. on Graphics*, vol. 1, no. 1, pp. 7–24, Jan. 1982. [Online]. Available: <http://doi.acm.org/10.1145/357290.357293>
- [27] R. Fisher, "Dispersion on a Sphere," *Royal Society of London Proceedings Series A*, vol. 217, pp. 295–305, May 1953.
- [28] R. O. Dror, A. S. Willsky, and E. H. Adelson, "Statistical Characterization of Real-World Illumination," *Journal of Vision*, vol. 4, no. 9, Sep. 2004.
- [29] R. R. Wilcoxon, *Introduction to Robust Estimation and Hypothesis Testing*, 2nd ed., ser. Statistical Modeling and Decision Science. Academic Press, Dec. 2004.
- [30] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *The Computer Journal*, vol. 7, no. 2, pp. 155–162, 1964.
- [31] P. Debevec, "Rendering synthetic objects into real scenes: bridging traditional and image-based graphics with global illumination and high dynamic range photography," in *Proc. of ACM SIGGRAPH*, 1998, pp. 189–198.



Stephen Lombardi is a PhD student at Drexel University in the Department of Computing. He received his B.S. in Computer Science in 2009 from The College of New Jersey. His main research interests lie in computer vision; specifically, he is interested in knowing how physical quantities can be inferred from image data in the real world. He is a student member of IEEE.



Ko Nishino is an associate professor in the Department of Computing in the College of Computing and Informatics at Drexel University. He is also an adjunct associate professor in the Computer and Information Science Department of the University of Pennsylvania and a visiting associate professor of Osaka University. He received a B.E. and an M.E. in Information and Communication Engineering in 1997 and 1999, respectively, and a PhD in Computer Science in 2002, all from The University of Tokyo. Before

joining Drexel University in 2005, he was a Postdoctoral Research Scientist in the Computer Science Department at Columbia University. His primary research interests lie in computer vision and include appearance modeling and synthesis, geometry processing, and video analysis. His work on modeling eye reflections received considerable media attention including articles in New York Times, Newsweek, and NewScientist. He received the NSF CAREER award in 2008.