Reflectance Capture using Univariate Sampling of BRDFs

Zhuo Hui,[†] Kalyan Sunkavalli,[‡] Joon-Young Lee,[‡] Sunil Hadap,[‡]
Jian Wang,[†] and Aswin C. Sankaranarayanan[†]
[†]Carnegie Mellon University [‡]Adobe Research

Abstract

We propose the use of a light-weight setup consisting of a collocated camera and light source — commonly found on mobile devices — to reconstruct surface normals and spatially-varying BRDFs of near-planar material samples. A collocated setup provides only a 1-D "univariate" sampling of a 3-D isotropic BRDF. We show that a univariate sampling is sufficient to estimate parameters of commonly used analytical BRDF models. Subsequently, we use a dictionary-based reflectance prior to derive a robust technique for per-pixel normal and BRDF estimation. We demonstrate real-world shape and capture, and its application to material editing and classification, using real data acquired using a mobile phone.

1. Introduction

Reflectance properties play an important role in the appearance of objects in a scene. For an opaque object, these properties are represented by the 4-D bidirectional reflectance distribution function (BRDF), which completely characterizes how a material interacts with incident light. Measuring the BRDF of a material often requires dense sampling of the 4-D space using precisely calibrated, and often prohibitively expensive, acquisition setups [9, 15–17].

More recently, researchers have looked at the problem of reflectance capture "in the wild", under relatively unconstrained conditions, and using commodity hardware. Because of the ill-posed nature of this problem, these methods rely on extra information like the presence of reference materials in the scene [23] or restrict themselves to BRDFs with stochastic, texture-like spatial variations [3].

The goal of our work is to enable the acquisition of the shape and spatially-varying BRDF (SV-BRDFs) of a wide range of real-world materials with using a practical, easy-to-deploy setup. To this end, we would like to use a mobile device — with a camera and a controllable flash — to take reflectance measurements. However, the position of the flash on these devices is fixed relative to the camera, and they are often nearly collocated. As a result, capturing images using this setup gives us only a sparse sampling of the BRDF. Even for the restricted set of isotropic materials (which are described by a 3-D BRDF), these measurements constitute only the 1-D slice of the 3-D BRDF that contains the specular lobe. We refer to this as a univariate sampling of the BRDF. The main contribution of our work is to show that such a univariate sampling of a material's appearance is, in fact, sufficient to recover per-pixel surface normals and BRDF estimates.

Real-world BRDFs can be well approximated as a linear combination of a small set of basis BRDFs [11, 17]. Based on this property, we show that while the sparse univariate samples are not sufficient by themselves, combining them with a dictionary-based prior [13] can lead to high-quality reflectance estimates. Further, we show that the parameters of many classical analytical BRDF models can be estimated purely from univariate sampling. This is because a collocated setup samples the specular lobe of the BRDF, which plays a major role in material appearance. Thus, when constrained to take a few sparse samples of the BRDF, instead of spreading these samples across the 4-D (or a 3-D isotropic or a 2-D bivariate) space, concentrating these samples in this 1-D slice is a better way to identify the BRDF.

We use the camera and flash unit on an iPhone 6S device to scan numerous near-planar (wrt depth) targets and subsequently estimate their shape and reflectance. For each target, we capture multiple images by moving the phone. For ease of calibration of the camera/light-source orientation, we place small checkerboard patterns on the near-vicinity of the target; the acquired images are aligned via a homography estimated using these checkerboard patterns. Using the aligned images, we estimate per-pixel surface normals and SV-BRDFs using a novel, robust method based on our univariate sampling strategy. We demonstrate this robustness on a wide range of scenes with complex SV-BRDFs and further, showcase the use of the proposed BRDF acquisition technique for reflectance rendering as well as material clustering. Figure 1 showcases our SV-BRDF and normal estimates for a real-world sample.

Contributions. Our specific contributions are as follows:

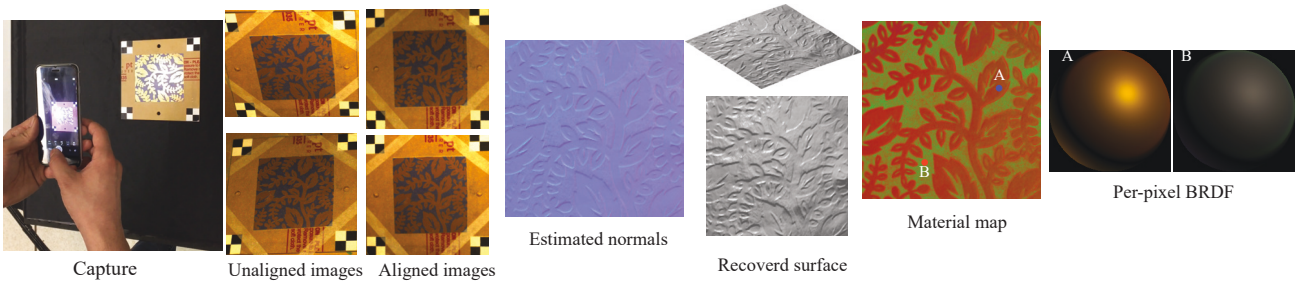


Figure 1. We acquire multiple images of a near-planar object using the camera and the flash unit on a mobile phone and subsequently, estimate the surface normals as well as the spatially-varying BRDF at each pixel.

BRDF identifiability analysis. We provide a comprehensive theoretical and empirical analysis of the identifiability of BRDFs given sparse samples from a collocated setup.

Practical shape and SV-BRDF estimation. We propose a robust optimization scheme to recover per-pixel normals and BRDFs of near-planar real-world materials from images captured with a collocated setup.

Limitations. Our method is limited to near-planar samples with little depth variation. This is because we rely on a planar geometric proxy to align the multiple captured images. We assume that the images are radiometrically calibrated. The light intensity across the material sample should be uniform and significantly greater than the ambient light levels. Our method requires alignment for the input sequence. Imprecise alignment may lead to the blurry of the results. Finally, our analysis will fail for complex BRDFs like anisotropic materials and in the modeling of the Fresnel effect at grazing incidence angles.

2. Prior work

Direct BRDF measurement. The BRDF is a function of four angles, two each for incident and outgoing directions, and classical BRDF measurement approaches [10, 16, 17] sample this 4D space by capturing images under varying lighting and viewing directions. Densely sampled measured BRDFs can provide faithful renditions of material appearance, but require specialized acquisition setups to capture large numbers of images.

Photometric stereo methods. Photometric stereo methods aim to recover shape from images captured with a fixed camera and varying lighting. While originally proposed for materials with known reflectance [30, 31], they have been extended to jointly infer shape and reflectance properties. This is done by using low-dimensional parametric reflectance models such as the isotropic Ward model [11], or directional statistics BRDF model [20–22]. Alternatively, the form of the reflectance function is constrained, typically by assuming that the BRDF is isotropic. Romeiro et al. [25] show that isotropic BRDFs are well-approximated by a 2-D bivariate representation and use this to recover BRDF from a single image of a known shape under known illumination. The bivariate representation has been used for shape and SVBRDF estimation from multiple images [4, 29], and blind reflectance recovery from a single image of a known shape [26]. Chandraker et al. [5–7] leverage motion cues to recover shape and reflectance for objects with dichromatic BRDFs. While more general than direct BRDF capture, these methods rely on restricted setups (calibrated, distant lighting and cameras) and/or extra information (known geometry, homogenous BRDFs). Our goal is to capture general SV-BRDFs using a light-weight capture setup.

Optimal BRDF sampling. Nielsen et al. [19] address the problem of identifying the optimal set of reflectance measurements required to recover a BRDF. This idea is further extended by Xu et al. [32] to consider near-field measurements. These papers show that a small set of images – in some cases, even two – are sufficient to estimate a BRDF. However, they are restricted to homogeneous materials and the nature of these measurements requires two pre-calibrated cameras and light sources. In contrast, we seek to recover SV-BRDFs using commodity hardware, and we demonstrate that this is possible using a collocated setup by enforcing a dictionary-based prior on the reconstruction.

BRDF acquisition using commodity devices. Higo et al. [12] capture images with a hand-held camera with an attached point light source and use a combination of near-light Photometric Stereo and multi-view stereo to reconstruct roughly Lambertian objects. Ren et al. [23] show that SV-BRDFs can be acquired using a fixed camera and a moving hand-held source by placing reference material tiles in the scene. While their results are impressive, the use of reference materials makes this setup less practical in real-world situations. Aittala et al. [3] propose to estimate SVBRDFs and normal maps from flash/no-flash image pairs captured using mobile devices. They extend this work to a single image using neural network-based texture

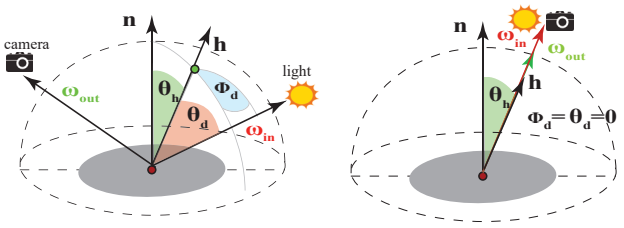
features [2]. However, these methods are restricted to stationary texture-like SVBRDFs and are aimed at reproducing plausible texture variations rather than accurate measured BRDF reconstruction. Riviere et al. [24] propose two prototypes using a mobile camera-flash or an LCD panel for reflectance capture. Their mobile camera solution can only handle rough specular surfaces and their shape and BRDF estimates are largely based on heuristics. In contrast, we can handle a wider range of materials because of our robust dictionary-based shape and reflectance estimation.

3. Univariate sampling of BRDFs

While arbitrary BRDFs are 4D functions of reflectance, many real-world materials are isotropic, in that, their BRDF is invariant to joint rotations of the incident and outgoing directions about the surface normal. The BRDF of such isotropic materials can be represented with a three-angle coordinates system, often using the half-angle parameterization [27] that is defined as follows. Given the surface normal $\mathbf{n}$, the incident direction $\boldsymbol{\omega}_i$ and the outgoing direction $\boldsymbol{\omega}_o$ — all unit-norm vectors — we first compute the half-angle $\mathbf{h} = (\boldsymbol{\omega}_i + \boldsymbol{\omega}_o)/2$. Next we define (θ_h, ϕ_h) as the elevation and azimuth, respectively, of the half-angle with respect to the surface normal, and (θ_d, ϕ_d) as the elevation and azimuth, respectively, of the outgoing direction with respect to the half-angle (see Figure 2). An isotropic BRDF, represented as $\rho(\theta_h, \theta_d, \phi_d)$, is represented as a function over θ_h, θ_d, and ϕ_d with $\theta_h, \theta_d \in [0, \pi/2)$ and $\phi_d \in [0, \pi)$. A subsequent reduction in dimensionality is provided by bivariate models [25] that further assume that the BRDF is invariant to changes in ϕ_d, and hence, the resulting reflectance is simply a function of θ_h and θ_d.

Collocated systems and univariate sampling. When the light source and the camera are collocated, then the incident and outgoing directions are the same, i.e., $\boldsymbol{\omega}_i = \boldsymbol{\omega}_o = \mathbf{h}$. Hence, $\theta_d = \phi_d = 0°$. Hence, any sampling of the BRDF is purely a function of θ_h. We refer to this paradigm as univariate sampling. Further, when there is a small, but fixed, offset between the light source and camera, then θ_d and ϕ_d are no longer zero, but are known constants independent of θ_h and ϕ_h, and hence can be pre-calibrated.

An important question to resolve upfront is whether univariate sampling can provide sufficiently rich measurements to be able to capture salient features of the measured BRDF, as well as enable stable reconstructions of the BRDF. We address this in two different ways. First, in Section 3.1, we show that the parameters of many analytical BRDF models are identifiable from noiseless univariate samples. Second, in Section 3.2, we provide empirical results characterizing accuracy of BRDFs, estimated from univariate samples.



Half-angle representation Collocated setup

Figure 2. Schematic of half-angle BRDF representation with respect to $(\theta_h, \theta_d, \phi_d)$ and univariate sampling on θ_h.

3.1. Identifiability under univariate sampling

We now address the question of identifiability of BRDFs from univariate samples, i.e., in the absence of noise, can there exist two distinct BRDFs that produce the same set of univariate samples? The answer is a resounding yes, if we do not further constrain the BRDF in some meaningful way. We do so by restricting ourselves to popular parametric BRDF models, and show that the parameters of the models are identifiable from noiseless univariate samples. Given the space constraints, we show this for the Cook-Torrance model [8] and provide justifications for other models including the Blinn-Phong, isotropic Ward and the Ashikhmin-Shirley model in the supplemental material.

Proposition. The parameters of the Cook-Torrance model are identifiable from noiseless univariate measurements.

Proof. BRDF measurements under the Cook-Torrance model are dependent on two parameters: m and F_0. Under univariate sampling, the BRDF can be written as:

$$\rho(\theta_h) = \rho_d + \frac{(1-\rho_d)DGF}{\pi\,(\mathbf{n}^\top \mathbf{l})\,(\mathbf{n}^\top \mathbf{v})} = \rho_d + \frac{(1-\rho_d)DGF_0}{\pi \cos^2\theta_h} \quad (1)$$

where

$$D = \frac{e^{-\tan^2\theta_h/m^2}}{m^2\cos^4\theta_h}, \quad G = \min(1, 2\cos^2\theta_h).$$

The term G is purely a function of θ_h and does not depend on any parameters, i.e. F_0 and m. Note that the Fresnel term, F, reduces to a constant F_0 for a collocated setup.

First, we observe that $\rho_d = \rho(\pi/2)$.[1] Second, we can now rearrange (1) to the following expression:

$$\log \frac{\pi(\rho(\theta_h) - \rho_d)\cos^6\theta_h}{(1-\rho_d)G} = \log\left(\frac{F_0}{m^2}\right) - \frac{\tan^2\theta_h}{m^2} \quad (2)$$

[1] In practice, due to fore-shortening, we cannot make an observation at $\theta_h = \pi/2$; however, this can easily be handled by sampling the BRDF at values close to $\pi/2$ and predicting the limiting value.

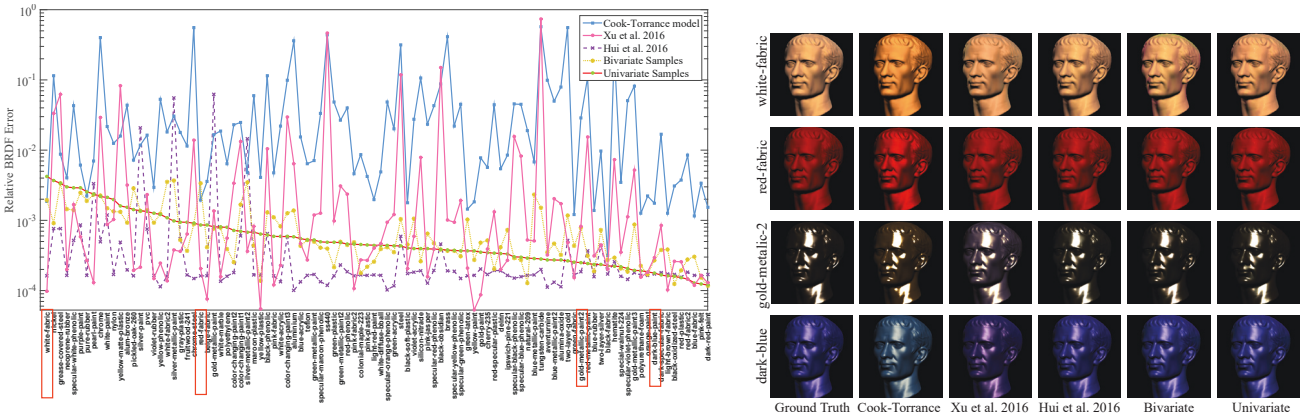


Figure 3. We demonstrate the performance of univariate sampling against a number of other sampling strategies. Shown are the relative BRDF errors of the reconstructed BRDF on MERL database for our technique, the data-driven method of Hui et al. [13], the bivariate model [25], the parametric model of Cook-Torrance, and optimal sampling model of Xu et al. [32]. We also compare against a 2D sampling strategy that we refer to as "bivariate sampling" that provides samples in a bivariate BRDF space (θ_h, θ_d). We observe that the method of Hui et al. [13] returns the best performance on an average. However, the univariate sampling with the proposed prior is able to compete against most of the state-of-the-art methods, both quantitatively as well as qualitatively.

Note that we have complete knowledge of the LHS term in (2). Further, if we plot the LHS as a function of $\tan^2\theta_h$, then the resulting plot is expected to be a straight line whose slope is $-1/m^2$ and whose intercept is $\log(F_0/m^2)$. Hence, we can identify all parameters of the model from the univariate measurements. □

3.2. Empirical validation

Next, we show that BRDFs can be estimated reliably from univariate measurements. Univariate samples provide a highly under-determined set of measurements and hence, recovering BRDFs from them requires the use of strong reflectance priors. We use a dictionary-based model for this purpose, borrowing an idea proposed recently in Hui et al. [13, 14].

Dictionary-based BRDF models. There have been many approaches [4, 11, 32] that model the BRDF at each pixel to lie in the non-negative span of a set of exemplar BRDFs, that we refer to as a dictionary. A dictionary D is simply a collection of exemplar BRDFs, often grouped together as a matrix $D = [\rho_1, \rho_2, \ldots, \rho_M]$, where each column is the BRDF of a measured material. Given D, we represent a BRDF ρ as:

$$\rho = D\mathbf{c}, \quad \mathbf{c} \geq 0.$$

Instead of estimating the high-dimensional vector ρ, we only need to estimate the abundances $\mathbf{c}$, whose dimension is proportional to the number of materials in the dictionary. Following Hui et al. [13], we further assume that $\mathbf{c}$ is sparse, suggesting that BRDF is the linear combination of a few dictionary atoms.

BRDF recovery. Univariate sampling measurements can be written as follows:

$$\begin{aligned} y(\theta_h) &= S(\theta_h)\rho + \eta \\ &= S(\theta_h)D\mathbf{c} + \eta, \end{aligned}$$

where $S(\theta_h)$ is the linear sampling operator that extracts the value at the input BRDF at $(\theta_h, 0, 0)$ and η is the measurement noise. Given M samples, corresponding to half-angle elevations in the set $\{\theta_h^1, \ldots, \theta_h^M\}$, we can compute the coefficients $\mathbf{c}$ by solving for the problem as

$$\widehat{\mathbf{c}} = \arg\min_{\mathbf{c}\geq 0} \sum_{i=1}^{M} \|y(\theta_h^i) - S(\theta_h^i)D\mathbf{c}\|_2^2 + \lambda\|\mathbf{c}\|_1. \quad (3)$$

We can now obtain the BRDF estimate $\widehat{\rho} = D\widehat{\mathbf{c}}$. The procedure illustrated above is from the dictionary-based modeling of BRDFs in Hui et al. [13], adapted to the univariate sampling scenario.

Evaluation. We evaluate the performance of the reconstruction technique with the state-of-the-art methods on the entire MERL database using a leave-one-out scheme. In particular, we compare against the parametric model of Cook-Torrance, the optimal sampling method in Xu et al. [32], and the isotropic sampling in Hui et al. [13]. For the Cook-Torrance model, we used the parameters reported in [18] — these parameters were optimized over the entire BRDF. For Hui et al. [13], we fix the surface normal at the north pole $[0,\ 0,\ 1]^\top$ and randomly sample the isotropic BRDF space for 20 combinations of lighting/view directions and reconstruct using a dictionary-based prior. For Xu et al. [32], we used the 20 optimal BRDF entries indicated

in their work. For the univariate sampling, we randomly sample the θ_h axis of the BRDFs and collect 20 samples with collocated lighting and view direction. Similarly, we also sample the bivariate BRDF space spanned by θ_h and θ_d with the 20 lighting/view combinations; we refer to this as bivariate sampling and use the same recovery algorithm as with univariate samples. For the results of Hui et al [13], univariate, and bivariate sampling, we perform 5 different random trials and report the average errors in Figure 3. The relative BRDF errors for these methods are shown in Figure 3, where we observe that univariate sampling is quite competitive to state-of-the-art models.

Remark. For many materials, the univariate sampling outperforms competing methods that sample in the bivariate space (θ_h and θ_d) or the isotropic space (θ_h, θ_d and ϕ_d). Given that we enforced a measurement budget for all methods, univariate sampling enjoys a denser sampling of the specular lobe. However, as we increase the number of measurements, univariate sampling has diminishing returns in reconstruction performance while competing methods that perform full sampling as well as bivariate sampling continue to observe significant gains. Our empirical evaluation also indicates that BRDFs of real-world materials are highly redundant and that the univariate sampling of an isotropic BRDF for $\theta_d = 0$ is often sufficient for high-quality reconstructions. This hypothesis is similar in spirit to bipolynomial BRDF model introduced by Shi et al. [28], providing the BRDF as the product of two univariate function over θ_d and θ_h, respectively.

4. Shape and reflectance estimation under univariate sampling

Acquisition setup and calibration. Our imaging setup consists of a nearly-collocated camera and light source, we assume the intrinsic matrix of the camera is known via a one-time pre-calibration. We acquire Q (typically, about 100) images at different viewpoints of a target. We assume the target is nearly planar, mainly for ease of registering the images across different viewpoints using homography-based methods. For each view, we use the four checker board patterns attached to the corners of the target to compute the homography. The checker board patterns also allow us to compensate the lighting variations within each captured image. Using the homography, we align pixels across different images and find world coordinates of all pixels. We now have a stack of intensity observations under known lighting and viewing directions for each pixel.

Problem statement. Given the aligned images, we can formulate the objective function that incorporates both surface normal and BRDF at pixel $\mathbf{p}$ as

$$\{\widehat{\mathbf{n}}_\mathbf{p}, \widehat{\mathbf{c}}_\mathbf{p}\} = \underset{\mathbf{c}\geq 0, \mathbf{n}}{\arg\min} \|\mathbf{I}_\mathbf{p} - B(\mathbf{n}, \mathbf{l}_\mathbf{p}, \mathbf{v}_\mathbf{p})\mathbf{c}\|_2^2 + \lambda\|\mathbf{c}\|_1 \quad (4)$$

where $\mathbf{I}_\mathbf{p} \in \mathbb{R}^Q$ denotes the image intensities observed at pixel $\mathbf{p}$ after alignment, $\mathbf{l}_\mathbf{p}$ and $\mathbf{v}_\mathbf{p}$ are the lighting and viewing directions for Q collected images, i.e. $\mathbf{l}_\mathbf{p} = [l_\mathbf{p}^1, l_\mathbf{p}^2, \dots l_\mathbf{p}^Q]$, $\mathbf{v}_\mathbf{p} = [v_\mathbf{p}^1, v_\mathbf{p}^2, \dots v_\mathbf{p}^Q]$. Note that $\mathbf{l}_\mathbf{p}$ and $\mathbf{v}_\mathbf{p}$ are known via the calibration. The term $B(\mathbf{n}, \mathbf{l}_\mathbf{p}, \mathbf{v}_\mathbf{p})$, an $Q \times M$ matrix, is given as

$$B(\mathbf{n}, \mathbf{l}_\mathbf{p}, \mathbf{v}_\mathbf{p}) = S(\mathbf{n}, \mathbf{l}_\mathbf{p}, \mathbf{v}_\mathbf{p})D,$$

where S has Q rows and a number of columns equal to the dimensionality of the BRDFs; here, S encodes the shading term as well as sampling of the BRDF. The estimates of the surface normal $\mathbf{n}_\mathbf{p}$ and the abundance $\mathbf{c}_\mathbf{p}$ amount to solving a quadratic cost function with ℓ_1-norm constraint.

Identifying BRDF exemplars. For computational efficiency, we enforce the sparsity prior on the abundances by first identifying a compact set of BRDF exemplars for a material sample. Specifically, we solve for the abundances at each pixel via (4) with initialized flat surface and sum the abundances across all pixels. Now, we obtain the summed result $\boldsymbol{\mathcal{C}} \in \mathbb{R}^M$, where M is the number of atoms in the dictionary. We empirically observe that only few atoms in $\boldsymbol{\mathcal{C}}$ have large values while the remaining entries are close to zero, which is consistent with the observation in [4, 13].

We retain only the K (in our case $K = 10$) BRDFs with the highest values of $\boldsymbol{\mathcal{C}}$ as our compact set of BRDF exemplars. This obviates the need for the sparsity constraint in subsequent iterations, thus speeding up computation. We denote $\widehat{B}$ as the dictionary with columns that corresponds to the exemplar set of atoms. We now solve for the normals and the coefficients:

$$\{\widehat{\mathbf{n}_\mathbf{p}}, \widehat{\mathbf{c}}_\mathbf{p}\} = \underset{\mathbf{c}\geq 0, \mathbf{n}}{\arg\min} \|\mathbf{I}_\mathbf{p} - \widehat{B}(\mathbf{n}, \mathbf{l}_\mathbf{p}, \mathbf{v}_\mathbf{p})\mathbf{c}\|_2^2. \quad (5)$$

Surface normal and SV-BRDF estimation. Given the initial estimate of $\mathbf{c}^{(0)}$ from flat surface and $\widehat{B}$, we use an iterative local search to solve for the surface normals. Specifically, we build a 2D grid with respect to the elevation and azimuth angles, and search in the grid for the normals which can best describe the intensity profile. In the first iteration, we initialize all the surface normals pointing toward the north pole, i.e., a flat surface, and solve for the abundances $\widehat{\mathbf{c}}^0$ via (4). Now, at the j-th iteration, we have normal estimate $\widehat{\mathbf{n}}_\mathbf{p}^{(j-1)}$ with elevation angle $\theta_\mathbf{p}^{(j-1)}$ and azimuth angle $\phi_\mathbf{p}^{(j-1)}$. The 2D grid for the (j)-th iteration is constructed as

$$\mathcal{N}^{(j)} = \{(\tilde{\theta}, \tilde{\phi}) || \tilde{\theta} - \theta_\mathbf{p}^{(j-1)}| \leq \mathcal{T}_\theta, |\tilde{\phi} - \phi_\mathbf{p}^{(j-1)}| \leq \mathcal{T}_\phi\},$$

where $\mathcal{T}_\theta$ and $\mathcal{T}_\phi$ are the thresholds to determine the cardinality of the candidate set. We can incorporate a coarse-to-fine search by specifying different values for $\mathcal{T}_\theta$ and $\mathcal{T}_\phi$, where $\mathcal{T}_\theta$ is varying from 5 to 0.1 degree while 50 to 1 degree for $\mathcal{T}_\phi$. For each element in $\mathcal{N}^{(j)}$, the candidate surface normal is computed as

$$\tilde{\mathbf{n}} = [\sin(\tilde{\theta})\cos(\tilde{\phi}),\ \sin(\tilde{\theta})\sin(\tilde{\phi}),\ \cos(\tilde{\theta})]$$

The estimate of the surface normal at a pixel $\mathbf{p}$ is given as

$$\widehat{\mathbf{n}}_{\mathbf{p}}^{(j)} = \arg\min_{\mathbf{n_p}\in\mathcal{N}^{(j)}} \|\mathbf{I_p} - \widehat{B}(\mathbf{n_p}, \mathbf{l_p}, \mathbf{v_p})\mathbf{c}_{\mathbf{p}}^{(j-1)}\|_2^2. \quad (6)$$

This is solved by scanning over all the elements in $\mathcal{N}^j$. Note that $\mathbf{c_p}$ has kept fixed with the values from the $(j-1)$-th iteration. Once we obtain $\widehat{\mathbf{n}}_{\mathbf{p}}^{(j)}$, we update the coefficients $\mathbf{c_p}$ by solving

$$\widehat{\mathbf{c}}_{\mathbf{p}}^{(j)} = \arg\min_{\mathbf{c_p}} \|\mathbf{I_p} - \widehat{B}(\widehat{\mathbf{n}}_{\mathbf{p}}^{(j)}, \mathbf{l_p}, \mathbf{v_p})\mathbf{c_p}\|_2^2 \quad s.t.\ \mathbf{c_p} \geq 0. \quad (7)$$

The algorithm typically converges within 10 iterations. The ultimate estimate of BRDF at each pixel is $\widehat{\rho}_{\mathbf{p}} = \widehat{D}\widehat{\mathbf{c}}_{\mathbf{p}}^{(J)}$, where $\widehat{D}$ corresponds to the selected columns for $\widehat{B}$ and J denotes for the number of iterations.

5. Results and Applications

In this section, we characterize the performance of our technique on a wide range of real-world scenes captured with iPhone 6s for a variety of tasks. We fix the target sample and move the phone while capturing images under the phone's flash illumination (see Figure 1). The images were captured with 2016×1512 pixels and we crop the regions with the target object for shape and BRDF estimation. We recover the per-pixel BRDFs with 1 degree for each angle in BRDF space, which leads to a $90 \times 90 \times 180 = 1,458,000$ dimensional vector. We direct the reader to the accompanying supplementary material for more results, comparisons, and analysis.

5.1. Shape and Reflectance Estimation

We process the captured images using the technique detailed in Section 4 to recover per-pixel surface normals and SV-BRDFs. We integrate the estimated normals using Poisson reconstruction [1] to obtain the 3D surface.

Shape estimation. To evaluate the performance of our shape estimation, we compare against the work of Riviere et al. [24], who use a similar mobile camera-based setup. While we model near-field camera and lighting, they assume that the camera and light are distant. In addition, their reflectance estimation is based on image heuristics, unlike our optimization-based framework with a BRDF prior. As demonstrated in Figure 4, our technique recovers more fine scale structures than [24]. In addition, our technique successfully separates reflectance effects from geometry, and as a result our reconstructions are largely planar. In contrast, their BRDF errors leak into the shape estimates leading to deviations from the planar structure of the samples. More comparisons with [24] on both real and synthetic scenes can be found in supplementary material.

Reflectance capture. Figures 1, 5 illustrate the performance of our method on datasets captured using an iPhone 6s. These four datasets — `leaf`, `leather`, `fur` and `characters` — have 123, 126, 70, and 138 input images, respectively. For each dataset, we show the estimated surface normals and recovered 3D shape under different viewpoints. The surface reconstructions show that we can recover fine-scale geometric details like yarn threads and leather patterns, even for samples with complex BRDFs. While we use a large number of input images to produce the results, our experience is that the performance degrades gracefully with a smaller number of images. We direct the reader to the supplementary materials where we include the BRDF/normal estimation error as a function of number of images on the synthetic dataset.

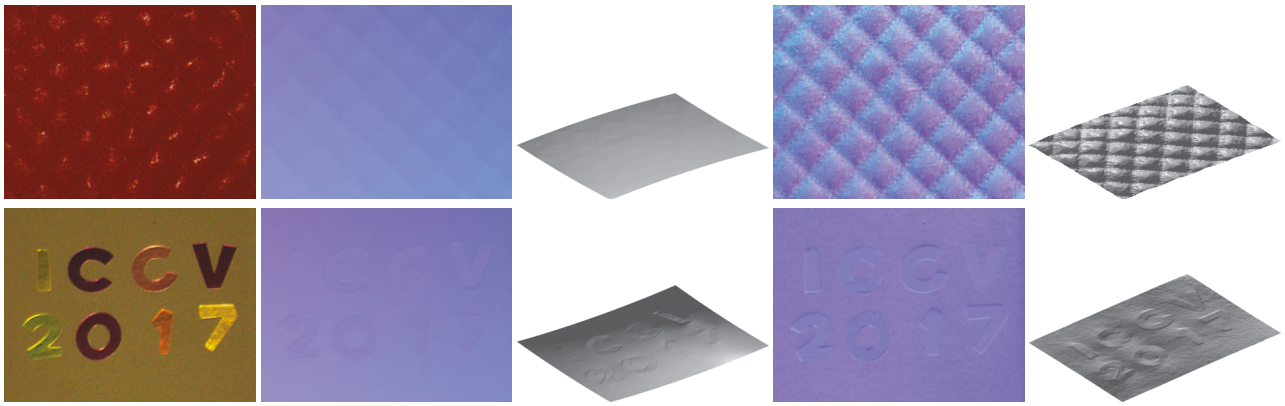
In addition to the images captured for shape and SVBRDF estimation, we capture additional images using a fixed camera and moving light source, i.e, a non-collocated setup. These "novel lighting" images are not part of the training dataset, and are used to visualize how accurately our shape and reflectance estimates generalize to directions that were not sampled. As shown in Figure 5, images rendered using our estimated normals and BRDFs under these novel lights closely resemble the actual captured photographs, indicating the robustness of our method.

5.2. Applications

Material editing. Once we reconstruct surface normals and SVBRDFs we can edit the material properties of the captured samples. This is demonstrated in Figure 6, where we a) swap specular and diffuse materials between two regions of the same sample, and b) transfer the specular material from one sample to a completely different sample. We re-render these edited BRDFs using the original estimated normals and view/lights. As can be seen here, our method is able to produce visually plausible results.

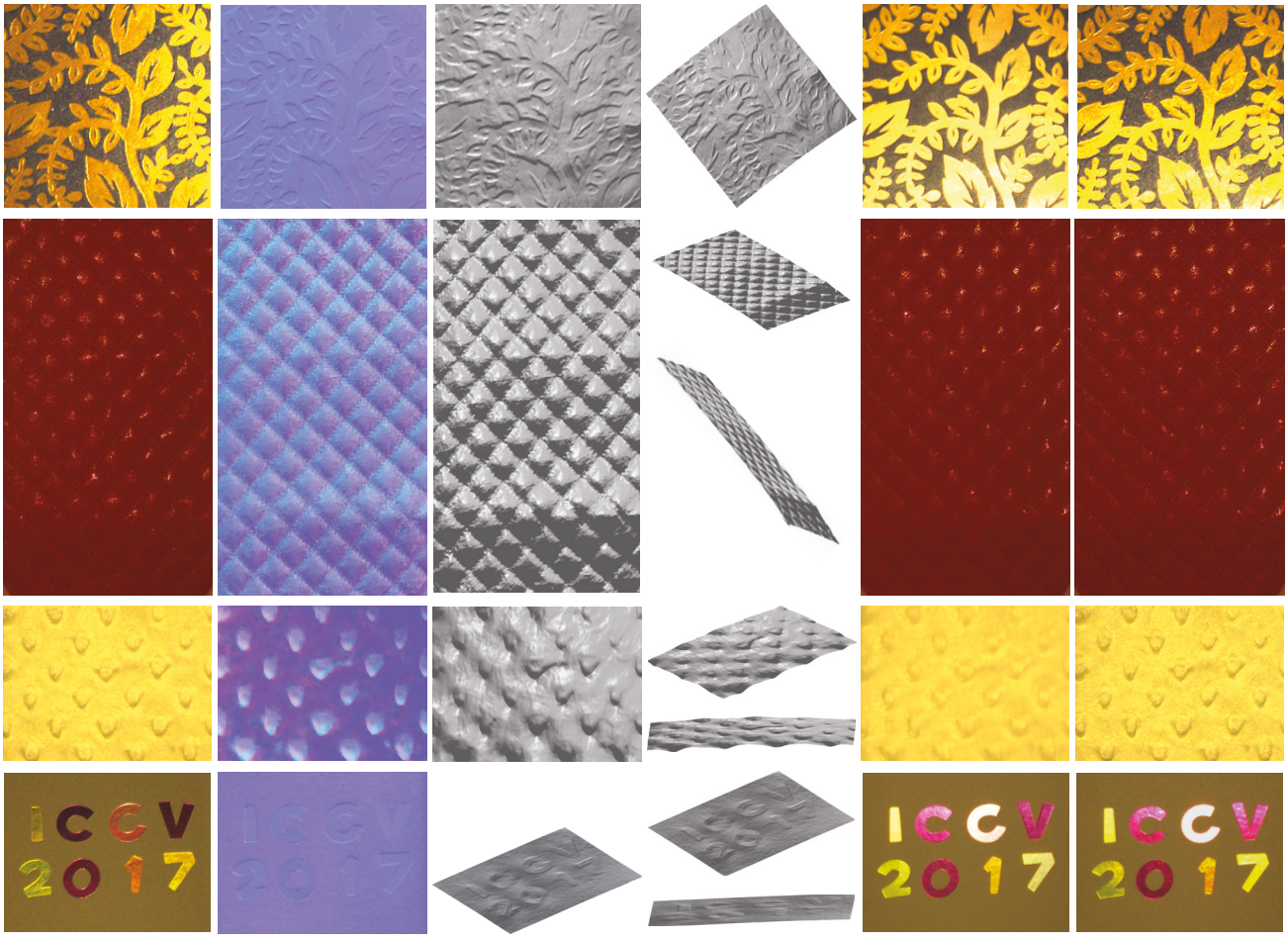
Material trait analysis. Previous work on recognizing material types uses specific optical setups [33] or projects raw BRDF measurements to a low-dimensional space [17]. However, these approaches are designed for objects with uniform reflectance or homogeneous BRDFs. In contrast, our technique estimates per-pixel BRDF abundances, and we can leverage this to estimate material traits at each pixel.

In order to do this, we first annotate all the materials in



Sample image Riviere et al. [24] Our results

Figure 4. We compare our performance on surface normal estimation with Riviere et al. [24] on two datasets. Shown are (left-right) one sample image, estimated normals and recovered 3D shape via Poisson reconstruction. Please note that our reconstructions, like the actual samples, are close to planar and contain more fine-scale detail.



(a) Input sample (b) Estimated normals (c) Recovered surface (d) Rendering (e) Photograph

Figure 5. We demonstrate shape and reflectance estimation on images captured using an iPhone 6S (a). We show the estimated normal map in false color (b) and recovered surface (c). We also compare rendered (d) results against actual captured photographs under novel lighting (e) that is not collocated with the camera.

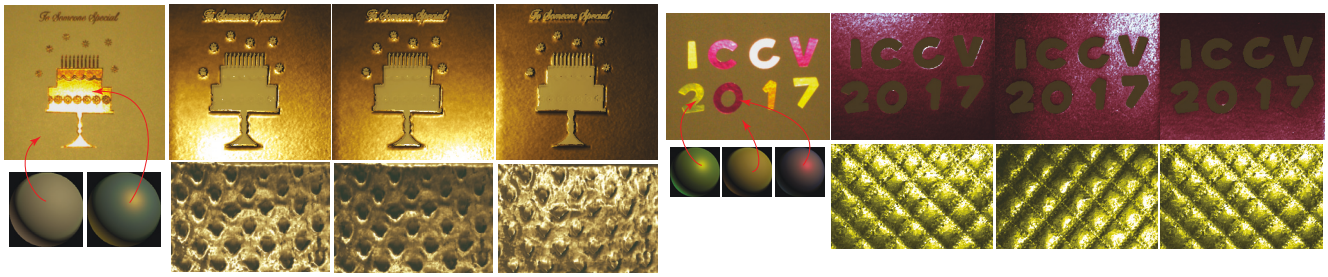


Figure 6. Material editing on two real samples. For the examples at the top, we compute the mean BRDF in the specular and diffuse regions of the samples (shown on the left), swap them and re-render them with the estimated normals, lights and cameras. For the examples at the bottom, we replace their SVBRDFs with the specular BRDFs from the top samples. These results are visually plausible, especially considering the fact that specular materials are likely to expose errors in geometry and material more clearly.

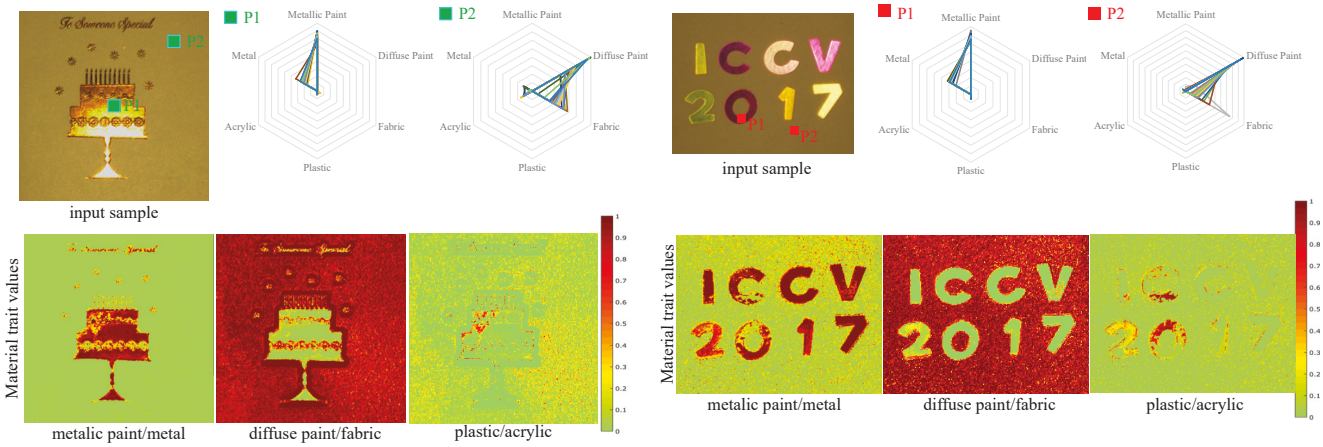


Figure 7. Material trait analysis on real captured data. (top) For two regions indicated by $\mathbf{p}_1$ and $\mathbf{p}_2$, we plot the associated material trait values (computed as described in Section 5.2). Pixels ($\mathbf{p}_1$) with metallic properties have large values in *metallic paint* and *metal* while pixels ($\mathbf{p}_2$) with diffuse Lambertian-like materials show large values in *diffuse paint* and *fabric*. (bottom) We visualize per-pixel material trait values for three material groups — *metallic paint+metal*, *diffuse paint+fabric*, and *plastic+acrylic*. This leads to clean, consistent material segmentations.

the MERL database with one of three unique material traits — *metal* + *metallic paint*, *fabric* + *diffuse paint* and *acrylic* + *plastic*. These three categories were chosen manually by visual inspection. We denote the i-th trait as $\mathcal{M}_i$. Given our abundance estimates $\widehat{\mathbf{c}}_{\mathbf{p}}$, we compute the per-pixel trait values by summing the abundances corresponding to materials with the same trait. Finally, we normalize these value so that they sum to 1:

$$\mathbf{m}_{\mathbf{p}}^{i} = \frac{\sum_{j \in \mathcal{M}_i} \widehat{\mathbf{c}}_{\mathbf{p}}(j)}{\sum_{k} \widehat{\mathbf{c}}_{\mathbf{p}}(k)}.$$

Figure 7 illustrates our proposed material trait analysis scheme for two datasets. Our predictions are consistent with the material properties of these samples – e.g., regions with metallic materials return high probabilities for the traits under *metal* + *metallic paint* – and accurately segment the samples into different materials.

6. Conclusion

This paper demonstrates the feasibility of reflectance capture using a collocated light source and camera, a hardware setup that is commonly found in mobile devices. We show that univariate sampling, commonly believed to be undesirable for reflectance estimation, can offer high-quality estimates of SV-BRDFs. While our results are applicable only to planar scenes, we believe that the ideas espoused in this paper are an important step towards reflectance capture in the wild.

Acknowledgments

Hui and Sankaranarayanan acknowledge support via the NSF CAREER grant CCF-1652569, the NGIA grant HM0476-17-1-2000, and a gift from Adobe Research.

References

[1] A. Agrawal, R. Chellappa, and R. Raskar. An algebraic approach to surface reconstruction from gradient fields. In *International Conference on Computer Vision (ICCV)*, 2005. 6

[2] M. Aittala, T. Aila, and J. Lehtinen. Reflectance modeling by neural texture synthesis. *ACM Transactions on Graphics (TOG)*, 35(4), 2016. 3

[3] M. Aittala, T. Weyrich, and J. Lehtinen. Two-shot svbrdf capture for stationary materials. *ACM Transactions on Graphics (TOG)*, 34(4):110, 2015. 1, 2

[4] N. Alldrin, T. Zickler, and D. Kriegman. Photometric stereo with non-parametric and spatially-varying reflectance. In *Computer Vision and Pattern Recognition (CVPR)*, 2008. 2, 4, 5

[5] M. Chandraker. On joint shape and material recovery from motion cues. In *European Conference on Computer Vision (ECCV)*, 2014. 2

[6] M. Chandraker. What camera motion reveals about shape with unknown brdf. In *Computer Vision and Pattern Recognition (CVPR)*, 2014. 2

[7] M. Chandraker. The information available to a moving observer on shape with unknown, isotropic brdfs. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 38(7):1283–1297, 2016. 2

[8] R. L. Cook and K. E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics (TOG)*, 1(1):7–24, 1982. 3

[9] K. J. Dana, B. van Ginneken, S. K. Nayar, and J. J. Koenderink. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics (TOG)*, 18(1):1–34, 1999. 1

[10] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *Computer Graphics and Interactive Techniques*, pages 145–156, 2000. 2

[11] D. B. Goldman, B. Curless, A. Hertzmann, and S. Seitz. Shape and spatially-varying brdfs from photometric stereo. In *International Conference on Computer Vision (ICCV)*, 2005. 1, 2, 4

[12] T. Higo, Y. Matsushita, N. Joshi, and K. Ikeuchi. A hand-held photometric stereo camera for 3-d modeling. In *International Conference on Computer Vision (ICCV)*, 2009. 2

[13] Z. Hui and A. Sankaranarayanan. Shape and spatially-varying reflectance estimation from virtual exemplars. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 1(99):1, 2016. 1, 4, 5

[14] Z. Hui and A. C. Sankaranarayanan. A dictionary-based approach for estimating shape and spatially-varying reflectance. In *International Conference on Computational Photography (ICCP)*, 2015. 4

[15] J. Lawrence, A. Ben-Artzi, C. DeCoro, W. Matusik, H. Pfister, R. Ramamoorthi, and S. Rusinkiewicz. Inverse shade trees for non-parametric material representation and editing. *ACM Transactions on Graphics (TOG)*, 25(3):735–745, 2006. 1

[16] S. R. Marschner, S. H. Westin, E. P. Lafortune, K. E. Torrance, and D. P. Greenberg. Image-based brdf measurement including human skin. In *Rendering Techniques*. 1999. 1, 2

[17] W. Matusik, H. Pfister, M. Brand, and L. McMillan. A data-driven reflectance model. *ACM Transactions on Graphics (TOG)*, 22(3):759–769, 2003. 1, 2, 6

[18] A. Ngan, F. Durand, and W. Matusik. Experimental analysis of brdf models. *Rendering Techniques*, 2005(16):2, 2005. 4

[19] J. B. Nielsen, H. W. Jensen, and R. Ramamoorthi. On optimal, minimal brdf sampling for reflectance acquisition. *ACM Transactions on Graphics (TOG)*, 34(6):186, 2015. 2

[20] G. Oxholm and K. Nishino. Shape and reflectance from natural illumination. In *European Conference on Computer Vision (ECCV)*, 2012. 2

[21] G. Oxholm and K. Nishino. Multiview shape and reflectance from natural illumination. In *Computer Vision and Pattern Recognition (CVPR)*, 2014. 2

[22] G. Oxholm and K. Nishino. Shape and reflectance estimation in the wild. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 38(2):376–389, 2016. 2

[23] P. Ren, J. Wang, J. Snyder, X. Tong, and B. Guo. Pocket reflectometry. 30(4):45, 2011. 1, 2

[24] J. Riviere, P. Peers, and A. Ghosh. Mobile surface reflectometry. In *Computer Graphics Forum*, 2016. 3, 6, 7

[25] F. Romeiro, Y. Vasilyev, and T. Zickler. Passive reflectometry. In *European Conference on Computer Vision (ECCV)*, 2008. 2, 3, 4

[26] F. Romeiro and T. Zickler. Blind reflectometry. In *European Conference on Computer Vision (ECCV)*, 2010. 2

[27] S. Rusinkiewicz. A new change of variables for efficient brdf representation. In *Eurographics Workshop on Rendering*, 1998. 3

[28] B. Shi, P. Tan, Y. Matsushita, and K. Ikeuchi. Bi-polynomial modeling of low-frequency reflectances. *IEEE transactions on pattern analysis and machine intelligence (TPAMI)*, 36(6):1078–1091, 2014. 5

[29] R. Shiradkar, L. Shen, G. Landon, S. Heng Ong, and P. Tan. A new perspective on material classification and ink identification. In *Computer Vision and Pattern Recognition (CVPR)*, 2014. 2

[30] W. M. Silver. *Determining shape and reflectance using multiple images*. PhD thesis, Massachusetts Institute of Technology, 1980. 2

[31] R. J. Woodham. Photometric method for determining surface orientation from multiple images. *Optical Engineering*, 19(1), 1980. 2

[32] Z. Xu, J. B. Nielsen, J. Yu, H. W. Jensen, and R. Ramamoorthi. Minimal brdf sampling for two-shot near-field reflectance acquisition. *ACM Transactions on Graphics (TOG)*, 35(6):188, 2016. 2, 4

[33] H. Zhang, K. Dana, and K. Nishino. Reflectance hashing for material recognition. In *Computer Vision and Pattern Recognition (CVPR)*, 2015. 6