OpenSVBRDF: A Database of Measured Spatially-Varying Reflectance

XIAOHE MA, State Key Lab of CAD&CG, Zhejiang University, China
XIANMIN XU, State Key Lab of CAD&CG, Zhejiang University, China
LEYAO ZHANG, State Key Lab of CAD&CG, Zhejiang University, China
KUN ZHOU, State Key Lab of CAD&CG, Zhejiang University and ZJU-FaceUnity Joint Lab of Intelligent Graphics, China
HONGZHI WU, State Key Lab of CAD&CG, Zhejiang University, China

Fig. 1. Rendering results of 108 (out of 1,000) high-quality SVBRDFs captured and reconstructed with our novel integrated system. Each row represents one category in our database. Please refer to the accompanying video for animated sequences.

We present the first large-scale database of measured spatially-varying anisotropic reflectance, consisting of 1,000 high-quality near-planar SVBRDFs,

*Corresponding authors: Kun Zhou & Hongzhi Wu ({kunzhou,hwu}@acm.org).
Authors' addresses: Xiaohe Ma, State Key Lab of CAD&CG, Zhejiang University, Hangzhou, 310058, China, xiaohema1998@gmail.com; Xianmin Xu, State Key Lab of CAD&CG, Zhejiang University, Hangzhou, 310058, China, xuxianmincs@gmail.com; Leyao Zhang, State Key Lab of CAD&CG, Zhejiang University, Hangzhou, 310058, China, herrmerid@gmail.com; Kun Zhou, State Key Lab of CAD&CG, Zhejiang University and ZJU-FaceUnity Joint Lab of Intelligent Graphics, Hangzhou, 310058, China, kunzhou@acm.org; Hongzhi Wu, State Key Lab of CAD&CG, Zhejiang University, Hangzhou, 310058, China, hwu@acm.org.

0730-0301/2023/12-ART254
https://doi.org/10.1145/3618358

spanning 9 material categories such as wood, fabric and metal. Each sample is captured in 15 minutes, and represented as a set of high-resolution texture maps that correspond to spatially-varying BRDF parameters and local frames. To build this database, we develop a novel integrated system for robust, high-quality and -efficiency reflectance acquisition and reconstruction. Our setup consists of 2 cameras and 16,384 LEDs. We train 64 lighting patterns for efficient acquisition, in conjunction with a network that predicts per-point reflectance in a neural representation from carefully aligned two-view measurements captured under the patterns. The intermediate results are further fine-tuned with respect to the photographs acquired under 63 effective linear lights, and finally fitted to a BRDF model. We report various statistics of the database, and demonstrate its value in the applications of material generation, classification as well as sampling. All related data, including future additions to the database, can be downloaded from https://opensvbrdf.github.io/.

CCS Concepts: • **Computing methodologies → Reflectance modeling**.

Additional Key Words and Phrases: SVBRDF, dataset, illumination multiplexing, appearance acquisition

ACM Reference Format:
Xiaohe Ma, Xianmin Xu, Leyao Zhang, Kun Zhou, and Hongzhi Wu. 2023. OpenSVBRDF: A Database of Measured Spatially-Varying Reflectance. *ACM Trans. Graph.* 42, 6, Article 254 (December 2023), 14 pages. https://doi.org/10.1145/3618358

1 INTRODUCTION

High-quality material appearance models the intricate physical interactions with light [Dorsey et al. 2010]. Often represented as a 6D Spatially-Varying Bidirectional Reflectance Distribution Function (SVBRDF), it is an indispensable component in visual computing, with a wide range of applications including cultural heritage, e-commerce, video games and visual effects. In computer graphics, high-quality digitized appearance can be rendered to faithfully reproduce the complex physical look that varies with location, lighting and view [Jakob et al. 2022]. On the other hand, material appearance helps machines better understand the real world from images [Li et al. 2020].

Over the past decades, considerable efforts have been made in building highly valuable reflectance databases, driven by the growing demand for accurate, diverse digital appearance both in academia and industry. This results in a few pieces of seminal work, including [Lawrence et al. 2006; Matusik et al. 2003] for scanned BRDFs/SVBRDFs and [Deschaintre et al. 2018] for synthetic SVBRDFs. However, the number of publicly available captured reflectance datasets remains limited nowadays, which hinders the development of related research in this data-hungry era. For example, the MERL database of 100 measured isotropic BRDFs [Matusik et al. 2003] remains in active use in research, 20 years after its publication.

Here the main reason is the technical difficulty in acquiring large-scale datasets with existing techniques. Despite its quality, exhaustively sampling the 6D physical domain of a single SVBRDF is prohibitively time-consuming [Dana et al. 1999; Lawrence et al. 2006], and thus cannot scale to building a large database. Strong-prior-based approaches trade reconstruction quality for acquisition efficiency [Lensch et al. 2003; Wang et al. 2008]; the quality of results cannot be guaranteed, whenever the assumptions are not satisfied. For illumination-multiplexing techniques with high quality and efficiency in general, even state-of-the-art work [Kang et al. 2018] is not sufficiently robust to handle challenging cases such as brushed metal and polished wood.

To address the above issues, we present a novel integrated system for *robust, high-quality and -efficiency capture* of near-planar anisotropic SVBRDFs. Our setup consists of 2 cameras and 16,384 independently controlled, high-brightness LEDs. We train a small set of lighting patterns for efficient acquisition, in conjunction with a network that predicts per-point reflectance in a neural representation from carefully aligned two-view measurements captured under the patterns. The intermediate results are further fine-tuned with respect to the photographs acquired under various effective linear light sources, and finally fitted to a BRDF model whose parameters are stored as texture maps. We validate the results against photographs and evaluate the design decisions.

Using our system, we build OpenSVBRDF, *the first large-scale database of measured spatially-varying anisotropic reflectance*. It consists of 1,000 high-quality near-planar SVBRDFs, spanning 9 categories of materials such as wood, fabric and metal. Each SVBRDF is captured in 15 minutes, and represented as a set of high-resolution texture maps that correspond to spatially-varying GGX BRDF parameters [Walter et al. 2007] and local frames. All related data (texture maps and photographs under learned patterns and linear lights) as well as future additions to the database can be freely downloaded from https://opensvbrdf.github.io/. Finally, we report various statistics of the database and explore its applications in material generation, classification as well as sampling. Superior results are demonstrated with the help of our database.

2 RELATED WORKS

2.1 Reflectance Acquisition

Here we mainly review acquisition techniques with *active illumination*, which are most related to this paper. Interested readers are directed to excellent surveys for a broader view of the topic [Dong 2019; Guarnera et al. 2016; Weinmann and Klein 2015; Weyrich et al. 2009].

A straightforward way to capture a general SVBRDF is to densely sample its 6D domain [Dana et al. 1999; Lawrence et al. 2006]. A spherical gantry moves a pair of a camera and a point light to sample at different combinations of the view and lighting directions. While the result quality is excellent, the acquisition time is extremely long. So other work aims to make acquisition more practical, which can be divided into the following two classes.

2.1.1 Prior-Based. This class of methods introduce various priors to regularize the reconstruction problem, therefore considerably reducing the number of measurements. Marschner et al. [2000] recover isotropic reflectance of a homogeneous convex object from a single view direction. Lensch et al. [2003] model the appearance as a linear combination of basis materials to constrain the reconstruction from a sparse number of flash-lit images. The idea is extended from a linear space to a low-dimensional manifold in [Dong et al. 2010]. Wang et al. [2008] exploit the spatial similarity of reflectance and spatial variation of local frames for single-view reconstruction of

microfacet BRDFs. Sparse lighting and view sampling directions are optimized by leveraging the MERL BRDF database and the spatial homogeneity for BRDF acquisition from a small number of images [Nielsen et al. 2015; Xu et al. 2016]. A dictionary-based reflectance prior is proposed in [Hui et al. 2017]. Nam et al. [2018] use hundreds of multi-view flash photographs to recover both a 3D shape and isotropic reflectance expressed as a linear combination of basis materials.

With the success of deep learning, priors gradually shift from heuristic ones to data-driven ones. Aittala et al. [2016] exploit structural similarity to estimate a stationary SVBRDF from a single flash image. A self-augmentation training process is employed to model SVBRDF from a single photograph [Li et al. 2017]. Deschaintre et al. [2018] propose a method that takes a single input photograph lit by a flash, and outputs an SVBRDF with a network trained over a dataset of procedural materials. Gao et al. [2019] learn a latent embedding to regularize the optimization for SVBRDF reconstruction with respect to an arbitrary number of input images. Guo et al. [2021] introduce highlight-aware convolution to estimate the saturated highlights from the adjacent unsaturated area in a single image.

In comparison, we emphasize result quality and choose to rely on as few assumptions as possible. For example, while the majority of existing techniques exploit spatial coherence, our approach works in a pixel-independent fashion to preserve high-frequency spatial details. On the other hand, our measured database might serve as a more powerful prior for existing learning-based work, which is often limited by isotropic training datasets: with our database, it might be possible to learn novel *anisotropic* reflectance estimation networks from a highly sparse number of input images.

2.1.2 Illumination Multiplexing. This highly successful class of approaches pack more information in a unit measurement by programming the intensities of multiple lights simultaneously over time, which improves the acquisition efficiency and signal-to-noise ratio. Lightstages take photographs of a sample under gradient illumination [Ghosh et al. 2009] or spherical harmonics [Tunwattanapong et al. 2013], and compute the reflectance from a manually derived inverse lookup table, which maps the observed radiance to anisotropic BRDF parameters. In Gardner et al. [2003] and Chen et al. [2014], a linear light source is regularly moved over a planar sample, and the reflectance is reconstructed from the corresponding appearance variations. The irregular motion of the linear light is supported in [Ren et al. 2011] with the help of pre-calibrated physical BRDF patches. Aittala et al. [2013] employ a camera and a near-field LCD panel as a programmable light source, to capture an isotropic reflectance using a frequency domain analysis. A sophisticated system is proposed in [Nam et al. 2016] to reconstruct micro-scale reflectance via an alternating optimization.

Recently, neural reflectance acquisition techniques map both the physical acquisition and computational processing to a single network, enabling the joint and automatic optimization of both the hardware and software. High acquisition efficiency and result quality are demonstrated on planar [Kang et al. 2018] and non-planar reflectance [Kang et al. 2019; Ma et al. 2021]. Our method is most similar to this line of work. In comparison, we further improve the reconstruction quality to robustly scale to a large number of physical samples, via an intermediate neural representation (Sec. 6) and additional fine-tuning (Sec. 7.1).

2.2 Measured Reflectance Datasets

As aforementioned, due to the difficulty of supporting robust, high -quality and -efficiency capture at once, public databases of accurately measured reflectance have always been a scarce and valuable resource over the past decades.

BRDF. Matusik et al. [2003] build a database of 100 measured isotropic BRDFs, expressed in the half-angle parameterization and discretized into 90×90×180 angular bins. The BRDFs of 8 automotive paints are acquired with an image-based setup in [Günther et al. 2005]. The UTIA dataset focuses on dense sampling of anisotropic reflectance and contains 150 measured BRDFs [Filip and Vávra 2014]. Recently, Dupuy and Jakob [2018] propose an adaptive BRDF parameterization for efficient sampling, and capture a dataset with 51 isotropic and 11 anisotropic BRDFs.

SVBRDF. Lawrence et al. [2006] acquire 5 anisotropic SVBRDFs and develop a decomposition algorithm to efficiently represent the data. Deschaintre et al. [2018] provide a dataset of 1,850 artist-created, parametric isotropic SVBRDFs, which are widely used in subsequent research on learning-based material estimation.

BTF. Despite its generality, Bidirectional Texture Function is an image-based representation, which does not record reflectance. Therefore, no physical knowledge in the angular domain can be exploited; exhaustive sampling of the view/lighting directions has to be conducted. Dana et al. [1999] capture 61 samples as the CUReT database. The UBO2014 dataset [Weinmann et al. 2014] consists of 7 material categories, each of which includes 12 samples, with an angular sampling rate of 151×151.

In comparison, OpenSVBRDF is the first large-scale database of measured spatially-varying anisotropic reflectance. We will make public all data related to this growing dataset, which hopefully will benefit both academia and industry.

3 PRELIMINARIES

We first list the single-channel relationship among the image measurement B from a surface point $\mathbf{p}$, the reflectance f and the intensity I of each LED on our device as follows:

$$B(I, \mathbf{x_p}, \mathbf{n_p}, \mathbf{t_p}) = \sum_l I(l) \int \frac{1}{||\mathbf{x_l} - \mathbf{x_p}||^2} \Psi(\mathbf{x_l}, -\omega_\mathbf{i}) V(\mathbf{x_l}, \mathbf{x_p}) \\ f(\omega_\mathbf{i}'; \omega_\mathbf{o}', \mathbf{p})(\omega_\mathbf{i} \cdot \mathbf{n_p})^+ (-\omega_\mathbf{i} \cdot \mathbf{n_l})^+ d\mathbf{x_l}. \quad (1)$$

Here l is the index of an LED, and $I(l)$ is its intensity in the range of $[0, 1]$, the collection of which will be referred to as a *lighting pattern*. Moreover, $\mathbf{x_p}/\mathbf{n_p}/\mathbf{t_p}$ is the position/normal/tangent of $\mathbf{p}$, while $\mathbf{x_l}/\mathbf{n_l}$ is the position/normal of a point on the light. We denote $\omega_\mathbf{i}/\omega_\mathbf{o}$ as the lighting/view direction, with $\omega_\mathbf{i} = \frac{\mathbf{x_l} - \mathbf{x_p}}{||\mathbf{x_l} - \mathbf{x_p}||}$. $\Psi(\mathbf{x_l}, \cdot)$ represents the angular distribution of the light intensity. V is a visibility function between $\mathbf{x_l}$ and $\mathbf{x_p}$. The operator $(\cdot)^+$ computes a dot product and clamps a negative result to zero. We employ the anisotropic GGX

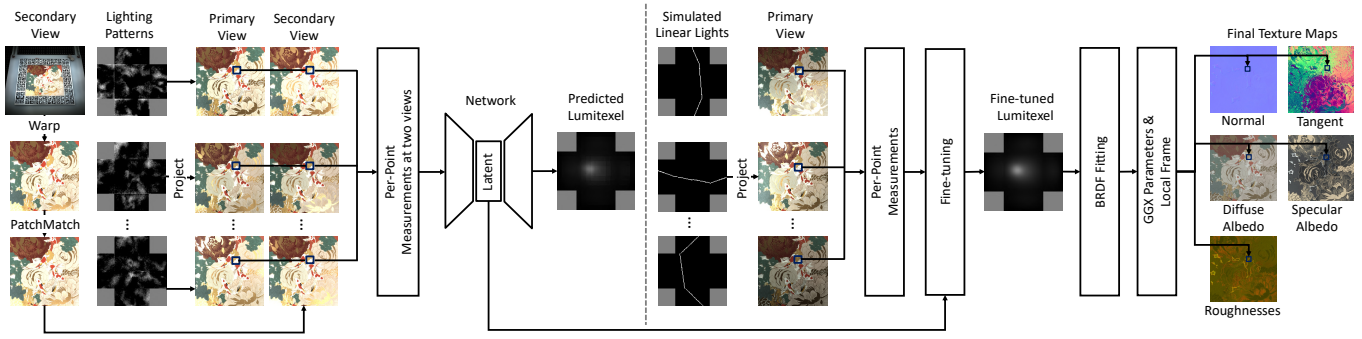
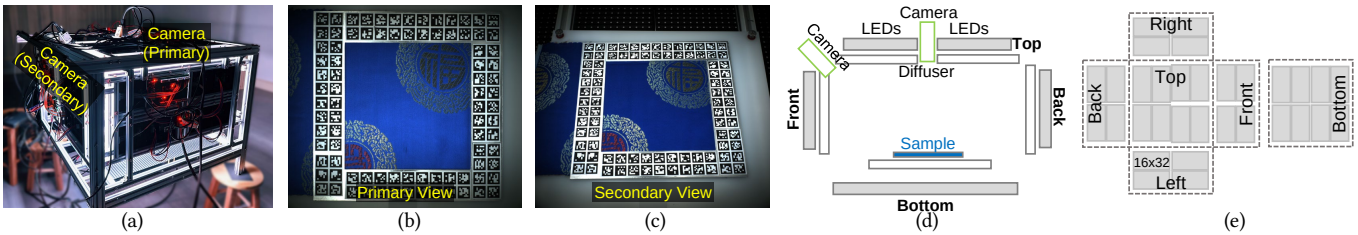


Fig. 2. Our pipeline. For each surface point on the sample, we physically encode its corresponding lumitexels at 2 views with 64 learned lighting patterns as photometric measurements; accurate correspondences between 2 views are computed by warping followed by dense SIFT features matching. A network then transforms these measurements to recover the lumitexel at the primary view, expressed in an intermediate neural representation. The neural representation is further optimized with respect to the image measurements under 63 effective linear light sources, and fitted to a GGX BRDF along with a local frame. The final results are stored as texture maps.



(a) (b) (c) (d) (e)

Fig. 3. Our acquisition setup. From left to right: (a) a photograph of our device, (b) the primary view, (c) the secondary view, (d) a side view, and (e) the light layout by unfolding all faces to the top plane. A total of 16,384 LEDs are installed.

model [Walter et al. 2007] to efficiently represent the BRDF f:

$$f(\omega_\mathbf{i};\omega_\mathbf{o},\mathbf{p}) = \frac{\rho_d}{\pi} + \rho_s \frac{D(\omega_\mathbf{h};\alpha_x,\alpha_y)F(\omega_\mathbf{i},\omega_\mathbf{h})G(\omega_\mathbf{i},\omega_\mathbf{o};\alpha_x,\alpha_y)}{4(\omega_\mathbf{i}\cdot\mathbf{n})(\omega_\mathbf{o}\cdot\mathbf{n})},$$

where ρ_d/ρ_s is the diffuse/specular albedo, α_x/α_y is the roughness, and ω_h is the half vector. In addition, D is the microfacet distribution function, F is the Fresnel term, and G is the geometry term for shadowing/masking effects. Also an index of refraction of 1.5 is used in all experiments. Next, a *lumitexel* m is defined as the collection of virtual measurements of BRDF f at $\mathbf{p}$, with one LED on at a time [Lensch et al. 2003]. It is a function of the light index l:

$$m(l;\mathbf{p}) = B(\{I(l)=1, \forall_{k\neq l} I(k)=0\},\mathbf{p}). \tag{2}$$

4 ACQUISITION OVERVIEW

We propose a novel integrated system for robust, high-quality and -efficiency capture of near-planar reflectance. First, we build an illumination multiplexing setup with 16,384 LEDs and 2 cameras. To scan a physical sample, we establish accurate correspondences between 2 views by matching dense SIFT features under uniform lighting. Next, for each surface point on the sample, we physically encode its corresponding lumitexels at 2 views with 64 learned lighting patterns as photometric measurements. A network then transforms these measurements to recover the lumitexel at the primary view, expressed in an intermediate neural representation. The neural representation is further fine-tuned with respect to the image measurements under 63 effective linear light sources, and fitted to a GGX BRDF along with a local frame. The final results are stored as texture maps. Please refer to Fig. 2 for a graphical illustration.

5 ACQUISITION SETUP

We build a near-field light-hemicube to scan physical appearance (Fig. 3). Its size is about 70cm×70cm×40cm. The material sample is placed on a translucent acrylic board, which can be slid in/out via attached drawer runners for rapid sample replacement in database construction. The maximum allowed sample size is 15cm×15cm. A hollow aluminum square covered with ARTags[Fiala 2005] is placed on top of the sample for fixation during acquisition.

We install two 24MP Basler a2A5328-15ucPRO vision cameras, pointing towards the sample from approximately 90° (primary) and 45° (secondary). The primary view is designed for direct imaging the sample at a standard view; otherwise, images warped from other views would have uneven sampling rates at different locations. The secondary view is for capturing interesting reflectance features in the angular domain, as is common in previous work [Kang et al. 2018]. There are 16,384 high-brightness LEDs around the sample, attached with diffusers and mounted to all six sides of our hemicube. The LED pitch is 1cm, and the intensity is quantized with 8 bits

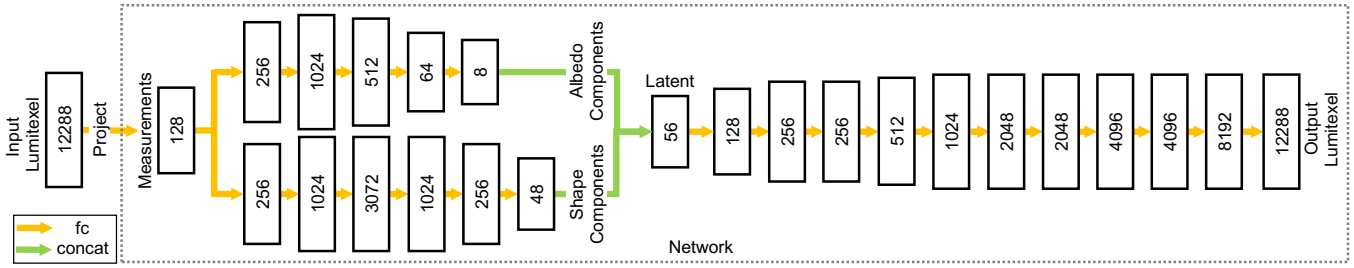


Fig. 4. Network architecture. The physical input lumitexel is transformed into a set of measurements by a single fc layer that represents all lighting patterns. Two fc networks convert the measurements to 8D albedo and 48D shape components of a latent vector. Next, an 11-layer fc network recovers the final lumitexel from the latent vector. Each fc layer before the latent is followed by a batch normalization layer and then a leaky ReLU activation layer. Each fc layer after the latent and before the last one is followed by a leaky ReLU activation layer.

using house-made circuits. Our device is similar to the one built in [Ma et al. 2023], with one extra camera and bottom-side lighting.

We calibrate the intrinsic and extrinsic parameters of cameras, as well as the positions, orientations and angular intensity distribution of each LED. In addition, vignetting is corrected with a flat field source, and color calibration is performed with an X-Rite ColorChecker Passport.

6 ACQUISITION NETWORK

Input/Output. The input to our network is a set of 128 single-channel physical measurements of a point on the material sample at 2 views, captured under 64 pre-optimized lighting patterns. Note that the original RGB channels of all measurements are averaged to a single gray-scale channel. We discuss the handling of RGB information in Sec. 7.1. The output is a gray-scale lumitexel at the primary view with a dimension of 12,288, the number of LEDs in the upper hemisphere of the sample plane. The lighting pattern number is set to 64, instead of 32 as in existing work [Kang et al. 2018, 2019], since we aim for improved quality at the cost of slightly increased acquisition time in database construction. The output of our network is the predicted lumitexel.

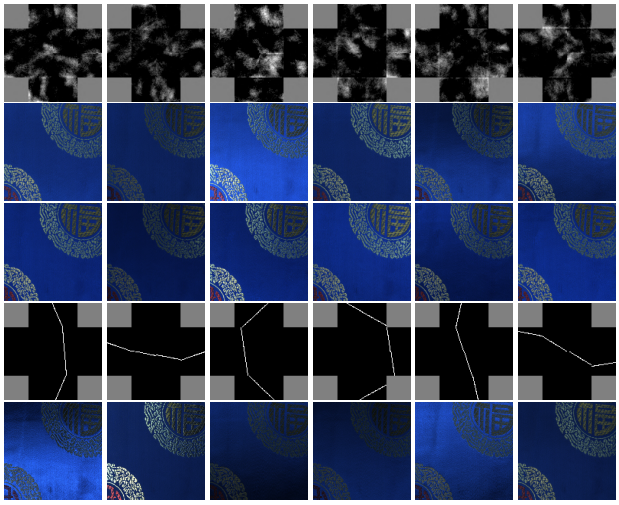


Fig. 5. Visualization of lighting patterns and corresponding photographs. From the top row to bottom: learned patterns, corresponding photographs at the primary and secondary view; linear lighting patterns, and corresponding photograph at the primary view. Each lighting pattern is parameterized on a cross, by unfolding all hemicube faces except the bottom to the top plane. Note that only a subset of all patterns are shown due to limited space.

Architecture. The network consists of 17 fc layers (Fig. 4). Its bottleneck is a 56D latent vector, which can be viewed as a low-dimensional neural representation of the lumitexel. To support subsequent fine-tuning step, this vector contains an 8D albedo component which relates to diffuse and specular albedo, and a 48D shape component that controls the albedo-independent part of a lumitexel. Alternatively, one can choose not to disentangle albedo and the rest in the latent vector. However, this brings in undesired extra degrees of freedom in the RGB case, since now the latent vector for each channel can correspond to e.g., different roughness parameters, which severely limits the reconstruction quality. Please refer to Sec. 7.1 for more details.

Similar to previous work on neural acquisition, we link the lighting patterns with the network in a differentiable fashion: measurements of the reflected radiances under projected lighting patterns are modeled as dot products between the physical lumitexel at the same view and the patterns (Eq. 1). This allows the joint optimization of illumination conditions and the network towards optimal reconstruction quality.

Loss. Our loss function measures the squared difference between the predicted lumitexel and its label:

$$L = \Sigma_l [\log(1 + m(l)) - \log(1 + \tilde{m}(l))]^2. \quad (3)$$

Here $m/\tilde{m}$ represents the predicted/ground-truth lumitexel, respectively. A log transform is performed to compress the high dynamic range in the reflectance.

Training. Our network is implemented with PyTorch and trained using the Adam optimizer, with mini-batches of 50, a momentum of 0.9 and a learning rate of 1×10^{-4} for 0.5 million iterations. Xavier initialization is applied. Based on the GGX BRDF model and the calibration data of the device, we generate 200 million virtual lumitexels as training data (Eq. 1), by randomly sampling the position, the local frame, as well as BRDF parameters. Please refer to [Kang et al. 2018] for details.

For robustness in physical acquisition, we apply dropout regularization with a rate of 30% to all layers except the last, and perturb the synthetic measurements with a multiplicative Gaussian noise ($\mu = 1, \sigma = 5\%$), similar to [Kang et al. 2019]. We add additional training noise to the secondary view to increase the robustness of the network, in case the correspondences are not accurate at test time. When generating the lumitexels in the secondary view, we replace 30% of parameters with perturbed ρ_d/ρ_s and α_x/α_y, by multiplying with a Gaussian noise ($\mu = 1$, $\sigma = 5\%$) and perturbing n with a random orthogonal vector, whose length is drawn from a Gaussian distribution ($\mu = 0$, $\sigma = 15\%$). We also replace 10% of parameters with completely random values, so that the network learns to deal with the challenging cases with inaccurate correspondences.

To decouple albedo and shape components in the latent vector, we train in pairs of samples. Suppose we randomly sample a pair of reflectance $\{\rho_d^i, \rho_s^i, \alpha_x^i, \alpha_y^i, ...\}_{i=1,2}$. We exchange their albedos to obtain two new sets of parameters $\{\rho_d^2, \rho_s^2, \alpha_x^1, \alpha_y^1, ...\}$ and $\{\rho_d^1, \rho_s^1, \alpha_x^2, \alpha_y^2, ...\}$. At the same time, we also swap the corresponding albedo components, and encourage the predicted lumitexels to be similar to the ones generated with the new sets of parameters. The sum of the losses on the above 4 lumitexels will be used as the total loss.

7 ACQUISITION PROCESSING

7.1 Fine-tuning

While the quality of the direct output of our network is often satisfactory, we further push it via sample-specific fine-tuning to achieve photo-grade reconstruction with a variety of view and lighting conditions. This is crucial in building a database, which requires robust, high-quality reconstructions for a large number of diverse samples.

Specifically, we capture additional photographs of the sample at the primary view under 64 randomly sampled effective linear lights, implemented with our setup (Fig. 5). Note that various types of lighting patterns are tested, including point, area and learned patterns. We find that linear lights strike a good balance between acquisition time and quality after fine-tuning (a comparison with learned patterns is shown in Fig. 7). Next, per-pixel fine-tuning of the latent vector is performed with respect to 63 photographs (the remaining one from all 64 photographs is left for validation): the loss is the error between the physical measurements and the simulated measurements computed from the latent vector with our network and Eq. 1. This process consists of 3 steps. (1) We optimize a gray-scale 8+48 = 56D latent vector against the gray-scale version of photographs, taking the latent vector computed by our network as the initialization. (2) For each RGB channel, we fix the shared 48D shape components, and use the same gray-scale albedo components as the initial values, to optimize against one channel of the photographs. (3) The shared 48D shape components and the RGB 8×3 = 24D albedo components are simultaneously refined, resulting in an RGB latent vector of 24+48 = 72D. Note that our latent dimension is determined via extensive experiments, ranging from 128 (16D albedo + 112D highlight shape) to 36 (6D albedo + 30D highlight shape). We find that 8D is needed to represent a single-channel albedo to avoid color artifacts, and 48D is needed to faithfully represent the highlight.

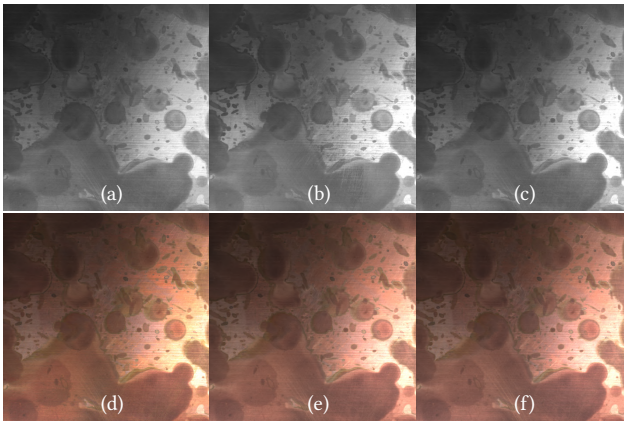


Fig. 6. Visualization of the results after each processing step. (a) the gray-scale version of the photograph, (b) network output, the result after each fine-tuning step (c-e), and (f) the original photograph.

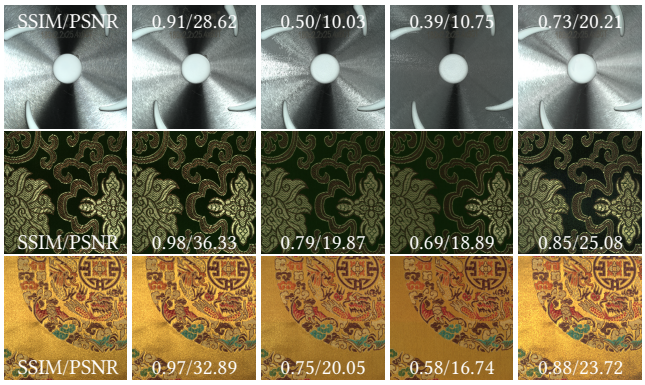


Fig. 7. Impact of lighting patterns and initialization for fine-tuning. From the left column to right: photographs, fine-tuned results under linear lights with network-inferred latent initialization (our approach), direct fine-tuned results under linear lighting patterns, direct fine-tuned results under learned lighting patterns, and direct network output without fine-tuning.

We use Adam optimizer with mini-batches of 100 and a momentum of 0.9 in all steps. The learning rate is 0.02, 0.05 and 0.02, and the iteration number is set to 100, 50 and 200, for 3 steps respectively. These hyper-parameters are determined from experiments. The intermediate results after each step are shown in Fig. 6.

Fig. 7 demonstrates the quality improvement of fine-tuning on challenging physical samples. The figure also shows that fine-tuning only leads to inferior results; it is crucial for our network to provide decent starting points. Therefore, both network inference and fine-tuning are essential to the quality of our database.

7.2 Fitting

The RGB latent vector from fine-tuning, essentially describing a 2D lumitexel, can be converted to a standard 4D BRDF for use in any conventional graphics pipeline. While we select GGX in this paper, other forms of analytical or neural representations can also be employed.

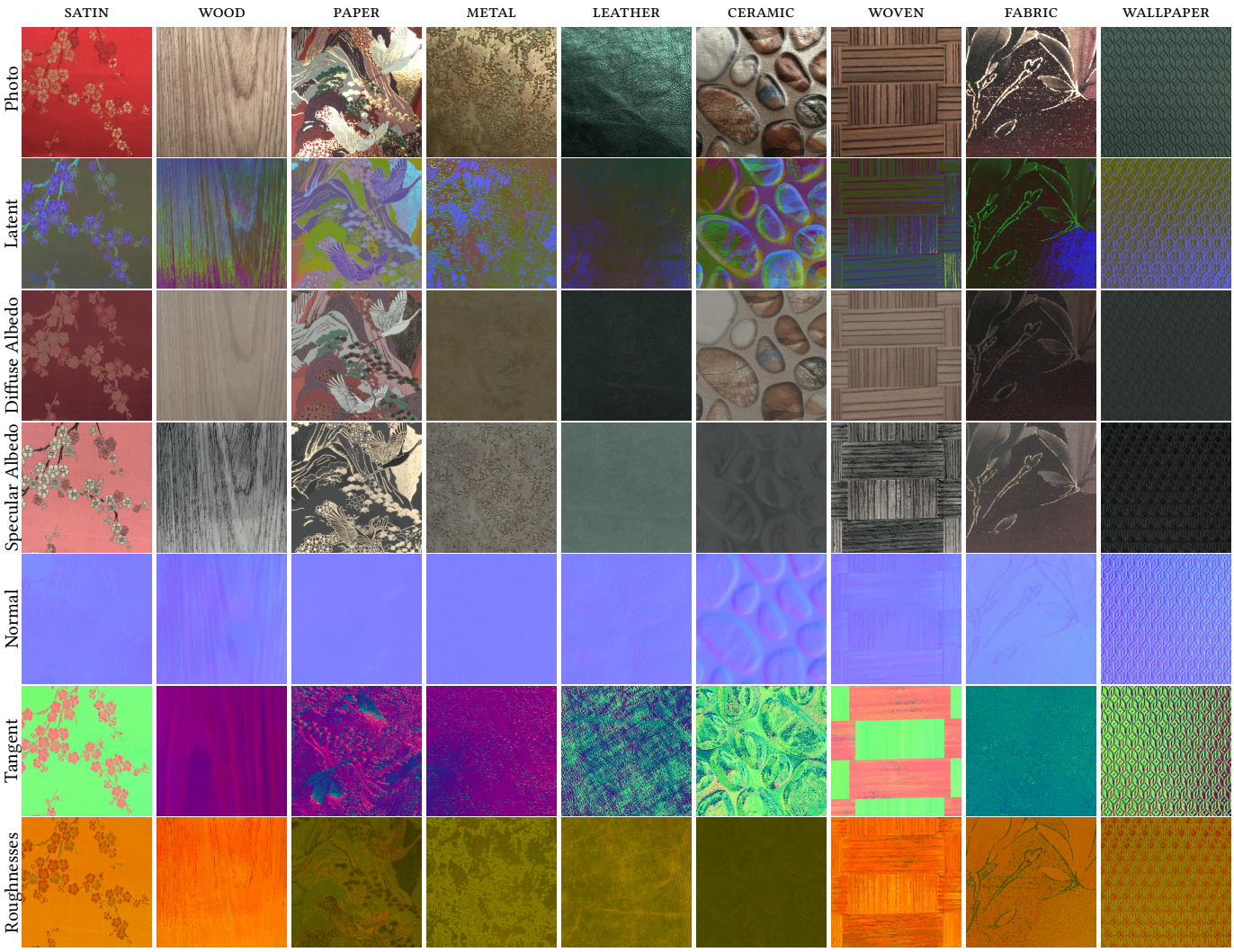


Fig. 8. Visualization of database examples in each category. For images from the top row to bottom in each column: a photograph of the physical sample, the visualization of latent vectors, and the fitting results to the GGX model. Latent vectors are projected to 3 dimensions via PCA for visualization. Each normal/tangent is added with $(1, 1, 1)$ and then divided by 2 to fit into the range of $[0, 1]^3$. The roughness α_x/α_y is visualized in the red/green channel.

We perform the fitting via differentiable rendering, with a loss that minimizes the differences between the lumitexel computed from the RGB latent vector and the one computed from GGX parameters. Note that we do not directly optimize GGX parameters. Instead, we reparameterize the GGX model plus the local frame with 16 neural parameters and jointly train a 4-layer fully-connected network that transforms the neural parameters into GGX ones and the local frame [Xu et al. 2023], for each database sample. Compared with the original model, this sample-specific neural BRDF reparameterization is more amenable for deep learning and results in higher-quality reconstructions. The network is trained for 80K iterations with a learning rate of 10^{-3} and a batch size of 128. By the end of the training, we obtain the fitting result as a by-product, and store them as final texture maps. Note that some discontinuities can be observed in the tangent maps (e.g., Fig. 8). This is mainly due to the isotropic nature at certain locations: in an ideally isotropic case, any unit vector orthogonal to the normal can be selected as the tangent, since it has no effect on the final rendering.

We observe that in certain cases, the reflectance quality slightly degrades, after fitting to the GGX model from the latent vectors, as shown in Fig. 9. Therefore, we store the intermediate latent vectors along with GGX parameters in the database, and leave it for future work to find the optimal representation for our captured data (e.g., specialized models [Jin et al. 2022; Liu et al. 2016]) or recent neural ones [Fan et al. 2022; Zheng et al. 2021].

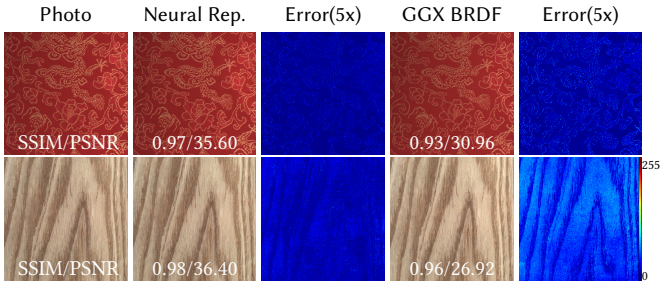


Fig. 9. Comparison of GGX BRDF and our neural representation. For each row, the 1st column is a photograph. The 2nd and 4th columns are the rendering results of neural and GGX BRDF representation. The 3rd and 5th columns are color-coded differences (multiplied by 5).

7.3 Discussion

While fitting our sparse measurements from scratch seems to be a viable option, it is not feasible in practice due to the highly under-constrained nature of the problem [Gao et al. 2019]. Bridging the information gap between sparse measurements and the final result requires additional prior knowledge. In our case, the trained network serves as an effective learned prior for the reconstruction. On the other hand, direct parameter regression is not suitable either. Existing techniques cannot regress anisotropic BRDF parameters with high precision, with tangent/roughnesses being particularly challenging, according to our experience.

8 ADDITIONAL DETAILS

Two-view Alignment. To prepare the input for our network, we align the secondary view to the primary in a coarse-to-fine manner. First, 2-view photographs of the sample under uniform lighting are taken. With the help of ARTags, we compute a homography matrix to roughly warp the secondary view to the primary. To further improve the precision, we compute dense 128D SIFT features [Liu et al. 2011] on the primary view and the warped image. Finally, PatchMatch [Barnes et al. 2009] based on the SIFT feature is applied (iterations=10, patch size=7, search radius=50) to establish more accurate correspondences between the two views. The importance of careful 2-view alignment over the final reconstructions is demonstrated in Fig. 10.

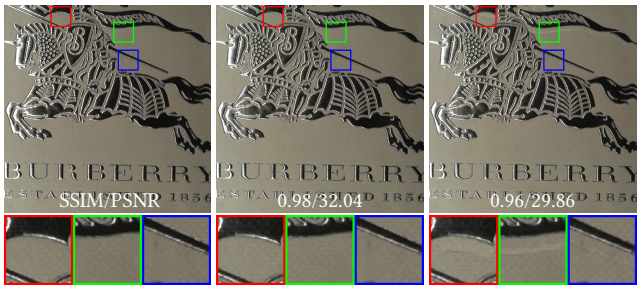


Fig. 10. Impact of accurate 2-view correspondences. From the left to right: a photograph, the rendering result of a reconstructed sample with/without dense SIFT feature matching.

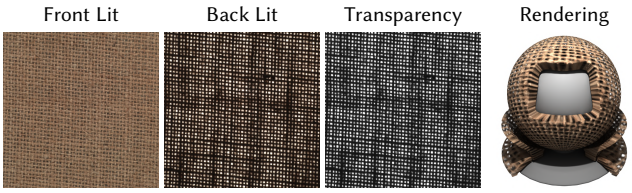


Fig. 11. Transparency Estimation. From left to right, the photograph lit from above the upper hemisphere of the sample/from the bottom side of the setup, the computed transparency map, and the rendering result on a material ball with the fitting results and the transparency map.

Transparency Estimation. We follow the work of [Gardner et al. 2003] to compute per-pixel transparency for each physical sample, by taking an additional image with all 4,096 LEDs from the bottom side of the setup switched on. The measurement at each pixel in this image, divided by its counterpart in another pre-captured photograph with the sample removed, yields the percentage of light transmitted. We store this result as a transparency map. An example is illustrated in Fig. 11.

9 DATABASE

With our integrated system, we build OpenSVBRDF, the first large-scale database of measured anisotropic SVBRDFs. Readers are strongly encouraged to first view the accompanying video, which demonstrates the rendering results of 135 samples (out of 1,000 total samples) with novel, changing view and lighting conditions.

Physical Samples. We collect them by purchasing sample books or individual samples from online vendors, most of which come from manufacturers that supply home decoration/furniture/fashion industry.

Timing. All computation is done on a workstation with dual Intel Xeon 4210 CPUs, 256GB DDR4 memory and 4 NVIDIA GeForce RTX 3090 GPUs. It takes about 40 hours to train our network, with 0.5 million iterations.

For each physical sample, it takes about 15 minutes to capture all 193 HDR photographs using exposure bracketing: 2 minutes for 64×2 (views) photographs under learned patterns, 12 minutes for 64 images under effective linear light sources and 15 seconds for 1 image with bottom-side LEDs on for transparency estimation. For processing, it takes on average 5 minutes for our network to decode all lumitexels from valid measurements, 1.5 hours for fine-tuning and 20 minutes for fitting GGX parameters and local frames.

Organization. We group all 1,000 samples into 9 categories (see Tab. 1 for details). For each sample in the database, we store 193 raw HDR photographs (15GB), its intermediate neural representation (290MB), as well as 6 texture maps (ρ_d,ρ_s,α_x/α_y,n,t and β) representing GGX

Table 1. Database sample distribution in different categories.

Category	#	Category	#	Category	#
SATIN	120	WOOD	300	PAPER	100
METAL	35	LEATHER	180	CERAMIC	34
WOVEN	30	FABRIC	105	WALLPAPER	96

parameters, local frames and transparency (55MB). The spatial resolution is 1,024×1,024 for both the neural representation and texture maps. Please refer to Fig. 8 for a visualization of examples in each category. In addition, we record the physical dimensions (width/height) as well as an online link that can be used to order the physical sample as meta-data.

Statistics. Our database makes it possible for us to compute and analyze the distributions of different GGX parameters across samples in each category/all samples, as shown in Fig. 12. Note that we have reconstructed 1,000×1,024×1,024 BRDFs in total, which is more than one billion. It is interesting to note the differences in the distributions. For example, the asymmetry between the distributions of α_x and α_y in SATIN/WOOD/WOVEN suggests the strong anisotropy in these categories; ρ_d for METAL is concentrated on the lower end, which agrees with common observations; the distributions of α_x, α_y and ρ_s suggest a narrow, relatively weak highlight in CERAMICS,

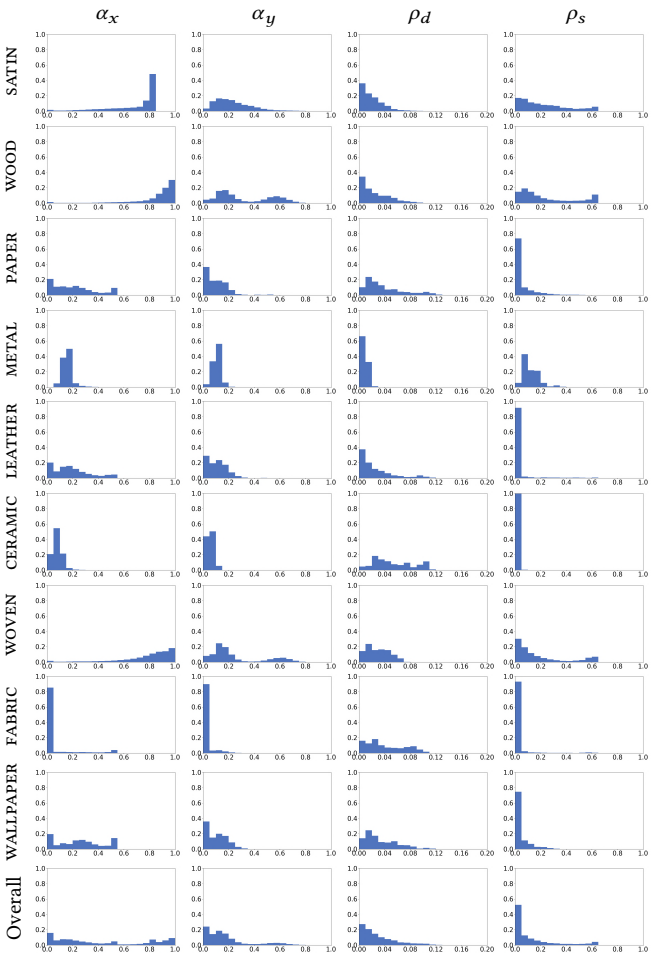


Fig. 12. Visualization of the statistics of the fitted GGX BRDF parameters. The x-axis is the range of each parameter, and the y-axis is the percentage. We calculate the statistics on each category separately, as well as over all samples in the database.

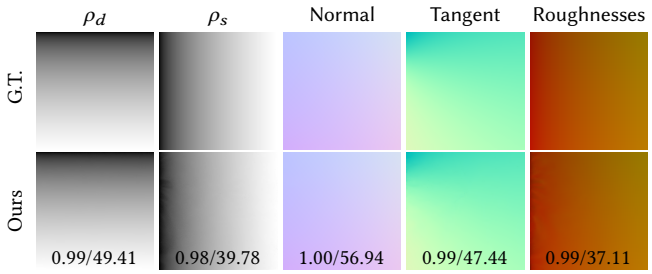


Fig. 13. Reconstruction results of a synthetic anisotropic SVBRDF. The first row shows various GGX parameter maps of the sample, and the second displays our reconstruction. The quantitative error of our results with respect to the ground-truth are reported in the bottom in SSIM/PSNR.

which actually comes from the glazing layer often found in these materials.

10 VALIDATIONS

10.1 Latent Optimization

We validate the quality of the latent optimization (Sec. 7.1) by comparing the rendering with the latent result against a photograph under a novel effective linear light, which is not used in training. The average SSIM between the photograph and the rendering across all 1,000 samples in our database is 0.98. For visual inspection, please refer to Fig. 15, where one sample from each category is displayed.

10.2 Reflectance Fitting

Correctness. We perform a virtual acquisition and reconstruction experiment on a challenging synthetic anisotropic SVBRDF to validate the correctness of the final fitting results. The synthetic sample consists of a mixture of different ramps of GGX parameters, as shown in the first row of Fig. 13. We apply physically-based rendering to simulate the measurements of this sample with a virtual device whose parameters are identical to the real one, and then reconstruct from the measurements and fit GGX parameters, according to the pipeline described in Fig. 2. Our approach almost exactly recovers all parameters (Fig. 13).

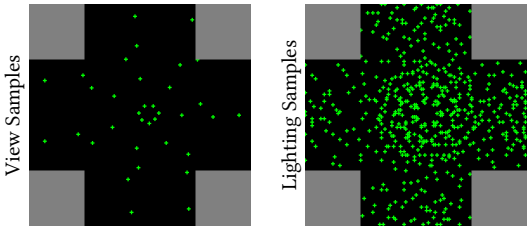


Fig. 14. Visualization of densely sampled, effective view and lighting conditions, which are used in the validation experiments described in the supplemental material. A hemi-cube parameterizaion of the unit upper hemisphere is used. Each sample is marked with a green cross. Note that for lighting, each sample represents a direction towards the center of a group of 8×8 LEDs.

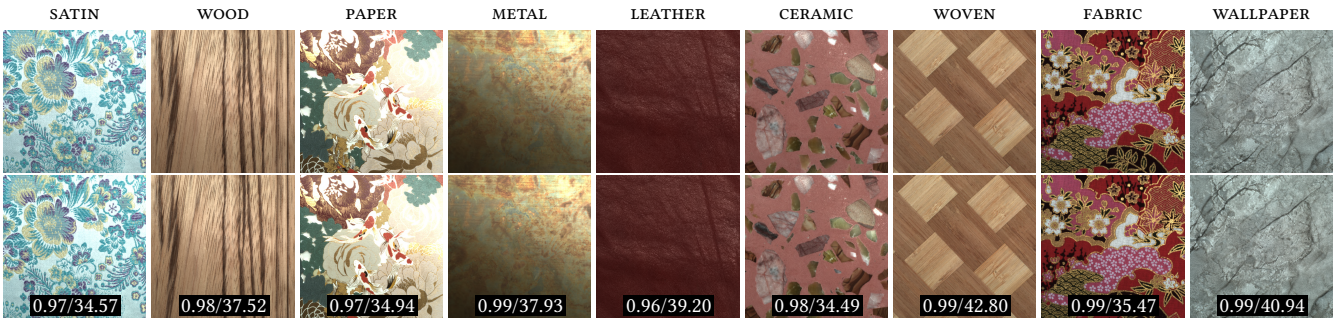


Fig. 15. Photo validation results. The top row shows photographs of physical samples, and the bottom displays the rendered images of the reconstructed results from our neural representation. The quantitative errors of our results with respect to the photographs are reported in the bottom row in SSIM/PSNR.

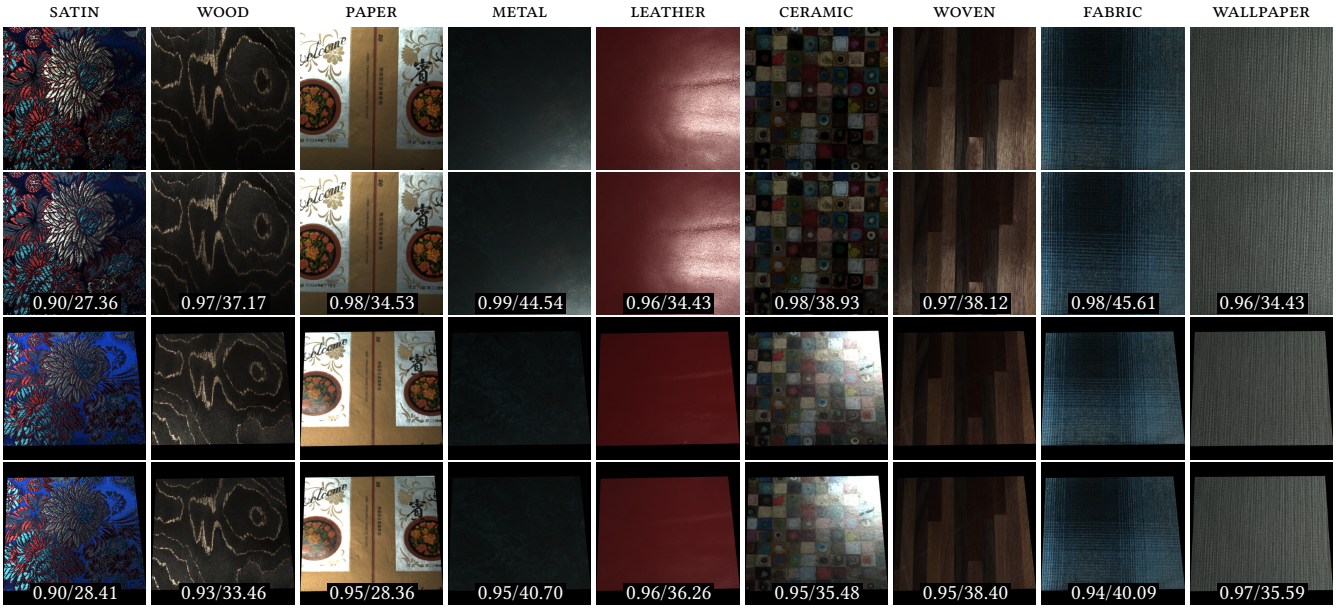


Fig. 16. Photo validation of our BRDF fitting results at two views. The first/third row shows photographs of physical samples at the primary/secondary view, and the second/fourth lists the corresponding renderings using our final BRDF parameters. The quantitative errors of our results with respect to the photographs are reported in the bottom of each rendered image in SSIM/PSNR.

Primary & Secondary Views. We also validate the rendering results of our reconstructions against photographs under novel lighting conditions from both the primary and secondary views in Fig. 16.

Additional View & Lighting Conditions. To more thoroughly assess the result quality of our approach, we further conduct photograph validation experiments on 560 pairs of view and lighting conditions that vary considerably over the upper hemisphere. The distributions of effective view/lighting conditions are visualized in Fig. 14. The average SSIM over all pairs of view and lighting conditions ranges from 0.90 to 0.97 among all 9 physical samples. Due to the space limit of the main paper, please refer to the supplemental material for more details on our experiments and results.

11 APPLICATIONS

We demonstrate the value of our database in the following three applications.

11.1 Material Generation

We follow MaterialGAN [Guo et al. 2020] to train StyleGAN2 [Karras et al. 2020] on our captured data, to implicitly learn an SVBRDF manifold for novel material generation and material morphing. Two modifications are made: first, we increase the output resolution from 256×256 to 512×512 to make better use of our high-resolution database; second, we extend their work to support anisotropic reflectance.

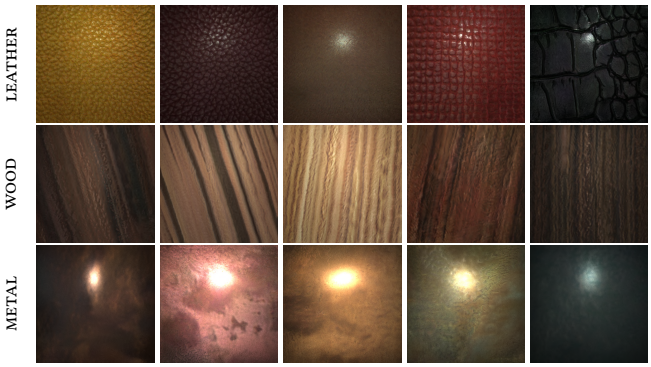


Fig. 17. Rendering results of the generated materials by randomly sampling in the latent space of each MaterialGAN trained on a different category.

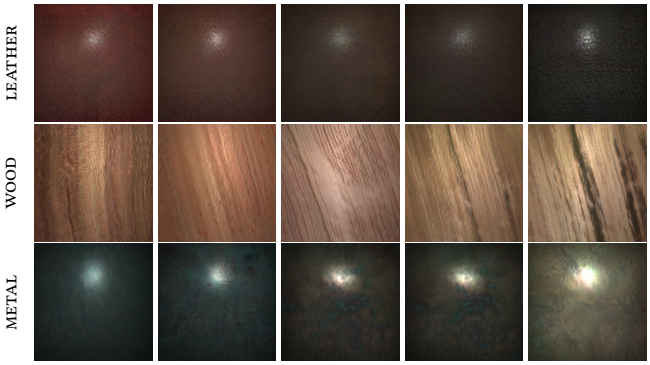


Fig. 18. Material interpolation in the latent space of each MaterialGAN trained on a different category. Each row shows plausible morphing results between pairs of generated materials.

Specifically, we train a StyleGAN2 for each of the three categories in OpenSVBRDF: LEATHER, WOOD and METAL. To generate sufficient training data, we take each database sample with a resolution of 1,024×1,024, and randomly crop, scale and rotate to obtain 30 smaller training SVBRDFs at a 512×512 resolution. The training texture maps are stacked in 14 channels (3 for ρ_d, 3 for ρ_d, 2 for α_x and α_y, 3 for n and 3 for t). We slightly modify the original StyleGAN2 to output 14 channels as well. For normal and tangent, we add an extra normalization layer to force the network to output unit vectors. We use the official PyTorch implementation of StyleGAN2 and the same training settings as MaterialGAN. The discriminator is shown 4 million images. It takes about 36 hours to train a StyleGAN2 for one category of samples.

Fig. 17 displays the rendering results of a variety of SVBRDFs, generated by randomly sampling the MaterialGAN latent space. Note that we use the Gram-Schmidt process to re-orthogonalize the output normal and tangent. In Fig. 18, we also show plausible morphing results between pairs of generated materials, by interpolating in the latent space of each category.

11.2 Material Classification

We further apply OpenSVBRDF to material classification with uncontrolled and controlled illumination.

First, we use our database to enhance one state-of-the-art single-image material classification technique [Mao et al. 2021], originally trained and tested on the Flickr Material Dataset (FMD) [Sharan et al. 2009]. Specifically, we follow [Weinmann et al. 2014] to synthesize training images, by rendering our database samples from 3 selected categories (detailed below) with randomly generated view and lighting conditions. In training a 3-category classifier, these additional images will be used with a probability of 30% and the original images from FMD with 70%. In comparison, we also train a classifier with augmented data from UBO2014 [Weinmann et al. 2014] in exactly the same way described above. Note that the categories FABRIC, WOOD and LEATHER are selected, because they are the only 3 common ones among FMD, UBO2014 and OpenSVBRDF. The average test accuracies on the same 10 test sets are 89.37% (original, trained on FMD only), 90.74% (augmented with UBO2014) and 92.80% (augmented with ours). The results show that the rich data in OpenSVBRDF is helpful to boost the performance of single-image material classification.

Next, we exploit OpenSVBRDF for material classification with controlled lighting, a task simply not possible with a conventional 2D image database like FMD. First, our database samples in the aforementioned 3 categories are split for training/test with a ratio of 1:1. We then compute the lighting patterns to perform one-vs-all classification among 3 categories as in [Gu and Liu 2012]. Specifically, we train 3 discriminative lighting patterns $w_{1,2,3}$ for 3 categories with simple linear SVM classifiers. All negative values are moved to a 4th pattern, resulting in $w^- = -\min(0, w_1, w_2, w_3)$, and $w_k^+ = w_k + w^-, k = 1, 2, 3$. Per-pixel classification can be conducted, taking as input the pixel measurements under 4 learned patterns $w_{1,2,3}^+$ and w^-. Finally, we determine the image classification result

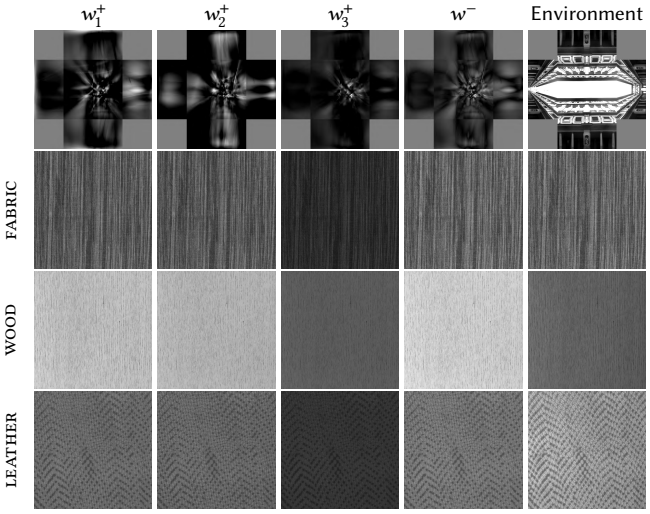


Fig. 19. Examples of material classification with active lighting. The top row is the computed lighting patterns to perform one-vs-all classification among 3 categories, and an environment map for rendering images with passive illumination. The next three rows are three cases, which a single-image deep classifier fails to produce correct results while a linear SVM classifier succeeds using 4 input images under active lighting for each case.

from the majority of per-pixel results. In comparison, a deep classifier [Mao et al. 2021] is solely trained on synthesized images from the same training samples with random lighting conditions. The average test accuracies are 83.25% (uncontrolled illumination, deep network) and 90.10% (learned lighting patterns, linear SVM). Moreover, three examples are shown in Fig. 19: while the deep classifier fails to classify any of them correctly, learned patterns probe more physical information which lead to correct results in all cases using simple SVM classifiers.

11.3 Optimized BRDF Sampling

Finally, we harness the implicit prior in our dataset for optimized anisotropic BRDF sampling. Two common forms of acquisition geometries are studied: directional sampling and illumination multiplexing.

We first investigate BRDF sampling using an extremely sparse number of lighting directions and a fixed view direction with an elevation angle of 45°. Specifically, we generate anisotropic BRDFs with random ρ_d, ρ_s, α_x and α_y sampled from our database. The normal/tangent/position is fixed to a canonical value, as is common in previous work [Filip et al. 2014]. These BRDFs are split into training/test with a ratio of 4:1. Next, we modify [Kang et al. 2018] with [Xu et al. 2018] to obtain an autoencoder that takes as input the BRDF values at as few as 4 learned lighting directions and outputs a 2D BRDF slice at the fixed view direction. After training, the average test loss (Eq. 3) is 12.67. In comparison, we replace the training data with ρ_d and ρ_s sampled uniformly at random, and α_x and α_y sampled uniformly at random on the log scale, in the range of the parameters computed from the entire database, similar to previous work like [Kang et al. 2018]. The autoencoder trained with such data results in an average test loss of 28.35. Fig. 20 visualizes the performance of two autoencoders: the one trained with the database distribution produces a considerably higher quality, both

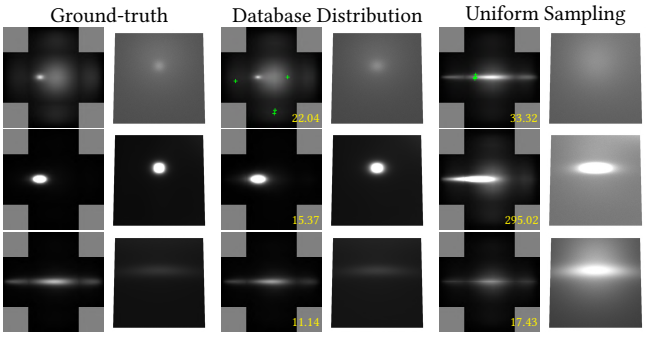


Fig. 20. Reconstruction results of two autoencoders using optimized directional sampling. For each pair of images, the left is a lumitexel and the right the rendering result from a BRDF fitted to the lumitexel. From the left to right: ground-truth, the reconstructions of [Kang et al. 2018] trained with the BRDF distribution in our database/GGX parameters sampled uniformly at random. The numerical errors, computed using Eq. 3, are listed in the bottom-right corner of each related image. All predicted lumitexels are the direct network output prior to fitting. The optimized lighting directions are marked as green crosses.

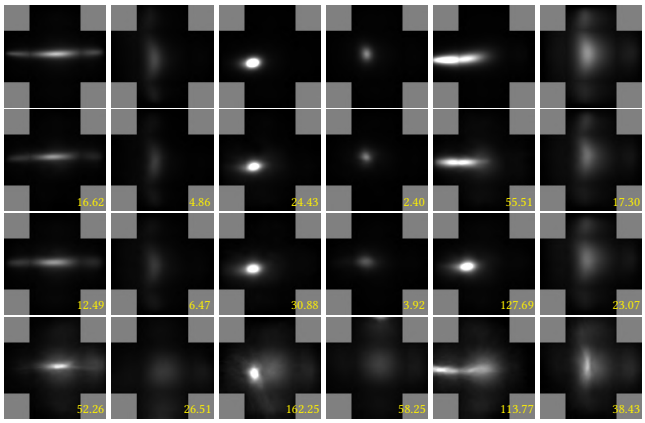


Fig. 21. Reconstruction results of three autoencoders using optimized illumination multiplexing. The top row is the ground-truths, and the 2nd/3rd row is the predicted lumitexels of autoencoders [Kang et al. 2018] trained using the database distribution with 32/10 lighting patterns. The 4th row is the results of the autoencoder trained with GGX parameters sampled uniformly at random with 32 lighting patterns.

qualitatively and quantitatively; its optimized lighting directions are also more evenly distributed in the angular domain.

Moreover, we apply a similar idea to improve illumination - multiplexed sampling of anisotropic BRDFs, when a dense set of light sources are available to be programmed simultaneously. Similar to the previous experiment on directional sampling, we generate two training sets and one test set, with the exception that the position is randomly sampled and the normal/tangent is either sampled according to our database distribution, or randomly sampled as is common in related work [Kang et al. 2018]. We optimize different versions of autoencoders [Kang et al. 2018] with the two training sets. The test losses are 27.69 (#patterns=32, uniform sampling), 23.81 (#patterns=10, database distribution), 10.78 (#patterns=32, database distribution). With the same number of lighting patterns, our database distribution helps improve the reconstruction quality considerably, compared with a uniform sampling; also the number of lighting patterns can be reduced to less than 1/3 for similar-quality results. Please refer to Fig. 21 for a visualization.

12 LIMITATIONS AND FUTURE WORK

This work is subject to a number of limitations. First, our samples are limited to be near-planar, as we cannot handle large structures that cause inter-view visibility changes. Increasing the number of cameras along with network modifications might be a viable solution. Next, following existing work, our modeling of sub-surface scattering is highly simplified. It will be useful to perform differentiable rendering with a more advanced model like BSSRDF. In addition, the current selection of lighting patterns for fine-tuning is empirical. It will be intriguing to systematically investigate optimal fine-tuning patterns, to further reduce the number of captured images and improve reconstruction quality.

In the future, we plan to expand the database with more samples that exhibit diverse appearances. We are also interested in setting

up open challenges (e.g., on material estimation, classification, generation): standardized benchmarks based on our growing dataset might be useful to quantitatively evaluate existing as well as future research on a common ground. Finally, we are excited to build a large-scale measured database with both high-quality appearance and shape in the near future.

ACKNOWLEDGMENTS

We would like to thank anonymous reviewers for their comments, Yaxin Yu and Kaizhang Kang for help in building the prototype, and Yue Dong, Xin Tong, Julie Dorsey and Holly Rushmeier for their support. This work is partially supported by NSF China (62022072 & 62227806), Zhejiang Provincial Key R&D Program (2022C01057), the Fundamental Research Funds for the Central Universities, the XPLORER PRIZE and Information Technology Center and State Key Lab of CAD&CG, Zhejiang University.

REFERENCES

Miika Aittala, Timo Aila, and Jaakko Lehtinen. 2016. Reflectance Modeling by Neural Texture Synthesis. *ACM Trans. Graph.* 35, 4, Article 65 (July 2016), 13 pages.

Miika Aittala, Tim Weyrich, and Jaakko Lehtinen. 2013. Practical SVBRDF Capture in the Frequency Domain. *ACM Trans. Graph.* 32, 4, Article 110 (July 2013), 12 pages.

Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B Goldman. 2009. PatchMatch: A randomized correspondence algorithm for structural image editing. *ACM Trans. Graph.* 28, 3 (2009), 24.

Guojun Chen, Yue Dong, Pieter Peers, Jiawan Zhang, and Xin Tong. 2014. Reflectance Scanning: Estimating Shading Frame and BRDF with Generalized Linear Light Sources. *ACM Trans. Graph.* 33, 4, Article 117 (July 2014), 11 pages.

Kristin J. Dana, Bram van Ginneken, Shree K. Nayar, and Jan J. Koenderink. 1999. Reflectance and Texture of Real-World Surfaces. *ACM Trans. Graph.* 18, 1 (jan 1999), 1–34.

Valentin Deschaintre, Miika Aittala, Frédo Durand, George Drettakis, and Adrien Bousseau. 2018. Single-Image SVBRDF Capture with a Rendering-Aware Deep Network. *ACM Trans. Graph.* 37, 128 (aug 2018), 15.

Yue Dong. 2019. Deep appearance modeling: A survey. *Visual Informatics* 3, 2 (2019), 59–68.

Yue Dong, Jiaping Wang, Xin Tong, John Snyder, Yanxiang Lan, Moshe Ben-Ezra, and Baining Guo. 2010. Manifold Bootstrapping for SVBRDF Capture. *ACM Trans. Graph.* 29, 4, Article 98 (July 2010), 10 pages.

Julie Dorsey, Holly Rushmeier, and François Sillion. 2010. *Digital modeling of material appearance.* Elsevier.

Jonathan Dupuy and Wenzel Jakob. 2018. An Adaptive Parameterization for Efficient Material Acquisition and Rendering. *ACM Trans. Graph.* 37, 6 (2018), 14 pages.

Jiahui Fan, Beibei Wang, Milos Hasan, Jian Yang, and Ling-Qi Yan. 2022. Neural Layered BRDFs. In *ACM SIGGRAPH 2022 Conference Proceedings*. Article 4, 8 pages.

Mark Fiala. 2005. ARTag, a fiducial marker system using digital techniques. In *CVPR*, Vol. 2. 590–596.

J. Filip and R. Vávra. 2014. Template-Based Sampling of Anisotropic BRDFs. *CGF* 33, 7 (2014), 91–99.

Jirí Filip, Radomír Vávra, and Michal Havlícek. 2014. Effective Acquisition of Dense Anisotropic BRDF. In *Pattern Recognition*. 2047–2052.

Duan Gao, Xiao Li, Yue Dong, Pieter Peers, Kun Xu, and Xin Tong. 2019. Deep Inverse Rendering for High-resolution SVBRDF Estimation from an Arbitrary Number of Images. *ACM Trans. Graph.* 38, 4, Article 134 (July 2019), 15 pages.

Andrew Gardner, Chris Tchou, Tim Hawkins, and Paul Debevec. 2003. Linear light source reflectometry. *ACM Trans. Graph.* 22, 3 (2003), 749–758.

Abhijeet Ghosh, Tongbo Chen, Pieter Peers, Cyrus A. Wilson, and Paul Debevec. 2009. Estimating Specular Roughness and Anisotropy from Second Order Spherical Gradient Illumination. *CGF* 28, 4 (2009), 1161–1170.

Jinwei Gu and Chao Liu. 2012. Discriminative illumination: Per-pixel classification of raw materials based on optimal projections of spectral BRDF. In *CVPR*. 797–804.

Darya Guarnera, Giuseppe C. Guarnera, Abhijeet Ghosh, Cornelia Denk, and Mashhuda Glencross. 2016. BRDF Representation and Acquisition. *CGF* 35, 2 (2016), 625–650.

Johannes Günther, Tongbo Chen, Michael Goesele, Ingo Wald, and Hans-Peter Seidel. 2005. Efficient acquisition and realistic rendering of car paint. In *VMV*, Vol. 5. 487–494.

Jie Guo, Shuichang Lai, Chengzhi Tao, Yuelong Cai, Lei Wang, Yanwen Guo, and Ling-Qi Yan. 2021. Highlight-Aware Two-Stream Network for Single-Image SVBRDF Acquisition. *ACM Trans. Graph.* 40, 4, Article 123 (jul 2021), 14 pages.

Yu Guo, Cameron Smith, Miloš Hašan, Kalyan Sunkavalli, and Shuang Zhao. 2020. MaterialGAN: reflectance capture using a generative SVBRDF model. *ACM Trans. Graph.* 39, 6 (2020), 1–13.

Zhuo Hui, Kalyan Sunkavalli, Joon-Young Lee, Sunil Hadap, Jian Wang, and Aswin C. Sankaranarayanan. 2017. Reflectance Capture Using Univariate Sampling of BRDFs. In *ICCV*.

Wenzel Jakob, Sébastien Speierer, Nicolas Roussel, Merlin Nimier-David, Delio Vicini, Tizian Zeltner, Baptiste Nicolet, Miguel Crespo, Vincent Leroy, and Ziyi Zhang. 2022. *Mitsuba 3 renderer.*

Wenhua Jin, Beibei Wang, Milos Hasan, Yu Guo, Steve Marschner, and Ling-Qi Yan. 2022. Woven fabric capture from a single photo. In *SIGGRAPH Asia 2022 Conference Papers*. 1–8.

Kaizhang Kang, Zimin Chen, Jiaping Wang, Kun Zhou, and Hongzhi Wu. 2018. Efficient Reflectance Capture Using an Autoencoder. *ACM Trans. Graph.* 37, 4, Article 127 (July 2018), 10 pages.

Kaizhang Kang, Cihui Xie, Chengan He, Mingqi Yi, Minyi Gu, Zimin Chen, Kun Zhou, and Hongzhi Wu. 2019. Learning Efficient Illumination Multiplexing for Joint Capture of Reflectance and Shape. *ACM Trans. Graph.* 38, 6, Article 165 (Nov. 2019), 12 pages.

Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020. Analyzing and improving the image quality of stylegan. In *CVPR*. 8110–8119.

Jason Lawrence, Aner Ben-Artzi, Christopher DeCoro, Wojciech Matusik, Hanspeter Pfister, Ravi Ramamoorthi, and Szymon Rusinkiewicz. 2006. Inverse shade trees for non-parametric material representation and editing. *ACM Trans. Graph.* 25, 3 (2006), 735–745.

Hendrik P. A. Lensch, Jan Kautz, Michael Goesele, Wolfgang Heidrich, and Hans-Peter Seidel. 2003. Image-based Reconstruction of Spatial Appearance and Geometric Detail. *ACM Trans. Graph.* 22, 2 (April 2003), 234–257.

Xiao Li, Yue Dong, Pieter Peers, and Xin Tong. 2017. Modeling Surface Appearance from a Single Photograph Using Self-augmented Convolutional Neural Networks. *ACM Trans. Graph.* 36, 4, Article 45 (July 2017), 11 pages.

Zhengqin Li, Mohammad Shafiei, Ravi Ramamoorthi, Kalyan Sunkavalli, and Manmohan Chandraker. 2020. Inverse rendering for complex indoor scenes: Shape, spatially-varying lighting and svbrdf from a single image. In *CVPR*. 2475–2484.

Albert Julius Liu, Zhao Dong, Miloš Hašan, and Steve Marschner. 2016. Simulating the Structure and Texture of Solid Wood. *ACM Trans. Graph.* 35, 6, Article 170 (Nov. 2016), 11 pages.

Ce Liu, Jenny Yuen, and Antonio Torralba. 2011. SIFT Flow: Dense Correspondence across Scenes and Its Applications. *IEEE PAMI* 33, 5 (2011), 978–994.

Xiaohe Ma, Kaizhang Kang, Ruisheng Zhu, Hongzhi Wu, and Kun Zhou. 2021. Free-Form Scanning of Non-Planar Appearance with Neural Trace Photography. *ACM Trans. Graph.* 40, 4, Article 124 (jul 2021), 13 pages.

Xiaohe Ma, Yaxin Yu, Hongzhi Wu, and Kun Zhou. 2023. Efficient Reflectance Capture With a Deep Gated Mixture-of-Experts. *IEEE TVCG* (2023), 1–12.

Shangbo Mao, Deepu Rajan, and Liang Tien Chia. 2021. Deep residual pooling network for texture recognition. *Pattern Recognition* 112 (2021), 107817.

Stephen Marschner, Stephen Westin, Eric Lafortune, and Kenneth Torrance. 2000. Image-Based Bidirectional Reflectance Distribution Function Measurement. *Applied optics* 39 (07 2000), 2592–600.

Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. 2003. A Data-driven Reflectance Model. *ACM Trans. Graph.* 22, 3 (July 2003), 759–769.

Giljoo Nam, Joo Ho Lee, Diego Gutierrez, and Min H Kim. 2018. Practical svbrdf acquisition of 3d objects with unstructured flash photography. *ACM Trans. Graph.* 37, 6 (2018), 1–12.

Giljoo Nam, Joo Ho Lee, Hongzhi Wu, Diego Gutierrez, and Min H. Kim. 2016. Simultaneous Acquisition of Microscale Reflectance and Normals. *ACM Trans. Graph.* 35, 6, Article 185 (Nov. 2016), 11 pages.

Jannik Boll Nielsen, Henrik Wann Jensen, and Ravi Ramamoorthi. 2015. On Optimal, Minimal BRDF Sampling for Reflectance Acquisition. *ACM Trans. Graph.* 34, 6, Article 186 (Oct. 2015), 11 pages.

Peiran Ren, Jiaping Wang, John Snyder, Xin Tong, and Baining Guo. 2011. Pocket reflectometry. *ACM Trans. Graph.* 30, 4 (2011), 1–10.

Lavanya Sharan, Ruth Rosenholtz, and Edward Adelson. 2009. Material perception: What can you see in a brief glance? *Journal of Vision* 9, 8 (2009), 784–784.

Borom Tunwattanapong, Graham Fyffe, Paul Graham, Jay Busch, Xueming Yu, Abhijeet Ghosh, and Paul Debevec. 2013. Acquiring Reflectance and Shape from Continuous Spherical Harmonic Illumination. *ACM Trans. Graph.* 32, 4, Article 109 (July 2013), 12 pages.

Bruce Walter, Stephen R. Marschner, Hongsong Li, and Kenneth E. Torrance. 2007. Microfacet Models for Refraction through Rough Surfaces. In *EGSR*.

Jiaping Wang, Shuang Zhao, Xin Tong, John Snyder, and Baining Guo. 2008. Modeling Anisotropic Surface Reflectance with Example-based Microfacet Synthesis. *ACM Trans. Graph.* 27, 3, Article 41 (Aug. 2008), 9 pages.

Michael Weinmann, Juergen Gall, and Reinhard Klein. 2014. Material classification based on training data synthesized using a BTF database. In *ECCV*. 156–171.

Michael Weinmann and Reinhard Klein. 2015. Advances in Geometry and Reflectance Acquisition. In *SIGGRAPH Asia Courses*. Article 1, 71 pages.

Tim Weyrich, Jason Lawrence, Hendrik P. A. Lensch, Szymon Rusinkiewicz, and Todd Zickler. 2009. Principles of Appearance Acquisition and Representation. *Found. Trends. Comput. Graph. Vis.* 4, 2 (2009), 75–191.

Xianmin Xu, Yuxin Lin, Haoyang Zhou, Chong Zeng, Yaxin Yu, Kun Zhou, and Hongzhi Wu. 2023. A unified spatial-angular structured light for single-view acquisition of shape and reflectance. In *CVPR*. 206–215.

Zexiang Xu, Jannik Boll Nielsen, Jiyang Yu, Henrik Wann Jensen, and Ravi Ramamoorthi. 2016. Minimal BRDF Sampling for Two-shot Near-field Reflectance Acquisition. *ACM Trans. Graph.* 35, 6, Article 188 (Nov. 2016), 12 pages.

Zexiang Xu, Kalyan Sunkavalli, Sunil Hadap, and Ravi Ramamoorthi. 2018. Deep image-based relighting from optimal sparse samples. *ACM Trans. Graph.* 37, 4 (2018), 126.

Chuankun Zheng, Ruzhang Zheng, Rui Wang, Shuang Zhao, and Hujun Bao. 2021. A Compact Representation of Measured BRDFs Using Neural Processes. *ACM Trans. Graph.* 41, 2 (2021), 15 pages.