On Neural BRDFs: A Thorough Comparison of State-of-the-Art Approaches

Florian Hofherr[1,2] Bjoern Haefner[† 1,2,3] Daniel Cremers[1,2]
[1]Technical University of Munich [2]Munich Center for Machine Learning [3]NVIDIA

Abstract

The bidirectional reflectance distribution function (BRDF) is an essential tool to capture the complex interaction of light and matter. Recently, several works have employed neural methods for BRDF modeling, following various strategies, ranging from utilizing existing parametric models to purely neural parametrizations. While all methods yield impressive results, a comprehensive comparison of the different approaches is missing in the literature. In this work, we present a thorough evaluation of several approaches, including results for qualitative and quantitative reconstruction quality and an analysis of reciprocity and energy conservation. Moreover, we propose two extensions that can be added to existing approaches: A novel additive combination strategy for neural BRDFs that split the reflectance into a diffuse and a specular part, and an input mapping that ensures reciprocity exactly by construction, while previous approaches only ensure it by soft constraints.

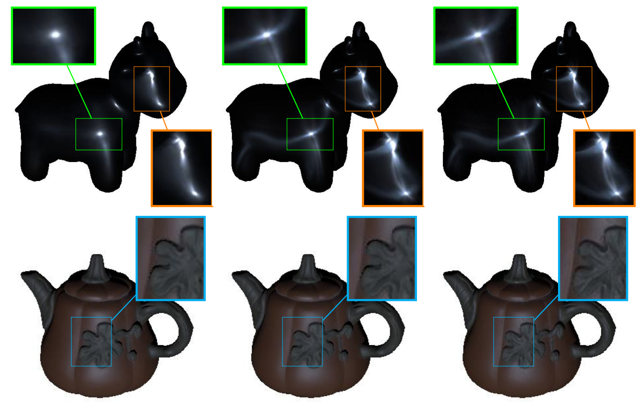

Figure 1. In this work, we present a thorough comparison of neural approaches for BRDF parametrization, including methods based on parametric models like the Disney BRDF [10] and purely neural approaches. We show that while purely neural approaches have advantages for highly specular materials (top row: semi-synthetic, chrome steel from MERL [29]), we find much less difference between the methods for the DiLiGenT-MV real-world dataset [28] (bottom row: pot2 from DiLiGenT-MV).

1. Introduction

The physical principles underlying the interaction of light with matter are diverse and complex. The electromagnetic waves are reflected and refracted at material interfaces, and they are scattered and absorbed within materials. These principles need to be taken into account to obtain realistic rendering results. The standard approach to describe an opaque surface is the *bidirectional reflectance distribution function* (BRDF), which for a given light direction describes the amount of light reflected in a specific view direction.

Realistic BRDF modelling is a long-standing research question and countless methods have been proposed. A popular class of approaches aims to explicitly replicate the physical behavior. Some works are mainly phenomenological [6, 34], others, particularly the microfacet approaches, carefully model the physical principles [11, 40, 41, 47, 48]. While these works rely on a thorough analysis of the underlying physics, they only include a subset of the governing phenomena, limiting them to the modeling choices.

Recently, neural fields have become popular for BRDF modeling, allowing for a continuous and resolution-free representation. While several works ultimately rely on parametric physical models, for which the neural field is used to predict the parameters [5, 7–9, 12, 15, 16, 42, 51, 52, 55], other approaches use neural networks to directly predict the value of the BRDF [13, 39, 43, 54]. Although all neural BRDFs yield impressive results, a thorough comparison between the different approaches is lacking in the literature.

Contribution In this work, we address this shortcoming and perform an exhaustive comparison of different approaches for neural BRDF modeling. In contrast to the setting in most previous works, we deliberately estimate the reflectance for *given* geometry and *calibrated* light to avoid cross-influences from the joint estimation of shape, material

[†]The contribution was done while at TUM.

and lighting conditions. This ensures that the capabilities of the approaches can be compared as unobscured as possible. The results suggest advantages for purely neural methods over approaches based on parametric models in particular for highly specular materials.

Moreover, we propose two extensions that can be added to existing purely neural approaches: A novel input mapping that ensures reciprocity of the BRDF by construction as well as an enhancement for split-based methods that aims at representing the physical behavior more faithfully.

We summarize our main contributions as follows:

1. We perform an extensive comparison of different neural BRDF approaches, including neural fields for parameters of different physical models and direct BRDF prediction by the neural network.
2. We analyze the energy and reciprocity constraint for all methods. The results show that neural approaches do not seem to learn reciprocity from data. We propose a modification to ensure reciprocity by construction.
3. We introduce a novel splitting scheme for the diffuse and specular part of additive neural BRDF models and show that it improves existing methods.

Please see florianhofherr.github.io/neural-brdfs and the supplementary material for additional architecture and evaluation details as well as more experiments.

2. Related Work

BRDF modeling has been studied extensively, and various approaches have been proposed. Comprehensive overviews can be found in [32] and [14].

Phenomenological and Physically-Based BRDF Models Classical works explicitly model the interaction between light and matter by parametric equations, often employing a split into a diffuse component like the Lambertian reflection [26] and a specular component. The Phong and the Blinn-Phong models [6, 34] are popular phenomenological approaches for the specular part, despite their physical incorrectness. Several modifications have been proposed, making them, *e.g.*, energy conserving and anisotropic [4, 25].

Most physically-based models for the specular part are microfacet models, introduced by Torrance and Sparrow [46]. They use statistically distributed small mirror-like microfacets in combination with the Fresnel reflection law. Several extensions to the microfacet model have been presented [11, 17, 40, 41, 47, 48], among those, the popular Disney BRDF [10] and the Oren-Nayar model that generalizes diffuse reflection by Lambertian microfacets [33]. Recently, Ichikawa *et al*. addressed shortcomings in the latter model and proposed an approach based on Fresnel reflectance and transmission that also models polarimetric behavior [19].

Neural Approaches Based on Parametric Models Several neural reflectance methods rely on the extensive research on parametric and physical BRDF models. The idea is to compute the parameters of those models with spatial MLPs for a given input location and to compute the reflectance using the model with the predicted parameters.

The well-established microfacet BRDF [11, 46, 48] has been successfully applied to neural scene reconstruction, where a spatial neural network jointly parametrizes the scene geometry and the parameters for the reflection model [5, 7, 42, 51]. Trained on multi-view images, this allows for novel view synthesis and other downstream tasks like relighting or the extraction of 3D assets. Other works employ the microfacet model for material estimation from a single flash image based on U-Nets [12, 16] and in combination with Generative Adversarial Networks to create a generative reflectance model for material reconstruction [15].

The Disney BRDF [10] is another parametric model that has been used successfully to design neural reflectance methods. Zhang *et al*. combine it with a neural SDF for joint estimation of scene geometry and appearance [52]. Also, it is used in the scene model proposed by Zhang *et al*. which includes indirect illumination [55]. Brahimi *et al*. employ it for scene reconstruction based on volumetric rendering [9] and for uncalibrated point-light photometric stereo [8].

Purely Neural Models In contrast, purely neural methods do not depend on parametric models but use MLPs to predict the BRDF values directly. Two main approaches can be distinguished: While some works use a single MLP to compute the complete BRDF value, others rely on an additive split into a diffuse and a specular part.

Sztrajman *et al*. [43] use a single neural network to represent a non-spatially-varying BRDF, which they train on the MERL data [29]. By using a variational-autoencoder, they compress their neural BRDFs into a latent space. Hu *et al*. [18] employ an autoencoder to compress the MERL BRDFs directly. Zhang *et al*. [54] use a neural BRDF based on an additive split with a specular component pre-trained on the MERL data [29] for neural scene reconstruction under unknown lighting conditions. They demonstrate a slight advantage of their purely neural approach over a microfacet BRDF-based neural model. Sarka *et al*. [39] employ several MLPs to compute an additive split as well as terms for indirect illumination to obtain highly realistic reconstructions of human heads. Among the other methods, they are the only ones who aim to ensure the reciprocity of the BRDF. Zheng *et al*. [50] use a single MLP to model the reflectance for volume rendering based scene reconstruction. In contrast to previous methods, they do not model effects like interreflection and Lambert's cosine law explicitly but have the neural network learn them, which they facilitate by additional *light transport hints* predicted by separate networks.

Neural Methods in Computer Graphics Neural approaches have also been explored in computer graphics to model highly complex materials. Fan *et al*. [13] propose learned operations in a latent space to model the layering of multiple materials. Rainer *et al*. [35, 36] and Rodriguez-Pardo *et al*. [37] employ autoencoders to compress Bidirectional Texture Functions (BTFs) into a latent space, enabling fast test-time encodings. Kuznetsov *et al*. extend an MLP decoder for BTFs by a generalization of classic mipmap pyramids and curvature awareness [23, 24]. Tg *et al*. have investigated neural approaches to model the subsurface scattering in translucent objects [44, 45].

3. Background: BRDF and Rendering

The *bidirectional reflectance distribution function* (BRDF) describes how a surface reflects incoming light. More precisely, the BRDF $f(x, l, v)$ describes the ratio of the radiance reflected in the viewing direction $v \in \mathbb{S}^2$ to the irradiance incident from the light direction $l \in \mathbb{S}^2$ at the position x. In this work, we focus on *spatially varying* BRDF (SVBRDF) where f depends on x. For simplicity, however, we use BRDF and SVBRDF interchangeably.

The *rendering equation* integrates the reflections of the incident irradiances $L_i \geq 0$ over the upper hemisphere $\mathbb{H}$ centered around the surface normal to obtain the total radiance $L_o(x, v)$ at position x in the viewing direction v,

$$L_o(x, v) = \int_{\mathbb{H}} f(x, l, v) L_i(x, l) \cos\theta_l \, \mathrm{d}l. \tag{1}$$

A plausible and physically realistic BRDF needs to fulfill three properties: *positivity* (Eq. (2)), *Helmholtz reciprocity* (Eq. (3)), and *energy conservation* (Eq. (4)),

$$f(x, l, v) \geq 0, \qquad \forall x, \forall l, \forall v, \tag{2}$$

$$f(x, l, v) = f(x, v, l), \qquad \forall x, \forall l, \forall v, \tag{3}$$

$$\int_{\mathbb{H}} f(x, l, v) \cos\theta_v \, \mathrm{d}v \leq 1, \qquad \forall x, \forall l. \tag{4}$$

The first two properties are fulfilled by most state-of-the-art BRDF models [10, 25, 46]. However, only very few models fulfill energy conservation by construction [25]. While for special algorithms like (bidirectional) path-tracing, reciprocity and energy conservation ensure convergence, in most cases, it is sufficient to fulfill them only approximately without noticeable artifacts [1].

Mainly two physical processes are responsible for the reflection of light. These are often modeled as two separate terms within the BRDF: Surface reflection and subsurface scattering, often called specular and diffuse term. In the first case, light is directly mirrored at the surface and creates a view-dependent reflection lobe. In the second case, light enters the surface and is scattered and partially absorbed until a fraction of light is re-emitted. While in general, subsurface scattering is not purely uniform, *i.e*., not Lambertian, but depending on the viewing direction [33], most models still assume a uniform diffuse reflection. Light being reflected at the surface is absent for subsurface scattering [1], *i.e*., one can disjointly split the light used for surface reflection and for subsurface scattering.

A common special case are isotropic BRDFs which we consider in this paper. In this case the reflectance at a point does not change if the object is rotated around the normal, or in other words, if the relative angle between the light and the view direction remains the same. In that case, three angles are sufficient for the BRDF parameterization.

4. Neural BRDFs

This work aims to compare different neural BRDF modeling approaches in a unified manner. We deliberately choose a setting with *given* geometry and *calibrated* lights to avoid confounding effects due to joint estimation of geometry, reflectance, and light. This ensures that the results capture the capabilities of the models as clearly as possible. We use meshes to represent the geometry since they allow for accurate normals. We choose to reconstruct the BRDF from posed images rather than from BRDF measurements since this is the more practical and widely used setting.

In Sec. 4.1, we describe the neural BRDF models considered in this work. Sec. 4.2 describes the angular parametrization of the directions and the input encodings used for the neural networks. Moreover, we propose two extensions for existing approaches to ensure reciprocity by construction and to enhance architectures that use an additive split of diffuse and specular parts in Sec. 4.3 and 4.4.

4.1. Approaches to Neural BRDF Modelling

Since several ideas of previous methods are similar in spirit but implemented with slightly different architectures, we did our best to tune one architecture for each idea to yield optimal results for our data. Note that we focus on the actual BRDF representation; therefore, we do not include parts of the models concerned with the geometry estimation, like *e.g*. additional regularizers.

4.1.1 Neural BRDFs Based on Parametric Models

This class of neural BRDFs is based on physically-based BRDF models to represent the reflective behavior and is used in several previous works [5, 7–9, 12, 15, 16, 42, 51, 52, 55]. A neural network predicts the (spatially varying) parameters of the reflection model, which are then combined with the viewing and light direction to compute the BRDF values using the analytical formula of the parametric model. See Fig. 2 for a visualization.

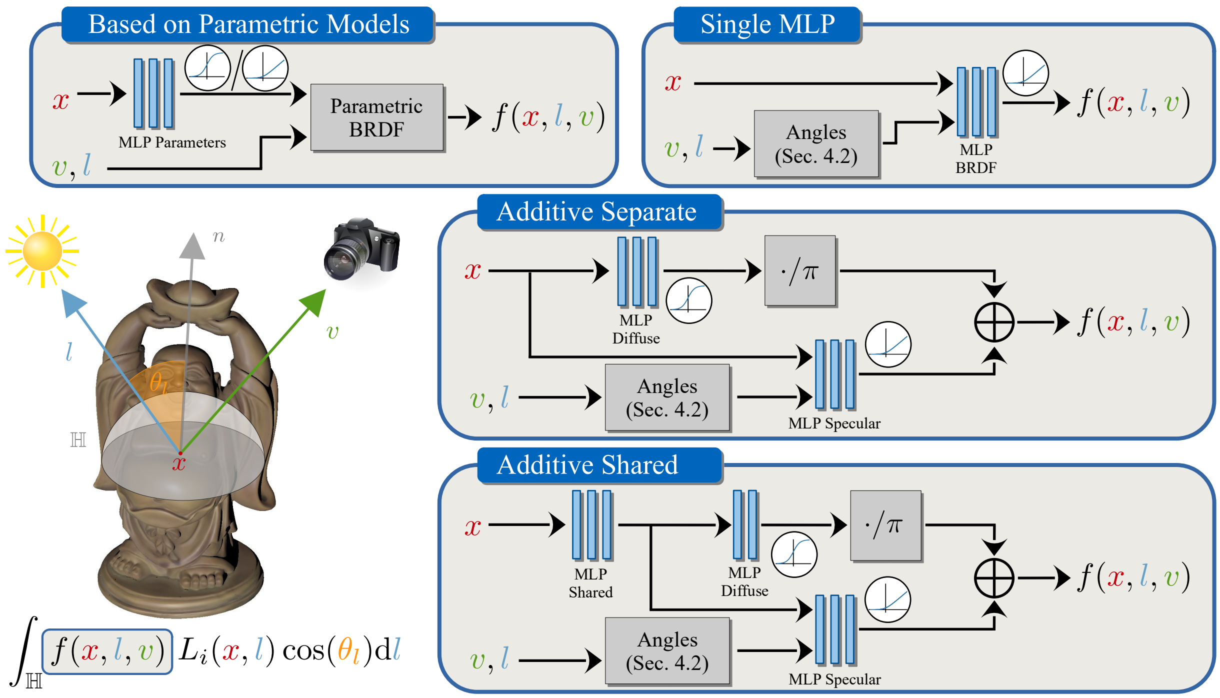

Figure 2. Overview of the neural models to compute the BRDF $f(x, l, v)$. Approaches based on a parametric model employ an MLP to predict the parameters for the respective model. A sigmoid or softplus output nonlinearity is used, depending on the range of the parameters. Methods based on a single MLP compute the BRDF value directly from the position and the view and light direction. Additive models split the reflection into a diffuse and a specular component. While the *separate* architecture uses independent MLPs for both, the *shared* architecture uses a common MLP with two separate heads. For all *purely* neural approaches, the Rusinkiewicz angles [38] are used to parametrize the directions. Moreover, we employ an intrinsic approach [22] to encode the position directly on the mesh as well as positional encoding for the angles. See Sec. 4.2 for details on the angles and the encoding.

In this work, we evaluate parametric neural BRDFs based on the Torrance-Sparrow (TS) model [46] and the isotropic variant of the Disney BRDF [10] as state-of-the-art parametric models. For reference, we also evaluate the energy-preserving variant of the Phong BRDF (RP) [25]. We refer to the appendix for more details on the models.

Finally, we compare against the recently proposed FM-BRDF model [19], which addresses shortcomings in the Oren-Nayar model. Since the normalization estimation for this model's normal distribution is computationally intractable for the spatially varying data, we estimate one single normalization per object. At the same time, all the other parameters depend on the position. Note that for the semi-synthetic MERL-based data, this makes the estimation less complex since the material is uniform over the mesh.

We use an MLP with 6 layers of width 128 and a skip connection to the third layer to predict the parameters for all parametric models. We use ReLU nonlinearities between the layers and a sigmoid or a softplus output nonlinearity, depending on the range of the respective parameter. For the Disney and the Microfacet model, we found that the convergence is more stable when the output for the roughness is scaled by 0.5 before the last sigmoid function.

4.1.2 Purely Neural BRDFs

In contrast to the models in the previous section, purely neural models do not rely on a parametric model but directly predict the BRDF value using neural networks.

Single MLP Several previous works employ a single MLP to predict the BRDF value from the position x and the light and view direction l and v directly [13, 43, 50]. The adapted architecture for our experiments is shown in Fig. 2. We map the directions to the Rusinkiewicz angles as described in Sec. 4.2 and employ a softplus output nonlinearity to ensure the positivity of the resulting BRDF value. Again, we use a 6-layer MLP of width 128 with ReLU activations and an input skip to layer 3.

Additive Split Other works predict diffuse and specular reflections, which are additively combined into the final BRDF value. NeRFactor [54] uses two separate MLPs, which we adapt as *Additive Separate*, see Fig. 2. We remove their albedo clamping and use a 3D RGB specular part instead of the proposed scalar term since both yielded worse results. See the appendix for details and the experimental evaluation. We use a per-object specular part instead of the pre-trained module to facilitate a fair comparison between all methods without the need for pre-training. For both MLPs, we use 4 layers of width 128 and a skip connection to the second layer, following the original work. For the diffuse MLP, we use a sigmoid output and divide by π to obtain the Lambertian diffuse term. For the specular MLP, a softplus output ensures positivity.

Instead of two separate MLPs, LitNeRF [39] employs a shared spatial MLP that extracts common features that are then used by a diffuse and specular head. We adopt this architecture as *Additive Shared*, see Fig. 2. We use 5 layers for the shared ReLU network of width 128 with a skip connection to the third layer, a single layer of width 128 for the diffuse MLP and two layers of width 128 for the specular output. Again, a sigmoid and a softplus output are used for the diffuse and the specular head, respectively. Both additive architectures use the angle parametrization introduced in the next section.

4.2. Angle Parametrization and Input Encodings

Following previous work [43, 54], we map the view and the light direction to the Rusinkiewicz angles [38] before feeding them into the network. The most important advantage of this reparametrization is that the specular peaks align with the coordinate axes. For an isotropic BRDF, the Rusinkiewicz angles read $(\theta_h, \theta_d, \phi_d) \in [0, \frac{\pi}{2}]^2 \times [0, 2\pi]$. Please see the appendix for a more detailed discussion of the angle parameterizations and an experimental comparison to the common view-light angles.

Previous work has shown that vanilla MLPs have difficulties representing high-frequency data. Therefore, applying some encoding function to the input is common practice. Since we parametrize the BRDF directly on the mesh, we encode the spatial position with the intrinsic encoding for neural fields on manifolds proposed by Koestler *et al.* [22], which has been shown to be advantageous compared to common extrinsic encodings. We use positional encoding [31] for the Rusinkiewicz angles. For more details on the encodings, we refer to the appendix.

4.3. A Novel Mapping to Ensure Reciprocity

While the neural BRDFs based on parametric models fulfill the reciprocity constraint in Eq. (3) by construction, LitNeRF [39] is the only purely neural approach that aims at fulfilling this constraint. During training, they randomly swap view and light direction to force the model to treat them interchangeably. However, this is a soft constraint, and it is not guaranteed that the reciprocity is fulfilled. In contrast, we propose a mapping of the Rusinkiewicz angles that ensures that Eq. (3) is fulfilled exactly by construction.

For an isotropic BRDF and the Rusinkiewicz angles, the reciprocity condition reads

$$f(x, \theta_h, \theta_d, \phi_d) = f(x, \theta_h, \theta_d, \phi_d + \pi). \tag{5}$$

We exploit this and map the Rusinkiewicz angles to

$$\Theta = [\theta_h, \theta_d, \phi_{d,\pi}, \phi_{d,\pi} + \pi], \tag{6}$$

where $\phi_{d,\pi}$ is short notation for ϕ_d modulo π and $\text{range}(\Theta) = [0, \frac{\pi}{2}]^2 \times [0, \pi] \times [\pi, 2\pi]$. We now have

$$\Theta(\theta_h, \theta_d, \phi_d) = \Theta(\theta_h, \theta_d, \phi_d + \pi). \tag{7}$$

Hence, if we use Θ as input to the downstream blocks, the reciprocity constraint in Eq. (3) is fulfilled by construction.

4.4. Enhancing the Additive Split

The existing purely neural models based on an additive split do not consider that light reflected at the surface cannot be used for subsurface scattering. We propose a novel extension to additive split approaches that can model this phenomenon. Besides the specular reflection $f_s(x, l, v)$, we use the specular MLP to predict a weight $\xi(x, l, v) \in [0, 1]^3$, which we use to reduce the diffuse part $f_d(x)$ of the BRDF channel-wise. The resulting additive split reads

$$f(x, l, v) = (1 - \xi(x, l, v)) \circ f_d(x) + f_s(x, l, v). \tag{8}$$

Since the weight ξ makes the diffuse summand view-dependent, which increases the ambiguity, we add two regularizers. We use an L1 loss between the diffuse part and the reference image to encourage the model to represent as much as possible with the diffuse part, and moreover, we use an L1 regularizer on the specular part to encourage sparsity. Please see the appendix for more details.

5. Methodology and Datasets

We estimate the BRDFs from a sparse set of HDR images from multiple viewpoints, each taken under changing directional lighting conditions. We assume mesh, camera poses, and light calibration to be known. This controlled setting ensures that the strengths of the reflection models can be assessed without confounding effects from estimating other quantities as well.

5.1. Rendering

For a single directional light, the irradiance L_i is independent of the position x, and the light direction l is constant for each view. In that case, the integral rendering equation in Eq. (1), reduces to a single evaluation and the rendering $L_o(x, v)$ for the pixel corresponding to x and v now reads

$$L_o(x, v) = f(x, l, v) L_i \mathbb{I}_s(x, l) \cos \theta_l. \tag{9}$$

The indicator function $\mathbb{I}_s(x, l)$ is used as a masking term to account for cast-shadows. Its value is computed by casting a secondary ray from the first intersection point x on the mesh in direction l and checking if the ray intersects the object. We apply a shadow bias to avoid self-shadow aliasing [1]. The intersection points x are computed by standard ray mesh intersection, and we compute the normal on the mesh by barycentric interpolation of the vertex normals.

5.2. Loss and Training

Since we are working with HDR images, which have a much larger dynamic range than sRGB images, a standard L_2 loss would be dominated by the bright regions, suppressing information from darker regions. To avoid this, we follow the analysis by Mildenhall *et al*. [30] and use a gamma correction function $\gamma : [0, 1] \mapsto [0, 1]$ to construct a tone-mapped MSE-based loss term

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N}(\gamma(L_o(x, v)) - \gamma(L_{GT}(x, v))^2, \qquad (10)$$

where $L_{GT}(x, v)$ is the ground truth color of the pixel corresponding to x and v in linear color space.

We train the models using the Adam optimizer [21] with a batch size of $N_b = 2^{15}$ color values. For more details on the loss and the training, see the appendix.

5.3. Datasets

We use the DiLiGenT-MV real-world multi-view dataset with calibrated lighting [28]. It consists of HDR images of 5 objects with spatially varying, complex reflectance behaviors, taken from 20 views, each captured under 96 calibrated directional lights. We use the included ground truth meshes.

Additionally, we create a semi-synthetic dataset based on real BRDF measurements to evaluate and analyze the performance in a more controlled setting. We render HDR images with 25 different *uniform* MERL BRDFs [29] on 9 common 3D test meshes [20]. We add Gaussian noise with $\sigma = 10^{-3}$ to the images. Besides the rendering, we also store the raw reflectance values to enable the direct evaluation of the BRDF prediction. We train all models on 10 views with 30 lights each and use 12 lights for each of the 10 remaining views as the unseen test set for the evaluation.

6. Experiments

In this section, we present a thorough evaluation of the neural BRDF approaches. First, we conduct a qualitative and quantitative comparison on the semi-synthetic and real-world datasets described in Sec. 5.3. Subsequently, we analyze several aspects of the models, including energy conservation and the approaches to reciprocity, in more depth. We refer to the appendix for more information on the models and the datasets, as well as several additional experiments.

Evaluation Metrics We transform the renderings of the test set from linear to sRGB space and evaluate PSNR, LPIPS [53] and DSSIM, where the latter is based on SSIM [49] by DSSIM $= (1 - \text{SSIM})/2$. Moreover, we report the HDR version of the ꟻLIP metric [2, 3]. For each metric, we average the results of the respective dataset.

Following [27], we report the root mean squared error between the cubic root of the predicted and the GT BRDF values for the synthetic experiments (RMSE $\sqrt[3]{}$). As analyzed in [27], we discard values close to the horizon from the data. Moreover, we exclude values from saturated pixels. This BRDF-space metric correlates well with perceptual similarity. We refer to the supplement for more details.

6.1. Comparison of the BRDF Models

We show a qualitative comparison of the models in Fig. 3 and report the metrics in Tab. 1. For the semi-synthetic examples based on the MERL data [29], the results show a significant advantage of the purely neural approaches, outperforming the methods based on parametric models by a large margin on all metrics. The examples in the top rows of Fig. 3 and Fig. 1 confirm that only the purely neural approaches can faithfully capture the complex reflection patterns of highly specular materials. For the real-world data, the difference between the models is much less significant. One reason is the absence of highly specular materials in the DiLiGenT-MV dataset. Another potential reason are interreflections in the real-world data, which we do not model. Finally, there might be noise in the real-world data not captured by the simple noise model used to create the semi-synthetic data. Indeed, we see a higher influence of noise on the purely neural models: If we remove the noise from the semi-synthetic data, the performance gains for those models are much higher than for the parametric models.

Among the parameter-based models, the Disney BRDF [10] performs best, which is expected since it is the most sophisticated model. For the FMBRDF [19], we observe a "glow" for the cow from the real-world data (see Fig. 3). We hypothesize that this originates from the scalar specular term since a similar effect is observed for the scalar additive architecture, as shown in the appendix. This effect does not appear for other objects.

For the purely neural approaches, we see a different behavior between the two datasets, with the ranking order being reversed. The reason seems to be the number of layers that the view and the light direction are fed through. While it is only 2 layers for the additive shared architecture, there are 4 for additive separate and 6 for the single MLP architectures. We investigate this further in the next section.

Our enhancement for the additive split (*cf*. Sec. 4.4) shows slight improvements for both additive approaches. This confirms that the additional freedom can help the model to adapt better to the underlying BRDF functions.

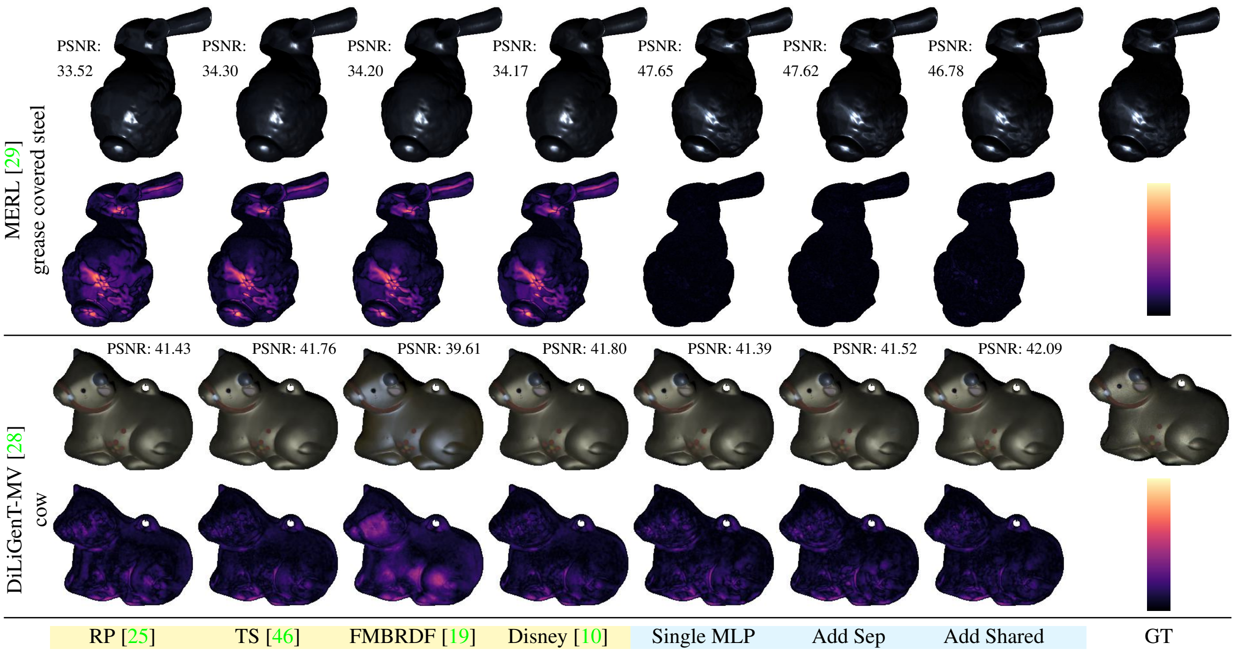

Figure 3. Qualitative evaluation of the reconstruction for the MERL grease covered steel BRDF [29] uniformly rendered on the bunny mesh from [20] and the cow object from the real-world DiLiGenT-MV data [28]. Shown are renderings in sRGB space with the corresponding PSNR values and the ꟻLIP error maps for the sRGB renderings. Only the purely neural approaches ( ) are able to reconstruct the intricate reflection patterns for the grease covered steal faithfully. For the cow, the results for the parametric models ( ) are much more on par, and all models show difficulties in similar areas – in particular in recesses, which suggests unmodeled interreflections as a potential reason.

	MERL [29]					DiLiGenT-MV [28]			
	RMSE $\sqrt[3]{}$ ↓	PSNR ↑	DSSIM ↓	LPIPS ↓	ꟻLIP ↓	PSNR ↑	DSSIM ↓	LPIPS ↓	ꟻLIP ↓
Realistic Phong [25]	1.632	44.05	0.779	1.656	6.025	41.59	0.775	1.771	3.119
Torrance-Sparrow [46]	1.216	46.06	0.620	1.174	5.717	41.83	0.757	1.763	3.054
Fresnel Microfacet BRDF [19]	1.167	46.40	0.654	1.310	5.876	41.42	0.752	1.781	3.228
Disney [10]	1.218	46.46	0.616	1.158	5.735	41.89	0.750	1.788	3.056
Single MLP	0.554	52.35	**0.411**	**0.565**	4.795	41.64	0.789	1.765	2.962
Additive Separate	0.562	52.28	0.413	0.571	4.802	41.77	0.758	1.702	2.924
Additive Shared	0.665	51.40	0.428	0.621	5.075	42.35	0.714	1.671	2.856
Additive Separate (enhanced)	**0.547**	**52.39**	0.412	0.565	**4.794**	41.88	0.744	1.663	2.883
Additive Shared (enhanced)	0.579	51.95	0.419	0.580	4.836	**42.38**	**0.709**	**1.659**	**2.827**

Table 1. Quantitative comparison of the BRDF models on the MERL-based semi-synthetic dataset [29] and the real-world data [28]. RMSE $\sqrt[3]{}$ is the RMSE of the cubic root of the BRDF values. PSNR, DSSIM, LPIPS are computed for the sRGB renderings. All quantities are first averaged over one object and then averaged over all objects in the respective dataset. RMSE $\sqrt[3]{}$, DSSIM, LPIPS and ꟻLIP are scaled by 100. While the purely neural approaches ( ) show superior results over the parametric models ( ) on the semi-synthetic data, the difference is much smaller on the real-world data. We attribute this to more materials with complex reflections in the MERL BRDFs and potentially unmodelled noise patterns and interreflections in the DiLiGenT-MV dataset.

6.2. Analysis of the BRDF Models

Number of Layers for View/Light Directions To confirm that, indeed, the difference in the performance of the purely neural models for the real-world and semi-synthetic data originates from the number of layers (NOL) for the directions, we perform experiments where we vary this parameter for each of the purely neural models.

The results are shown in Tab. 2, where we indicate the change in the NOL for the directions per model. Please see the appendix for the detailed architecture changes. The numbers confirm the trend that reducing the NOL decreases the reconstruction quality for the MERL examples while increasing the quality for the DiLiGenT-MV data while increasing the NOL has the opposite effect. This may suggest that more layers for the directions are beneficial for complex reflective patterns as contained in the MERL data, but at the same time, make the model less robust to potentially noisy measurements of real-world data.

	MERL [29]					DiLiGenT-MV [28]			
	ΔRMSE $\sqrt[3]{}$	ΔPSNR	ΔDSSIM	ΔLPIPS	ΔꟻLIP	ΔPSNR	ΔDSSIM	ΔLPIPS	ΔꟻLIP
Single MLP* (NOL dirs 6 → 3)	+0.01	-0.15	+0.01	+0.01	+0.02	+0.51	-0.06	-0.10	-0.09
Additive Sep.* (NOL dirs 4 → 2)	+0.02	-0.09	+0.01	+0.01	+0.02	+0.26	-0.01	+0.02	+0.01
Additive Shared* (NOL dirs 2 → 4)	-0.09	+0.71	-0.01	-0.05	-0.25	-0.37	+0.01	+0.01	+0.02

Table 2. Quantitative changes for varying the number of layers (NOL) for the directions. Shown is the difference to the results in Tab. 1; RMSE $\sqrt[3]{}$, DSSIM, LPIPS and ꟻLIP are scaled by 100. The experiments confirm that reducing the NOL for the directions tends to increase the reconstruction quality for the real-world data while simultaneously decreasing it slightly for the semi-synthetic data. Increasing the NOL has the opposite effect. We hypothesize, that more layers for the directions enable the model to learn the more complex reflection patterns in the MERL data, while simultaneously making the model less robust to noise that might be contained in the real-world data.

	MERL [29]	DiLiGenT-MV [28]
Single MLP	$2.93 \cdot 10^{-2}$	2.52
Add Sep	$2.62 \cdot 10^{-2}$	1.94
Add Shared	$4.70 \cdot 10^{-3}$	$6.78 \cdot 10^{-1}$
Single MLP (rnd. in-swap)	$5.42 \cdot 10^{-4}$	$2.21 \cdot 10^{-2}$
Add Sep (rnd. in-swap)	$4.87 \cdot 10^{-4}$	$1.61 \cdot 10^{-2}$
Add Shared (rnd. in-swap)	$1.60 \cdot 10^{-4}$	$1.11 \cdot 10^{-3}$

Table 3. RMSE of the violation of the reciprocity constraint Eq. (3). The error is significant without any reciprocity strategy, in particular for the real-world data. While the random input swap strategy in [39] reduces the error, the constraint is still violated. In contrast, our mapping ensures reciprocity exactly by construction (and therefore is not shown here). The RMSE is *not* scaled here.

Reciprocity Mapping To evaluate the random input swap training strategy used in LitNerf [39] as well as our input mapping proposed in Sec. 4.3, we quantify how much the symmetry constraint is violated by computing the RMSE between the BRDF values with changed positions, i.e. for the pairs $f(x, l, v)$ and $f(x, v, l)$. We report the results for the random input swap and the models without any reciprocity strategy in Tab. 3. Note that our approach fulfills the constraint exactly by construction and is therefore not included in the table. We see, that without any reciprocity strategy, the constraint is violated significantly, particularly for the real-world data. While the random input swap does reduce this error by several orders of magnitude, the constraint is still violated, which might cause problems for particular algorithms. In contrast, Eq. (3) is fulfilled by construction by our approach. The influence on the reconstruction quality is analyzed in the appendix. Apart from the single MLP architecture, for which the random input swap performs slightly better on the real-world data, both approaches have a similar effect, and we obtain only slightly worse results as a trade-off for the ensured reciprocity.

Energy Conservation To analyze how well the models fulfill the energy conservation in Eq. (4), we analyze $50k$ randomly sampled point-light pairs for each object from the test data. We approximate the energy integral in Eq. (4) by Monte-Carlo (MC) integration, where we use cosine-weighted hemisphere sampling and draw $20k$ view direction samples for each integral to ensure convergence. A more detailed discussion can be found in the appendix.

	MERL [29]		DiLiGenT-MV [28]	
	$\% > 1$	$\mu_{.5} > 1$	$\% > 1$	$\mu_{.5} > 1$
RP [25]	**none**	**all valid**	**none**	**all valid**
TS [46]	0.02%	1.1	0.91%	1.1
FMBRDF [19]	0.01%	1.1	1.29%	1.1
Disney [10]	0.05%	1.1	1.64%	1.1
Single MLP	0.05%	1.5	5.17%	3.7
Add Sep	0.02%	1.2	8.01%	2.4
Add Shared	0.01%	1.2	4.43%	1.5

Table 4. Analysis of energy conservation (Eq. (4)) on the MERL-based synthetic dataset [29] and the real-world data from DiLiGenT-MV [28]. Shown are the percentage of the 50k sampled point-light pairs violating the energy conservation ($\% > 1$) and the median of the energies larger than 1 ($\mu_{.5} > 1$), each averaged over all experiments of the respective dataset.

The results in Tab. 4 show that for the semi-synthetic data, all approaches fulfill energy conservation almost everywhere. For the real-world data, we observe more significant violations; in particular for the purely neural methods, which might again indicate measurement noise. The fully separated additive approach seems to be particularly disadvantageous for energy conservation.

7. Conclusion

In this work, we have presented an exhaustive comparison of different approaches to neural BRDF modeling. The results show that while purely neural approaches have advantages for materials with complex reflective patterns, the performance of methods based on parametric models on real-world data with less complex reflection patterns is comparable. We found signs that purely neural methods cannot learn reciprocity and energy conservation from data, particularly for real-world images, and we have analyzed approaches to ensure reciprocity, including a newly proposed input mapping that ensures reciprocity by construction. Finally, we have presented an extension to models based on additive splits that aims to capture the physical behavior better and shows improvements for existing approaches.

Acknowledgements This work was supported by the ERC Advanced Grant SIMULACRON.

References

[1] Tomas Akenine-Möller, Eric Haines, Naty Hoffman, Angelo Pesce, Michał Iwanicki, and Sébastien Hillaire. *Real-Time Rendering 4th Edition*. A K Peters/CRC Press, 2018.

[2] Pontus Andersson, Jim Nilsson, Tomas Akenine-Möller, Magnus Oskarsson, Kalle Åström, and Mark D. Fairchild. ꟻLIP: A Difference Evaluator for Alternating Images. *ACM Computer Graphics and Interactive Techniques*, 2020.

[3] Pontus Andersson, Jim Nilsson, Peter Shirley, and Tomas Akenine-Möller. Visualizing Errors in Rendered High Dynamic Range Images. In *Eurographics Short Papers*, 2021.

[4] Michael Ashikhmin and Peter Shirley. An anisotropic phong BRDF model. *Journal of Graphics Tools*, 2000.

[5] Sai Bi, Zexiang Xu, Pratul Srinivasan, Ben Mildenhall, Kalyan Sunkavalli, Miloš Hašan, Yannick Hold-Geoffroy, David Kriegman, and Ravi Ramamoorthi. Neural reflectance fields for appearance acquisition. *arXiv*, 2020.

[6] James F. Blinn. Models of light reflection for computer synthesized pictures. In *International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 1977.

[7] Mark Boss, Raphael Braun, Varun Jampani, Jonathan T. Barron, Ce Liu, and Hendrik P. A. Lensch. Nerd: Neural reflectance decomposition from image collections. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

[8] Mohammed Brahimi, Bjoern Haefner, Zhenzhang Ye, Bastian Goldluecke, and Daniel Cremers. Sparse views, near light: A practical paradigm for uncalibrated point-light photometric stereo. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[9] Mohammed Brahimi, Bjoern Haefner, Tarun Yenamandra, Bastian Goldluecke, and Daniel Cremers. Supervol: Super-resolution shape and reflectance estimation in inverse volume rendering. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2024.

[10] Brent Burley and Walt Disney Animation Studios. Physically-based shading at disney. In *International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 2012.

[11] Robert L. Cook and Kenneth E. Torrance. A reflectance model for computer graphics. *ACM Transactions on Graphics (TOG)*, 1982.

[12] Valentin Deschaintre, Miika Aittala, Frédo Durand, George Drettakis, and Adrien Bousseau. Single-image SVBRDF capture with a rendering-aware deep network. *ACM Transactions on Graphics (TOG)*, 2018.

[13] Jiahui Fan, Beibei Wang, Milos Hasan, Jian Yang, and Ling-Qi Yan. Neural layered BRDFs. In *International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 2022.

[14] Dar'ya Guarnera, Giuseppe Claudio Guarnera, Abhijeet Ghosh, Cornelia Denk, and Mashhuda Glencross. BRDF representation and acquisition. *Computer Graphics Forum*, 2016.

[15] Yu Guo, Cameron Smith, Milos Hasan, Kalyan Sunkavalli, and Shuang Zhao. MaterialGAN: reflectance capture using a generative SVBRDF model. *ACM Transactions on Graphics (TOG)*, 2020.

[16] Philipp Henzler, Valentin Deschaintre, Niloy J. Mitra, and Tobias Ritschel. Generative modelling of BRDF textures from flash images. *ACM Transactions on Graphics (TOG)*, 2021.

[17] Nicolas Holzschuch and Romain Pacanowski. A two-scale microfacet reflectance model combining reflection and diffraction. *ACM Transactions on Graphics (TOG)*, 2017.

[18] Bingyang Hu, Jie Guo, Yanjun Chen, Mengtian Li, and Yanwen Guo. DeepBRDF: A deep representation for manipulating measured BRDF. In *Computer Graphics Forum*, 2020.

[19] Tomoki Ichikawa, Yoshiki Fukao, Shohei Nobuhara, and Ko Nishino. Fresnel microfacet BRDF: Unification of polari-radiometric surface-body reflection. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[20] Alec Jacobson. Common 3D test models. https://github.com/alecjacobson/common-3d-test-models, 2020.

[21] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.

[22] Lukas Koestler, Daniel Grittner, Michael Möller, Daniel Cremers, and Zorah Lähner. Intrinsic neural fields: Learning functions on manifolds. In *European Conference on Computer Vision (ECCV)*, 2022.

[23] Alexandr Kuznetsov, Krishna Mullia, Zexiang Xu, Milos Hasan, and Ravi Ramamoorthi. NeuMIP: multi-resolution neural materials. *ACM Transactions on Graphics (TOG)*, 2021.

[24] Alexandr Kuznetsov, Xuezheng Wang, Krishna Mullia, Fujun Luan, Zexiang Xu, Milos Hasan, and Ravi Ramamoorthi. Rendering neural materials on curved surfaces. In *International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 2022.

[25] Eric P Lafortune and Yves D Willems. Using the modified phong reflectance model for physically based rendering. *Katholieke Universiteit Leuven. Departement Computerwetenschappen*, 1994.

[26] Johann Heinrich Lambert. *Photometria Sive De Mensura Et Gradibus Luminis, Colorum Et Umbrae*. Klett, 1760.

[27] Guillaume Lavoué, Nicolas Bonneel, Jean-Philippe Farrugia, and Cyril Soler. Perceptual quality of BRDF approximations: dataset and metrics. *Computer Graphics Forum*, 2021.

[28] Min Li, Zhenglong Zhou, Zhe Wu, Boxin Shi, Changyu Diao, and Ping Tan. Multi-view photometric stereo: A robust solution and benchmark dataset for spatially varying isotropic materials. *IEEE Transactions on Image Processing (TIP)*, 2020.

[29] Wojciech Matusik, Hanspeter Pfister, Matt Brand, and Leonard McMillan. A data-driven reflectance model. *ACM Transactions on Graphics (TOG)*, 2003.

[30] Ben Mildenhall, Peter Hedman, Ricardo Martin-Brualla, Pratul P. Srinivasan, and Jonathan T. Barron. NeRF in the dark: High dynamic range view synthesis from noisy raw images. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[31] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *European Conference on Computer Vision (ECCV)*, 2020.

[32] Rosana Montes and Carlos Ureña. An overview of BRDF models. *University of Grenada, Technical Report*, 2012.

[33] Michael Oren and Shree K. Nayar. Generalization of lambert's reflectance model. In *International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 1994.

[34] Bui Tuong Phong. Illumination for computer generated pictures. *Communications of the ACM*, 1975.

[35] Gilles Rainer, Abhijeet Ghosh, Wenzel Jakob, and Tim Weyrich. Unified neural encoding of BTFs. In *Computer Graphics Forum*, 2020.

[36] Gilles Rainer, Wenzel Jakob, Abhijeet Ghosh, and Tim Weyrich. Neural BTF compression and interpolation. In *Computer Graphics Forum*, 2019.

[37] Carlos Rodriguez-Pardo, Konstantinos Kazatzis, Jorge Lopez-Moreno, and Elena Garces. NeuBTF: Neural fields for BTF encoding and transfer. *Computers & Graphics*, 2023.

[38] Szymon M Rusinkiewicz. A new change of variables for efficient BRDF representation. In *Eurographics Symposium on Rendering (EGSR)*, 1998.

[39] Kripasindhu Sarkar, Marcel C. Bühler, Gengyan Li, Daoye Wang, Delio Vicini, Jérémy Riviere, Yinda Zhang, Sergio Orts-Escolano, Paulo F. U. Gotardo, Thabo Beeler, and Abhimitra Meka. LitNeRF: Intrinsic radiance decomposition for high-quality view synthesis and relighting of faces. In *SIGGRAPH Asia*, 2023.

[40] Christophe Schlick. An inexpensive BRDF model for physically-based rendering. In *Computer Graphics Forum*, 1994.

[41] Bruce Smith. Geometrical shadowing of a random rough surface. *IEEE Transactions on Antennas and Propagation*, 1967.

[42] Pratul P Srinivasan, Boyang Deng, Xiuming Zhang, Matthew Tancik, Ben Mildenhall, and Jonathan T Barron. NeRV: Neural reflectance and visibility fields for relighting and view synthesis. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[43] Alejandro Sztrajman, Gilles Rainer, Tobias Ritschel, and Tim Weyrich. Neural BRDF representation and importance sampling. In *Computer Graphics Forum*, 2021.

[44] Thomson TG, Jeppe Revall Frisvad, Ravi Ramamoorthi, and Henrik Wann Jensen. Neural BSSRDF: Object appearance representation including heterogeneous subsurface scattering. *arXiv*, 2023.

[45] Thomson TG, Duc Minh Tran, Henrik W Jensen, Ravi Ramamoorthi, and Jeppe Revall Frisvad. Neural SSS: Lightweight object appearance representation. In *Computer Graphics Forum*, 2024.

[46] Kenneth E. Torrance and Ephraim M. Sparrow. Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America (JOSA)*, 1967.

[47] TS Trowbridge and Karl P Reitz. Average irregularity representation of a rough surface for ray reflection. *Journal of the Optical Society of America (JOSA)*, 1975.

[48] Bruce Walter, Stephen R Marschner, Hongsong Li, and Kenneth E Torrance. Microfacet models for refraction through rough surfaces. In *Eurographics Symposium on Rendering (EGSR)*, 2007.

[49] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing (TIP)*, 2004.

[50] Chong Zeng, Guojun Chen, Yue Dong, Pieter Peers, Hongzhi Wu, and Xin Tong. Relighting neural radiance fields with shadow and highlight hints. In *International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 2023.

[51] Kai Zhang, Fujun Luan, Zhengqi Li, and Noah Snavely. IRON: inverse rendering by optimizing neural sdfs and materials from photometric images. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[52] Kai Zhang, Fujun Luan, Qianqian Wang, Kavita Bala, and Noah Snavely. PhySG: Inverse rendering with spherical gaussians for physics-based material editing and relighting. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[53] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[54] Xiuming Zhang, Pratul P. Srinivasan, Boyang Deng, Paul E. Debevec, William T. Freeman, and Jonathan T. Barron. NeRFactor: neural factorization of shape and reflectance under an unknown illumination. *ACM Transactions on Graphics (TOG)*, 2021.

[55] Yuanqing Zhang, Jiaming Sun, Xingyi He, Huan Fu, Rongfei Jia, and Xiaowei Zhou. Modeling indirect illumination for inverse rendering. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.