BSDF Importance Baking: A Lightweight Neural Solution to Importance Sampling Parametric BSDFs

YAOYI BAI, University of California, Santa Barbara, USA
SONGYIN WU, University of California, Santa Barbara, USA
ZHENG ZENG, University of California, Santa Barbara, USA
BEIBEI WANG, Nankai University, China
LING-QI YAN, University of California, Santa Barbara, USA

Parametric BSDFs (Bidirectional Scattering Distribution Functions) are pervasively used because of their flexibility to represent a large variety of material appearances by simply tuning the parameters. While efficient evaluation of parametric BSDFs has been well-studied, high-quality importance sampling techniques for parametric BSDFs are still scarce. Existing sampling strategies either heavily rely on approximations and result in high variance, or solely perform sampling on a portion of the whole BSDF slice. Moreover, many of the sampling approaches are specifically paired with certain types of BSDFs. In this paper, we seek an efficient and general way for importance sampling parametric BSDFs. We notice that the nature of importance sampling is the mapping between a uniform distribution and the target distribution. Specifically, when BSDF parameters are given, the mapping that performs importance sampling on a BSDF slice can be simply recorded as a 2D image that we name as importance map. Following this observation, we accurately precompute the importance maps using a mathematical tool named optimal transport. Then we propose a lightweight neural network to efficiently compress the precomputed importance maps. In this way, we have completely brought parametric BSDF importance sampling to the precomputation stage, avoiding heavy runtime computation. Since this process is similar to light baking where a set of images are precomputed, we name our method importance baking. Together with a BSDF evaluation network and a PDF (probability density function) query network, our method enables full multiple importance sampling (MIS) without any revision to the rendering pipeline. Our method essentially performs perfect importance sampling. Compared with previous methods, we demonstrate reduced noise levels on rendering results with a rich set of appearances, including both conductors and dielectrics with anisotropic roughness.

CCS Concepts:
- Computing methodologies → Rendering.

Additional Key Words and Phrases: physically based rendering, importance sampling, neural rendering

1 INTRODUCTION

BSDFs (Bidirectional Scattering Distribution Function) are key to realistic appearances. They are the combination of BRDFs (Bidirectional Reflectance Distribution Function) for reflection and BTDFs (Bidirectional Transmittance Distribution Function) for transmission. Among all BSDFs, parametric BSDFs are pervasively used because of their flexibility to represent a large variety of materials' optical properties using a few parameters.

In the modern Monte Carlo rendering framework, BSDFs need not only be evaluated, but also importance sampled. The core of BSDF importance sampling is to choose an outgoing direction according to a probability density function (PDF) that closely resembles the 2D BSDF slice jointly determined by BSDF parameters and an incident direction. However, parametric BSDFs barely have accurate analytical importance sampling strategies. Existing solutions tend to sample part of the BSDF slice. For instance, single scattering microfacet BSDFs sample the normal distribution functions (NDFs) [Beckmann and Spizzichino 1963; Walter et al. 2007] or visible normal distribution functions (VNDFs) [Heitz and d'Eon 2014]. Multiple-scattering microfacet BSDFs (e.g., Heitz et al. [Heitz et al. 2016]) perform random walk, resulting in high variance.
Recently, neural approaches have been proposed to represent measured BRDFs [Sztrajman et al. 2021; Zheng et al. 2021], which are mostly tied with compression of measured BSDFs and seldomly study the problem of accurate importance sampling. Recent works that use neural methods to solve importance sampling do exist, but few work well with the naturally high-dimensional parametric BSDFs with many parameters, especially those utilizing normalizing flow [Müller et al. 2019; Xie et al. 2019].

In this work, we focus on accurately solving the importance sampling problem for parametric BSDFs. Our insight is that, importance sampling, regardless of the specific methods (e.g., inverse sampling) to achieve it, is in essence a mapping between a uniform distribution and the target distribution. Specifically, importance sampling a parametric BSDF slice is a 2D to 2D mapping which can be simply recorded as a 2D image that we name as importance map. Therefore, if we can precompute and compress all the importance maps for all combinations of parameters and incident directions, we will be able to completely bring accurate parametric BSDF importance sampling to the precomputation stage, avoiding heavy runtime computation.

However, we found that the commonly used 2D importance sampling method—the marginalized inverse transform sampling, a.k.a. row-column sampling—produces problematic discontinuous importance maps that are not only difficult to compress but also prone to artifacts when interpolated. To solve this issue, we introduce the Optimal Transport (OT) theory to BSDF sampling for the first time, which is by nature suitable to provide a smooth mapping between distributions. Once we have precomputed the importance maps, we propose a lightweight neural network to store and compress them efficiently. In addition to this importance sampling network, we also provide a BSDF evaluation network and a pdf query network, leading to a complete neural solution for general parametric BSDFs, supporting full multiple importance sampling (MIS) without any revision to the rest of the rendering pipeline.

Since the process of computing and compressing importance maps is similar to light baking, in the sense that a set of images are precomputed and queried during runtime, we name our method BSDF importance baking. It essentially performs perfect importance sampling, and can be used for any general parametric BSDFs, even those lacking analytic importance sampling solutions. Compared with previous methods, we demonstrate reduced noise levels on rendering results with a rich set of appearances, including both conductors and dielectrics with anisotropic roughness.

2 RELATED WORK

Parametric BSDFs. Parametric BSDFs represent different materials with several parameters. Two major groups of parametric BSDFs include empirical models (e.g., [Ashikhmin and Shirley 2001; Phong 1975]) and physically-based models (e.g., microfacet models). The microfacet BRDF was introduced into computer graphics by Cook and Torrance [Cook and Torrance 1982] and was extended to handle refraction [Stam 2001; Walter et al. 2007]. The microfacet model includes a normal distribution function (NDF), a shadowing-masking function, and a Fresnel term. Several common NDFs include Beckmann [Beckmann and Spizzichino 1963], GGX [Walter et al. 2007] and Generalized-Trowbridge-Reitz (GTR) [Burley and Studios 2012].

The shadowing-masking function [Heitz 2014] shows the probability that the light ray could reach and leave the microsurface, including Smith-based models [Smith 1967] and V-groove models [Cook and Torrance 1982]. Traditional microfacet models only consider the single scattering, leading to energy loss, especially for rough surfaces. Under the Smith assumption, Heitz et al. [2016] extended the previous microfacet BSDF model to multiple-bounce and compensate for energy loss. Later, two methods [Lee et al. 2018; Xie and Hanrahan 2018] propose analytic solutions for multiple bounces under the less accurate V-groove assumption. In this paper, we demonstrate our idea on Smith-based microfacet models, although our method is not restricted to any specific kind of parametric BSDF.

BSDF Importance Sampling. Under the Monte Carlo rendering framework, the BSDF sampling strategy is crucial to variance reduction. However, sampling all the components within the BSDF is non-trivial. In the microfacet model, sampling the NDF gives a good approximation but can produce significant variance when the incident direction is from a grazing angle. Heitz et al. [2017; 2018; 2014] reduced sampling variance by sampling the distribution of visible normals (VNDF). However, VNDF sampling only works for NDF that is stretch-invariant. Thus, some NDFs (e.g., GTR) do not allow for accurate VNDF sampling and only use an approximation. Even worse, when considering multiple scattering in the microfacet model [Heitz et al. 2016], the importance sampling becomes a random walk due to the absence of the closed-form formulation, leading to low performance.

Unlike these methods, our method considers the entire BSDF and allows for perfect importance sampling. Our model does not require that the BSDF has a closed-form formulation, thus can be used for arbitrary parametric BSDF models, including multiple scattering models.

Neural network for BRDF representation. Neural networks (NN) have been recently used for BRDF representation on measured materials by representing one spatially-varying BRDF (SVBRDF) per network [Kuznetsov et al. 2021; Rainer et al. 2019], all BRDFs within a unified network [Hu et al. 2020; Rainer et al. 2020; Zheng et al. 2021], or each BRDF as a standalone decoder network [Sztrajman et al. 2021]. Besides compressing the measured BRDFs, Sztrajman et al. [Sztrajman et al. 2021] also support importance sampling by mapping a measured BRDF to an approximate parametric BRDF, but the differences between them result in imperfect importance sampling. Xie et al. [2019] proposed to use a RealNVP network to learn multiple scattering equivalent NDFs in the slope space, enabling importance sampling. However, their method cannot support the high-dimensional parameter space for parametric BRDFs (e.g., spatially-varying Fresnel and anisotropic roughness) because of the notoriously bulky structure of normalizing flow structures and is considered too slow [Müller et al. 2019] for practical use.

Our method focuses on importance sampling general parametric BSDFs (including BTDFs). By completely decoupling the computation of the importance maps to the precomputation stage, we only require using a lightweight neural network to perform compression. Furthermore, since our importance maps are generally smooth thanks to optimal transport, they are naturally suitable for a neural network to compress.
BSDF Importance Baking: A Lightweight Neural Solution to Importance Sampling Parametric BSDFs

3 BSDF IMPORTANCE BAKING

3.1 Background

Parametric BSDFs. As mentioned in Sec. 2, we focus on microfacet BSDFs, typical and widely-used parametric BSDFs. For single-bounce microfacet BSDFs where the light bouncing off the microfacets will either be occluded or exit directly, they can be analytically written as:

\[ f_s(\omega_i, \omega_o) = \frac{F(\omega_i)G(\omega_i, \omega_o)D(\omega_o)}{4|\omega_i \cdot n||\omega_o \cdot n|}, \]

where \( F \) is the Fresnel term deciding the overall amount of light reflected, often approximated with a base color \( R_0 \) and an incident direction \( \omega_i \). \( G \) is the shadowing-masking term accounting for self-obscuration, and \( D \) is the normal distribution function (NDF), queried along the half vector \( \omega_{ho} = \frac{\omega_i + \omega_o}{|\omega_i + \omega_o|} \).

For multiple-bounce microfacet BSDFs where the light may bounce multiple times before they exit the surface, there is still no analytical expression. Moreover, heavy random walk simulations will be needed. In this work, we consider both single- and multiple-bounce BSDFs. Without loss of generality, we focus on GGX NDFs and describe conductors in the main text. We provide more information for dielectrics in the supplementary document.

Monte Carlo (MC) integration and importance sampling. The MC method provides an unbiased estimator to any definite integral in the domain \( \Omega \) with an arbitrary integrand \( f(x) \):

\[ F = \int_{\Omega} f(x) \, dx \approx \frac{1}{N} \sum_{k=1}^{N} f(X_k) \frac{p(X_k)}{p(x)} X_k \sim p(x). \]

The MC estimator approximates the integral \( F \) by drawing samples according to a probability density function (PDF) \( p(x) \). The \( f(X_k)/p(X_k) \) is known as the sampling weight. To minimize the estimation variance, importance sampling is desired – the closer between the “shapes” of \( f(x) \) and \( p(x) \), the lower variance does the MC estimator have.

BSDF importance sampling. MC integration is the core of the modern rendering pipeline, solving the rendering equation at each shading point:

\[ L(\omega_o) = \int L(\omega_i) f_s(\omega_i, \omega_o)(\omega_o, n) \, d\omega_i. \]

Given the BSDF parameters \( \epsilon = \{R_0, \alpha_s, \alpha_g\} \) (anisotropic roughness) and the viewing direction \( \omega_o \), the BSDF becomes a 2D slice, and BSDF importance sampling seeks on a good PDF similar to this 2D BSDF slice, together with a sampling technique that generates samples according to this PDF.

Moreover, the modern rendering pipeline usually supports multiple importance sampling (MIS). In the MIS framework, a renderer should implement not only BSDF sampling, but also BSDF evaluation (returning the BSDF value given the incident and outgoing directions) and PDF query (returning the PDF value similarly).

3.2 Understanding BSDF importance sampling

In this subsection, we analyze the essence of BSDF importance sampling, and present insight and motivation for our importance baking scheme.

We start with two issues from current BSDF importance sampling solutions.

(1) Not all parametric BSDFs can be analytically sampled. This includes certain difficult-to-sample NDFs such as GTR, certain sampling methods such as visible NDF (VNDF) sampling only working with specific NDFs, as well as multiple-bounce BSDFs in general.

Even analytic sampling methods are not perfect. In order to achieve the best quality, BSDF sampling requires a reasonable pdf \( p(x) \) close to \( f_s(\omega_i, n) \), but most analytical methods only sample the VNDF (ignoring the Fresnel term) or even just NDF (ignoring everything else).

To deal with these issues and achieve perfect importance sampling, we propose our understanding in the essence of the sampling process – drawing samples according to a specific PDF is actually mapping from a uniform distribution to that PDF, viz., \( f : (\xi(0), \xi(1)) \mapsto (u, v) \) (in 2D). As an example, we visualize such a
mapping acquired on a 2D BSDF slice in Fig. 2. In this visualization, we assume the pixels represent a uniform grid on the unit square \([0,1]^2\), and each pixel \(i\) at position \((\xi_i^{(0)}, \xi_i^{(1)})\) stores its mapped position \((u_i, v_i)\) as red and green. In this way, we are able to generate an image for any 2D mapping, which we name the importance map.

With this understanding, we immediately come up with the following insights.

1. For a 2D BSDF slice, it is not crucial whether it can be sampled analytically or not because analytic sampling only corresponds to a quick lookup on the importance map. Instead, being able to acquire and query a high-quality importance map is the actual key to perfect importance sampling. Fortunately, we demonstrate that the importance map can be precomputed with the help of optimal transport (Sec. 3.3).

2. Suppose one importance map can be obtained from a 2D BSDF slice defined with BSDF parameters \(\epsilon\) and the incident direction \(\omega_i\). In order for the full parametric BSDF to be importance sampled, we have to collect all importance maps of all combinations of \(\epsilon\) and \(\omega_i\). This requires heavy storage, and therefore compression is needed. Fortunately, we demonstrate that compression is not only possible but also efficient with the help of a lightweight neural network (Sec. 3.4).

Since we propose to precompute and compress the importance maps, the entire process is similar to the concept of light baking in real-time rendering. Therefore, we name it BSDF importance baking.

### 3.3 Optimal transport for precomputed importance

In this subsection, we focus on attaining the importance maps. We start from an important fact that is often ignored. That is, potential mappings that produce the same importance sampled PDF are not unique. Consider a toy example of sampling a truncated, and normalized 1D Gaussian defined on \([0,1]\) shown on the right. Suppose we uniformly subdivide \([0,1]\) into four segments \(A, B, C, D\), each integrates to a probability of 0.1, 0.4, 0.4 and 0.1. Then we subdivide the uniform \([0,1]\) into four segments with lengths 0.1, 0.4, 0.4 and 0.1, and name them 1, 2, 3 and 4, respectively. Then the mapping \(1 \mapsto A, 2 \mapsto B, 3 \mapsto C, 4 \mapsto D\) is a valid mapping that importance samples the Gaussian, but the mapping \(1 \mapsto A, 2 \mapsto C\) and \(3 \mapsto B, 4 \mapsto D\) is also perfect importance sampling.

Moreover, from this example, one can immediately tell which sampling strategy is better: the first one is much smoother, and the second one suffers from discontinuity. This conclusion is never trivial because it directly proves that one pervasively used solution to obtain the importance maps — the marginalized inverse transform sampling, a.k.a. row-column sampling — is not suitable for generating good importance maps (Fig. 3). This is because a small perturbation in \(\xi_0\) may result in a different row with a different 1D distribution, then even for similar \(\xi_1\), the resulting column can be far away.

Therefore, a good property that we prefer is locality, i.e., the neighborhood of a point also maps to the neighborhood of its target according to the same mapping \(f: N(\xi^{(0)}, \xi^{(1)}) \mapsto N(u, v)\). To satisfy this requirement, we refer to optimal transport (OT) — specifically, discrete optimal transport from the Lagrangian view [Feydy et al. 2019], which is able to find an optimal 1-to-1 mapping between two point distributions with the same number of points. In our case, this is to map from the unit square to the 2D BSDF slice for each combination of BSDF parameters \(\epsilon\) and incident directions \(\omega_i\). To conduct optimal transport, we first discretize both distributions into ordered point sets

\[
\alpha = \bigcup_{i=1}^{n} \delta(\xi_i^{(0)}, \xi_i^{(1)}), \quad \beta = \bigcup_{j=1}^{n} \delta(u_j, v_j),
\]

where \(n\) is the total number of points, and the \(\delta(\cdot)\) is Dirac delta impulse at different positions. We weigh each point the same, which immediately indicates that it is the local densities of those points that represent the values of the original continuous distributions. In other words, the continuous-to-discrete conversion itself is exactly importance sampling. We manually convert the unit square into a regular grid \(\alpha\) (pixels) and use row-column sampling to convert the 2D BSDF slice into \(\beta\).

Then we conduct optimal transport, giving a 1-to-1 correspondence between any two point distributions \(\alpha\) and \(\beta\), minimizing
the Euclidean distance between them:
\[
\arg \min_{\psi} \sum_{i=1}^{n} \| \mathbf{a}(i) - \mathbf{b}(\psi(i)) \|_2,
\]
where \(\psi\) is a permutation of the sequence \(1, 2, \ldots, n\), computed by the optimal transport process. After this, we record the positions of each \(\mathbf{b}(\psi(i))\) into each pixel's red and green channel, which completes the computation of an importance map.

Note that during the conversion of \(\mathbf{b}\), we used row-column sampling. However, this is in essence different from using that to find the mapping – we only use row-column sampling to discretize an image into points, and the mapping is found by optimal transport. More conversion tools can be explored in the research area of image stippling [Kim et al. 2008].

Also, note specifically that we focus on the black box usage of optimal transport as a general mathematical tool. However, we do not intend to compare or improve specific optimal transport solvers. We also do not extend further discussion on specific accurate/approximate distance metrics (e.g., earth mover’s distance, Wasserstein distance, Kullbach–Leibler divergence, Sinkhorn distance, etc.). In the supplementary document, we provide our choices for implementation.

3.4 A lightweight neural network for importance baking

With the acquired importance maps, we present our neural solution to compress them, completing our importance baking scheme in this subsection.

We start with a few design principles from a series of observations.

1. We notice a significant amount of similarities between the importance maps when the BSDF parameters and incident directions change. The smooth change of these importance maps inspire us to use a neural network to compress them.

2. As mentioned in Sec. 2, since runtime performance is crucial to core rendering, the highest level design of our neural network is to keep it as lightweight as possible, thus allowing for fast inference during rendering.

3. During BSDF importance sampling, only one outgoing direction needs to be sampled at a time. Therefore, the importance map should be point queried instead of being output on the whole. This also further reduces the complexity of our neural network.

4. For Monte Carlo estimation, the BSDF sampling process is expected to output not only an outgoing direction but also its sampling weight, which is the BSDF value divided by the PDF value. In our case, the sampling weight is a 3-channel value and is almost constant since we design our PDFs to have the same shapes with the BSDF slices converted to grayscale.

5. For completeness to run in modern renderers with MIS framework, we also compress the BSDFs and PDFs using separate neural networks. Since we have full control of the precomputed sampling process, the BSDF and PDF values are readily available.

Based on these design principles, we propose three lightweight neural networks for the three different tasks.

**BSDF sampling network.** Aside from the BSDF parameters \(\epsilon\) and the incident direction \(\omega_i\), our importance sampling network takes two random numbers and outputs the sampled outgoing direction together with its sampling weight:

\[
I(\epsilon, \omega_i, \xi_0, \xi_1) = (\omega_o, sw_R, sw_G, sw_B).
\]

To validate the functionality of our BSDF sampling network, we compare the importance maps and the binning results between the ground truth and our network’s prediction, as shown in Fig. 4. Thanks to our use of optimal transport that results in smooth importance maps, our network is able to learn to compress them well.

**BSDF evaluation network.** For BSDF evaluation, our evaluation network \(\mathcal{E}\) takes an additional outgoing direction \(\omega_o\) and outputs the BSDF value as a 3-channel RGB value:

\[
\mathcal{E}(\epsilon, \omega_i, \omega_o) = f_{\mathcal{E}}(\omega_i, \omega_o, \omega_o) \langle \omega_i, n \rangle.
\]

**PDF query network.** Our PDF query network \(\mathcal{P}\) has a very similar definition compared to the BSDF evaluation network. It also takes in the combinations of BSDF parameters as well as the outgoing directions and returns the PDF value of sampling that direction under solid angle measurement:

\[
\mathcal{P}(\epsilon, \omega_i, \omega_o) = \text{PDF}(f_{\mathcal{E}}(\omega_i, \omega_o, \omega_o) \langle \omega_i, n \rangle),
\]
Yaoyi Bai, Songyin Wu, Zheng Zeng, Bei Bei Wang, and Ling-Qi Yan

NDF, 170 spp, 15.20 s
relMSE 0.0036

VNDF, 170 spp, 15.23 s
relMSE 0.0025

Ours, 128 spp, 15.05 s
relMSE 0.0030

Fig. 6. Equal-time comparison between our method, NDF sampling and VNDF sampling for single-bounce anisotropic microfacet BSDFs. Our method has much less noise than NDF sampling and has similar noise as VNDF sampling, especially around the grazing angles.

and we leave the PDF conversion details to the supplementary document.

All the three neural networks share a similar lightweight structure, as illustrated in Fig. 5. Note specifically that the emphasis on the neural network design in core rendering [Fan et al. 2022; Zhu et al. 2022] is different from that in deep learning. We use neural networks only as a general tool for efficiently compressing high-dimensional data. Please find more details regarding the training (e.g., loss function, data generation, hyperparameters) and inference (e.g., renderer integration) of our lightweight neural networks in the supplementary document.

4 RESULTS AND COMPARISON

We have implemented our network in Mitsuba renderer [Jakob 2010]. The network inferences are integrated with C++ and Mitsuba using Eigen [Guennebaud et al. 2010]. All the rendering performance is measured on an Intel 8-core i9-9900K machine.

In this section, we validate our results on both rough conductors and rough dielectric with the GGX NDF. We also compare with previous works on both single scattering importance sampling and multiple scattering importance sampling.

We use relative mean square error (relMSE) to measure the difference with the ground truth. Since our method and the other methods might converge to different ground truths, we use their own converged results as the ground truth.

Matpreview Scene. Fig. 6 shows a Matpreview scene with $\alpha_x = 0.1$ and $\alpha_y = 0.04$ under an environment lighting. In this scene, we only consider the single scattering in the microfacet model. We compare our sampling model with NDF sampling [Cook and Torrance 1982], and VNDF sampling [Heitz and d’Eon 2014] with equal time. We only use BSDF sampling and direct lighting to show the results clearly. By comparison, our result is less noisy than the NDF sampling, while our result is quite similar to VNDF sampling, since VNDF is already a good approximation of the BRDF slice. Note that VNDF sampling is not general and can only work with stretch-invariant NDFs, like Beckmann and GGX distribution functions.

Elephant scene. In Fig. 7, we show an elephant with $\alpha_x = 0.5$ and $\alpha_y = 0.1$ under an environment light. In this scene, we consider multiple scattering in the microfacet model and focus on direct lighting with BSDF sampling. We use our network trained on the multiple-bounce dataset for BSDF sampling and compare against Heitz et al. [2016], and Xie et al. [2019] with equal time. The result by Heitz et al. [2016] shows high variance, while Xie et al. [2019] produces a biased result. Our result has less noise than Heitz et al. and shows the best quality.

Ginkgo ornament scene. Our model provides a full suit, including BSDF eval, sample, and PDF networks, to enable MIS for Monte Carlo rendering. We validate these three networks, by comparing the light sampling, BSDF sampling, and MIS in our method on different roughness ($\alpha = 0.08$ for the top and $\alpha = 0.3$ for the bottom). BSDF sampling works better on low-roughness materials and light sampling works better on high roughness materials, while their combination (MIS) produces the highest quality.
work for sampling should not be more complex than training for with different indices of refraction \( \eta \) which include both reflection and refraction. The Bathroom Shelf Scene in Fig. 9 shows some jars and bottles under environment lighting, which have multiple-bounce dielectric microfacet materials as well.

larger roughness show the opposite status. MIS always produces the highest-quality results.

**Display Shelf Scene.** Fig. 1 combines various multiple-bounce conductor microfacet BSDFs as a teaser to show the capability of our neural network. The scene is illuminated by an area light. The materials rendered with our neural BSDFs are marked in the figure. Compare with Heitz et al, our results achieve even lower noise level with fewer spp.

**Bathroom Shelf Scene.** Our network also supports dielectric BSDFs, which include both reflection and refraction. The Bathroom Shelf Scene in Fig. 9 shows some jars and bottles under environment lighting, which have multiple-bounce dielectric microfacet materials with different indices of refraction \( \eta \) and roughness.

5 DISCUSSION AND LIMITATIONS

**Unified representation of BSDF evaluation and sampling.** In our BSDF evaluation network, when the BSDF parameters are specified, one incident direction \( \omega_i \) will produce a corresponding 2D BSDF slice, where each pixel is a BSDF value. And this BSDF slice is queried using the 2D outgoing direction \( \omega_o \). Similarly, given the BSDF parameters and the incident direction, a 2D importance map is determined, where a pixel contains the mapped sample position and the sampling weight (optional). And it is queried with a 2D random number \((x_0, x_1)\). With this similarity identified, we have the evidence to claim that BSDF evaluation and sampling (and PDF query) are in essence very similar. Therefore, training a neural network for sampling should not be more complex than training for evaluation. This observation also give us the insight that prohibitively expensive neural network structures, such as RealNVP, could be avoided.

**More complex BSDFs.** Constrained by the length of the paper, we only use single- and multiple-bounce microfacet conductors and dielectrics as representatives for parametric BSDFs. But we would like to point out that there are many other interesting parametric BSDFs, such as Disney principled materials [Burley and Studios 2012], iridescent materials [Guillén et al. 2020], layered materials [Guo et al. 2018] and so on. These parametric BSDFs have even higher-dimensional but similarly smooth parameter spaces. Hence, we believe that using a lightweight neural network to sample these parametric BSDFs is still potentially a good solution. We would like to leave this as an interesting and important future work that may further improve the practicality of our method.

**Non-parametric/Measured BSDFs.** Apart from parametric BSDFs, there are also possibilities for our method to be used for non-parametric or measured BSDFs. Intuitively, non-parametric BSDFs are often large data blocks. However, they are in fact of much lower dimensions than the parametric BSDFs. For example, the measured 4D BRDF collection or the 6D bidirectional texture function (BTF) have even smaller parameter spaces than our training data. In this case, we believe that our method can show even better capability for data compression and sampling calculation with the same data generation and training routine.

**Performance.** Our results correctly reconstruct the parametric BSDFs’ appearance with comparatively less noise, but the performance influence cannot be neglected even though our networks are small. Since we simply use inline CPU integration to fully integrate our neural networks into the renderer with minimum revision to the rendering pipeline, the inference of our neural networks is far from optimized. There are many ways to further improve inference efficiency, such as using a GPU inference framework, e.g., TensorRT, and/or devoting to considerable engineering optimization [Müller et al. 2021].

**Bias.** Neural network prediction will inevitably introduce bias even though the original training data are unbiased. Strict applications such as white furnace test will expose the bias issue immediately. Our method does have visible bias (for example, the sharp angle in the crop of ours in Fig. 6), but we do not see apparent problems in these results from our practical applications. Nevertheless, we believe that it is still meaningful to study other methods that enable unbiased importance map compression.

6 CONCLUSION AND FUTURE WORK

We have introduced BSDF importance baking, a lightweight neural solution to perform perfect importance sampling of parametric BSDFs. We start from the observation that the mapping that performs importance sampling on a BSDF slice can be simply recorded as a 2D importance map. Following this observation, we propose to use optimal transport to accurately precompute the importance maps, then we use a lightweight neural network to efficiently compress them. Together with a BSDF evaluation network and a PDF query network, our method enables full multiple importance sampling (MIS) without any revision to the rendering pipeline. Compared with previous methods, we demonstrate reduced noise levels on rendering results with a rich set of appearances, including both conductors and dielectrics with anisotropic roughness.
We believe that we have brought about novel contributions: our method is the first to utilize optimal transport in rendering applications that are not affected by its heavy computation during runtime; our method is the first practical method that has comparable or better performance compared with existing methods; and our method is the first complete neural alternative that has potential to fully replace parametric BSDFs with MIS.

In the future, an immediate research direction is to improve the quality of our lightweight neural networks and further improve their performance. It is also useful to extend current support from microfacet BSDFs to other parametric and non-parametric BSDFs. Sampling other forms of appearance representation could also be interesting, for example, using our importance baking scheme to investigate the sampling problem of 4D light fields data or 5D neural radiance fields (NeRF) data. Apart from neural compression, other data compression strategies could be explored as well, until a more efficient or unbiased method is found.

REFERENCES

Michael Ashikhmin and Peter Shirley. 2001. An Anisotropic Phong Light Reflection Model. Journal of Graphics Tools 5 (01 2001).

P. Beckmann and A. Spizzichino. 1963. The scattering of electromagnetic waves from rough surfaces. Pergamon Press.

Nicolas Bonneel and David Coeurjolly. 2019. SPOT: Sliced Partial Optimal Transport. ACM Transactions on Graphics (SIGGRAPH) 38, 4 (2019).

Nicolas Bonneel, Gabriel Peyré, and Marco Cuturi. 2013. Sinkhorn distances: Lightspeed computation of optimal transport. Commun. ACM 66, 6 (Jun 2013), 31–37. https://doi.org/10.1145/2461912.2461913

Nicolas Bonneel, Michel Van De Panne, Sylvain Paris, and Wolfgang Heidrich. 2011. Displacement interpolation using Lagrangian mass transport. In Proceedings of the 2011 SIGGRAPH/Asia conference. 1–12.

Brent Burley and Walt Disney Animation Studios. 2012. Physically-based shading at Disney. In ACM SIGGRAPH, Vol. 2012. vol 2012, 1–7.

Robert I. Cook and Kenneth E. Torrance. 1982. A reflectance model for computer graphics. ACM Transactions on Graphics (ToG) 1, 1 (1982), 7–24.

Nicolas Bonneel, Gabriel Peyré, and Marco Cuturi. 2016. Wasserstein barycentric coordinates: histogram regression using optimal transport. ACM Trans. Graph. 35, 4 (2016), 71–1.

Nicolas Bonneel, Guilherme Vaz De Panne, Sylvain Paris, and Wolfgang Heidrich. 2011. Displacement interpolation using Lagrangian mass transport. In Proceedings of the 2011 SIGGRAPH Asia conference. 1–12.

Eric Heitz. 2014. Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs. Journal of Computer Graphics Techniques (JCGT) 3, 2 (June 2014), 85–107.

Eric Heitz. 2017. A Simpler and Exact Sampling Routine for the GXX Distribution of Visible Normals. Ph.D. Dissertation. Unity Technologies.

Eric Heitz. 2018. Sampling the GXX distribution of visible normals. Journal of Computer Graphics Techniques 7, 4 (2018), 1–15.

Eric Heitz and Eugene d’Eon. 2014. Importance sampling microfacet-based BSDFs using the distribution of visible normals. In Computer Graphics Forum. Vol. 33. Wiley Online Library, 103–112.

Eric Heitz, Johannes Hanika, Eugene d’Eon, and Carsten Dachsbacher. 2016. Multiple-scattering microfacet BSDFs with the Smith model. ACM Transactions on Graphics (TOG) 35, 4 (2016), 14–14.

Bingyang Hu, Jie Guo, Yanjun Chen, Mengtian Li, and Yanwen Guo. 2020. DeepBRDF: A Deep Representation for Manipulating Measured BRDF. In Computer Graphics Forum, Vol. 39. Wiley Online Library, 157–166.

Wenzel Jakob. 2010. Mitsuba renderer.

Leonid V Kantorovich. 1942. On the translocation of masses. In Dokl. Akad. Nauk. USSR (NS), Vol. 37. 199–201.

Dongyeon Kim, Minjun Son, Yunjin Lee, Henry Kang, and Seungyong Lee. 2008. Feature-guided image stitching. In Computer Graphics Forum, Vol. 27. Wiley Online Library, 1209–1216.

Alexandr Kuznetsov, Krishna Mullia, Zexiang Xu, Milosi Haian, and Ravi Ramamoorthi. 2021. NeuMIP: Multi-Resolution Neural Materials. Transactions on Graphics (Proceedings of SIGGRAPH) 40, 4, Article 175 (July 2021), 13 pages.

Joo Lee, Adrian Jarabo, Daniel Jesi, Diego Gutierrez, and Min Kim. 2018. Practical multiple scattering for rough surfaces. ACM Trans. Graph. 37, Article 175 (Dec. 2018), 12 pages.

Gaspard Monge. 1781. Mémoire sur la théorie des déblais et des remblais. Histoire de l’Académie Royale des Sciences de Paris (1781).

Eric Heitz. 2018. Sampling the GGX distribution of visible normals.

Eric Heitz. 2014. Understanding the Masking-Shadowing Function in Microfacet-Based BRDFs. Journal of Computer Graphics Techniques (JCGT) 3, 2 (June 2014), 85–107.

Eric Heitz. 2017. A Simpler and Exact Sampling Routine for the GXX Distribution of Visible Normals. Ph.D. Dissertation. Unity Technologies.

Eric Heitz. 2018. Sampling the GXX distribution of visible normals. Journal of Computer Graphics Techniques 7, 4 (2018), 1–15.

Eric Heitz and Eugene d’Eon. 2014. Importance sampling microfacet-based BSDFs using the distribution of visible normals. In Computer Graphics Forum. Vol. 33. Wiley Online Library, 103–112.

Eric Heitz, Johannes Hanika, Eugene d’Eon, and Carsten Dachsbacher. 2016. Multiple-scattering microfacet BSDFs with the Smith model. ACM Transactions on Graphics (TOG) 35, 4 (2016), 1–14.

Bingyang Hu, Jie Guo, Yanjun Chen, Mengtian Li, and Yanwen Guo. 2020. DeepBRDF: A Deep Representation for Manipulating Measured BRDF. In Computer Graphics Forum, Vol. 39. Wiley Online Library, 157–166.

Wenzel Jakob. 2010. Mitsuba renderer.

Leonid V Kantorovich. 1942. On the translocation of masses. In Dokl. Akad. Nauk. USSR (NS), Vol. 37. 199–201.

Dongyeon Kim, Minjun Son, Yunjin Lee, Henry Kang, and Seungyong Lee. 2008. Feature-guided image stitching. In Computer Graphics Forum, Vol. 27. Wiley Online Library, 1209–1216.

Gilles Rainer, Abhijit Ghosh, Wenzel Jakob, and Tim Weyrich. 2019. Neural BRDF Representation and Importance Sampling. In Computer Graphics Forum, Vol. 40. Wiley Online Library, 332–346.

Bruce Welford, Stephen R. Marschner, Hongsong Li, and Kenneth E Torrance. 2007. Microfacet Models for Refraction through Rough Surfaces. Rendering techniques 2007 (2007), 18th.

Feng Xie and Pat Hanrahan. 2018. Multiple Scattering from Distributions of Specular Y-Grooves. ACM Trans. Graph. 37, 6, Article 276 (2018), 14 pages.

Feng Xie, Anton Kaplanayan, Warren Hunt, and Pat Hanrahan. 2019. Multiple scattering using machine learning. In ACM SIGGRAPH 2019 Talks. 1–2.

Chunhong Zheng, Ruzhang Zheng, Rui Wang, Shuang Zhao, and Hujun Bao. 2021. A Compact Representation of Measured BRDFs Using Neural Processes. ACM Transactions on Graphics (TOG) 41, 2 (2021), 1–15.

Junqiu Zhu, Si Zhai, Lu Wang, Yanning Xu, and Ling-Qi Yan. 2022. Practical Level-Of-Detail Aggregation of Fur Appearance. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2022) 41, 4 (2022).