Inverse Rendering of Translucent Objects with Shape-Adaptive Importance Sampling

Jooeun Son[1] , Yucheol Jung[1] , Gyeongmin Lee[1] , Soongjin Kim[1] , Joo Ho Lee[2] , Seungyong Lee[†1]

[1]POSTECH, South Korea
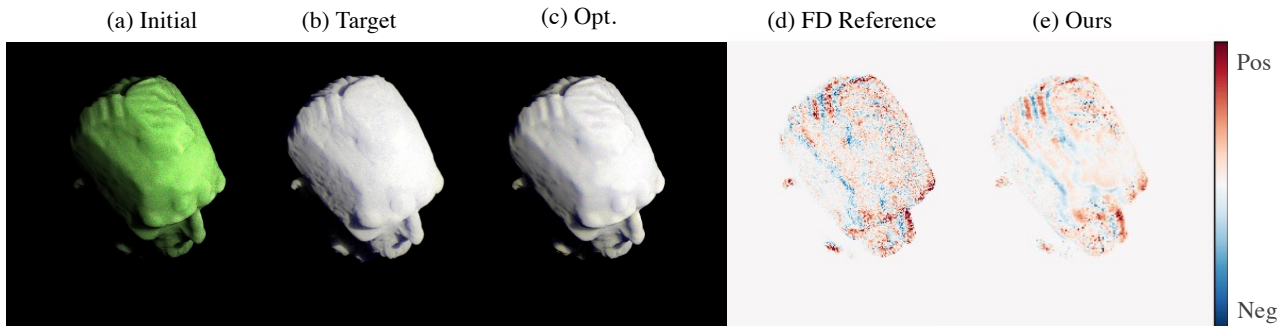[2]Sogang University, South Korea

Figure 1: *Inverse rendering with shape-adaptive importance sampling. (a-c) We jointly optimize scattering parameters σ_t and α of a homogeneous translucent object using the shape-BSSRDF model [VKJ19]. We address the challenges in differentiability posed by the importance sampling framework by approximating gradients using offset samples. (d, e) Our method computes highly accurate gradients with low variance.*

Abstract

Subsurface scattering is ubiquitous in organic materials and has been widely researched in computer graphics. Inverse rendering of subsurface scattering, however, is often constrained by the planar geometry assumption of traditional analytic Bidirectional Surface Scattering Reflectance Distribution Functions (BSSRDF). To address this issue, a shape-adaptive BSSRDF model has been proposed to render translucent objects on curved geometry with high accuracy. In this paper, we leverage this model to estimate parameters of subsurface scattering for inverse rendering. We compute the finite difference of the rendering equation for subsurface scattering and iteratively update material parameters. We demonstrate the performance of our shape-adaptive inverse rendering model by analyzing the estimation accuracy and comparing to inverse rendering with plane-based BSSRDF models and volumetric methods.

CCS Concepts
• Computing methodologies → Rendering;

1. Introduction

Inverse rendering of 3D objects has been studied intensively for decades to bridge the gap between the real and digital worlds. Beyond 3D reconstruction, accurate material estimation facilitates seamless transmission of real objects to virtual scenes and vice versa via 3D printing. Most inverse rendering works [ZSD*21, ZLLS22, JLX*23] focus on reconstruction of objects with basic

materials, such as those exhibiting diffuse or specular surface reflection. Relatively few studies address material estimation for objects with subsurface scattering phenomena, which are prevalent in the appearance of everyday objects, such as human skin and fruits.

Subsurface scattering is a physical event of light transport where light penetrates the surface, scatters inside the translucent material, and exits at the adjacent surface point. To compute the contribution of subsurface scattering numerically, volumetric path tracing methods simulate the light transport of photons stochastically [GZB*13, ZYZ21, CLZ*20]. However, such brute-force sim-

---

† Corresponding Author

ulation requires a large number of path-tracing samples for the rendering to converge.

In contrast, analytic subsurface scattering models such as BSS-RDFs [JMLH01, DJ05, HCJ13] successfully render translucent objects using a much smaller number of path-tracing samples by estimating the contribution of adjacent surface points based on a strong geometric assumption that the surface is planar and has infinitely thick volume below. Still, the BSSRDF suffers from compromises in the rendering accuracy caused by assuming planar geometry and ignoring the shape of the local neighborhood.

To address this issue, Vicini et al. [VKJ19] proposed a shape-adaptive subsurface scattering method, where a sampling network learns how to sample adjacent surface points according to the shape of local neighborhood and medium parameters, resulting in accurate rendering. However, adopting the shape-adaptive BSSRDF model to inverse rendering is not straightforward.

In this work, we analyze and solve the difficulties of differential rendering with the shape-adaptive BSSRDF model to build a differentiable shape-adaptive BSSRDF framework. First, the Monte Carlo importance sampling required for rendering with the shape-adaptive BSSRDF is affected by the medium parameters, but the derivative of this sampling operation with respect to the medium parameter is difficult to compute. To resolve this problem, we propose an offset sample-based gradient estimation. Second, while the absorption by the medium affects the image, the stochastic behavior of absorption makes the differentiation non-trivial. We handle this problem by utilizing simple reparameterization to obtain gradients stably with standard automatic differentiation (Fig. 1).

Compared to the previous approach using BSSRDFs, our method using the differentiable shape-adaptive BSSRDF estimates material parameters for translucent objects more accurately by removing unrealistic planar geometric assumptions, while retaining stable optimization with a small number of samples per pixel (spp). We demonstrate our numerical accuracy and stable convergence via thorough experiments with synthetic data.

2. Related Work

2.1. Rendering of Translucent Objects

A prominent method for achieving photorealistic rendering is Monte Carlo path tracing [KVH84, Rus88], which simulates explicit light paths resulting from scattering events. While this approach can produce highly realistic images, it is computationally expensive. This challenge is exacerbated for participating media with high scattering albedos, leading to prolonged random walks. To address this issue, several works have enhanced the sampling efficiency of Monte Carlo rendering techniques by integrating Dwivedi sampling, which guides random walks toward surfaces for early termination, thus reducing computational costs [Kd14, MHD16].

Another significant approach involves the use of BSSRDFs. BSSRDFs mathematically model how incoming light is transported across a surface by solving the Radiative Transfer Equation (RTE) at the surface boundary. Incorporating diffusion theory simplifies the RTE, enabling straightforward solutions using dipole [JMLH01], multipole [DJ05], or photon beam [HCJ13] configurations.

Although diffusion-based methods are computationally efficient, they fall short in physical accuracy compared to full light path simulations, particularly for low-scattering materials. This limitation arises because diffusion-based models typically assume a planar geometry, which does not accurately represent complex surfaces. Empirical solutions have been developed to better model low-scattering materials [DLR*09] and to integrate single and multiple scattering effects without separate modeling [Chr15].

To further improve accuracy, Vicini et al. [VKJ19] proposed a shape-adaptive BSSRDF model for importance sampling that accounts for surface geometry variations, enhancing the realism for translucent objects. The shape-adaptive BSSRDF model utilizes a variational autoencoder (VAE) [KW13] to model the high-dimensional space of scattering distributions conditioned on the shape of local neighborhoods and material information. The VAE performs shape-adaptive importance sampling of exit positions of lights. In our work, we design an inverse rendering framework by building a differentiable shape-adaptive BSSRDF to infer material parameters from images of translucent objects.

2.2. Differentiable Rendering

Differentiable rendering is an extensively studied topic in the graphics and vision community for estimating scene parameters from images, such as object shape, lighting, texture, and material parameters. Most rasterizer-based differentiable renderers [LB14, KUH18, GCM*18, LLCL19] are interested in handling simplified lighting models, such as the Phong model [Pho75], that do not consider the effect of subsurface scattering. Recent neural renderering works [ZSD*21, ZLLS22, JLX*23] showcase joint reconstructions of geometry and physically-based material parameters, but these works also assume the objects do not exhibit subsurface scattering. Recent differentiable renderers focusing on realistic appearance are based on Monte Carlo path tracing [TSG19, LADL18, NDVZJ19] to handle global lighting or adopt more realistic materials such as microfacet models [WMLT07]. Still, many path-tracing differentiable renderers are designed for opaque objects. Our work builds an accurate differentiable renderer for translucent objects based on Monte Carlo path tracing.

One of the key challenges in differentiable rendering is obtaining meaningful gradients from non-differentiable discontinuities so that the scene parameters are updated according to the position of discontinuities, e.g., deforming object shape to fit object silhouette. Zhang et al. [ZWZ*19] identified three types of edges responsible for such parameter-dependent discontinuities: boundary, silhouette, and sharp edges, which are inevitable when optimizing the scene geometry. Parameter-dependent discontinuities have been handled via additional sampling at the discontinuity [LADL18, ZMY*20, ZYZ21], reparameterization of rendering integrals [LHJ19, BLD20, XBLZ23], and importance sampling of boundary light paths [YLB*22]. Our work also addresses similar parameter-dependent discontinuities; a material parameter for subsurface scattering affects the sampling distribution for Monte Carlo path tracing. Depending on the material parameter, some samples

can be occluded by object surface, creating discontinuity. Handling this parameter-dependent discontinuity enhances stability and accuracy of the parameter estimation in scenes with complex geometries. In our work, we demonstrate that we can simply use forward differences to effectively handle this issue.

2.3. Inverse Rendering of Translucent Materials

Accurate rendering of translucent objects can be achieved via Monte Carlo volumetric path tracing. Thus, a natural approach for inverse rendering of translucent objects is to utilize differentiable renderers based on volumetric path tracing [GZB*13, CLZ*20, ND-SRJ20, ZYZ21, VSJ21]. However, explicit simulation of light paths within the internal volume is more computationally expensive than standard path tracing using only surface samples. Volumetric path tracing requires a higher number of samples to achieve convergence.

Several alternatives to volumetric path tracing for inverse rendering of translucent materials have been proposed. One approach involves learning relightable neural scene representations by approximating single and multiple scattering effects [ZSS21] or using Microflake theory [JAM*10, HDCD15]. Another approach uses BSSRDFs for various applications, such as estimating medium parameters from a single image [MES*11] or handling heterogeneous media [ZIK*17].

Recently, Deng et al. [DLW*22] demonstrated that the generalized differential path tracing framework [ZMY*20] can be extended to support BSSRDF models by incorporating the simple dipole BSSRDF model [JMLH01]. This approach circumvents volumetric path tracing by using BSSRDFs, but makes unrealistic geometric assumptions of planar geometry, thereby compromising the accuracy of material parameter estimation. In contrast, our work employs a shape-adaptive BSSRDF, which does not impose the planar geometry assumption and yields a more accurate simulation of subsurface scattering without the need for volumetric path tracing.

3. Preliminary

3.1. Inverse Rendering for Translucent Objects

The physical processes that contribute to the translucent appearance of an object are absorption and scattering. They determine the distribution of paths that light travels in and affect the appearance of an translucent object. The medium parameters that affect absorption and scattering are the extinction coefficient σ_t, scattering albedo α, and anisotropy parameter g [NGHJ18]. Assuming point lighting, the rendering equation using BSSRDF is given as:

$$L_o(x_o, \omega_o) = \int_A S(x_o, x_i, \omega_i, \omega_o) \, L_i(x_i, \omega_i) \, |n \cdot \omega_i| \, dx_i, \quad (1)$$

where L_o is the outgoing radiance at surface point x_o in the direction ω_o, S is the BSSRDF, L_i is the incident radiance at x_i along direction ω_i, and n is the surface normal at x_i. For a point light source, $L_i(\mathbf{x}_i, \omega_i) = V(x_i, \omega_i) \, I/r^2$, where V is the visibility term, I is the radiant intensity, and r is the distance between x_i and the light source. With the shape-adaptive importance sampling framework, both the BSSRDF S and the sampling of x_i are affected by scattering parameters.

The estimation of the medium parameters can be achieved via inverse rendering that consists of two stages: 1) a primal rendering phase that produces the image from parameters, 2) a differential rendering phase that computes the gradients with respect to parameters. The computed gradients are used to minimize the error between primal renderings and target images using an iterative optimization method such as gradient descent.

In this work, we reconstruct two scalar-valued medium parameters $\sigma_t \in \mathbb{R}$ and $\alpha \in \mathbb{R}$ from target images. We make a few assumptions in the problem formulation for simplicity. We assume a homogeneous material, and the anisotripy parameter $g = 0$. The geometry and lighting are given in advance. We ignore the effect of indirect illumination and assume light paths $\bar{x}$ are ordered sequences of points in the form $\bar{x} = (x_{light}, x_i, x_o, x_{cam})$, where x_{light} and x_{cam} are the positions of the point light and camera, respectively.

3.2. Shape-adaptive BSSRDF

The key idea of shape-adaptive BSSRDF [VKJ19] is to adapt the BSSRDF S to the medium and shape of the local neighborhood:

$$S(x_o, x_i, \omega_i, \omega_o) = S_p(x_o, x_i, \omega_i; \theta_{med}, \theta_{shape}) S_d(\omega_o), \quad (2)$$

where S_p represents a learnable term conditioned on medium parameters $\theta_{med} = (\alpha, \sigma_t, g)$ and local polynomial shape descriptors θ_{shape}, and S_d is an analytically implemented term for addressing the effect of Fresnel transmission. Compared to using a fixed BSSRDF for all points on the surface, this shape-adaptive formulation provides accurate rendering of translucent materials by removing the unrealistic assumption of a planar surface.

Instead of directly implementing the learnable function S_p, Vicini et al. [VKJ19] trained a sampling network, which is implemented as a VAE, to sample positions from a distribution proportional to S_p. Sample points x_i are generated by first sampling 3D positions x_{VAE} near the point x_o, and projecting the sampled points onto the surface using a projection operation Π:

$$x_i(\theta_{med}, \theta_{shape}) = \Pi(x_o + \sigma_n(\theta_{med}) \, x_{VAE}(\theta_{med}, \theta_{shape})), \quad (3)$$

where $\sigma_n(\theta_{med})$ is the standard deviation determined by the medium parameters θ_{med}.

Since the sampling density of x_i is proportional to S_p, the Monte Carlo integration in rendering cancels the S_p term and the probability density function $p(x_i)$:

$$\sum_{x_i} \frac{L_i(x_i, \omega_i) \, S_p(x_o, x_i, \omega_i; \theta_{med}, \theta_{shape}) \, S_d(\omega_o) \, V(x_i, \omega_i)}{p(x_i)} \quad (4)$$

$$= \sum_{x_i} L_i(x_i, \omega_i) \, S_{abs} \, S_d(\omega_o) \, V(x_i, \omega_i), \quad (5)$$

where S_{abs} is a scaling factor determined by a separate absorption network $Absorb(x_o; \theta_{med}, \theta_{shape})$, which we elaborate in Section 4.3. For more details of shape-adaptive BSSRDF, such as the formulation of σ_n, refer to the supplementary document.

4. Differentiable Shape-adaptive BSSRDF

For any scene parameter θ (e.g., medium parameters), the image derivative with respect to the parameter θ can be computed via the

differential rendering equation:

$$\partial_\theta L_o(x_o, \omega_o) = \partial_\theta \int f(x_i; \theta) \, dx_i, \tag{6}$$

where $f(x_i) = S(x_o, x_i, \omega_i, \omega_o) \, L_i(x_i, \omega_i) \, |n \cdot \omega_i|$ is the integrand of Eq. (1). In the following, we analyze the challenges in differential rendering and design our solutions for building a differentiable shape-adaptive BSSRDF.

4.1. Challenges in Differentiation of Importance Sampling

The shape-adaptive BSSRDF samples exiting positions x_i from a density proportional to S_p. Since the sampled points x_i are influenced by the medium parameters via σ_n (Eq. (3)), differentiation of the importance sampling iprocess is required. However, calculating the derivatives for the importance-sampled positions is not trivial.

We consider the implementation of importance sampling where samples in a canonical domain $u \in \mathcal{U}$, generated from an easily sampled distribution, are transformed into samples in the target domain $x_i \in \mathcal{X}$ with a desired probability density $p(x)$. The transformation is expressed as

$$x_i = T(u, \theta). \tag{7}$$

Here, we assume that the importance sampling function T is injective and differentiable for the purposes of theoretical analysis. We can reformulate the primal rendering equation as

$$L_o(x_o, \omega_o) = \int_{\mathcal{U}} f(T(u, \theta); \theta) |J(u)| \, du, \tag{8}$$

where $|J(u)|$ is the determinant of the Jacobian of T. As a result, the differential rendering equation (Eq. (6)) can be written as

$$\partial_\theta L_o(x_o, \omega_o) = \partial_\theta \int_{\mathcal{U}} f(T(u, \theta); \theta) |J(u)| \, du. \tag{9}$$

When T is a function of θ, the boundary $\partial \mathcal{U}$ of the canonical domain $\mathcal{U}$ may move. For example, a sample x_i originally visible for some θ may end up in an occluded region for a different value of θ. In this case, the boundary $\partial \mathcal{U}$ moves to exclude the point u corresponding to x_i out of the visible region. The accurate computation of gradients in the presence of such moving discontinuities has been addressed through the introduction of a boundary term.

Reynolds transport theorem [ZMY*20, ZYZ21] addresses the parameter-dependent discontinuities by separately handling the boundary integral:

$$\partial_\theta L_o(x_o, \omega_o) = \int_{\mathcal{U}} \partial_\theta \big[f(T(u, \theta); \theta) |J(u)| \big] \, du \tag{10}$$

$$+ \int_{\partial \mathcal{U}} \Delta \big[f(T(u, \theta); \theta) |J(u)| \big] \cdot v_{dis} \, du, \tag{11}$$

where v_{dis} is the velocity of the discontinuity boundary $\partial \mathcal{U}$ and Δ is the difference across the discontinuity boundary. Eqs. (10) and (11) are referred to as the interior and boundary terms, respectively. Below, we analyze the challenges in the calculation of each term for our case of shape-adaptive BSSRDF.

Interior term Previous differentiable rendering works have observed significant challenges when importance sampling is involved. We apply the analysis of Zeltner et al. [ZSGJ21] and refer interested readers to their work for details.
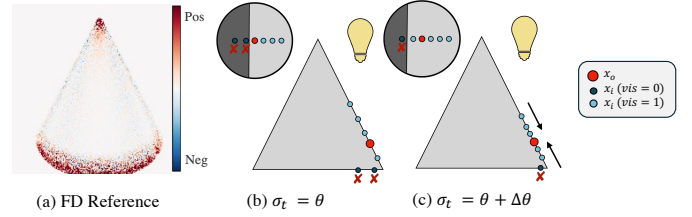

(a) FD Reference   (b) $\sigma_t = \theta$ (c) $\sigma_t = \theta + \Delta\theta$

Figure 2: *Illustration of parameter-dependent discontinuities. (a) We consider a point x_o on the lower lateral surface of the cone, where the pixel intensity increases with increasing σ_t. (b, c) As σ_t increases, samples x_i move closer to x_o. This shift causes an initially invisible sample (in dark blue) nearest to the discontinuity boundary $\mathbf{x}_B$ to become visible, thereby increasing the pixel value at x_o. This phenomenon is a form of moving discontinuities in the sample domain (inset with black edges).*
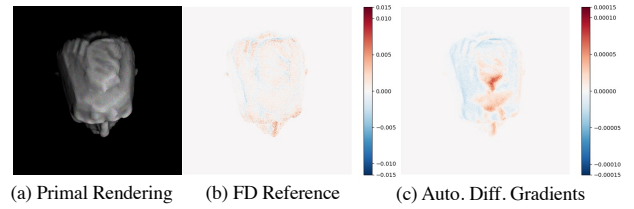

(a) Primal Rendering   (b) FD Reference   (c) Auto. Diff. Gradients

Figure 3: *Biased gradients due to the absence of the boundary term. We compare image derivatives with respect to σ_t computed by naïve automatic differentiation against those obtained by finite differences. (c) Naïve autodiff gradients fail to account for parameter-dependent discontinuities, resulting in biased gradients that differ significantly from (b) the finite difference reference.*

Zeltner et al. [ZSGJ21] argue that the transformation T can introduce significant variance to the interior term. This holds true even when the sampling density is perfectly proportional to the integrand. We assume the target domain is parameterized by a unit hypercube $\mathcal{U} = [0, 1]^n$ as in the original paper. With perfectly proportional importance sampling, we can formulate the interior term as:

$$\int_{\mathcal{U}} \partial_\theta f(T(u, \theta); \theta) |J(u)| \, du = \int_{\mathcal{U}} \partial_\theta \left[\frac{f(T(u, \theta); \theta)}{p(T(u, \theta); \theta)} \right] du \tag{12}$$

$$= \int_{\mathcal{U}} \partial_\theta \big[g(T(u, \theta); \theta) \big] \, du \tag{13}$$

$$= \int_{\mathcal{U}} \big[\partial_x g(u; \theta) \, \partial_\theta T(u, \theta) + \partial_\theta g(u; \theta) \big] \, du, \tag{14}$$

where the determinant of the Jacobian $|J(u)| = 1/p(x_i; \theta)$ and the integrand f is written as a product of the sampling distribution p and the remaining terms g: $f = p \cdot g$.

This analysis is relevant to our case, as we assume the sampling distribution is perfectly proportional to the integrand. Zeltner et al. [ZSGJ21] observed that the derivative $\partial_\theta T(u, \theta)$ of the transformation T in the first term of Eq. (14) can be highly sensitive to small changes in θ, which can lead to high-variance estimates. Our implementation of T involves the projection operation Π (Eq. (3)), which makes the shape-adaptive BSSRDF sensitive to this is-

sue. Projection can introduce distortion by mapping originally close points to distant surface points, e.g., at corners and micro-detailed surfaces.

Boundary term We note that the boundary term is needed in our case even with static geometry, due to parameter-dependent importance sampling. Figure 2 illustrates this issue. As we do not update the geometry, the spatial position of the visibility boundary $x_B = \mathrm{T}(u_B; \theta)$ does not change. The corresponding boundary u_B in the sample domain, however, changes with parameter θ. We visualize the domain $\mathcal{U}$, which is the latent space of the VAE in our case, as a 2D domain for illustration purpose.

The evaluation of the boundary term requires sampling of points in the boundary $\partial \mathcal{U}$. This requires mapping the visibility boundary in $\mathcal{X}$ back to $\mathcal{U}$. However, as the VAE is not invertible due to non-linear layers (e.g. ReLU activation), transformation T is not invertible. Biased gradients due to the absence of the boundary term are illustrated in Figure 3.

4.2. Offset Sample-based Gradient Estimation

From the above analysis, we identify two main challenges in applying the state-of-the-art differentiable rendering frameworks to the shape-adaptive model: 1) high variance in the interior term and 2) difficulty of sampling the boundary term. Rather than separately modeling each term, we take an orthogonal approach.

We propose to compute gradients by computing the finite difference of *offset* samples. A similar idea has been explored by [KMA*15], where image gradients are approximated by finite differences of adjacent light paths for primal path tracing. Our offset samples are obtained by slightly perturbing θ used to compute x_i while fixing the sample u in the canonical domain:

$$x_i^+ = \mathrm{T}(u; \theta^+), \ x_i^- = \mathrm{T}(u; \theta^-), \qquad (15)$$

where θ^+, θ^- are 1-dimensional parameters perturbed in both directions.

For the extinction coefficient σ_t, its main contribution to the rendering pipeline is determining the scale of scattering σ_n (Eq. (3)). In our setting where $g = 0$, the extinction coefficient σ_t does not affect x_{VAE}, but only the scale σ_n. Hence, we can calculate the offset samples efficiently without evaluating VAE twice using the same samples x_{VAE} for both θ^+ and θ^-:

$$x_i^+ = \Pi(x_o + \sigma_n(\theta^+) x_{\mathrm{VAE}}), \ x_i^- = \Pi(x_o + \sigma_n(\theta^-) x_{\mathrm{VAE}}). \ (16)$$

Finally, the gradient with respect to the parameter θ is approximated by the sum of the differences of offset samples:

$$\partial_\theta L_o \approx \sum_{x_i} \frac{g(x_i^+) - g(x_i^-)}{2\Delta\theta}. \qquad (17)$$

4.3. Gradients of Scattering Albedo α

The scattering albedo α's contribution to the appearance of translucency is different from σ_t in the primal rendering pipeline. α contributes to the absorption probability estimation using the absorption network $Absorb(x_o; \theta_{\mathrm{med}}, \theta_{\mathrm{shape}})$. In the primal rendering phase, the validity of samples is stochastically determined via

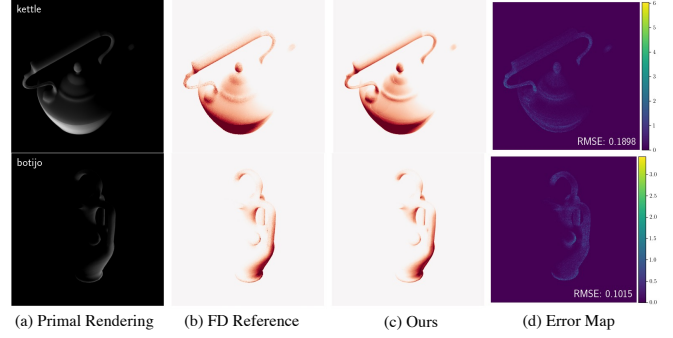(a) Primal Rendering   (b) FD Reference   (c) Ours   (d) Error Map

Figure 4: α *gradients. Reparameterizing the estimated absorption probability as a scale factor allows the differentiation of α using naïve automatic differentiation.*

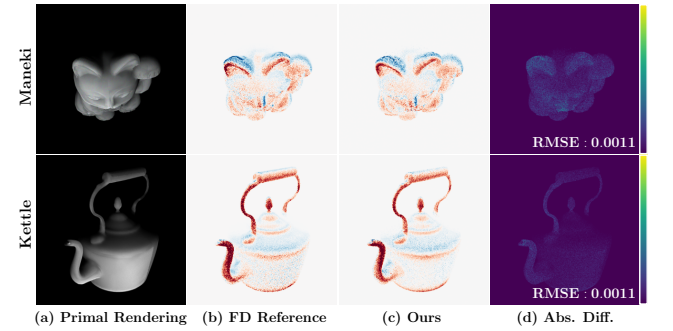(a) Primal Rendering   (b) FD Reference   (c) Ours   (d) Abs. Diff.

Figure 5: *Comparison of image derivatives. We compare image derivatives computed with (b) finite differences (128 spp) and (c) our method (32 spp).*

the estimated absorption probability. For example, light entering a highly absorptive object, hence having high absorption probability, will be less likely to exit and contribute to the pixel intensity.

In the differential rendering phase, we reparameterize the estimated absorption probability as a scale factor $S_{\mathrm{abs}} = 1 - Absorb(x_o; \theta_{\mathrm{med}}, \theta_{\mathrm{shape}})$ applied to each sample (Eq. (5)), rather than using it as termination probability. This design choice serves two purposes. First, α becomes differentiable as it now contributes to the intensity (color) of each sample (Fig. 4). Second, samples are not wasted due to termination.

In summary, while the gradient for the extinction coefficient σ_t is computed with offset samples, the gradient for the scattering albedo α is computed with automatic differentiation of a reparameterized rendering equation.

5. Experiments

We jointly optimize the extinction coefficient σ_t and scattering albedo α in our experiments, unless explicitly stated otherwise. To evaluate the accuracy of our estimations, we measure the Root Mean Square Error (RMSE) between the estimated parameters and their corresponding ground truth values after 300 iterations. All plots for σ_t have their values scaled by a factor of 0.01 for visualization purpose. We use the Adam Optimizer [KB17].
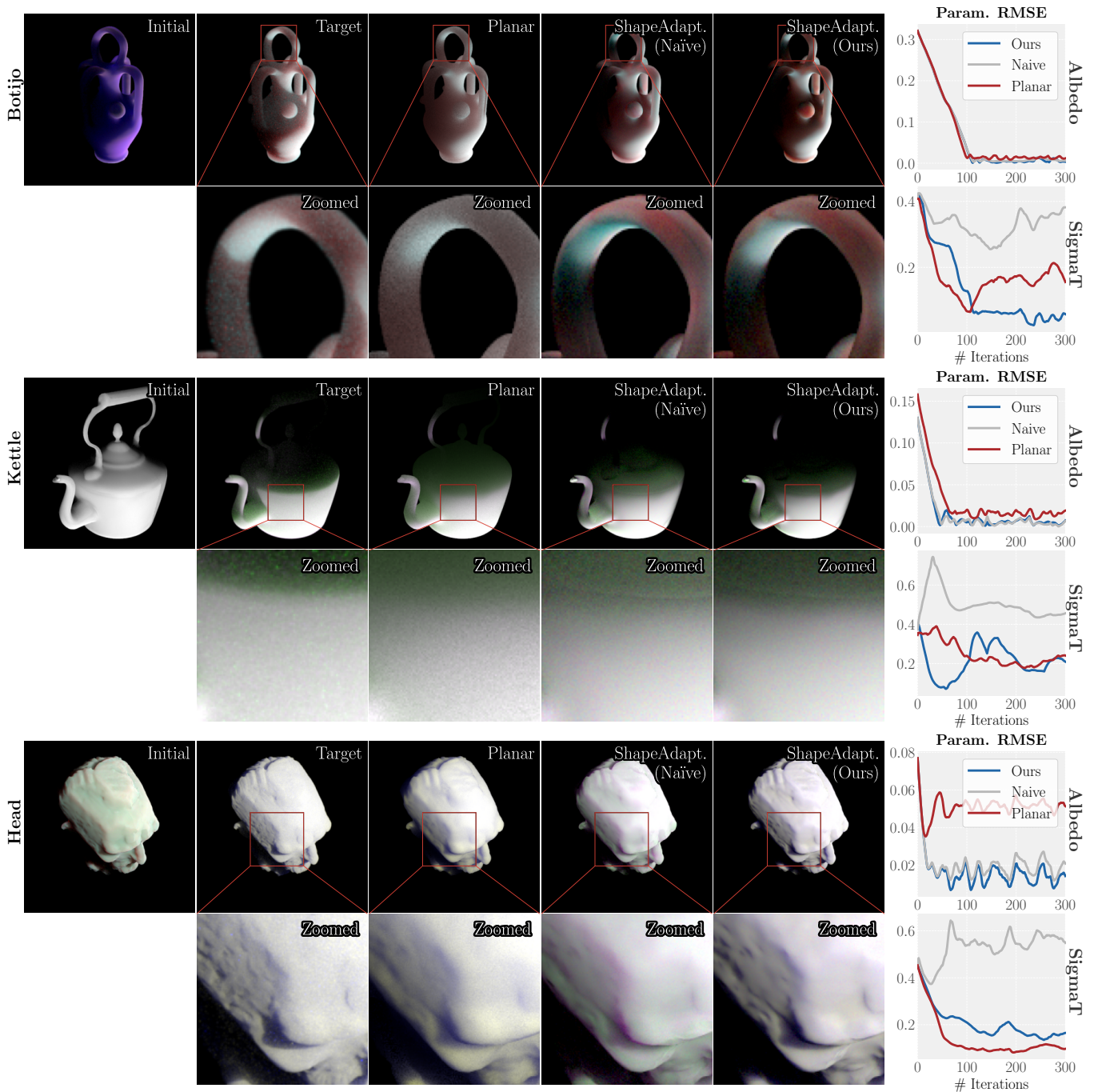
Figure 6: *Shape-adaptive BSSRDF vs. Planar BSSRDF. We optimize the scattering albedo and extinction coefficient using the shape-adaptive BSSRDF and a planar BSSRDF-based method [DLW*22]. The shape-adaptive model ("Naïve") with naïve automatic differentiation fails to converge to correct parameter values. The shape-adaptive model ("Ours") tends to preserve details better (e.g., dents in the Head scene), whereas the planar model ("Planar") tends to generate more blurry results.*

Dataset We render roughly 40 images per object using the volumetric path tracer Mitusba3 [JSR*22]. We use the volpath integrator with max_depth set as 32, and set the BSDF as null material to discard the effect of interreflection. The ground-truth parameter values for the extinction coefficient and scattering albedo are sampled from the range [30,100] and [0.8,0.98], respectively.

Implementation details We utilize automatic differentiation and GPU ray-tracing provided by the differentiable rendering frame-

work PSDR-CUDA [ZMY*20, DLW*22]. We trained the neural networks of the forward model (e.g., VAE, Absorption Network) with a slightly more lightweight architecture than that described in the original paper [VKJ19] for computational efficiency. Details of the modified architecture are in the supplementary. The forward model takes as input reduced medium parameters α' and σ'_t to decouple the effect of the anisotropy parameter g based on similarity theory [WPW89]. In all experiments including ours and compared methods, we set $g = 0$, and hence assume $\sigma_t = \sigma'_t, \alpha = \alpha'$.

Validation In Figure 5, we compare the accuracy of gradients computed with our method to gradients obtained via finite differences. For a more precise comparison, random seeds were fixed when obtaining the FD result, resulting in low noise. Our gradient computation algorithm correctly approximates the FD reference without the need for an explicit discontinuity handling.

Ours vs. Naïve automatic differentiation In Figure 6, we compare our method ("Ours") with naïve automatic differentiation ("Naïve"). The parameter-dependent discontinuities inherent in the shape-adaptive BSSRDF lead to inaccurate gradients (Figure 3), resulting in training divergence. Additional experiments are included in the supplementary material.

Shape-adaptive BSSRDF vs. Planar BSSRDF In Figure 6, we compare our method against an inverse rendering method [DLW*22] using a planar BSSRDF [JMLH01]. Our experiments demonstrate that the shape-adaptive importance sampling framework yields lower parameter error compared to the planar-based model. Qualitatively, our results exhibit sharper details. On the other hand, the projection algorithm accompanied by the planar-BSSRDF model tends to blur out the rendered images.

Robustness to initialization To demonstrate the robustness of our method, we conducted experiments using eight randomly sampled initial values for both scattering parameters, σ_t and α (Figure 7). As optimization proceeds, the shaded area becomes more narrow, indicating stable convergence regardless of initialization.

Shape-adaptive BSSRDF vs. Volumetric rendering In Figure 8, we optimize σ_t using our method and PSDR-VOL [ZYZ21]. While PSDR-VOL is capable of joint shape optimization, we disabled components specific to handling shape discontinuities for fair comparison. The Head scene is monochromatic and α is fixed to 0.9. At 32 spp, our method converges to a value close to the ground truth after 300 iterations. Both at equal (32 spp) and increased spp (128 spp), PSDR-VOL diverges.

In Figure 9, we jointly estimate σ_t and α for each RGB channel using our method and Path Replay Backpropagation (PRB) [VSJ21] implemented in Mitsuba3 [JSR*22]. Because the synthetic data was rendered using the same forward rendering model employed by PRB, it has an advantage in optimization. Still, we observed that our method is capable of converging earlier to a solution close to ground-truth parameters, especially for objects characterized by a low σ_t and α.

Time efficiency In Table 1, we report the average computation time per iteration for the Head scene in Figure 8, measured on a single NVIDIA GeForce RTX 4090 graphics card. The max depth of both PSDR-VOL and PRB were set to 6 during training. Our method demonstrates better time efficiency compared to PSDR-VOL and
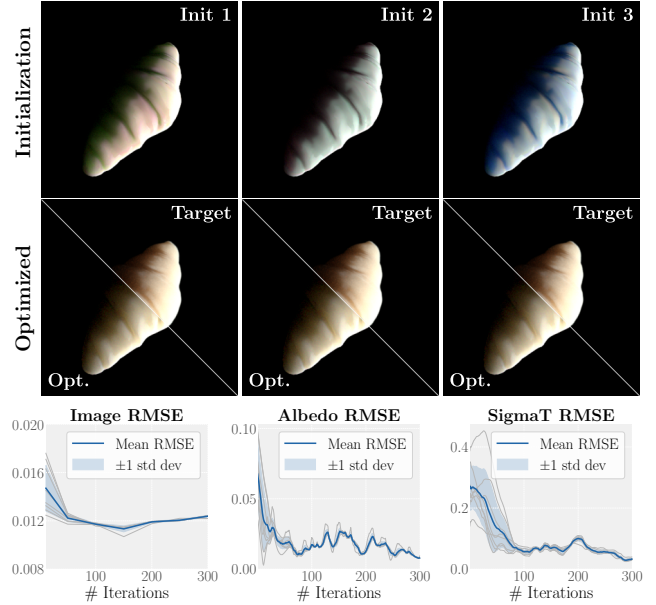

Figure 7: *Robustness to initialization. We conducted experiments using eight randomly sampled initial parameter values on the* Croissant *scene. The top two rows show results of three samples in the experiments. In the bottom row, individual error plots are shown in* gray lines. *The* blue lines *represent the mean across all experiments, while the lightly shaded area indicates ± 1 standard deviation from the mean. As optimization proceeds, the shaded area narrows, reflecting the convergence of the experiments regardless of the initialization.*

PRB, as BSSRDF methods do not require explicit sampling of internal paths within the volume.

Table 1: *Average time per iteration (sec.) of different approaches.*

Ours (32 spp)	Ours (64 spp)	PSDR-VOL (32 spp)	PRB (32 spp)
1.40	2.80	2.28	4.07

6. Discussion

Comparison to boundary integral methods By implicitly accounting for boundary effects, our approach eliminates the need for explicit sampling of discontinuity boundaries required in other methods [ZMY*20, ZYZ21]. Our method computes the finite difference of each offset sample independently, allowing those near boundaries to naturally exhibit values greater in magnitude.

Comparison to finite differences Our approach to differentiating importance sampling involves using offset samples for gradient computation, similar to the finite difference method. However, unlike naïve finite differences, which require multiple sampling passes and introduce noise due to uncorrelated samples, our method shares samples between offset evaluations. This reduces computational cost and yields more stable gradients.
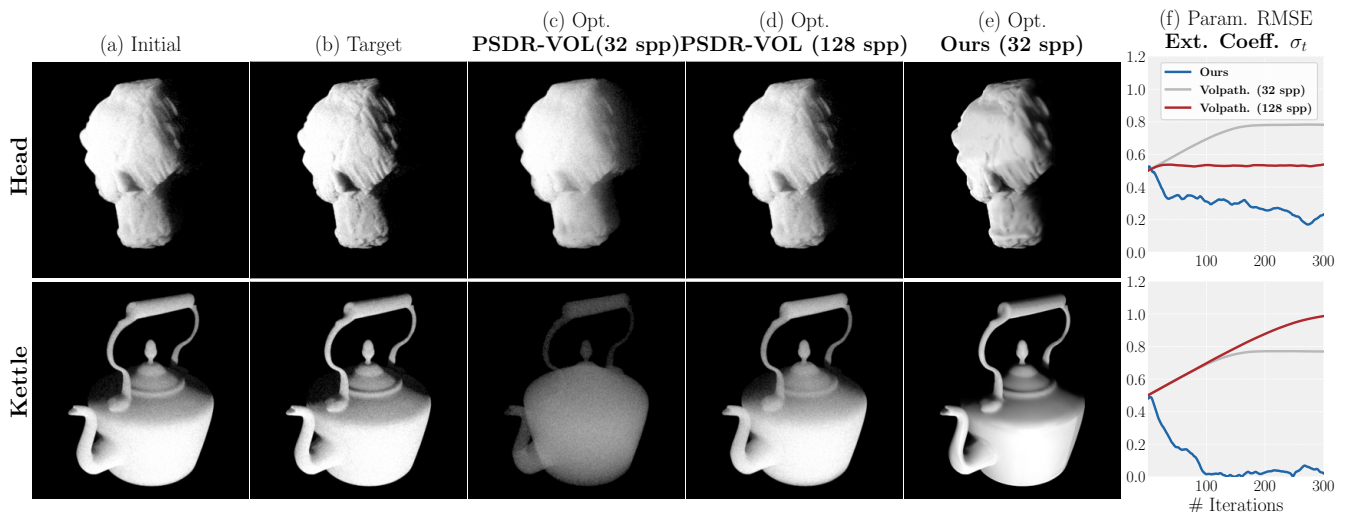
Figure 8: *Comparison of shape-adaptive inverse rendering with PSDR-VOL [ZYZ21]. While the forward model of PSDR-VOL generates more physically accurate renderings, our results indicate that (c, e) the shape-adaptive BSSRDF achieves more accurate reconstructions of the extinction coefficient σ_t when evaluated at equivalent samples per pixel (spp). Notably, (e) the shape-adaptive BSSRDF demonstrates convergence at a low spp (32 spp), whereas (d) PSDR-VOL fails to converge even when utilizing a higher number of samples (128 spp).*
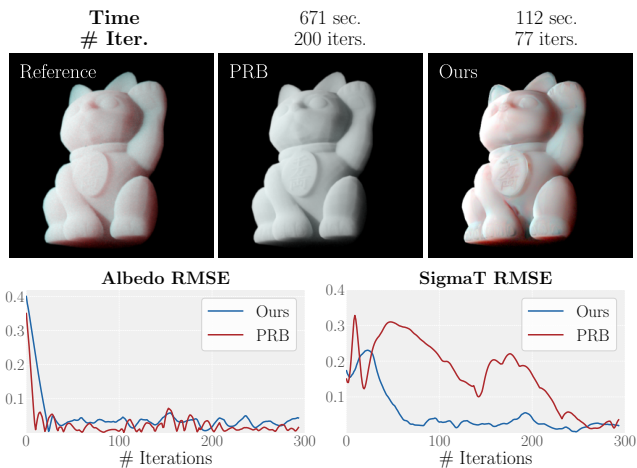

Figure 9: *Comparison of shape-adaptive inverse rendering with PRB [VSJ21]. We jointly optimize σ_t and α using our method and PRB at 32 spp. Despite our method employing an approximate forward model, our method converges in fewer iterations. The plots show that our model converges in less than 80 iterations, nearly six times faster than PRB, which requires more than 200 iterations.*

7. Conclusion

In this paper, we presented a practical inverse rendering method for 3D objects with homogeneous subsurface scattering materials by developing a differentiable shape-adaptive BSSRDF model. Our analysis identified two primary challenges in differentiating the shape-adaptive importance sampling framework: (1) difficulty of computing derivatives involving the importance sampling operation and (2) the stochastic behavior of the absorption probability. We addressed the first issues by developing a robust finite difference-based approach using correlated offset samples for extinction coefficient σ_t. The second issue is handled by reformulating the effect of absorption in the rendering process from Russian Roulette path termination to a differentiable scale factor, facilitating naïve automatic differentiation for scattering albedo α. By combining finite differences (σ_t) with automatic differentiation (α), our approach achieves reliable reconstruction, even with a relatively small number of samples.

Limitations & Future work Our method is well-suited for simple problem settings where the optimized parameters are low-dimensional. Extending the finite difference-based approach to more general settings, which involve more complex light paths or a larger set of optimized parameters, provides a promising direction for future research. For instance, we anticipate that heterogeneity can be approximated by textured spatial variation without increasing the complexity of the finite difference method. Extending the finite difference-based approach to a general importance sampling framework or exploring a different sampling model architecture (e.g., Normalizing Flows) would be an another interesting future research direction.

References

[BLD20] BANGARU S. P., LI T.-M., DURAND F.: Unbiased warped-area sampling for differentiable rendering. *ACM Transactions on Graphics (TOG) 39*, 6 (2020), 1–18. 2

[Chr15] CHRISTENSEN P. H.: An approximate reflectance profile for efficient subsurface scattering. In *ACM SIGGRAPH 2015 Talks*. 2015, pp. 1–1. 2

[CLZ*20] CHE C., LUAN F., ZHAO S., BALA K., GKIOULEKAS I.: Towards learning-based inverse subsurface scattering. In *2020 IEEE International Conference on Computational Photography (ICCP)* (2020), IEEE, pp. 1–12. 1, 3

[DJ05] DONNER C., JENSEN H. W.: Light diffusion in multi-layered translucent materials. *ACM Transactions on Graphics (ToG) 24*, 3 (2005), 1032–1039. 2

[DLR*09] DONNER C., LAWRENCE J., RAMAMOORTHI R., HACHISUKA T., JENSEN H. W., NAYAR S.: An empirical bssrdf model. In *ACM SIGGRAPH 2009 papers*. 2009, pp. 1–10. 2

[DLW*22] DENG X., LUAN F., WALTER B., BALA K., MARSCHNER S.: Reconstructing translucent objects using differentiable rendering. In *ACM SIGGRAPH 2022 Conference Proceedings* (2022), pp. 1–10. 3, 6, 7

[GCM*18] GENOVA K., COLE F., MASCHINOT A., SARNA A., VLASIC D., FREEMAN W. T.: Unsupervised training for 3d morphable model regression. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 8377–8386. 2

[GZB*13] GKIOULEKAS I., ZHAO S., BALA K., ZICKLER T., LEVIN A.: Inverse volume rendering with material dictionaries. *ACM Transactions on Graphics (TOG) 32*, 6 (2013), 1–13. 1, 3

[HCJ13] HABEL R., CHRISTENSEN P. H., JAROSZ W.: Photon beam diffusion: A hybrid monte carlo method for subsurface scattering. In *Computer Graphics Forum* (2013), vol. 32, Wiley Online Library, pp. 27–37. 2

[HDCD15] HEITZ E., DUPUY J., CRASSIN C., DACHSBACHER C.: The sggx microflake distribution. *ACM Transactions on Graphics (TOG) 34*, 4 (2015), 1–11. 3

[JAM*10] JAKOB W., ARBREE A., MOON J. T., BALA K., MARSCHNER S.: A radiative transfer framework for rendering materials with anisotropic structure. In *ACM SIGGRAPH 2010 papers*. 2010, pp. 1–13. 3

[JLX*23] JIN H., LIU I., XU P., ZHANG X., HAN S., BI S., ZHOU X., XU Z., SU H.: Tensoir: Tensorial inverse rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2023), pp. 165–174. 1, 2

[JMLH01] JENSEN H. W., MARSCHNER S. R., LEVOY M., HANRAHAN P.: A practical model for subsurface light transport. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques* (2001), pp. 511–518. 2, 3, 7

[JSR*22] JAKOB W., SPEIERER S., ROUSSEL N., NIMIER-DAVID M., VICINI D., ZELTNER T., NICOLET B., CRESPO M., LEROY V., ZHANG Z.: Mitsuba 3 renderer, 2022. https://mitsuba-renderer.org. 6, 7

[KB17] KINGMA D. P., BA J.: Adam: A method for stochastic optimization, 2017. URL: https://arxiv.org/abs/1412.6980, arXiv:1412.6980. 5

[Kd14] KŘIVÁNEK J., D'EON E.: A zero-variance-based sampling scheme for monte carlo subsurface scattering. In *ACM SIGGRAPH 2014 Talks*. 2014, pp. 1–1. 2

[KMA*15] KETTUNEN M., MANZI M., AITTALA M., LEHTINEN J., DURAND F., ZWICKER M.: Gradient-domain path tracing. *ACM Transactions on Graphics (TOG) 34*, 4 (2015), 1–13. 5

[KUH18] KATO H., USHIKU Y., HARADA T.: Neural 3d mesh renderer. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2018), pp. 3907–3916. 2

[KVH84] KAJIYA J. T., VON HERZEN B. P.: Ray tracing volume densities. *ACM SIGGRAPH computer graphics 18*, 3 (1984), 165–174. 2

[KW13] KINGMA D. P., WELLING M.: Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013). 2

[LADL18] LI T.-M., AITTALA M., DURAND F., LEHTINEN J.: Differentiable monte carlo ray tracing through edge sampling. *ACM Transactions on Graphics (TOG) 37*, 6 (2018), 1–11. 2

[LB14] LOPER M. M., BLACK M. J.: Opendr: An approximate differentiable renderer. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VII 13* (2014), Springer, pp. 154–169. 2

[LHJ19] LOUBET G., HOLZSCHUCH N., JAKOB W.: Reparameterizing discontinuous integrands for differentiable rendering. *ACM Transactions on Graphics (TOG) 38*, 6 (2019), 1–14. 2

[LLCL19] LIU S., LI T., CHEN W., LI H.: Soft rasterizer: A differentiable renderer for image-based 3d reasoning. In *Proceedings of the IEEE/CVF international conference on computer vision* (2019), pp. 7708–7717. 2

[MES*11] MUNOZ A., ECHEVARRIA J. I., SERON F. J., LOPEZ-MORENO J., GLENCROSS M., GUTIERREZ D.: Bssrdf estimation from single images. In *Computer Graphics Forum* (2011), vol. 30, Wiley Online Library, pp. 455–464. 3

[MHD16] MENG J., HANIKA J., DACHSBACHER C.: Improving the dwivedi sampling scheme. In *Computer Graphics Forum* (2016), vol. 35, Wiley Online Library, pp. 37–44. 2

[NDSRJ20] NIMIER-DAVID M., SPEIERER S., RUIZ B., JAKOB W.: Radiative backpropagation: An adjoint method for lightning-fast differentiable rendering. *ACM Transactions on Graphics (TOG) 39*, 4 (2020), 146–1. 3

[NDVZJ19] NIMIER-DAVID M., VICINI D., ZELTNER T., JAKOB W.: Mitsuba 2: A retargetable forward and inverse renderer. *ACM Transactions on Graphics (TOG) 38*, 6 (2019), 1–17. 2

[NGHJ18] NOVÁK J., GEORGIEV I., HANIKA J., JAROSZ W.: Monte carlo methods for volumetric light transport simulation. In *Computer graphics forum* (2018), vol. 37, Wiley Online Library, pp. 551–576. 3

[Pho75] PHONG B. T.: Illumination for computer generated pictures. *Communications of the ACM 18*, 6 (1975), 311–317. 2

[Rus88] RUSHMEIER H. E.: *Realistic image synthesis for scenes with radiatively participating media.* Cornell University, 1988. 2

[TSG19] TSAI C.-Y., SANKARANARAYANAN A. C., GKIOULEKAS I.: Beyond volumetric albedo–a surface optimization framework for non-line-of-sight imaging. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (2019), pp. 1545–1555. 2

[VKJ19] VICINI D., KOLTUN V., JAKOB W.: A learned shape-adaptive subsurface scattering model. *ACM Transactions on Graphics (TOG) 38*, 4 (2019), 1–15. 1, 2, 3, 7

[VSJ21] VICINI D., SPEIERER S., JAKOB W.: Path replay backpropagation: Differentiating light paths using constant memory and linear time. *ACM Transactions on Graphics (TOG) 40*, 4 (2021), 1–14. 3, 7, 8

[WMLT07] WALTER B., MARSCHNER S. R., LI H., TORRANCE K. E.: Microfacet models for refraction through rough surfaces. In *Proceedings of the 18th Eurographics conference on Rendering Techniques* (2007), pp. 195–206. 2

[WPW89] WYMAN D. R., PATTERSON M. S., WILSON B. C.: Similarity relations for the interaction parameters in radiation transport. *Applied optics 28*, 24 (1989), 5243–5249. 7

[XBLZ23] XU P., BANGARU S., LI T.-M., ZHAO S.: Warped-area reparameterization of differential path integrals. *ACM Transactions on Graphics (TOG) 42*, 6 (2023), 1–18.

[YLB*22] YAN K., LASSNER C., BUDGE B., DONG Z., ZHAO S.: Efficient estimation of boundary integrals for path-space differentiable rendering. *ACM Transactions on Graphics (TOG) 41*, 4 (2022), 1–13. 2

[ZIK*17] ZENG X., IWAGUCHI T., KUBO H., FUNATOMI T., MUKAIGAWA Y.: Estimating parameters of subsurface scattering using directional dipole model. In *2017 Nicograph International (NicoInt)* (2017), IEEE, pp. 41–48. 3

[ZLLS22] ZHANG K., LUAN F., LI Z., SNAVELY N.: Iron: Inverse rendering by optimizing neural sdfs and materials from photometric images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2022), pp. 5565–5574. 1, 2

[ZMY*20] ZHANG C., MILLER B., YAN K., GKIOULEKAS I., ZHAO S.: Path-space differentiable rendering. *ACM Transactions on Graphics (TOG) 39*, 4 (2020), 143–1. 2, 3, 4, 7

[ZSD*21] ZHANG X., SRINIVASAN P. P., DENG B., DEBEVEC P., FREEMAN W. T., BARRON J. T.: Nerfactor: Neural factorization of shape and reflectance under an unknown illumination. *ACM Transactions on Graphics (ToG) 40*, 6 (2021), 1–18. 1, 2

[ZSGJ21] ZELTNER T., SPEIERER S., GEORGIEV I., JAKOB W.: Monte carlo estimators for differential light transport. *ACM Transactions on Graphics (TOG) 40*, 4 (2021), 1–16. 4

[ZSS21] ZHENG Q., SINGH G., SEIDEL H.-P.: Neural relightable participating media rendering. *Advances in Neural Information Processing Systems 34* (2021), 15203–15215. 3

[ZWZ*19] ZHANG C., WU L., ZHENG C., GKIOULEKAS I., RAMAMOORTHI R., ZHAO S.: A differential theory of radiative transfer. *ACM Transactions on Graphics (TOG) 38*, 6 (2019), 1–16. 2

[ZYZ21] ZHANG C., YU Z., ZHAO S.: Path-space differentiable rendering of participating media. *ACM Transactions on Graphics (TOG) 40*, 4 (2021), 1–15. 1, 2, 3, 4, 7, 8