Supplementary Material: 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

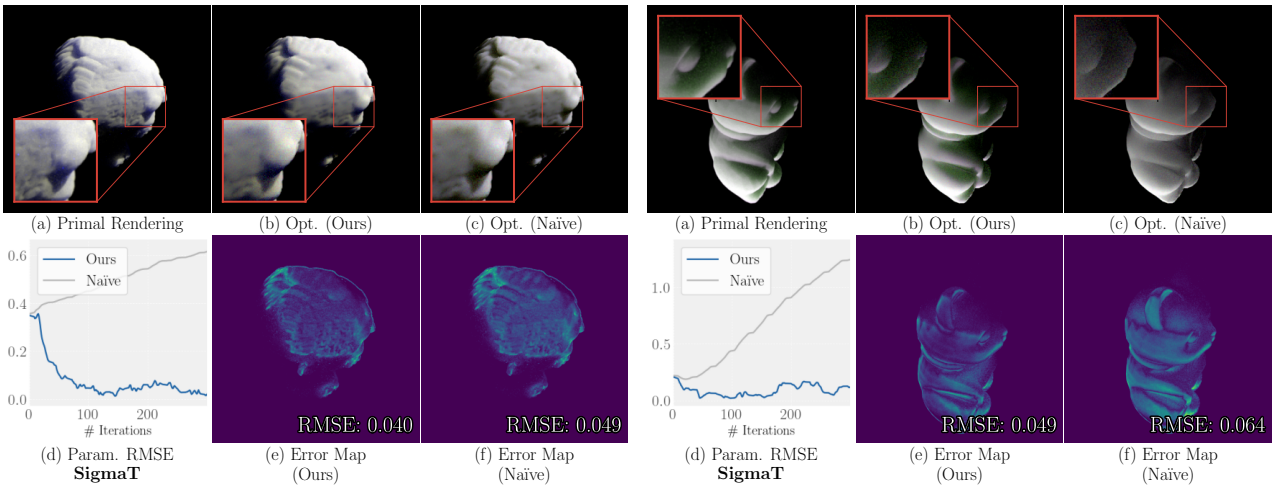

Figure 1: *Ours vs. Naïve autodiff. We compare optimization results using our offset-based gradient computation against those obtained with naïve automatic differentiation (autodiff). Naïve autodiff fails due to biased gradients caused by parameter-dependent discontinuities in the importance sampling framework. Consequently, (a,c) it is unable to accurately learn the basic tint of the object and often continues to (d) update parameters in erroneous directions.*

1. Details of Shape-adaptive BSSRDF

We describe more details of the shape-adaptive BSSRDF [VKJ19] here. Given a point x_o on the surface, the outgoing radiance L_o is computed with Monte Carlo integration of samples x_i generated via importance sampling. Medium parameters and polynomial shape descriptors $\theta = (\theta_{\text{med}}, \theta_{\text{shape}})$ are processed and fed into a VAE, which transforms a random variable $z \sim \mathcal{N}(0, \mathbf{I})$ into an exit position in the canonical space.

$$x_{VAE}(\theta) = VAE(z \mid \theta) \tag{1}$$

Positions x_{VAE} are then scaled by the standard deviation σ_n of its neighborhood, which is a function of medium parameters.

$$\sigma_n(\theta_{med}) = \sigma_n(\alpha, \sigma_t, g) = 2(\frac{1}{4}g + \frac{1}{4}\alpha + \alpha_{eff}(\alpha))/\sigma_t, \tag{2}$$

where

$$\alpha_{eff}(\alpha) = 1 - \frac{1}{8}\log(e^8 + \alpha(1 - e^8)). \tag{3}$$

In the calculation of shape-adaptive BSSRDF, the term S_d is implemented as

$$S_d(\omega_o) = F_t(\theta)/\pi, \tag{4}$$

where $F_t(\theta)$ is the Fresnel transmission [JMLH01].

Model architecture We modified the architecture of the neural networks used in the forward model (e.g., the Feature, Absorption, and Scatter Networks) to improve the rendering speed. The modifications are as follows: 1) The widths of the fully-connected layers were halved (from 64 to 32) and 2) The depths of both the scattering and feature networks were reduced by one layer. All other training details (e.g., dataset, training loss) remain identical to those described in the original paper [VKJ19]. We observed no significant difference in rendering quality with the compressed model.

2. Additional Experiments

Ours vs. Naïve autodiff. In Figure 1, we present additional results for the comparison between naïve automatic differentiation and our offset-bsed gradients. Our method achieves accurate inverse rendering by properly handling gradients from parameter-dependent discontinuities.

3. Experimental Setting Details

Implementation details The forward rendering algorithm (e.g., projection algorithm) was implemented as closely as possible to the original paper [VKJ19]. Polynomial descriptors are fit only once before training starts, and not updated afterwards. The step size $\Delta\theta$ in computing the σ_t gradient is set to 1.0.

Figure 5 in the main paper We test on two monochromatic objects ($\sigma_t = 50.0, \alpha = 0.90$) and compute image derivatives with respect to σ_t. For this experiment, the step size $\Delta\theta$ used for finite differences and our method is both set to 5.

Figure 8 in the main paper Initial parameter values are $\sigma_t = 50.0$, $\alpha = 0.9$, where the scattering albedo α is fixed throughout training. The target is $\sigma_t = 100.0$.

Figure 9 in the main paper Initial parameter values are $\sigma_t = 50.0$, $\alpha = 0.5$ for all RGB channels. Target values are $\sigma_t = (20.0, 50.0, 50.0)$, $\alpha = (0.9, 0.9, 0.9)$. The optimized image results are rendered using 512 spp for all methods. For the Path Replay Backpropagation (PRB) [VSJ21] experiments, we experimented with four different learning rates. The plots show the result of the experiment with fastest convergence, where the learning rate is 0.05.

References

[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. 1

[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

[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. 2