WildLight: In-the-wild Inverse Rendering with a Flashlight (Appendix)

Ziang Cheng, Junxuan Li, Hongdong Li
Australian National University
{ziang.cheng,junxuan.li,hongdong.li}@anu.edu.au

1. Principled BRDF formulation

Our intrinsic network outputs a spacial distribution of texture elements (textels) Θ that parameterize the BRDF at that location. Since we are only interested in recovering opaque objects, we use a submodel of the original paper [1] without the refractive glass lobe. This parameterization can be expressed as:

$$\Theta = \{base\_color, roughness, clearcoat\_glossiness \\ , subsurface, metallic, dieletric, clearcoat\} \in [0,1]^9 \tag{1}$$

where $base_color \in [0,1]^3$ is a 3D vector that defines the RGB base color of material, and all other terms are scalars in range $[0,1]$. Since the camera and light angle are always aligned in a co-located setup, hereafter we denote them by a single direction $\mathbf{h}$.

The model is a linear combination of diffuse and specular lobes, defined as follows:

$$\begin{aligned}\rho(\mathbf{n},\mathbf{h};\Theta) =& (1-metallic)\rho_{\text{diffuse}}(\mathbf{n},\mathbf{h};\Theta)+ \\ & metallic\rho_{\text{metallic}}(\mathbf{n},\mathbf{h};\Theta)+ \\ & 0.08\times(1-metallic)dieletric\rho_{\text{dieletric}}(\mathbf{n},\mathbf{h};\Theta)+ \\ & 0.25\times clearcoat\rho_{\text{clearcoat}}(\mathbf{n},\mathbf{h};\Theta)\end{aligned} \tag{2}$$

The diffuse component have two lobes: one for base diffuse and one for subsurface scattering

$$\rho_{\text{diffuse}} = (1-subsurface)\rho_{\text{base\_diffuse}} + subsurface\rho_{\text{subsurface}} \tag{3}$$

where the two diffuse lobes are defined as

$$\rho_{\text{base\_diffuse}} = \frac{base\_color}{\pi}\Big(1+(2roughness-0.5)(1-\mathbf{n}^\top\mathbf{h})^5\Big)^2 \tag{4}$$

and $\rho_{\text{subsurface}}$ is defined as

$$\frac{1.25base\_color}{2\pi}\Big(\big(1+(roughness-1)(1-\mathbf{n}^\top\mathbf{h})^5\big)^2(\frac{1}{\mathbf{n}^\top\mathbf{h}}-1)+1\Big) \tag{5}$$

respectively.

The metallic and dielectric lobes share the same GGX distribution, and both can be factorized into three terms:

$$\rho_{\text{specular}} = \frac{DGF}{4(\mathbf{n}^\top\mathbf{h})^2} \tag{6}$$

where D and G are the microsurface distribution function and mask-shadowing term, respectively:

$$D = \frac{roughness^4}{\pi\Big((roughness^4-1)(\mathbf{n}^\top\mathbf{h})^2+1\Big)^2} \tag{7}$$

and

$$G = \frac{2}{\sqrt{1+\frac{roughness^4(1-(\mathbf{n}^\top\mathbf{h})^2)}{(\mathbf{n}^\top\mathbf{h})^2}}+1} \tag{8}$$

The only place where metallic and dielectric lobes differ is the Fresnel term F: the metallic lobe is chromatic while the dielectric lobe is not.

$$F_{\text{metallic}} = base\_color \tag{9}$$

$$F_{\text{dieletric}} = 1 \tag{10}$$

The clearcoat term can also be factorized similar to before

$$\rho_{\text{clearcoat}} = \frac{D_cG_cF_c}{4(\mathbf{n}^\top\mathbf{h})^2} \tag{11}$$

where

$$D_c = \frac{roughness_c^2-1}{2\pi\log(roughness_c)(1+(roughness_c^2-1)(\mathbf{n}^\top\mathbf{h})^2)} \tag{12}$$

$$G_c = \frac{2}{\sqrt{1+\frac{0.25^2(1-(\mathbf{n}^\top\mathbf{h})^2)}{(\mathbf{n}^\top\mathbf{h})^2}}+1} \tag{13}$$

$$F_c = 0.2 \tag{14}$$

Here the clearcoat roughness $roughness_c$ is confined within range $[0.001, 0.1]$ and is linearly parameterized by $clearcoat\_glossiness$

$$roughness_c = 0.1 - 0.099clearcoat\_glossiness. \tag{15}$$

2. Surface-to-surface distance for geometry evaluation

We used a commutative mesh-to-mesh distance for evaluating the surface geometry produced from different method in Table 1. This distance metric is defined as follows:

$$D(S_1, S_2) = \frac{1}{2}\,\mathbb{E}\left(d(\mathbf{x}_1, S_2) + d(\mathbf{x}_2, S_1)\right) \tag{16}$$

where d is the point-to-manifold Euclidean distance, S_1 and S_2 are the visible surface regions, and $\mathbf{x}_1$ and $\mathbf{x}_2$ are two mutually independent random points uniformly sampled from S_1 and S_2 respectively. We compute the mean distance by repeatedly sampling $\mathbf{x}_1, \mathbf{x}_2$ until the standard variation of sampling mean is no greater than 10^{-5}, or until an excessive number of one million pairs of points have been sampled. For the median distance, we replace the expectation operator $\mathbb{E}$ with the sample median.

References

[1] Brent Burley. Physically-based shading at disney. In *ACM SIGGRAPH Course Notes. Practical physically-based shading in film and game production.*, volume 2012, pages 1–7, 2012. 1