

A Data-Driven Light Scattering Model for Hair

Leonid Pekelis*, Christophe Hery†, Ryusuke Villemin‡, Junyi Ling§

Pixar Technical Memo #15-02



Figure 1: A series of snap-shot renders from importance sampling our hair reflectance functions. Our model supports efficient and seamless rendering of fibers with elliptical cross section. This is key to producing time-homogeneous, realistic caustics. We produced these images by rotating an outdoor IBL gathered in midday, and using a ray tracer with multiple importance sampling. Each image is approximately 64 light samples and 256 brdf samples per pixel.

Abstract

We present an implementation of the [Marschner et al. 2003] model for importance sampling light reflected from hair. The implementation makes use of a version of Adaptive Importance Sampling (AIS), specialized to fit easily sampled distributions to BCSDFs. Our model is novel among importance sampling implementations in that it includes all the features of Marschner such as eccentricity for elliptical cross-sections, and extends them by adding azimuthal roughness control, and natural fiber torsion. It is also fully energy preserving. We compare with the implementations in both [d’Eon et al. 2013] and [Ou et al. 2012] as well as a ground truth, physical model which directly evaluates light-cylinder interactions. Our model well approximates the ground truth, and is significantly faster than other implementations.

Keywords: hair, fibers, light scattering, rendering, importance sampling, density estimation

*Stanford University - Department of Statistics, lpekelis@stanford.edu

†Pixar Animation Studios, chery@pixar.com

‡Pixar Animation Studios, rvillemin@pixar.com

§Pixar Animation Studios, jling@pixar.com

1 Introduction

A quarter-century after the pioneering work of [Kajiya and Kay 1989], rendering hair and fur in 3D animation remains a challenge.

Hair rendering is computationally expensive. While a uniform structure is apparent, minute differences in between fibers are what make the overall picture convincing, especially in today’s ever increasing demand for rendering fidelity, such as in cinema. This requires fiber level detail which can compound render time, as it is standard to have millions of fibers on a single groom.

Furthermore, each hair can be physically demanding. As a dielectric with eccentric cross-section, hair allows for the possibility of many internal bounces of light refraction and reflection. The work of [Marschner et al. 2003] is famous for describing the geometric equations governing these phenomena for circular cross-sections, as well as showing that they are *not* enough to reproduce the characteristic look of hair. It is the elliptic shape which causes the desired caustics, as seen in figure 1, and also complicates the equations.

We present a hair scattering model and importance sampling scheme that incorporates features such as elliptic cross-sections and minor-axis rotation, which help give the hair its distinctive look, while still optimizing for fast computation times. The key is a data-centric optimization scheme which fits easily sampled distributions to physically complicated scattering behavior. The result is an efficient model, capable of rendering physically realistic hair.

While specified to rendering hair, our general approach - fitting complicated BRDFs through Adaptive Importance Sampling (AIS) - may be useful in other areas of shading as well.

2 Background and Related Work

In this section we review the Marschner model, set-up mathematical notation, as well as describe related work. For an excellent, comprehensive survey of hair modeling see [Ward et al. 2007].

2.1 Scattering theory: the Marschner model

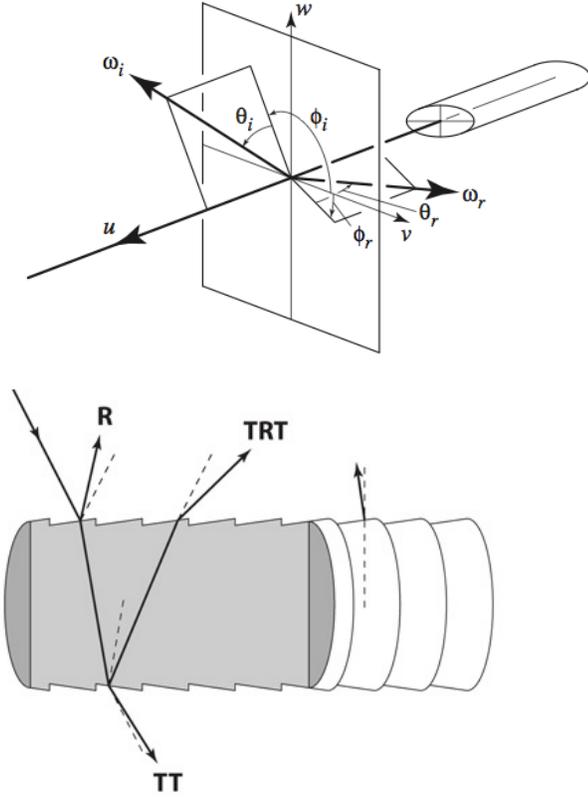


Figure 2: The coordinate system and representation of first 3 internal bounces. From [Marschner et al. 2003].

Our scattering model is built on the foundation of [Marschner et al. 2003]. In order to properly treat hair torsion (section 4.2), we extend their spherical coordinate system to be both locally well-defined and globally consistent. Figure 2 from [Marschner et al. 2003] is a starting diagram. The following section reviews their approach. Our extensions are noted where relevant.

Three orthonormal vectors form a right handed basis: u - the tangent vector, v - the major axis, bisecting the thickest cross-section, and w - the minor axis orthogonal to the others. We distinguish v from w as scattering functions produced by elliptical hair fibers are not rotationally invariant. Global consistency follows from defining both u and v from hair geometry, and requiring v to be a smoothly varying function of its distance from the root of the hair fiber.

Given an incoming light direction, ω_i , we want to find an outgoing, or reflected, direction, ω_r . Both directions are described by their longitudinal and azimuthal angles, $(\theta, \phi) \in [-\pi/2, \pi/2] \times [-\pi, \pi]$. The scattering function is defined as differential radiance over differential energy and described by a sum over all internal bounces resulting in the emission of a light ray - (R)eflection, (T)ransmission(T)ransmission, TRT, TRRT, and so forth,

$$S(\omega_i, \omega_r) = \frac{dL_r(\omega_r)}{dL_i(\omega_i)} = \sum_{\rho \in \{R, TT, TRT, \dots\}} S_\rho(\theta_i, \phi_i, \theta_r, \phi_r).$$

Any terms after the first three are disregarded as higher bounces are assumed to be largely attenuated and thus their contribution to total energy negligible. Note that [d'Eon et al. 2011] has since shown that for light colored hair, there is non-negligible energy, particularly at grazing angles in higher terms. Extending the approach in section 3.2 to higher bounces is straightforward, and we leave this for future work.

The third term is separated into two further terms, TRT and $GLINT$, treating the singularities - points of infinite differential radiance - caused by three bounces separately. This allows attenuation to be treated differently for either term. Hence $\rho \in \{R, TT, TRT, GLINT\}$.

Each S_ρ term is partitioned into the product of a marginal, longitudinal scattering function (M) and a conditional, azimuthal scattering function (N). Both products are further conditioned on a vector of user-defined parameters, β . For example,

$$S_R(\theta_i, \phi_i, \theta_r, \phi_r) = M_R(\theta_i | \theta_r, \phi_r, \beta) N_R(\phi_i | \theta_i, \theta_r, \phi_r, \beta) \quad (1)$$

A list of the user defined parameters contained in β is available in the Appendix as Table 2.

The BRDF, S , is pre-multiplied by Fresnel attenuation and volume absorption, which we represent by a single term, $A_\rho(\theta_i, \phi_i, \theta_r, \phi_r, \beta)$, and then integrated over all incoming light directions to calculate outgoing energy,

$$L_r(\omega_r) = \int L_i(\omega) \tilde{S}(\omega, \omega_r | \beta) V(\omega) \cos(\theta) / \cos^2(\theta_r) d\omega$$

$$\tilde{S}(\omega, \omega_r | \beta) = \sum_{\rho} A_\rho(\omega, \omega_r, \beta) M_\rho(\theta | \omega_r, \beta) N_\rho(\phi | \theta, \omega_r, \beta) \quad (2)$$

In Equation 2, V is a visibility function, and the \cos terms account for solid angle change of measure.

Additional notation needed from the Marschner model is $\theta_d = (\theta_r - \theta_i)/2$, $\phi = \phi_r - \phi_i$, $\theta_h = (\theta_i + \theta_r)/2$, and e as the ratio of minor to major axis fiber diameters. All units for the sequel are in centimeters (cm).

2.2 Importance sampling of scattering functions

A standard approach to solve (2) shared by most contemporary hair rendering models including [Hery and Ramamoorthi 2012], [Ou et al. 2012], and [d'Eon et al. 2013], is through importance sampling. If we denote the entire integrand by $p(\omega)$, and we sample incoming light directions by $\omega_j \sim q(\omega)$, $j = 1, \dots, n$, then the importance sampled estimate is

$$L_r(\omega_r) = \sum_{j=1}^n p(\omega_j) / q(\omega_j).$$

As $\tilde{S}(\omega, \omega_r | \beta)$ is represented by a linear sum of terms, it is easiest to importance sample each term separately and then sum the resulting estimates.

Efficient importance sampling of scattering functions helps to reduce computation time by concentrating sampled rays on high energy areas. In other words, $q(\omega) \approx p(\omega)$. Our model takes a

related view from [Hery and Ramamoorthi 2012] which assumes $p(\omega) = f(\omega)q(\omega)$ where f denotes scene specific quantities, leaving the estimate as $\sum_j f(\omega_j)$ now with optimal uniform importance weights, and pushing errors to model approximations.

Searching for less intensive sampling alternatives, such as [Ou et al. 2012] found for Gaussian lobes, further increases efficiency by reducing the cost of sampling. This idea is explored further, along with tradeoffs, in section 3.1.

Finally, scattering functions should be energy conserving. Without this property, changing model parameters can have unintended and unforeseen effects. For example, energy loss at grazing angles for the original formulation of [Marschner et al. 2003] was demonstrated in [d’Eon et al. 2011]. We apply a white furnace test to all our lobes to directly observe preservation of energy. This test has been rigorously justified for microfacet BRDFs in [Heitz 2014].

2.3 Physical properties, simplifying approximations

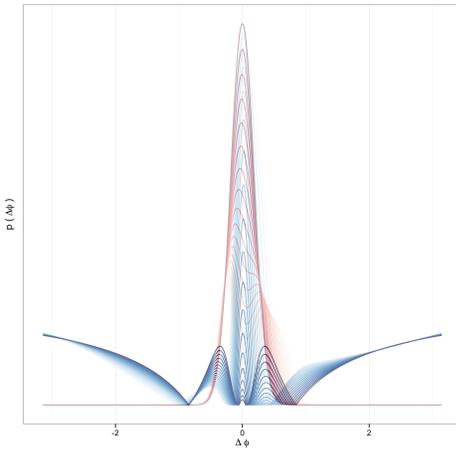


Figure 3: Scattering functions of [Marschner et al. 2003] reacting to varying eccentricity. N_{TRT} is in blue, and N_{GLINT} is in red. Darker colors denote more circular cross-sections $e \rightarrow 1$, and are characterized by more symmetry.

Research into biological and physical properties of hair fibers, usually with application in cosmetics, has helped to produce more accurate hair renders. In [Robbins 2012] hair fiber shape is described as depending on the cross-sectional distribution of different types of cortical cells, causing some hair types to both bow out in an ellipse and curl about their tangent axis. Elliptical cross-sections can have large impacts on the resulting scattering function. Figure 3 shows terms from the scattering function of [Marschner et al. 2003] changing drastically under varying amounts of eccentricity.

Yet as long as computing time is a scarce resource, a painter’s illusion [Kajiya and Kay 1989] - implied detail higher than actual resolution - is desired over potentially costly modeling of individual fibers. This often leads to approximation and simplification such as in [Goldman 1997], [Sadeghi et al. 2010]. These models tend to experience runtime reductions and have less numerical instability, but also may unduly lose accuracy due to subjective choices of functional forms.

Our approach is to first fully define the physical properties of light scattering on hair, and then rely on numerical algorithms, such as AIS (section 5.1), to solve explicit objective functions which retain the most discernible features of light scattering, while reducing

complexity to a space of fast, and stable distributions.

We note that [d’Eon et al. 2013] has a sampling scheme that circumvents the need for any approximation of the Marschner model, but their approach does not handle extension to elliptical fibers, and can be very costly (see section 7).

The rest of the paper proceeds as follows. Section 4 describes noteworthy features of our model. Section 3 gives details of our scattering functions. A description of Adaptive Importance Sampling, our approach to fitting BRDFs, is given in section 5.1. Section 6 gives implementation details. We compare our model to two others from the literature, as well as a path traced ground truth, in section 7, and conclude in section 8.

3 Scattering Functions

3.1 Longitudinal scattering, M_ρ

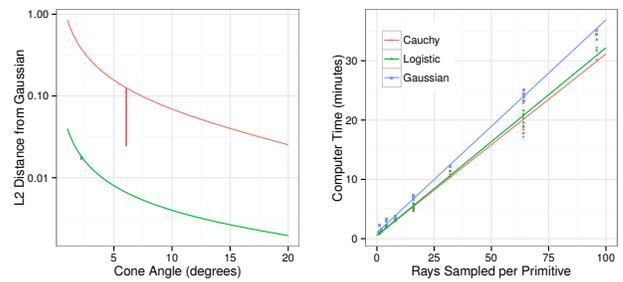


Figure 4: We compare three proposals for longitudinal lobes - Gaussian, Cauchy and Logistic. The left shows minimal L2 distance of Gaussian from the other two. Note the y-axis is on log scale. While the right gives estimates of average processor time as a function of BRDF samples per primitive. We find the logistic distribution is a closer fit to the Gaussian and has indistinguishable cost from Cauchy.

It is natural to approximate the aggregate impact of microfacet roughness by a Gaussian lobe. It is theoretically justifiable by the Central Limit Theorem; the impact of many small, rough displacements on a system results in a smooth bell curve of outputs. Unfortunately Gaussian sampling is costly, especially in the case of hair reflectance, as the range of potential longitudinal angles is closed, depends on the reflected direction, and non symmetric around its center. Even Box-Muller transforms are unable to efficiently take advantage of both samples with a non-symmetric range.

We compare two alternative distributions which are symmetric and have smoothly decaying tails - Cauchy and Logistic. Two tests determined a winner: a shape analysis finding the closest fit to a Gaussian, and cost analysis for impact on render time. For the second, we ran a small experiment, comparing each distribution on total CPU time to render a swath of hair across multiple numbers of samples and cone widths. Results are shown in figure 4.

We find the logistic to have closest match in L2 distance by two orders of magnitude. Using our implementation of either Cauchy or Logistic decreases initialization cost by roughly 30% over Gaussian, and the per-sample cost of Logistic is slightly less than Cauchy, which is in turn 25% less than a Box-Muller implementation of Gaussian sampling. The speed increase of logistic and Cauchy are due to their invertible, closed form CDFs.

Logistic is therefore used for longitudinal scattering in our model,

$$M_\rho(\theta|\omega_r, \beta) = l(\theta_h; \alpha_\rho, \lambda_\rho) \quad (3)$$

where α_ρ and λ_ρ are the longitudinal offset and cone width. The logistic functional form is given in the Appendix in Table 3.

3.2 Azimuthal scattering, N_ρ

The objective of our model is to find azimuthal scattering functions, N_R , N_{TT} , N_{TRT} , and N_{GLINT} , which closely match their theoretical definitions (section 2.1) while still being easily sampled.

Definition 1. A function, $f : \mathcal{X} \subseteq \mathbb{R} \rightarrow \mathbb{R}$ is *easily sampled* if it has a bijective, everywhere finite definite integral with a closed-form inverse, or is a linear combination of such functions.

The following lemma shows why easily sampled scattering functions are useful for ray tracing.

Lemma 1. *Easily sampled functions are fast to sample, stable, and pass the furnace test.*

Proof. Let $f(x)$ be an easily sampled function, and $F(x) = \int_a^x f(s)ds$, where $\mathcal{X} = [a, b]$. Then $Y = F^{-1}((1 - U)F(a) + UF(b))$, has distribution $Y \propto f$ when U is uniformly distributed on $[0, 1]$. The cost to sample depends on the cost to evaluate F^{-1} , which is low.

Second, since f is integrable, both F and F^{-1} are continuous since F is bijective. Hence clamping and numerical overflow produce expected, stable, results.

Finally, we can rewrite $f(x) = (F(b) - F(a)) \frac{f(x)}{F(b) - F(a)}$, the right part of which can be interpreted as a probability density and has integral 1, passing the furnace test. The left part can be interpreted as the total energy of f . Extending the lemma to linear combinations of such f s is straightforward. \square

Returning to azimuthal scattering functions, the first term, N_R , can be solved analytically as a cosine. This function is easily sampled, and an efficient sampling schemes is described in [Hery and Ramamoorthi 2012]. The other terms do not have exact, easily sampled solutions.

The second term, N_{TT} , depends on the longitudinal θ only through its cone width, and can be shown to be symmetric about $-\pi$ with domain $[0, 2\pi]$. Following the run time results of the previous section, we approximate N_{TT} by a logistic distribution, with cone angle a function of θ_D found by Taylor approximation (details given in the Appendix, table 3). It is straightforward to show that the logistic distribution is easily sampled.

The last two scattering functions, N_{TRT} and N_{GLINT} are the most complex. Besides dependence on θ , they often have multiple lobes, and sometimes singularities. Scattering is also no longer symmetric when fibers have elliptical cross-section ($e < 1$). For example, see figure 3.

Here we employ our data-centric, numerical optimization approach.

Recall β_ρ is a parameter vector defining the shape of N_ρ . The components of β_ρ for each ρ are listed in table 2. Consider a set of basis densities, $p_d(\phi; b_d)$, with b_d parameters for p_d , and a_d non-negative weights, for $d = 1, \dots, D$, which are easily sampled for all b_d . Then for any sample of azimuthal angles ϕ_j , $j = 1, \dots, n$,

and suitable distance metric h on \mathcal{F} , the space of distributions, $h : \mathcal{F} \times \mathcal{F} \times [-\pi, \pi] \rightarrow \mathbb{R}_+$, we directly solve the optimization

$$\begin{aligned} a^*, b^* \in \arg \min_{\{a, b\}_{d=1}^D} \\ \sum_{j=1}^n h \left(N_\rho(\phi_j | \beta_\rho) \left\| \sum_{d=1}^D a_d p_d(\phi_j; b_d) \right. \right) N_\rho(\phi_j | \beta_\rho) \quad (4) \\ \text{s.t. } a_d \geq 0 \quad , \quad b_d \in B, \end{aligned}$$

with B a possible set of constraints on b .

We use AIS (section 5.1) to solve (4) for $\rho = TRT$ and $GLINT$. The basis distributions p_d we chose are a combination of skew logistics and monomials for TRT and solely skew logistics for $GLINT$. These distributions are defined in Table 3. Since each basis distribution we consider is easily sampled, N_{TRT} and N_{GLINT} are as well.

To summarize the section, we present explicit solutions for N_ρ . All parameter definitions are in the Appendix.

$$N_R(\phi | \sigma_R) = \frac{1}{4\sigma_R} \cos\left(\frac{\phi}{2\sigma_R}\right)$$

$$N_{TT}(\phi | \theta_d, \sigma_{TT}) = l(\phi; \pi, s(\theta_d)\sigma_{TT})$$

$$N_{TRT}(\phi | \beta_{TRT}) = \sum_{d=1}^5 a_{TRT,d}(\beta_{TRT}) p_{TRT,d}(\phi | b_d(\beta_{TRT}))$$

$$N_{GLINT}(\phi | \beta_G) = \sum_{d=1}^3 a_{GL,d}(\beta_G) p_{GLI,d}(\phi | b_d(\beta_{GLI}))$$

4 Additional Model Features

4.1 Attenuation and volume absorption

As in [Marschner et al. 2003] and [d'Eon et al. 2011], we model Fresnel attenuation at each internal scattering event, and take the corresponding product of reflection and refraction terms as Fresnel for each scattering component. A F_ρ for each S_ρ term.

Figure 5 shows the effect of Fresnel attenuation on the first term, F_R . We compare two competing approaches - the standard model of [cite - Cook and Torrance] adapted by [d'Eon et al. 2011] to fit a cylindrical coordinate system, and a derivation using the Bravais index developed by [Marschner et al. 2003]. We find that the standard approach is more sensitive to changes in the longitudinal angle, and prefer its appearance.

As angles of reflection inside the fiber do not, in general, coincide with the final scattering direction, $w_o - w_i$, there is no obvious extension to higher order terms. We examined both the Bravais formulation of [Marschner et al. 2003] and a more ad hoc function used in [d'Eon et al. 2011]. We prefer the softened look of [d'Eon et al. 2011], after incorporating the suggestion of [Marschner et al. 2003] that Fresnel attenuation for GLINTs should always be calculated at the exact location of singularity.

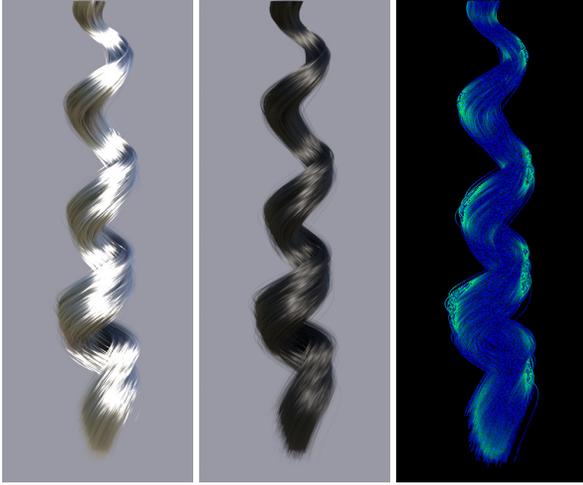


Figure 5: Using a Fresnel term accounting for azimuthal angle correctly attenuates grazing angles. The far left image shows the R term without Fresnel. The second adds a Fresnel term accounting for both longitudinal and azimuthal offsets. The third image is an $L2$ difference between image rendered with this Fresnel and one accounting for only longitudinal angles.

Our volume absorption term is a softened exponential decay in the angle of azimuthal reflectance

$$T_\rho(\theta, \phi) = e^{-\rho_\# \zeta(C) \left| \frac{\cos(\gamma)}{\cos(\theta_D)} \right|} \quad (5)$$

where $\rho_\#$ accounts for the increasing length of the internal light pathway in higher terms $\{R-0, TT-1, TRT-2, GLINT-2\}$, and $\gamma = \gamma(\phi)$ an approximation of the internal, azimuthal angle of reflection given incident and outgoing angles. The term $\zeta(C)$ matches overall wavelength to a user supplied color, C , when the rest of the exponent is maximized. It is straightforward to show that the volume absorption in [Marschner et al. 2003] reduces to ours.

We expose $A_\rho = F_\rho T_\rho$ directly for artist manipulation, and sample from $\rho \sim A_\rho$ for lobe selection.

4.2 Eccentricity and curl



Figure 6: The effect of increasingly elliptical hair fibers, from left to right, top to bottom: $e = 0.85, 0.90, 0.95, 1.00$.

The distribution of cortical cells in a hair fiber cross-section affect how much fibers curl in on themselves. [Robbins 2012] A distribu-

tion of cells in symmetric, concentric circles produces straight hair, while asymmetric distributions produce curly hair. This distribution of cell types also creates eccentricity. More asymmetry results in both curlier and more eccentric fibers. Curl affects light scattering as the position of fiber axis relative to camera direction rotates down the length of the hair, and this position matters for elliptical cross-sections.

We achieve a perception of fiber twisting without modifying hair geometry by defining $\phi_0(l)$, the rotation of hair major axis as a function of distance from the root, l . Our underlying assumption is that ϕ_0 rotates 360° for every full curl a fiber completes, and any partial curls induce proportional partial rotation.

It would be onerous to ask an artist to input the number of curls completed for every hair fiber. We instead estimate reasonable values from eccentricity. The STAM is a robust classification of human hair into 8 curl types from over 1,500 subjects spanning 18 different racial subgroups. [De La Mettrie et al. 2007] STAM is also very correlated with eccentricity. Regressed curl type as a function of eccentricity, explains over 99.9% of the variation in curl type across numerous sources [Robbins 2012]. We use these findings and invert STAM to compute a predicted curl diameter from eccentricity, denoted $CD(e)$.

If we let C_l denote the proportion of curl completed at l , the equation for ϕ_0 is simply

$$\phi_0(l; e) = 2\pi C_l = \frac{2l}{CD(e)} \pmod{2\pi}.$$

Further details are in the Appendix, section 9.2.

5 Fitting Azimuthal Scattering for TRT and $GLINT$

In this section we describe our AIS algorithm for fitting easily sampled distributions to azimuthal scattering for TRT and $GLINT$ terms, and detail the following extensions which improve convergence closer to a global optimum: periodicity constraints, alternating optimization schemes, and early stopping criteria.

5.1 Solving optimization with adaptive importance sampling

The optimization (4) has many local minima and is sensitive to initial conditions. We re-purpose an adaptive importance sampling (AIS) scheme to solve (4). For the rest of this section, we replace the target function N_ρ with a generic g for ease of notation.

AIS mitigates sensitivity to initial conditions as the grid of ϕ_i on which to approximate g is recomputed at every step and focused on the most important areas of g . We tackle scalability by an explicit choice of h as $L2$ distance, which allows us to decouple a_{ds} from b_{ds} . To our knowledge, the idea of AIS was first introduced in [Oh and Berger 1992].

The AIS algorithm consists of 4 steps. We give an overview of each step with a mathematical description, and general pseudocode. Each function used below has implementations in a variety of languages widely available in open source.

- I. Given values of $\{a_1, b_1, \dots, a_D, b_D\}$, set

$$P_d = \int p_d(\phi; b_x) d\phi \quad , \quad \pi_d \propto a_d P_d$$

1: \triangleright Compute total energy, P_d, π_d for each d .

II. Sample

$$\phi_1, \dots, \phi_m \sim \sum_{d=1}^D \pi_d p_d(\phi, b_d) / P_d$$

- 1: \triangleright Sample ϕ from the mixture distribution.
- 2: float $d = \text{SAMPLEDISCRETE}(D, (\pi_1, \dots, \pi_D))$
- 3: float $b_d = \text{GETPARAMS}(\beta)$
- 4: float $\phi = \text{SAMPLEFROMP}(d, b_d)$

III. Compute weights

$$w_i = \frac{g(\phi_i)}{\sum_{d=1}^D a_d p_d(\phi_i; b_d)} / \sum_{j=1}^m \frac{g(\phi_j)}{\sum_{d=1}^D a_d p_d(\phi_j; b_d)}$$

- 1: **for** ($j = 0; j < m; j += 1$) **do**
- 2: float $g_j = \text{PDFG}(\phi_j)$
- 3: float $w_j = g_j / \sum_{d=1}^D a_d p_{dj} / \sum_j w_j$
- 4: **end for**

IV. Solve optimization

$$\{a', b'\}_{d=1}^D \in \arg \min_{\{a, b\}_{d=1}^D} \sum_{i=1}^m h \left(g(\phi_i) \left\| \sum_{d=1}^D a_d p_d(\phi_i; b_d) \right\| \right) w_i$$

s.t. $a_d \geq 0, b_d \in B$.

- 1: \triangleright For $h(g||f) = (g - f)^2$, we decouple a_d and b_d as follows.
- 2: \triangleright With $g_i = g(\phi_i), P_{ij} = p_i(\phi_j; b_i), d^k P_{ij} = (\partial_k b_i) p_i(\phi_j; b_i)$, and $A = \text{DIAG}(a), G = \text{DIAG}(g)$, etc.
- 3: \triangleright Solve gradient for b
- 4: array $\nabla_b = -2(g - a^t P) W d^2 P A$
- 5: \triangleright Solve Hessian for b
- 6: matrix $H_b = 2(d P A)^t W (d P A) - 2 \text{DIAG}((g - a^t P) W d^2 P A)$
- 7: float $b = \text{NEWTONSTEPS}(\nabla_b, H_b)$
- 8: \triangleright Weighted regression of g on p_d gives a
- 9: array $a' = (P^t G P)^{-1} P^t G g$

V. Repeat until weighted objective in (4) is below desired threshold, ϵ .

- 1: float $O = 0$
- 2: **for** ($j = 0; j < m; j += 1$) **do**
- 3: $O += (g_j - \sum_{d=1}^D a_d p_{dj})^2 w_j$
- 4: **end for**
- 5: **if** $O > \epsilon$ **then**
- 6: GOTO(1)
- 7: **end if**

A few notes on the above algorithm:

1. When sampling as in step II, each component of the mixture is chosen according to its total energy, π_d , and P_d is the normalizing constant required to sample p_d . These are also the

exact factors needed to make the sampling distribution equal to the minimizing mixture distribution.

2. We importance sample the fitted azimuthal scattering functions in the exact same way. The importance weights are constant and all equal to $\sum_d a_d P_d$ to account for the normalization of π_d above.
3. The solution for b in step IV is purposefully vague. In fact full optimization at each step is not necessary, and a single gradient descent step is sufficient. Empirical results show that taking roughly 10 gradient steps and then resampling achieves the best global performance.
4. It is also possible to use other functions for h . In particular, [Cappé et al. 2008] consider $h(g||f) = \log(g/f)$, the relative entropy.
5. The IS in AIS leads to a natural inclusion of algorithm iterations in tandem with ray tracing. We allude to an online solution to (4) in section 9.1.

5.2 Periodicity

As azimuthal radiance, N_ρ , has periodicity 2π , we further ensure the condition

$$\sum_{d=1}^D a_d p_d(-\pi; b_d) = \sum_{d=1}^D a_d p_d(\pi; b_d)$$

by adding a Lagrange term to the objective,

$$O + \tau g(\pi) \left(\sum_{d=1}^D a_d p_d(\pi) - \sum_{d=1}^D a_d p_d(-\pi) \right)^2$$

where $\tau > 0$ is a tuning parameter. We have found that $\tau \geq 10$ works well.

5.3 Optimization Schemes

In the previous section it was noted that it may not be necessary, or even advantageous to solve the optimization above fully at each AIS step. Since AIS optimizes on a random sample of domain points, full optimization may lead to spurious fits that plateau at a local optimum. Two other optimization schemes designed to “shake up” the overall solution, and attain fits closer to the global optimum are:

1. Alternate between sequentially optimizing $\{a_d, b_d\}$ for each d , keeping the other $d - 1$ parameter sets constant, and completely optimizing all parameters on all D densities at once. The benefit of sequential optimization is separability and reduction in the computational complexity in each step, leading to local optimization in d . The benefit of complete optimization is full consideration of correlation between the coefficients.
2. Alternate the distribution of ϕ_i , the sampled ϕ grid, between the target distribution $N_\rho(\phi)$ and the current proposal, p_t . Sampling from the first is achieved by sampling from the second with importance weights $N_\rho(\phi)/p_t(\phi)$. This approach cycles two complementary objectives: ensure good fits where N_ρ has large density, and reduce bad fits where p_t has high density but N_ρ does not.

5.4 Early Stopping

Last, we detail a hypothesis test that can stop the AIS procedure at some time $t' < T$, declaring the algorithm converged and further improvements on the objective, O_t , unlikely.

With varied complexity of distributions to fit, finding a universal optimal ϵ for convergence in step V would either limit AIS from achieving the best fit possible, or waste many computing cycles. Another approach is to look at successive objective ratios, O_{t+1}/O_t , but since AIS is stochastic, the objective O_t is not monotone decreasing. We could have a lull in the objective ratio owing to an in-favorable sample of ϕ_i s, rather than actual convergence. Our solution finds overwhelming evidence for convergence in face of such noise by borrowing from the hypothesis testing framework.

Let $o_t = O_t/L$ be the relative objective value, where L is the L_2 norm of the normalized target BRDF, $\|g(\phi) - \int g(\phi)\|^2$, $x_t = \log(o_t)$, and $y_t = x_t - x_{t-1}$. Note that because AIS is allowed to find arbitrarily good and bad fits, we have $y_t \in (-\infty, \infty)$. We make the following assumptions on y_t

1. $y_t = \mu_t + e_t$ with $e_t \sim N(0, \sigma_t)$,
2. there exists t' such that $\mu_t \in \log(1 \pm \epsilon)$ for $t > t'$,
3. $\sigma_{t+1} \leq \sigma_t$ and $\mathbb{E}(e_i e_j) = 0$ if $|i - j| > k$.

In other words, we see a noisy version of the true test error, but the noise does not increase over time, and noise is only correlated for k lagged observations. We then formulate a null hypothesis for lack of convergence by time t as

$$H_{0t} : |e^{\mu t} - 1| > \epsilon \text{ for } t \leq t.$$

The null hypothesis is rejected, and thus AIS is terminated at the first time $T = Jk$ when an estimate of the false positive probability p_T falls below a predetermined level α .

The p-value, $p_T = \prod_{j=1}^J p_{jk}$ can be calculated as follows, starting with the partial sums,

$$S_{jk,t} = \sum_{t-jk+1}^{t-(j-1)k} y_t = x_{t-(j-1)k} - x_{t-jk} \quad (6)$$

, $\hat{\sigma}_{jk,t}^2$ as the empirical standard deviation of the j th step back k terms, $\mu_l = \log(1 + \epsilon)$, $\mu_h = \log(1 - \epsilon)$,

$$\nu_i = \max_{v \in \{\mu_l, \mu_h\}} \mu_i - \frac{\hat{\sigma}_{2k,t}^2}{\hat{\sigma}_{2k,t}^2 + \hat{\sigma}_{k,t}^2} (S_{2k,t}/k - v)$$

with the max replaced by min if $i = h$, $\sigma_t^2 = \hat{\sigma}_{k,t}^2 + \frac{\hat{\sigma}_{3k,t}^2 \hat{\sigma}_{2k,t}^2}{\hat{\sigma}_{3k,t}^2 + \hat{\sigma}_{2k,t}^2}$, and finally,

$$p_t = 1 - \Phi\left(\sqrt{k}(S_{k,t}/k - \nu_l)/\sigma_t^2\right) + \Phi\left(\sqrt{k}(S_{k,t}/k - \nu_h)/\sigma_t^2\right).$$

The following result validates the early stopping approach.

Lemma 2. *Under the 3 assumptions above, the sequential test, rejecting when $p_T < \alpha$, controls the probability of stopping while H_0 is true at level α for all t .*

Proof. First, from multivariate normal calculations, the conditional distribution of $S_{k,t}|S_{2k,t}$ is seen to be normal with moments

$$\begin{aligned} \mu_t | S_{2k,t} &= \mu_t - \frac{\sigma_{t-k-1}}{\sigma_t + \sigma_{t-k-1}} (S_{2k,t} - \mu_{t-k-1}) \\ \sigma_t^2 | S_{2k,t} &= \sigma_t^2 + \frac{\sigma_{t-k-1} \sigma_{t-2k-1}}{\sigma_{t-k-1} + \sigma_{t-2k-1}} \end{aligned}$$

Next a test for H_{0t} , $t = jk$, and test statistics $S_{jk,t}$, is characterized by the rejection region

$$S_{jk,t} \in R_j = ([r_{1j}, \infty) \cap \{\mu_j \leq \mu_L\}) \cup ((-\infty, r_{2j}] \cap \{\mu_j \geq \mu_R\})$$

for some values of r_{ij} . By our independence assumption we have

$$Pr(S_j \in R_j, j \leq J | H_0) = Pr(S_1 \in R_1 | H_0) \prod_{j=2}^J Pr(S_j \in R_j | H_0, S_{j-1})$$

We can bound the first term by

$$\begin{aligned} Pr(S_1 \in R_1 | H_0) &\leq Pr(S_1 \leq r_{1j} | H_0, \mu_1 = \mu_L) \\ &\quad + Pr(S_1 \geq r_{2j} | H_0, \mu_1 = \mu_H) \end{aligned}$$

and similarly for the other $j = 2, \dots, J$ terms, including an additional supremum over the mean of the $j-1$ th term. This shows that a test for H_{0T} , $T = Jk$, rejects when the product of J probability bounds are below α . Plug in estimates of these bounds are exactly p_t evaluated at $t = jk$.

Finally, the sequential test for stopping at the first $t = jk$ such that H_{0t} is rejected will only proceed to the $j+1$ st test if all previous tests have failed to reject. Hence we can have at most 1 false rejection at any time t , and the overall type 1 error of the sequential test is preserved. \square

We have found the values of $\epsilon = 0.05$ and $\alpha = .01$ to work well in practice, and to be reasonably robust. Numerical simulations suggest setting k between 15 and 20.

6 Implementation Details

In our current implementation, we use AIS to precompute a map between parameters β_ρ and those of our optimal mixture distribution a_d and b_d , $d = 1, \dots, D$. The map is stored in a lookup table for each ρ , and is referenced during render time for every ray sampled. There is one row in the table for each variable in β_ρ that must be precomputed. These values are $\phi_r \in [0, \pi/2]$, $\theta_D \in [0, \pi/2]$, $\sigma_\rho \in [10, 25]$ and $e \in [0.85, 1]$, which are taken from a suitably fine grid. We found that angles in 2 degree increments, σ_ρ in 5 degree increments, and eccentricity in 0.5 increments worked well, resulting in a total of 25,572 rows, or scattering function fits, for each *TRT* and *GLINT*. All of our lookup tables for *TRT* and *GLINT* combined take up 20.8 MB of disk space.

We select a row of the lookup table by a function of $(\phi_r, \theta_D, \sigma_\rho, e)$. Our function first finds an ideal row and then returns the closest

row in the table. All other parameters in β_ρ , such as hair color, are defined by the user and evaluated at runtime.

The real time implementation in section 9.1 removes the need for pre-computation and is the subject of current work.

7 Results and Comparison

7.1 Errors in BRDF Fits

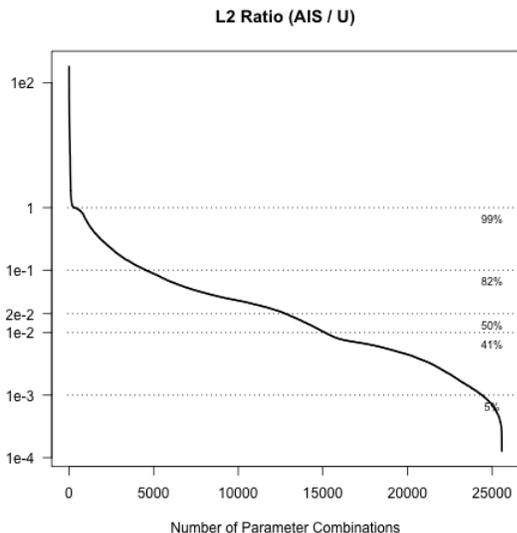


Figure 7: Histogram of L2 errors across all parameters values evaluates for $\rho = TRT$ azimuthal scattering functions.

To measure the performance of the AIS algorithm we estimate the relative L2 error of our algorithm in estimating azimuthal scattering functions,

$$err(\beta_\rho) = \frac{\int_\phi \left\| N_\rho(\phi|\beta_\rho) - \sum_{d=1}^D a_d p_d(\phi; b_d) \right\|_2^2}{\left\| N_\rho(\phi|\beta_\rho) - \int N_\rho(s|\beta_\rho) ds \right\|_2^2},$$

by evaluating the integrands over a fixed grid of 3601 equally spaced angles ϕ_i . For brevity, we only include results for estimating $\rho = TRT$, since they are harder to fit than *GLINT* scattering functions across all parameters. As described in section 6, we fit 25,572 total different scattering functions across a range of parameters. Figure 7 gives a histogram of relative error, showing that 82% of all fits achieve at least a 10 fold closer fit than a uniform distribution. Figure 10 shows a heatmap of all fits showing strong dependence of the performance of the algorithm on parameter regions. Subsequent analysis found that the dark red band of poor fits coincides with very high *GLINT* energy, negating the visual impact of poorer *TRT* fits for these parameter combinations.

Finally, figure 9 shows a heat map of errors further specifying to parameters $e = 0.85$, and $\sigma_{TRT} = 10$, as well as the fits achieved by AIS (red) compared to the truth (black) for the minimum, 25th quantile, 75th quantile, and maximum L2 errors for this parameter range. Fine structure, such as in the center of the fourth image can produce irregular fits. Further study into the space of appropriate basis functions should produce more accurate fits, though we reiterate that the current fits achieve current demands for visual realism.

7.2 Rendering Results

| | R | TT | TRT + GLINT |
|---------------------|-------|-------|----------------|
| [d'Eon et al. 2011] | 13.42 | 14.42 | 31.22 - 163.13 |
| Ours | 13.13 | 13.49 | 16.07 |

Table 1: Render time in minutes using 10 Intel Xeon X5660 2.80Ghz processors. The large range in render time for [d'Eon et al. 2011] is due to a requirement for numerical quadrature evaluations of their Gaussian Detector pdf. The lower bound is setting the number of partitions in the numerical integral to 10, while the upper bound is setting partitions to 100. As can be seen in 8, the numerical quadrature approximation is not converged at 10 partitions.

We have implemented the output from our algorithm described above in commercially available Pixar’s RenderMan ray tracer (<http://renderman.pixar.com/>). For comparison we also rendered all example images using two existing shading models, introduced in [Ou et al. 2012] and [d’Eon et al. 2013] as well as a ground truth path tracer solving for light bounces directly inside a cylinder.

We created two test groomes with Pixar’s proprietary grooming plugin for Maya which were then exported to RenderMan. These groomes were designed to be visual approximations of the tests presented by [Marschner et al. 2003] and [d’Eon et al. 2013]. The curly lock is composed of 5,513 individual strands and the straight groom of 19,226 strands. A supplemental turntable movie of the curly lock under various lighting conditions to demonstrate the ability of our model to handle smooth variation is also supplied.

We tasked ourselves with reproducing the most similar lighting possible between our model, [Ou et al. 2012], and [d’Eon et al. 2013], compared to a ground truth path tracer on the first groom. Figure 11 shows the results of our best efforts, while the following is a log of our experience.

With the exception of eccentricity, [d’Eon et al. 2013] was very similar to our model and therefore fairly easy to match. We found that the longitudinal distribution of [d’Eon et al. 2013] was softer than both Cauchy and logistic alternatives. We found a cone angle difference of about 3° worked well for λ_ρ between 5 and 8.

On the other hand, [Ou et al. 2012] was more difficult to set. Mainly, the model is not normalized and does not natively account for Fresnel or volume absorption. Because Fresnel is especially critical to *R*, we added a common Fresnel term to it. For the other lobes, we first tweaked the numbers to visually emulate our *TT* and *TRT* without absorption or Fresnel. We then lowered their overall energy to match energy left after full attenuation.

Table 1 give runtimes comparing our model and [d’Eon et al. 2013]. We find ours outperforms [d’Eon et al. 2013], especially for TRT + GLINT renders.

The consensus is that our model achieves more realism than [Ou et al. 2012] with less work required of the animator, and renders faster than [d’Eon et al. 2013].

8 Conclusion

In this paper, we presented a data-centric approach to building a light scattering model for hair. We have demonstrated a reasonable fit to the scattering functions of [Marschner et al. 2003], and renders using [d’Eon et al. 2011] as well as a ground truth path tracer, while keeping computation time low. We are also able to account for individual fiber details, such as eccentricity and curl, which all serve to enhance the visual perception of hair.

There are a number of extensions and generalizations that could be addressed in future work. First, the basis of functions we used for optimization could likely be more general. This poses an interesting research problem as standard Fourier expansions may not yield easily sampled approximations. Second, it is straightforward to extend our algorithm to focus on higher order bounces, as [d'Eon et al. 2011] points out that solutions at least exist for the first five. This would allow for a more accurate representation of light, or white colored hair. Third, we have not considered multiple scattering in our model. One direction is to follow the approach of [Zinke et al. 2008]. Fourth, we feel there is room for our model to exist as one of a sequence of models, allowing a smooth degradation in quality for shots further from the camera. This principle of rendering complexity proportional to perceived complexity is discussed in [Goldman 1997].

We finally note that our algorithm could be successfully applied to many other kinds of BRDFs as well, and look forward to extending stochastic optimization principles more broadly to light scattering.

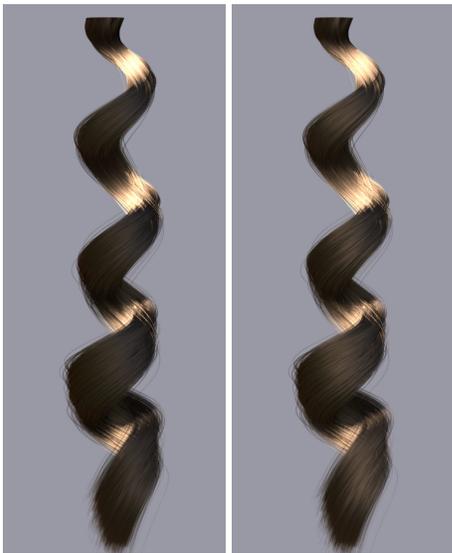


Figure 8: Two rendered images of TRT/GLINT using [d'Eon et al. 2013]. The left uses 10 quadrature partitions, while the right uses 100.

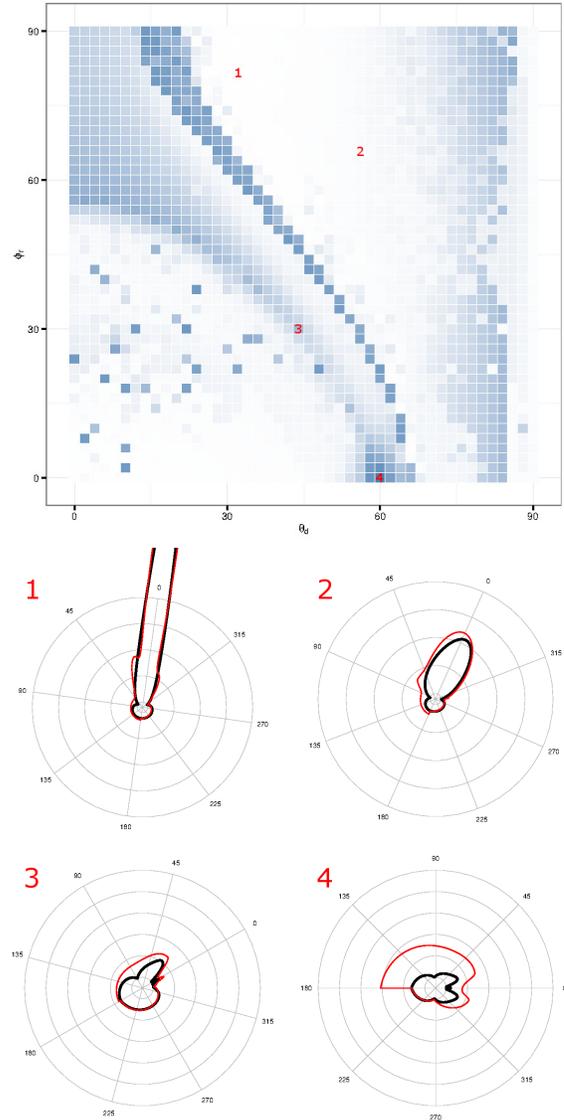


Figure 9: A heat map of the optimized objective in (4). We chose the most peaked, and therefore hardest to fit distributions, $\rho = TRT$, $e = 0.85$, $\sigma_{TRT} = 10$. Below are four fits achieved by our algorithm (red) compared to the truth (black) for 1 - minimum, 2 - 25th quantile, 3 - 75th quantile, and 4 - maximum of errors in the heat map. The 0 radial axis in the bottom images is aligned with ϕ_r from the heat map.

Acknowledgements

Obviously this work would not have happened without the pioneering research from Steve Marschner. The authors would like to thank Brent Burley for all the relevant discussions and ultimately for the suggestion of the logistic function. We are also grateful for the quick adoption by look development and lighting artists at Pixar for the productions of *The Good Dinosaur* and *Finding Dory*. And finally, we would like to acknowledge the Renderman group in Seattle, which even decided to distribute our hair shader as part of Renderman 20.0.

References

- CAPPÉ, O., DOUC, R., GUILLIN, A., MARIN, J.-M., AND ROBERT, C. P. 2008. Adaptive importance sampling in general mixture classes. *Statistics and Computing* 18, 4, 447–459.
- DE LA METTRIE, R., SAINT-LÉGER, D., LOUSSOUARN, G., GARCEL, A.-L., PORTER, C., AND LANGANEY, A. 2007. Shape variability and classification of human hair: a worldwide approach. *Human biology* 79, 3.
- D’EON, E., FRANCOIS, G., HILL, M., LETTERI, J., AND AUBRY, J.-M. 2011. An energy-conserving hair reflectance model. In *Computer Graphics Forum*, vol. 30, Wiley Online Library, 1181–1187.
- D’EON, E., MARSCHNER, S., AND HANIKA, J. 2013. Importance sampling for physically-based hair fiber models. In *SIGGRAPH Asia 2013 Technical Briefs*, ACM, 25.
- GOLDMAN, D. B. 1997. Fake fur rendering. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, ACM Press/Addison-Wesley Publishing Co., 127–134.
- HEITZ, E. 2014. Understanding the masking-shadowing function in microfacet-based brdfs.
- HERY, C., AND RAMAMOORTHI, R. 2012. Importance sampling of reflection from hair fibers. *Journal of Computer Graphics Techniques (JCGT)* 1, 1, 1–17.
- KAJIYA, J. T., AND KAY, T. L. 1989. Rendering fur with three dimensional textures. In *ACM Siggraph Computer Graphics*, vol. 23, ACM, 271–280.
- MARSCHNER, S. R., JENSEN, H. W., CAMMARANO, M., WORLEY, S., AND HANRAHAN, P. 2003. Light scattering from human hair fibers. In *ACM Transactions on Graphics (TOG)*, vol. 22, ACM, 780–791.
- OH, M.-S., AND BERGER, J. O. 1992. Adaptive importance sampling in monte carlo integration. *Journal of Statistical Computation and Simulation* 41, 3-4, 143–168.
- OU, J., XIE, F., KRISHNAMACHARI, P., AND PELLACINI, F. 2012. Ishair: importance sampling for hair scattering. In *Computer Graphics Forum*, vol. 31, Wiley Online Library, 1537–1545.
- ROBBINS, C. R. 2012. Chemical and physical behavior of human hair.
- SADEGHI, I., PRITCHETT, H., JENSEN, H. W., AND TAMSTORF, R. 2010. An artist friendly hair shading system. *ACM Transactions on Graphics (TOG)* 29, 4, 56.

- WARD, K., BERTAILS, F., KIM, T.-Y., MARSCHNER, S. R., CANI, M.-P., AND LIN, M. C. 2007. A survey on hair modeling: Styling, simulation, and rendering. *Visualization and Computer Graphics, IEEE Transactions on* 13, 2, 213–234.
- ZINKE, A., YUKSEL, C., WEBER, A., AND KEYSER, J. 2008. Dual scattering approximation for fast multiple scattering in hair. In *ACM Transactions on Graphics (TOG)*, vol. 27, ACM, 32.

9 Appendix

9.1 Real time adaptive importance sampling

Here we motivate an alternate implementation of our algorithm to compute the mapping, $\beta_{TRT} \leftrightarrow \{a_1, b_1, \dots, a_D, b_D\}$, on-the-fly. A key idea is to leverage the information in sampling $w_i \sim S(w_i, w_r)$ during ray tracing across the entire scene.

Sampling ϕ still uses a lookup table for parameter mapping, but we no longer assume the weights are uniform. Instead explicitly compute $w(\phi) \propto N_\rho(\phi) / \sum_a a_d p_d(\phi; b_d)$. The optimization in step (3) has a particularly easy solution when $M = 1$, and $h(g||f)$ the L_2 distance.

- 1: \triangleright Pre-defined variables: ϵ a small, threshold parameter.
- 2: **if** $(N_\rho(\phi) - \sum_a a_d p_d(\phi; b_d))w(\phi) \geq \tau$ **then**
- 3: float $b'_i \in \{b : (\delta b)p_i(\phi; b) = 0\}$
- 4: $a'_i = (N_\rho(\phi) - \sum_a a_d p_d(\phi; b'_d)) / p(\phi; b'_i)$
- 5: **end if**
- 6: \triangleright Store the objective value $N_\rho(\phi) - \sum_a a'_d p_d(\phi; b'_d)$

Under mild conditions, averaging every new entry in a lookup table row will converge to the same optimum as we find by iterating the full algorithm. Once the running average of objective values is below ϵ , we can consider that row converged, sample any similar rays with weight 1, and avoid further optimization steps or computation of $g(\phi)$.

9.2 Curl details

We compute curl type, T , from eccentricity using the following fitted regression from [Robbins 2012]

$$T(a) = \exp \{ 1.105 + 1.977(a^{-1} - 1.474) + 4.032(a^{-1} - 1.474)^2 + 10.039(a^{-1} - 1.474)^3 \}.$$

Curl diameter, $CD(T)$, is then calculated by linearly interpolating the values in table 4 found from the means of STAM endpoints.

| T | CD(T) |
|------|-------|
| 1.00 | 45.00 |
| 2.00 | 8.25 |
| 3.00 | 4.40 |
| 4.00 | 2.15 |
| 5.00 | 1.20 |
| 8.00 | 0.20 |

Table 4: Curl Diameter (cm) end points for hair types (T).

| Parameter | Description | Typical Values |
|------------------------------|---|-------------------------|
| Fiber properties | | |
| C | RGB hair color vector | 0 to 1 |
| e | eccentricity, ratio minor to major cross-section axis | 0.85 to 1 |
| CD | average curl diameter, estimated from e | 0.20 to 45 |
| κ | gain | 1 |
| Scattering function | | |
| ρ | lobe index | $T, TT, TRT, GLINT$ |
| α_ρ | longitudinal shift. α_ρ in Marschner | |
| λ_ρ | longitudinal width. β_ρ in Marschner | |
| σ_ρ | azimuthal width factor. $\sigma_{TRT} = \sigma_{GLINT}$ | 0.5 to 10 |
| β | Wrapper for user input parameters. | |
| Basis distributions | | |
| μ | mean value | $-\pi$ to π |
| σ_l, σ_r | left and right side widths (skew logistic) | 0 to ∞ |
| s | monomial exponent | 0 to 1 |
| Adaptive importance sampling | | |
| $h(f g)$ | distance function on densities f and g | $(g - f)^2$ |
| D | number of basis components | 3 to 5 |
| p_d | basis component | monomial, skew logistic |
| ϵ | threshold for numerical optimization | $1e - 4$ |

Table 2: All parameters of our shading model

| Name | Functional Form |
|------------------------------|---|
| Distributions | |
| Logistic | $l(\theta_h; \mu, \sigma) = e^{\frac{\theta_h - \mu}{\sigma}} / \sigma \left(1 + e^{\frac{\theta_h - \mu}{\sigma}}\right)^2$ |
| Skew Logistic | $p_{sl}(\phi; \mu, \sigma_l, \sigma_r) = \sigma_l l(\phi; \mu, \sigma_l) 1_{\phi \leq \mu} + \sigma_r l(\phi; \mu, \sigma_r) 1_{\phi > \mu}$ |
| Monomial | $p_w(\phi; \mu, s) = \frac{1}{s+1} (\phi - \mu)^s$ |
| Others | |
| Indicator of set A | 1_A |
| N_{TT} scale approximation | $s(\theta_D) = -\frac{\pi}{2} (1 - \eta(\theta_D)^{-1}) \log(2d - 1 - 2\sqrt{d(d-1)})$ $d = (\sqrt{2} - \tau) / (1 - \tau)$ $\tau = \text{asin}(\eta(\theta_D)^{-1})$ |

Table 3: A list of functions and distributions

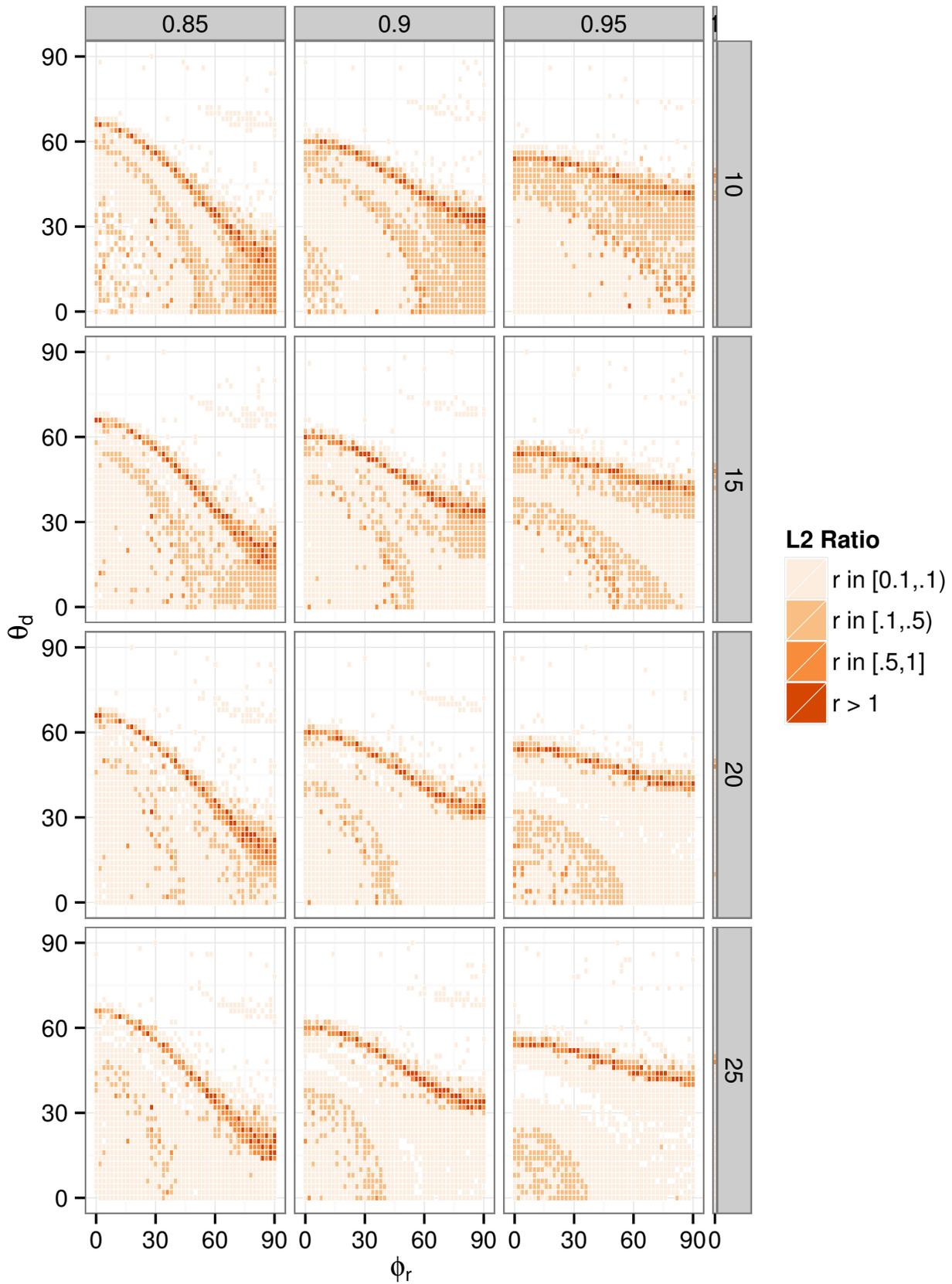


Figure 10: L2 relative errors across all parameter values for $\rho = \text{TRT}$ azimuthal scattering functions.



Figure 11: Comparison of four models for sampling hair, from top left to bottom right - ground truth path tracer , [Ou et al. 2012] , [dEon et al. 2013] , and ours.

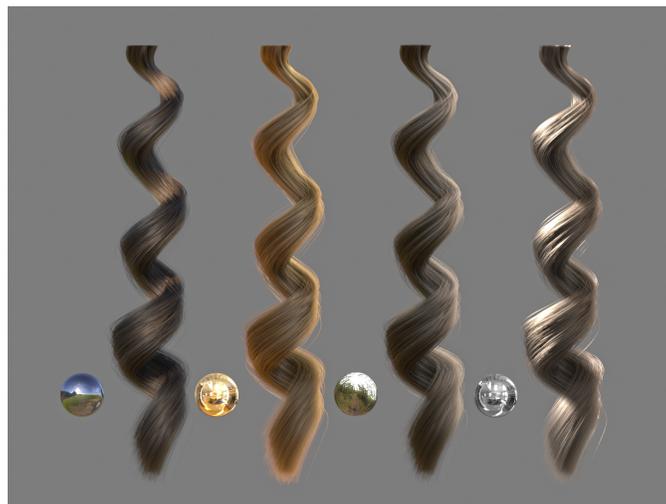


Figure 12: Our hair model under various IBLs.