Path Traced Subsurface Scattering using Anisotropic Phase Functions and Non-Exponential Free Flights

Magnus Wrenninge   Ryusuke Villemin   Christophe Hery

1   Abstract

With the recent move to path tracing for both surface and volume rendering, subsurface scattering has been one of the last light transport modes to rely on empirical or approximate models. Although rendering of subsurface scattering using path tracing is conceptually simple, making the model artist friendly is not.

Our new model is an unbiased subsurface scattering estimator that retains the controllability of the previous models, while introducing both anisotropic (directional) scattering as well as non-exponential free flights, which enables artistic control over the depth of light bleed without losing fine detail.

2   Diffusion models

In computer graphics, the first practical subsurface scattering model was introduced by Jensen et al. in 2001, and used a dipole model on a semi-infinite plane [JMLH01]. This model, along with a recent alternative by Burley [Bur15] which is based on function fitting, have been available in RenderMan for several versions.

An important distinction of diffusion models is that they only model *multiple scattering*. We generally distinguish between scatter events at different path depths: *Zero-scatter* events are photons that enter a surface and immediately reach an exit point without having scattered in the medium itself. *Single scatter* events do the same with a single interaction inside the volume. *Multiple scatter* events have two or more scatter events before exiting through the surface of the object. Diffusion models only capture the multiple scattering contribution, leaving zero and single scatter events to be treated explicitly as separate BxDF lobes.

Figure 1: Left: Burley's SSS model. Middle: Our path traced SSS model. Right: Our non-exponential SSS model. *Head data courtesy of Infinite Realities via Creative Commons.*

2.1 Artist control

Although the appearance of subsurface scattering is due to interactions that take place inside the skin, it is often intuitive for an artist to describe the appearance in terms of *surface albedo*, i.e. the color seen at the surface of the object. For a diffuse surface, the surface albedo can trivially be painted into a texture map, but in the subsurface scattering case, the apparent color of the surface depends on the *scattering albedo* of the participating medium, which is different from the surface albedo.

In this context, diffusion-based subsurface scattering models have an advantage: because the exitant radiance is a closed-form function of the scattering albedo of the medium, it is possible to invert the function such that a user can specify the desired surface color of an object, which the renderer can convert internally into a scattering albedo [Her03]. This inversion process is referred to as *albedo inversion* and is a key feature in making subsurface scattering controllable by artists.

3 Path traced subsurface scattering

While diffusion models provide a closed-form and accurate solution where the surface is flat, areas with high curvature (for example, the nose or ears of a human face) present a problem in that the semi-infinite slab approximation fails. If we instead consider the medium embedded in the surface to be just another volume, we can use the same path tracing technique for subsurface scattering as for general volume rendering. In this context, we perform a random walk starting at the entry point on the surface, through the interior of the object, until we finally intersect the exterior interface and

proceed to gather illumination.

There are several benefits to this approach. The geometric features of the surface are explicitly tracked, resulting in more accurate shadowing and details in high-curvature areas. It is also possible to account for zero- and single-scatter events, making the model better at preserving energy than the diffusion models. Finally, we can also model the event where the ray hits the surface and enters the medium with BxDFs rather than ad-hoc methods.

When entering the subsurface object, we choose the initial direction from a lambertian distribution[1] around the inverse surface normal. With the path initialized, the random walk is constructed by advancing along the ray by a given distance, and at end of each step updating the direction of the next step.

Because the medium is homogeneous, the transmittance (i.e. the ratio of uncollided photons N_u to initial photons N) for a given extinction cross section σ and a given distance s is the exponential function known as Beer's Law:

$$T = \frac{N_u}{N} = 1 - e^{-\sigma s} \tag{1}$$

To determine the length of a given step, we draw samples from transmittance and invert the function to yield the corresponding distance

$$\xi = 1 - e^{-\sigma s}$$
$$s = -\frac{\ln(1 - \xi)}{\sigma}$$

In the isotropic scattering case, the direction of the next step is unrelated to the previous step, and we can simply draw a sample from a spherical distribution.

With each forward step along the walk, a ray is traced against the geometry to see if the step resulted in crossing the exterior interface. If this happens, the random walk is complete and a connection between the entry and the exit point is formed. We proceed to compute direct illumination at the exit point and weigh the incoming radiance by a lambertian response, without applying any surface coloring effects.

If the path continues to a fixed depth (we set this to 256 by default) without finding an exit point, we deem the photon lost and set the final contribution to zero. In practice, this happens very seldom, as we use Russian Roulette to terminate paths with low weight stochastically.

---

[1]The lambertian model simplifies sampling, especially when exiting the surface, but other BxDF models could also be used.
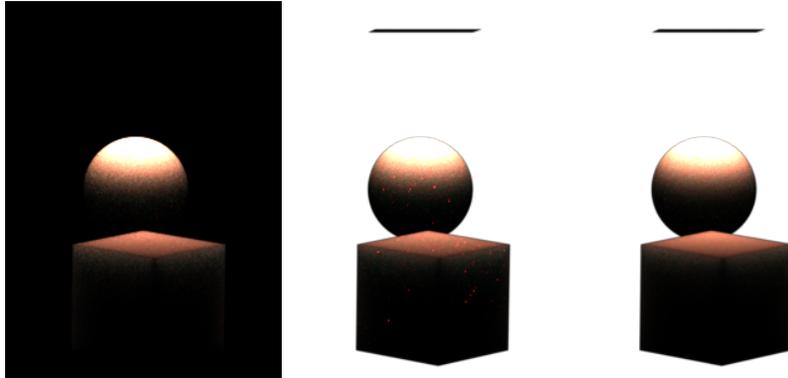
Figure 2: Left: No MIS. Center: Albedo-based MIS. Right: Throughput-based MIS.

3.1 Chromatic effects

As seen in Equation 2, distance sampling uses the extinction coefficient. The most trivial way to extend this to non monochromatic media, would be to first pick a wavelength, and then apply distance sampling based on this particular wavelength. This is obviously not very efficient, particularly considering SSS is just a part of the full light path connecting the camera to a light source. Constraining to a single wavelength in the SSS sampling means that all the subsequent bounces will also operate on that wavelength only. One easy improvement is to use Multiple Importance Sampling (MIS) between channels (RGB):

$$P(s) = P(\lambda) * P_\lambda(s) \tag{2}$$

with

$$P_\lambda(s) = \sigma_\lambda e^{(-s\sigma_\lambda)} \tag{3}$$

and the MIS weight, in case of standard RGB renders using the balance heuristic become:

$$W_\lambda(s) = P(\lambda)/(P(\lambda_R) + P(\lambda_G) + P(\lambda_B)). \tag{4}$$

The only remaining question is how to choose $P(\lambda)$. We could use the single scatter albedo for each channel as a weight, but this gave us occasional fireflies when having highly tinted extinction due to the throughput spiking for very deep path. One possible fix we used, is to use the current throughput instead. That will bound the throughput as the highest contribution channel will also have the higher probability of being picked.

4   Anisotropic scattering

As light scatters in a participating medium, the outgoing direction can be more or less dependent on the incoming direction. For isotropically scattering media, there is no correlation, i.e. with each bounce there is an equal probability for light to scatter in any direction. Sometimes there is a
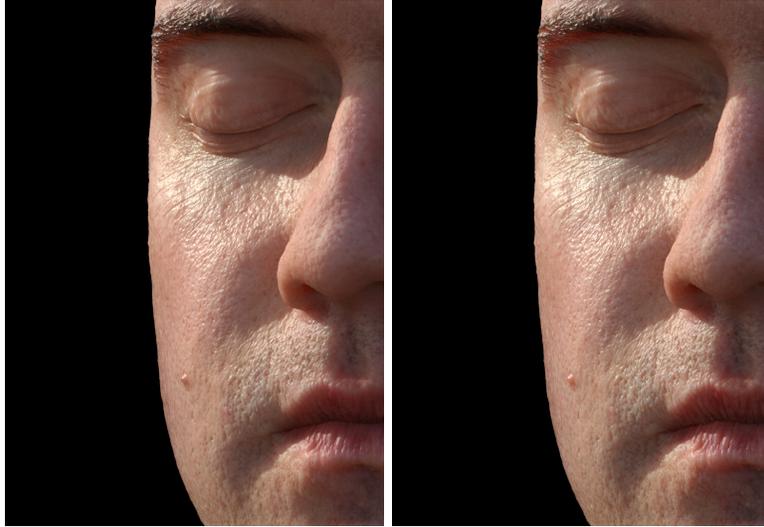
Figure 3: Left: Isotropic scattering. Right: Forward scattering with $g = 0.8$. Note the subtle differences in the eyelids and nostrils. *Head data courtesy of Infinite Realities via Creative Commons.*

pronounced directionality, where light tends to bounce in a direction that is close to the original direction; this process is called *anisotropic scattering*. Media such as smoke are relatively isotropic, whereas water droplets (for example in steam or clouds) are strongly forward-scattering.

The relationship between incoming and outgoing directions is represented by a *phase function*, which assigns a probability to each incoming/outgoing direction pair. Since only the relative angle between the two directions is relevant, phase functions map one-dimensional angle quantities to a one-dimensional probability. Phase functions may be wavelength-dependent, e.g. the Mie scattering function, in which case the probability needs to be evaluated in a spectral fashion.

The diffusion models for subsurface scattering generally do not capture anisotropic scattering, instead assuming isotropic behavior. However, real-world measurements of skin have shown the anisotropy value to be $g = 0.8$, and it would be desirable to incorporate this behavior in a subsurface scattering model for the purpose of increasing realism.

The most popular phase function is Henyey-Greenstein, often abbreviated as HG, which was introduced in the context of astrophysics in 1941 [HG41]. The parameter g controls the anisotropic bias, with $g = 0$ being an isotropic distribution, $g = 1$ being a forward Dirac response and $g = -1$ being a backward Dirac response.

$$p(\theta) = \frac{1}{4\pi} \frac{1 - g^2}{(1 + g^2 - 2g\cos\theta)^{3/2}} \tag{5}$$

The HG function is perfectly invertible, so in path tracing it is possible to generate "perfect" BxDF sampling by choosing outgoing directions (ϕ, θ) based on an incoming direction vector ω.
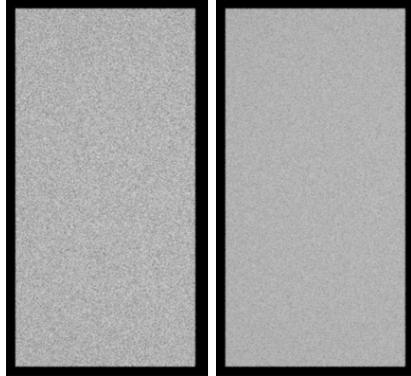
Figure 4: Left: normal sampling. Right: Dwivedi sampling.

5   Dwivedi sampling

Comparing path traced subsurface scattering and our previous diffusion-based subsurface scattering, we found that the former is usually slower to converge. There are a lot of factors to this (we are performing a full simulation now, without approximation!), but one of the main issue is that when doing the random walk, a lot of the paths get lost inside the object. Using dwivedi sampling, we can improve this by "forcing" the walk to come back to the surface. For the technical details, we defer to [D'E16]. We implemented a somewhat simplified version that fits in our renderer. When starting a random walk inside a SSS object, we quickly determine its thickness, and "bias" the sampling to increase the chance to come back to the entry side, or get through to the other side:

- Shoot a probe ray to estimate the thickness of the object. We crudely approximate the object to be a slab.

- When sampling the distance and/or the phase function, use [Kd14] distance and direction sampling. We mix these additional sampling methods to the existing classical sampling. In the end, we MIS between 6 methods (for RGB renders).

- The MIS weighting between classical and dwivedi is based on the heuristic from [Kd14], but we modified it to be slightly different for distance and direction sampling.

If your renderer permits it, you can go one step further and add even more additional sampling, like light sampling, to guide the random walk to the light sources, or even improve the geometric slab estimate.

6 Numerical Albedo Inversion

For diffusion models, it is possible to find a closed form inversion of the scattering-to-surface-color function. One example is the inversion for the Better Dipole model [Her12]. This would be useable in the path tracing framework as well, but it does not handle the anisotropic scattering case.

Over the last few years, Disney have performed numerical inversion of both hair and subsurface scattering by simply rendering images of a single slab of geometry with varying scattering albedo. Once an image is produced, a mapping from final surface color to scattering albedo can be found. Given a large enough set of images, it is then possible to fit a parametric function to the measured values, resulting in the same type of inversion function as for a diffusion model.

In our case, we want to invert the albedo-to-surface-color while also varying the anisotropy parameter g, implying a two-dimensional function. We perform the curve fit in two stages: first we find parameters to a function that computes an albedo α given a desired surface color x. The function is a linear blend between two arctan functions, which was necessary to provide a close enough fit for the low end of the function's range.

$$\alpha(x) = (1 - x^{0.25}) \cdot A \tan^{-1}(Bx)^C + x^{0.25} \cdot D \tan^{-1}(Ex)^F \tag{6}$$

This process was repeated using 40 values for the scattering albedo α and 40 values for the anisotropy g, and for each of the 1600 combinations, a tuple of (A, B, C, D, E, F) was recorded. Then, a polynomial was used to fit A, C, D and F for each value of g:

$$A(g) = a_0 + a_1 g + a_2 g^2 + a_3 g^3 + a_4 g^4 + a_5 g^5 + a_6 g^6 + a_7 g^7, \tag{7}$$

and an exponential-of-polynomial was used for B and E:

$$B(g) = a_0 + a_1 \exp(a_2 g + a_3 g^2 + a_4 g^3 + a_5 g^4 + a_6 g^5 + a_7 g^6 + a_8 g^7). \tag{8}$$

Once the second set of coefficients was found, we could simply make each parameter in Equation 6 a function of g:

$$\alpha(x, g) = (1 - x^{0.25}) \cdot A(g) \tan^{-1}(B(g)x)^{C(g)} + x^{0.25} \cdot D(g) \tan^{-1}(E(g)x)^{F(g)} \tag{9}$$

The function used for fitting by Chiang [CKB16] is a third order exponential-of-polynomial, which works well for the isotropic case, but as g increases, we found that the arctangent gave better results. We also note that solving these jointly would be possible, however we got the best fit from using the two-step process.

Figure 5: Left: Exponential free-flight model. Center: Non-exponential free-flight with bleed 0.5 ($a = 2$). Right: Non-exponential free-flight model with bleed 1.0 ($a = 1$). *Head data courtesy of Infinite Realities via Creative Commons.*

7 Non-Exponential Free Flight

As light moves through a participating medium, the proportion of unscattered to scattered light is a function of the distance traveled. The organization (i.e. distribution) of scattering particles affects the behavior of this function. For *random media*, where the distribution is entirely uncorrelated, the behavior is defined by Beer's Law (Section 3 and Equation 1). However, most participating media found in nature are not entirely uncorrelated, leading to non-exponential free path distributions. For media with positive correlation, the distributions have a much longer *tail*, allowing a small fraction of photons to travel much deeper than an exponential distribution would allow. Although this behavior is not necessarily occuring in skin, we found that the long-tail distribution idea provides a well-defined artistic control over the distance that light bleeds into the subsurface medium.

We found one such model developed by Davis and Xu [DX14] to be especially promising, as it parameterized the non-exponential behavior in a way that makes the visual result intuitive and predictable. In the described model, transmittance (denoted T_a to distinguish it from the exponential-case T in Equation 1) takes the form

$$T_a = \frac{1}{(1 + \sigma s/a)^a} \tag{10}$$

It is important to understand that the non-exponential models change a fundamental behavior in the medium: for the exponential case, the probability of scattering is proportional only to the scattering cross section of the medium; in the non-exponential case, this probability now becomes a function

of both the scattering cross section *as well as the distance traveled in the medium*. Mathematically, this makes it appear as if each photon now has "memory" where past non-interactions influence the probability of future interactions. The geometric interpretation is perhaps more intuitive: if we consider that the medium has positive correlation in the distribution of scattering particles, we imply that there are structural "holes" in the medium where there are fewer particles. For a photon travelling through the medium, it is more likely to scatter in the correlated clusters, and if no interaction is found in the cluster, the subsequent void presents a lower probability for finding the next interaction.

Incorporating the non-exponential free path model into the MC random walk framework requires the follow relationships:

transmittance	$T_a(s) = 1/(1 + \sigma_e s/a)^a$
pdf	$p(s) = \sigma_e/(1 + \sigma_e s/a)^{a+1}$
cdf	$c(s) = 1/(1 + \sigma_e s/a)^a$
sample	$s(\xi) = a \cdot (\xi^{-1/a} - 1)/\sigma_e$
scatter	$S(s) = \sigma_s/(1 + \sigma_e s/a)$

In order to expose the parameter a to the user in a more reasonable range, we have chosen to name the parameter *bleed*, and defining it such that the unit range is meaningful: $a = 1/b$. Figure 5 shows the effect of increasing the bleed parameter.

Although we need to change the underlying equations when incorporating the new model, the albedo inversion is not affected, for the same reason that it isn't affected by changes to MFP. Consequently, artists are free to manipulate bleed without losing control over the apparent surface color.

8   Further reading

For a survey of non-graphics literature related to light transport in participating media, Eugene D'Eon's *A Hitchhiker's Guide to Multiple Scattering* [D'E16] is an excellent resource.

References

[Bur15]   Brent Burley. Extending disney's physically based brdf with integrated subsurface scattering. *ACM SIGGRAPH 2015 Physically Based Shading Course Notes*, 2015.

[CKB16]   Matt Jen-Yuan Chiang, Peter Kutz, and Brent Burley. Practical and controllable subsurface scattering for production path tracing. In *ACM SIGGRAPH 2016 Talks*, page 49. ACM, 2016.

[D'E16]   Eugene D'Eon. A hitchhiker's guide to multiple scattering. Technical report, 2016.

[DX14]   Anthony B Davis and Feng Xu. A generalized linear transport model for spatially correlated stochastic media. *Journal of Computational and Theoretical Transport*, 43(1-7):474–514, 2014.

[Her03]   Christophe Hery. Renderman, theory and practice. *ACM SIGGRAPH 2003 Course Notes*, pages 73–88, 2003.

[Her12]   Christophe Hery. Texture mapping for the better dipole model. *Pixar Internal Tech Memo #12-11*, 2012.

[HG41]   Louis G Henyey and Jesse Leonard Greenstein. Diffuse radiation in the galaxy. *The Astrophysical Journal*, 93:70–83, 1941.

[JMLH01] Henrik Wann Jensen, Stephen R Marschner, Marc Levoy, and Pat Hanrahan. A practical model for subsurface light transport. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 511–518. ACM, 2001.

[Kd14]   Jaroslav Křivánek and Eugene d'Eon. A zero-variance-based sampling scheme for monte carlo subsurface scattering. In *ACM SIGGRAPH 2014 Talks*, page 66. ACM, 2014.