Importance Sampling Spherical Harmonics

Wojciech Jarosz[1,2]   Nathan A. Carr[2]   Henrik Wann Jensen[1]

[1]University of California, San Diego,   [2]Adobe Systems Incorparated

Abstract

In this paper we present the first practical method for importance sampling functions represented as spherical harmonics (SH). Given a spherical probability density function (PDF) represented as a vector of SH coefficients, our method warps an input point set to match the target PDF using hierarchical sample warping. Our approach is efficient and produces high quality sample distributions. As a by-product of the sampling procedure we produce a multi-resolution representation of the density function as either a spherical mip-map or Haar wavelet. By exploiting this implicit conversion we can extend the method to distribute samples according to the product of an SH function with a spherical mip-map or Haar wavelet. This generalization has immediate applicability in rendering, e.g., importance sampling the product of a BRDF and an environment map where the lighting is stored as a single high-resolution wavelet and the BRDF is represented in spherical harmonics. Since spherical harmonics can be efficiently rotated, this product can be computed on-the-fly even if the BRDF is stored in local-space. Our sampling approach generates over 6 million samples per second while significantly reducing precomputation time and storage requirements compared to previous techniques.

Categories and Subject Descriptors (according to ACM CCS): Computer Graphics [I.3.7]: Raytracing—Numerical Analysis [G.1.2]: Approximation—Probability and Statistics [G.3]: Probabilistic algorithms (including Monte Carlo)—

1. Introduction

Spherical harmonics are a frequency-space basis for representing functions defined over the sphere. They have been studied extensively and have widespread applicability to many physical problems ranging from atomic electron configurations in physical chemistry, to the representation of gravitational and magnetic fields of planetary bodies in geodesy. They also appear in quantum mechanics as the solutions of the Schrödinger equation in spherical coordinates.

In computer graphics many quantities are naturally defined over the spherical or hemispherical domain, making spherical harmonics a natural basis for these computations. Furthermore, spherical harmonics have many convenient properties which lead to efficient implementations for convolutions, rotations, and double product integrals. Consequently, spherical harmonics have found recent popularity in computer graphics in the area of precomputed radiance transfer (PRT) techniques [SKS02, KSS02] and have also been used heavily in related domains such as radiative transfer [Cha60] for some time. We refer the reader to recent surveys [Slo08, Gre03] for an in-depth introduction to available techniques.

Another class of basis functions which are often used in computer graphics are wavelets. Whereas spherical harmonics decompose functions into independent frequency bands, wavelets hierarchically decompose functions into components which are localized in both frequency and space. Due to their simplicity, Haar wavelets have received much attention in computer graphics research. Haar wavelets have been successfully applied to PRT [NRH03, NRH04, LSSS04, SM06] and radiosity computation [GSCH93]. Furthermore, Haar wavelets support efficient sampling using a simple hierarchical warping technique [CJAMJ05].

Generating samples according to the intensity distribution of a function is a useful operation in many disciplines. In Monte Carlo techniques, this takes the form of importance sampling and can be used to significantly reduce variance in numerical integration. Importance sampling techniques also have widespread use outside of Monte Carlo rendering. They can be used in non-photorealistic rendering, artificial stippling, and procedural geometry placement. It is often beneficial to also enforce certain quality measures [Coo86, Mit91, Shi91a] on the resulting sample distribution in order
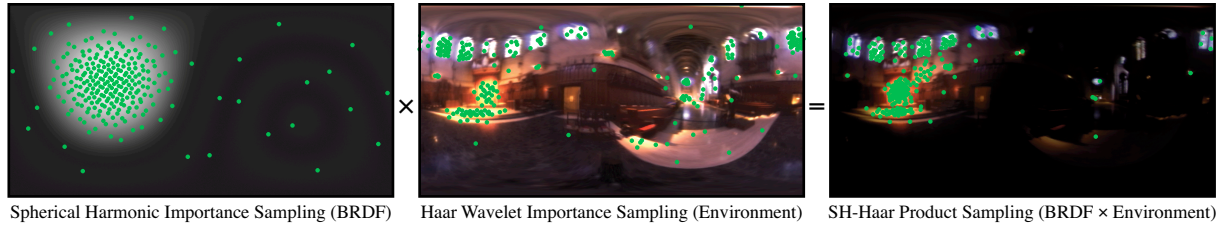
Spherical Harmonic Importance Sampling (BRDF) × Haar Wavelet Importance Sampling (Environment) = SH-Haar Product Sampling (BRDF × Environment)

Figure 1: *In this paper we present a method for importance sampling spherical harmonic functions such as the BRDF on the left. Our technique can also be combined with wavelets to importance sample the product of a spherical harmonic function and a Haar wavelet function (right).*

to reduce variance further or simply to obtain more visually-pleasing distributions.

Haar wavelets and spherical harmonics share many of the same strengths. Haar wavelets can more efficiently approximate high-frequency functions; however, they do not support efficient rotations and may produce distracting aliasing artifacts. Spherical harmonics, on the other hand, support efficient rotation and are efficient for representing low-frequency functions. Spherical harmonics do not suffer from the same aliasing artifacts as Haar wavelets but instead exhibit error in the form of "ringing". To date, however, we are unaware of any efficient technique for importance sampling spherical harmonics. In this paper, we address this limitations by presenting a number of techniques which bridge the gap between the capabilities of spherical harmonics and Haar wavelets.

Contributions. The first contribution of this paper is a high-quality importance sampling technique for spherical harmonics. We draw inspiration from hierarchical wavelet importance sampling [CJAMJ05] and develop a novel memory efficient sampling approach for spherical harmonics which requires little or no precomputation. We bridge the gap between spherical harmonics and Haar wavelets by showing how this new sampling scheme can be used to sample the *product* of a spherical harmonic function and a Haar wavelet function. This product can be evaluated efficiently on-the-fly and allows for high-frequency content in one function and efficient rotation of the other. Such a situation is common in computer graphics when, for example, convolving a smooth BRDF with a high frequency environment map (see Figure 1).

In summary, we present the following novel contributions:

- An importance sampling technique for spherical harmonic functions.
- An analytic conversion between spherical harmonics and spherical mip-maps/Haar wavelets.
- A practical method for sampling the product of spherical harmonics and Haar wavelets.

2. Related Work

A significant amount of work has been invested in efficient sampling techniques for computer graphics. We briefly summarize the most relevant work and focus on importance sampling in the context of rendering.

BSDF Sampling. A commonly used technique for increasing the efficiency of ray tracing based algorithms is to importance sample the reflected rays based on the scattering function. Several important BRDF models can be directly importance sampled, including the Phong model [Shi91b], the Ward model [War92], the Lafortune model [LFTG97], and many others [AS00,PH04]. Complex BRDF models such as the Torrance-Sparrow [TS67] and Cook-Torrance [CT82] models cannot be analytically inverted and require numerical approximations. However, parametric models which are directly invertible, such as the Ward, Lafortune, and half-way vector disk BRDFs [EBJ*06] have been successfully used to approximate complex measured BRDF data [MPBM03]. Several other researchers have addressed the problem of importance sampling measured BRDFs [Lal97,CPB03,CBP04, MPBM03,Mat03,LRR04]. Though spherical harmonics can be used to efficiently represent low frequency BRDFs and indirect lighting [KGPB05], previous researchers have been unsuccessful at using this representation directly for importance sampling.

In radiative transfer within participating media, phase function are generally smooth and commonly represented using spherical harmonics. Unfortunately, since spherical harmonics do not support importance sampling, most techniques rely on simpler, analytically invertible phase functions for rendering. Many natural phase functions are circularly symmetric and contain prominent forward and backward scattering components. Our approach is particularly well-suited for such 1D zonal functions since it gains further efficiency when using only the zonal subset of spherical harmonics.

Environment Map Sampling. Several techniques have been developed for sampling complex distant lighting by placing pre-integrated direction lights at the brightest locations in environment maps [CD01, ARBJ03, KK03, ODJ04, PH04]. These methods work well within diffuse scenes but become increasingly inefficient on glossy and specular surfaces. Cabral [CMS87] and Ramamoorthi and Hanrahan [RH02] addressed this issue by using spherical harmonics to pre-filter the environment map according to the BRDF. However, since efficient sampling techniques were not available for spherical harmonics, these methods could not account for occlusions from other objects. Our importance sampling scheme directly addresses this limitation of spherical harmonics.

Product Sampling. Approaches have also been developed which try to use information about both the BRDF and the lighting in the sampling process. Multiple importance sampling [VG95] can be used to combine several sampling strategies. More recently, approaches have been developed which try to distribute samples according to the product of the lighting and the BRDF. The simplest techniques first draw samples according to one of the terms and then adjust these samples to match the product distribution. This can be done using importance resampling [TCE05], or by rejection sampling [BGH05]; however, these become inefficient if the two functions have peaks in different directions. Furthermore, none of these techniques can preserve stratification.

Clarberg et al. [CJAMJ05] developed a wavelet importance sampling approach which can *directly* generate samples according to the product of two or more functions. They introduced a novel sampling procedure which hierarchically warps a uniform input point set to match the wavelet product. Since warping is performed hierarchically and in small increments, stratification properties of the input set are preserved in the output distribution. Unfortunately, since the two functions need to lie in the same coordinate system, either a large set of pre-rotated copies or pre-computed wavelet rotation matrices [WNLH06] needs to be stored. Due to memory constraints, this limits the method to relatively low resolution lighting and may still require hours of precomputation. Recent work has tried to address this limitation by either heuristically splitting the environment map based on peaks in the BRDF [CETC06] or by constructing a wavelet approximation of the BRDF on-the-fly [CAM08]. Our importance sampling scheme relies on the same sample warping procedure to efficiently produce high quality distributions for spherical harmonics. It can also be applied to product sampling and addresses the limitations of wavelet importance sampling by relying on efficient rotations in the spherical harmonic domain. This makes it the first technique that seamlessly combines and exploits the complementary properties of spherical harmonics and Haar wavelets.

Subr and Arvo [SA07] developed an algorithm for importance sampling the product of distant environment map lighting with an orientable clamped-cosine lobe. They create a triangulated representation of the environment map and store the illumination premultiplied by each of the first nine spherical harmonic basis functions at every vertex. This forms a steerable basis where the clamped-cosine can be efficiently rotated to any orientation. Sampling this triangulation is done using an approach much like Clarberg et al.'s [CJAMJ05] hierarchical sample warping. Though similar, our approach has several advantages to their method. First, Subr and Arvo only considered a circularly symmetric cosine lobe, which they approximate using order 3 spherical harmonics. Our importance sampling technique works for arbitrary spherical harmonic functions of any order. Though generalizing their approach to higher orders is possible, since a full 2D environment triangulation must be stored for each additional SH

Spherical Harmonics
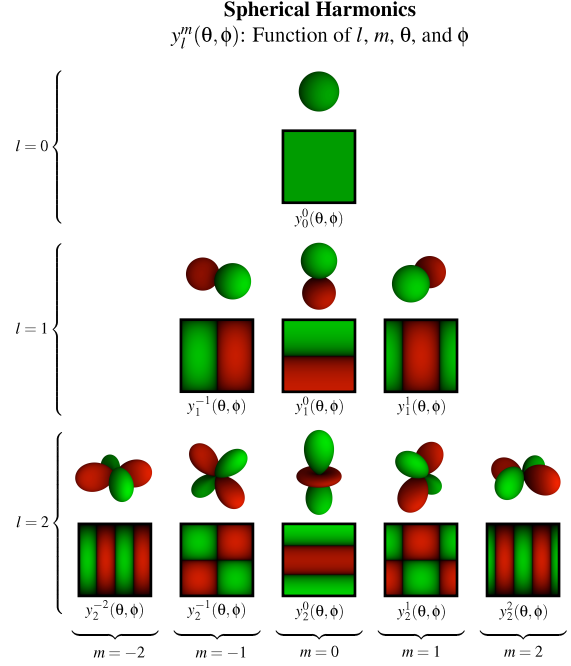$y_l^m(\theta, \phi)$: Function of l, m, θ, and ϕ

Figure 2: Visualizations of the first 9 spherical harmonic basis functions in Cartesian and spherical coordinates, where green indicates positive values and red is negative.

basis, memory usage becomes an issue. Our method, on the other hand, requires no additional storage if importance sampling a single SH function and only utilizes a 1D mipmap per SH basis for product sampling. Furthermore, our approach can exploit the fast wavelet product to multiply two functions on-the-fly, while their method requires a fixed, premultiplied environment map.

3. Background
Since our importance sampling approach relies on both spherical harmonics and hierarchical sample warping, we first provide a brief introduction to these two topics. In the subsequent sections, we describe our importance sampling approach by building on these definitions.

3.1. Spherical Harmonics
Definition. Spherical harmonics are a frequency-space basis for representing functions defined over the sphere. If we represent a direction vector $\vec{\omega}$ using the standard spherical parameterization, $\vec{\omega} = (x, y, z) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$, then the real-valued spherical harmonic basis functions are defined as:

$$y_l^m(\theta, \phi) = \begin{cases} K_l^m P_l^{|m|}(\cos\theta)\cos(|m|\phi) & \text{if } m \geq 0, \\ K_l^m P_l^{|m|}(\cos\theta)\sin(|m|\phi) & \text{if } m < 0, \end{cases} \quad (1)$$

where P_l^m are the associated Legendre polynomials and K_l^m are normalization constants.

The basis functions are indexed according to two integer constants, the *order* or *band*, l, and the *degree*, m. These
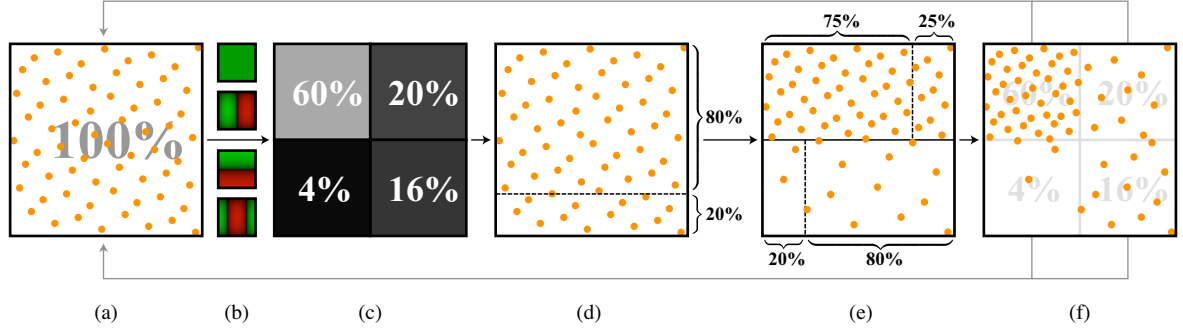
Figure 3: To perform hierarchical sample warping, we start with a uniform set of points (a) and integrate the SH basis functions (b) to obtain the quadrant probabilities (c). The point set is first partitioned vertically according to the marginal row probabilities (d) and then each row is divided horizontally according to the probabilities of its cells (e). The whole process repeats on the four quadrants recursively.

satisfy the constraint that l is a nonnegative integer and $-l \leq m \leq l$. The basis functions can be visually arranged into a pyramid of increasing order as shown in Figure 2. Often it is convenient to reformulate the indexing scheme to use a single parameter $i = l(l+1)+m$, which flattens the pyramid using a level-order traversal.

Properties and Operations. With spherical harmonics, any real-valued spherical function $f(\vec{\omega})$ may be approximated as $\tilde{f}(\vec{\omega})$, by using a linear combination of SH basis functions:

$$f(\vec{\omega}) \approx \tilde{f}(\vec{\omega}) = \sum_{l=0}^{n-1} \sum_{m=-l}^{l} y_l^m(\vec{\omega}) f_l^m, \qquad (2)$$

The coefficients f_l^m are computed by *projecting f* onto each basis function y_l^m.

Spherical harmonics are orthonormal and rotation invariant, allowing for efficient, alias-free least-squares projection and reconstruction of spherical functions. These properties lead to a number of efficient operations for computing rotations [CIGR99, IR96, PH07], convolutions, and double product integrals, all in the SH domain. Spherical harmonics also have a number of other useful properties which are well studied in the literature [MS67].

Limitations. Though spherical harmonics are a natural basis for many problems in computer graphics, they unfortunately do not currently support an effective importance sampling procedure. We address this gap in previous work and draw inspiration for our method from the hierarchical sample warping technique for importance sampling Haar wavelets.

3.2. Hierarchical Sample Warping
In 2D, both mip-maps and Haar wavelets can be seen as a quad-tree hierarchy of coefficients. In a mip-map, each tree node (past the root) points to 4 children and contains exactly 4 scaling coefficients, while in a Haar wavelet each node points to 4 children but contains at most 3 detail coefficients. If we traverse either type of tree in top-down order, at each node we can extract the average value within the four child regions. This observation is responsible for the efficient hierarchical sample warping scheme developed by Clarberg et al. [CJAMJ05]. This technique warps an initially uniform set of points by first computing the marginal probability of intensity in the top and bottom half of the node. The input point set is split according to this ratio and scaled to fit within the two halves. This process is repeated independently on each of the two rows, and the whole process repeats hierarchically on the 4 children. It is easy to generalize this procedure to mip-maps or wavelets in any dimension. We illustrate this process in Figure 3. For simplicity we will use the mip-map basis in our subsequent discussion, but a Haar basis could easily be used instead.

4. Approach
The key insight behind our approach is that we can apply hierarchical sample warping to an SH function $\tilde{f}$ as long as we have a suitable mapping onto the sphere and can compute the average value, or integral, of $\tilde{f}$ at each level in the traversal. The second observation, which makes this idea practical, recognizes that since any SH function is a weighted sum of basis functions y_l^m, the integral of $\tilde{f}$ over any region Ω_R is simply the weighted sum of *integrated* basis functions $\hat{y}_l^m$:

$$\int_{\Omega_R} \tilde{f}(\vec{\omega})\, d\vec{\omega} = \int_{\Omega_R} \left(\sum_{l=0}^{n-1} \sum_{m=-l}^{l} f_l^m y_l^m(\vec{\omega}) \right) d\vec{\omega},$$

$$= \sum_{l=0}^{n-1} \sum_{m=-l}^{l} f_l^m \underbrace{\int_{\Omega_R} y_l^m(\vec{\omega})\, d\vec{\omega}}_{\hat{y}_l^m}. \qquad (3)$$

Equation 3 says that if we have access to a collection of pre-integrated basis functions $\hat{y}_l^m$ then we can apply hierarchical sample warping to any arbitrary SH function by performing a weighted sum with its projection coefficients. A naive application of this concept could be to create a 2D mip-map or Haar wavelet representation of each basis function in a pre-process and then perform a weighted sum of these hierarchical representations on-the-fly to perform warping. This works quite well, but we can do significantly better by exploiting the structure of the basis functions.
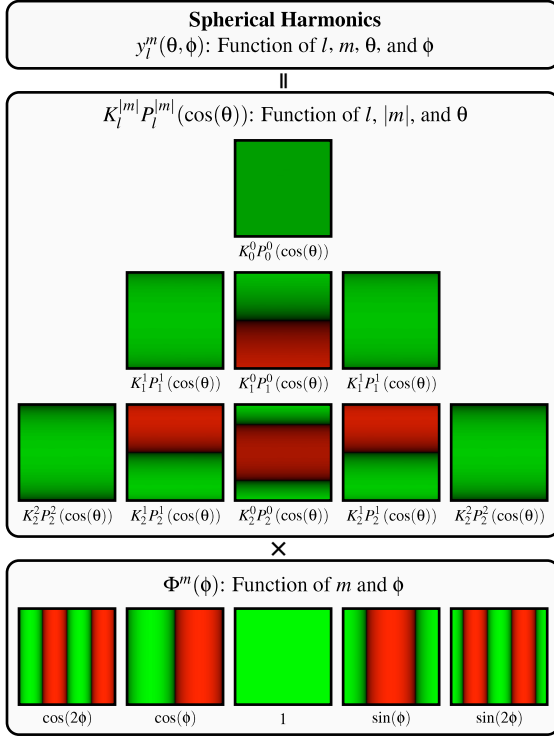
Spherical Harmonics
$y_l^m(\theta,\phi)$: Function of l, m, θ, and ϕ

$\parallel$

$K_l^{|m|} P_l^{|m|}(\cos(\theta))$: Function of l, $|m|$, and θ

$K_0^0 P_0^0(\cos(\theta))$

$K_1^1 P_1^1(\cos(\theta))$ $K_1^0 P_1^0(\cos(\theta))$ $K_1^1 P_1^1(\cos(\theta))$

$K_2^2 P_2^2(\cos(\theta))$ $K_2^1 P_2^1(\cos(\theta))$ $K_2^0 P_2^0(\cos(\theta))$ $K_2^1 P_2^1(\cos(\theta))$ $K_2^2 P_2^2(\cos(\theta))$

$\times$

$\Phi^m(\phi)$: Function of m and ϕ

$\cos(2\phi)$ $\cos(\phi)$ 1 $\sin(\phi)$ $\sin(2\phi)$

Figure 4: *Spherical harmonics (top) are separable into a θ-dependent component (middle) and a ϕ-dependent component (bottom).*

4.1. Computing Basis Function Integrals

If we use a standard spherical mapping, the average value of a basis function y_l^m over a mip-map node is proportional to the following integral in spherical coordinates:

$$\hat{y}_l^m = \iint y_l^m(\theta,\phi) \sin\theta \, d\theta \, d\phi. \qquad (4)$$

This expression can be simplified by observing that the SH basis functions are separable. This can be expressed as:

$$y_l^m(\theta,\phi) = K_l^m P_l^{|m|}(\cos\theta) \, \Phi^m(\phi), \qquad (5)$$

where,

$$\Phi^m(\phi) = \begin{cases} \cos(m\phi) & \text{if } m \geq 0, \\ \sin(|m|\phi) & \text{if } m < 0, \end{cases} \qquad (6)$$

and is illustrated in Figure 4. We can therefore transform the two-dimensional integral in Equation 4 into two one-dimensional integrals:

$$\hat{y}_l^m = \iint K_l^m P_l^{|m|}(\cos\theta) \, \Phi^m(\phi) \sin\theta \, d\theta \, d\phi,$$
$$= K_l^m \left(\int P_l^{|m|}(\cos\theta) \sin\theta \, d\theta \right) \left(\int \Phi^m(\phi) \, d\phi \right). \quad (7)$$

The first term is simply a normalization constant. We describe how to compute the two integrals in the following sections.

4.1.1. Integrals of Associated Legendre Functions

The middle term in Equation 7 involves computing the integral of an expression containing two trigonometric functions

and an associated Legendre function. Solving this integral directly is difficult. We can however simplify the problem by performing a change of variables to cylindrical coordinates and observing that,

$$\int_{\theta_-}^{\theta_+} P_l^{|m|}(\cos\theta) \sin\theta \, d\theta = \int_{z_-}^{z_+} P_l^{|m|}(z) \, dz = \hat{P}_l^{|m|}, \qquad (8)$$

where $z = \cos\theta$, $z_- = \cos(\theta_-)$, and $z_+ = \cos(\theta_+)$. We therefore need a method to compute the integral of the associated Legendre functions, $\hat{P}_l^{|m|}$.

Since associated Legendre functions are just polynomials in z we could symbolically integrate them using a computer algebra system such as Maple or MATLAB and hardcode these functions up to a sufficiently large order. A similar approach was taken by Mousa et al. [MCA06] for integrating the spherical harmonics over a spherical triangle. Though this approach is exact, it does have a few drawbacks. Firstly, a set of hardcoded functions requires a predetermined maximum number of bands in the SH approximation. This may be sufficient for some applications, but reduces the generality of the approach. More importantly, in our application, when evaluating the integral of an SH function using Equation 3 we need to compute the integrals of all basis functions up to some band n, which is very inefficient with this method.

A more efficient approach is to derive recurrence relations which compute the integral $\hat{P}_l^{|m|}$ directly from the integrals at lower values of l and m. Two such methods have been independently developed in the geodesy field [Pau78, DiD82]. We use the approach by DiDonato since it is more efficient and compact. This involves two recurrence relations. The first is a recurrence on l:

$$\hat{P}_l^m = \frac{(l-2)(l-1+m)}{(l+1)(l-m)} \hat{P}_{l-2}^m$$
$$- \frac{2l-1}{(l+1)(l-m)}(1-z^2)P_{l-1}^m(z)\Big|_{z_-}^{z_+}, \qquad (9)$$

which is valid for $0 \leq m < l$. When $m = l-1$, the first term above drops out since $\hat{P}_{l-2}^{l-1}$ is implicitly zero. This relation has a singularity when $m = l$, so we use a second recurrence for this case:

$$\hat{P}_l^l = \frac{1}{l+1}\left[l(2l-3)(2l-1)\hat{P}_{l-2}^{l-2} + zP_l^l(z)\Big|_{z_-}^{z_+}\right]. \quad (10)$$

To start the recurrence we can use:

$$\hat{P}_0^0 = z_+ - z_-, \qquad (11)$$

$$\hat{P}_1^0 = \frac{z_+^2 - z_-^2}{2}, \qquad (12)$$

$$\hat{P}_1^1 = \frac{1}{2}\left[z\sqrt{1-z^2} + \sin^{-1}(z)\right]\Big|_{z_-}^{z_+}. \qquad (13)$$

These two recurrence relations are numerically robust and are *linear* in the total number of basis functions (which is n^2 for all basis functions up to order n). In contrast, symbolic evaluation is *quadratic* in the number of basis functions and can suffer from catastrophic floating-point cancellation

---

Algorithm 1: AssocLegendreIntegral(*n*, *Pa*, *Pb*, *a*, *b*)

Data: n is the number of bands to evaluate.
Data: $Pa[l, m]$ is an array of all associated Legendre functions evaluated at a.
Data: $Pb[l, m]$ is an array of all associated Legendre functions evaluated at b.
Data: $[a, b)$ is the integration interval.
Result: $\hat{P}[l, m]$ is an array of integral results, evaluated for $l \geq 0$ and $0 \leq m \leq l$.

```
   // Evaluate Equation 11:
```
1 $\hat{P}[0, 0] = b - a$;
2 **if** $n \equiv 1$ **then**
3 $\quad$ **return**;
```
   // Evaluate Equations 12 and 13:
```
4 $\hat{P}[1, 0] = \frac{b^2 - a^2}{2}$;
5 $\hat{P}[1, 1] = \frac{1}{2}\left(b\sqrt{1 - b^2} + \operatorname{asin}(b) - a\sqrt{1 - a^2} + \operatorname{asin}(a)\right)$;
6 **for** $l = 2$ **to** $n - 1$ **do**
7 $\quad$ **for** $m = 0$ **to** $l - 2$ **do**
```
          // Evaluate Equation 9:
```
8 $\quad\quad$ $g(x) := (2l - 1)(1 - x^2)Px[l - 1, m]$;
9 $\quad\quad$ $\hat{P}[l, m] = \frac{(l-2)(l-1+m)\hat{P}[l-2,m] - g(b) + g(a)}{(l+1)(l-m)}$;
```
       // Evaluate special case of Equation 9:
```
10 $\quad$ $\hat{P}[l, l - 1] = \frac{2l-1}{l+1}\left((1 - a^2)Pa[l - 1, l - 1] - (1 - b^2)Pb[l - 1, l - 1]\right)$;
```
       // Evaluate Equation 10:
```
11 $\quad$ $\hat{P}[l, l] = \frac{1}{l+1}\left(l(2l - 3)(2l - 1)\hat{P}[l - 2, l - 2] + bPb[l, l] - aPa[l, l]\right)$;
12 **return**;

---

when implemented naïvely. We provide pseudo-code for this recurrence procedure in Algorithm 1.

4.1.2. Integral Along Polar Direction

The last term in Equation 7 is trivial to compute since it is the integral of a trigonometric function:

$$\hat{\Phi}^m(\phi) = \int \Phi^m(\phi)\, d\phi = \frac{1}{m}\begin{cases} \sin(m\phi) & \text{if } m \geq 0, \\ \cos(m\phi) & \text{if } m < 0. \end{cases} \quad (14)$$

This can be evaluated directly; however, to compute integrals for all basis functions up to some band n, this requires $O(n)$ calls to expensive trigonometric functions. We can make this more efficient by using the following multiple-angle recurrence relations:

$$\sin(m\phi) = 2\sin((m - 1)\phi)\cos(\phi) - \sin((m - 2)\phi), \quad (15)$$
$$\cos(m\phi) = 2\cos((m - 1)\phi)\cos(\phi) - \cos((m - 2)\phi). \quad (16)$$

This recurrence can be initialized with:

$$m = 0: \quad \cos(0) = 1, \quad\quad\quad \sin(0) = 0,$$
$$m = 1: \quad \cos(\phi) = \cos(\phi), \quad \sin(\phi) = \sin(\phi),$$

which requires only a single evaluation of the two trigonometric functions for any value of n.

5. Spherical Harmonic Importance Sampling

We can easily apply the concepts from the previous section into a practical approach for importance sampling spherical harmonics. First, we need to map the 2D mip-map or Haar wavelets onto the spherical domain. This can be done using either cylindrical or spherical coordinates. Hierarchical warping needs to evaluate Equation 3 over the area of each node in the traversal. We therefore need to compute the integral, $\hat{y}_l^m$,

of the SH basis functions, which, in cylindrical coordinates, becomes:

$$\hat{y}_l^m = K_l^m \left(\int_{z_-}^{z_+} P_l^{|m|}(z)\, dz\right)\left(\int_{\phi_-}^{\phi_+} \Phi^m(\phi)\, d\phi\right). \quad (17)$$

These two integrals can be efficiently evaluated for all values of l and m using Algorithm 1 and the multi-angle recurrence relations from Section 4.1.2. The same procedure can be used for spherical mappings after applying a change of variables to the above equation.

To perform warping, we start at the root and evaluate the definite integrals of all basis functions, $\hat{y}_l^m$, over the four quadrants of the sphere. The weighted sum of these integrals according to Equation 3 gives us the *exact* integral of the spherical harmonic function for the four quadrants. We warp the points according to this 2×2 importance function, and repeat hierarchically for each of the four child nodes. At each stage of warping we evaluate Equation 17 with integration bounds defined by the extents of the four child nodes. At the end of the procedure the points are distributed with a PDF proportional to the spherical harmonic function. The PDF of each sample is simply the integral of the containing node divided by the integral of the root node.

5.1. Using 1D Mip-maps

The warping technique described so far requires no precomputation and has infinite precision since we can continuing warping to an arbitrary depth. However, infinite precision may not be necessary, especially for low-order spherical harmonic functions which vary smoothly. To improve performance, we could precompute a cylindrical mip-map representation of all the needed basis functions and then perform warping on the weighted sum of these mip-maps. However, Equations 7 and 17 show that, due to separability, we can compute the 2D integral of the spherical harmonic bases by only performing two 1D integrals. We can therefore implicitly construct each of these 2D mip-maps by storing only two 1D mip-maps. This significantly reduces memory requirements and also improves performance due to better cache coherence. This leads to an efficient importance sampling scheme for spherical harmonics which only requires $(n + 1)n/2 + 2n - 1$ 1D mip-maps for order n SH functions. Note that these mip-maps need to be created only once, since they are not function-dependent and can be used for any spherical harmonic function by weighting using a different set of coefficients f_l^m.

5.2. Product Sampling

Clarberg et al. [CJAMJ05] showed how to directly importance sample the product of two functions stored as Haar wavelets by exploiting a fast wavelet product [NRH04]. The spherical harmonic sampling technique described in the previous section implicitly creates a mip-map or Haar wavelet representation of the function. We can therefore exploit the fast wavelet product to importance sample the product of a spherical harmonic function and a mip-map or Haar wavelet
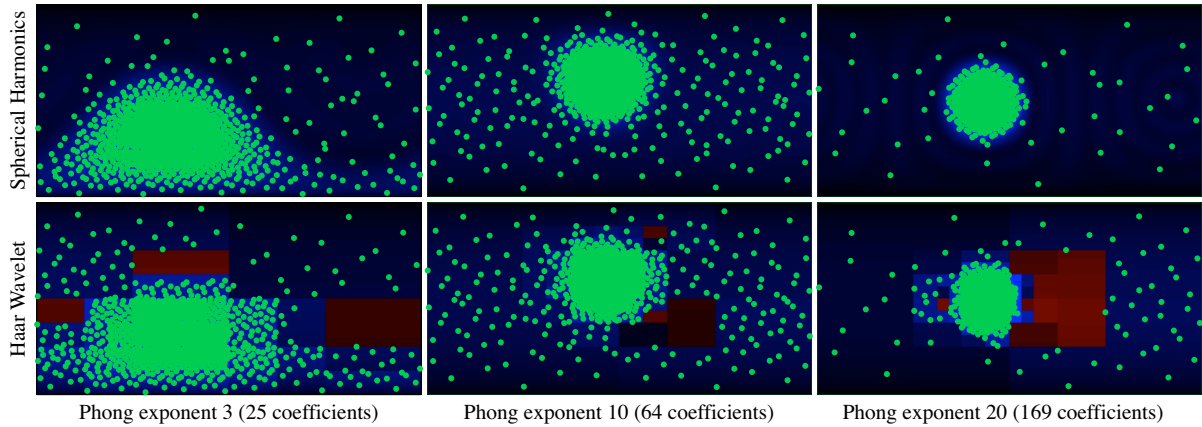
Phong exponent 3 (25 coefficients)	Phong exponent 10 (64 coefficients)	Phong exponent 20 (169 coefficients)

Figure 5: *Comparison of BRDF importance sampling using spherical harmonics (top) and Haar wavelets (bottom). We compare these representations using a Phong lobe with specular exponent 3, 10, and 20. Haar wavelets exhibit significant artifacts when compressed using a low number of coefficients and this results in areas with a negative function value (shown in red). These areas receive no samples, which would introduce bias if used in rendering. Furthermore, the Haar wavelet results in a blocky appearance of the samples. Spherical harmonics provide a significantly smoother representation of the BRDF and consequently also a smoother sample distribution.*
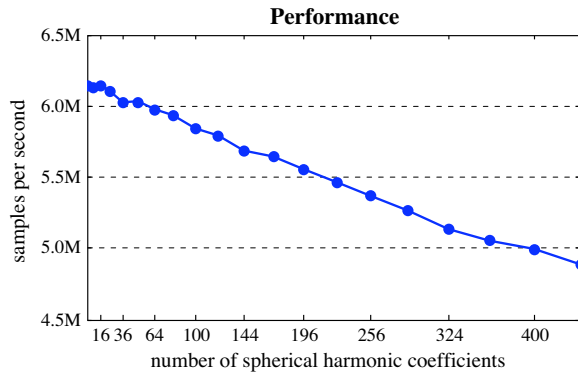

Figure 6: *Performance of spherical harmonic importance sampling. The graph shows samples per seconds as a function of the number of coefficients. The timing is for a 2.4GHz Core2 Duo using one core.*

on-the-fly. This bridges the gap between the spherical harmonic and Haar wavelet bases, allowing us to use spherical harmonics for smooth functions such as BRDFs and Haar wavelets for high-frequency environment maps. Furthermore, the spherical harmonic function can be efficiently rotated, freeing us to use a local coordinate system and a single, high-resolution environment map.

6. Results

We implemented our spherical harmonics sampling algorithm with support for wavelet products. For comparison we also implemented the wavelet importance sampling algorithm by Clarberg et al. [CJAMJ05]. All our results were produced on an Intel 2.4GHz Core 2 Duo using one core.

In Figures 5 and 7 we provide visual comparisons of importance sampling Haar wavelets and spherical harmonics using the same number of coefficients. Figure 5 shows the sample distribution resulting from importance sampling a BRDF compressed using these two representations. We use Phong

lobe BRDFs with specular exponents set to 3, 10 and 20. Both representations use 25, 64 and 169 coefficients respectively. With this extreme compression, the Haar approximation contains many blocky artifacts which are directly visible in the sample distribution. Moreover, the Haar compression introduces several areas with a negative function value (shown in red), which results in areas without any samples.

In Figure 7 we compare the two representations for product sampling of a BRDF and an environment map. Note how the negative regions of the Haar-compressed BRDF directly impact the sample distribution. The product receives no sample points even if the environment contains bright lighting within these regions.

Though the results in Figures 5 and 7 use the same number of SH and Haar coefficients, for rendering, a Haar wavelet implementation would need to store significantly more data. Spherical harmonics support efficient on-the-fly rotations, while a Haar representation would need to be pre-rotated for a large set of normal directions, or a large set of pre-computed rotation matrices would need to be stored. This means that for moderately glossy BRDFs, our spherical harmonic sampling scheme provides significant memory reduction while providing higher quality sample distributions.

Figure 8 shows comparison renderings using our sampling method and wavelet importance sampling in a scene containing a glossy teapot on a diffuse floor within Grace Cathedral. We show results for 8 and 128 shadow rays per pixel for both methods. Our approach produces visually equivalent results to wavelet importance sampling in less time, even though it uses six orders of magnitude less memory (0.3 KB vs. 256 MB). We also show the artifacts that would occur in wavelet importance sampling if the total budget for wavelet data was limited to 1.0 KB. This extreme compression not only produces significantly more noise, but also introduces
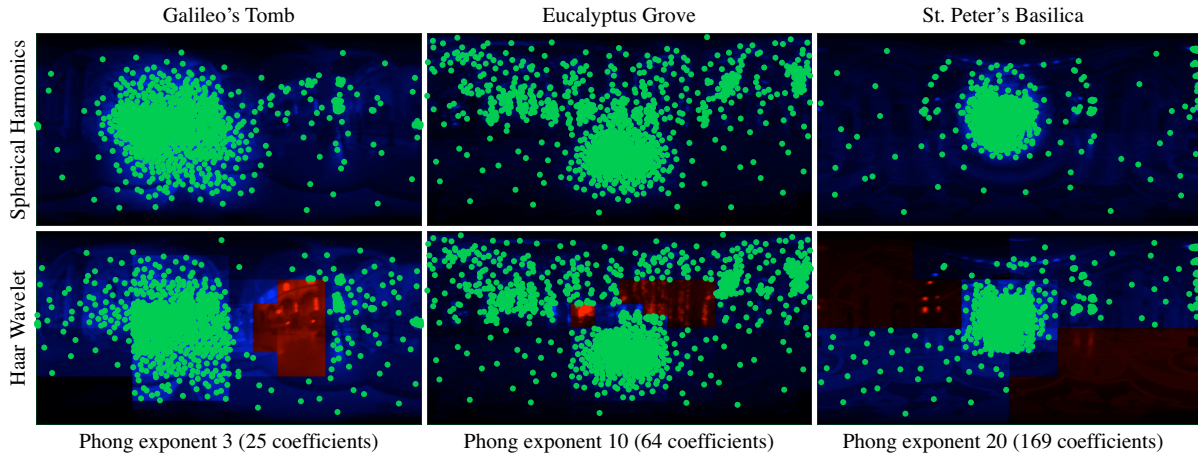
Figure 7: *Comparison of sample distributions computed from the product of an environment map with the Phong lobes shown in Figure 5. Note how the highly compressed BRDF representation introduces negative values which results in areas without any samples even though the environment contains bright sources in these regions. Our product of spherical harmonics for the BRDF and Haar wavelets for the environment results in a smoother sample distribution without the large holes and blocky appearance of the pure wavelet product sampling approach.*

bias and missing lighting features from negative values in the wavelet reconstruction. Note that the render time only increases marginally when moving from 8 to 128 samples per pixel. This is because there is an overhead to compute the product and tracing rays is very fast for this simple scene.

In Figure 6 we plot the performance of our algorithm as a function of the number of spherical harmonic coefficients. Using order 5 spherical harmonics, our importance sampling technique runs as fast as wavelet importance sampling, producing over 6 million samples per second. This is also competitive with other state-of-the-art importance sampling techniques [DH06, Ost07]. The runtime of our algorithm is linearly dependent on the number of spherical harmonics coefficients in the approximation; however, the rate of increase is extremely slow and we are able to generate 5 million samples per second even when using 400 spherical harmonic coefficients.

7. Discussion & Future Work

It may initially seem problematic that the spherical harmonic basis functions (and their integrals) take on both positive and negative values. However, this is not a problem for sample warping since the Haar basis functions also contain negative values. Hierarchical sample warping uses the average value of the *reconstructed* function $\tilde{f}$ within each child node. With Haar wavelets, each node's average is computed using a weighted sum of Haar basis functions; whereas in our approach, each node's average is computed using a weighted sum of integrated SH basis functions. The warping procedure works properly as long as the *reconstructed* function $\tilde{f}$ is non-negative.

Unfortunately, if a function is compressed excessively, negative values may appear in the reconstructed function as well. This is true for both Haar and SH approximations. Since functions that take on negative values are not valid PDFs,

they introduce bias in rendered images and gaps in the sample distributions. Though none of the SH approximations in Figures 5 and 7 exhibit this behavior, lower order approximations or higher frequency functions could introduce negative values. For low-frequency BRDFs, our results show that spherical harmonics out-perform Haar wavelets. Spherical harmonics are not, however, without limitations. Higher specular exponents produce higher frequencies and require more SH coefficients for accurate reconstruction without ringing. Though very high-frequency BRDFs cannot be efficiently represented using spherical harmonics, multiple importance sampling [VG95] works quite well for such BRDFs. An interesting avenue of future work would be to determine the transition point at which Haar wavelets or spherical harmonics become more appropriate.

One way to avoid negative values in the reconstruction is to apply a positive offset to the whole function. Subr and Arvo [SA07] proposed an offset and reflection technique which could easily be applied within our algorithm; however, automatically finding a sufficiently large offset to ensure non-negative values for general functions is challenging. A more practical approach could be to perform a constrained least-squares projection that ensures the approximation remains positive.

8. Conclusion

In this paper, we presented a novel method for sampling spherical harmonic functions as well as sampling the product of spherical harmonics with Haar wavelets/mip-maps. We exploit the separability of the basis functions and utilize efficient recurrence relations to provide an implicit conversion to the mip-map basis. Our approach is able to generate over 6 million samples per second while significantly reducing the memory requirements of previous product sampling techniques.
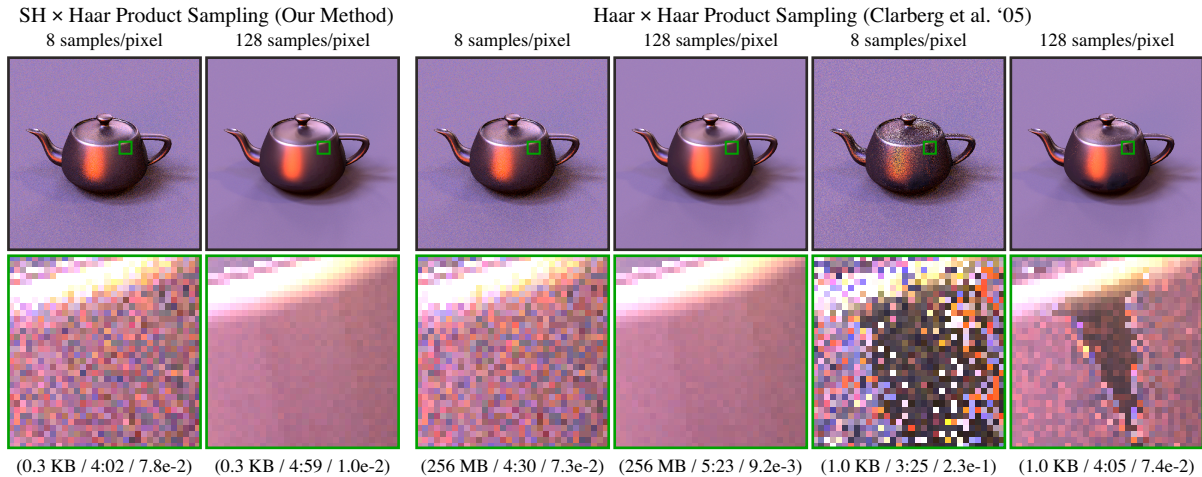
SH × Haar Product Sampling (Our Method)		Haar × Haar Product Sampling (Clarberg et al. '05)			
8 samples/pixel	128 samples/pixel	8 samples/pixel	128 samples/pixel	8 samples/pixel	128 samples/pixel
(0.3 KB / 4:02 / 7.8e-2)	(0.3 KB / 4:59 / 1.0e-2)	(256 MB / 4:30 / 7.3e-2)	(256 MB / 5:23 / 9.2e-3)	(1.0 KB / 3:25 / 2.3e-1)	(1.0 KB / 4:05 / 7.4e-2)

Figure 8: *Rendering comparison between our approach (left) and standard wavelet importance sampling (right). We report the memory usage, render time (in minutes), and resulting RMS error under each image. Our method produces visually equivalent results using significantly less memory. We use order 3 and order 9 zonal harmonics for the Lambertian floor and glossy teapot respectively, whereas nearly 70 million pre-rotated BRDF coefficients are needed to obtain similar quality with wavelets. With 256 wavelet coefficients for the BRDF (far right), wavelet importance sampling suffers from significant noise and lighting artifacts even when using 128 shadows rays per pixel.*

References

[ARBJ03] AGARWAL S., RAMAMOORTHI R., BELONGIE S., JENSEN H. W.: Structured importance sampling of environment maps. *ACM Transactions on Graphics 22*, 3 (2003), 605–612.

[AS00] ASHIKHMIN M., SHIRLEY P.: An anisotropic phong brdf model. *J. Graph. Tools 5*, 2 (2000), 25–32.

[BGH05] BURKE D., GHOSH A., HEIDRICH W.: Bidirectional importance sampling for direct illumination. In *Eurographics Symposium on Rendering* (2005), Eurographics Association, pp. 147–156.

[CAM08] CLARBERG P., AKENINE-MÖLLER T.: Practical product importance sampling for direct illumination. *Computer Graphics Forum (Proceedings of Eurographics 2008) 27*, 2 (2008), 681–690.

[CBP04] CLAUSTRES L., BOUCHER Y., PAULIN M.: Wavelet projection for modelling of acquired spectral brdf. *Optical Engineering 43*, 10 (2004), 2327–2339.

[CD01] COHEN J., DEBEVEC P.: LightGen, HDRShop plugin. http://gl.ict.usc.edu/HDRShop/lightgen/, 2001.

[CETC06] CLINE D., EGBERT P. K., TALBOT J. F., CARDON D. L.: Two stage importance sampling for direct lighting. In *Rendering Techniques 2006: 17th Eurographics Workshop on Rendering* (June 2006), pp. 103–114.

[Cha60] CHANDRASEKHAR S.: *Radiative Transfer*. Dover Publications, New York, 1960.

[CIGR99] CHOI C. H., IVANIC J., GORDON M. S., RUEDENBERG K.: Rapid and stable determination of rotation matrices between spherical harmonics by direct recursion. *Journal of Chemical Physics 111*, 19 (1999), 8825–8831.

[CJAMJ05] CLARBERG P., JAROSZ W., AKENINE-MÖLLER T., JENSEN H. W.: Wavelet importance sampling: efficiently evaluating products of complex functions. *ACM Transactions on Graphics 24*, 3 (2005), 1166–1175.

[CMS87] CABRAL B., MAX N., SPRINGMEYER R.: Bidirectional reflection functions from surface bump maps. In *ACM SIGGRAPH* (1987), pp. 273–281.

[Coo86] COOK R. L.: Stochastic sampling in computer graphics. *ACM Transactions on Graphics 5*, 1 (Jan 1986), 51–72.

[CPB03] CLAUSTRES L., PAULIN M., BOUCHER Y.: Brdf measurement modelling using wavelets for efficient path tracing. *Computer Graphics Forum 22*, 4 (Dec. 2003), 701–716.

[CT82] COOK R. L., TORRANCE K. E.: A reflectance model for computer graphics. *ACM Transactions on Graphics 1*, 1 (Jan. 1982), 7–24.

[DH06] DUNBAR D., HUMPHREYS G.: A spatial data structure for fast poisson-disk sample generation. *ACM Trans. Graph. 25*, 3 (2006), 503–508.

[DiD82] DIDONATO A. R.: Recurrence relations for the indefinite integrals of the associated legendre functions. *Mathematics of Computation 38*, 158 (April 1982), 547–551.

[EBJ*06] EDWARDS D., BOULOS S., JOHNSON J., SHIRLEY P., ASHIKHMIN M., STARK M., WYMAN C.: The halfway vector disk for brdf modeling. *ACM Trans. Graph. 25*, 1 (2006), 1–18.

[Gre03] GREEN R.: Spherical harmonic lighting: The gritty details. *Archives of the Game Developers Conference* (March 2003).

[GSCH93] GORTLER S. J., SCHRÖDER P., COHEN M. F., HANRAHAN P.: Wavelet radiosity. In *ACM SIGGRAPH* (New York, NY, USA, 1993), ACM, pp. 221–230.

[IR96] IVANIC J., RUEDENBERG K.: Rotation matrices for real spherical harmonics. Direct determination by recursion. *Journal of Physical Chemistry 100*, 15 (1996), 6342–6347.

[KGPB05] KŘIVÁNEK J., GAUTRON P., PATTANAIK S., BOUATOUCH K.: Radiance caching for efficient global illumination

computation. *IEEE Transactions on Visualization and Computer Graphics 11*, 5 (2005), 550–561.

[KK03] KOLLIG T., KELLER A.: Efficient illumination by high dynamic range images. In *Eurographics Symposium on Rendering* (2003), Eurographics Association, pp. 45–50.

[KSS02] KAUTZ J., SLOAN P.-P., SNYDER J.: Fast, arbitrary BRDF shading for low-frequency lighting using spherical harmonics. In *Eurographics Workshop on Rendering* (2002), pp. 291–296.

[Lal97] LALONDE P.: *Representations and Uses of Light Distribution Functions*. PhD thesis, University of British Columbia, 1997.

[LFTG97] LAFORTUNE E. P. F., FOO S.-C., TORRANCE K. E., GREENBERG D. P.: Non-linear approximation of reflectance functions. In *ACM SIGGRAPH* (1997), pp. 117–126.

[LRR04] LAWRENCE J., RUSINKIEWICZ S., RAMAMOORTHI R.: Efficient brdf importance sampling using a factored representation. *ACM Transactions on Graphics 23*, 3 (2004), 496–505.

[LSSS04] LIU X., SLOAN P.-P., SHUM H.-Y., SNYDER J.: All-frequency precomputed radiance transfer for glossy objects. In *Eurographics Workshop on Rendering* (June 2004), pp. 337–344.

[Mat03] MATUSIK W.: *A Data-Driven Reflectance Model*. PhD thesis, MIT, 2003.

[MCA06] MOUSA M., CHAINE R., AKKOUCHE S.: Direct spherical harmonic transform of a triangulated mesh. *Journal of graphics tools 11*, 2 (2006), 17–26.

[Mit91] MITCHELL D. P.: Spectrally optimal sampling for distribution ray tracing. In *ACM SIGGRAPH* (1991), pp. 157–164.

[MPBM03] MATUSIK W., PFISTER H., BRAND M., MCMILLAN L.: A data-driven reflectance model. In *ACM SIGGRAPH* (2003), pp. 759–769.

[MS67] MACROBERT T. M., SNEDDON I. N.: *Spherical harmonics: an elementary treatise on harmonic functions, with applications*, third ed. Pergamon Press, Oxford, England, 1967.

[NRH03] NG R., RAMAMOORTHI R., HANRAHAN P.: All-frequency shadows using non-linear wavelet lighting approximation. *ACM Trans. Graph. 22*, 3 (2003), 376–381.

[NRH04] NG R., RAMAMOORTHI R., HANRAHAN P.: Triple product wavelet integrals for all-frequency relighting. In *ACM SIGGRAPH* (2004), pp. 477–487.

[ODJ04] OSTROMOUKHOV V., DONOHUE C., JODOIN P.-M.: Fast hierarchical importance sampling with blue noise properties. *ACM Transactions on Graphics 23*, 3 (2004), 488–495.

[Ost07] OSTROMOUKHOV V.: Sampling with polyominoes. *ACM Trans. Graph. 26*, 3 (2007), 78.

[Pau78] PAUL M. K.: Recurrence relations for integrals of associated legendre functions. *Bulletin Geodesique 52* (1978), 177–190.

[PH04] PHARR M., HUMPHREYS G.: *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann, 2004.

[PH07] PINCHON D., HOGGAN P. E.: Rotation matrices for real spherical harmonics: general rotations of atomic orbitals in space-fixed axes. *Journal of Physics A: Mathematical and Theoretical 40*, 7 (2007), 1597–1610.

[RH02] RAMAMOORTHI R., HANRAHAN P.: Frequency space environment map rendering. *ACM Transactions on Graphics 21*, 3 (2002), 517–526.

[SA07] SUBR K., ARVO J.: Steerable importance sampling. *Interactive Ray Tracing, 2007. RT '07. IEEE Symposium on* (Sept. 2007), 133–140.

[Shi91a] SHIRLEY P.: Discrepancy as a quality measure for sample distributions. In *Eurographics '91*. Elsevier Science Publishers, Amsterdam, North-Holland, 1991, pp. 183–94.

[Shi91b] SHIRLEY P. S.: *Physically Based Lighting Calculations for Computer Graphics*. PhD thesis, University of Illinois at Urbana–Champaign, 1991.

[SKS02] SLOAN P.-P., KAUTZ J., SNYDER J.: Precomputed radiance transfer for real-time rendering in dynamic, low-frequency lighting environments. *ACM Trans. Graph. 21*, 3 (2002), 527–536.

[Slo08] SLOAN P.-P.: Stupid spherical harmonics (SH) tricks. Game Developers Conference, February 2008.

[SM06] SUN W., MUKHERJEE A.: Generalized wavelet product integral for rendering dynamic glossy objects. *ACM Transactions on Graphics 25*, 3 (2006), 955–966.

[TCE05] TALBOT J., CLINE D., EGBERT P. K.: Importance resampling for global illumination. In *Rendering Techniques* (2005), Eurographics Association, pp. 139–146.

[TS67] TORRANCE K. E., SPARROW E. M.: Theory for off-specular reflection from roughened surfaces. *Journal of the Optical Society of America (1917-1983) 57* (Sept. 1967), 1105–+.

[VG95] VEACH E., GUIBAS L. J.: Optimally combining sampling techniques for Monte Carlo rendering. In *ACM SIGGRAPH* (1995), ACM New York, NY, USA, pp. 419–428.

[War92] WARD G. J.: Measuring and modeling anisotropic reflection. *SIGGRAPH Computer Graphics 26*, 2 (1992), 265–272.

[WNLH06] WANG R., NG R., LUEBKE D., HUMPHREYS G.: Efficient Wavelet Rotation for Environment Map Rendering. In *Eurographics Symposium on Rendering* (2006), pp. 173–182.