Bidirectional Importance Sampling for Direct Illumination

David Burke, Abhijeet Ghosh and Wolfgang Heidrich

The University of British Columbia[†]

Abstract

Image-based representations for illumination can capture complex real-world lighting that is difficult to represent in other forms. Current importance sampling strategies for image-based illumination have difficulties in cases where both the illumination and the surface BRDF contain important high-frequency detail – for example, when a specular surface is illuminated by an environment map containing small light sources.
We introduce the notion of bidirectional importance sampling*, in which samples are drawn from the product distribution of both the surface reflectance and the light source energy. While this approach makes the sample selection process more expensive, we drastically reduce the number of visibility tests required to obtain good image quality. As a consequence, we achieve significant quality improvements over previous sampling strategies for the same compute time.*

Keywords: Methods and Applications – Monte Carlo Techniques; Rendering – Ray Tracing; Rendering – Global Illumination.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [COMPUTER GRAPHICS]: Three-Dimensional Graphics and Realism, Raytracing.

1. Introduction

Image-based representations for illumination, such as environment maps, textured area lights, and light fields, have received considerable attention in recent years. The main reason for this attention is that images can capture complex real-world illumination that is difficult to represent in other forms.

When integrating image-based lighting such as environment maps into a rendering system, the use of a good sampling strategy for illumination is paramount. While several researchers have recently worked on this problem, the approach taken in most of that work is an importance sampling strategy based on the energy distribution in the image. Unfortunately, such an approach performs poorly for highly specular surfaces, since samples chosen this way have a low probability of residing within the specular lobe. Similarly, if importance sampling is based solely on the BRDF of the

surface, then the sampling will not perform well for high-frequency illumination. In either case, costly visibility tests are required for directions that contribute little to the surface illumination for a particular viewpoint.

This paper introduces *bidirectional importance sampling*, a method that samples visibility according to an importance derived from the product of BRDF and environment map illumination.[†] The challenge of this approach is to develop an efficient means of drawing samples from this product distribution. The task is complicated by the fact that the 2D BRDF slice varies from point to point on the surface. Furthermore, the environment map is usually represented relative to a global coordinate frame, while the BRDF is expressed in a local frame that changes with surface orienta-

---

[†] E-mail:{dburke, ghosh, heidrich}@cs.ubc.ca

[†] In our discussion we will refer to environment maps although the method also applies to texture-mapped area light sources as illustrated in Figure 7.

tion. For these reasons, precomputation approaches, such as storing a table of the product distribution, are infeasible.

Our solution to this problem is a two-step approach. In the first step, the product distribution is estimated based solely on light source and BRDF information, but not visibility. This estimate is implemented using either rejection sampling (Section 4.1) or sampling-importance resampling (SIR, Section 4.2). Since no visibility tests are involved in this first step, it can be performed rapidly. The second step then uses the distribution generated in the first step for importance sampling of the direct illumination, including visibility tests. This approach has the following benefits:

- Visibility tests are restricted to directions that can contribute significantly to the illumination. The number of visibility tests can be reduced drastically as a result.
- While our method increases the cost of sample generation, we achieve significant quality improvements for the same compute time under the assumption of BRDF representations that support efficient evaluation and sampling.
- The performance gains increase for more complex scenes, since sample generation is independent of scene complexity whereas visibility tests are not.
- Our method creates samples on the fly and does not require expensive precomputation.

The rest of this paper is structured in the following manner. Section 2 reviews some of the relevant work in sampling from images and environment maps. Section 3 gives an overview of our approach, before describing two realizations in Section 4. Stratification as well as an extension adding a solid angle term to the importance are discussed in Section 5. We conclude with results and a discussion in Section 6.

2. Related Work

All rendering systems, both global and local, must at some point compute the direct illumination in the scene. Unfortunately, this task remains expensive, especially for complex light sources such as environment maps and other image-based representations. Much effort has focused on the development of more efficient techniques for completing this task.

2.1. Sampling from Environment Maps

Illumination from environment maps has been a topic of much recent research. Most of this work focuses on interactive applications and therefore uses expensive precomputation [Gre86,HS99,KM00,KVHS00]. In some recent work, the illumination and/or BRDF are projected into finite bases such as spherical harmonics (e.g., [RH01, RH02, SKS02]) and wavelets [NRH03].

Other researchers have used importance sampling techniques to distribute samples according to the energy distribution in the environment map. The importance sampling is often implemented using a point relaxation scheme [CD01, KK03] related to Lloyd's clustering algorithm [Llo83]. This method has also been used in a stippling context for importance sampling from image data [DHvOS00]. These relaxation methods have the disadvantage of requiring time-consuming precomputation. Also, Lloyd's algorithm is not proven to converge in dimensions higher than 1, and in practice these algorithms can miss high-frequency detail in the images.

Ostromoukhov et al. [ODJ04] presented a technique for distributing 2D point samples that is much faster than relaxation-based approaches and also appears to produce a good spatial distribution for the points. In the context of stippling, Secord et al. [SHS02] described an algorithm based on computing the cumulative density function by pre-integrating and inverting the image intensities. Afterwards, samples can be drawn from the cumulative density function in constant time. This is a simple and efficient method, a variant of which we use in our work.

Agarwal et al. [ARBJ03] introduced a sampling method for environment maps in which the sampled importance takes into account both the energy distribution in the environment map and the solid angle separating the samples. In this way, close clustering of environment map samples is avoided, which reduces redundant shadow tests. In our work, we can also choose to include the solid angle in the importance term (Section 5). Like other algorithms, Agarwal et al.'s method is based on point relaxation, but in addition they require quantization of the environment map.

As an extension to their work, Agarwal et al. [ARBJ03] sort the samples for each shading operation by the magnitude of their contribution to the final illumination. They sample all point lights deterministically, in order of contribution, until the contrast that the remaining lights can add falls below a predetermined threshold. This use of the product of BRDF and environment map value is one step towards our approach of drawing samples according to an importance that is the product of BRDF and light distribution.

However, like other point-relaxation methods, Agarwal et al. generate only one sampling pattern and use it throughout the scene. This technique essentially replaces the environment map with a set of directional light sources. On the one hand, this approach eliminates noise, but on the other hand it introduces aliasing visible in the form of quantized penumbra regions, banding, or missing highlights from smaller light sources. Kollig and Keller [KK03] propose to use interleaved sampling [KH01] of multiple precomputed patterns to reduce this problem.

In our work, we use different random sampling patterns for every surface location. While this introduces noise, it gets rid of aliasing and helps avoid visibility tests for directions that are not important for a given BRDF.

2.2. Sampling from BRDFs

Importance sampling from the BRDF is a common operation. The exact mechanics of it, however, depend on the specific representation used. Simple analytical models such as diffuse, Phong, or generalized cosine models can be sampled analytically (see e.g., [Shi00]).

For tabulated BRDFs, McCool and Harwood [MH97] proposed a kd-tree representation that can be efficiently traversed for importance sampling. Recently, Lawrence et al. [LRR04] introduced a method that works on a factored representation instead. This also reduces the memory footprint of the BRDF representation.

In the case of procedural shaders, importance sampling is difficult, but it can be done if the shader provides additional information. For example, Slusallek et al. [SPS95] propose that the shader should provide cosine lobes approximating the full reflectance function for a given point. These lobes can then be sampled analytically.

In our current implementation we use only Phong and diffuse reflection models. However, our method could easily be extended to incorporate more sophisticated materials using any of the above methods.

2.3. Multiple Sampling Approaches

There has also been work on multiple sampling approaches. Veach and Guibas [VG95] weight samples drawn from both the light sources and the BRDF to reduce the variance of the results. Before rendering, a decision is made as to how many samples to draw from each distribution. The resulting variance is therefore a simple blend between the variances of the individual distributions (see Section 3). Our method reduces variance further by sampling directly from the product distribution, rather than just mixing samples taken from the individual distributions.

The recent work of Szecsi et al. [SSSK04] is based on correlated sampling, in which the unoccluded illumination is computed separately, and only the difference due to visibility is sampled. This method generally performs well in fully visible regions, but rather poorly in occluded or partially occluded regions, since the sampling of visibility does not follow a special sampling pattern. Our work, by contrast, focuses visibility tests on directions for which we can expect major contributions to the illumination. Our approach is mathematically straightforward and allows for direct sampling of the product distribution of illumination and BRDF without any guesswork.

3. Bidirectional Importance Sampling

As mentioned in the introduction, we propose a bidirectional sampling approach in which both the energy distribution in the environment map and the reflectance of the BRDF are taken into account. This is a two-step approach: we initially create samples according to either the BRDF alone or the environment map alone, and then adjust these samples to be proportional to the product distribution. The adjusted samples are then used for visibility testing.

We operate on the assumption that creating samples from only the environment map or only the BRDF model is inexpensive, and that the visibility test dominates the cost. This assumption holds for scenes with complex geometry and for BRDF models optimized for sampling. In this scenario, one can benefit from extra time spent in attaining a good sample distribution that takes both the BRDF and environment map into account. Such a distribution selects only those directions for visibility testing that contribute significantly to the reflected radiance of the surface under evaluation.

Consider the direct illumination at a point for a given observer direction ω_r:

$$L_r(\omega_r) = \int_\Omega f_r(\omega_i \to \omega_r) \cos\theta_i L_i(\omega_i) V(\omega_i) d\omega_i, \quad (1)$$

with L_i denoting the incident illumination from an environment map, f_r representing the BRDF, and V being the binary visibility term.

Our approach is to perform importance sampling using the product of the incident light distribution and the BRDF as the importance function:

$$p(\omega_i) := \frac{f_r(\omega_i \to \omega_r) \cos\theta_i L_i(\omega_i)}{\int_\Omega f_r(\omega_i \to \omega_r) \cos\theta_i L_i(\omega_i) d\omega_i}. \quad (2)$$

Observe that the normalization term in the denominator is the direct illumination integral with the visibility term $V(\omega_i)$ omitted. In other words, this term is the exitant radiance in the absence of shadows. We refer to it as L_{ns} ("radiance, no shadows"):

$$L_{ns} := \int_\Omega f_r(\omega_i \to \omega_r) \cos\theta_i L_i(\omega_i) d\omega_i. \quad (3)$$

If we draw sample directions $\omega_{i,j} \sim p(\omega_i)$ according to the product distribution in Equation 2, we can estimate Equation 1 with $L_{N,p}$, where

$$
\begin{aligned}
L_{N,p}(\omega_r) &= \frac{1}{N}\sum_{j=1}^{N} \frac{f_r(\omega_{i,j} \to \omega_r)\cos\theta_{i,j} L_i(\omega_{i,j}) V(\omega_{i,j})}{p(\omega_{i,j})}; \\
&= \frac{L_{ns}}{N}\sum_{j=1}^{N} V(\omega_{i,j}). \quad (4)
\end{aligned}
$$

We refer to $L_{N,p}$ as the *bidirectional estimator* for the direct illumination integral. The evaluation of Equation 4 can be interpreted as taking the unoccluded reflected radiance L_{ns} and scaling it by the average result of N visibility tests performed along directions that contribute most significantly to the radiance.

We can compute the variance of this estimator using standard results for importance sampling (e.g. [Shi00]), and obtain

$$\omega_{i,j} \sim p(\omega_i) \rightarrow \text{var}(L_{N,p}) = \frac{L_{ns}^2}{N} \text{var}(V(\omega_i)).$$

Note that the variance of the bidirectional estimator for the reflected radiance depends only on the variance in the visibility function. By contrast, conventional approaches perform importance sampling either solely from the intensity in the lighting or solely from the BRDF. In the former case, we get the importance function

$$q_L(\omega_i) := \frac{L_i(\omega_i)}{\int_\Omega L_i(\omega_i)d\omega_i} \qquad (5)$$

with the corresponding Monte Carlo estimator

$$
\begin{aligned}
L_{N,L}(\omega_r) &= \frac{1}{N} \sum_{j=1}^{N} \frac{f_r(\omega_{i,j} \rightarrow \omega_r)\cos\theta_{i,j}L_i(\omega_{i,j})V(\omega_{i,j})}{q_L(\omega_{i,j})} \\
&= \frac{\int_\Omega L_i(\omega_i)d\omega_i}{N} \sum_{j=1}^{N} f_r(\omega_{i,j} \rightarrow \omega_r)\cos\theta_{i,j}V(\omega_{i,j}).
\end{aligned}
$$

The resulting variance using this estimator $\omega_{i,j} \sim q_L(\omega_i)$ is then:

$$\text{var}(L_{N,L}) = \frac{\int L_i^2}{N} \text{var}(f_r(\omega_i \rightarrow \omega_r)\cos\theta_i V(\omega_i)).$$

In other words, when proposing samples from the environment only, the resulting variance is proportional to the variance in the BRDF. Similarly, when proposing solely from the BRDF, variance is proportional to the lights. It follows that the greatest reduction in image noise occurs when samples are drawn from the function with greater variance. This is consistent with intuition. If the BRDFs are diffuse but the lighting contains high frequencies, then directions should be chosen according to the importance of the lights. On the other hand, if light sources in the environment map are relatively broad but the surfaces are glossy or shiny, then proposing from the BRDF will be the better approach.

Either approach will produce significant noise if both the BRDF *and* the illumination contain any high frequency information. The solution of Veach and Guibas [VG95] was to combine samples drawn exclusively from either the lights or the BRDF. However, a mix of samples still suffers from dependence on the variances of the individual techniques.

Figure 1 shows angular plots of the probability densities corresponding to the various proposal distributions. The top image depicts samples drawn from a Phong BRDF overlaid onto the energy distribution of an environment map. It is obvious that sampling from the BRDF alone misses the bright lights in the environment. The center image shows samples drawn from an environment map, rendered into the importance function for the Phong BRDF at a specific viewing direction. It can be seen that most of these samples are placed outside the specular lobe of the BRDF. Finally, the
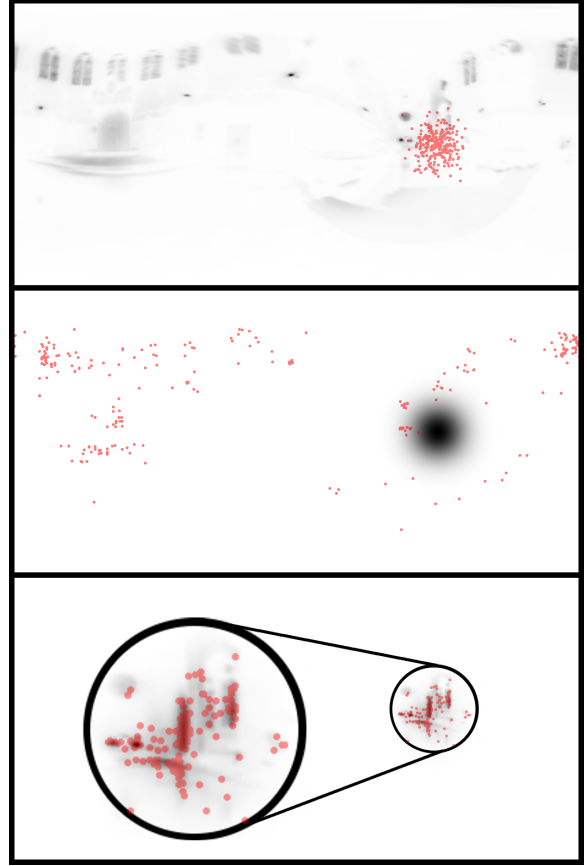

Figure 1: *From top to bottom: angular plots of the importance function of the Grace Cathedral EM, a specular Phong BRDF of exponent 50, and their product. Samples (red discs) drawn solely from the BRDF or the environment vastly undersample the product distribution. The sample set in the bottom image was generated with our SIR technique (described in Section 4.2).*

bottom image represents samples drawn form the product distribution, as well as the product distribution itself. With this method, all samples reside on bright spots of the environment map but also inside the specular lobe.

4. Realizing Bidirectional Sampling

The challenge in realizing bidirectional importance sampling is that the product distribution of the BRDF and environment map is not only too expensive to compute on the fly when drawing samples, but also too high-dimensional to precompute. The BRDF is a 4D function that maps from incoming directions to outgoing directions. The relevant 2D slice of the BRDF, corresponding to a specific outgoing light direction ω_r, varies from point to point in the scene due to changes in the local surface orientation. Directional illumi-

nation such as an environment map is two-dimensional, and thus the BRDF-EM product has six dimensions. Even with a coarse discretization of the BRDF, which might cause high frequency features in the BRDF to be lost, precomputing the product distribution and storing it in a table for sampling is prohibitively expensive.

We suggest the following process for sampling from the product of the lights and the BRDF. First, we create samples according to either the environment map or the BRDF. Then, we adjust the sample distribution such that the directions chosen for visibility testing will be proportional to the product distribution.

We have developed two solutions that realize this redistribution of samples, one based on rejection sampling and the other on the sampling-importance resampling (SIR) algorithm. Note that the overall algorithm is a two-stage approach. That is, the local illumination integral is always estimated with importance sampling, but the subproblem of creating the appropriate samples is solved with either rejection sampling or SIR.

Our two realizations of bidirectional importance sampling are detailed in the following two sections.

4.1. Sample Generation through Rejection

Our first approach for sampling from the product distribution is through *rejection sampling*. To create samples $\omega_{i,j} \sim p(\omega_i)$, we can approximate $p(\omega_i)$ with a PDF $q(\omega_i)$, such that $p(\omega_i) < c \cdot q(\omega_i)$ for some constant c and all directions ω_i. We then generate random samples $\omega_{i,j} \sim q(\omega_i)$ and accept them with a probability of $p(\omega_{i,j})/(c \cdot q(\omega_{i,j}))$.

In our particular case, a simple way of bounding $p(\omega_i)$ from Equation 2 is to use q_L, the energy distribution of the light sources (Equation 5), as the approximation. The bounding constant in this case is $f_{\max} := \max_{\omega_i} q_f(\omega_i)$, the largest value of the BRDF distribution over all incident light directions but for a given fixed exitant direction. Clearly, $p(\omega_i) < f_{\max} \cdot q_L(\omega_i)$. Figure 2 illustrates rejection sampling using this approach.

Since q_L is just the usual importance from the environment map alone, we can sample from it in constant time by pre-computing the cumulative density function through integration and inversion, as described by Secord et al. [SHS02]. This precomputation step needs to be performed only once per environment map, and only requires a fraction of a second, so that even dynamic changes of the environment map in an interactive ray-tracer should be feasible.

In order to accept N visibility samples, on average we have to create $M \approx f_{\max} \cdot N$ environment map samples $\omega_{i,j}$ through importance sampling, and then accept each sample individually with probability
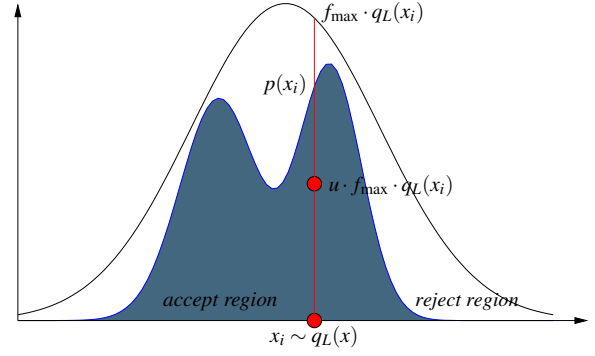

Figure 2: *Sample generation by rejection sampling. A sample $x_i \sim q_L(x)$ is accepted as being a valid sample of the target distribution $p(x)$ if a uniform sample in $[0, f_{\max} \cdot q_L(x)]$ falls under the product distribution $p(x_i)$.*

$$\frac{p(\omega_{i,j})}{f_{\max} \cdot q_L(\omega_{i,j})} = \frac{f_r(\omega_{i,j})\cos\theta_{i,j} \cdot \int_\Omega L_i(\omega_i)d\omega_i}{f_{\max} \cdot L_{ns}}.$$

Both this formula and the final radiance estimate from Equation 4 require the normalization term L_{ns} from Equation 3. We can estimate this term using information that has already been computed during the rejection sampling: since we already evaluate both the BRDF and environment map for the M directions $\omega_{i,j} \sim q_L(\omega_i)$, we can approximate L_{ns} as

$$L_{ns} \approx \frac{\int_\Omega L_i(\omega_i)d\omega_i}{M} \sum_{j=1}^{M} f_r(\omega_{i,j} \to \omega_o)\cos\theta_{i,j}. \quad (6)$$

Another interpretation of this method is that we estimate the unoccluded illumination L_{ns} with M samples, using importance sampling from the environment map. However, we evaluate the visibility for only N of those samples for which the BRDF is large enough to amount to a significant light contribution. The directions for the visibility tests are chosen in an unbiased fashion.

So far, we have bounded the actual target PDF as a constant times the environment map PDF. This is appropriate if the BRDF contains mostly low frequencies, i.e., if $f_{\max}$ is a close bound of the real BRDF distribution. If this is not the case, then most samples will be rejected, and the rejection sampling will become inefficient. In that case, we can perform the same rejection sampling algorithm by approximating the environment map with a conservative bound and then selecting samples according to the real BRDF. Under this scheme, we now have $p(\omega_i) < L_{\max} \cdot q_f(\omega_i)$, which amounts to generating samples from the BRDF alone and then rejecting them according to the product distribution as before.

Given these two ways of rejection sampling, we usually want to draw the initial samples in such a way that the

bounding constant is minimized. That is, if $f_{\max} < L_{\max}$ we importance sample from the environment map; otherwise, we importance sample from the BRDF. In practice, we randomly choose which of the two methods to use. The method with the smaller bounding constant is chosen with a higher probability.

As demonstrated in Section 6, our rejection sampling approach has worked well in our experiments. However, the inherent downside of using rejection sampling is that one cannot guarantee bounds on the execution time for creating a new sample. If the area between $c \cdot q(\omega_{i,j})$ and $p(\omega_{i,j})$ is large, the probability of sample acceptance will be low.

One way of dealing with this is to choose a maximum number of sample attempts in the rejection sampling. If no samples are accepted, a possible strategy could be to test visibility for a random subset. A less expensive but biased possibility is to use the unoccluded illumination wherever visibility has not been tested at all. The rationale behind this approach is that the rejection process will fail mostly in very dark areas, where the product of illumination and BRDF is very small. In these areas, the visibility term will not have significant impact anyway. In practice, we have not found it necessary to resort to these biased methods, since the rejection sampling acceptance probability has been sufficiently high even in the presence of highly specular BRDFs and complex environments.

4.2. Sample Generation through SIR

Our second method from sampling the product distribution does not suffer from the unbounded execution time of the rejection sampling. This method uses the so-called *sampling-importance resampling* (SIR) algorithm [Tan96, GCSR95, SG92].

SIR first draws a set of M samples $X = \{x_1, \ldots, x_M\}$ from a simple distribution $q(x)$. The actual target distribution $p(x)$ is evaluated at these M samples, and the resulting values are used to approximate p. In a second step, a smaller set of N samples $Y = \{y_1, \ldots, y_N\}$ is drawn from X with sample probabilities $w(x_i)$ proportional to their importance ratio $p(x_i)/q(x_i)$. As the number of first-round samples M approaches infinity, the sample set Y can be shown to have been drawn directly from p. The closer q approximates p, the faster the method converges.

We can apply SIR to the problem of drawing samples from the bidirectional distribution. We can use either q_L (i.e., sampling from the light sources) or q_f (i.e., sampling from the BRDF) for the first stage. As in the rejection sampling approach, starting with q_L is advantageous if the illumination contains higher frequencies than the BRDF and vice versa, since the higher frequency factor better approximates the shape of the product distribution.

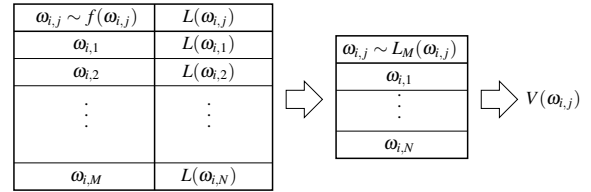Figure 3 summarizes the approach. The total number of

Figure 3: *Sampling-importance resampling (SIR). First, M samples are proposed from q_f, the PDF of the BRDF. The candidate directions are then resampled based on the incoming light along those directions, producing N samples for visibility testing. N is generally much less than M.*

samples generated for each pixel is exactly $M + N$. This is an improvement over rejection sampling for two reasons. First, execution time is tightly bounded. We no longer have to wait an indeterminate time for the rejection criterion to accept a sample. Using the SIR algorithm, samples can be drawn directly from the product distribution in constant time.

The second improvement over rejection sampling is that the sample sizes M and N can be chosen freely, yielding fine control over the tradeoff between quality and time. For example, the BRDF sample size can be adjusted based on the expense of sampling from the BRDF model. The sample size M dictates the quality of the estimate of L_{ns}, and hence the quality of unoccluded regions. Also, it is possible to directly select N — the target number of visibility rays traced per pixel — based on, for example, scene complexity.

As the cost of ray tracing typically dominates rendering time, our general approach for generating results has been to fix N and adjust M so as to increase or decrease the variance in resampled directions. Typical values of M are one to two orders of magnitude larger than N. Note that conventional importance sampling from either the BRDF or the illumination alone are just special cases of the SIR technique where $M = N = 1$.

5. Enhancements

Stratification. It is also possible to stratify bidirectional sampling, although this would only make a difference for low frequency illumination and BRDFs. Sampling from the BRDF or environment map is just importance sampling, and can hence be stratified.

One way of achieving approximate stratification is to sample from the cumulative density function based on a low-discrepancy series rather than Poisson distributed samples [SHS02]. This is the approach we take in our implementation. If the samples in the first stage are stratified, the visibility rays in the rejection sampling case are automatically also stratified, because we are just using a subset of the original samples. In the SIR algorithm, the resampling stage is again essentially an importance sampling step and hence can also be stratified.

Solid Angle Weighting. So far, we have used the light intensities directly as an importance function whenever we sampled from the light sources. Agarwal et al. [ARBJ03] pointed out that the variance of the visibility tests can be reduced by introducing a solid angle term that prevents the rays from clustering in small regions, since the visibility test is likely to yield the same result for similar rays.

This solid angle weighting is easy to incorporate into our sampling strategy, simply by adjusting the probability function generated from the environment map. This proceeds as follows. Like Agarwal et al., we first quantize the image into k intensity levels. This binning is performed on the logarithm of pixel intensity to account for HDR representations. Next, connected components are found in the quantized image by running a breadth-first search. The solid angle of each connected component is found by summing the solid angles of each pixel making up the connected component. Importances L_i of the pixels, originally taken just from intensity, are now scaled by this solid angle area. The new importance is given $q'_L = L \cdot (\min(0.01, \Delta\omega))^b$ as discussed by Agarwal et al. [ARBJ03], where b is in the range $[0.1, 0.2]$, depending on the average size of the light sources in the environment map.

The time complexity of the area-weighting algorithm described above is linear in the size of the environment map, and can be performed in negligible time during loading. Sampling is still constant in time using the cumulative density function of the new importance distribution. In our experiments, we found that this additional solid angle weighting does *not* measurably improve our results. We believe that this is due to the fact that we work with very small sample sizes, which makes clustering of visibility rays unlikely, even without the solid angle weighting. However, since including the term is cheap, we use it anyway.

6. Results

In the following, we compare the results of our techniques with previous sampling strategies for rendering from environment maps. In our tests, illumination comes from image-based representations of illumination, such as environment maps and texture-mapped area light sources. Images were generated with a reasonably well-optimized ray tracer using a voxel grid as the acceleration data structure for intersection queries. Our comparisons examine the output quality of the various rendering algorithms for a fixed amount of computing time. We performed these tests on a 3.0 GHz P4 running Linux.

Figures 8, 4, and 6 contain images of Michelangelo's David in the Grace Cathedral environment. We use the version of David with 700k-triangles acquired from the Digital Michaelangelo Project [Sta01]. In our implementation, intersecting a ray with the David model takes, on average, 6.1 μs on our test machine. The Grace Cathedral environment is a 1024×512 HDR map with a contrast ratio of $10^7 : 1$. In all tests, each algorithm was given 13.0 seconds to render a 176×248 image. This small image resolution was chosen in order distinguish differences between the images when presented in print form.

In a first test, we compared rejection sampling and SIR (Figure 8). Both the algorithms produced images of indistinguishable quality at the same computing time for a variety of combinations of materials and environment maps. In the rest of this section, we therefore compare previous sampling techniques only to our SIR algorithm, which we prefer because of its deterministic performance characteristics.

Figure 6 compares bidirectional sampling to earlier methods: sampling only from either the lights or BRDFs, and Veach&Guibas' multiple importance sampling [VG95]. In the latter case, the weights for choosing between lights and BRDF were optimized manually through trial and error. For bidirectional importance sampling, we used SIR with $M = 800$ primary samples and $N = 15$ final samples for which visibility was tested.

The first row of the figure uses a glossy Phong BRDF with an exponent of 10. In this case, sampling from the environment map only (left column) is still preferable to sampling from the BRDF (center left), since the environment map contains higher frequencies than BRDF. Even so, sampling from the environment map only results in visible noise. Multiple importance sampling produces a result comparable to environment map sampling, while bidirectional sampling clearly outperforms all other methods.

The second row of Figure 6 shows the same scene with a shinier BRDF (Phong exponent of 50). Now, sampling from the BRDF produces better results than sampling from the environment map. Multiple importance sampling further improves on this result. However, bidirectional sampling again outperforms all other methods. In the last row of the figure, we added a diffuse component. This significantly lowers the quality of BRDF sampling. Again, bidirectional sampling is superior to the other strategies without having to adjust weights as in the case of multiple importance sampling.

Figure 7 shows more comparisons between bidirectional sampling and importance sampling from light sources. In the left image pair, the illumination is from an HDR environment of lower frequency than the Grace Cathedral while the BRDF of the Buddha model has significant specular (Phong exponent of 50, $k_s = 0.5$) as well as diffuse ($k_d = 0.5$) components. In this case, sampling only according to either the BRDF or the illumination performs particularly poorly compared to bidirectional sampling. In the right image pair, the light source is now a texture-mapped area light. Note how the reflections of the windows on the shiny floor are smoother with bidirectional sampling.

Figure 4 presents a quality comparison between bidirectional sampling and the best case scenarios for sampling

Figure 4: *Quality comparison between SIR algorithm and the best cases for importance sampling techniques. Top row: David with a purely diffuse BRDF in high frequency lighting. Bottom row: David with a purely specular BRDF (Phong exponent 50) in low-frequency lighting. Top left: Importance sampling according to EM. Top right: SIR algorithm proposing samples according to EM and resampling according to BRDF. Bottom left: Importance sampling according to BRDF. Bottom right: SIR algorithm proposing samples from the BRDF and resampling according to EM. 176×248 images computed in 13.0 seconds.*

from either the lights or the BRDF only. The top row shows the David model with a purely diffuse BRDF in high frequency lighting of the Grace Cathedral. This is the best case for importance sampling from the lights, as the environment map contains all the high frequency information, whereas the BRDF is very smooth. Veach&Guibas' multiple importance sampling more or less reduces to pure importance sampling from illumination in this case. Bidirectional sampling does better than purely sampling from the illumination even in this case, since it accounts for the cosine falloff of the diffuse material.

The bottom row of Figure 4 shows a highly specular David in the comparatively low frequency lighting of the Uffizi Gallery. This is the opposite scenario, where it makes sense to sample according to the BRDF, which is a high fre-

quency function. Even here, bidirectional sampling outperforms pure importance sampling, resulting in a higher quality image for the same compute time.
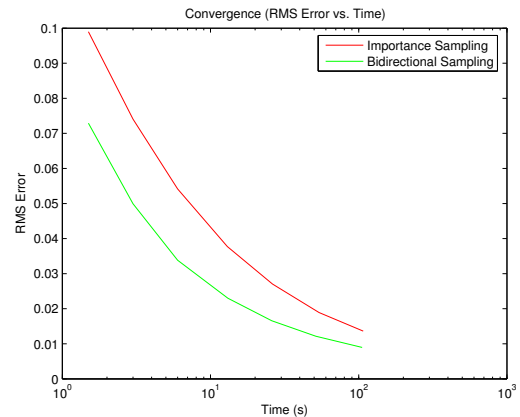

Figure 5: *Convergence plots of RMS errors for importance sampling and bidirectional sampling. Note how the RMS error reduces faster for bidirectional sampling.*

Finally in Figure 5, we present a comparison of the convergence in terms of RMS errors for importance sampling and bidirectional sampling. The plot here was computed for the David model (Phong exponent 50, $k_s = 0.5, k_d = 0.5$) in the Grace Cathedral environment, with first round sampling from the illumination and resampling based on the BRDF. It is clear from the figure that the RMS error converges faster for bidirectional sampling. We found similar behavior for other materials and environment maps.

In summary, the results presented here clearly demonstrate that the approach of sampling directly from the product distribution outperforms previous sampling strategies. What is more, we are able to achieve comparable quality with far fewer rays, meaning that our techniques are particularly beneficial to rendering complex scenes where ray-scene intersection queries are expensive.

7. Conclusions

We presented two Monte Carlo strategies for sampling the incident illumination from environment maps, taking into account both the light distribution and the surface reflectance. By providing a means of sampling from a more complex target distribution, our methods achieve lower variance, especially in renderings of scenes with high frequency lighting or specular BRDFs, as compared to traditional importance sampling strategies.

Although our proposed bidirectional methods take longer to generate samples than simpler approaches, the number of samples required to achieve good quality is considerably less than when sampling according to a simple function. For

Figure 6: *David in Grace Cathedral –* 176×248 *images rendered in* 13.0 *seconds. Left column: Importance sampling purely from the illumination (100 samples). Center left: Importance sampling purely from the BRDF (75 samples). Center right: Combined sampling (Veach&Guibas) with manually fine-tuned weights. Right: Bidirectional importance sampling with SIR (15/800 samples). Top row: Phong exponent 10, $k_s = 1.0, k_d = 0.0$. Center: Phong exponent 50, $k_s = 1.0, k_d = 0.0$. Bottom row: Phong exponent 50, $k_s = 0.5, k_d = 0.5$.*

large datasets with complex structures, the time required to trace shadow rays will dominate the rendering time. In such cases, our methods provide greater benefit over importance sampling from the EM or BRDF alone.

Future work in this direction could be the examination of other sampling strategies that exist in the literature, such as iterative SIR, Metropolis-Hastings, and particle filtering [AdFDJ03]. The general idea behind these strategies is to use samples that have already been drawn as a basis for proposing further, fitter samples. It would be interesting to

explore how these methods of sampling from more complicated distributions could be applied to other problems in computer graphics.

8. Acknowledgements

References

[AdFDJ03] ANDRIEU C., DE FREITAS N., DOUCET A., JORDAN M.: An introduction to MCMC for machine learning, 2003. 9

[ARBJ03] AGARWAL S., RAMAMOORTHI R., BELONGIE S., JENSEN H. W.: Structured importance sampling of environment maps. *ACM Transactions on Graphics (Proc. Siggraph) 22*, 3 (July 2003), 605–612. 2, 7

[CD01] COHEN J., DEBEVEC P.: Light-gen. http://www.ict.usc.edu/~jcohen/lightgen/lightgen.html, 2001. 2

[DHvOS00] DEUSSEN O., HILLER S., VAN OVERVELD C., STROTHOTTE T.: Floating points: A method for computing stipple drawings. In *Proc. of Eurographics* (Aug. 2000), pp. 41–50. 2

[GCSR95] GELMAN A., CARLIN J., STERN H., RUBIN D.: *Bayesian data analysis*. Chapman and Hall, London, 1995. 6

[Gre86] GREENE N.: Environment mapping and other applications of world projections. *IEEE CG&A 6*, 11 (1986), 21–29. 2

[HS99] HEIDRICH W., SEIDEL H.-P.: Realistic, hardware-accelerated shading and lighting. In *Proc. of ACM Siggraph '99* (Aug. 1999), pp. 171–178. 2

[KH01] KELLER A., HEIDRICH W.: Interleaved sampling. In *Eurographics Workshop on Rendering* (June 2001), pp. 269–276. 2

[KK03] KOLLIG T., KELLER A.: Efficient illumination by high dynamic range images. In *Eurographics Symposium on Rendering* (June 2003), pp. 45–51. 2

[KM00] KAUTZ J., MCCOOL M.: Approximation of glossy reflection with prefiltered environment maps. In *Proc. of Graphics Interface* (May 2000), pp. 119–126. 2

[KVHS00] KAUTZ J., VÁZQUEZ P.-P., HEIDRICH W., SEIDEL H.-P.: Unified approach to prefiltered environment maps. In *Eurographics Workshop on Rendering* (2000), pp. 185–196. 2

[Llo83] LLOYD S.: An optimization approach to relaxation labelling algorithms. *Image and Vision Computing 1*, 2 (1983), 85–91. 2

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

[MH97] MCCOOL M. D., HARWOOD P. K.: Probability trees. In *Proc. Graphics Interface* (1997), pp. 37–46. 3

[NRH03] NG R., RAMAMOORTHI R., HANRAHAN P.: All-frequency shadows using non-linear wavelet lighting approximation. *ACM Transactions on Graphics (Proc. Siggraph) 22*, 3 (July 2003), 376–381. 2

[ODJ04] OSTROMOUKHOV V., DONOHUE C., JODOIN P.-M.: Fast hierarchical importance sampling with blue noise properties. *ACM Transactions on Graphics (Proc. Siggraph) 23*, 3 (Aug. 2004), 488–495. 2

[RH01] RAMAMOORTHI R., HANRAHAN P.: An efficient representation for irradiance environment maps. In *Proc. of ACM Siggraph '01* (2001), pp. 497–500. 2

[RH02] RAMAMOORTHI R., HANRAHAN P.: Frequency space environment map rendering. In *Proc. of ACM Siggraph '02* (2002), pp. 517–526. 2

[SG92] SMITH A., GELFAND A.: Bayesian statistics without tears: a sampling-resampling perspective. vol. 46, pp. 84–88. 6

[Shi00] SHIRLEY P.: *Realistic Ray Tracing*. A K Peters, Natick, 2000. 3, 4

[SHS02] SECORD A., HEIDRICH W., STREIT L.: Fast primitive distribution for illustration. In *Eurographics Workshop on Rendering* (June 2002), pp. 215–226. 2, 5, 6

[SKS02] SLOAN P., KAUTZ J., SNYDER J.: Precomputed radiance transfer for real-time rendering in dynamic, low-frequency environments. In *Proc. of ACM Siggraph '02* (2002), pp. 527–536. 2

[SPS95] SLUSALLEK P., PFLAUM T., SEIDEL H.-P.: Using procedural renderman shaders for global illumination. In *Proc. of Eurographics* (1995), vol. 14, pp. 311–324. 3

[SSSK04] SZECSI L., SBERT M., SZIRMAY-KALOS L.: Combined correlated and importance sampling in direct light source computation and environment mapping. In *Proc. of Eurographics* (2004), vol. 23, pp. 585–593. 3

[Sta01] STANFORD GRAPHICS GROUP: Digital Michaelangelo project, 2001. http://graphics.stanford.edu/projects/mich/. 7

[Tan96] TANNER M.: *Tools for statistical inference*, 3rd ed. Springer Verlag, New York, 1996. 6

[VG95] VEACH E., GUIBAS L.: Optimally combining sampling techniques for monte carlo rendering. In *Proc. of ACM Siggraph '95* (Aug. 1995), pp. 419–428. 3, 4, 7

Figure 7: *Quality comparison of our method against standard importance sampling for the same compute time. Left image pair: illumination from an environment map. Right image pair: illumination from an area light source. Left column: traditional importance sampling from the light source. Right column: bidirectional importance sampling.*
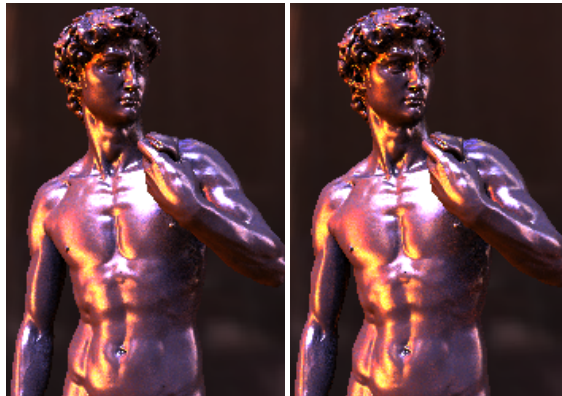

Figure 8: *Quality comparison between our two proposed bidirectional sampling methods. Left: Rejection sampling. Right: Sampling-importance resampling (SIR).* 176×248 *images computed in* 13.0 *seconds using* $15/800$ *rejection and SIR samples.*