EUROGRAPHICS 2012 / P. Cignoni, T. Ertl
(Guest Editors)

Volume 31 (2012), Number 2

Importance Caching for Complex Illumination

Iliyan Georgiev[†1] Jaroslav Křivánek[‡2] Stefan Popov[†1] Philipp Slusallek[†1,3]

[1]Saarland University and Intel VCI, Saarbrücken [2]Charles University, Prague [3]DFKI, Saarbrücken

Abstract

Realistic rendering requires computing the global illumination in the scene, and Monte Carlo integration is the best-known method for doing that. The key to good performance is to carefully select the costly integration samples, which is usually achieved via importance sampling. Unfortunately, visibility is difficult to factor into the importance distribution, which can greatly increase variance in highly occluded scenes with complex illumination.
In this paper, we present importance caching – a novel approach that selects those samples with a distribution that includes visibility, while maintaining efficiency by exploiting illumination smoothness. At a sparse set of locations in the scene, we construct and cache several types of probability distributions with respect to a set of virtual point lights (VPLs), which notably include visibility. Each distribution type is optimized for a specific lighting condition. For every shading point, we then borrow the distributions from nearby cached locations and use them for VPL sampling, avoiding additional bias. A novel multiple importance sampling framework finally combines the many estimators. In highly occluded scenes, where visibility is a major source of variance in the incident radiance, our approach can reduce variance by more than an order of magnitude. Even in such complex scenes we can obtain accurate and low noise previews with full global illumination in a couple of seconds on a single mid-range CPU.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Raytracing, Radiosity; G.3 [Mathematics of Computing]: Probability and Statistics—Probabilistic algorithms (including Monte Carlo)

1. Introduction

Many rendering applications, such as architectural visualization, require computing global illumination in the scene. Solving this problem involves costly evaluation of multi-dimensional integrals of complex functions. While the utility of Monte Carlo ray tracing algorithms in such problems has been widely demonstrated, the stochastic nature of the approach entails a rather slow convergence. This issue can be alleviated by using variance reduction techniques, such as importance sampling. However, effectively capturing the important features of the integrands of concern remains a challenging problem, which we address in this work in the context of virtual point light (VPL) rendering [Kel97].

To this end, the state-of-the-art unbiased importance sampling methods most often construct an explicit representation of a probability density function (PDF) for sampling incident illumination. Ideally, this PDF would be the product of the individual terms under the reflection integral. In practice, not all of these terms are known, so the PDF often includes only some of them, visibility in particular being commonly omitted [CJAMJ05, BGH05, WA09]. Moreover, most importance sampling methods do not take advantage of the fact that illumination is often piece-wise smooth over surfaces, and build a PDF for each shading point independently, which can be too expensive. As an alternative, (ir)radiance caching based approaches perform accurate lighting computations at isolated points and interpolate the values in between [WRC88, KGPB05]. This aggressive interpolation can bring large efficiency gains, alas at the cost of bias.

In this paper, we propose a method for efficient importance sampling of the incident illumination at a point. Our approach tries to sample proportionally to the actual local contributions of all VPLs distributed in the scene. It maintains efficiency by using an importance caching scheme that exploits coherence without introducing additional bias. First, at a sparse set of locations in the scene we evaluate accurate

† e-mail: {georgiev, popov, slusallek}@cs.uni-saarland.de
‡ e-mail: jaroslav.krivanek@mff.cuni.cz

lighting, and store several types of probability distributions derived from these evaluations. The distributions, each optimized for a specific lighting condition, are then used for sampling VPL contributions at other locations. We finally combine the many resulting estimators using a novel bilateral multiple importance sampling framework that uses a new aggressive weighting heuristic.

Our method handles direct and indirect illumination simultaneously, and the sampling distributions consider all terms of the reflection integrand, including visibility. We demonstrate that this can deliver significant variance reduction in occluded scenes with complex lighting, where visibility is often a major source of variance. Since our approach does not introduce bias, high-quality results can be obtained progressively with a fixed memory footprint. We also show that a simplified version can produce highly accurate low-noise previews interactively even in such complex scenes.

2. Related Work

The goal of global illumination algorithms is to compute the reflected radiance L^o at every shading point, usually resorting to Monte Carlo integration with PDF p:

$$\begin{aligned} L^o &= \int_S B(r)L(r)G(r)V(r)\,dA(r) \\ &\approx \frac{1}{N}\sum_{k=1}^{N}\frac{B(R_k)L(R_k)G(R_k)V(R_k)}{p(R_k)} = \frac{1}{N}\sum_{k=1}^{N}\frac{f(R_k)}{p(R_k)}, \end{aligned} \tag{1}$$

where S is the set of all scene surface points ($R_k \in S$), including light sources, and B, L, G, and V are the BRDF, incident radiance, geometric, and visibility terms between the point and R_k. If shared among all shading points, the radiance samples R_k are often called virtual point lights (VPLs) [Kel97].

Importance sampling. Ideally, we would sample with the PDF p proportional to the integrand f. However, obtaining such a PDF in practice is unfeasible, particularly because the visibility function V is not given in a closed form, and/or the incoming radiance may not be known.

The state-of-the-art importance sampling methods build distributions that include some of the terms in f, usually the product BL, for sampling environment [CJAMJ05, CETC06, CAM08b] and indirect [WA09] illumination from VPLs. Rouselle et al. [RCL*08] include a conservative visibility term. Importance resampling methods [BGH05, TCE05] use approximate product distributions, and generally can handle direct and indirect illumination simultaneously.

The photon map has been successfully used for illumination importance sampling [Jen95, Hei01, Pha05, SL06], These methods sample ray directions from distributions derived from the photons cached at nearby shading points.

All above approaches build the distribution for each shading point independently. In contrast, the distributions we use are shared among points, and include all integrand terms.

A single distribution is often not a good match for the whole integrand, therefore samples can be drawn from multiple distributions. Veach [Vea97] developed multiple importance sampling (MIS) as a framework for combining different distributions in one estimator. We sample from multiple distributions and propose a novel MIS weighting heuristic.

VPL methods. VPL rendering was introduced by Keller [Kel97]. Some methods use adaptive clustering to render the many VPLs necessary to handle scenes with complex lighting. Lightcuts [WFA*05, WABG06] compute an error-bound pixel estimate from a cut in a VPL tree. The cut is chosen by bounding the tree cluster contributions locally, assuming full visibility. Hašan et al. [HPB07] compute a global VPL clustering for the whole image that groups lights with similar contributions, including visibility. Similarly to our approach, Ou et al. [OP11] further refine the clustering on small surface regions in order to capture the local VPL importance. Interactive VPL methods achieve efficiency through aggressive visibility approximations and illumination interpolation, coupled with an optimized GPU implementation [RGK*08, DGR*09, REH*11].

Importance-driven VPL generation has recently received attention [WBS03, SIMP06, SIP07, GS10]. These methods sample VPLs based on their total image contribution, but do not adapt to individual locations. This is why they can be inefficient in scenes with complex non-uniform lighting. We use the method of Georgiev and Slusallek [GS10] as an initial global VPL resampling step, and apply our new importance sampling strategy on the resulting VPL set.

Exploiting coherence. Taking advantage of inter-integral correlation is another way to increase rendering efficiency. Irradiance caching [WRC88] and radiance caching [KGPB05, GKB09] compute precise illumination estimates at a set of adaptively chosen locations in the scene, and interpolate (ir)radiance at points outside this set. Kontkanen et al. [KRK04] extend the approach to adaptive irradiance filtering over surfaces. These methods assume smooth illumination, and rely on geometric features to discover discontinuities. Křivánek et al. [KBPv06] improved the caching and extrapolation strategies of (ir)radiance caching. However, since caching is based on adaptive supersampling, strong illumination features can still be missed. This adds bias to the solution which is impossible to prevent.

On top of product importance sampling of BRDF and environment map, Clarberg et al. [CAM08a] add a control variate term that includes interpolated visibility. This reduces noise in occluded regions, but since the sampling distribution remains unchanged, the effect of variance reduction is limited. Finally, table-driven adaptive importance sampling [CAE08] reduces variance by mixing standard importance sampling with shared importance information across neighboring pixels. While this mixture is conservative, it still increases variance at discontinuities, which is then clamped in a biased way. Our method addresses this problem by sampling from distributions that remain robust at discontinuities.

3. Algorithm Overview

Sampling and integrating the incident illumination can be too expensive to perform for each pixel individually. We take a common approach to amortizing the path sampling cost by sharing one set of light sub-paths among all pixel integrals. Before rendering, we sample the light sources and trace light paths in the scene, creating a set of *virtual point lights* (VPLs) that approximate the direct and indirect radiance distribution in the scene [Kel97, WFA*05]. We additionally use the method of Georgiev and Slusallek [GS10] for initial global VPL resampling, which equalizes the total camera contribution of all VPLs, discarding those that are irrelevant for the current viewpoint.

After distributing N_{VPL} VPLs R_k with corresponding PDFs $p(R_k)$, rendering an image reduces to summing up the contribution of all VPLs at every shading point x:

$$L_x^o \approx \sum_{k=1}^{N_{VPL}} B_x(R_k)L_x(R_k)G_x(R_k)V_x(R_k) = \sum_{k=1}^{N_{VPL}} f_x(R_k), \quad (2)$$

where the subscript x denotes evaluation at point x. Here, $L_x(R_k)$ includes the normalization factor $1/(p(R_k)N_{VPL})$.

Equation (2) can be prohibitively expensive to evaluate for the common case where the number of VPLs is on the order of thousands or more. Our solution is based on an unbiased Monte Carlo estimation of the sum in Equation (2):

$$\widetilde{L}_x^o = \frac{1}{n}\sum_{m=1}^{n}\frac{f_x(R_m)}{p_x(R_m)} \approx \sum_{k=1}^{N_{VPL}} f_x(R_k) \approx L_x^o, \quad (3)$$

evaluating the contribution of a small number of $n \ll N_{VPL}$ VPLs, chosen from a probability distribution p_x over R_k. Making p_x exactly proportional to f_x can result in a zero-variance estimate, which can be obtained with just one VPL per pixel. This, however, can be very difficult to achieve, especially because p_x has to include the visibility term, the evaluation of which involves the costly ray casting operation.

A common approach to decreasing the distribution construction cost is to exclude visibility from it [BGH05, CJAMJ05]. Unfortunately, as Figure 1 demonstrates, doing so in occluded scenes can destroy proportionality and lead to dramatic variance increase. Moreover, the total rendering time is still dominated by the distribution construction.

Following these observations, we propose a solution that (1) samples VPLs with a distribution that accounts for visibility, (2) avoids the costly distribution construction at each shading point, and (3) remains robust at discontinuities.

3.1. Importance Caching

Our approach is based on the idea to take a set of *evaluated* VPL contributions at one location and reuse them in the form of *importance* at other locations. For each frame, we first perform full VPL evaluation at a sparse set of surface points I_j generated by tracing random rays from the camera. Each of these *importance records* (IRs) stores the contributions $f_j(R_k)$ of all VPLs R_k at its position. This set of contributions f_j is then normalized to a probability distribution $\overline{f}_j$ for sampling VPLs at nearby shading points. $\overline{f}_j$ represents the *fully* evaluated integrand, including visibility, and allows for exploiting illumination coherence in an unbiased way. Thus, instead of constructing costly and only approximate distributions at each shading point, we build perfect distributions at a small number of points – the IRs. They are stored in an easy to build one-dimensional cumulative distribution function (CDF) form. The distribution $\overline{f}_j$ at each IR provides perfect importance sampling of VPL contributions at its associated location I_j (see Figure 1).

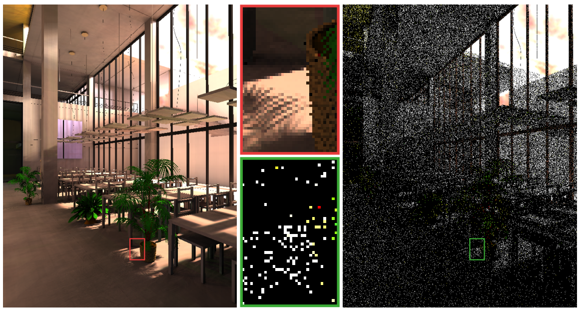

Figure 1: *A scene illuminated by an environment map, with radiance field approximated by 6000 VPLs. Using only one VPL per pixel, chosen from a distribution proportional to its actual contribution, would result in perfect importance sampling, giving the exact solution (left). Excluding visibility from the distribution leads to excessive variance (right).*

During final rendering, at each shading point x we first find the M nearest records I_j with shortest distance according to the metric $d(x, I_j) = ||x - I_j|| + \lambda\sqrt{1 - N_x \cdot N_{I_j}}$, where N_x and N_{I_j} are the surface normals at x and I_j, respectively. Similar metrics are common in illumination interpolation methods [WRC88], where λ, trades off the importance of Euclidean and orientation distance. The IRs are organized in a range-search kd-tree. Lookups are performed in a two-step filtering approach, similar to [CAM08a]. We use $M = 3$ and $\lambda = 0.5/D$, where D is the scene's bounding box diagonal.

We use the distributions from the closest IRs for evaluating $\widetilde{L}_x^o$ (Equation (3)). This way, the pre-sampling setup for x is reduced to a nearest neighbor search. Moreover, since choosing a VPL involves only a CDF binary search, we effectively exploit the whole importance information stored at the IRs, but access just a small fraction of the actual data.

When reused at a spatially close location x, we can expect that, while not perfect anymore, $\overline{f}_j$ will most often be closely proportional to f_x, leading to robust importance sampling. However, proportionality decreases at discontinuities, in turn increasing variance. We combat this problem in two ways.

We first ensure that sampling remains robust when reusing distributions across illumination discontinuities. In such cases $\overline{f}_j$ might assign low, and even zero, probabilities to important VPL contributions at x. Therefore, in addition to

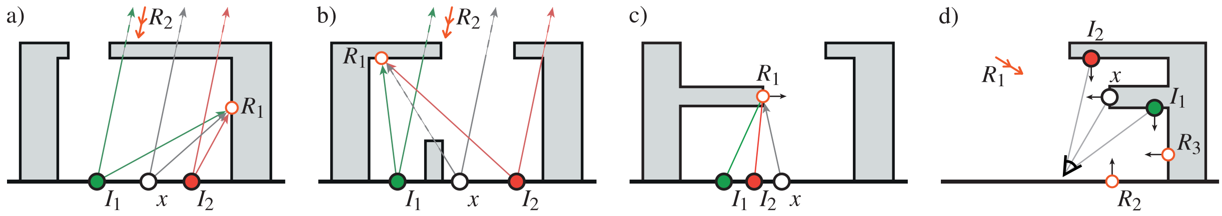

Figure 2: *Four illumination conditions, encountered when reusing information from importance records (IRs) I_1 and I_2 at shading point x. At each IR we define four distributions, designed to discover VPL contributions under a different condition. a) In the case of smooth illumination in the local neighborhood, full contribution sampling ($\mathcal{F}$) can achieve close proportionality to the integrand. b) Unoccluded contribution sampling ($\mathcal{U}$) is robust to VPL contribution changes due to varying occlusion with position. c) Bounded contribution sampling ($\mathcal{B}$) in addition discovers new contributions due to orientation changes. d) Conservative uniform sampling ($\mathcal{C}$) handles situations where the IR importance information is irrelevant at the shading point x.*

this aggressive distribution, at each IR I_j we build three increasingly conservative distributions derived from f_j that are less likely to miss important new VPL contributions in the vicinity of I_j. We design these sampling techniques by identifying the situations that cause changes in VPL contribution due to variations in position and orientation (Section 4).

Second, we combine the many distributions borrowed from nearby importance records around each shading point, such that salient information is extracted from the individual estimates. We perform a bilateral multiple importance sampling combination using a novel α-max heuristic to weight distributions based on a specified prioritization (Section 5).

Our method does not distinguish between the different kinds of VPLs, e.g. finite and infinite, direct and indirect [WFA*05, HPB07]. This way we effectively importance sample the *total* incident illumination at the particular point. In addition, thanks to the unbiased exploitation of coherence, we can discard all cached data at the end of each frame, and obtain a high-fidelity solution via progressive averaging. Convergence is thus achieved with a fixed memory footprint.

4. Sampling Distributions

Recall that each importance record (IR) I_j stores the normalized local contributions $\bar{f}_j(R_k)$ of all VPLs R_k. Given a nearby shading point x, if the illumination is locally smooth, it is likely that $\bar{f}_j$ will be closely proportional to f_x, leading to robust importance sampling. However, in regions around illumination discontinuities there can be little or no correlation, which may result in excessive variance. This happens mainly due to high-energy regions in f_x not being present in f_j and consequently sampled with low probabilities.

We identify the causes for VPL contribution changes between surface points, and at each IR build four increasingly conservative distributions. They help the sampling remain robust when reusing importance across discontinuities.

$\mathcal{F}$**: Full contribution sampling.** As discussed above, the most straightforward distribution to define at each IR I_j is the normalized $\bar{f}_j$. This full contribution distribution (Figure 2a), denoted $\mathcal{F}$, is our most aggressive one and also the one with highest variance reduction potential. It can in fact achieve perfect proportionality, e.g. on flat unoccluded diffuse surfaces illuminated by distant light sources, where f_x is independent of position x. $\mathcal{F}$ often discovers the largest fraction of the reflected radiance, as we show in Section 6.

$\mathcal{U}$**: Unoccluded contribution sampling.** It sometimes happens that all I_j in a region agree on the importance of a particular R_k, though inconsistently with its contribution at x. In scenes with occlusion, these inconsistencies are often caused by a change in the visibility function, i.e. $V_x(R_k) \neq V_j(R_k)$. False positives (e.g. R_1 in Figure 2b) increase variance but are usually not too problematic, as their contribution at x is zero. On the other hand, if false negatives (R_2 in Figure 2b) are sampled with low probability, variance can explode.

The unoccluded contribution distribution ($\mathcal{U}$) is more conservative than $\mathcal{F}$, and is designed to discover VPL contributions potentially missed due to occlusion. The distribution is built from the VPL contributions at I_j, but with $V_j(R_k) = 1$, i.e. from the unoccluded $f_j^u(R_k) = B_j(R_k)L_j(R_k)G_j(R_k)$. The technique is particularly robust at finding small-scale illumination features, such as a bright spot produced by a thin light beam passing through a small window (see Figure 7).

$\mathcal{B}$**: Bounded contribution sampling.** Sometimes all I_j may falsely suggest a low contribution of R_k at x due to differences in surface orientation, i.e. $G_x(R_k) > G_j(R_k)$. Figure 2c shows an example of such inconsistency. The actual contribution $f_x(R_k)$ might in fact be small, but if sampled with a low probability, variance can increase dramatically.

The bounded contribution distribution ($\mathcal{B}$) is a conservative extension of $\mathcal{U}$ that targets orientation-induced false negatives. Knowing that each IR will be used in a small neighborhood, we build the $\mathcal{B}$ distribution from the contributions $f_j^b(R_k) = B_j(R_k)L_j(R_k)G_j^{max}(R_k)$. Here, $G_j^{max}(R_k)$ is the upper-bound geometric term for points inside the region of influence of I_j, computed as described in Appendix A. As noted above, $\mathcal{B}$ usually finds features with a rather small contribution to the outgoing radiance. It does so, however, with high enough probabilities that avoid excessive variance.

$\mathcal{C}$**: Conservative sampling.** When the nearest IRs around x are at a large distance in position and orientation, false negatives can occur that $\mathcal{F}, \mathcal{U}$ and $\mathcal{B}$ do not handle. An example is shown in Figure 2d. Such situations are most notable when strong illumination falls on a surface that has a small screen-space footprint and unique orientation in its neighborhood. Since no useful data can be reliably extracted from any nearby IR, we handle such rare cases with the most conservative uniform sampling distribution ($\mathcal{C}$) over all VPLs.

5. Bilateral Combination of Sampling Distributions

To produce a final image, at each shading point x we use the four distributions at each of the M nearest importance records (IRs) for VPL sampling. We combine all distributions, transforming $\widetilde{L}^o_x$ (Equation (3)) into a multiple importance sampling (MIS) estimator. While MIS was originally developed for continuous distributions, it can be shown to work in the discrete case too. In order to avoid bias, we must ensure that every VPL R_k is chosen with a non-zero probability, however all $\mathcal{F}, \mathcal{U}$ and $\mathcal{B}$ can have zeros. Fortunately, MIS does not require each distribution to be non-zero everywhere [Vea97]. Thus, including $\mathcal{C}$ is sufficient to guarantee the unbiasedness of the resulting combined estimator.

Having all $4M$ distributions, we can construct a MIS estimator by choosing a weighting heuristic [Vea97]. Figure 3 shows a matrix arrangement of the distributions at the $M = 3$ nearest IRs around a point. As expected, the M distributions in a row, each corresponding to an IR, correlate closely, while the four increasingly conservative distributions in each column have different structure. Veach's MIS heuristics are solely based on probabilities, hence unable exploit this additional information that may hint to when particular distributions achieve closer proportionality. Choosing one of these heuristics to weight all distributions can thus be sub-optimal.

We perform a bilateral two-stage combination of the distributions in the matrix, which uses different weighting schemes for the rows and columns, in order to better preserve the qualities of each distribution. We first construct a MIS estimator that mixes the M distributions in each row. This way, we conceptually collapse all columns into one. The four distributions in this resulting column are subsequently combined to form the final MIS estimator.

5.1. Column Combination

The estimator that combines all M distributions in row i using a column weighting heuristic $w^{col}_{i,j}$ reads:

$$\widetilde{L}^o_{x,i} = \sum_{j=1}^{M} \frac{1}{n_i} \sum_{k=1}^{n_i} w^{col}_{i,j}(R_{i,j,k}) \frac{f_x(R_{i,j,k})}{p_{i,j}(R_{i,j,k})}, \qquad (4)$$

where $p_{1,j} = \mathcal{F}_j$, $p_{2,j} = \mathcal{U}_j$, $p_{3,j} = \mathcal{B}_j$, $p_{4,j} = \mathcal{C}_j$, taking the same number of samples n_i from each distribution in row i.

Since we do not have reliable means to detect all discontinuities, shadows in particular, we cannot determine which IRs correlate most with the shading point x. Therefore, we

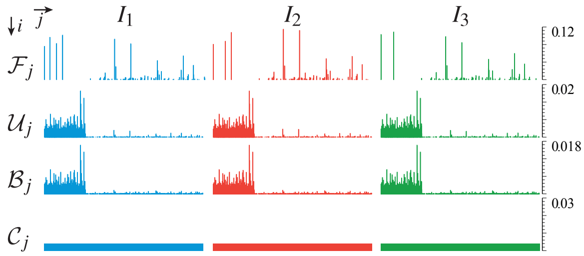

Figure 3: *A matrix arrangement of the four distributions at the three closest records to the red-marked point in the left image in Figure 7. Notice the correlation among the columns. Direct illumination VPLs are in the beginning of the VPL list, hence the higher probabilities in the left.*

avoid aggressive weighting for combining the columns j and take the safest available heuristic – the balance heuristic: $w^{col}_{i,j}(R) = p_{i,j}(R)/\sum_{l=1}^{M} p_{i,l}(R)$.

The balance heuristic corresponds to directly sampling from a mixture distribution: $p_i(R) = \frac{1}{M}\sum_{l=1}^{M} p_{i,l}(R)$ [Vea97]. (In our case this can also be interpreted as importance interpolation at the shading point x.) To achieve this, we construct the mixture CDF corresponding to p_i implicitly while sampling, during a synchronous traversal of the individual CDFs by averaging their elements. The direct mixture sampling requires M times fewer random numbers and facilitates stratification. Equation (4) now simplifies to a standard estimator:

$$\widetilde{L}^o_{x,i} = \frac{1}{n_i} \sum_{k=1}^{n_i} \frac{f_x(R_{i,k})}{p_i(R_{i,k})}.$$

5.2. Row Combination

In order to obtain the final estimator, we combine the four row estimators $\widetilde{L}^o_{x,i}$ using a row weighting heuristic w^{row}_i:

$$\widetilde{L}^o_x = \sum_{i=1}^{4} \frac{1}{n_i} \sum_{k=1}^{n_i} w^{row}_i(R_{i,k}) \frac{f_x(R_{i,k})}{p_i(R_{i,k})}. \qquad (5)$$

For w^{row}_i the balance heuristic is not always the best choice. Each of the $\mathcal{F}$, $\mathcal{U}$, $\mathcal{B}$, and $\mathcal{C}$ techniques is designed to be locally the most proportional under different conditions, i.e. to make the problem *low-variance* [Vea97]. Figure 4 left illustrates that averaging a locally proportional and a uniform distribution can destroy the qualities of both over the whole domain. In such cases, more aggressive heuristics can perform better, e.g. the power or max heuristics [Vea97]. They weight techniques proportionally to their probabilities under the assumption that higher probabilities result in lower variance. While this approach avoids extreme variance effectively, it can still destroy local proportionality. Notice in Figure 4 right that the uniform distribution p_u has higher probability around the borders in the middle region, yet p alone results in a zero-variance estimator for the region.

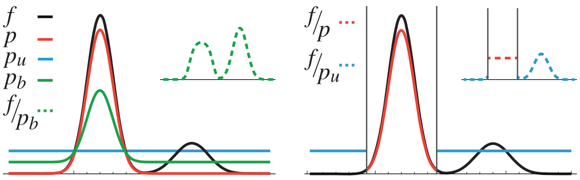

Figure 4: *Left: The balance heuristic mixture p_b ruins the local proportionality of the aggressive distribution p to f by averaging it with the defensive uniform distribution p_u, while still resulting in lower than optimal probabilities outside the peak. Right: By adaptively partitioning the domain, the novel α-max heuristic avoids mixing, selecting the distribution with closest proportionality in a user-controlled way.*

The α-max heuristic. We often have additional knowledge about the distributions, e.g. we expect a certain distribution to achieve closer proportionality more often than others. We have designed the α-max heuristic to exploit this knowledge:

$$w_s^{\alpha}(x) = \begin{cases} 1, & w_i^{\alpha}(x) = 0, \ \text{for } 1 \le i < s, \ \text{ and} \\ & p_s(x) \ge \max\limits_{s<i\le m} \alpha_i p_i(x), \\ 0, & \text{otherwise}, \end{cases} \tag{6}$$

where x is sampled from p_s, and each of the m distributions p_i has an associated priority index i and confidence value $\alpha_i \in (0;1]$. Note that setting all $\alpha_i = 1$ gives Veach's max heuristic, while α_1 is redundant, thus we always set it to 1.

The α-max heuristic partitions the sampling domain by assigning distributions to those parts of the domain where their probability is larger than all lower-priority distributions scaled by their confidence values. This controlled partitioning can better preserve proportionality if we have some a-priori information about the distributions. This can in turn result in more effective variance reduction than the original Veach heuristics. Figure 4 right demonstrates this, with ordering $(\alpha_p, \alpha_{p_u}) = (1, 0.2)$. It also illustrates how α-max with α_{p_u} as a tradeoff parameter can be used as an alternative to defensive importance sampling [Hes95], which also suffers from mixture proportionality losses. These can be remedied effectively by α-max, as we also show in Section 6.

For the sampling techniques from Section 4 we use the ordering $\mathcal{F}, \mathcal{U}, \mathcal{B}, \mathcal{C}$. This prioritization follows the increasing distribution conservativeness, setting highest priority to $\mathcal{F}$.

5.3. Distribution Optimization

Veach [Vea97] argues that zero-weight heuristics waste computation on samples that then get ignored. However, α-max can actually be more computationally efficient than, e.g., the power heuristic. Since weights are based on cheap to compute sampling probabilities, we can avoid the costly f_x evaluation by directly discarding zero-weight samples.

This observation can be further exploited to avoid generating samples that would end up being discarded in the first place. This can be achieved by reducing the redundancy, i.e. overlap, of the distributions. We directly modify them at each I_j as they are constructed. For every VPL R_k and $p_{i,j}$ (corresponding to $\mathcal{F}_j$, $\mathcal{U}_j$, $\mathcal{B}_j$, $\mathcal{C}_j$), we multiply $p_{i,j}(R_k)$ by $w_{i,j}^{\alpha}(R_k)$ before accumulating it into its CDF. This way, we partition the sampling domain, so that exactly one of the four distributions at I_j is non-zero for any VPL. This in turn increases probabilities after re-normalization. Note that since the optimization is performed at each record independently, it does not eliminate all redundancies in the mixtures used for rendering. It also slightly alters them, which is not a problem in practice, as α-max is still evaluated during rendering.

6. Results

We implemented the importance caching method in a standard ray tracer, and performed tests on a mid-range 4-core Intel Core i7-860 CPU. The images in each comparison are rendered in equal time of 20 sec at 1024×768 resolution.

Technique comparison. Figure 7 shows the STUDY HALL scene illuminated by the high frequency St. Peter's Cathedral environment map, rendered with full global illumination using different methods. This scene allows us to test and compare all sampling distributions and combination strategies, as it synthesizes a variety of lighting conditions. These include smooth direct and indirect illumination, small-scale illumination and geometric features, and occlusion.

The image labels in Figure 7 indicate the distributions and weighting heuristics used. β denotes the power heuristic [Vea97, p. 273] with $\beta = 2$; d – defensive importance sampling with 0.4 uniform weight; α – the α-max heuristic with $(\alpha_{\mathcal{F}}, \alpha_{\mathcal{U}}, \alpha_{\mathcal{B}}, \alpha_{\mathcal{C}}) = (1, 0.5, 0.5, 0.3)$ (Equation (6)).

Overall, full contribution sampling ($\mathcal{F}$) performs remarkably well in smooth illumination regions, as it exploits any available coherence, sampling the whole incident illumination with close proportionality, resulting in a low-variance estimate. This proportionality, however, is ruined when mixing $\mathcal{F}$ with other distributions. The power heuristic and especially defensive importance sampling increase variance over α-max in such regions by 45% and 330% respectively.

When $\mathcal{F}$ misses important contributions, it is often due to changes in occlusion. While $\mathcal{C}$ does not provide high enough sampling probabilities, $\mathcal{U}$ is robust at discovering new unoccluded VPLs, as seen when comparing $\mathcal{FC}^{\alpha}$ and $\mathcal{FUC}^{\alpha}$.

The $\mathcal{FUBC}^{\beta}$ and $\mathcal{FUBC}^{\alpha}_{unopt}$ images do not use distribution optimization. As a result, $\mathcal{B}$ cannot robustly capture the orientation-induced false negatives, as it is too conservative initially. After optimization and re-normalization, the average increase in probability for $\mathcal{F}$, $\mathcal{U}$, $\mathcal{B}$, $\mathcal{C}$ is respectively 0.05%, 20%, 454%, 50%. $\mathcal{B}$ benefits significantly mostly due to the removed redundancies, an effect of overlap with $\mathcal{U}$ (see Figure 3). The best result with low-variance smooth regions and no spikes is then produced by $\mathcal{FUBC}^{\alpha}$.

The benefit of including occlusion in the VPL sampling can be seen when comparing $\mathcal{FUBC}^{\alpha}$ against $\mathcal{UBC}^{\alpha}$ which increases variance by up to $10\times$. We also compare against resampled importance sampling (RIS), as well as against uniform VPL sampling (i.e. $\mathcal{C}$ alone) which produces a similar result to path tracing. For RIS, we first uniformly pick S candidates at a point and shade them without visibility, and then draw $S/20$ samples from the on-the-fly constructed CDF. Even though $\mathcal{UBC}^{\alpha}$ relies on sometimes inaccurate borrowed distributions, it is still superior to RIS, as it carries importance information for *all* VPLs, avoiding the initial blind RIS sampling. Increasing the RIS candidate count reduces performance without improving quality noticeably.

Figure 8 visualizes the fractional contributions of the individual distributions to the reference image in Figure 7. $\mathcal{F}$ achieves close proportionality almost everywhere. As a result, it captures about 94% of the illumination. $\mathcal{U}$ and $\mathcal{B}$ discover almost all that has been missed (3% each), leaving a very small fraction to $\mathcal{C}$. Moreover, $\mathcal{U}$, $\mathcal{B}$ and $\mathcal{C}$ together require $6\times$ more samples than $\mathcal{F}$ due to their increasing conservativeness (which is slightly ameliorated by the optimization), while having rather insignificant overall contribution.

Pre-sampling, rendering and performance. The 20-second comparison images were rendered by progressively averaging 3 frames, each using 2700 importance records and 8000 VPLs per frame. Table 1 summarizes the average fractional time spent in different steps of $\mathcal{FUBC}^{\alpha}$ and RIS.

Numerical convergence. Figure 5 shows the ability of our method to simultaneously handle different types of radiance samples, as well as many light sources. It also compares the convergence of path tracing (PT), resampled importance sampling (RIS) and importance caching (IC) for direct and full global illumination from environment and area lights.

The time plots (in seconds) show the root mean squared error (RMSE). Due to the necessary geometric clamping of indirect VPLs, the path tracer has been slightly modified so that all three methods converge to the same biased solution on the right. In both cases, IC consistently outperforms PT and RIS with an average variance reduction of $9\times$ and $4.8\times$ respectively for the left and $25\times$ and $9\times$ for the right images.

Glossy materials. Figure 6 shows three Buddha statues with increasingly specular Phong materials, illuminated by the St. Peter's environment map. We compare against resampled importance sampling (RIS) with the same configuration as for the STUDY HALL scene, as well as to BRDF and illumination sampling combined with the power heuristic (MIS).

This is a worse case scenario for our method, due to the lack of occlusion and less available coherence to exploit, a result of the high frequencies in geometry, BRDFs and illumination. We still notice that in occluded regions and on low-to-mid frequency reflection or geometry distributions, IC outperforms the other two methods, producing very smooth results. However, with narrow glossy BRDF lobes, even slight orientation changes can render the importance record (IR) data invalid at nearby shading points. In such situations the RIS and MIS algorithms perform better than IC.

Based on the above observations, we developed the following combination of IC and RIS. We first compute the average "BRDF distance" between point x and its M nearest records I_j: $d_x = 1 - \frac{1}{M}\sum_{j=1}^{M} \frac{f_r(\omega_x^o, x, \overline{\omega_j^o})}{f_r(\omega_x^o, x, \overline{\omega_x^o})}$, where ω^o and $\overline{\omega^o}$ denote an outgoing radiance direction and its reflection, respectively. This metric captures orientation differences between specular lobes, while treating the BRDFs as black boxes. If d_x is below 0.5 we compute the outgoing radiance using IC, otherwise with RIS. The combined IC+RIS algorithm (Figure 6 top right) preserves the qualities of both IC and RIS by adaptively detecting the better performing one at the particular shading point. Note that noise is further reduced on the statues, as now RIS can take more samples in the given time.

High-quality preview rendering. The results in Figure 8 suggest that we can obtain a quick, highly accurate, and low-variance solution using a small number of samples drawn only from $\mathcal{F}$. The implementation then can be simplified to storing only one distribution at each IR, and evaluating the α-max heuristic only w.r.t. $\mathcal{C}$, effectively discarding samples with probabilities below a constant threshold. While this introduces some bias, a unique feature of our method is that, if desired, it can be compensated for in a controlled way by selectively including $\mathcal{U}$, $\mathcal{B}$ and $\mathcal{C}$ in the sampling process.

Figure 9 demonstrates the high quality achieved in just 2 seconds (0.5 FPS) using the described method – IC $\mathcal{F}^{\alpha}$, on the STUDY HALL and SPONZA scenes, compared to RIS and PT, as well as to classic instant radiosity [Kel97]. We also show direct importance visualization (DIV) at each shading point, which is equivalent to irradiance interpolation. DIV is only slightly faster than IC $\mathcal{F}^{\alpha}$, since the latter samples only 4 VPLs per pixel, which now takes only about 30% of the frame time. Although noise-free, DIV blurs away all fine details, while instant radiosity suffers from severe aliasing. On the other hand, RIS and PT produce very noisy images.

7. Discussion

Although our sampling framework is unbiased, the use of VPLs incurs a small systematic error when geometric clamping is used. In addition, the assumption that the scene radiance field can be well approximated by a small set of radiance samples generated from the light sources becomes invalid in the presence of high-frequency BRDFs with glossy inter-reflections. Handling such materials with VPLs is a general problem, although solutions exist [KK04, HKWB09, DKH*10]. Additionally, the number of VPLs can be efficiently increased by using a spatial hierarchy [WFA*05].

Our implementation uses uniform distribution of ISs over the image plane. We experimented with distributions that adapt to geometric discontinuities. However, this caused a decrease in IR density around sharp discontinuity regions, e.g. shadow penumbrae, increasing variance noticeably. A better adaptive strategy would be to utilize actual statistics from previous frames, which we leave for future work.

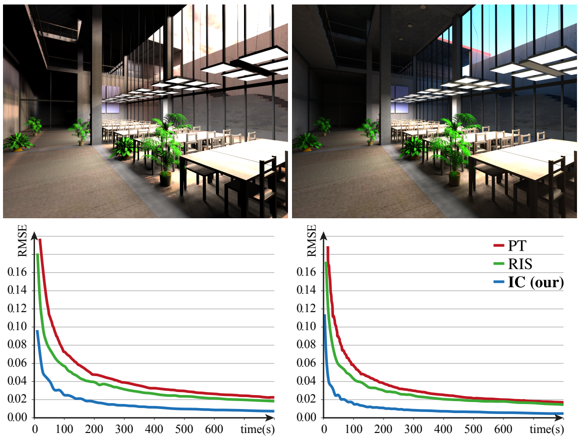

Figure 5: *Top row: The* STUDY HALL *scene illuminated by 36 area lights and an environment map. Left: Direct illumination from St. Peter's map. Right: Full global illumination from a sunset map. Bottom row: total image RMSE time convergence plots of path tracing (PT), resampled importance sampling (RIS) and importance caching (IC).*

	Preprocess		Rendering			
	VPL	CDF	Sampl.	Shad.	RT	Other
IC	**3%**	**14%**	**30%**	**3%**	**42%**	**8%**
RIS	0%	-	3%	70%	24%	3%

Table 1: *Fractional time break-down for the $\mathcal{FUBC}^{\alpha}$ and RIS images in Figure 7 rendered in 20 sec with importance caching (IC) and resampled importance sampling (RIS).*

To render a frame, our algorithm needs $N_{VPL} \times N_{IR} \times 4 \times$ sizeof(float) bytes of memory, where N_{VPL} and N_{IR} are the number of VPLs and importance records (IRs). For the tests in figures 5, 6 and 7 this amounted to about 330MB. Memory consumption can be decreased by splitting the image into tiles which are rendered independently. Alternatively, N_{VPL} and/or N_{IR} can be reduced at the cost of increasing variance.

8. Conclusions

We present importance caching – a global illumination algorithm that improves importance sampling of VPL contributions, while maintaining efficiency by exploiting the illumination coherence in the scene. The idea is to compute exact VPL contributions at a sparse set of locations, and reuse these evaluations in the form of importance for selecting the few most relevant VPLs at other locations. We design a set of suitable importance distributions to ensure robustness at illumination discontinues, and propose a novel α-max heuristic to combine these distributions in a low-variance estimator.

Our algorithm is best suited to architectural visualization of scenes with occlusion and complex non-uniform lighting, where visibility is often a major source of variance. It can deliver a good initial approximation in a matter of seconds

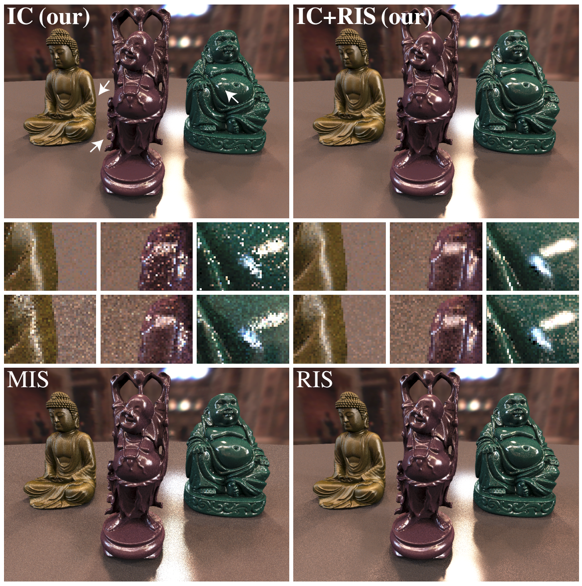

Figure 6: *Glossy Buddha statues with Phong exponents, from left to right, 10, 80, 300 on a glossy ground with exponent 500, rendered in 20 sec using our importance caching algorithm (IC), resampled importance sampling (RIS), MIS-combined BRDF/illumination sampling (MIS), and combined IC and RIS (IC+RIS). In this worst case scenario for IC it still outperforms RIS and MIS in regions with occlusion and/or smooth illumination, while the combined IC+RIS algorithm achieves the best overall quality.*

on a standard CPU, while a simplified version interactively renders accurate high quality previews. Future work can improve scalability with the number VPLs, e.g. via similarity clustering. The algorithm will also likely benefit from a GPU implementation, as its stages are inherently parallel.

Acknowledgements: The first author would like to thank Vincent Pegoraro for the fruitful discussions.

References

[BGH05] BURKE D., GHOSH A., HEIDRICH W.: Bidirectional Importance Sampling for Direct Illumination. *Computer Graphics Forum (Proceedings of Eurographics)* (2005). 1, 2, 3

[CAE08] CLINE D., ADAMS D., EGBERT P.: Table-driven Adaptive Importance Sampling. *Proceedings of EGSR* (2008). 2

[CAM08a] CLARBERG P., AKENINE-MÖLLER T.: Exploiting Visibility Correlation in Direct Illumination. *Computer Graphics Forum (Proceedings of EGSR) 27*, 4 (2008). 2, 3

[CAM08b] CLARBERG P., AKENINE-MÖLLER T.: Practical Product Importance Sampling for Direct Illumination. *Computer Graphics Forum (Proceedings of Eurographics) 27*, 2 (2008). 2

[CETC06] CLINE D., EGBERT P. K., TALBOT J. F., CARDON D. L.: Two Stage Importance Sampling for Direct Lighting. In *Eurographics Symposium on Rendering* (2006). 2

[CJAMJ05] CLARBERG P., JAROSZ W., AKENINE-MÖLLER T., JENSEN H. W.: Wavelet Importance Sampling: Efficiently Eval-

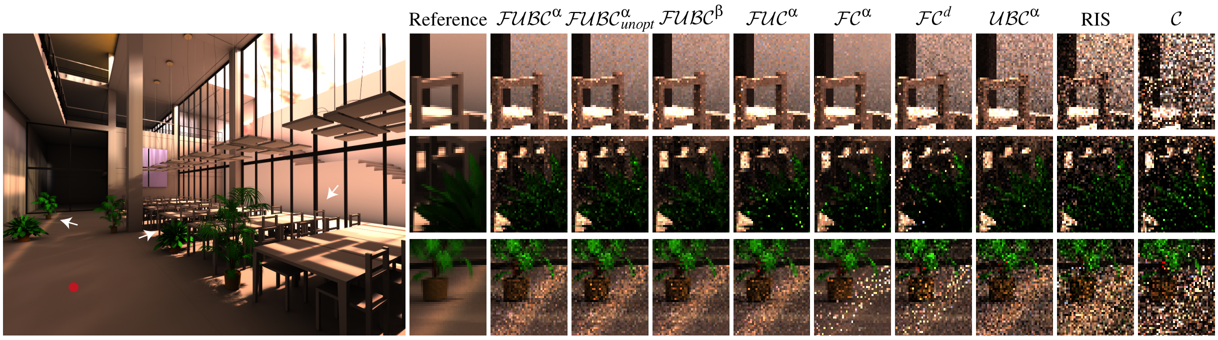

Figure 7: *A comparison of different rendering setups. The label on each image denotes the techniques used for sampling and combination (as superscript). Each image uses 35 samples per pixel per frame, distributed manually among the sampling distributions to achieve the best result. The reference image was rendered using $\mathcal{FUBC}^{\alpha}$ in 1000s, the other images took 20s.*

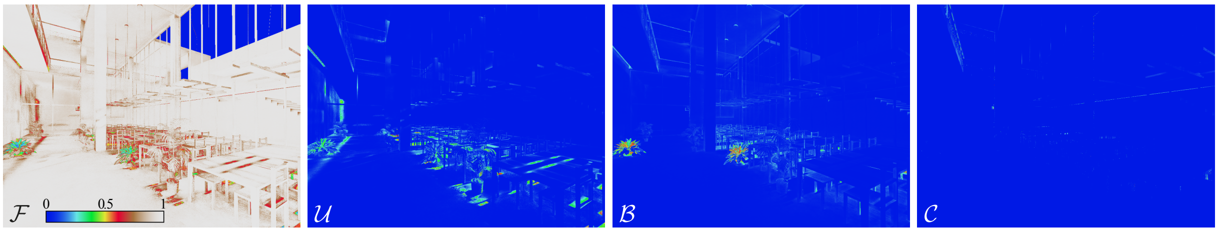

Figure 8: *Fractional contributions of the sampling distributions used for the reference image in Figure 7. With the so chosen α-max parameters, $\mathcal{F}$ captures* 93.8% *of the illumination, $\mathcal{U}$ and $\mathcal{B}$ account for* 3% *each, leaving* 0.2% *of the contribution to $\mathcal{C}$.*

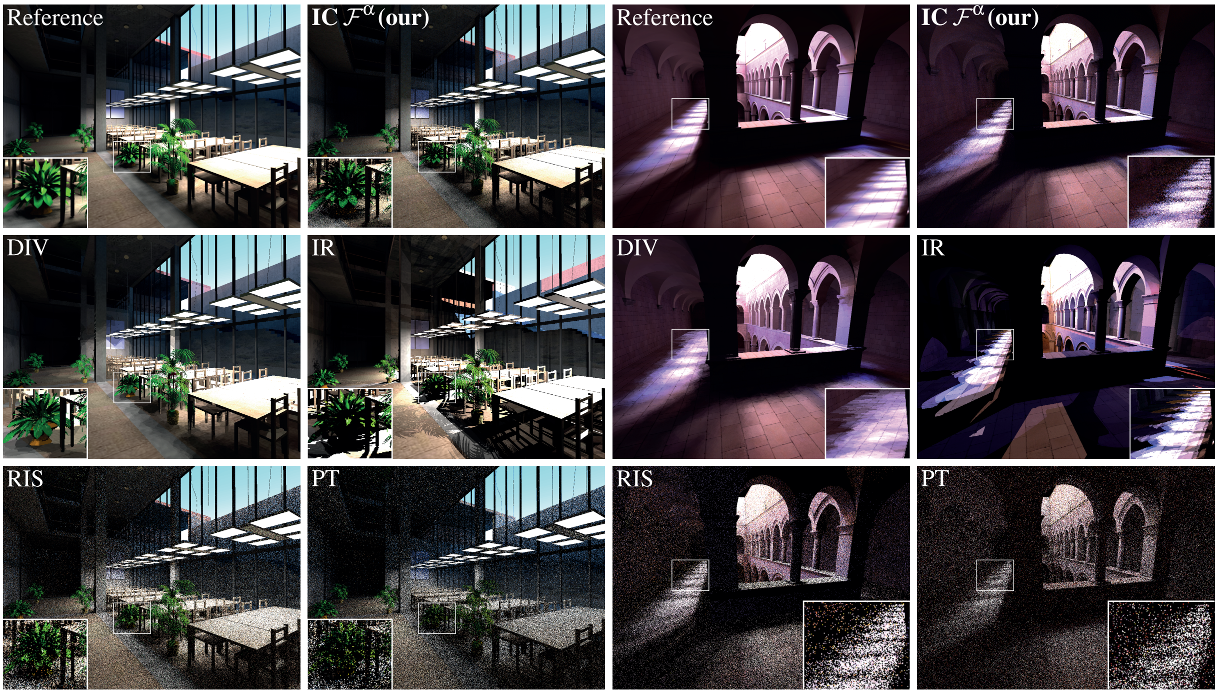

Figure 9: *A 2-second full global illumination preview quality comparison on the* STUDY HALL *(left) and* SPONZA *(right) scenes between our importance caching algorithm using α-max clamped $\mathcal{F}$ only (IC $\mathcal{F}^{\alpha}$), direct importance visualization (DIV), instant radiosity (IR), resampled importance sampling (RIS) and path tracing (PT). Our algorithm (IC $\mathcal{F}^{\alpha}$) produces a very low noise image close to the reference, and even without the $\mathcal{U}$, $\mathcal{B}$ and $\mathcal{C}$ techniques handles well orientation and visibility discontinuities, which are blurred away by DIV. IR suffers from aliasing, while RIS and PT both produce very noisy results.*

uating Products of Complex Functions. *ACM Transactions on Graphics (Proceedings of SIGGRAPH) 24*, 3 (2005). 1, 2, 3

[DGR*09] Dong Z., Grosch T., Ritschel T., Kautz J., Seidel H.-P.: Real-time Indirect Illumination with Clustered Visibility. In *VMV Workshop* (2009). 2

[DKH*10] Davidovič T., Křivánek J., Hašan M., Slusallek P., Bala K.: Combining Global and Local Virtual Lights for Detailed Glossy Illumination. *ACM TOG (Proceedings of SIGGRAPH Asia) 29* (2010). 7

[GKB09] Gassenbauer V., Křivánek J., Bouatouch K.: Spatial Directional Radiance Caching. *Computer Graphics Forum 28*, 4 (2009), 1189–1198. 2

[GS10] Georgiev I., Slusallek P.: Simple and Robust Iterative Importance Sampling of Virtual Point Lights. *Proceedings of Eurographics (short papers)* (2010). 2, 3

[Hei01] Heinrich Hey and Werner Purgathofer: *Importance Sampling with Hemispherical Particle Footprints*. Tech. Rep. TR-186-2-01-05, Vienna University of Technology, 2001. 2

[Hes95] Hesterberg T.: Weighted Average Importance Sampling and Defensive Mixture Distributions. *Technometrics 37*, 2 (1995). 6

[HKWB09] Hašan M., Křivánek J., Walter B., Bala K.: Virtual Spherical Lights for Many-Light Rendering of Glossy Scenes. In *ACM SIGGRAPH Asia 2009 papers* (2009). 7

[HPB07] Hašan M., Pellacini F., Bala K.: Matrix Row-Column Sampling for the Many-Light Problem. In *ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM. 2, 4

[Jen95] Jensen H. W.: Importance Driven Path Tracing using the Photon Map. In *Eurographics Rendering Workshop* (1995), Springer-Verlag, pp. 326–335. 2

[KBPv06] Křivánek J., Bouatouch K., Pattanaik S. N., Žára J.: Making Radiance and Irradiance Caching Practical: Adaptive Caching and Neighbor Clamping. *Computer Graphics Forum (Proceedings of EGSR)* (2006). 2

[Kel97] Keller A.: Instant Radiosity. *SIGGRAPH '97 (Proceedings of the 24th annual conference on Computer graphics and interactive techniques)* (1997). 1, 2, 3, 7

[KGPB05] Křivánek J., Gautron P., Pattanaik S., Bouatouch K.: Radiance Caching for Efficient Global Illumination Computation. *IEEE TVCG 11*, 5 (2005). 1, 2

[KK04] Kollig T., Keller A.: Illumination in the Presence of Weak Singularities. In *MCQMC Methods* (2004). 7

[KRK04] Kontkanen J., Räsänen J., Keller A.: Irradiance Filtering for Moite Carlo Ray Tracing. In *Monte Carlo and Quasi-Monte Carlo Methods 2004* (2004). 2

[OP11] Ou J., Pellacini F.: LightSlice: Matrix Slice Sampling for the Many-Lights Problem. *ACM Trans. Graph. 30* (2011). 2

[Pha05] Pharr M.: Extended Photon Map Implementation in PBRT, 2005. 2

[RCL*08] Rousselle F., Clarberg P., Leblanc L., Ostromoukhov V., Poulin P.: Efficient Product Sampling using Hierarchical Thresholding. *The Visual Computer (Proceedings of CGI 2008), 24*, 7-9 (2008). 2

[REH*11] Ritschel T., Eisemann E., Ha I., Kim J. D., Seidel H.-P.: Making Imperfect Shadow Maps View-Adaptive: High-Quality Global Illumination in Large Dynamic Scenes. In *EGSR 2011* (2011). 2

[RGK*08] Ritschel T., Grosch T., Kim M. H., Seidel H.-P., Dachsbacher C., Kautz J.: Imperfect Shadow Maps for Efficient Computation of Indirect Illumination. In *SIGGRAPH ASIA* (2008), vol. 27. 2

[SIMP06] Segovia B., Iehl J.-C., Mitanchey R., Péroche B.: Bidirectional Instant Radiosity. In *Proceedings of the 17th Eurographics Workshop on Rendering* (2006). 2

[SIP07] Segovia B., Iehl J.-C., Péroche B.: Metropolis Instant Radiosity. In *Proceedings of Eurographics* (2007). 2

[SL06] Steinhurst J., Lastra A.: Global Importance Sampling of Glossy Surfaces Using the Photon Map. *Symposium on Interactive Ray Tracing 0* (2006), 133–138. 2

[TCE05] Talbot J. F., Cline D., Egbert P. K.: Importance resampling for global illumination. In *Proc. EGSR* (2005). 2

[Vea97] Veach E.: *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford University, 1997. 2, 5, 6

[WA09] Wang R., Akerlund O.: Bidirectional Importance Sampling for Unstructured Illumination. In *Proceedings of Eurographics 2009* (2009). 1, 2

[WABG06] Walter B., Arbree A., Bala K., Greenberg D. P.: Mulltidimensional lightcuts. In *ACM SIGGRAPH 2006 Papers* (New York, NY, USA, 2006), ACM, pp. 1081–1088. 2

[WBS03] Wald I., Benthin C., Slusallek P.: Interactive Global Illumination in Complex and Highly Occluded Environments. *14th Eurographics workshop on Rendering* (2003). 2

[WFA*05] Walter B., Fernandez S., Arbree A., Bala K., Donikian M., Greenberg D. P.: Lightcuts: a Scalable Approach to Illumination. In *ACM SIGGRAPH* (2005). 2, 3, 4, 7

[WRC88] Ward G. J., Rubinstein F. M., Clear R. D.: A Ray Tracing Solution for Diffuse Interreflection. In *Proc. ACM SIGGRAPH* (1988). 1, 2, 3

Appendix A: VPL Contribution Bounds

For the $\mathcal{B}$ distribution (Section 4) we need a geometric term bound between VPL R_k and receiver points x in the neighborhood of importance record (IR) I_j (see Figure 10). If R_k is infinite, we simply have $G_k^{max}(R_k) = \cos\theta_{min}$. Otherwise, we have to compute $G_k^{max}(R_k) = \frac{\cos\theta_{min}\cos\theta_{min}^{R_k}}{d_{min}}$, for which we also need the radius r_k of the region of influence of I_k.

We bound each term individually. We estimate r_k by multiplying the radius of the camera ray footprint at I_k by the screen-space IR density. The distance bound is then $d_{min} = max(0, d - r_k)$, where $d = ||R_k - I_j||$. For $\theta_{min}^{R_k}$ we find the maximum change $\Delta\theta^{R_k} = -asin(r/d)$. For θ_{min} we use a maximum change of $\Delta\theta = -30°$, which worked better in our tests than the conservative $\theta_{min} = 0°$.

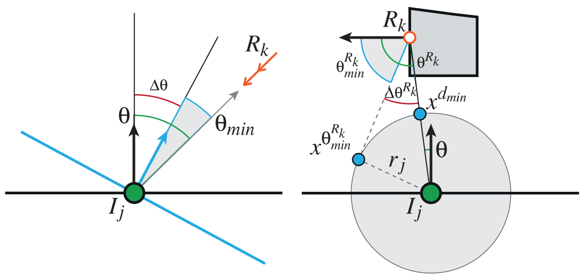

Figure 10: *Bounding the geometric term in a neighborhood of record I_j for infinite (left) and finite (right) VPLs.*