Sequential Monte Carlo Adaptation in Low-Anisotropy Participating Media

Vincent Pegoraro [1]   Ingo Wald [1,2]   Steven G. Parker [1,3]

[1]University of Utah   [2]Intel   [3]NVIDIA

Abstract

This paper presents a novel method that effectively combines both control variates and importance sampling in a sequential Monte Carlo context. The radiance estimates computed during the rendering process are cached in a 5D adaptive hierarchical structure that defines dynamic predicate functions for both variance reduction techniques and guarantees well-behaved PDFs, yielding continually increasing efficiencies thanks to a marginal computational overhead. While remaining unbiased, the technique is effective within a single pass as both estimation and caching are done online, exploiting the coherency in illumination while being independent of the actual scene representation. The method is relatively easy to implement and to tune via a single parameter, and we demonstrate its practical benefits with important gains in convergence rate and competitive results with state of the art techniques.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism - Ray-tracing G.3 [Probability and Statistics]: Probabilistic Algorithms

1. Introduction

Participating media are used to model a wide variety of elements ranging from translucent materials such as skin or marble where subsurface scattering plays a major role, to various gaseous volumes such as smoke, clouds, and the atmosphere. The ability to accurately and efficiently simulate their properties has considerable scientific implications. Although realistic rendering of these media is often of concern to the movie and gaming industries, such interest also emerged among safety oriented research where typical scenarios entail predicting the visibility of traffic signs in a foggy weather or exit signs in a smoke-filled room.

Despite their low order of convergence, Monte Carlo methods are a very general and robust technique for stochastically estimating multi-dimensional integrals and have consequently been heavily used in path-tracing to render complex global illumination effects. Several techniques were developed to reduce the variance of such estimates including importance sampling, control variates and (ir)radiance caching, sometimes trading noise for bias perceptually less noticeable in order to yield plausible renderings with practical computation times. Integrating the product of the cached radiance and the phase function must in general be done on the fly via resampling. While realistic for a few coefficients, this becomes prohibitive for refined representations making it hard to predict whether the reduction in variance will actually overcome the considerable computational overhead.

Building on the previous concepts, this paper presents a novel method targeting physically meaningful renderings and exploiting the low dependence of the phase function on the direction of the incoming ray. Radiance estimates computed during rendering are cached in a 5D data structure designed as a spatial octree where leaf nodes contain a directional grid of adaptive resolution. This caching scheme provides dynamically refined representations of two predicate functions allowing both control variates and importance sampling to be used in a sequential Monte Carlo context. This context allows for increases in the *order* of convergence (not just a constant noise reduction factor) of the estimation process. Since each new estimate is evaluated according to a fixed snapshot of the two predicates, no bias is introduced, while allowing the functions to evolve between samples.

This document starts by providing an overview of the related work and theoretical background. The method is then presented followed by both quantitative and qualitative results along with a discussion of its limitations.

2. Related Work

Several approaches have been proposed in the computer graphics literature to reduce the variance of global illumination estimates computed using Monte Carlo integration techniques. Such methods include bidirectional path-tracing [LW96], photon-mapping [JC98] and Metropolis light transport [PKK00]. A good survey is provided in [CPCP*05].

Irradiance caching was first introduced by Ward et al. [WRC88, WH92] and subsequently improved in [SM02, TL04, CB04]. Extensions to radiance were proposed including stepwise representations, such as fixed-size data structures [CZS96, GSHG98] and Haar wavelets [CSSD96], as well as (hemi)spherical harmonics [AFO05, KG05, KBPZ06]. Voxel-based [BSS93] and envelope-based methods [BSS94] were also introduced for volumetric media. The structure is typically initialized in a pre-processing pass.

Other methods focused on using the cached values to guide the sampling process more efficiently, formulating the probability density functions (PDFs) via 2D k-D trees [DW94], fixed grids [DW95], photon maps [Jen95], Haar wavelets [PP99] and particle-footprints [HP02]. In order to guarantee non-zero stepwise representations, the PDF values are customarily artificially clamped to a minimal threshold.

Lafortune et al. [LW95] proposed to cache surface radiance values in a duotricenary tree (the direct 5D extension of an octree uniformly handling spatial and directional coordinates) refined based on the density of primary samples. In addition to guiding the sampling process, the stepwise fixed-grid hemispherical representations built via resampling are also used as control variates. Besides the undesirable discontinuous nature of the resulting integrand, this double usage is actually of no benefit. Also, while they reported reductions in variance but with unaffected convergence rates, the linear cost of resampling induced large computational overheads.

Recently, sequential methods focusing on adaptation during rendering received some attention in the graphics community and the works of Fan [Fan06] and Ghosh et al. [GDH06] showed promising applications of this framework.

Although for participating media, the method presented in this paper extends Lafortune's work by addressing its various issues. The proposed data structure was designed for efficient estimations on the integration domain of concern and consists of a spatial tree of directional entities requiring no resampling. It provides symbiotic continuous control variates and stepwise PDF predicates guided by both visual importance and the features of the scene, combined in a sequential context to yield increases in the convergence rate.

3. Theoretical Background

This section provides an overview of the related theoretical background, including Monte Carlo integration and the main concepts of radiative energy transfer in participating media while referring the reader to classic texts for further details.

3.1. Monte Carlo Integration

Monte Carlo methods are a general and robust technique for stochastically evaluating multi-dimensional integrals. To reduce the variance of the estimates, several techniques were developed [HH64, KW86]. The control variates method assumes the knowledge of a function g approximating the integrand f and analytically integrable as G, while importance sampling assumes a normalized PDF $p \geq 0$ correlated with f and such that $p \neq 0$ whenever $f \neq 0$. Defining N samples, a continuous random variable $\vec{X}$ distributed according to p and the expectation E, both techniques can be combined [OZ00] to compute the integral F of f on a domain D, yielding the unbiased estimator $\widehat{F}$ of standard deviation $\sigma[\widehat{F}]$

$$F = E\left[\frac{f(\vec{X}) - g(\vec{X})}{p(\vec{X})}\right] + G \Rightarrow \widehat{F} = \sum_{i=1}^{N} \frac{f(\vec{x_i}) - g(\vec{x_i})}{Np(\vec{x_i})} + G \quad (1)$$

$$\sigma\left[\widehat{F}\right] = \sqrt{\frac{1}{N}V\left[\frac{f(\vec{X}) - g(\vec{X})}{p(\vec{X})}\right]} = \frac{1}{N^{\frac{1}{2}}}\sigma\left[\frac{f(\vec{X}) - g(\vec{X})}{p(\vec{X})}\right] \quad (2)$$

Equation 1 shows that p should now resemble $f - g$ rather than f. As the sign of the integrand $f - g$ might here vary while $p \geq 0$ must hold, an alternative is to correlate p with $|f - g|$ instead [Bek99]. Also, if g is proportional to p, the previous estimator becomes identical to the one with importance sampling alone. This implies that if a function is used for importance sampling, using it as a control variate as well will not yield any further variance reduction [Vea97].

Equation 2 shows that when using static predicate functions p (as when importance sampling from the phase function) and g (as when $g = 0$ and $p = 1/\|D\|$ for basic Monte Carlo integration), the method exhibits an order of converge of $1/2$, meaning that n^2 times as many samples are necessary to reduce the expected error by $1/n$. In this context, these techniques yield a reduction of variance if $V[(f - g)/p] < V[f\|D\|]$ (both constant with respect to N) corresponding to a vertical translation on a log scale of the convergence curves shown in figure 3 (division by a constant factor on a linear scale). To affect the slope of these curves, i.e. the convergence rate, sequential Monte Carlo methods are adequate.

In Markov chain Monte Carlo (MCMC) methods, the next state solely depends on the present state, i.e. every future state is conditionally independent of every prior state. Sequential Monte Carlo (SMC) methods split the computation in stages such that the estimator in a subsequent stage is adapted based on the information gained during previous stages in the sequence. While this dependent sampling may appear to introduce bias, it can be proven that the result is unbiased and that the method can considerably increase the rate of convergence of the estimation process [Hal62]. This can be illustrated by assuming adaptive predicates g and p such that $V[(f - g)/p]$ decreases with an order 2α with respect to $V[f]$. The standard deviation then becomes

$$\sigma\left[\widehat{F}\right] = \sqrt{\frac{1}{N}\frac{1}{N^{2\alpha}}V\left[f(\vec{X})\right]} = \frac{1}{N^{\frac{1}{2}+\alpha}}\sigma\left[f(\vec{X})\right]. \quad (3)$$

3.2. Radiative Transfer in Participating Media

The evolution of (spectral) radiance L as light travels through a medium is defined by the Radiative Transport Equation (RTE). For a given wavelength λ, position in space $\vec{x}$, and direction $\vec{\omega}$, this integro-differential equation reads [SH81]

$$(\vec{\omega} \cdot \nabla) L(\lambda, \vec{x}, \vec{\omega}) = \sigma_a(\lambda, \vec{x})(L_e(\lambda, \vec{x}, \vec{\omega}) - L(\lambda, \vec{x}, \vec{\omega})) +$$
$$\sigma_s(\lambda, \vec{x})(L_i(\lambda, \vec{x}, \vec{\omega}) - L(\lambda, \vec{x}, \vec{\omega})), \quad (4)$$

where σ_a is the absorption coefficient, σ_s the scattering coefficient, L_e the emitted radiance, and L_i the in-scattered radiance which depends on a normalized phase function Φ

$$L_i(\lambda, \vec{x}, \vec{\omega}) = \int_{4\pi} L(\lambda, \vec{x}, \vec{\omega}_i) \Phi(\lambda, \vec{\omega}, \vec{\omega}_i) d\vec{\omega}_i. \quad (5)$$

Defining the boundary condition as the background radiance L_b, the extinction coefficient as $\sigma_t = \sigma_a + \sigma_s$ and the optical thickness as $\tau(\lambda, \vec{x}_a, \vec{x}_b) = \int_{\vec{x}_a}^{\vec{x}_b} \sigma_t(\lambda, \vec{x}) d\vec{x}$, the RTE accepts an analytical solution which reads

$$L(\lambda, \vec{x}, \vec{\omega}) = e^{-\tau(\lambda, \vec{x}, \vec{x}_0)} L_b(\lambda, \vec{x}_0, \vec{\omega}) + \quad (6)$$
$$\int_{\vec{x}}^{\vec{x}_0} e^{-\tau(\lambda, \vec{x}, \vec{x}')} (\sigma_a(\lambda, \vec{x}') L_e(\lambda, \vec{x}', \vec{\omega}) + \sigma_s(\lambda, \vec{x}') L_i(\lambda, \vec{x}', \vec{\omega})) d\vec{x}'.$$

A first unbiased estimation method often used in homogeneous media defines the source radiance $L_t = (\sigma_a L_e + \sigma_s L_i)/\sigma_t$ and uses $e^{-\tau} \sigma_t$ as a PDF to sample the integrand obtained by rewriting equation 6 as a single integral [LW96]

$$L(\lambda, \vec{x}, \vec{\omega}) = \int_{\vec{x}}^{\vec{x}_{\inf}} e^{-\tau(\lambda, \vec{x}, \vec{x}')} \sigma_t(\lambda, \vec{x}') \quad (7)$$
$$\left(\|\vec{x}' - \vec{x}\| < \|\vec{x}_0 - \vec{x}\| ? L_t(\lambda, \vec{x}', \vec{\omega}) : L_b(\lambda, \vec{x}_0, \vec{\omega}) \right) d\vec{x}'.$$

A second approach often used in inhomogeneous media approximates equation 6 by assuming that the properties of the medium are constant over a set of non-overlapping intervals along the ray, as in [LW96, JC98, PKK00]. The ray-marching algorithm yields for a single interval $\|\Delta \vec{x}\|$

$$L(\lambda, \vec{x} + \Delta \vec{x}, \vec{\omega}) = e^{-\sigma_t(\lambda, \vec{x}) \|\Delta \vec{x}\|} L(\lambda, \vec{x}, \vec{\omega}) + \quad (8)$$
$$(1 - e^{-\sigma_t(\lambda, \vec{x}) \|\Delta \vec{x}\|}) L_t(\lambda, \vec{x}, \vec{\omega}).$$

While the other terms are locally defined, estimating the in-scattered radiance and transmitted radiance L requires new rays to be traced. To evaluate both while tracing a single ray, we propose to rewrite the associated terms on the right-hand side of equation 8 as a single integral I which is a straightforward juxtaposition onto a single virtual spherical domain of extent $4\pi + 1$. We define $\kappa_t = e^{-\sigma_t \|\Delta \vec{x}\|}$, $\kappa_s = (1 - e^{-\sigma_t \|\Delta \vec{x}\|})\omega$, the single-scattering albedo $\omega = \sigma_s / \sigma_t$, the solid angles $\Omega_t = 1$ and $\Omega_s = 4\pi$ such that $\Omega_t \cap \Omega_s = \emptyset$, and $\delta(\vec{\omega}, \Omega) = 1$ if $\vec{\omega} \in \Omega$ and 0 otherwise.

$$I(\lambda, \vec{x}, \vec{\omega}) = \kappa_t(\lambda, \vec{x}) L(\lambda, \vec{x}, \vec{\omega}) + \kappa_s(\lambda, \vec{x}) L_i(\lambda, \vec{x}, \vec{\omega})$$
$$= \int_{\Omega_t \cup \Omega_s} \delta(\vec{\omega}_i, \Omega_t) \kappa_t(\lambda, \vec{x}) L(\lambda, \vec{x}, \vec{\omega}) \quad (9)$$
$$+ \delta(\vec{\omega}_i, \Omega_s) \kappa_s(\lambda, \vec{x}) L(\lambda, \vec{x}, \vec{\omega}_i) \Phi(\lambda, \vec{\omega}, \vec{\omega}_i) d\vec{\omega}_i$$

4. SMC Adaptation for Participating Media Rendering

This section describes how to carry the evaluation of integrals 5 (for unbiased estimations 7) and 9 (for ray-marching 8) in a sequential Monte Carlo context using both control variates and importance sampling. As the integrand f in both integrals is expressed in terms of the incoming radiance L, we introduce a 5D data structure allowing for efficient estimations of integrals over solid angles in which the samples computed during rendering are cached. The following subsections detail the directional caching schemes for each variance reduction technique as well as the spatial and directional adaptive refinement strategies before explaining how this information is used in the estimate evaluation process.

4.1. Caching for Control Variates

The representation defining predicate g should provide low-cost read/write access and efficient integration to compute G. B-splines meet both criteria as their basis functions have local support and their integral evaluates to a simple sum of their coefficients regardless of their order, except at the domain boundaries. For this property to hold in 2D, we regularly partition the normalized spherical coordinates $s = \phi/2\pi$ and $t = (1 - cos(\theta))/2$ as to yield uniform solid angles.

Control variates lead to the new integrand $f - g$ of which properties must be analyzed in correlation with the complexity of evaluating g. While order 0 B-splines are the cheapest (involving 1 coefficient), their piecewise constant representation artificially introduces undesirable high-frequency discontinuities in the integrand, therefore decreasing the potential benefit of the method. Order 1 B-splines (piecewise linear) consequently provide higher quality estimates for a modest overhead (4 coefficients) while remaining natural interpolants. B-splines of order 2 (piecewise quadratic) and higher obviously entail a higher cost while being smoother and less tight to the control points as the support of the basis functions increases, usually yielding lower quality estimates. Order 1 B-splines are consequently most suitable and a grid representation allows for efficient interpolation.

We exploit the periodicity in s and introduce in t two polar values computed as the average of the boundary coefficients at $t = 0$ and $t = 1$ respectively. These allow to eliminate discontinuities at the poles when reconstructing g and to regularize the boundaries with respect to integration, yielding a more efficient computation of G which evaluates to a simple average assuming an isotropic phase function. As shown in figures 1(a)-1(b), each directional cell holds a color of which channels represent the coefficients of the 2D B-splines defining predicate g. Whenever a new sample is estimated, its color is averaged with the corresponding cell's coefficients while incrementing its counter of cached records C which determines the respective weights $1/(C+1)$ and $C/(C+1)$. The spherical integral and polar averages are maintained and updated at each write operation, allowing the constant time computation of both g and G during the estimation process.
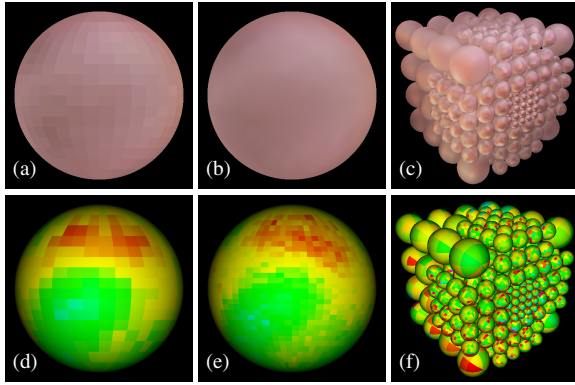
Figure 1: *(a) Per-color-channel coefficients stored in each directional cell of a radiance cache (b) 2D B-spline reconstruction of the incoming radiance defining predicate g (d-e) Color-mapped scalar values of predicate p at 2 different refinement stages (c-f) Illustration of the hierarchical data structure designed as a spatial octree of which leaf nodes hold adaptive directional grids defining both control variates and importance sampling 5D predicate functions.*

4.2. Caching for Importance Sampling

For efficiency reasons, the resolution used to represent predicate p is set to be the same as the one for g. Drawing samples from a given PDF can be done by inverting its cumulative distribution function (CDF) defined as its partial integral. This favors cheap low-orders while continuity is not crucial here. Order 0 B-splines are therefore adequate. In addition to the radiance coefficients and records counter, each cell contains a scalar estimate of the value of $|f-g|$ over the associated solid angle as shown in figures 1(d)-1(e). When a new sample of f is added to a cell, the value of g is determined and $f-g$ computed. Since the latter is a color, a scalar PDF sample is generated by averaging the absolute values of its channels and merged with the cell's PDF coefficient.

To make the sampling process inexpensive, each cache maintains a logical tree of partial sums [Vea97] similar in spirit to a Huffman tree, stored in a flat array of size $2N-1$ with N being the number of cells. Each node of this complete binary tree holds the sum of its 2 children, starting with the cells' values of p as the leaves up until the root holding the sum of all PDF coefficients. While write operations need to traverse the $log_2(2N)$ nodes of a branch, the space of basis functions can now be sampled in logarithmic time given a random number. Normalization is achieved by multiplying the random number by the value of the root node. If the random quantity is greater than the value of the first child of the current node, its PDF value is subtracted from the quantity and the second child becomes the current node, the latter being set to the first child otherwise. The process is recursively repeated as to traverse an entire branch until a cell is reached and a random direction is drawn from the linear CDF.

4.3. Adaptive refinement

The proposed hierarchical structure provides an adaptive representation permanently refining in correlation with the current records population. Given that positional interpolation mainly impacts continuity in the efficiency of the estimation process which we do not seek, the structure was designed as an octree of spatial partitions, each holding an adaptive radiance cache as shown in figures 1(c)-1(f).

The structure is initialized as a single octree node holding a radiance cache with only two cells (in s) of which radiance B-spline coefficients, PDF values and records counters default to zero. Since the PDF is not relevant at this stage, a uniform directional sampling strategy is used. For each ray traced, the octree is traversed in logarithmic time to reach the node corresponding to the new sample's position while its direction determines the cache's cell which should be updated. If the refinement criterion is met, the resolution of the cache is doubled in both polar and azimuthal coordinates while duplicating previous records to preserve the data repartition. The octree node then subdivides and its radiance cache is replaced by eight new copies of it. Therefore, only leaf nodes contain a cache of which resolution is correlated with their octree depth. The cells' records counters of the eight new caches are then divided by the dimensionality of the split, i.e. 32. This effectively reduces the weight of ancient coarse records and allows future locally relevant samples to be more influential. Inheritance is enforced by preventing the counters from being rounded down to zero which would cause a new record to overwrite rather than being merged with ancestral information. While each cell of the radiance cache has to be processed, the linear cost of refining is however not prohibitive as its frequency of occurrence is low compared to other read/write accesses.

This inheritance strategy allows a PDF to always contain a portion of its ancestors'. By prohibiting the refinement of the initial root node until its two cells contain non-zero PDF coefficients, all PDFs are guaranteed to be non-zero as well. This allows the PDFs to tend freely towards zero where needed while remaining implicitly well-behaved without the need for an artificial bound as in previous approaches.

4.4. Refinement Criterion

The refinement criterion is defined as a threshold on the average value of the records counters also maintained in each radiance cache to yield a constant time access. Since the density of rays is defined by both the camera location and the PDFs guiding the sampling process, such criterion will adaptively promote deeper refinement in highly sampled regions most crucial to the variance reduction techniques. This threshold actually controls the inertia of the system and decreasing it will increase the versatility of the structure requiring a smaller population before refining. This induces predicate functions quickly morphing into the target functions, yielding improved convergence rates and lower variance.

However, if the threshold is too low, the structure might evolve while being under-populated and yield unreliable predicates. These will generate estimates of increased variance themselves affecting the subsequent predicates and lowering their reliability, causing the method to diverge while still producing statistically correct results. Hence, the optimal criterion is the lowest one guaranteeing that the structure contains meaningful information before refining. In our experiments, it was determined empirically by conducting a few trial-and-error tests on down-sampled images.

4.5. Estimate evaluation

Considering unbiased estimations (7), the associated PDF provides a means of defining a 3D sample position along the ray and depending on its location, either L_b or L_i will be evaluated. In the latter case, the octree is traversed to reach the node corresponding to the estimate position. The term G used for control variates is directly read from the spherical integral of the associated radiance cache. Using its tree of partial PDF sums, a direction of associated p is then importance sampled while the associated radiance approximation is computed from the B-spline coefficients and weighted by an isotropic phase function to yield g. This term is then subtracted from f estimated as a newly ray-traced sample times the actual phase function value. The result is finally divided by p and added to G to form the final low-variance estimate.

For ray-marching (8), the octree is traversed at each step to find the corresponding node. The spherical integral approximating L_i is read from the radiance cache and multiplied by κ_s. The result is added to κ_t times the radiance approximation to L in the transmission direction to yield G. The spherical integral of the PDF, computed as a simple division of the value at the root of the tree of partial PDF sums by the number of cells in the cache, is then multiplied by κ_s while the value of the PDF in the transmission direction is multiplied by κ_t. The probability tree is therefore completed on the fly and used to draw a sample direction of associated p. If it lies in the scattering interval Ω_s, the value of g and f are computed as before and multiplied by κ_s, whereas if the sample direction lies in Ω_t, the value of g is reconstructed in the transmission direction and a newly ray-traced sample estimates f, both being multiplied by κ_t. In either case, g is subtracted from f and the result is again divided by p and added to G to form the final low-variance estimate.

4.6. Pseudo-Code

Figure 2 provides a high-level pseudo-code illustration (for $L_e = 0$) of the integration of the various steps individually presented. Line 5. corresponds to importance sampling as described in section 4.2, lines 6. and 7. to the control variates step from section 4.1 and lines 8. to 11. to the actual estimation process from section 4.5. Finally, line 13. populates the structure as explained in sections 4.1 and 4.2 while line 14. corresponds to the refinement step from section 4.3.

```
1. EstimateRayIntegral()
2. (position, weight) = GetSamplePositionFromRayPDF();
3. if (position < mediumBoundary)
4.    cache = octree.GetCache(position);
5.    (direction, p) = cache.GetSampleDirection();
6.    G = cache.GetIntegralForIsotropicPhaseFunction();
7.    g = cache.GetRadiance(direction);
8.    g *= isotropicPhaseFunction.GetWeight(direction);
9.    radiance = TraceRay(position, direction);
10.   f = radiance * phaseFunction.GetWeight(direction);
11.   estimate = G + (f - g) / p;
12.   node = octree.GetNode(position);
13.   node.AddRecordToCache(direction, radiance);
14.   if (node.CriterionIsMet()) node.Refine();
15. else
16.   estimate = TraceBackgroundRay();
17. return estimate * weight;
```

Figure 2: *Pseudo-code for unbiased estimations (7)*

5. Results

In order to demonstrate the convergence characteristics of the method, we experimented with a test-bed consisting of a box of homogeneous isotropic participating medium illuminated by a gradient background. The solutions were computed using Markov chain Monte Carlo integration with a number of samples several orders of magnitude greater than the ones used for the test cases. The quantitative results for various albedos are shown in figure 3 where the number of samples per pixel on the abscissa increases by a factor 4. The slopes of the root mean squared error (RMSE) curves illustrate the 0.5 convergence rate of MCMC integration compared to the higher order of SMC here ranging in between 0.62 and 0.63 ($\sim$ 25% gain). Considering the rightmost vertices of the intermediate-albedo graph, the MCMC approach would require about 24 times as many samples in order to reach the error level achieved by the SMC method which in contrast only requires a 57% overhead (factor of 1.57) in computational time, yielding a 15.3X speed-up.

While it is constant for MCMC integration, the efficiency ($(variance * cost)^{-1}$) of the SMC method keeps increasing with the sampling rate. For the high-albedo case, control variates provide most of the gain whereas importance sampling has a more prominent impact as the albedo diminishes. This illustrates their respective strengths. While control variates approximate well smooth variations, importance sampling performs better for higher frequencies by focusing on strong contributions. When combined, control variates allow importance sampling to focus on hard features rather than smooth high contributions, yielding increased efficiencies. The only exception arises in the high-albedo case where the cost of importance sampling here supersedes its benefit, making control variates alone more efficient. However, extrapolating from the slopes of these curves, the method may become beneficial at higher sampling rates.
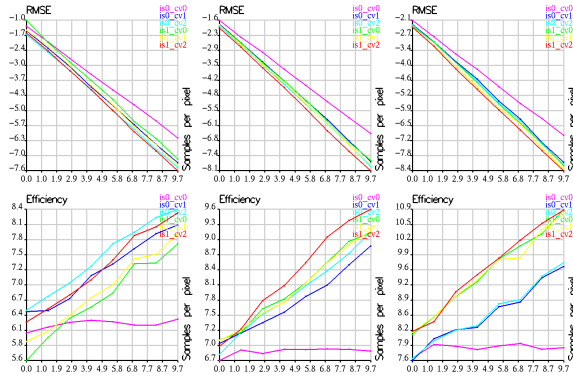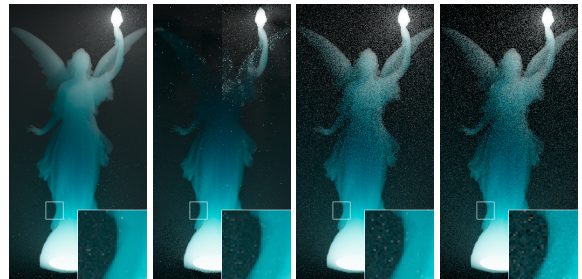
Figure 3: *Logarithmic plots of root mean squared error and efficiency versus number of samples per pixel for experiments with an albedo of 1, 0.8 and 0.5 (from left to right). Results are shown for several strategies: no control variate (cv0), piecewise constant control variates (cv1), piecewise linear control variates (cv2), no importance sampling (is0) and piecewise constant PDFs (is1). is1_cv2 corresponds to the proposed SMC method and is0_cv0 to MCMC.*

| Figure	4	5	6
Criterion	512	256	256
Depth	8	7	6
Nodes	12577	2489	6913
Leaves	11005	2178	6049
Cells	23810048	3887712	6525440
Memory	0.978 GB	163.5 MB	274.8 MB

Table 1: *Characteristics of the data structure after rendering the listed figures, including the value used for the refinement criterion, the maximal depth of the octree, its total number of nodes, its number of leaf nodes (i.e. of radiance caches), the total number of cells for all radiance caches, and the total memory usage (using double-precision).*

Figure 4: *Lucy in the fog illuminated by two implicitly sampled lights, rendered using (from left to right) SMC importance sampling and control variates (4096 spp), SMC importance sampling alone (3025 spp), SMC control variates alone (8836 spp), and MCMC (10816 spp) all in 32.5 hours.*

Table 1 details the memory requirements of the data structure for various qualitative experiments generated at a resolution of 512x512 pixels on an Intel Xeon 3.00GHz processor desktop with 2GB of RAM. To evaluate the ability of the method to autonomously capture energy radiators, we experimented with a scene containing implicitly sampled (i.e. not a priori identified) light sources of about equal contribution, the first one being large and dim and the second relatively localized and intense. Figure 4 shows that control variates alone adequately capture and reduce variance from the larger light (the bluish component in the fog and near Lucy's feet; see insert) but have little impact on evaluating contributions from the small light (the whitish spots in the fog and near Lucy's head). While importance sampling alone also reasonably evaluates the larger light, it however dramatically increases variance for the small light causing many pixels to under-estimate its contribution (the darker regions) and a few to largely over-estimate it (the bright spots). This result is nonetheless statistically correct in the sense that the total energy distributed across the image is preserved, and illustrates the divergence discussed in section 4.4. Discontinuities in efficiency due to the lack of positional interpolation are also clearly visible. However, when combining both techniques, control variates mainly handle the large source and allow importance sampling to focus on the smaller one, reinforcing the statement made earlier. Because the resulting sampling strategy differs from the one with importance sampling alone, both computational overheads and variations in path-length will impact the overall rendering cost. Also, since the structure captures the global radiance distribution in the scene, a single instance can be shared by several objects with different materials (e.g. the fog and Lucy's body).

Although the method is optimal for purely isotropic phase functions, it also provides substantial improvements in low-anisotropy media. Figure 5 shows an inhomogeneous cloud with a forward Henyey-Greenstein phase function of asymmetry coefficient 0.1 illuminated by a directional light source. The superiority of the combination of both variance reduction techniques in the SMC framework is here again illustrated. Also, since main contributions are due to relatively short paths which are favored by importance sampling, the latter allows greater sampling rates for an equal rendering time despite its computational overhead.

Finally, figure 6 shows an isotropic homogeneous medium illuminated through a stained glass. Although both bidirectional path-tracing [LW96] and photon-mapping [JC98] can render such effects, these methods will be somehow inefficient without additional information about the scene since many connections will be blocked by the wall or many photons will lie on the other side of it [PKK00]. This issue is addressed in [PP98] and [KW00] yet requiring a third pass to the initially two-stage photon-mapping technique. In contrast, the proposed SMC method provides substantial variance reductions over bidirectional path-tracing while gathering information in a single rendering pass.
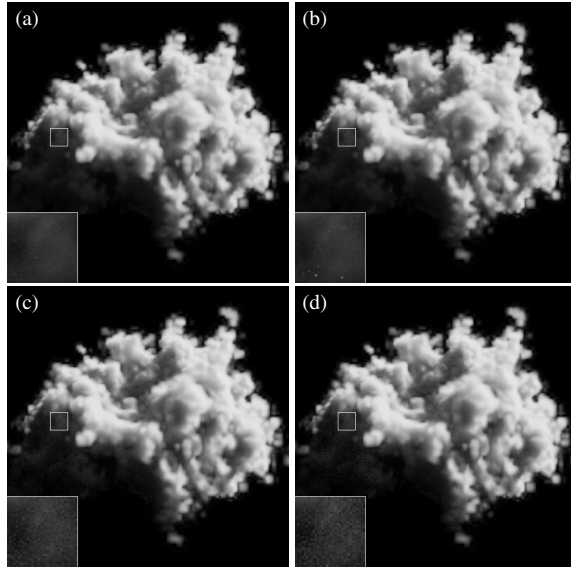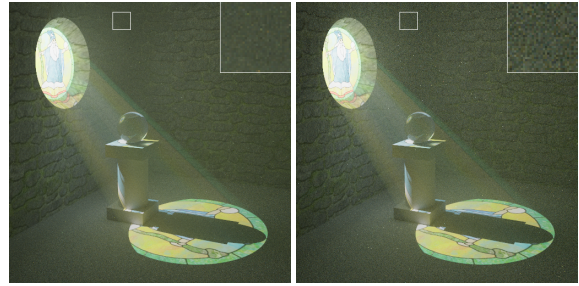
Figure 5: *An anisotropic inhomogeneous cloud rendered using (a) SMC importance sampling and control variates (256 spp), (b) SMC importance sampling alone (250 spp), (c) SMC control variates alone (166 spp), and (d) MCMC phase function importance sampling (174 spp) all in 4.7 hours.*

Figure 6: *A dust-filled room illuminated through a stained glass rendered using (from left to right) SMC importance sampling and control variates (1024 spp) and bidirectional path-tracing (1064 spp) both in 3.8 hours.*

6. Discussion and Future Work

While remaining statistically correct, the quality of the estimates will degrade as the anisotropy of the media increases and our assumption gets gradually violated. Further investigation is required to alleviate this limitation while preserving the efficiency of the method. Although setting the refinement criterion requires little effort, further investigation is also needed to determine an optimal intrinsically divergence-inhibitive formulation which adapts to the local complexity in lighting rather than being global to the scene.

Another limitation occurs in rendering volume caustics which are visually more appealing when using a point light source. The probability of implicitly sampling such source is null while refraction/reflection prevents explicit sampling. Because they start paths directly from the light source, bidirectional approaches [LW96, JC98] should here be preferred instead. The Metropolis light transport algorithm [PKK00] also performs well in such setting with focused high contribution paths. As noted in [JC98], Metropolis however tends to perform no better than pure bidirectional path-tracing in high-albedo media where the illumination is mainly soft, and where the proposed SMC method performs well. The latter can consequently be regarded as complementary to the aforementioned bidirectional approaches and future directions of research could explore ways of combining these techniques to exploit their respective strengths.

7. Conclusion

We have presented a novel method which effectively combines both control variates and importance sampling in a symbiotic sequential Monte Carlo context. The method yields continually increasing efficiencies thanks to a modest computational overhead achieved by exploiting the low-anisotropy of the participating media and implicitly guarantees non-zero PDFs via its inheritance strategy.

A main advantage is that no pre-computation is needed as both estimation and caching are done online, allowing the sampling process to be driven by both visual importance and features of interest in the scene while remaining unbiased. The algorithm exploits the coherency in illumination of the latter while being independent of its actual representation. The technique is also relatively easy to implement in a general Monte Carlo path-tracer and easy to tune via a single refinement parameter compared to the choice of good mutation strategies for Metropolis light transport or optimally balancing computation in a multi-pass approach.

In addition to important gains in the convergence rate, the quantitative and qualitative results showed that this combined model outperforms the individual variance reduction techniques on which it is based, and is competitive with state of the art techniques. The method consequently appears as a promising step towards efficiently simulating accurate light transport in participating media via self-tuning estimators that learn to become effective based on the information previously collected during the rendering process itself.

8. Acknowledgments

References

[AFO05] ARIKAN O., FORSYTH D. A., O'BRIEN J. F.: Fast and Detailed Approximate Global Illumination by Irradiance Decomposition. In *SIGGRAPH* (2005), pp. 1108–1114.

[Bek99] BEKAERT P.: *Hierarchical and Stochastic Algorithms for Radiosity*. PhD thesis, Leuven, Belgium, 1999.

[BSS93] BLASI P., SAËC B. L., SCHLICK C.: A Rendering Algorithm for Discrete Volume Density Objects. *EG Computer Graphics Forum 12*, 3 (1993), 201–210.

[BSS94] BLASI P., SAËC B. L., SCHLICK C.: An Importance Driven Monte-Carlo Solution to the Global Illumination Problem. In *EGWR* (1994), pp. 177–187.

[CB04] CHRISTENSEN P. H., BATALI D.: An Irradiance Atlas for Global Illumination in Complex Production Scenes. In *EGSR* (2004), pp. 133–141.

[CPCP*05] CEREZO E., PEREZ-CAZORLA F., PUEYO X., SERON F., SILLION F.: A Survey on Participating Media Rendering Techniques. *The Visual Computer 21*, 5 (2005), 303–328.

[CSSD96] CHRISTENSEN P. H., STOLLNITZ E. J., SALESIN D. H., DEROSE T. D.: Global Illumination of Glossy Environments Using Wavelets and Importance. *Transactions on Graphics 15*, 1 (1996), 37–71.

[CZS96] CHIU K., ZIMMERMAN K., SHIRLEY P.: The Light Volume: an Aid to Rendering Complex Environments. In *EGWR* (1996), pp. 1–10.

[DW94] DUTRÉ P., WILLEMS Y. D.: Importance-driven Monte Carlo Light Tracing. In *Proceedings of Eurographics Workshop on Rendering* (1994), pp. 185–194.

[DW95] DUTRÉ P., WILLEMS Y. D.: Potential-driven Monte Carlo Particle Tracing for Diffuse Environments with Adaptive Probability Functions. In *Eurographics Workshop on Rendering* (1995), pp. 306–315.

[Fan06] FAN S.: *Sequential Monte Carlo Methods for Physically-Based Rendering*. PhD thesis, University of Wisconsin-Madison, USA, 2006.

[GDH06] GHOSH A., DOUCET A., HEIDRICH W.: Sequential Sampling for Dynamic Environment Map Illumination. In *EGSR* (2006), pp. 115–126.

[GSHG98] GREGER G., SHIRLEY P., HUBBARD P. M., GREENBERG D. P.: The Irradiance Volume. *Computer Graphics and Applications 18*, 2 (1998), 32–43.

[Hal62] HALTON J. H.: Sequential Monte Carlo. *Cambridge Philosophical Society 58* (1962), 57–78.

[HH64] HAMMERSLEY J. M., HANDSCOMB D. C.: *Monte Carlo Methods*. Chapman and Hall, 1964.

[HP02] HEY H., PURGATHOFER W.: Importance Sampling with Hemispherical Particle Footprints. In *Proceedings of SCCG* (2002), pp. 107–114.

[JC98] JENSEN H. W., CHRISTENSEN P. H.: Efficient Simulation of Light Transport in Scenes with Participating Media using Photon Maps. In *Proceedings of SIGGRAPH* (1998), pp. 311–320.

[Jen95] JENSEN H. W.: Importance Driven Path Tracing Using the Photon Map. In *EGWR* (1995), pp. 326–335.

[KBPZ06] KŘIVÁNEK J., BOUATOUCH K., PATTANAIK S. N., ZÁRA J.: Making Radiance and Irradiance Caching Practical: Adaptive Caching and Neighbor Clamping. In *Proceedings of EGSR* (2006), pp. 127–138.

[KG05] KŘIVÁNEK J., GAUTRON P.: Radiance Caching for Efficient Global Illumination Computation. *TVCG 11*, 5 (2005), 550–561.

[KW86] KALOS M. H., WHITLOCK P. A.: *Monte Carlo Methods, Volume I: Basics*. Wiley, 1986.

[KW00] KELLER A., WALD I.: Efficient Importance Sampling Techniques for the Photon Map. In *Proceedings of VMV* (2000), pp. 271–279.

[LW95] LAFORTUNE E. P., WILLEMS Y. D.: A 5D Tree to Reduce the Variance of Monte Carlo Ray Tracing. In *EGWR* (1995), pp. 11–20.

[LW96] LAFORTUNE E. P., WILLEMS Y. D.: Rendering Participating Media with Bidirectional Path Tracing. In *EGWR* (1996), pp. 91–100.

[OZ00] OWEN A., ZHOU Y.: Safe and Effective Importance Sampling. *Journal of the American Statistical Association 95*, 449 (2000), 135–143.

[PKK00] PAULY M., KOLLIG T., KELLER A.: Metropolis Light Transport for Participating Media. In *Rendering Techniques* (2000), pp. 11–22.

[PP98] PETER I., PIETREK G.: Importance Driven Construction of Photon Maps. In *EGWR* (1998), pp. 269–280.

[PP99] PIETREK G., PETER I.: Adaptive Wavelet Densities for Monte Carlo Ray Tracing. In *WSCG Conference Proceedings* (1999), pp. 217–224.

[SH81] SIEGEL R., HOWELL J. R.: *Thermal Radiation Heat Transfer*. Hemisphere Publishing Corp., 1981.

[SM02] SMYK M., MYSZKOWSKI K.: Quality Improvement for Indirect Illumination Interpolation. In *ICCVG* (2002), pp. 685–692.

[TL04] TABELLION E., LAMORLETTE A.: An Approximate Global Illumination System for Computer Generated Films. *TOG 23*, 3 (2004), 469–476.

[Vea97] VEACH E.: *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford University, USA, 1997.

[WH92] WARD G. J., HECKBERT P.: Irradiance Gradients. In *EGWR* (1992), pp. 85–98.

[WRC88] WARD G. J., RUBINSTEIN F. M., CLEAR R. D.: A Ray Tracing Solution for Diffuse Interreflection. In *SIGGRAPH* (1988), pp. 85–92.