Path Throughput Importance Weights

Jendersie, Johannes

TU Clausthal, Germany

September 7, 2018

Abstract

Many Monte Carlo light transport simulations use multiple importance sampling (MIS) to weight between different path sampling strategies. We propose to use the path throughput to compute the MIS weights instead of the commonly used probability density per area measure. This new formulation is equivalent to the previous approach and results in the same weights as well as implementation. However, it is more intuitive and can help in understanding the effects of modifications to the weight function. We show some examples of required modifications which are often neglected in implementations. Also, our new perspective might help to derive MIS strategies for new samplers in the future.

1 Introduction

Beginning with *Path Tracing* [Kaj86] there are different solutions to the light transport simulation problem. In all cases the integral equation (*Rendering Equation*)

$$L(\mathbf{x}_i, \overrightarrow{\mathbf{w}_i}) = \int_\Omega L(\mathbf{x}_{i+1}, \overleftarrow{\mathbf{w}_i}) \rho(\mathbf{x}_i, \overrightarrow{\mathbf{w}_i}, \overleftarrow{\mathbf{w}_i}) \langle \mathbf{n}_i, \overleftarrow{\mathbf{w}_i} \rangle \mathrm{d}\overleftarrow{\mathbf{w}_i}$$

is solved numerically by sampling. For the notation please refer to table 1. Dependent on the direction of tracing we talk from light transport (paths starting at the light source/photons) or importance transport (paths starting at the observer). Both define a different sampler for the same paths. Since they cover different light effects more successfully, the combination to *Bidirectional Path Tracing* (BPT) by Veach [VG95a] gives a more robust solution.

The key idea in BPT is to weight each path from each sampler using *Multiple Importance Sampling* (MIS). The weights form a partition of unity, such that the weighted sum of all samplers is again an unbiased estimate of the *Rendering Equation*. The goal of that weights is to find a minimal variance solution. I.e. if one of the samplers has a lower variance than others, it should be preferred, otherwise an average of multiple equal samplers will also result in a lower variance due to higher sample count. The state of the art weight function is called the *Balance Heuristic* (introduced by Veach [VG95a]) and is explained in Section 2.

$\mathbf{x}_i$	A path vertex with index i; indices are ascending and start with 0 at the observer
$\mathcal{X}_a$	A path $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{s+t}$ from sampler a with s vertices on the view sub-path and t vertices on the light sub-path
a, b	Indices used to depict different samplers
$\mathbf{n}_i$	The surface normal at vertex $\mathbf{x}_i$ with $\|\mathbf{n}_i\| = 1$
$\overrightarrow{\mathbf{w}_i}$	A direction from $\mathbf{x}_{i\text{-}1}$ to $\mathbf{x}_i$ with $\|\overrightarrow{\mathbf{w}_i}\| = 1$
$\overleftarrow{\mathbf{w}_i}$	A direction from $\mathbf{x}_{i+1}$ to $\mathbf{x}_i$ with $\|\overleftarrow{\mathbf{w}_i}\| = 1$
$\overrightarrow{p_i}, \overleftarrow{p_i}$	Sampling PDF at vertex $\mathbf{x}_i$ in importance transport direction $\overrightarrow{p_i} = p(\mathbf{x}_i, \overrightarrow{\mathbf{w}_i}, \overleftarrow{\mathbf{w}_i})$ and light transport direction $\overleftarrow{p_i} = p(\mathbf{x}_i, \overleftarrow{\mathbf{w}_i}, \overrightarrow{\mathbf{w}_i})$
ρ_i	The bidirectional scattering or reflectance distribution function (BxDF); Same in both transport directions due to reciprocity
$\langle \cdot, \cdot \rangle$	Scalar product of two vectors (equals the $\cos\theta$ between the two vectors if both are normalized)
$S(\mathcal{X})$	Sampled value (radiance) of a path $\mathcal{X}$

We introduce a new way to think of the *Balance Heuristic* in Section 3. Our approach is to use the path throughput (the sampled quantity) instead of probabilities. This new perspective helps in understanding the implications of modifications to the renderer and allow more intuitive extensions towards other samplers (See Section 4). For an already existing and correct implementation there is nothing to be changed.

An example of such an extension is the combination of Photon Mapping [Jen96] with BPT. In photon mapping two sub-paths are merged by searching end points in a local neighborhood at one path end, instead of connecting sub-paths only. The difficulty here is to find a compatible probability description for both methods to be able to compute the MIS. The solution was discovered by [GKDS12] and [HPJ12] in parallel. Using our perspective the solutions becomes trivial.

A further family of samplers are Marcov Chain Monte Carlo methods which conditionally exchange sub-paths (light, importance or both) to sample an arbitrary target function. In [ŠOHK16] MCMC was combined with BPT using MIS, too. For one chain their approach uses the unmodified path sampling probability, regardless of its optimality, since computing the true probability is unfeasible. Our approach suggests that parts of the throughput calculation (like acceptance probability) should be included into the MIS computation.

2 Path Probability Importance Weights

Given the *Probability Density Functions* (PDFs) p_a of a sampler and the number of samples n_a drawn from this sampler, the *Balance Heuristic* is

$$w_a = \frac{n_a p_a}{\sum_b n_b p_b} = \frac{1}{1 + \sum_{b \neq a} \frac{n_b p_b}{n_a p_a}}. \tag{1}$$

According to Elvira et al. [EMLB17] it is the best known strategy for sampling a mixture of PDFs. Also, Veach [VG95b] states it is the optimal choice, if sampling PDF mixtures with random decisions between PDFs (one-sample model). To obtain a variance optimal combination, it is also necessary to choose the number of samples n_a optimally (multi-sample model) and then use a weighting without the n_a [SHSK16]. The second form of Eq. (1) is used in practice to avoid the computation the numerically challenging probability sums.

However, in light transport simulation we do not know the PDF p_b of all possible samplers when sampling a vertex of a path. The difficulty is that paths from the adjoint quantity (importance $\leftrightarrow$ light) depend on the scene globally. I.e. the PDF of paths which randomly reach the current vertex is not given by local properties.

A solution was given by Veach [VG95a] who used the probability density per unit area instead. For each segment of the path this measure is

$$p(\mathbf{x}_i \rightarrow \mathbf{x}_{i+1}) = \frac{\overrightarrow{p_i} \cdot |\langle \mathbf{n}_{i+1}, \overleftarrow{\mathbf{w}_i} \rangle|}{\|\mathbf{x}_i - \mathbf{x}_{i+1}\|^2} \tag{2}$$

where $\overrightarrow{p_i}$ is the sampling PDF at the source vertex, $\langle \mathbf{n}_{i+1}, \overleftarrow{\mathbf{w}_i} \rangle$ is the $\cos\theta$ at the target vertex and $\|\mathbf{x}_i - \mathbf{x}_{i+1}\|^2$ is the squared distance between the two surfaces. Analogously, the probability in the inverse direction $p(\mathbf{x}_i \leftarrow \mathbf{x}_{i+1})$ is defined by inverting all directional sizes on the right side and replacing the indices of $\overrightarrow{p}$ and $\mathbf{n}$.

The probability of a path $\mathcal{X}$ is given by the product of all its segment probabilities and the probabilities to sample the first and last vertex.

$$p(\mathcal{X}) = p(\mathbf{x}_0) \cdot \prod_{i=0}^{s-2} p(\mathbf{x}_i \rightarrow \mathbf{x}_{i+1}) \cdot \prod_{i=s}^{s+t-2} p(\mathbf{x}_i \leftarrow \mathbf{x}_{i+1}) \cdot p(\mathbf{x}_{s+t-1}). \tag{3}$$

Whenever two sub-paths, one from the observer and one from the light source, are connected, the probability measure for that sampler is the product of the two parts as defined by Eq. (3). The measures for all other possible samplers are obtained by replacing forward and backward direction for the other path segments recursively. In the unidirectional case either the left or the right half of the product vanishes, including the sampling probability of the end vertex.

Figure 1 shows some of the possibilities for sampling a certain path. For example, to get the probability for the second path from the first

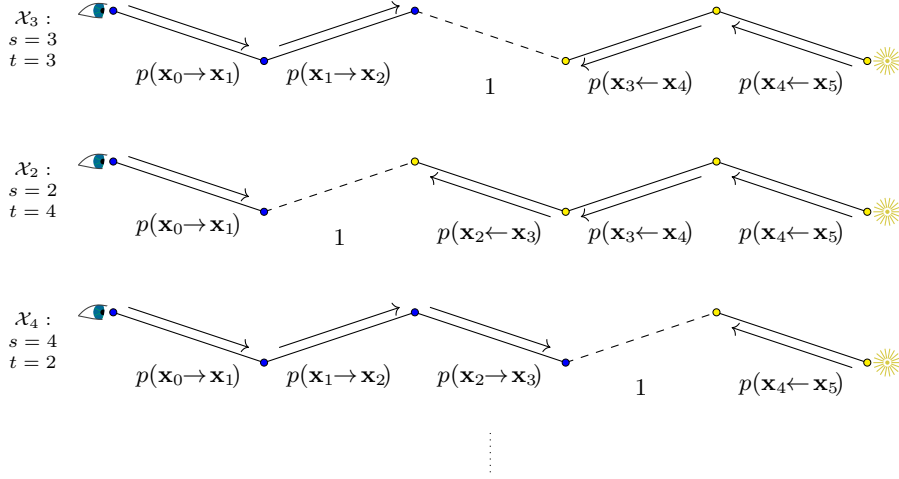$$\frac{p(\mathbf{x}_2 \leftarrow \mathbf{x}_3)}{p(\mathbf{x}_1 \rightarrow \mathbf{x}_2)}$$

Figure 1: Path probabilities for different examples of sampling possibilities for the same path.

must be multiplied. Hence, the ratio between two path probabilities has the form

$$\frac{p(\mathcal{X}_b)}{p(\mathcal{X}_a)} = \prod_{k=b}^{a-1} \frac{p(\mathbf{x}_k \leftarrow \mathbf{x}_{k+1})}{p(\mathbf{x}_{k-1} \rightarrow \mathbf{x}_k)} \tag{4}$$

where k ranges from the end vertex of the connection $s_b = b$ in $\mathcal{X}_b$ to the start vertex of the connection $(s_a = a)$-1 in $\mathcal{X}_a$. Here, w.l.o.g. the connection appears earlier on the path in $\mathcal{X}_b$ than in $\mathcal{X}_a$. Otherwise, the fraction must be inverted.

3 Path Throughput Importance Weights

To calculate the sample value (radiance throughput) $S(\mathcal{X}_a)$ of a sampled path we have

$$S(\mathcal{X}_a) = \frac{W(\mathbf{x}_0)}{p(\mathbf{x}_0)} \cdot \prod_{i=1}^{s-2} \frac{\overrightarrow{\rho_{\perp i}}}{\overrightarrow{p_i}} \cdot \frac{\overrightarrow{\rho_{\perp s-1}} \cdot \overleftarrow{\rho_{\perp s}}}{\|\mathbf{x}_s - \mathbf{x}_{s-1}\|^2} \cdot \prod_{i=s+1}^{s+t-2} \frac{\overleftarrow{\rho_{\perp i}}}{\overleftarrow{p_i}} \cdot \frac{L(\mathbf{x}_{s+t-1})}{p(\mathbf{x}_{s+t-1})}. \tag{5}$$

It depends on the sensor weight W, the emitted radiance L, multiple sampling events in the form $\rho_\perp/p$ and the transport evaluation of the connection. Here $\rho_\perp$ is the BxDF ρ multiplied with the outgoing cosine $\overrightarrow{\rho_{\perp i}} = \rho_i \cdot |\langle \mathbf{n}_i, \overleftarrow{\mathbf{w}_i} \rangle|$ and $\overleftarrow{\rho_{\perp i}} = \rho_i \cdot |\langle \mathbf{n}_i, \overrightarrow{\mathbf{w}_i} \rangle|$.

The true radiance estimate of a pixel $L(\mathbf{x}_0, \overrightarrow{\mathbf{w}_0})$ is then

$$E[S(\mathcal{X})] = \lim_{N \to \infty} \frac{1}{\sum_{k=1}^{N} w_k} \sum_{k=1}^{N} w_k S(\mathcal{X}_k),$$

4

where N is the total number of samples (increases with iterations) and $\mathcal{X}_k$ are different sampled paths. The weight w_k is artificially introduced and can be any weight like the *Balance Heuristic* from Eq. (1).

Our goal is to minimize the variance $\mathrm{V}[S(\mathcal{X})] = E[S(\mathcal{X})^2] - E[S(\mathcal{X})]^2$. Since $E[S(\mathcal{X})]$ is the fixed (true) value we only need to minimize $E[S(\mathcal{X})^2]$. Hence, preferring samplers with a small S must lead to a smaller variance. Therefore, we require

$$\hat{w}_a \propto S(\mathcal{X}_a)^{-1}.$$

This leads us to an equivalent formulation of the balance heuristic

$$\hat{w}_a = \frac{n_a \cdot S(\mathcal{X}_a)^{-1}}{\sum_b n_b \cdot S(\mathcal{X}_b)^{-1}} = \frac{1}{1 + \sum_{b \neq a} \frac{n_b S(\mathcal{X}_a)}{n_a S(\mathcal{X}_b)}}. \tag{6}$$

3.1 Equivalence to Probability Weights

Theorem 1. *The balance heuristic w_a using path probabilities Eq.(1) is equivalent to the heuristic $\hat{w}_a$ using inverse throughputs Eq.(6).*

Proof. In the calculation of $\hat{w}_a$ we need ratios of sampled throughputs which are:

$$\begin{aligned}
\frac{S(\mathcal{X}_a)}{S(\mathcal{X}_b)} &= \frac{\|\mathbf{x}_{b-1} - \mathbf{x}_b\|^2}{\overrightarrow{\rho_{\perp b-1}} \cdot \overleftarrow{\rho_{\perp b}}} \cdot \frac{\overrightarrow{\rho_{\perp a-1}} \cdot \overleftarrow{\rho_{\perp a}}}{\|\mathbf{x}_{a-1} - \mathbf{x}_a\|^2} \cdot \frac{\prod_{b-1}^{a-2} \overrightarrow{\rho_{\perp k}}/\overrightarrow{p_k}}{\prod_{b+1}^{a} \overleftarrow{\rho_{\perp k}}/\overleftarrow{p_k}} \\
&= \frac{\|\mathbf{x}_{b-1} - \mathbf{x}_b\|^2}{\|\mathbf{x}_{a-1} - \mathbf{x}_a\|^2} \cdot \frac{\prod_b^{a-1} \overrightarrow{\rho_{\perp k}}}{\prod_b^{a-1} \overleftarrow{\rho_{\perp k}}} \cdot \frac{\prod_{b+1}^{a} \overleftarrow{p_k}}{\prod_{b-1}^{a-2} \overrightarrow{p_k}} \\
&= \frac{\|\mathbf{x}_{b-1} - \mathbf{x}_b\|^2}{\|\mathbf{x}_{a-1} - \mathbf{x}_a\|^2} \cdot \prod_{k=b}^{a-1} \left| \frac{\langle \mathbf{n}_k, \overleftarrow{\mathbf{w}_k} \rangle}{\langle \mathbf{n}_k, \overrightarrow{\mathbf{w}_k} \rangle} \right| \cdot \prod_{k=b}^{a-1} \frac{\overleftarrow{p_{k+1}}}{\overrightarrow{p_{k-1}}}
\end{aligned}$$

using Eq. (5). In step one we split the products into an BxDF and a probability part and then moved the BxDFs into the product adjusting indices. In the second step we canceled the reciprocal BxDF arriving at a product of cosines and substituted the indices in the probabilities to unify the product range.

We get a similar transformation for Eq. (4) by inserting Eq. (2):

$$\begin{aligned}
\frac{p(\mathcal{X}_b)}{p(\mathcal{X}_a)} &= \left(\prod_{k=b}^{a-1} \frac{\overleftarrow{p_{k+1}} |\langle \mathbf{n}_k, \overrightarrow{\mathbf{w}_{k+1}} \rangle|}{\|\mathbf{x}_k - \mathbf{x}_{k+1}\|^2} \frac{\|\mathbf{x}_{k-1} - \mathbf{x}_k\|^2}{\overrightarrow{p_{k-1}} |\langle \mathbf{n}_k, \overleftarrow{\mathbf{w}_{k-1}} \rangle|} \right) \\
&= \prod_{k=b}^{a-1} \frac{\|\mathbf{x}_{k-1} - \mathbf{x}_k\|^2}{\|\mathbf{x}_k - \mathbf{x}_{k+1}\|^2} \cdot \prod_{k=b}^{a-1} \left| \frac{\langle \mathbf{n}_k, \overleftarrow{\mathbf{w}_k} \rangle}{\langle \mathbf{n}_k, \overrightarrow{\mathbf{w}_k} \rangle} \right| \frac{\overleftarrow{p_{k+1}}}{\overrightarrow{p_{k-1}}} \\
&= \frac{\|\mathbf{x}_{b-1} - \mathbf{x}_b\|^2}{\|\mathbf{x}_{a-1} - \mathbf{x}_a\|^2} \cdot \prod_{k=b}^{a-1} \left| \frac{\langle \mathbf{n}_k, \overleftarrow{\mathbf{w}_k} \rangle}{\langle \mathbf{n}_k, \overrightarrow{\mathbf{w}_k} \rangle} \right| \frac{\overleftarrow{p_{k+1}}}{\overrightarrow{p_{k-1}}}
\end{aligned}$$

In step one we used $\overrightarrow{\mathbf{w}_{k+1}} = -\overleftarrow{\mathbf{w}_k}$ and reordered the terms. In the second step, the first product was expanded and equal terms were canceled out. Comparing the last line of each ratio we see their equivalence. This applies to unidirectional cases ($s = 0$ or $t = 0$) in the same way. $\square$

4 Implementation Consequences

In practice, the terms in Eq. (4) are computed recursively allowing to store intermediate results in sub-paths and having higher robustness. With this approach both weights lead to an identical implementation, because expanding the ratios leads to the same expression as shown in the proof.

Instead, path throughput weights $\hat{w}_i$ are a tool to check whether the computed weights are correct or optimal. They show that all modifications to the estimated radiance need to be accounted for in the weight computation. To name some of them:

Russian Roulette To increase the efficiency of a renderer it is common to randomly terminate paths. This termination probabilities must be included into the segment probabilities.

Shading Normal Correction Shading normals[1], which are different from geometric normals, need a correction factor for a reciprocal light transport (see Veach's thesis [Vea97] chapter 5.3). It must be applied inversely to the probability densities for MIS computation, too.

Random Connection The most used implementation of BPT traces one view-path for every light-path and connects them deterministically in all possible ways. However, in GPU implementations often a single random decision is made [DKHS14]. Then, a connection has a probability of $1/\bar{l}$ with $\bar{l}$ being the average path length of paths which can be chosen by the process.

While some of these events are included in other renderers (e.g. Russian Roulette in the VCM implementation (see supplemental of [GKDS12])) we have never seen that the shading normal correction is included in the MIS computation.

Using $\hat{w}$ as a tool can also help in unifying different sampling approaches. For example, combining BPT with Vertex Merging (i.e. photon gathering) is more intuitive than for the probability based approach. It took several years and multiple authors [GKDS12, HPJ12] to derive a unified weight computation which allowed the combination of BPT with merges. In its essence, the required modification is the multiplication of the query area πr^2 to the path probability. In the path throughput this quantity is contained naturally. Thus, comparing connection paths with merge paths using S unifies the two sampling approaches trivially.

However, there are also cases where our formulation does not help. For example, another possibility of using random connections was used by Popov et al. [PRDD15]. They connected each view sub-path to multiple light sub-paths to increase the path reuse. The problem here is the correlation between paths due to the choice of connections. The authors derived an upper bound for the variances to obtain a more robust MIS calculation. This problem remains equally hard regardless of the MIS formulation.

---

[1]Smooth interpolated normals on a triangle or bump mapping modified normals

5   Results

We derived a new form of the *Balance Heuristic* (Eq. (6)) which opens new perspectives to the path combination problem. By using path throughputs, many effects, which must be included in the MIS weight, become more obvious than before. Whenever a new quantity is multiplied with the path throughput, the path probability should be divided with this quantity.

In future the new perspective may also help in finding better combination heuristics for old and new sampling techniques such as MCMC.

References

[DKHS14] Tomáš Davidovič, Jaroslav Křivánek, Miloš Hašan, and Philipp Slusallek. Progressive Light Transport Simulation on the GPU: Survey and Improvements. *ACM Transactions on Graphics (TOG)*, 33(3):29, 2014.

[EMLB17] Víctor Elvira, Luca Martino, David Luengo, and Mónica F Bugallo. Generalized Multiple Importance Sampling. *arXiv:1511.03095v2*, 2017.

[GKDS12] Iliyan Georgiev, Jaroslav Křivánek, Tomáš Davidovič, and Philipp Slusallek. Light Transport Simulation with Vertex Connection and Merging. *ACM Transactions on Graphics (TOG)*, 31(6):192:1–192:10, 2012.

[HPJ12] Toshiya Hachisuka, Jacopo Pantaleoni, and Henrik Wann Jensen. A Path Space Extension for Robust Light Transport Simulation. *ACM Transactions on Graphics (TOG)*, 31(6):191:1–191:10, November 2012.

[Jen96] Henrik Wann Jensen. Global Illumination using Photon Maps. In *Proc. of Eurographics Workshop on Rendering (EGWR)*, pages 21–30. Springer, 1996.

[Kaj86] James T. Kajiya. The Rendering Equation. In *Computer Graphics (Proc. SIGGRAPH)*, SIGGRAPH, pages 143–150. ACM, 1986.

[PRDD15] Stefan Popov, Ravi Ramamoorthi, Fredo Durand, and George Drettakis. Probabilistic Connections for Bidirectional Path Tracing. *Computer Graphics Forum (Proc. of EGSR)*, 34(4):75–86, 2015.

[SHSK16] M. Sbert, V. Havran, and L. Szirmay-Kalos. Variance Analysis of Multi-sample and One-sample Multiple Importance Sampling. *Computer Graphics Forum (CGF)*, 35(7):451–460, October 2016.

[ŠOHK16] Martin Šik, Hisanari Otsu, Toshiya Hachisuka, and Jaroslav Křivánek. Robust Light Transport Simulation via Metropolised Bidirectional Estimators. *ACM Transaction on Graphics (TOG)*, 35(6):245:1–245:12, 2016.

[Vea97] Eric Veach. *Robust Monte Carlo Methods for Light Transport Simulation*. PhD thesis, Stanford University, 1997.

[VG95a] Eric Veach and Leonidas J. Guibas. Bidirectional Estimators for Light Transport. In *Photorealistic Rendering Techniques*, pages 145–167. Springer Berlin Heidelberg, 1995.

[VG95b] Eric Veach and Leonidas J. Guibas. Optimally Combining Sampling Techniques for Monte Carlo Rendering. In *Proceedings of SIGGRAPH '95*, SIGGRAPH, pages 419–428. ACM, 1995.