多重重要性采样

Dezeming Family

2022 年 9 月 15 日

DezemingFamily 系列文章和电子书**全部都有免费公开的电子版**，可以很方便地进行修改和重新发布。如果您获得了 DezemingFamily 的系列电子书，可以从我们的网站 [https://dezeming.top/] 找到最新的版本。对文章的内容建议和出现的错误也欢迎在网站留言。

本文是对 [1] 论文的详细解读，该论文可以说是渲染必读论文，但有些符号表示和描述可能对初学者并不友好，由于里面介绍的几种重要技术，例如双向路径追踪、多重重要性采样以及 Metropolis 方法都是非常重要的，因此我打算写一下本论文的解读，作为构建“高端图形渲染学习体系”的一个重要组成部分。

由于论文 [1] 篇幅过长，为了减少 Latex 编译的时间以及更好把控不同部分的内容，我将整个论文划分为了多本小册子来进行讲解。

本文的预备知识：蒙特卡洛方法、蒙特卡洛光线追踪（可以看 Peter Shirley 的光线追踪三本小书）、BSDF 模型、路径追踪、向量空间。

目录

一　基本内容介绍

被积函数有时是非连续、高阶以及奇异的（可以搜索“奇异积分”了解），一种采样策略很难非常合适。多重重要性采样就是把多种采样方法以一定权重组合起来，即把不同方法采样到的样本进行组合（分配不同的权重），我们的目标是通过合适地选择权重函数找到一个最小化方差的估计器。

粗糙的表面和小光源的情况适合直接对光采样，光滑的表面和大光源的情况适合对 BSDF 采样，但我们很难在场景中去合理定义粗糙和光滑的界限，以及很难定义光源大小的界限。一个很合适的方法是采样某个点上的直接光照时，分别采样一次表面 BSDF 和光源点，然后把两次估计的值取平均。但在我们学习了多重重要性采样以后，我们会发现这种取平均的方式并不是最好的方法，因为多重重要性采样得到的方差会更小一些。

二　光传输的问题

2 1　光泽高光 (glossy highlights) 问题

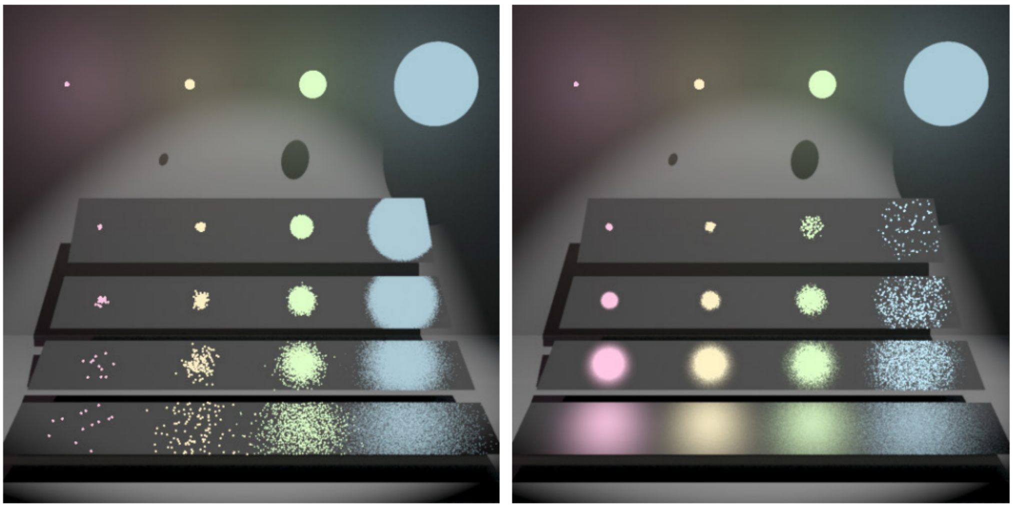
设当前表面点为 $\mathbf{x}'$。采样 BSDF，即根据预定义的概率密度 $p(\omega_i')$ 采样一个方向 ω_i'：

$$p(\omega_i') \propto f_s(\mathbf{x}', \omega_i' \to \omega_o') \tag{二.1}$$

采样光源，就是从光源上采样一个点 $\mathbf{x}$，然后采样点的概率密度 $p(\mathbf{x})$ 转化为采样到这个方向的概率密度 $p(\omega_i')$。注意这个概率密度的计算（见《光传输的路径积分与符号表示》）：

$$p(\mathbf{x}') = p(\omega_i')\frac{d\sigma^{\perp}(\omega_i')}{dA(\mathbf{x})} = p(\omega_i')\frac{|\cos\theta_o\cos\theta_i'|}{||\mathbf{x}-\mathbf{x}'||^2} \tag{二.2}$$

下面左右两幅图分别是采样 BSDF 和采样光源：



材质从上到下光滑度不断降低，也就是说最下面的板子最粗糙，最上面的板子最光滑。光滑的板子适合采样 BSDF，而粗糙的板子更适合采样光源。

多重重要性采样就是不但采样 BSDF，还采样光源。它旨在寻找最好的方式来将采样得到的样本组合起来得到方差更小的结果（不仅仅是将采样 BSDF 与采样光源得到的结果取平均）。

三　多重重要性采样

多采样模型 (multi-sample model) 将多种采样技术组合起来，组合权重需要保证结果是无偏的。可以得到无数种组合方法，但其中按理来说应该有一种方法是最好的。我们试图找到这种最好的方法。

3 1　多采样模型

考虑对 $f:\Omega \to \mathbb{R}$，我们计算积分：

$$\int_{\Omega} f(x) d\mu(x) \tag{三.1}$$

我们使用在 Ω 上的 n 个不同的采样技术，概率密度函数分别是 $p_1, ..., p_n$。我们假设仅下面的几种操作能够实现：

- 给定任何点 $x \in \Omega$，$f(x)$ 和 $p_i(x)$ 都能计算得到。
- 对于给定 p_i 分布，能够产生符合该分布的样本 X。

设 n_i 表示从 p_i 分布采样到的样本数，$n_i \geq 1$，我们令 $N = \sum n_i$ 表示样本总数。假设采样样本数是固定已知的（暂时先不考虑如何给不同的技术分配不同的样本量，放在本章后面介绍）。从技术 i 中采样到的样本标记为 $X_{i,j}$，$j = 1, ..., n_i$（第 i 种技术产生 n_i 个样本）。所有的样本假设都是独立的。

设允许不同的样本有不同的权重，对于从 p_i 采样得到的样本权重为 w_i，多采样估计器就是：

$$F = \sum_{i=1}^{n} \frac{1}{n_i} \sum_{j=1}^{n_i} w_i(X_{i,j}) \frac{f(X_{i,j})}{p_i(X_{i,j})} \tag{三.2}$$

估计器相当于把 $f(X_{i,j})/p_i(X_{i,j})$ 求加权和。注意这个权重并不是一个常量，而是随着 $X_{i,j}$ 变化。

这个权重需要满足两点：

$$\begin{aligned} &\sum_{i=1}^{n} w_i(x) = 1 \ whenever \ f(x) \neq 0 \\ &w_i(x) = 0 \ whenever \ p_i(x) = 0 \end{aligned} \tag{三.3}$$

上面两个式子暗示了在任何 $f(x) \neq 0$ 的位置至少有一个 $p_i(x)$ 必须是正的（也就是说并不需要所有的采样技术的采样区间都能覆盖整个函数的区间，所有的采样技术的区间之和覆盖整个采样区间即可）。

当 $n_i \geq 1$，可以得到上面的估计器是无偏的：

$$\begin{aligned} E[F] &= \sum_{i=1}^{n} \frac{1}{n_i} \sum_{j=1}^{n_i} E\big[w_i(X_{i,j}) \frac{f(X_{i,j})}{p_i(X_{i,j})}\big] \\ &= \sum_{i=1}^{n} \frac{1}{n_i} \sum_{j=1}^{n_i} \int_{\Omega} \frac{w_i(x) f(x)}{p_i(x)} p_i(x) d\mu(x) \\ &= \int_{\Omega} \sum_{i=1}^{n} w_i(x) f(x) d\mu(x) \\ &= \int_{\Omega} f(x) d\mu(x) \end{aligned} \tag{三.4}$$

注意上式中，对于一个估计器而言，$\int_{\Omega} f(x) d\mu(x)$ 就是 $\frac{f(x)}{p_i(x)}$ 的期望。

3 2　一个权重函数的例子

假设有三个采样技术 p_1，p_2，p_3。每种采样技术采样一个样本，假设权重函数是常量，因此：

$$F = w_1 \frac{f(X_{1,1})}{p_1(X_{1,1})} + w_2 \frac{f(X_{2,1})}{p_2(X_{2,1})} + w_3 \frac{f(X_{3,1})}{p_3(X_{3,1})} \tag{三.5}$$

由于：

$$V[F] = w_1V[F_1] + w_2V[F_2] + w_3V[F_3]$$
$$w_1 + w_2 + w_3 = 1 \tag{三.6}$$

因此，若一种采样技术方差很大，则总方差就会很大。所以用常量来作为权重并不是一种很好的方式（正如前面所说，可以同时采样光源和 BSDF，然后把这两种采样的结果取平均，这种直接取平均的权重方式方差很高，不建议使用）。

注意需要保证：

$$w_i(x) = 0 \; whenever \; p_i(x) = 0$$

也就是说当 w_i 都是常数时，要保证每种采样技术下采样到的样本的概率密度 p_i 的值都大于 0。

3 3 多采样模型的泛化

重写多采样估计器：

$$F = \sum_{i=1}^{n}\sum_{j=1}^{n_i} C_i(X_{i,j}) \tag{三.7}$$

其中，$C_i(X_{i,j})$ 表示对样本 $X_{i,j}$ 的采样贡献。要想使 F 是无偏的，则需要满足：

$$f(x) = \sum_{i=1}^{n} n_iC_i(x)p_i(x) \tag{三.8}$$

多采样模型可以表示任何无偏的组合策略，仅仅要求样本是独立的即可。我们回顾一下之前的公式：

$$F = \sum_{i=1}^{n}\frac{1}{n_i}\sum_{j=1}^{n_i} w_i(X_{i,j})\frac{f(X_{i,j})}{p_i(X_{i,j})} \tag{三.9}$$

也就是说：

$$C_i(X_{i,j}) = \frac{w_i(X_{i,j})f(X_{i,j})}{n_ip_i(X_{i,j})}$$
$$\sum_{i=1}^{n} n_iC_i(X_{i,j})p_i(X_{i,j}) = \sum_{i=1}^{n} w_i(X_{i,j})f(X_{i,j})$$
$$\Longrightarrow \sum_{i=1}^{n} n_iC_i(x)p_i(x) = \sum_{i=1}^{n} w_i(x)f(x) = f(x)$$

注意公式里的 i 表示第 i 种采样技术。

为了更清楚表示多采样模型，我们把式子展开，设 $n = 3$，$n_1 = 2$，$n_2 = 3$ 以及 $n_3 = 4$：

$$\begin{aligned}
F =& \sum_{i=1}^{n}\frac{1}{n_i}\sum_{j=1}^{n_i} w_i(X_{i,j})\frac{f(X_{i,j})}{p_i(X_{i,j})} \\
=& \quad \frac{1}{2}\left(w_1(X_{1,1})\frac{f(X_{1,1})}{p_1(X_{1,1})} + w_1(X_{1,2})\frac{f(X_{1,2})}{p_1(X_{1,2})}\right) \\
&+ \frac{1}{3}\left(w_2(X_{2,1})\frac{f(X_{2,1})}{p_2(X_{2,1})} + w_2(X_{2,2})\frac{f(X_{2,2})}{p_2(X_{2,2})} + w_2(X_{2,3})\frac{f(X_{2,3})}{p_2(X_{2,3})}\right) \\
&+ \frac{1}{4}\left(w_3(X_{3,1})\frac{f(X_{3,1})}{p_3(X_{3,1})} + w_3(X_{3,2})\frac{f(X_{3,2})}{p_3(X_{3,2})} + w_3(X_{3,3})\frac{f(X_{3,3})}{p_3(X_{3,3})} + w_3(X_{3,4})\frac{f(X_{3,4})}{p_3(X_{3,4})}\right)
\end{aligned}$$

3 4 平衡启发式 (balance heuristic)

3 4.1 基本原理

平衡启发式 (balance heuristic) 就是选用下面的权重函数：

$$\hat{w}_i(x) = \frac{n_ip_i(x)}{\sum_k n_kp_k(x)} \tag{三.10}$$

我们可以把这个策略叫做平衡启发式，这种组合策略的优势在于它的方差足够小：

$$V[\hat{F}] - V[F] \le \Big(\frac{1}{\min_i n_i} - \frac{1}{\sum_i n_i}\Big)\mu^2 \tag{三.11}$$

其中，$\hat{F}$ 表示使用平衡启发式方法组合权重；F 表示任意一种其他组合方法；$\mu = E[\hat{F}] = E[F]$ 表示被估计量。上式说明了平衡启发式方法组合权重比最优的权重组合方法之间的方差不会大于某个值（该值即 $\frac{\mu^2}{8}$）。证明可见论文原文。

对给一个更符合直觉的理解，假设有两个采样技术，每个采样技术采样的样本量为 $n_1 = n_2 = 4$，则：

$$V[\hat{F}] - V[F] \le \Big(\frac{1}{4} - \frac{1}{4+4}\Big)\mu^2 = \frac{\mu^2}{8} \tag{三.12}$$

也就是说平衡启发式方法比最优方法在方差上的差距不会大于 $\frac{\mu^2}{8}$。这个方差水平相当于发射 8 条阴影光线到面光源，其中会有 50% 的概率被遮挡（设被遮挡为 0，没有被遮挡为 1，$\mu = 0.5$）。

注意当两种采样技术采样的样本量都增加时，这个差距会趋近于 0，相当于会逼近最优采样器。

3 4.2 平衡启发式与 MC 的关系

用平衡启发式方法重写一下多采样估计器：

$$\begin{aligned}
F &= \sum_{i=1}^{n} \frac{1}{n_i} \sum_{j=1}^{n_i} w_i(X_{i,j}) \frac{f(X_{i,j})}{p_i(X_{i,j})} \\
&= \sum_{i=1}^{n} \frac{1}{n_i} \sum_{j=1}^{n_i} \Big(\frac{n_i p_i(X_{i,j})}{\sum_k n_k p_k(X_{i,j})}\Big) \frac{f(X_{i,j})}{p_i(X_{i,j})} \\
&= \sum_{i=1}^{n} \sum_{j=1}^{n_i} \frac{f(X_{i,j})}{\sum_k n_k p_k(X_{i,j})} \\
&= \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{n_i} \frac{f(X_{i,j})}{\sum_k c_k p_k(X_{i,j})}
\end{aligned}$$

其中，$N = \sum_i n_i$ 表示样本总数，$c_k = n_k/N$ 表示从采样技术 k 中采样的样本占总样本的比例。该式可以对应于标准的蒙特卡洛估计器形式 f/p，它可以通过重写分母来得到：

$$\hat{p}(x) = \sum_{k=1}^{n} c_k p_k(x) \tag{三.13}$$

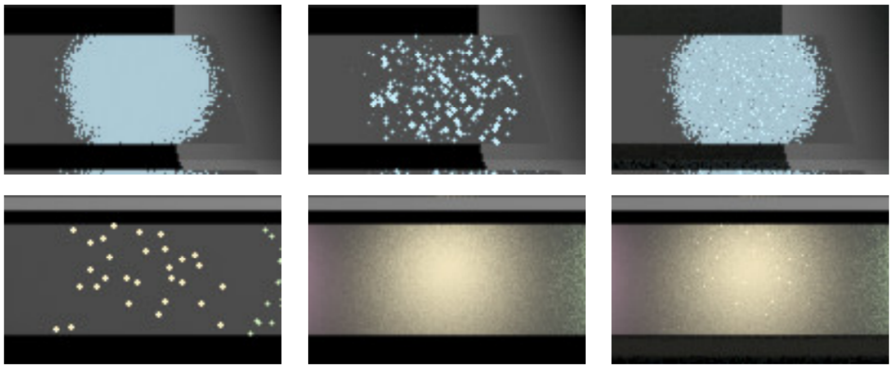
$\hat{p}(x)$ 被称为“组合样本概率密度”

四　更好的组合策略

我们上一章节知道平衡启发式方法与最优估计器相比还是多一些方差 $\frac{\mu^2}{8}$，本节研究更好的组合策略。

4 1　平衡启发式的一些问题

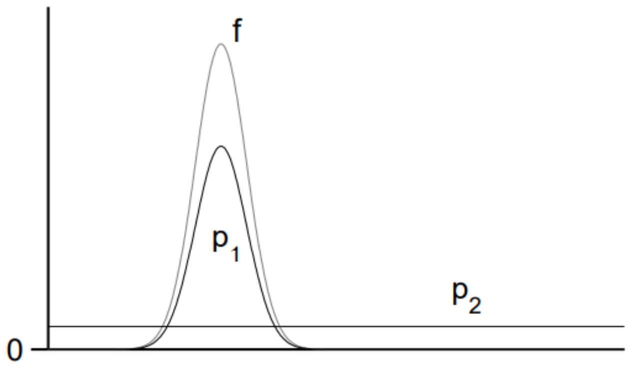
下图表示使用原来的采样方法和平衡启发式以后分别得到的效果，每种采样技术只采样一个样本（共两种采样技术：采样 BSDF 和采样光源）。上面这一排是大光源 + 光滑表面；下面这一排是小光源 + 粗糙表面：



(a) Sampling the BSDF　　**(b)** Sampling the lights　　**(c)** The balance heuristic

我们可以看出，平衡启发式会比最差的方法好，但是并不会比最好的方法好，这种现象叫做“低方差问题 (low-variance problems)”，这种问题出现在当某个 p_i 和被积函数 f 非常相符时，这会生成一个方差接近于 0 的估计器，优于平衡启发式方法。由于实际采样中我们很难确定哪种采样方式是最优的，因此无法将最优采样的样本权重置 1 而其他样本权重置 0。

考虑下图，有两种采样技术 p_1 和 p_2，概率密度 p_1 与 f 成比例关系，而 p_2 只是一个常数函数，此时，最优采样很明显应该是：$w_1^* \equiv 1$，$w_2^* \equiv 0$，这会得到一个方差为 0 的估计器。



我们下面看一下平衡启发式为什么会引入方差。

4 2　更好的组合策略

本节由于时间关系暂时没有安排写作内容，以前写作的分析可以参考博客：
https://feimo.blog.csdn.net/article/details/109675133

4 2.1　组合策略

4 2.2　方差界定

4 3　单采样模型

单采样模型 (one-sample model) 中，随机从 n 种采样技术中抽取一种并从该技术中采样一个样本。我们需要通过给单采样技术采样到的样本赋以不同的权重。

设 n 种采样技术 $p_1, ..., p_n$，每一种技术被抽中的概率分别是 $c_1, ..., c_n$，且 $\sum_i c_i = 1$。把选择一个采样技术、采样一个样本以及计算权重的过程叫做单采样估计器：

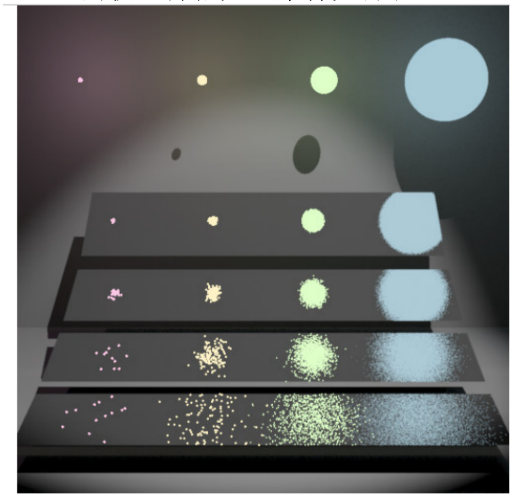
$$F = \frac{w_I(X_I) f(X_I)}{c_I p_I(X_I)} \tag{四.1}$$

w_i 还是要满足前面提出的无偏条件。在这里应用平衡启发式将是最优的权重计算方法。

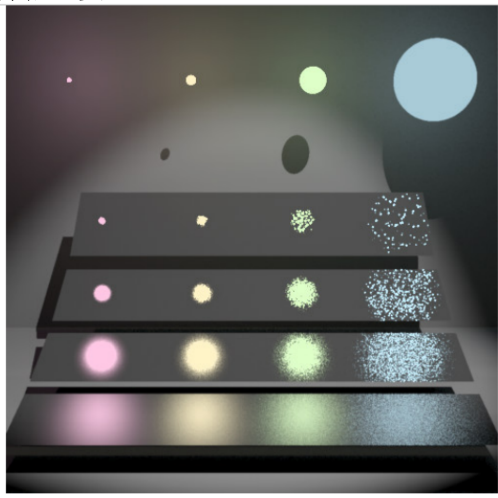
4 4 应用启发式方法的效果

多重重要性采样可以应用于两个重要领域：分布式光线追踪 (distribution ray tracing) 的光滑表面高光问题和最终的聚集过程 (final gather pass) 问题。我们只介绍分布式光线追踪，最终聚集过程问题我没有接触过，所以不进行讲解。

下图的图示中，表面粗糙度从上到下依次增加。在图 (d) 中用来将权重分配可视化，对于绿色表示采样光源的权重更大，红色表示采样 BSDF 的权重更大；由于红色 + 绿色会显示出黄色，因此黄色表示采样光源和 BSDF 的权重都很大，采样光源和 BSDF 都很重要。

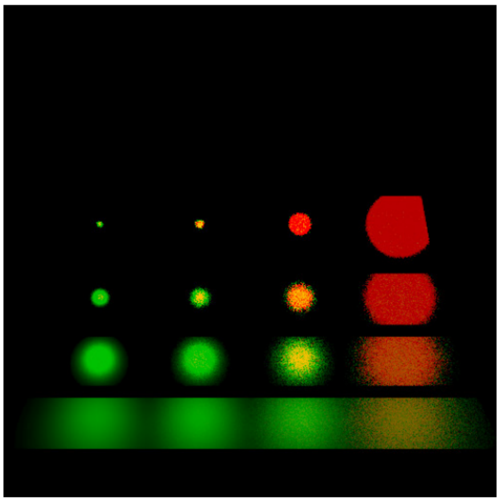


(a) Sampling the BSDF



(b) Sampling the light sources

(c) The power heuristic with $\beta = 2$.



(d) The weights used by the power heuristic.

从论文原文 [1] 给出的曲线图中也可以感受到，平衡启发式总体来说效果都比较好。

五　其他问题分析与讨论

关于 MIS 方法与基本的 MC 方法和分层抽样方法 (stratified sampling) 的联系，MIS 可以被认为是重要性采样和分层抽样的泛化。

因为它可以应用多种采样技术，所以相当于重要性采样的泛化。由于 MIS 从 n 个给定的区域 Ω_i 中抽取样本，所以它是分层抽样的泛化，注意这些区域不一定是不相交的，只需要保证这些区域加起来能构成整个函数区间即可。这种分层策略很有用，比如当 BSDF 是由多个组件（比如漫反射组件、镜面组件）构成时，从每个组件抽取几个样本就是分层抽样的思想。

给采样技术分配样本量

假设总共 N 个样本被均分到 n 种采样技术中，设 $\hat{F}$ 表示每种采样技术分配均匀的样本量这种采样技术，F 表示任意一种技术。根据论文证明：

$$V[\hat{F}] \leq nV[F] + \frac{n-1}{N}\mu^2$$

也就是说除了后面的一个附加项 $\frac{n-1}{N}\mu^2$，前面的方差项顶多会比最好的方法多出 n 倍；而如果权重选择不当，则根据我们前面章节的描述，可能方差会大任意数量级。

也就是说，n_i 的选择，即给一个采样技术分配多少样本，并不如如何适当选择权重函数那么重要。

5 1　本文小结

本文中介绍的多重重要性采样 (MIS) 技术是提高后面介绍的双向路径追踪方法的效率的关键方法。后面的文章会详细介绍双向路径追踪技术的基本方法，以及一些需要注意的细节。

参考文献

[1] Veach E . Robust Monte Carlo Methods for Light Transport Simulation[J]. Ph.d.thesis Stanford University Department of Computer Science, 1998.

[2] Arvo, J. [1995]. Analytic Methods for Simulated Light Transport, PhD thesis, Yale University.