PDF method report

谭钧豪

July 4, 2025

1 引言

在尝试进行照片级的真实渲染中，我们利用光路可逆的特点，从相机出发投射光线，模拟光路，较为精确的模拟出了包括软阴影、颜色、焦散等各种光学现象，得到了物理上较为逼真的结果。但实现的过程中有一个不可忽视的问题：如果存在一个小且高亮的光源，因为我们对漫反射的处理是仅考虑漫反射本身的特性，采用Lambertian 分布，绝大多数随机投射的光线会错过微小的光源，导致采样效率极低，图像充满难以收敛的噪点。不过如果我们在计算光的散射时合理的更偏向于光源，我们就可以处理这种情况，减少噪点，提升图片质量，这也就是PDF（概率密度函数）所在做的事。

2 PDF(probability density function)

概率密度函数（PDF）定义为[锈雀, 2024]：对于连续型随机变量X，若存在非负可积函数$f(x)$ 满足其累积分布函数$F(x)$ 的导数为$f(x)$，即：

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$$

则$f(x)$ 称为X 的概率密度函数。注意到按照定义，应该有$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$

3 蒙特卡洛积分

蒙特卡洛积分是一种使用随机抽样来估算定积分数值的方法，其形式可以表示为：

$$\int_a^b f(x)\mathrm{d}x \approx \frac{1}{N}\sum_{i=1}^{N}\frac{f(x_i)}{p(x_i)}$$

其中$f(x_i)$ 是被积函数在x_i 处的值，$p(x_i)$ 是在x_i 处的概率密度函数（PDF）的值。且随着N 的增大，右侧的值会趋近于左侧

举个例子，我们在光线追踪时可能需要计算一个点收到的来自周围的所有方向的光照综合。那这对应于在一个球面上对函数进行积分。那此时我们使用蒙特卡洛积分，我们随机采样球面上的单位向量，然后因为单位球面表面积为4π，所以每个位置被抽中的概率是一样的，也就是$p(s) = \frac{1}{4\pi}$，所以积分的结果可以有：

$$\int f(x)\mathrm{d}x \approx \frac{4\pi}{N}\sum_{i=1}^{N} f(x_i)$$

前面我们使用Lambertian 分布来处理漫反射，我们可以给其散射光赋上PDF p 来描述对光源的偏向性，那我们就可以用蒙特卡洛积分来估计其Lambertian 反射的结果。如果一条光线X 进行散射，出射光是x_i，$scatter(x_i)$ 是产生了x_i 这条出射光的概率，$f(x)$ 是得到x 这条光线的颜色，$A(x)$ 是x 这条光线的反照率。那我们只需要找到$A(x_i)scatter(x_i)f(x_i)$ 的PDF 函数$p(x_i)$，就可以有：

$$f(X) = \int A(x_i)scatter(x_i)f(x_i)\mathrm{d}x \approx \frac{1}{N}\sum_{i=1}^{N}\frac{A(x_i)scatter(x_i)f(x_i)}{p(x_i)}$$

4 光源PDF

我们先讨论平行四边形的光源。

对于一个面积为A 的光源，我们考虑光源表面均匀采样的情况，那光源表面的PDF 就是$\frac{1}{A}$。

为了求出一个点P 附近散射光的PDF，我们求这个光源到P 点的邻域球的投影：

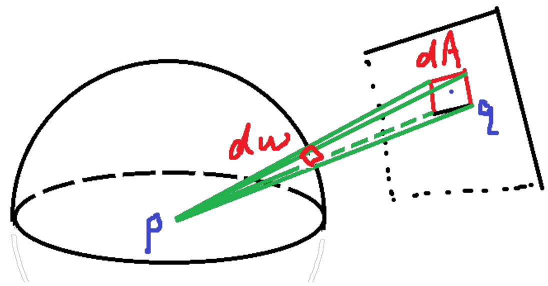

Figure 1: Projection of light shape onto PDF

对于光源上的一个小区域$\mathrm{d}A$，取到其中的概率为$p_q(q)\cdot\mathrm{d}A$，其投影到球面的小区域为$\mathrm{d}\omega$，其取到的概率可以表示为$p(\omega)\cdot\mathrm{d}\omega$。根据投影，有$\mathrm{d}\omega=\frac{\mathrm{d}A\cdot\cos(\theta)}{d^2(p,q)}$[Peter Shirley, 2025]

而由于二者的对应关系，取到其中的两个概率必须相等，也就是$p_q(q)\cdot\mathrm{d}A=p(\omega)\cdot\mathrm{d}\omega$。

而根据光源表面的PDF，$p_q(q)=\frac{1}{A}$。最终可以解得$p(\omega)=\frac{d^2(p,q)}{\cos(\theta)\cdot A}$

根据上面的式子可以看出，如果光源表面不均匀采样，我们所需要做的就是修改$p_q(q)$，然后再解出式子即可。

5 混合PDF

注意到，上面的光源PDF 并没有使用漫反射的Lambertian 分布，而实际上，我们仍然需要Lambertian 分布的实际规律。

于是我们类似的构造出Lambertian 分布的PDF，并希望结合两个PDF，使得其既有漫反射的实际规律，又能很好的减少噪点，提高效率。

关于p 函数的混合，计算颜色都并不难解决，直接通过平均即可，但是我们要如何生成出反射的光线呢？因为我们并不能很容易的知道我们要用哪个PDF 来生成我们的光线方向。

实际上，我们只需要使用随机，随机的从两种PDF 生成光线方式里选一个生成即可。

6 利用PDF生成随机元素

如果我们有一个PDF：$f : [0,1] \to R$，那我们希望通过这个PDF 来生成一个$[0,1]$ 的随机数，他的生成满足PDF 的分布。

一个直观的想法是直接将PDF 积分后取反函数，具体来说，令$F(x) = \int_0^x f(t)\mathrm{d}t$，令$G(x) = F^{-1}(x)$，我们生成一个$[0,1]$ 等概率生成的随机数y，然后$G(y)$ 就可以作为一个满足PDF 分布的随机数了。

不过实际应用上并没有这么简单，因为实际的函数可能难以求反函数，甚至可能难以积分。所以实际应用上，我们考虑吧函数线性化。

我们通过足够多的数据样本x_i 可以得到足够多的$f(x_i)$，然后得到其中的中位数$f(x_{mid})$，于是就可以拿$(0,0),(0.5,f(x_{mid})),(1,1)$ 三个点连线做反函数的近似。

当然，注意到这是可以进一步化归的，所以可以通过左右递归进一步近似。

在对于前面生成的漫反射PDF，我们可以使用前者的方法:

对于球面，我们可以使用极坐标系(φ,θ)，得到如下表示:

$$\begin{cases} x = \cos(\varphi)\sin(\theta) \\ y = \sin(\varphi)\sin(\theta) \\ z = \cos(\theta) \end{cases}$$

注意到散射出来的光线，如果和平面法向量的夹角相同，其概率应该是相同的，所以可以令z 轴与法向量平行，并一次建立坐标系，则可以有φ 的PDF 是$a(\varphi) = \frac{1}{2\pi}$

再根据Lambertian 分布，有$r_2 = \int_0^\theta 2\pi f(\theta')\sin(\theta')\mathrm{d}\theta$,其中$f$ 应该为常数，解得$\cos(\theta) = 1 - \frac{r_2}{2\pi C}$

如果知道θ 的最大角度$\theta_{\max}$，由于此时应该有$r_2 = 1$，可以解得：$C = \frac{1}{2\pi(1-\cos(\theta_{\max}))}$

于是可以进一步表示:

$$\begin{cases} z = 1 + r_2(\cos(\theta_{\max}) - 1) \\ x = \cos(2\pi r_1)\sqrt{1-z^2} \\ y = \sin(2\pi r_1)\sqrt{1-z^2} \end{cases}$$

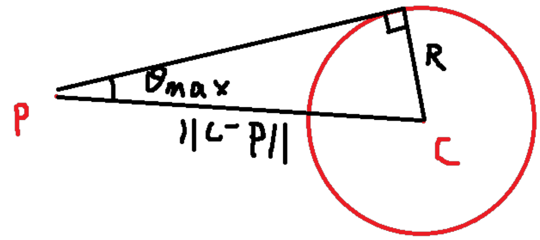

Figure 2: A sphere-enclosing cone

那么我们所还需要做的就是求出$\theta_{\max}$：

根据Figure 2，我们自然能表示出：$\theta_{\max} = \sqrt{1 - \frac{R^2}{||\mathbf{c}-\mathbf{p}||^2}}$

References

Steve Hollasch Peter Shirley, Trevor David Black. Ray tracing: The rest of your life, 2025. URL https://raytracing.github.io/books/RayTracingTheRestOfYourLife.html.

锈雀. 关于概率密度函数的一些解释, 2024. URL https://zhuanlan.zhihu.com/p/693841710.