PDF Method Intro

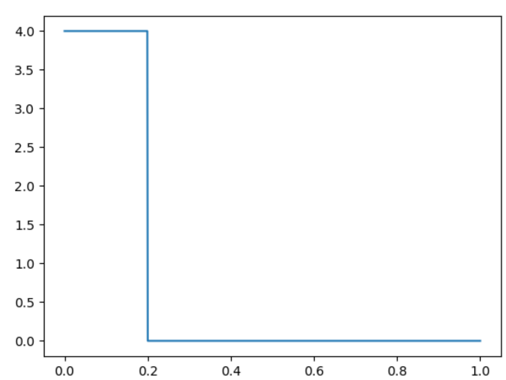
Jinxuan Zhu

July 11, 2024

1 What PDF aims to solve

In the first two books, we have learned how to render images using the ray tracing method. However, after lights are involved, the rendering yields noisy dark pixels unless we use a large sample count. This is because the light is emitted in all directions, and the probability of a ray hitting the light source is very low. The PDF method aims to solve this problem by providing a way to sample the light source more efficiently while following optics laws.

2 Monte Carlo method and importance sampling



The figure above briefly illustrates how light is distributed in all directions for a certain point in the scene. As we can see, in most directions, the light is very dim, while it's very bright in the direction of the light source. The final result is the sum of all the light in the scene, a.k.a. the integral of the light distribution function. To calculate the integral, the simplest way is to pick random directions for many times and calculate the average light intensity, which is known as the Monte Carlo method.

Consider what will happen if we choose directions uniformly. Most time you will just get 0.25, and the result will be very low; if you are lucky enough to get 4.0, the result can be very high. It takes an average of 5 samples to get the reasonable result 1.0.

What if we choose directions according to the light distribution function? We know that the result is high in the interval [0.0, 0.2]. If we choose more samples in the interval, we will be more likely to get 4.0. This is known as importance sampling.

However, by the method above we have a bias towards the light source, which will result in a result lighter than usual. Here comes the key point: we need to divide the result by the probability of choosing the direction. Now, while we choose more samples towards the light source, the result will also be divided by a larger probability, and we get an unbiased result.

3 Probability density function

The method discussed above is just an intuitive explanation. What is the "probability" of a direction? In fact, theoretically, the probability is always zero for a single direction. However, we have a mathematical way to describe the probability: probability density function (PDF).

The PDF function describes the probability of a random variable falling within a certain range: its integral over the range is the probability of a random variable falling within that. Its integral over the whole space is always 1.0.

With a given PDF we can generate a random number following the distribution; we can also calculate the PDF value (you can think of it as a normalized probability or weight) for a given number. Math details can be found in any probability theory book, and we will not discuss it here.

PDF also have two-dimensional versions, which use multiple integral to describe the probability of a random variable falling within a certain region. That's useful when sampling a direction in 3D space, which is actually two-dimensional. You project the direction onto a unit sphere and use solid angle to measure it.

4 Light model and calculation

Let's dive deeper into how light is scattered. Here is the formula given in the book:

$$\mathrm{Color}_o(\mathbf{x}, \omega_o, \lambda) = \int_{\omega_i} A(\mathbf{x}, \omega_i, \omega_o, \lambda) \cdot \mathrm{pScatter}(\mathbf{x}, \omega_i, \omega_o, \lambda) \cdot \mathrm{Color}_i(\mathbf{x}, \omega_i, \lambda)$$

We can see the function to be integrated is not only related to the input ray and albedo, but also the probability of the direction, which involves another PDF. If a direction is unlikely for scattering, the result will be multiplied by a smaller weight.

To calculate the integral, we need to sample direction according to a good PDF. Don't confuse this PDF we choose (sample PDF) with the PDF in the formula (scattering PDF). The sample PDF is determined by ourselves when choosing the direction, while the scattering PDF is determined by the material itself.

To avoid blowing up in calculation, we sample a single direction for the integral. The formula can be written as:

$$\mathrm{Color}_o(\mathbf{x}, \omega_o, \lambda) \approx \frac{A(\dots) \cdot \mathrm{pScatter}(\dots) \cdot \mathrm{Color}_i(\dots)}{p(\mathbf{x}, \omega_i, \omega_o, \lambda)}$$

a higher pScatter will result in a lighter result, which means the direction is more likely to be scattered by such material. On the contrary, a higher p will lead to a darker result: if a direction is chosen more frequently, we have to eliminate the bias.

5 How to choose the PDF

A good PDF should have higher values where the function to be integrated is higher. The albedo is generally constant, so we only need to consider pScatter and input color.

With respect to pScatter, we can directly use the scattering PDF given by the material. Sometimes we only need to consider pScatter when it is extreme and decisive, e.g. for a mirror which only reflects light in one direction. Here we simply follow the ray generated by the material: PDF is not needed.

For the input color, we should sample towards the light source: the PDF is a uniform sampling on the surface of light sources in the scene (implementation details can be found in the book). In fact, objects which do not emit light may also be considered as light sources, e.g. dielectric materials. They can reflect light from other sources intensively, and we should sample them as well. Conversely, lambertian materials should not be sampled, as they just reflect light in all directions.

If we only use pScatter for PDF, pScatter and p will cancel out in the formula, and we will get the same result as the original method; if we only consider sampling towards the light, diffuse reflection will be neglected, resulting in a high contrast image. The final PDF is the mixture of two PDFs. The book use a simple linear combination to mix them with same weight.

6 Conclusion and discussion

The PDF method is not only a fast way to sample the light source, but also a good way to understand the light distribution in the scene. The method is a good example of how to use probability theory in computer graphics.

Note that the PDF method is a probability-based method and may or may not yield a good result. We naively sample the ray directly to the light source with no consideration of occlusion and reflection. The output pictures also have some random noise.

The PDF method can be improved by choosing sampling PDF more carefully; a more complex mixture mode can be used.