Volume 2021 (4002), Number 1 pp. 1–9 COMPUTER GRAPHICS *seminar*

Volumetric Ray Tracing

1062[1]

[1]OTH Regensburg, Germany

Abstract

1. Introduction

In Computer Graphics, objects usually are represented as a set of geometric primitives (e.g. triangles), displaying the surface of the object. However, this approach is not always suitable. For example, if the original data representation of the object is volumetric data (which might be produced by medical 3D scans [Lev88]), the traditional rendering technique would necessitate the creation of an intermediate surface representation that might introduce unwanted artifacts [Lev88]. Another issue arises if the object has no well-defined surfaces to which geometric primitives could be fitted, such as a cloud or fog [KVH84]. In such cases, volumetric ray tracing might be used, a technique in which rays are cast through a volume which contains information about its optical properties(e.g color and opacity), sampled at various points within the volume, accumulated and projected on a 2D image (see figure 1) [Lev88]. A description of this algorithm is presented in this paper, alongside several strategies for improving computational speed, such as (probabilistic) early termination of a ray, and sampling at lower resolutions within the volume.

2. Derivation Of The Rendering Equation

Volumetric ray tracing follows the same basic principle as classical ray tracing in which an image is rendered by spanning a pixel plane in front of the camera and casting one or multiple rays through each pixel. That means for a point $\vec{x}$ on the plane we cast a ray in direction $-\omega$ and calculate the amount of light $\vec{x}$ receives from direction ω, called $L(\vec{x},\omega)$ [Gla89]. To do this, we calculate $\vec{y}$, the closest intersection point of the ray with a piece of geometry. At $\vec{y}$, we calculate the amount of light transported from $\vec{y}$ in direction ω (either through reflection or emission), called $L_e(\vec{y},\omega)$ (The exact method for calculating L_e is a topic of classical raytracing [Gla89] and will not be further evaluated in this paper. In the following, we assume L_e to be a known quantity). Thus, we get the equation

$$L(\vec{x},\omega) = L_e(\vec{y},\omega) \tag{1}$$

However, this equation assumes that the light has no interactions between $\vec{x}$ and $\vec{y}$, which is only true if the ray travels through a vacuum. If the light travels through a medium that interacts with it (called a participating medium), the interactions change the light along the ray [LW96]. In most cases, these interactions are small enough to be ignored, but in certain scenarios (e.g if there is fog, clouds or smoke present) they might have a noticeable effect. The types of interactions we need to consider are absorption, emission, in-scattering and out-scattering [Max95].

Absorption and out-scattering attenuate the light intensity along the ray, in-scattering and emission add to it.

2.1. Absorbtion And Out-Scattering

At first, let us consider only out-scattering and absorption, which occur due to tiny particles floating around in the volume (such as water droplets in clouds and fog, or dust particles in the air). Of course, these particles are far too numerous to be directly simulated, but their distribution in 3D space can be stochastically modeled (similar to how detailed surface structures can be modeled by micorfacets in classical ray tracing). QUESTION! We will follow Max [Max95] in our development of the stochastical model. To do so, consider a close-up look at a ray section traveling through the volume. This section can be assumed to be a cylinder with a base area E and a height Δh, through which the ray travels from top to bottom. Within this cylinder, there exists a certain number of out-scattering and absorbing particles, defined as $n_s = \rho_s E \Delta h$ and $n_a = \rho_a E \Delta h$ respectively, where ρ_s and ρ_a are the densities of the particles. From the top-down view, the particles will occupy an area of $n_s A = \frac{\rho_s A E \Delta h}{E}$ and $n_a A = \frac{\rho_a A E \Delta h}{E}$, respectively, if we assume that the particles do not overlap each other, which is reasonable if the densities and the height do not become too large. This can be simplified to $\rho_s A \Delta h$ and $\rho_s A \Delta h$, which gives us the fractions of light which are stopped in the cylinder, either through absorbtion or out-scattering. By letting Δh approach 0, we see that for each infinitessimaly small cylinder slice with height dh, the change in intensity is proportionally to $-(\rho_s A + \rho_a A)dh$. Formulated as a differential equation, this results in

$$dL = -(\rho_s(h)A + \rho_a(h)A)L(h)dh \tag{2}$$

At this point, we define the scattering coefficient $\mu_s = \rho_s(h)A$, the absorption coefficient $\mu_a = \rho_a(h)A$ and the extinction coeffcient

$\mu_t = \mu_s + \mu_a$, which give a measure of how much light is lost due to scattering, to absorption, and in total. Thus, equation [REFERNCE] can be simplified to

$$dL = -\mu_t(h)L(h)dh \tag{3}$$

which solution is

$$L(h) = L_0 e^{-\int_0^h \mu_t(s)ds} \tag{4}$$

Rearranging this equation gives the equation

$$\frac{L(h)}{L_0} = e^{-\int_0^h \mu_t(s)ds} \tag{5}$$

describing the ratio of the light that travels a distance h unimpeded. In the following, we call this quantity the transmittance, and refer to it as $\tau(\vec{x},\vec{x'})$ (the ratio of light that arrives at $\vec{x}$ from $\vec{x'}$). Thus, going back to equation 1], we can now specify how much L_e gets attenuated, and get the equation

$$L(\vec{x},\omega) = \tau(\vec{x},\vec{y})L_e(\vec{y},\omega) \tag{6}$$

This equation is still not acurate, since we also need to consider in-scattering and emission along the ray.

2.2. Emission And In-Scattering

For this purpose, let us consider any arbitrary point $\vec{u}$ on the ray from $\vec{x}$ to $\vec{y}$. At $\vec{u}$, the particles of the medium emit some light towards ω [Max95], which we call $\varepsilon(\vec{u},\omega)$. In our cylindrical model, the probability of finding an absorbing particle at position $\vec{u}$ is $\mu_a(\vec{u})$. We assume that the particles responsible for absorbtion are also responsible for emission [Max95], meaning the amount of light emitted at $\vec{u}$ towards ω is $\mu_a(\vec{u})\varepsilon(\vec{u},\omega)$. In-scattered light can arrive at $\vec{u}$ from every direction, which means we need to integrate over the sphere Ω surrounding $\vec{u}$ [JC98]:

$$\int_\Omega L(\vec{u},\omega`)d\omega` \tag{7}$$

However, not all light arriving at $\vec{u}$ is scattered towards ω. The so called phase function $f_p(\vec{u},\omega,\omega`)$ gives us the probality that light hitting a particle at $\vec{u}$ from $\omega`$ is reflected towards ω [JC98]. In this sense, it is analogous to the BRDF in classical ray tracing. Furthermore, the amount of light scattered towards ω also depends on the probability that in-scattering particles are present at $\vec{u}$ [JC98]. We assume that in-scattering particles are the same as out-scattering particles and get the equation

$$L_e^s(\vec{u},\omega) = \mu_s(\vec{u})\int_\Omega f_p(\vec{u},\omega,\omega`)L(\vec{u},\omega`)d\omega` + \mu_a(\vec{u})\varepsilon(\vec{u},\omega) \tag{8}$$

describing the total amount of light being added to the ray at $\vec{u}$. We call this quantity L_e^s and use the superscript s to differentiate it from L_e(which describes light reflection at solid surfaces). Not all of $L_e^s(\vec{u},\omega)$ arrives at $\vec{x}$, since the newly added light also needs to travel through the volume where it is attenuated by the transmittance $\tau(\vec{x},\vec{u})$. To get the total amount of light added to the ray, we sum up the light contribution of all points along the ray by integrating over it [ZRL*07]:

$$\int_{\vec{x}}^{\vec{y}} \tau(\vec{x},\vec{u})L_e^s(\vec{u},\omega)d\vec{u} \tag{9}$$

2.3. The Rendering Equation

Adding this to the light arriving from $\vec{y}$ gives the total amount of light arriving at $\vec{x}$ from ω.

$$L(\vec{x},\omega) = \int_{\vec{x}}^{\vec{y}} \tau(\vec{x},\vec{u})L_e^s(\vec{u},\omega)d\vec{u} + \tau(\vec{x},\vec{y})L_e(\vec{y},\omega) \tag{10}$$

This integral is complicated to solve, especially since the expression $L(\vec{x},\omega)$ appears on both sides of the equation. In the following, we will describe several methods for approximating a solution.

3. Ray Marching Algorithm

In this chapter we will discuss an algorithmic solution [Lev88, DH92] to a simplified version of the rendering equation we derived in the previous section.

3.1. Simplified Rendering Equation

In the following, we will ignore all scattering effects in the medium (meaning μ_s is 0) and assume that the medium only absorbs and emitts light [DH92], which can be understood as a scenario where all external light has been completly evenly scattered within the volume. This is analogous to only rendering ambient lighting in classical ray tracing, which assumes that the light is evenly distributed in the scene. Furthermore, we will also ignore all other geometry in the scene and only focus on the volume. The term $\tau(\vec{x},\vec{y})L_e(\vec{y},\omega)$, which describes the light contribution from the nearest intersection with a piece of geometry, is omitted. Thus, we arrive at the equation

$$L(\vec{x},\omega) = \int_{\vec{x}}^{\vec{y}} \tau(\vec{x},\vec{u})\mu_a(\vec{u})\varepsilon(\vec{u},\omega)d\vec{u} \tag{11}$$

For this chapter, we assume that the medium is contained within a cuboid boundary [DH92, Lev90]. Since there is no geometry, we can't cast the ray until it intersects with the geometry. Instead, the endpoint $\vec{y}$ of the ray will be the point where the ray leaves the cuboid boundary.

3.2. Discrete Approximation To The Simplified Equation

$L(\vec{x},\omega)$ is approximated by casting a ray in direction $-\omega$ and sampling it at evenly spaced points along the ray and summing up their contributions [Lev88]. The distance s between the sampling points should be smaller than the volumes nyquist limit[CITE], otherwise important details are missed[CITE]. We start sampling at the first point within the volume, which has a distance of s_0 to $\vec{x}$. The i-th sample point then is described by

$$p_i = \vec{x} + s_0(-\omega) + si(-\omega) \tag{12}$$

The light contribution of a point p_i is considered to be the light contribution of the ray segment between p_{i-1} and p_i.

$$\int_{p_{i-1}}^{p_i} \tau(p_0,p_{i-1})\tau(p_{i-1},p_{i'}\mu_a(p_{i'})\varepsilon(p_{i'},\omega)dp_{i'} \tag{13}$$

As a simplification, we assume all variables to be piecewise constant [DH92] on the ray segments. This yields

$$\prod_{1\le j\le i} (\tau(p_{j-1},p_j)) \cdot \mu_a(p_i)\varepsilon(p_i,\omega)s \tag{14}$$

for one line segment at i. The quantity $\mu_a(p_i)\varepsilon(p_i,\omega)s$ describes the emitted light from p_i and will from now on be refered to as the color $c(i)$. The product is the total attenuation (or transperancy) from $\vec{x}$ to p_i, $\tau(p_{j-1},p_j)$ is the attenuation from p_{j-1} to p_j and can be calculated by

$$\tau(p_{j-1},p_j) = e^{-s\mu_t(j)} \tag{15}$$

However, in computer graphics it is more common to use the opacity instead of the transperancy. Therefore, we will refer to $\tau(p_{j-1},p_j)$ as $1-\alpha(j)$ from now on, where $\alpha(j)$ is the opacity. Inserting this in formula for $\tau(\vec{x},p_i)$ yields $\prod_{1\leq j\leq i}(1-\alpha(j))$.

Again, we can refer to this quantity in terms of opacity rather than transperancy be using the equality

$$1-\beta(i) = \prod_{1\leq j\leq i}(1-\alpha(j)) \tag{16}$$

where $\beta(i)$ is the accumulated opacity between $\vec{x}$ and p_i. Using these results, we can approximate the integral in equation 11 as the sum

$$L(\vec{x},\omega) = \sum_{1\leq i\leq n}(1-\beta(i))c(i) \tag{17}$$

which can be calculated in a single for loop. Since $\beta(i)$ equals $\beta(i-1)(1-\alpha(i))$, it is not necessary to completly recalculate β for every step [DH92].

3.3. Underlying Volume Data

Due to out assumption that the variables μ_s, μ_a, μ_t and ε are piecewise constant, the opacity and color can easily be computed, if the volume provides such information. Often though, the volume is already defined in terms of color and opacity, in which case α and c can be sampled directly from it. In some cases, the volume might contain other information such as density (e.g. for medical 3D scans), in which case a preprocessing step [Lev88] must produce color and opacity. If the volume is defined as a ternary function, sampling a point x works simply by calculating the value of the function at x. If the volume is defined as a 3D array of voxels, the value must be found through interpolation (usually through trilinear interpolation [Lev88], but other methods such as monte carlo interpolation [HHCM21] are possible as well).

3.4. Direct Illumination

Until now, we have ignored scattering in the ray marching algorithm. In the following, we describe an approach to simulate single scattering (that is all light rays that are scattered only once) developed by Kajiya and Von Herzen [KVH84] based on the work of Blinn [**?**]. This works as a two-step process. The first step is computed on a volume $\mathcal{V}$ containing opacity and albedo information and a set of light sources $\{l_1,\dots,l_m\}$. Based on $\mathcal{V}$, one voxelized volume $\mathcal{V'}_k$ is created for every light source l_k in the following manner: For a voxel $\vec{x}$ in $\mathcal{V'}_k$ and light source l_k the amount of light $\vec{x}$ receives from l_k is calculated by attenuating the emmited light εl_k by the transmittance $\tau(\vec{x},l_k) = \int_{\vec{x}}^{l_k}(1-\alpha(\vec{x'}))d\vec{x'}$, which can be calculated as described above. The albedo $a(\vec{x})$ regulates how much of the arriving light is reflected at $\vec{x}$. Thus, the color of $\vec{x}$ is

$$c_k(\vec{x}) = a(\vec{x})\tau(\vec{x},l_k)\varepsilon(l_k) \tag{18}$$

Thus, the volume $\mathcal{V'}_i$ contains information about the amount of light, that is reflected from l_i at every point in space.

In the second step works very similar to ray marching as described by equation 17, with the only difference being the calculation of $c(i)$. For this, the values of $c_k(p_i)$ needs to be known for all k, which can be calculated by sampling and interpolating in $\mathcal{V'}_k$. $c_k(p_i)$ is the amount of light reflected at p_i, but only a portion is reflected towards ω (the direction to the camera). Thus, by scaling all c_k by the phase function and summing up their contribution, we can define $c(i)$ as

$$c(i) = \sum_{1\leq k\leq m} f_p(\omega,\omega_k,p_i)c_k(p_i) \tag{19}$$

with ω_k being the direction from p_i to l_k (ω_k is different for every l_k, which is why the different c_k have to be stored seperately).

This approach could be generalized to calculate multi-scattering (that is light that need two or more scattering events to reach the camera), but this would very quickly become to costly. Therefore, when rendering images with global illumination, more sophisticated algorithms are necessary.

4. Monte Carlo Ray Tracing and Global Illumination

A common technique in classical ray tracing to generate unbiased images with global illumination is the so-called monte carlo ray tracing (MCRT) approach which solves the rendering equation by stochastically sampling it (as opossed to uniform sampling used by the ray marching algorithm described above).

In this chapter, we describe a typical MCRT, explain certain techniques necessary for its formulation and describe further improvements to the MCRT.

4.1. General Monte Carlo Algorithm

We describe a typical MCRT such as used by Hofman et al. [CITE!]. The algorithm starts like[simiilar to?] the already mentioned bycasting a ray from the camera p_0 in direction $-\omega_0$. Then, a distance d_0 is sampled stochastically and a so called path vertex p_1 is created.From p_1, a new direction ω_1 is sampled. This process is repeated until a light source is hit or until the path gets terminated (ususally by Russian roulette). Furthermore, at every path vertex p_i, next event estimation is done by casting a shadow ray towards a randomly sampled point on a light source, called l_i. An illustration of the process can be seen in figure ???. Thus, equation ??? is approximated by

$$L(\vec{x},\omega_0) = \tag{20}$$

To compute this result, following quantities need to be known: The transmittance τ between the various path vertices and light sources, the directions ω_i, and the locations of l_i and p_i. The sampling of ω_i and l_i works like in classical ray tracing, using multipe importance sampling[CITE] of the phase function and light distribution. Sampling p_i and estimating the transmittance require novel approaches not used in normal ray tracing, two of which are described below.

4.2. Delta Tracking

To calculate $p_{i+1} = p_i - \omega t_t$ from a given p_i and ω_i, the distance t_i between those two points must be sampled. Recall that $\tau(p_i, p_{i+1})$ describes the ratio of light arriving from p_{i+1} at p_i without colliding with a particle. Therefore, the probability that a collision occurs between p_i and p_{i+1} is

$$1 - \tau p_i, p_{i+1} \tag{21}$$

This can be used to importance sample the distance t_i by inverting equation **??** and applyingthe resulting function to ξ (a random variable uniformly sampled from $[0,1]$). However, τ is defined as $e^{\int_0^{t_i} \mu_t(s)ds}$, which can't be easily inverted if the medium is heterogenous (if μ_t is not constant). Delta tracking (alos known as woodcock tracking) solves this problem by introducing so called fictitious particles, which neither scatter nor absorb light rays. Their distribution μ_f is chosen so that the sum of μ_t and μ_f add up to a globaly constant value, the so called majorant $\overline{\mu}$. The orignal problem of finding a real collision is then reformulated to finding any collision (fictitious or real). Since $\overline{\mu}$ is a constant, equation **??** can easily inverted and we get

$$t_i = -\frac{ln(1-\xi)}{\overline{\mu}} \tag{22}$$

as the sampled distance to the next collision, which might be with a real of a fictitious particle (a so called null collision). At position $p_i - \omega_i t_i$, the probability of hitting a real particle is $\frac{\mu_t(p_i - \omega_i t_i)}{\overline{\mu}}$. Thus, we can decide if a real particle has been hit by sampling another ξ and comparing it $\frac{\mu_t(p_i - \omega_i t_i)}{\overline{\mu}}$. If ξ is smaller, a real particle has been hit, if not, there is a null collision. In this case, we repeat the process from the point $p_i - \omega_i t_i$ until a real collision is detected. The distance to this real coollision is then the sum of all distances previously computed. The position of the real collision is then assumed to be the path vertex p_{i+1}.

Delta tracking can also be used to estimate the transmmittance τ between two points p' and $p + \omega t$. This is done by using delta tracking to find a real collision on the ray from p to p'. If this collision occurs before p', τ is assumed to be 0, if the collision occurs after p', τ is estimated as 1. However, there exists other, more sophisticated methods for etimating the transmittance, such as ratio tracking.

4.3. Ratio Tracking

Ratio tracking is used to estimate the transmittance between two already known points p and p'. This is done by calculating various collision points between p and p', as described above, and stops once a collision point behind p' is reached. At each collision point p^i, the value $\frac{\mu_f}{\overline{\mu}} = 1 - \frac{\mu_t}{\overline{mu}}$ is saved. This fraction between fictious particles and all particles describes the probability, that a light ray can pass p^i without real collision. The transmittance τ can then be estimated as

$$\tau p, p' = \prod_{1 \leq i \leq} (1 -$$