VolSDF笔记

> Volume Rendering of Neural Implicit Surfaces(Nips2021)

Abstract

神经渲染目前取得了比较大的成功，然而，由于表面是用是随机水平集重建的，所以会导致噪声多、低保真的重建。与过往以密度代表几何表面不同，本文使用集合表面来代表密度。具体来说是将密度函数的CDF转换成SDF，这种方法有三个好处，提供了偏纳归置，提供了提渲染公式的误差上下界，可以无监督高效解出表面。

NeuS 是通过一个具有可学习的标准差 s 的转换函数来实现 SDF-guided Volume Rending。与之相对的，VolSDF 则是通过控制采样的过程来实现 SDF-guided Volume Rending。

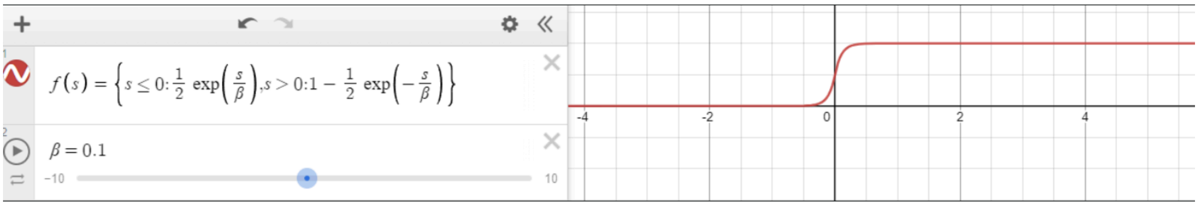
SDF guided Volume Rendering

VolSDF 的思路和 NeuS 一致，即将 SDF 值 (f(x)) 过一个转换函数得到σ然后参与积分。这里 VolSDF 采用了拉普拉斯分布的 CDF (Cumulative Distribution Function of the Laplace distribution)。给定 SDF $d_\Omega(x)$，有（其中α, β均为可学习参数）：

$$\sigma(x) = \alpha \Psi_\beta(-d_\Omega(x))$$

$$\Psi_\beta(s) = \begin{cases} \frac{1}{2}\exp\left(\frac{s}{\beta}\right) & \text{if } s \leq 0 \\ 1 - \frac{1}{2}\exp\left(-\frac{s}{\beta}\right) & \text{if } s > 0 \end{cases}$$

$\Psi_\beta(s)$的可视化如下：



$$f(s) = \begin{cases} \frac{1}{2}\exp\left(\frac{s}{\beta}\right), & s \leq 0 \\ 1 - \frac{1}{2}\exp\left(-\frac{s}{\beta}\right), & s > 0 \end{cases}$$

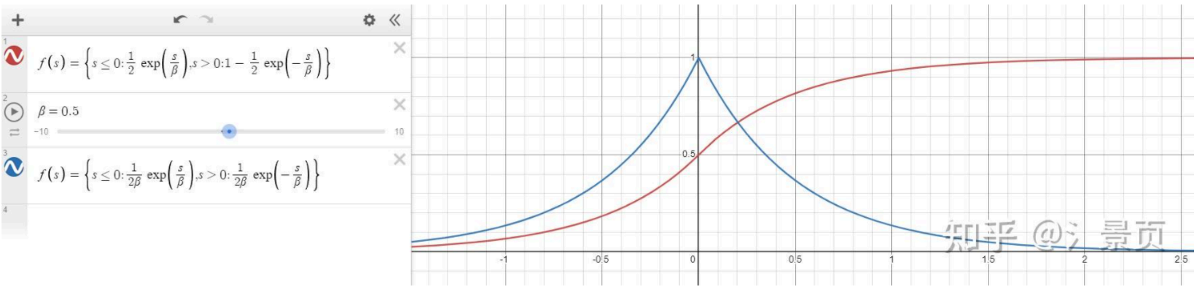
当β越接近 0 时，则越接近阶跃函数。请注意，VolSDF 的 SDF$-d_\Omega(x)$输入到$\Psi_\beta(s)$有个符号，意味着当在表面外时 $(d_\Omega(x) > 0)$，其结果很小；而当进入表面后$d\Omega(x) < 0$，其结果接近 1。这其实和 NeuS 提出的$\rho(t)$性质是一样的。

那么可以看到，SDF 与σ的幅值关系（这意味着每个 SDF 可能由适合自己的σ场）由α控制（论文中说的是同质固体 - homogeneous object），而接近表面时的平滑程度则由β控制。

以上，作者认为这样的设计引入了一个有利于学习形状的归纳偏置（inductive bias），同时还有助于限制 **opacity** 的误差边界。这句话可以说是 VolSDF 最重要的观点。

这里额外给出拉普拉斯分布的概率密度函数 PDF 的可视化（可视化令$\beta = 0$):

$\Phi(s) = \frac{1}{2\beta}\exp\left(-\frac{|x-\mu|}{\beta}\right)$



首先 VolSDF 的 opacity 与 NeuS 中提到的 opacity 概念是完全不同的（NeuS 的其实并不严格符合物理理，而 VolSDF 的还是符合物理的）。VolSDF 的 opacity 就是非常直白的：

$$O(t) = 1 - T(t)$$

而 (T(t)) 的概念也是最原始的体渲染中的 transparency。VolSDF 不关注 NeuS 中所提到的无偏性，相反 VolSDF 更关注 **实现体渲染的离散积分结果与连续积分结果的偏差有多大**。VolSDF 认为体渲染没有问题，问题出在 NeRF 采用蒙特卡洛积分的方式实现体渲染已经引入了 **积分近似误差**，损害了场的学习，而这一误差再引入了 SDF-guided (\sigma) 的 inductive bias 会更进一步损害重建的结果。这个误差是贯穿全文的。

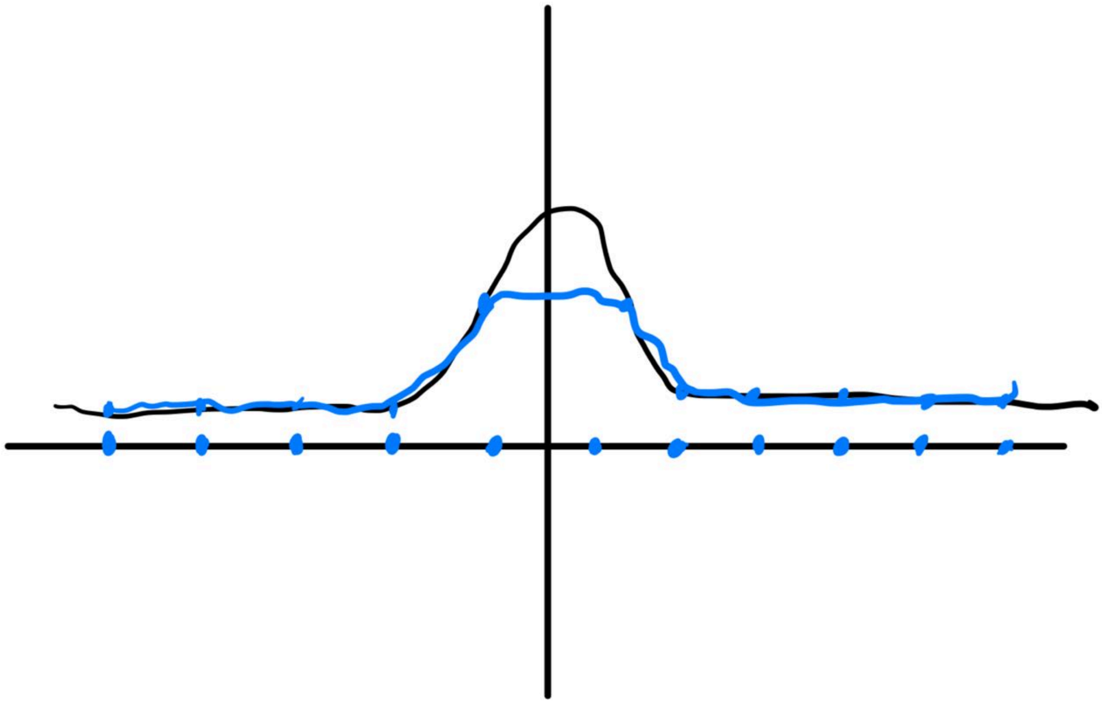
而 VolSDF 要做的就是 **通过设计相应的采样过程来减少这一误差**。反过来说，作者根据误差的推导，**通过使误差变小的方式来指导采样过程。**

> 作者还说 NeRF 采用的简单的基于 CDF 和直方图的 Coarse-to-Fine Sampling 存在问题：Using a naive or crude approximation of (O) would lead to a sub-optimal sample set (S) that misses, or over extends non-negligible (\tau) values.

以上就是 VolSDF 整个文章的中心点。总的来说，VolSDF 的目的也只是为了减少蒙特卡洛积分由于采样和离散化所带来的积分近似误差。相比于 NeuS 机械降神且“离经叛道”地从重建的角度提出两个原则并从一个归一化权重的公式开始推导；VolSDF 以更加忠诚的姿态从减少积分近似误差边界的角度开始推导，并最终专注于采样策略的优化。

Sampling

如果按我们之前的均匀采样（coarse），看下图，如果黑色是真实的分布，蓝色是我们的采样，最后得到的结果会和真实值差别很大



并且如果在fine采样之后，其实我们的密度分布也和第一次采样的很不同。在所以我们要解决这个问题。

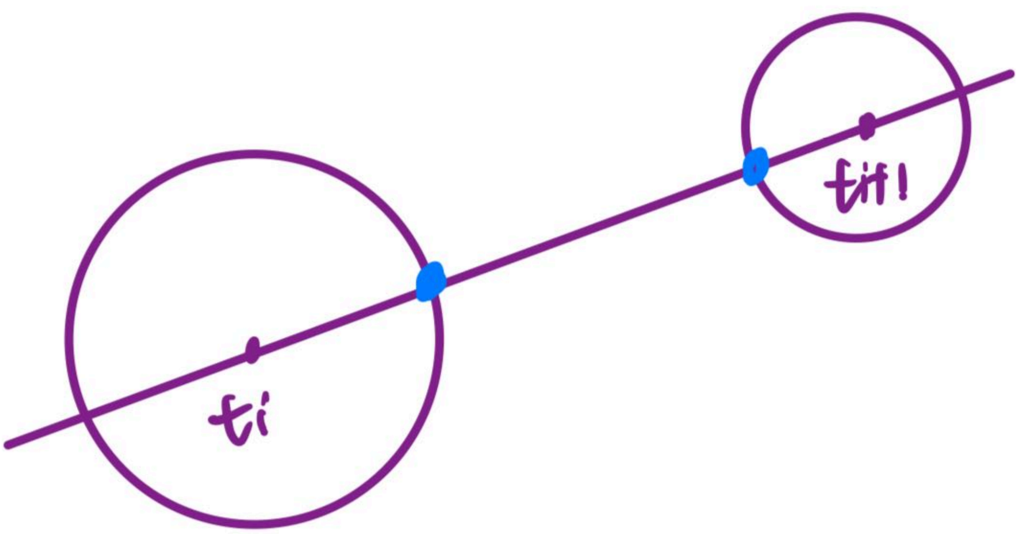
接着我们先记下这两个不等式。

- 第一个不等式：$|1-e^r| \leq e^{|r|}-1$
- 第二个不等式：$|f(x)-f(y)| \leq |f'_{\max}(x)||x-y|$

然后我们求区间$ti-ti+1$内SDF距离最小的点。

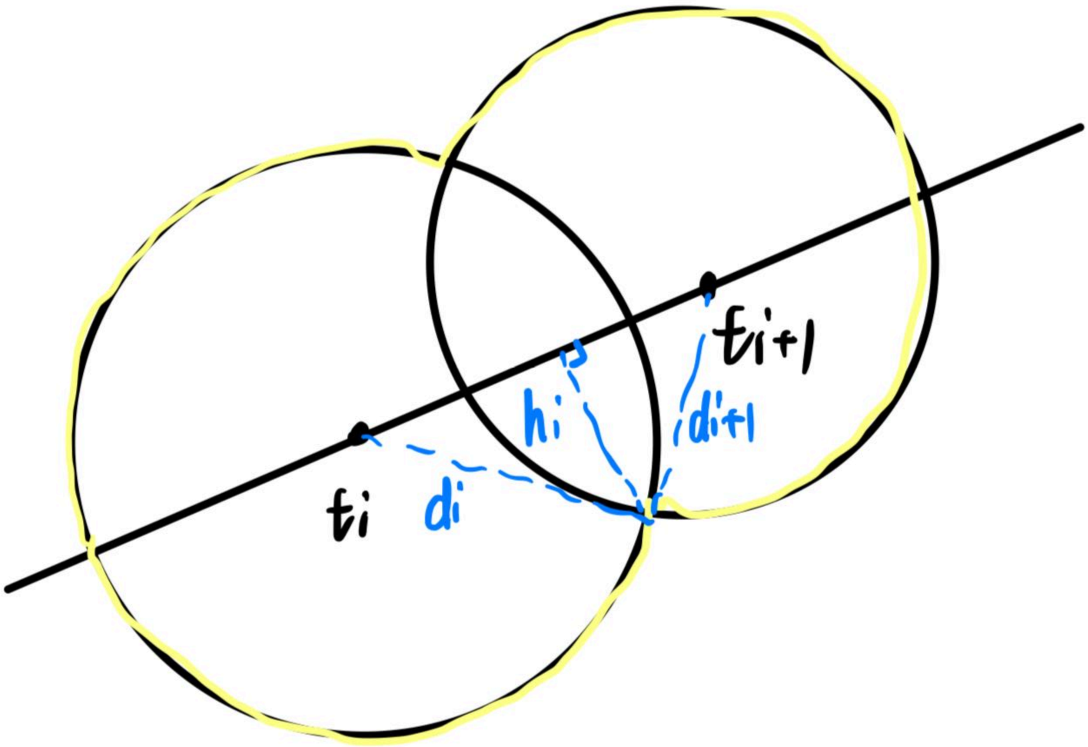
这里分两种情况

- 第一种



由于此时两个球壳内没有东西，最短距离点即为两个蓝点，最短距离为0

- 第二种



此时两个球壳相交，最短距离即为hi

我们可以用海伦公式计算，给定三角形的三条边a,b,c，其面积计算：

$$s=\frac{a+b+c}{2}$$

$$A=\sqrt{s(s-a)(s-b)(s-c)}$$

最后得到

$$d_i^{\star}=\begin{cases}0 & |d_i|+|d_{i+1}|\le\delta_i\\ \min\{|d_i|,|d_{i+1}|\} & \left||d_i|^2-|d_{i+1}|^2\right|\ge\delta_i^2\\ h_i & \textit{otherwise}\end{cases}, \tag{21}$$

$d_i=d_\Omega(\boldsymbol{x}(t_i)),\ h_i=\frac{2}{\delta_i}\sqrt{s(s-\delta_i)(s-|d_i|)(s-|d_{i+1}|)},\ s=\frac{1}{2}(\delta_i+|d_i|+|d_{i+1}|),\ \delta_i=t_{i+1}-t_i.$

继续推导，原来NeRF公式长这样。

$$\hat{C}(t)=\sum_{i=1}^{M}T_i\sigma_i c_i$$

一个采样点是好的采样点的话，就需要满足：

$$||C(t)-\hat{C}(t)||\le\epsilon$$

但是一段时间颜色和密度和透明度一般不会是常量，因此会有误差估计，主要是透明度，因为透明度是积分的，误差也会累计。因此造成估计误差的主要因素。

一个采样点是好的采样点的话，就主要需要满足：

$$||T(t) - \hat{T}(t)|| \le \epsilon$$

文章使用了不透明度的概念，$O(t) = 1 - T(t)$，表示某一点前面被物体挡住的概率。

$$||O(t) - \hat{O}(t)|| \le \epsilon$$

其中$\hat{O}(t)$是我们近似得到的值

$$O(t) = 1 - e^{-\int_0^t \sigma(s)ds}$$

进行离散化

$$\hat{O}(t) = 1 - e^{-\sum_{i=1}^{k-1} \sigma_i \delta_i + (t - t_k)\sigma_k}$$

令：

$$\hat{R}(t) = \sum_{i=1}^{k-1} \sigma_i \delta_i + (t - t_k)\sigma_k$$

$$E(t) = \int_0^t \sigma(s)ds - \hat{R}(t)$$

则原公式化为：

$$O(t) = 1 - e^{-(\hat{R}(t) + E(t))}$$

$$\hat{O}(t) = 1 - e^{-\hat{R}(t)}$$

$$|O(t) - \hat{O}(t)| = |e^{-\hat{R}(t)}(1 - e^{-E(t)})|$$

$$= e^{-\hat{R}(t)}|1 - e^{-E(t)}|$$

$$\le e^{-\hat{R}(t)}(e^{|E(t)|} - 1)$$

$$\le e^{-\hat{R}(t)} e^{\hat{E}(t)} - 1$$

设$t \in (t_k, t_{k+1})$，随着射线点的增加，$\hat{R}(t)$ 是增加的，$\hat{E}(t)$ 是减少的，因此$e^{-\hat{R}(t)}$是减少的，$e^{\hat{E}(t) - 1}$是增加的，因此：

$$|O(t) - \hat{O}(t)| \le e^{-\hat{R}(t_k)} e^{\hat{E}(t_{k+1})} - 1$$

$$\hat{E}(t_{k+1}) = \left| \int_0^{t_{k+1}} \sigma(s)ds - \hat{R}(t_{k+1}) \right|$$

$$= \left| \int_0^{t_{k+1}} \sigma(s)ds - \sum_{i=1}^{k} \sigma_i \delta_i - (t_{k+1} - t_k)\sigma_k \right|$$

$$\le \left| \sum_{i=1}^{k} \int_{t_i}^{t_{i+1}} \sigma(s)ds - \sigma_i \delta_i \right|$$

$$\le \sum_{i=1}^{k} \left| \int_{t_i}^{t_{i+1}} \sigma(s)ds - \sigma_i \delta_i \right|$$

由于：

$$\left|\int_{t_i}^{t_{i+1}}\sigma(s)ds-\sigma_i\delta_i\right|=\left|\int_{t_i}^{t_{i+1}}(\sigma(s)-\sigma(t_i))ds\right|\leq\int_{t_i}^{t_{i+1}}|\sigma(s)-\sigma(t_i)|ds\leq\int_{t_i}^{t_{i+1}}\sigma'(s)|s-t_i|ds$$

由于：

$$\sigma'(x)=\alpha\frac{1}{2\beta}e^{-\frac{|x|}{\beta}}$$

则最小的梯度为：

$$\sigma'=\alpha\frac{1}{2\beta}e^{-\frac{|d^*|}{\beta}}$$

其中，d^* 为(t_i,t_{i+1})之中的粒子离 SDF 最短的距离。
怎么来求d^*呢，可以通过t_i和t_{i+1}的 SDF 来求。

所以：

$$\left|\int_{t_i}^{t_{i+1}}\sigma(s)ds-\sigma_i\delta_i\right|\leq\int_{t_i}^{t_{i+1}}\sigma'|s-t_i|ds\leq\int_{t_i}^{t_{i+1}}\sigma'ds\leq\alpha\frac{1}{4\beta}e^{-\frac{|d_i^*|}{\beta}}\delta^2$$

代回式子：

$$\hat{E}(t_{k+1})\leq\sum_{i=1}^{k}\left|\int_{t_i}^{t_{i+1}}\sigma(s)ds-\sigma_i\delta_i\right|$$

$$\leq\alpha\frac{1}{4\beta}\sum_{i=1}^{k}e^{-\frac{|d_i^*|}{\beta}}\delta_i^2$$

因此：

$$|O(t)-\hat{O}(t)|\leq e^{-\hat{R}(t_k)}e^{\hat{E}(t_{k+1})-1}$$

$$\leq e^{-\hat{R}(t_k)}e^{\frac{\alpha}{4\beta}\sum_{i=1}^{k}e^{\frac{|d_i^*|}{\beta}}\delta_i^2-1}$$

$$\leq e^{\frac{\alpha}{4\beta}\sum_{i=1}^{k}e^{\frac{|d_i^*|}{\beta}}\delta_i^2-1}$$

$$\leq e^{\frac{\alpha}{4\beta}\sum_{i=1}^{k}\delta_i^2-1}$$

因此，不透明度预测的正确性依赖α, β, 以及δ_i。令δ_m为最大分段长度，一共为M段，则：

$$|O(t)-\hat{O}(t)|\leq e^{\frac{\alpha}{4\beta}\sum_{i=1}^{k}\delta_i^2-1}$$

要使误差控制在ϵ之内，一种是减少δ_m，一种是增加β。

假设：

$$\sum_{i=1}^{k}\delta_i^2=\frac{M^2}{n-1}$$

（这个是估计这些分段长度的平方和。）

要使：

$$|O(t) - \hat{O}(t)| \leq \epsilon$$

只需要其上界小于ϵ。

$$e^{\frac{\alpha}{4\beta}\sum_{i=1}^{k}\delta_i^2} - 1 \leq \epsilon$$

$$\frac{\alpha}{4\beta}\sum_{i=1}^{k}\delta_i^2 \leq \log(1+\epsilon)$$

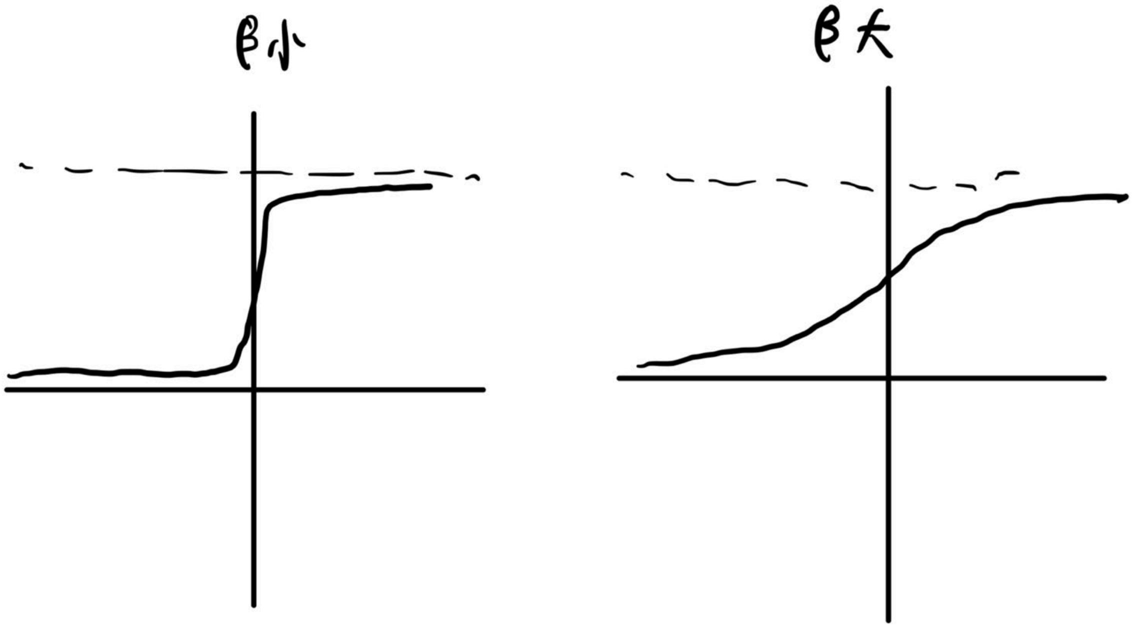
$$\frac{\alpha}{4\beta}\frac{M^2}{n-1} \leq \log(1+\epsilon)$$

$$\beta \geq \frac{\alpha M^2}{4(n-1)\log(1+\epsilon)}$$

所以现在我们减小误差的方式就有两种。一种是增加采样点，一种是增大β。

采样点数量的话由于增大会导致运算量水涨船高，所以不太现实。

但是的话，我们也不是希望β越大越好



我们比较正常正常的情况是，在物体外面时CDF应该为0，进入物体后CDF急剧上升为1；所以β太大，曲线比较平滑是不合理的。我们要找个平衡点。

有了上文铺垫后，看一下本文的采样方法。

Algorithm 1: Sampling algorithm.

Input: error threshold $\epsilon > 0$; β

Initialize $\mathcal{T} = \mathcal{T}_0$

Initialize β_+ such that $B_{\mathcal{T},\beta_+} \leq \epsilon$

while $B_{\mathcal{T},\beta} > \epsilon$ *and not max_iter* **do**

  upsample $\mathcal{T}$

  if $B_{\mathcal{T},\beta_+} < \epsilon$ **then**

    Find $\beta_\star \in (\beta, \beta_+)$ so that $B_{\mathcal{T},\beta_\star} = \epsilon$

    Update $\beta_+ \leftarrow \beta_\star$

  end

end

Estimate $\widehat{O}$ using $\mathcal{T}$ and β_+

$\mathcal{S} \leftarrow$ get fresh m samples using $\hat{O}^{-1}$

return $\mathcal{S}$

相当于我们先设置一个大点的β，保证不超过误差，然后我们慢慢增加采样点，减小β，看看能不能找到平衡点。最后蒙特卡洛采样，采用逆采样。

再提一嘴，相比于 NeRF 基于 CDF 和直方图的上采样，VolSDF 通过二分查找的方式找到新的β得到更小的 error bound。下图展示一下给定一个光线上 $SDF s(t)$（仅有一个交点），对应的$\sigma(t), T(t), \tau(t)$：

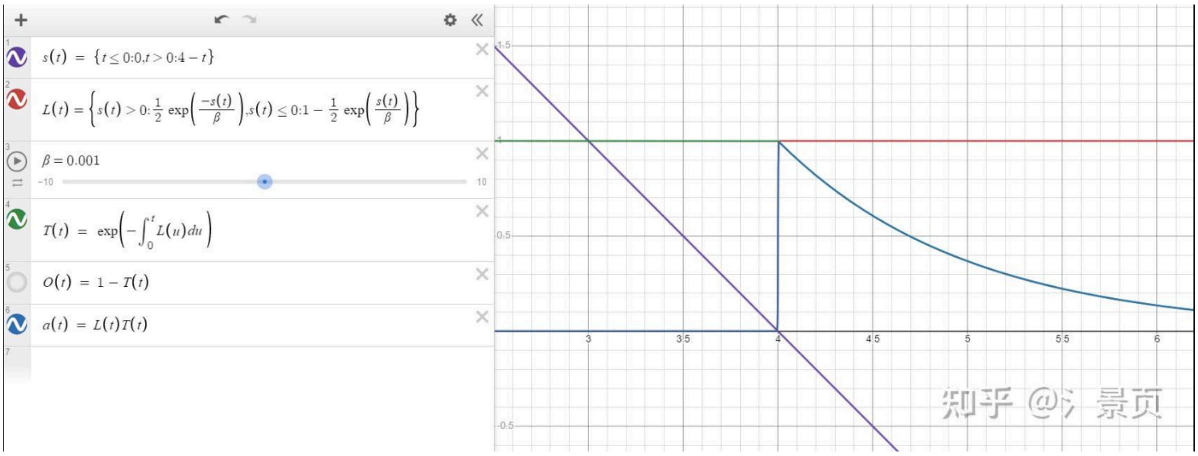
$$s(t) = \{t \leq 0 : 0.4 - t\}$$

$$L(t) = \begin{cases} \frac{1}{2}\exp\left(\frac{-s(t)}{\beta}\right) & t \leq 0 \\ 1 - \frac{1}{2}\exp\left(\frac{-s(t)}{\beta}\right) & t > 0 \end{cases}$$

$$T(t) = \exp\left(-\int_0^t L(u)\,du\right)$$

$$O(t) = 1 - T(t)$$

$$a(t) = L(t)\tau(t)$$



s(t) = {t ≤ 0:0, t > 0:4 − t}
L(t) = {s(t) > 0: 1/2 exp(−s(t)/β), s(t) ≤ 0: 1 − 1/2 exp(s(t)/β)}
β = 0.001
T(t) = exp(−∫_0^t L(u)du)
O(t) = 1 − T(t)
a(t) = L(t)T(t)
知乎 @沉景页