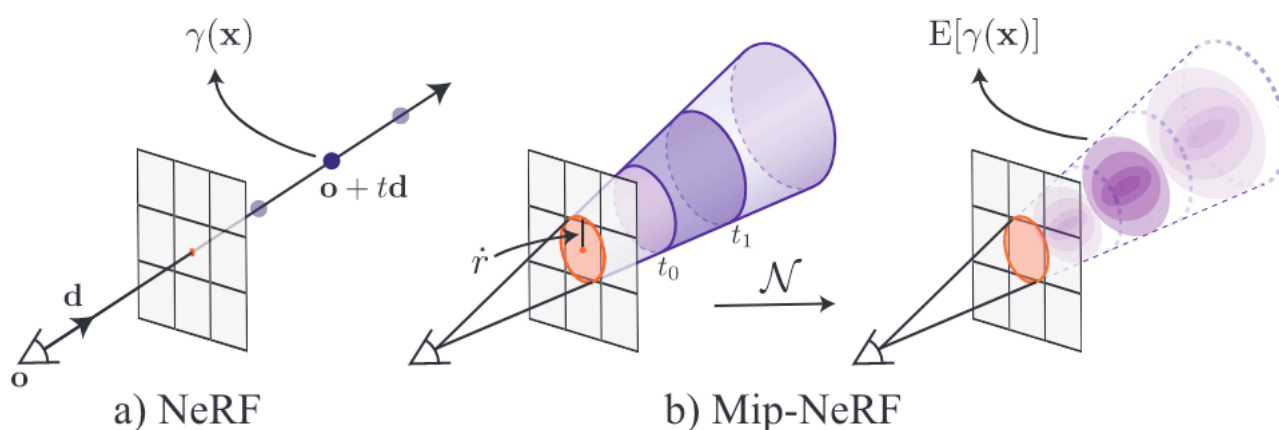


Mip-NeRF解读

Mip-NeRF: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields (ICCV2021)

Abstract:

传统的NeRF对一个像素点的渲染只采用单条光线，所以在低分辨率下训练，高分辨率下测试时，离子数量就会不够，产生模糊；使用高分辨率训练，低分辨率下测试时，会产生锯齿。



Mip-NeRF采用了图形学中Mip-map方法（详情可见Games101）。

如果使用图形学中的Mip-map的方法，需要先预训练不同分辨率下的体素分布，这种可行性不高(可能训练数据只有一个分辨率下的)，而且时长久。

如果使用图形学的超采样方法，虽然工作，但是对于每个像素点需要采样多根光线，训练时间直接翻倍。作者使用了的方法如下：

每个像素点的颜色，由圆锥内的所有粒子所决定，这样像素点的颜色不再依赖于一根射线上的粒子，可以较好的解决锯齿和模糊的问题。

在nerf原始的论文里面，神经网络中输入 (x,y,z) 以及观察方向，输出相应的位置的粒子的颜色与密度，我们在Mip- NeRF如果也使用这种方法，意味着我们要对一个锥体内的粒子在三维空间进行采样，使用采样的粒子来计算像素点的颜色与密度，这样会使计算量爆炸！

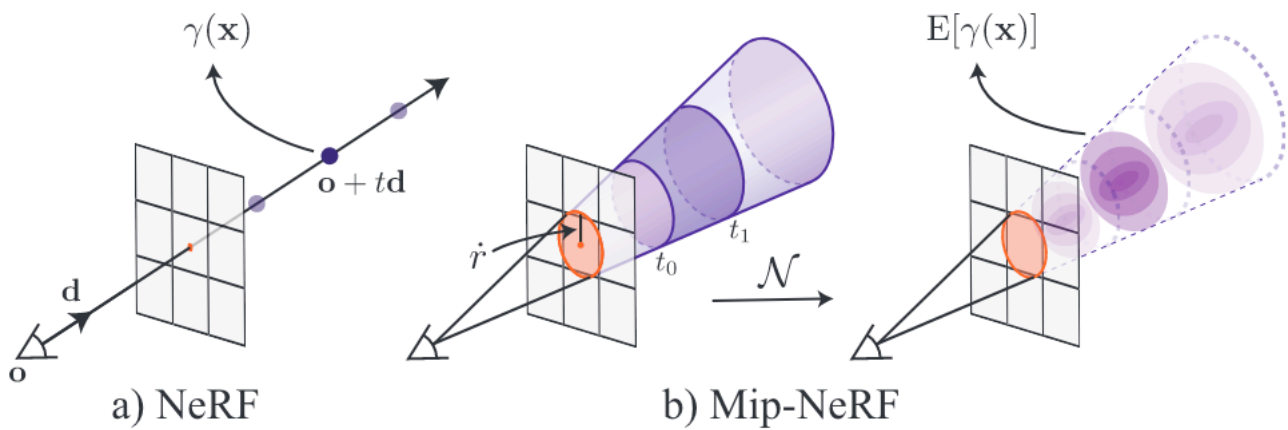
所以论文里将锥体分为一小段一小段，使用三维高斯分布来进行这些锥体段的空间分布，最后在神经网络中输入位置(均值与协方差)与观察角度，可以得到这一坨粒子的颜色期望与密度期望。

论文里涉及的两个最主要问题如下：

- 1.如何使用高斯分布对这些截锥体分布进行模拟？
- 2.在nerf论文里我们使用了位置编码，把 (x,y,z) 以及view dir分布进行了高频编码，在mip-nerf里输入变成了分布的均值与方差以及观察方向，如何对这些数据进行相应的位置编码？

Method

空间模拟分布



首先我们需要将一个均匀分布的截头锥形模拟为一个高维的高斯分布。

为方便计算，首先进行坐标转换，将原来的xyz坐标转换为t,r, θ 坐标系，t表示沿射线方向上两个截面的距离，r表示粒子离中心射线的距离， θ 表示粒子沿中心射线旋转的角度。

$$\begin{aligned}x &= r t \cos \theta \\y &= r t \sin \theta \\z &= t\end{aligned}$$

由坐标变换的公式可知：

$$dxdydz = \begin{vmatrix} \frac{dx}{dr} & \frac{dy}{dr} & \frac{dz}{dr} \\ \frac{dx}{dt} & \frac{dy}{dt} & \frac{dz}{dt} \\ \frac{dx}{d\theta} & \frac{dy}{d\theta} & \frac{dz}{d\theta} \end{vmatrix} dr d\theta dt = \begin{vmatrix} t \cos \theta & t \sin \theta & 0 \\ r \cos \theta & r \sin \theta & 1 \\ -rt \sin \theta & rt \cos \theta & 0 \end{vmatrix} dr d\theta dt = r t^2 dr d\theta dt$$

其截锥体的体积为：

$$\begin{aligned}V &= \iiint dxdydz = \int_{t_0}^{t_1} \int_0^{2\pi} \int_0^{r^*} r t^2 dr d\theta dt \\&= \frac{1}{3} \pi r^2 (t_1^3 - t_0^3)\end{aligned}$$

均匀分布的截锥体的密度为 $\frac{1}{V}$.

这一个均匀分布在t方向上的一阶矩为：

$$\begin{aligned}E(t) &= \frac{1}{V} \iiint t dxdydz \\&= \frac{1}{V} \int_{t_0}^{t_1} \int_0^{2\pi} \int_0^{r^*} t r t^2 dr d\theta dt \\&= \frac{3(t_1^4 - t_0^4)}{4(t_1^3 - t_0^3)}\end{aligned}$$

在t方向上的二阶矩是：

$$\begin{aligned}E(t^2) &= \frac{1}{V} \iiint t^2 dxdydz \\&= \frac{1}{V} \int_{t_0}^{t_1} \int_0^{2\pi} \int_0^{r^*} t^2 r t^2 dr d\theta dt\end{aligned}$$

$$= \frac{3(t_1^5 - t_0^5)}{5(t_1^3 - t_0^3)}$$

在x方向上的一阶矩（均值）一定为0，在x方向上的二阶矩为：

$$\begin{aligned} E(x^2) &= \frac{1}{V} \iiint (rt \cos \theta)^2 dx dy dz \\ &= \frac{1}{V} \iiint r^2 t^2 \cos^2 \theta r t^2 dr d\theta dt \\ &= \frac{3r^2(t_1^5 - t_0^5)}{20(t_1^3 - t_0^3)} \end{aligned}$$

因此，在t(z)方向上的均值为：

$$u_t = \frac{3(t_1^4 - t_0^4)}{4(t_1^3 - t_0^3)}$$

在t(z)方向上的方差为：

$$\sigma_t^2 = \frac{3(t_1^5 - t_0^5)}{5(t_1^3 - t_0^3)} - u_t^2$$

在x方向上和y方向上的均值为：

$$u_x = u_y = 0$$

在x方向上和y方向上的均值为：

$$\sigma_x^2 = \sigma_y^2 = \frac{3r^2(t_1^5 - t_0^5)}{20(t_1^3 - t_0^3)}$$

当 t_1 和 t_0 离得比较近的情况下，以上式子中分子和分母都会很小，进行除法时，稳定性不好，因此引入

$$t_\mu = (t_0 + t_1)/2, t_\delta = (t_1 - t_0)/2.$$

将其替换到原式子中，增加运算时的稳定性，替换结果如下：

$$\begin{aligned} \mu_t &= t_\mu + \frac{2t_\mu t_\delta^2}{3t_\mu^2 + t_\delta^2}, \quad \sigma_t^2 = \frac{t_\delta^2}{3} - \frac{4t_\delta^4(12t_\mu^2 - t_\delta^2)}{15(3t_\mu^2 + t_\delta^2)^2}, \\ \sigma_r^2 &= r^2 \left(\frac{t_\mu^2}{4} + \frac{5t_\delta^2}{12} - \frac{4t_\delta^4}{15(3t_\mu^2 + t_\delta^2)} \right). \end{aligned} \quad (41)$$

其均值与协方差为：

$$\begin{aligned} u &= \begin{bmatrix} 0 \\ 0 \\ u_t \end{bmatrix} \\ \Sigma &= \begin{bmatrix} \sigma_x & 0 & 0 \\ 0 & \sigma_y & 0 \\ 0 & 0 & \sigma_t \end{bmatrix} \end{aligned}$$

但是这个是在相机坐标下的，我们需要把它转换到世界坐标系下，设摄像机现在摆放的坐标为 \mathbf{O} ，则在世界坐标系下原点的位置为 $\mathbf{O} + u_t \mathbf{d}$ ， \mathbf{d} 为观察方向（单位向量）。

对xyz坐标先进行归一化处理，归一化后协方差矩阵不发生变化。

当观察方向为 \mathbf{d} 时，则原坐标(归一化后)与新坐标(归一化后)的对应关系为：

$$\begin{aligned} x &\rightarrow ? \\ y &\rightarrow ? \end{aligned}$$

$$z - > d$$

假定新坐标的另外两个基为 \mathbf{u} 与 \mathbf{v} ,则旧坐标与新坐标之间的变换矩阵为：

$$P = [\mathbf{u} \quad \mathbf{v} \quad \mathbf{d}]$$

由公式：

$$\text{cov}(Ax, Ay) = A\text{cov}(x, y)A^T$$

新的协方差矩阵为：

$$\begin{aligned} \Sigma' &= \begin{bmatrix} \text{cov}(u, u) & \text{cov}(u, v) & \text{cov}(u, w) \\ \text{cov}(v, u) & \text{cov}(v, v) & \text{cov}(v, w) \\ \text{cov}(w, u) & \text{cov}(w, v) & \text{cov}(w, w) \end{bmatrix} \\ &= \begin{bmatrix} \text{cov}(Px, Px) & \text{cov}(Px, Py) & \text{cov}(Px, Pz) \\ \text{cov}(Py, Px) & \text{cov}(Py, Py) & \text{cov}(Py, Pz) \\ \text{cov}(Pz, Px) & \text{cov}(Pz, Py) & \text{cov}(Pz, Pz) \end{bmatrix} \\ &= P \begin{bmatrix} \text{cov}(x, x) & \text{cov}(x, y) & \text{cov}(x, z) \\ \text{cov}(y, x) & \text{cov}(y, y) & \text{cov}(y, z) \\ \text{cov}(z, x) & \text{cov}(z, y) & \text{cov}(z, z) \end{bmatrix} P^T \\ &= P \Sigma P^T \\ &= [\mathbf{u} \quad \mathbf{v} \quad \mathbf{d}] \begin{bmatrix} \sigma_x & 0 & 0 \\ 0 & \sigma_y & 0 \\ 0 & 0 & \sigma_z \end{bmatrix} \begin{bmatrix} \mathbf{u}^T \\ \mathbf{v}^T \\ \mathbf{d}^T \end{bmatrix} \\ &= \sigma_x \mathbf{u} \mathbf{u}^T + \sigma_y \mathbf{v} \mathbf{v}^T + \sigma_z \mathbf{d} \mathbf{d}^T \end{aligned}$$

令：

$$\sigma_x = \sigma_y = \sigma_r$$

$$\Sigma' = \sigma_r (\mathbf{u} \mathbf{u}^T + \mathbf{v} \mathbf{v}^T) + \sigma_z \mathbf{d} \mathbf{d}^T$$

由于

$$\begin{aligned} &= [\mathbf{u} \quad \mathbf{v} \quad \mathbf{d}] \begin{bmatrix} \mathbf{u}^T \\ \mathbf{v}^T \\ \mathbf{d}^T \end{bmatrix} = \mathbf{I} \\ &\mathbf{u} \mathbf{u}^T + \mathbf{v} \mathbf{v}^T = \mathbf{I} - \mathbf{d} \mathbf{d}^T \end{aligned}$$

原式为：

$$\Sigma' = \sigma_r (\mathbf{I} - \mathbf{d} \mathbf{d}^T) + \sigma_z \mathbf{d} \mathbf{d}^T$$

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 2 & 0 & 0 & \cdots & 2^{L-1} & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 & \cdots & 0 & 2^{L-1} & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 & \cdots & 0 & 0 & 2^{L-1} \end{bmatrix}^T, \gamma(\mathbf{x}) = \begin{bmatrix} \sin(\mathbf{P}\mathbf{x}) \\ \cos(\mathbf{P}\mathbf{x}) \end{bmatrix}$$

最后求得 Encoding 期望：

$$\begin{aligned} E_{x \sim \mathcal{N}(\mu, \sigma^2)}[\sin(x)] &= \sin(\mu) \exp(-(1/2)\sigma^2) \\ E_{x \sim \mathcal{N}(\mu, \sigma^2)}[\cos(x)] &= \cos(\mu) \exp(-(1/2)\sigma^2) \end{aligned}$$

因此 IPE 编码可以表示为：

$$\begin{aligned} \gamma(\mu, \Sigma) &= E_{x \sim \mathcal{N}(\mu_\gamma, \Sigma_\gamma)}[\gamma(\mathbf{x})] \\ &= \begin{bmatrix} \sin(\mu_\gamma \circ \exp(-(1/2)\text{diag}(\Sigma_\gamma))) \\ \cos(\mu_\gamma \circ \exp(-(1/2)\text{diag}(\Sigma_\gamma))) \end{bmatrix} \end{aligned}$$

最后还有一个不足，就是协方差矩阵 Σ_γ 的计算成本比较高，同时也只有对角线的值需要考虑，因此在实际实现中可表示为：

$$\begin{aligned} \text{diag}(\Sigma_\gamma) &= [\text{diag}(\Sigma), 4\text{diag}(\Sigma), \dots, 4^{L-1}\text{diag}(\Sigma)]^T \\ \text{diga}(\Sigma) &= \sigma_t^2(\mathbf{d} \circ \mathbf{d}) + \sigma_r^2\left(\mathbf{1} - \frac{\mathbf{d} \circ \mathbf{d}}{\|\mathbf{d}\|_2^2}\right) \end{aligned}$$

Mip-NeRF 准确描述（进行了合理的建模）了像素包含区域随物体远近变化的关系（近大远小）。对像素来说，越远的区域截断视锥越大，即积分区域越大，此时 Encoding 高频部分的均值迅速衰减到 0（等价于 MipMap 的 Prefiltered 的功能），避免了远处样本点的突然出现的高频信号的影响，参考原文 3.1 最后一段的表述。

位置编码

Mip-NeRF 准确描述（进行了合理的建模）了像素包含区域随物体远近变化的关系（近大远小）。对像素来说，越远的区域截断视锥越大，即积分区域越大，此时 Encoding 高频部分的均值迅速衰减到 0（等价于 MipMap 的 Prefiltered 的功能），避免了远处样本点的突然出现的高频信号的影响，参考原文 3.1 最后一段的表述。

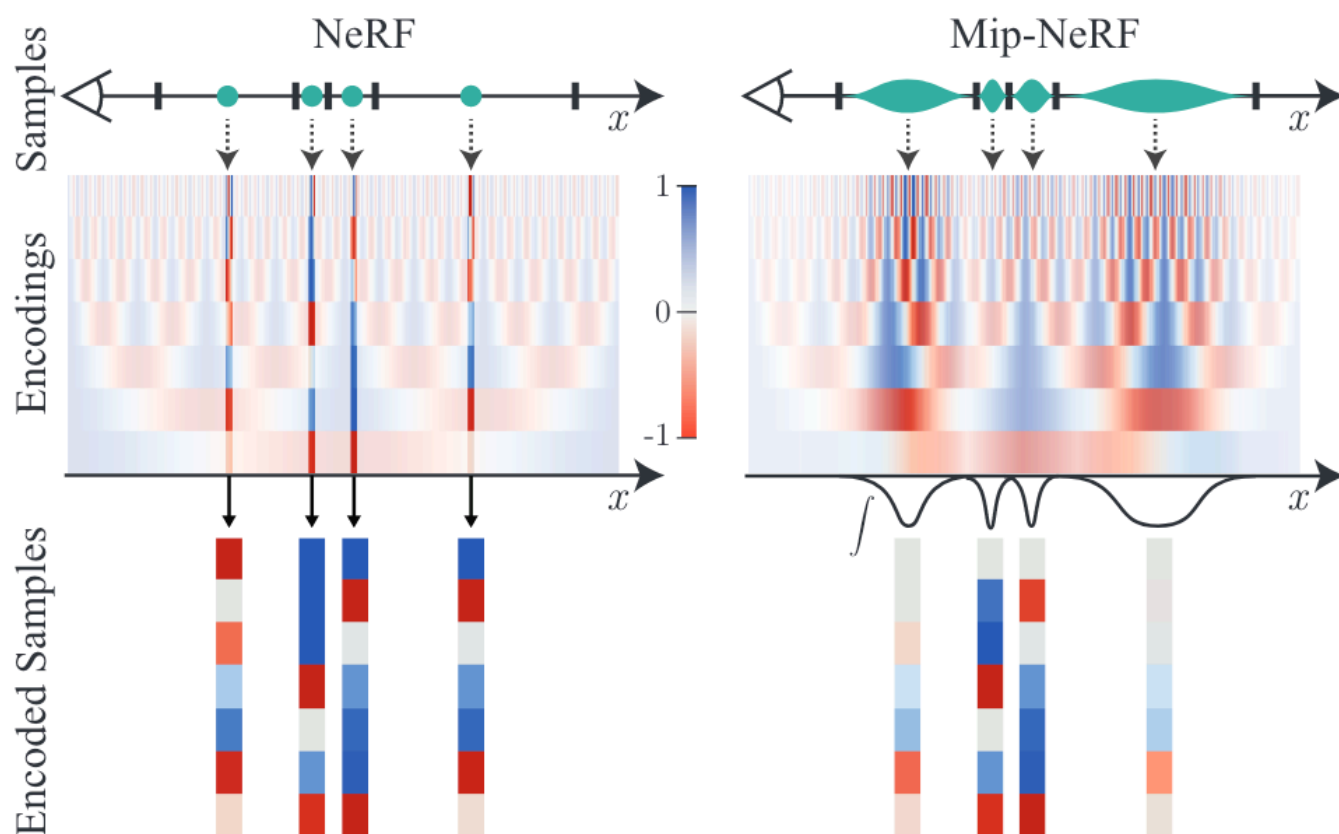
传统nerf中的位置编码

以位置(x,y,z)为例
令

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \\ 2^2 & 0 & 0 \\ 0 & 2^2 & 0 \\ 0 & 0 & 2^2 \\ \dots & & \\ 2^{L-1} & 0 & 0 \\ 0 & 2^{L-1} & 0 \\ 0 & 0 & 2^{L-1} \end{bmatrix}$$

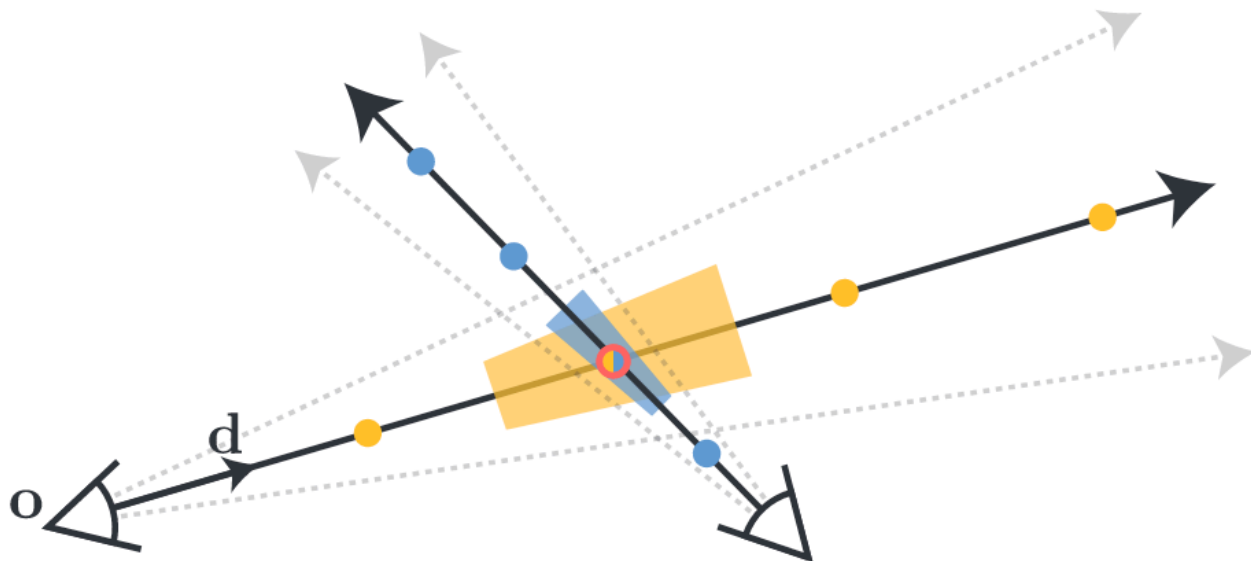
$$\begin{aligned} r(x) &= \begin{bmatrix} \sin(P\mathbf{x}) \\ \cos(P\mathbf{x}) \end{bmatrix} \\ &= \begin{bmatrix} \sin(x) \\ \sin(y) \\ \sin(z) \\ \sin(2x) \\ \sin(2y) \\ \sin(2z) \\ \dots \\ \sin(2^{L-1}x) \\ \sin(2^{L-1}y) \\ \sin(2^{L-1}z) \\ \cos(x) \\ \cos(y) \\ \cos(z) \\ \cos(2x) \\ \cos(2y) \\ \cos(2z) \\ \dots \\ \cos(2^{L-1}x) \\ \cos(2^{L-1}y) \\ \cos(2^{L-1}z) \end{bmatrix} \end{aligned}$$



NeRF 则没有这个概念（可以理解为相比 Mip-NeRF，NeRF 对深度没有那么敏感）。NeRF 的 PE 会导致同一个点在不同光线/同一光线不同远近上也会产生一样的 Encoding 结果表示这一段光线的特征。当这一样本点处于较远位置，但它又具有高频信号时，则不利于 NeRF 的学习（因为越远的点应当提供越少的信息量，但这种采样编码的结果违背了这一原则，可以理解为编码信号的走样）

所以文章 Intro 就指出当数据集中有距离目标物体远近不一以及分辨率不一的图片时，NeRF 就容易产生失真。

得益于 IPE，Mip-NeRF 能够只用一个 MLP 来实现 Coarse-to-Fine，因为此时采样不再会导致信号的失真，MLP “看到”的是一样的频率信号。



mip-nerf中的位置编码

传统的nerf的编码是离散的形式，输入为(x,y,z)的离散点，输出为对这个点的编码，但是mip nerf里我们的输入是一坨高斯分布（均值与协方差），如何对这个协方差进行编码？

我们从一维高斯分布来考虑这个问题，假设

x 服从 $N(u, \sigma)$

则：

$$\begin{aligned}
 E(\sin x) &= \int_{-\infty}^{+\infty} \sin x \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-u)^2}{2\sigma^2}} dx \\
 &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \sin x e^{-\frac{(x-u)^2}{2\sigma^2}} dx \\
 &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \sin(t+u) e^{-\frac{t^2}{2\sigma^2}} dt \\
 &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \sin t \cos u e^{-\frac{t^2}{2\sigma^2}} dt + \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{+\infty} \sin u \cos t e^{-\frac{t^2}{2\sigma^2}} dt
 \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2\pi\sigma}} \cos u \int_{-\infty}^{+\infty} \sin t e^{-\frac{t^2}{2\sigma^2}} dt + \frac{1}{\sqrt{2\pi\sigma}} \sin u \int_{-\infty}^{+\infty} \cos t e^{-\frac{t^2}{2\sigma^2}} dt \\
&= \frac{1}{\sqrt{2\pi\sigma}} \sin u \int_{-\infty}^{+\infty} \cos t e^{-\frac{t^2}{2\sigma^2}} dt \\
&= \frac{1}{\sqrt{2\pi\sigma}} \sin u R \left[\int_{-\infty}^{+\infty} e^{it} e^{-\frac{t^2}{2\sigma^2}} dt \right] \\
&= \frac{1}{\sqrt{2\pi\sigma}} \sin u R \left[\int_{-\infty}^{+\infty} e^{i\sqrt{2}\sigma u} e^{-u^2} \sqrt{2\sigma} du \right] = \\
&= \frac{1}{\sqrt{\pi}} \sin u R \left[\int_{-\infty}^{+\infty} e^{-(u^2 - i\sqrt{2}\sigma u + \frac{1}{2}\sigma^2 i^2)} e^{\frac{1}{2}\sigma^2 i^2} du \right] \\
&= \frac{1}{\sqrt{\pi}} \sin u R \left[\int_{-\infty}^{+\infty} e^{-(u^2 - i\sqrt{2}\sigma u - \frac{1}{2}\sigma^2)} e^{-\frac{1}{2}\sigma^2} du \right] \\
&= \frac{1}{\sqrt{\pi}} \sin u e^{-\frac{1}{2}\sigma^2} R \left[\int_{-\infty}^{+\infty} e^{(u - \sqrt{2}\sigma i)^2} du \right] \\
&= \frac{1}{\sqrt{\pi}} \sin u e^{-\frac{1}{2}\sigma^2} \sqrt{\pi} \\
&= \sin u e^{-\frac{1}{2}\sigma^2}
\end{aligned}$$

同样计算：

$$E(\cos x) = \cos u e^{-\frac{1}{2}\sigma^2}$$

把这个公式推广到高维的高斯分布上:

$$\begin{aligned}
E(\sin x) &= \sin u e^{-\frac{1}{2}|\Sigma|} \\
E(\cos x) &= \cos u e^{-\frac{1}{2}|\Sigma|}
\end{aligned}$$

如何得到 $E(\sin 2^i \mathbf{x})$,只需要设

$$\mathbf{y} = P\mathbf{x}$$

其中

$$P = \begin{bmatrix} 2^i & 0 & 0 \\ 0 & 2^i & 0 \\ 0 & 0 & 2^i \end{bmatrix}$$

$$\begin{aligned}
E(\sin 2^i \mathbf{x}) &= E(\sin \mathbf{y}) \\
&= \sin u_y e^{-\frac{1}{2}|\Sigma_y|} \\
&= \sin(Pu) e^{-\frac{1}{2}|P\Sigma_y P^T|}
\end{aligned}$$

#####