Some basic Computer Graphic for NERF:

Homogeneous Coordinates: represent a 4x4 matrix

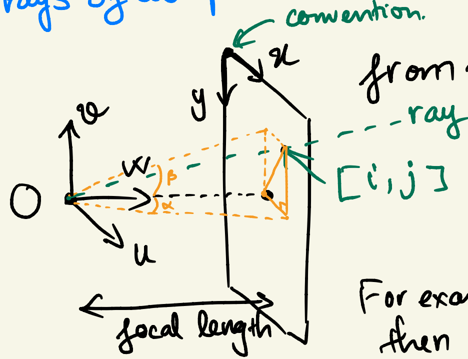
$$\begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

1 indicates point.

0 indicates vector.

where $\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$ is called rotation matrix and $\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}^T$ is the original point where camera is put.

Get rays of all pixel:



from this figure → $O = [d_1 \ d_2 \ d_3]$

$$[u \ v \ w] = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

For example, we have an image w/ width = height = 10 & focal = 1.

then create $i(10,10) = \begin{bmatrix} -5 & \cdots & 4 \\ \vdots & & \\ -5 & & \end{bmatrix}$ & $j(10,10) = \begin{bmatrix} 5 & \cdots & 5 \\ \vdots & & \\ -4 & & \end{bmatrix}$ & $Z(10,10)\begin{bmatrix} f & \cdots & f \\ \vdots & & \\ f & & \end{bmatrix}$

$$\hookrightarrow \left[\left(i/f;\ j/f;\ -f/f\right)\right]_{(10,10,3)} \Rightarrow \begin{bmatrix} (-5,5,-1) & \cdots & (4,5,-1) \\ \vdots & & \vdots \\ (-5,-4,-1) & \cdots & (4,-4,-1) \end{bmatrix}$$

each element represents a $(\tan\alpha, \tan\beta, -1)$ of an pixel

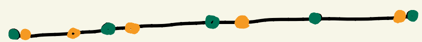
this is pinhole camera feature.

After find $(\tan\alpha, \tan\beta, -1)$ we can find the ray through any pixel by:

$$r_d = u.\tan\alpha + v.\tan\beta + w$$ where this is possible because $\|u\| = \|v\| = \|w\|$.

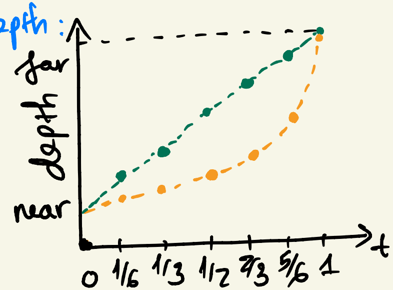
& r_o is the position of O.

Stratified Sampling: along the ray. we take n samples.



- same linearly between near & far.
- the near part is more informative (inverse depth).

Inverse depth:



with n = 7.

Instead of sample depth linearly by:

$$\text{sample} = (\text{far} - \text{near})t + \text{near}.$$

we can sample by inverse depth to sample more sample near the camera & less for far objects.

$$\text{sample} = \frac{1}{\left(\frac{1}{\text{far}} - \frac{1}{\text{near}}\right)t + \frac{1}{\text{near}}}$$

From these sample, we can sampling points on the rays by:

$$\text{points} = \text{ray}_{\text{origin}} + \text{ray}_{\text{direction}} \otimes \text{sample}.$$

like mul matrix but we keep the vector after mul instead of sum all.

For example: $O = [0\ 0\ 0]$ & $d = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}^T_{1\times 3}$ & $\text{sample} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 5 \end{bmatrix}_{4\times 1}$

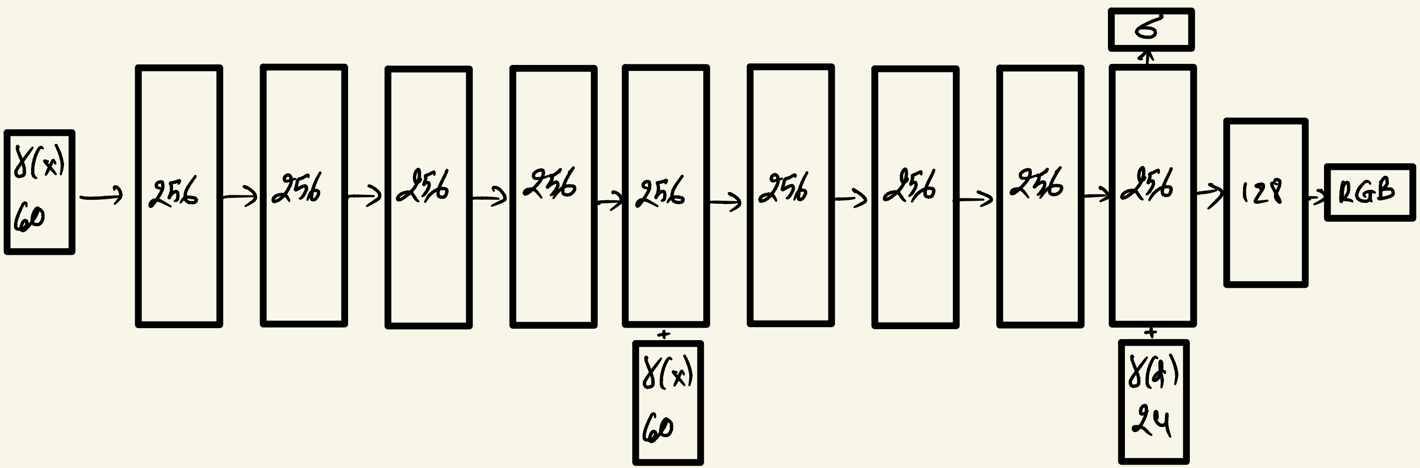
$\hookrightarrow$ $\text{points} = \begin{bmatrix} 2 & 4 & 6 \\ 3 & 6 & 9 \\ 4 & 8 & 12 \\ 5 & 10 & 15 \end{bmatrix}_{4\times 3}$ $\to$ 4 points.

Positional encoder:

- This technique is used tranfroms the input to higher dimension (a unique vector), $f: \mathbb{R} \mapsto \mathbb{R}^{2L}$. In practice, this make the input more easily approximate higher frequency function. (why?)

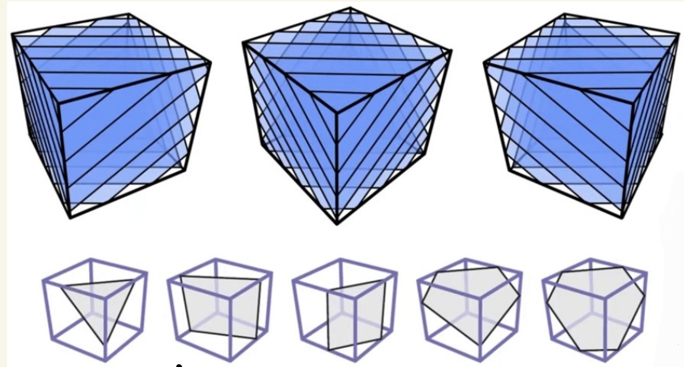
- $\gamma : x \mapsto [\sin(2^0 x), \cos(2^0 x), \ldots, \sin(2^{L-1} x), \cos(2^{L-1} x)]$

NeRF Model:



Volume Rendering

Volume Rendering w/ Slides:



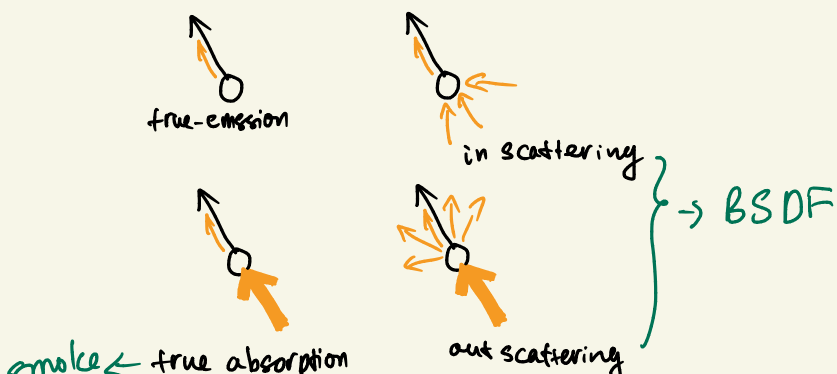
- This is one of the method use for rendering, it generate plane interpreted as slide and then clip the parts that does not belong to the volume.

- One of the way of clipping we can consider is set $\alpha = 0$ to part related to outside of the volume where α value in CGraphic is represent for the opacity or transparency.

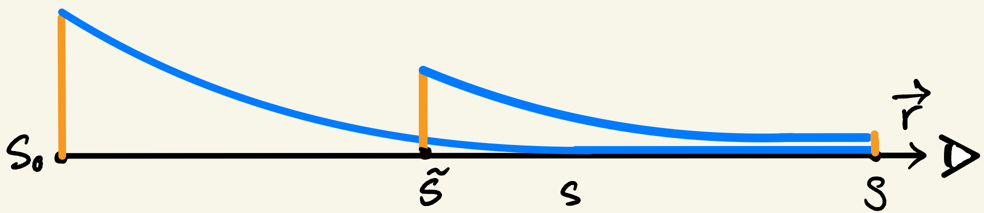
* Note: Transfer Functions -> use to lookup for some part of the data.

Volumetric Scattering:

Consider we ooding at the object as particles in the air, then consider one particle.



Volumetric Visibility:



The intensity in s can be caculated:

$$I(s) = \underbrace{I(s_0)}_{\substack{\uparrow \\ \text{intensity} \\ \text{at } s_0}} \cdot \underbrace{e^{-T(s_0,s)}}_{\substack{\uparrow \\ \text{exp decay}}} + \int_{s_0} \underbrace{q(\tilde{s})}_{\substack{\uparrow \\ \text{emission at} \\ \tilde{s}}} e^{-T(\tilde{s},s)} ds$$

where $\underset{\substack{\uparrow \\ \text{extinction}}}{T(s_1,s_2)} = \int_{s_1}^{s_2} \underset{\substack{\uparrow \\ \text{absorption}}}{k(s)} ds$

Consider 1 point then the color that we expect from that point is caculate as:

$$C(s) = C(s_0) \cdot \sigma(s_0) \cdot e^{-\tau(s_0,s)} \quad \text{where} \quad \tau(s_1,s_2) = \int_{s_1}^{s_2} \sigma(s) ds.$$

$\hookrightarrow$ $$C(r) = \int_{t_n}^{t_f} c(r(t),d) \cdot \sigma(r(t)) \cdot e^{-\tau(t_n,t)} dt.$$

$\hookrightarrow$ exp decay

$\hookrightarrow$ emission of previous points.

Inverse Transform Method:

Suppose $U \sim Unif[0; 1]$ & F is a 1-D cdf. Then:

$$X = F^{-1}(U)$$

has distribution F. Here we define:

$$F^{-1}(u) = \inf\{x : F(x) \geq u\}$$

$$\hat{C}(r) = \sum_{i=1}^{N} T_i(1 - \exp(-\sigma_i \delta_i))c_i \text{ , where } T_i = \exp(-\sum_{j=1}^{i-1} \sigma_j \delta_j)$$

$$\delta_j = t_{i+1} - t_i$$