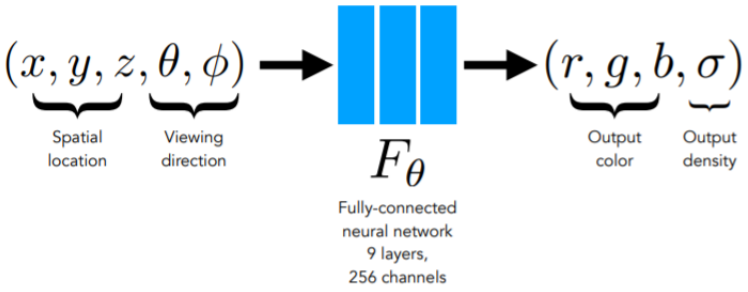
CS231n Lecture 16: Neural Radiance Fields

Sinkoo

March 27, 2022

1 NeRF

A neural radiance field is a fully connected network trained to reproduce input views of a single scene using a rendering loss. The network directly maps from spatial location and viewing direction (x, y, z, θ, ϕ) to color and opacity (r,g, b, σ), acting as the "volume" so we can use volume rendering to differentiably render new views.



To render this NeRF, there are 3 steps:
1. march camera rays through the scene to sample 3D points
2. use points from step1 and their corresponding 2D viewing directions(,) as input to the MLP to produce an output set of colors (c = (r, g, b)) and densities σ.
3. use volume rendering techniques to accumulate those colors and densities into a 2D image

2 Rendering

Divide the ray into N parts, sample one point from every part and get the color with:

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

Volume rendering is trivially differentiable

Rendering model for ray r(t) = o + td:

$$C \approx \sum_{i=1}^{N} T_i\alpha_i c_i$$

colors

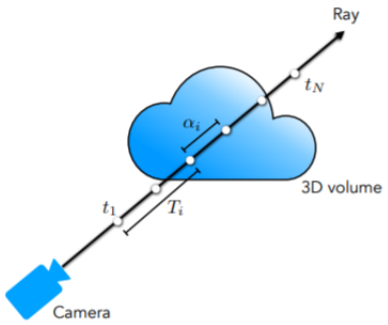
weights

How much light is blocked earlier along ray:

$$T_i = \prod_{j=1}^{i-1}(1 - \alpha_j)$$

How much light is contributed by ray segment *i*:

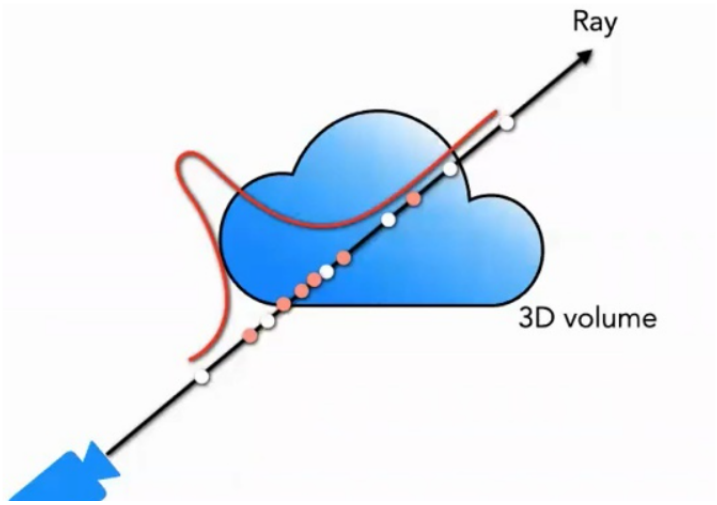
$$\alpha_i = 1 - e^{-\sigma_i\delta t_i}$$



But this model cannot up to state-of-the-art. Always produce image with blurring.

3 Improvements

Naive implementation produces blurry results.
1.Position encoding:Project 5D inputs into higher dimension.(Previous studies have shown that neural network is sensitive to low-frequency information and easy to ignore high-frequency texture information)
2.Hierarchical sampling procedure:



Use two networks training at the same time (coarse and fine networks). The input points of coarse network are obtained by uniformly sampling the light. The distribution of light is estimated according to the predicted by coarse network,sample from the estimated distribution, and then all samples are input into fine network for prediction.