**三维重建**

# 1 三维重建概述

## 博客综述

### 1.1.1 NeRF与三维重建专栏（一）领域背景、难点与数据集介绍

1. 网址： <https://juejin.cn/post/7232499180659458109>
2. 个人总结：
   1. **三维重建的方法**
      1. 传统基于几何的方法：使用NCC算法计算相关性然后记录视差，进而相似三角形做无尺度深度恢复
         1. 相关链接： <https://blog.csdn.net/weixin_44177123/article/details/105770779>
      2. 基于网络的：
         1. MVSNet为baseline的深度学习方法
         2. 基于NERF的
   2. **相关三维重建数据集：**
      1. 介绍了一堆 暂不研究
   3. 重建的评估指标：
      1. 图像渲染方面：
         1. PSNR 峰值信噪比
         2. SSIM 结构相似度
         3. LPIPS(Learned Perceptual Image Patch Similarity)
      2. 三维形状方面：
         1. CD(Chamfer Distance) 是一种常用的点云或mesh重建模型评估指标
            1. 涉及到几个概念

Nerf网络试过一次关于为SDF( signal distance file)场

需要将SDF、使用marching cubes算法来提取mesh

再进一步才能使用CD来度量点云/mesh

### 1.1.2 NeRF与三维重建专栏（二）NeRF原文解读与体渲染物理模型

1. 网址： <https://blog.csdn.net/KANG157/article/details/131010242>
2. 简要总结：
   1. 该部分相当于对2.1所述论文做了一个解释和概括

### 1.1.3 NeRF与三维重建专栏（三）nerf\_pl源码部分解读与colmapcuda算子使用

1. 网址： <https://blog.csdn.net/KANG157/article/details/131010803>

## NERF所用相关背景知识记录(辅助理解1.1)

**Nerf 四大核心点：**

**体渲染**

**MLP**

**Position encoding**

**Hierarchical sample**

### 体渲染理论

1. 神经辐射场的抽象理解辅助：
   1. 博客 <https://zhuanlan.zhihu.com/p/380015071>
   2. 总结：
      1. 基于全局坐标系下空间中一个点X（x,y,z）在某一个特定的观测方向D(dx,dy, dz)上的辐射值（具象为颜色贡献值）。可以由该点的 **体密度估计值** σ 和该点在该视角下的颜色**估计值**c=(r,g,b) 计算而得
2. 体素渲染 volume rendering的理论公式推导
   1. 博客推导：<https://blog.csdn.net/YuhsiHu/article/details/124318473>
   2. 视频讲解 <https://www.bilibili.com/video/BV1Qd4y1r7ZX/?spm_id_from=333.337.search-card.all.click&vd_source=4bd410628a2ee4d81ad9e6cdb964ee7f> 11min50s处

## 1.3 nerf++作者 基于nerf三维重建视频概述

视频链接：<https://www.bilibili.com/video/BV1d34y1n7fn/?spm_id_from=333.337.search-card.all.click&vd_source=4bd410628a2ee4d81ad9e6cdb964ee7f>

### 1.3.1 基于NeRF的三维内容生成

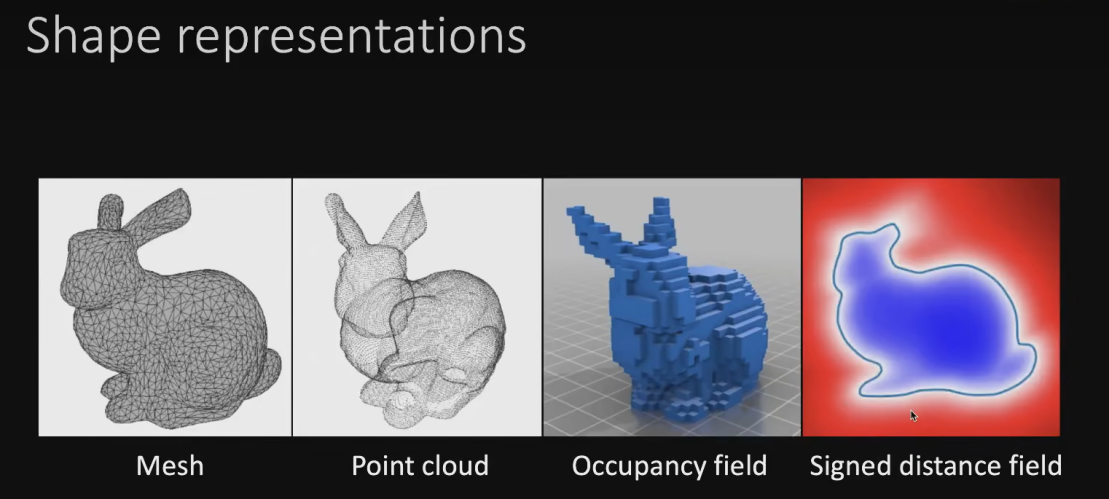
**Outline:**

1. 为什么要从图片中生成三维内容？
   1. 什么是三维内容
   2. 渲染反渲染
   3. 反渲染中的三个重要因素
2. Nerf为啥是个big thing（重大意义的东西）
   1. Netf三大组成成分
   2. 在nerf中总结出来的有用的技术点
   3. Nerf拓展：
      1. Unbonded scene 无边界的场景
      2. anti-aliasing 抗锯齿
3. 可编辑的三维内容
   1. 重新打光 材料可编辑
   2. 风格化的三维内容生成
   3. 其他的三维内容可编辑操作

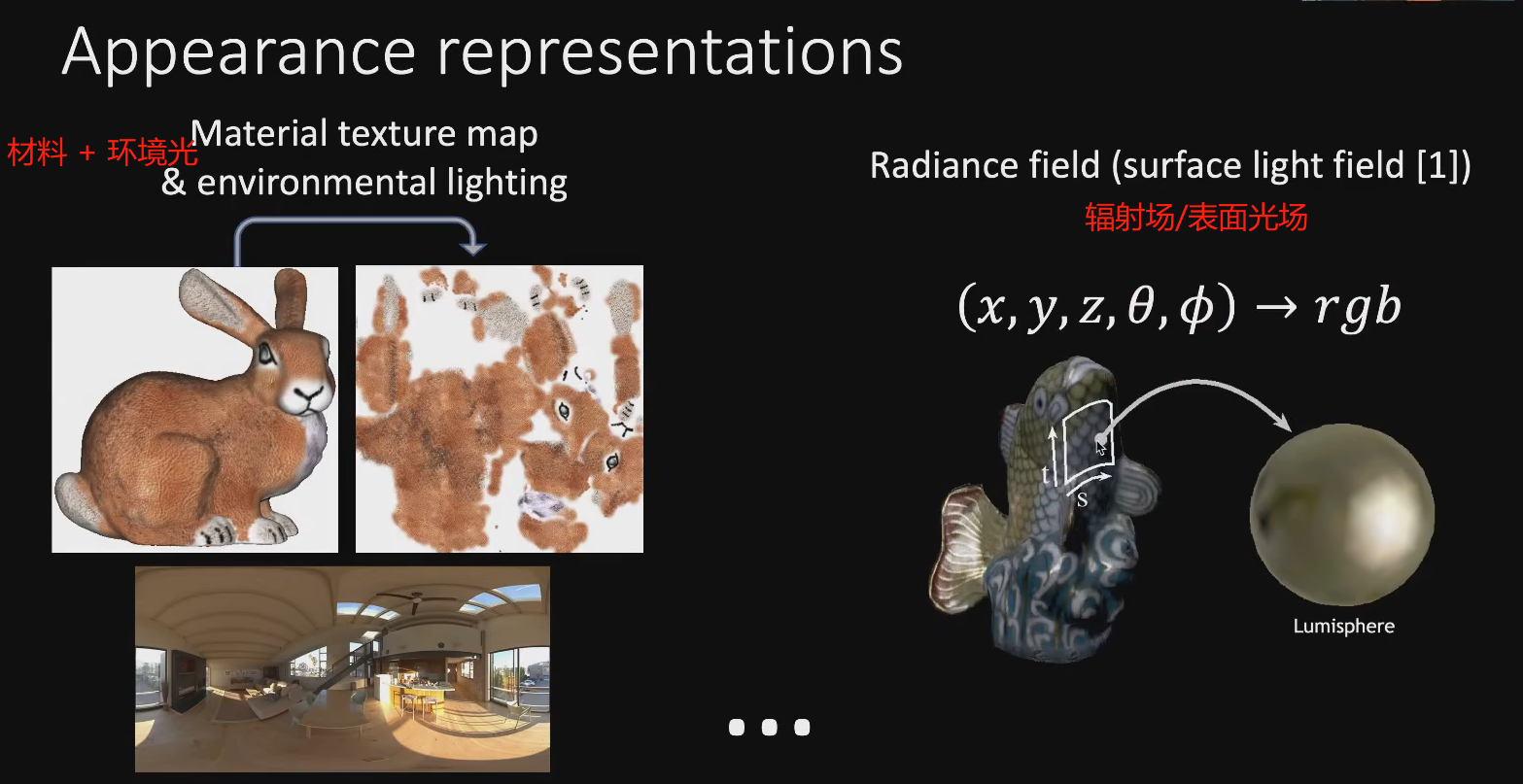
**开整:**

#### 为啥要生成三维内容

1. 什么是三维内容： shape + apperance 并且可以在不同的视角上渲染乘图片
2. \*
3. 反渲染的三个重要因素
   1. 用什么表示shape？ Volume density
   2. 用什么表述外观（appearance） C(r)
   3. 用啥方法 完成渲染？ 体素渲染公式



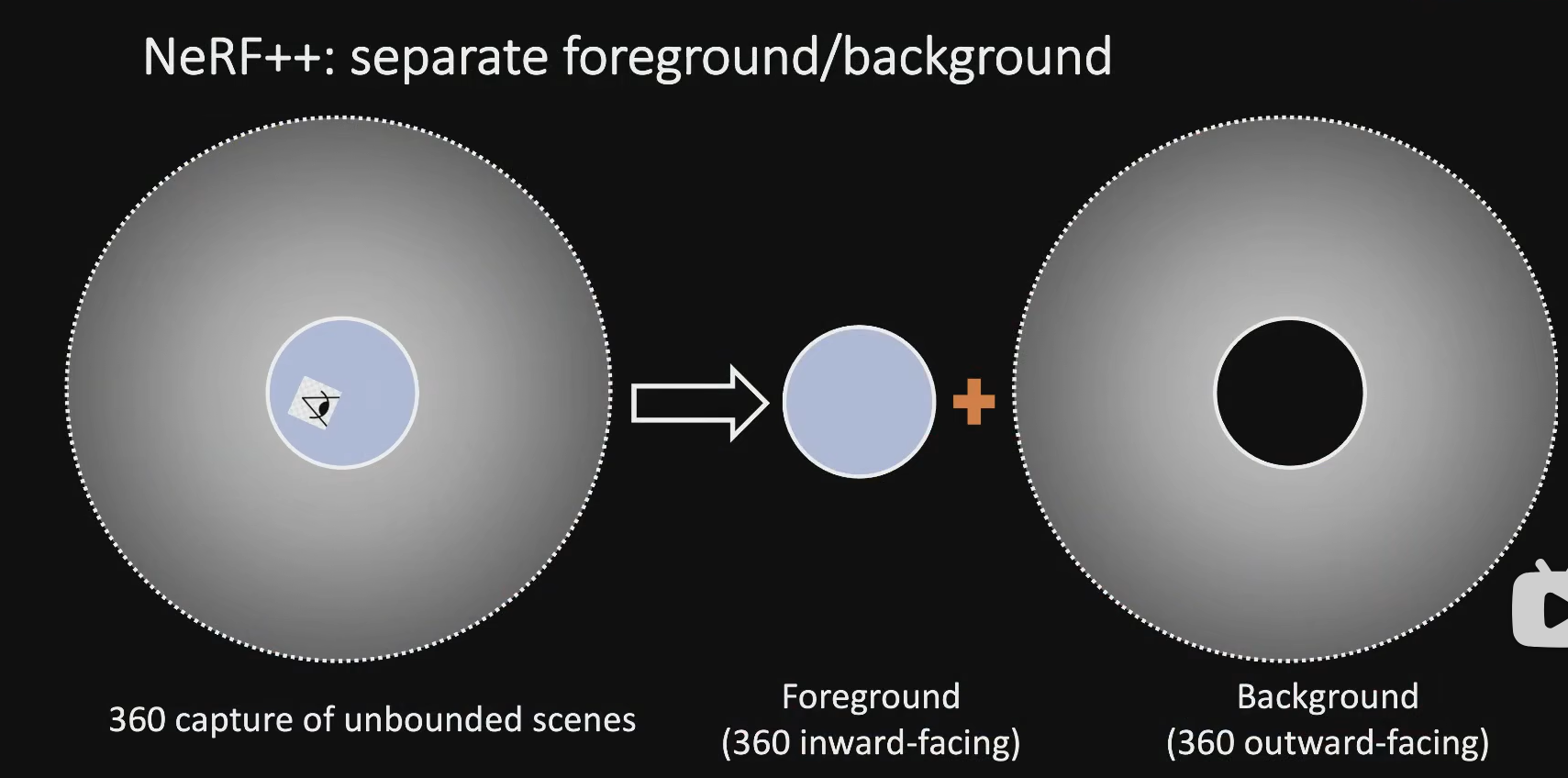
三维内容中形状的4种表示



三位内容中appearance representaion的2中表示方案

#### Nerf的重要贡献

1. Nerf三要素
   1. 体素密度
   2. 特定方向下的 点 emit radiance
   3. 渲染公式
2. Take home message
   1. Soft shape 相对于描述一个位置xyz 这种硬性描述（gard geometry）
      1. Soft shape 缺点一： expensive
      2. 缺点二： 没有可编辑性
   2. 位置编码部分的 Fourier features
3. Nerf拓展： 改变采样方式
   1. 当前nerf问题：360度拍摄时，无边界的场景下分辨率问题
      1. 背景很大，如果focus在整个场景 前景就糊了，反之背景就糊了
   2. 解决方案：nerf ++ --> 两个网络分别去拟合前景和背景



* 1. 如上方案前景的resolution问题可以解决但是背景很大 仍然会模糊（resolution低）

# 2 相关论文阅读总结

## 2.1 NeRF:原文 2020

**疑问点：位置编码公式怎么理解？**

### 2.1.0 Abstract--摘要

1. 本文贡献：
   1. 提出了一个新算法：在复杂场景下的新视角合成 方面取得了state-of-art的成果
2. 上述算法的概述：
   1. 使用了一个全连接的深度网络来represent 一个场景
   2. 该网络的**输入**：一个连续的5D coordinate (spatial location (x; y; z) and viewing direction (θ; φ))
   3. 该网络**输出**：在该空间位置上的体积密度 和 该视角下的发射辐射值output is the volume density and view-dependent emitted radiance at that spatial location
   4. 渲染时：将新视角下射线上的点使用传统的体素渲染公式 生成颜色和体积（深度信息）

总言之该算法的网络部分可以干这个事儿：Our method optimizes

a deep fully-connected neural network without any convolutional layers (often

referred to as a multilayer perceptron or MLP) to represent this function by

regressing from a single 5D coordinate (x; y; z; θ; φ) to a single volume density

and view-dependent RGB color.

### 2.1.1 Instruction

有两个概念结合博客的叙述对一下：

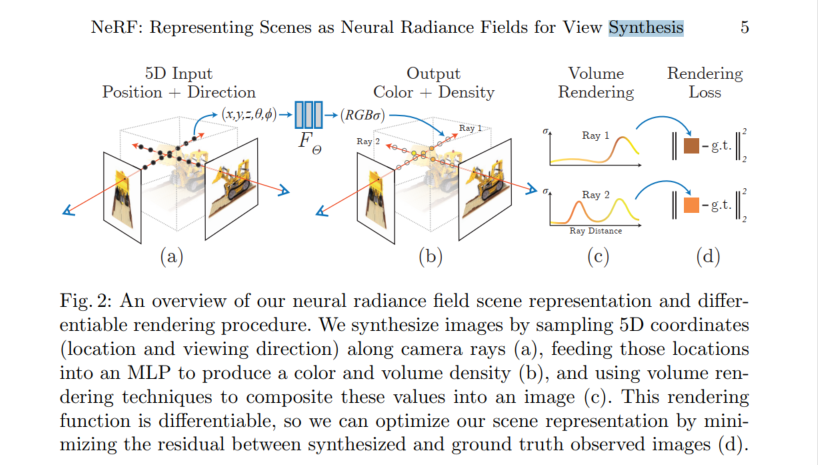
σ： Volume **density 点**体密度，其物理意义为射线在( x , y , z ) (x,y,z)(x,y,z)处击中粒子的概率 https://blog.csdn.net/KANG157/article/details/131010242

c: color：该点在**该（特定）视角下的**颜色估计值 <https://zhuanlan.zhihu.com/p/380015071>

#### 算法步骤分为三大步

* 1. march camera rays through the scene to generate a sampled set of 3D points
     1. 相机射线上生成3D采样点集
  2. use those points and their corresponding 2D viewing directions as input to the neural network to produce an output set of colors and densities
     1. 使用这些点和他们对应的2维（经纬度表示等价于三位方向向量）的角度方向作为输入。经过一个神经网络 获得这些点在该射线方向上的颜色和点体密度值
  3. use classical volume rendering techniques to accumulate those colors and densities into a 2D image
     1. 使用传统体素渲染的技术计算这些颜色和密度 得到2D图像上该点的颜色值 至此，一个新的视角的图像就可以渲染出一个点 推及全图

概述图：



#### 作者发现直接按照上述步骤干这个活有俩问题

We find that the basic implementation of optimizing a neural radiance field representation for a complex scene does not converge to a sufficiently highresolution representation and is inefficient in the required number of samples per camera ray

* 1. 复杂场景下高分辨率外观不能有效收敛（就是复杂场景做的烂）
     1. 论文提出解决方案：将输入的5D coordinate使用位置编码转成类似频域的高纬表示
  2. 相机设想上一定量的采样点会使得效率低（这个活干的慢）
     1. 方案：分层采样

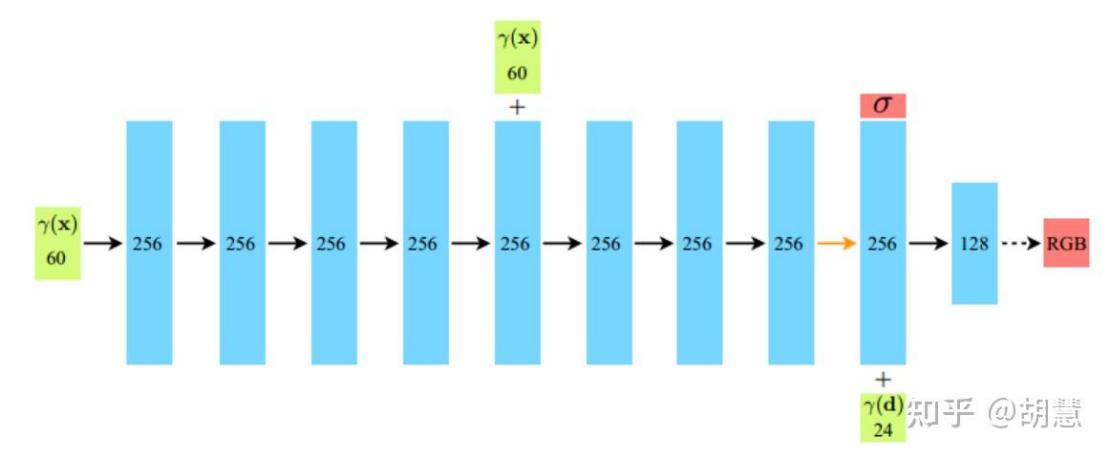
### 2.1.2 Related work

总：this section, we review these two lines of work and contrast them with our  
approach,

### 2.1.3 Neural Radiance Field Scene Representation

讲了输入 一个点的位置X(x,y,z)和方向向量d(dx, dy, dz)是如何使用论文所述神经辐射场来计算出该点的volume density和该视角下的 color的

1. 原文如此描述
   1. To accomplish this, the MLP FΘ first processes the input 3D coordinate x with 8 fully-connected layers (using ReLU activations and 256 channels per layer), and outputs σ and a 256-dimensional feature vector. This feature vector is then concatenated with the camera ray’s viewing direction and passed to one additional fully-connected layer (using a ReLU activation and 128 channels) that output the view-dependent RGB color
   2. 神经辐射场MLP结构配图：<https://zhuanlan.zhihu.com/p/466217848>



### 2.1.4 Volume Rendering with Radiance Fields

（使用神经辐射场做渲染）

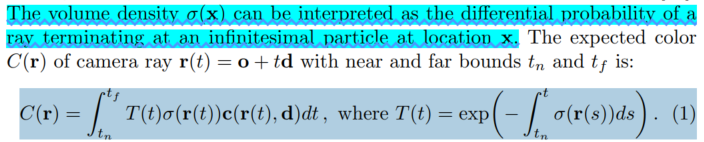
The volume density σ(x) can be interpreted as the differential probability of a

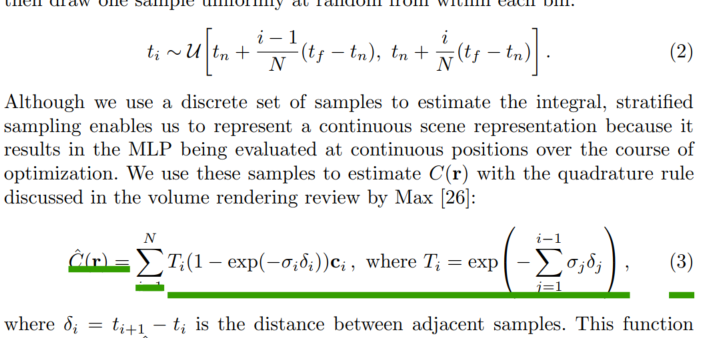
ray terminating at an infinitesimal particle at location x

渲染时本文将volume density解释为：光线在位于x处的无穷小粒子上被终止的可微概率

人话：在该点处不透光的概率 这个概率还是可微的

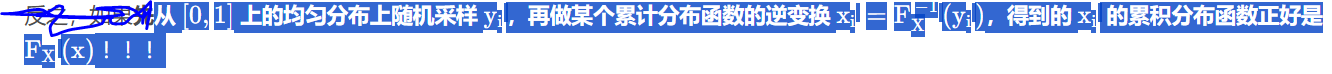
积分公式表述为：





### 2.1.5 Optimizing a Neural Radiance Field

神经辐射场 两个优化点

1. 问题一：
   1. we found that having the network FΘ directly operate on xyzθφ input coordinates results in renderings that perform poorly at representing high-frequency variation in color and geometry
   2. 解决方案：位置编码（就是用频域表示坐标值？不知道如何转的）
2. 问题二：
   1. Our rendering strategy of densely evaluating the neural radiance field network at N query points along each camera ray is inefficient:
   2. 解决方案： 分层采样
      1. 粗网路（Nc 粗采样点）+细网络（Nf 系采样点） Nc + Nf总的采样点放到渲染函数里一起算渲染值
      2. 逆变换采样概念 <https://zhuanlan.zhihu.com/p/622443806>
         1. 公式推导：下方两个博客总之推导出这么一个结论：任意分布，将其累积分布函数做为映射函数后得到一个新的分布，该映射后的分布就是均匀分布即其概率密度函数为均匀分布密度函数。故而
            1. <https://goodgoodstudy.blog.csdn.net/article/details/105962516>
            2. <https://blog.csdn.net/itnerd/article/details/105968943>

### 2.1.6 Implementation details

具体实现细节：