

nerf应用于多模态匹配中

匹配要求获得两张或多张图像之间的相互映射关系，

nerf应用于三维重建，输入的是多张图像，输出对应的rgbs，

要解决的难点在于：少量图片是否能够生成三维图像。

输入图像是有确定的位姿关系，但是在匹配中却是要求这个匹配关系的。

或者换一个思路，参考nerf的推导过程，

nerf不是基于大数据模式的，nerf只考虑单个场景的问题，用mlp隐式表示三维场景。通过单个场景的多个离散图像，恢复出连续的场景信息图。

我的理解是nerf中的mlp是用于储存该场景内光线映射的关系，输入的不再是多个图片信息，而是位置编码

对于匹配问题：是否可以通过对两张图像的已知位姿关系，求解出神经匹配场，获得连续的匹配信息图，从而获得位置匹配关系？

1. 要考虑这个模式是否可行：

要让匹配信息储存在网络里面，输入的不再是图片，而是？

用mlp隐式表示匹配场景？

2. 模仿视频插帧的工作，在重建场景下完成匹配的任务，任务网络在重建的过程中就隐式完成了匹配的表示。有点类似nerf隐式表示三维场景。

首先通过一个传统计算得到粗略的匹配关系，

网络输入像素坐标和时间点 $[x,y,t]$ ，输出对应的像素值 $[r,g,b]$ ，通过插值中间的关系输出矩阵的groundtruth。

后期可以使用cycle mlp的设计完成跨模态的匹配任务。

[1]NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis

[2] MVSNerF: Fast Generalizable Radiance Field Reconstruction from Multi-View Stereo ICCV 2021 训练一个具有泛化性能的先验网络，在推理的时候只用3张输入图片就重建一个新的场景。

[3] NerfingMVS: Guided Optimization of Neural Radiance Fields for Indoor Multi-view Stereo