

基于NeRF的物体重建和新视图合成
任务3

神经网络和深度学习

-期末作业-

马超

2024 年 6 月 29 日

目录

1	实验简介	1
2	NeRF基本原理	1
3	COLMAP简介	2
4	数据集和预处理	2
4.1	colmap导出相机参数	3
5	渲染3D图片	3
6	NeRF训练设置	4
6.1	评价指标	4
7	实验流程	5
8	github链接	5

1 实验简介

本实验利用NeRF（神经辐射场）模型，通过对物体的多角度拍摄图片/视频，并使用COLMAP进行相机参数估计，实现了物体的重建和新视图合成。本报告详细描述了实验的数据采集、模型训练与评估过程，并展示了训练中的损失曲线以及在测试集上的定量评价结果。

2 NeRF基本原理

NeRF利用神经网络表示场景中每个空间点的辐射场（Radiance Field），即每个点的颜色和密度随空间位置和视角的变化而变化。具体来说，NeRF模型包括两个核心部分：

- NeRF Encoder: 将输入的多视角图像转换为每个空间点的隐含表示。
- NeRF Decoder: 根据相机参数和空间点位置，生成像素的颜色值和密度。

NeRF技术已经在多个领域得到应用，包括但不限于：

- 虚拟现实（VR）和增强现实（AR）：生成逼真的虚拟场景，用于虚拟现实应用和增强现实体验。
- 电影和动画制作：用于生成复杂场景的特效和背景。
- 计算机视觉与图形学研究：用于场景重建、物体识别和3D渲染等方面。

尽管NeRF在生成逼真图像方面取得了显著进展，但仍面临一些挑战，如模型训练速度、场景动态性的处理以及硬件需求等。未来的研究可以集中在以下方向：

- 提升模型的速度和效率，使其能够处理更大规模和动态的场景。
- 结合传感器数据和深度学习模型，进一步提升场景重建的精度和实时性。
- 探索NeRF在医学成像、工业检测和自动驾驶等领域的应用潜力。

NeRF作为一种新兴的3D重建技术，展示了在生成高保真度图像合成方面的潜力和优势。随着技术的进步和应用场景的拓展，NeRF有望在多个领域发挥重要作用，并推动虚拟现实、增强现实以及计算机视觉研究的进步。

3 COLMAP简介

COLMAP (COLlaborative Mapping and Analysis Platform) 是一种流行的结构光学成像软件, 用于从图片集合中重建三维场景, 并估计相机参数和场景结构。使用COLMAP导出估计的相机参数可以通过以下步骤完成:

1. 安装和配置COLMAP: 确保已正确安装和配置COLMAP软件, 可以通过官方网站或GitHub获取最新版本。
2. 导入图片集合: 在COLMAP中导入包含多角度拍摄图片的数据集。
3. 运行图像匹配和三维重建: 使用COLMAP进行图像匹配和三维重建, 生成稀疏或稠密的点云和相机位姿。
4. 导出相机参数: 从COLMAP界面中导出估计的相机参数, 通常保存为文本文件或相机参数文件 (如.json或.txt格式)。

在导出相机参数时, 需要注意以下几点:

- 相机参数格式: 确保导出的相机参数文件格式与后续使用的3D重建或视图合成框架兼容。
- 精度和误差: COLMAP估计的相机参数可能存在误差, 这可能会影响后续应用的精度和稳定性。
- 后续处理: 根据具体需求, 可能需要对导出的相机参数进行后续处理或调整, 以满足特定应用的需求。

COLMAP是一种强大的工具, 用于从图片集合中估计相机参数并重建三维场景。通过正确使用和导出相机参数, 可以为后续的3D重建、视图合成和其他计算机视觉任务奠定良好的基础。

4 数据集和预处理

我们拍摄盆栽植物的多角度照片作为数据集进行渲染。然后使用COLMAP工具从多角度拍摄的物体图片中导出相机参数和场景点云。数据集被划分为训练集和测试集, 用于训练和评估NeRF模型。

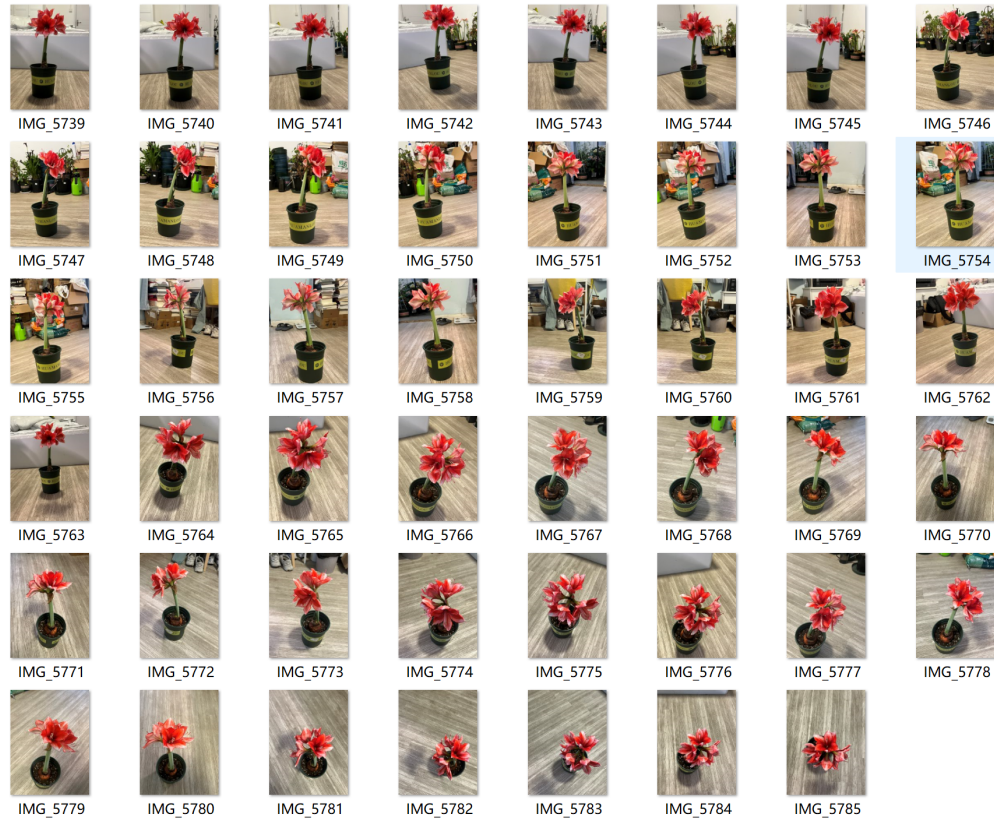


图 1: 拍摄的盆栽植物图片作为数据集

4.1 colmap导出相机参数

```
# Camera list with one line of data per camera:
# CAMERA_ID, MODEL, WIDTH, HEIGHT, PARAMS[]
# Number of cameras: 6
1 PINHOLE 3024 4032 3769.3841822335967 3642.0682354584624 1512 2016
6 PINHOLE 3024 4032 3604.139929509352 3580.9514299381949 1512 2016
7 PINHOLE 3024 4032 3626.8808029572015 3556.9591072049702 1512 2016
10 PINHOLE 3024 4032 3817.0478533753844 1416.7908162556146 1512 2016
11 PINHOLE 3024 4032 3387.6198707171507 4382.1160342047351 1512 2016
13 PINHOLE 3024 4032 4168.4828665899495 7438.9959424983981 1512 2016
```

图 2: colmap导出相机参数示例

5 渲染3D图片

colmap3D重建的盆栽图片如下:



图 3: 盆栽图片3D重建

6 NeRF训练设置

- **Batch Size:** 设置为8或16, 根据显存大小调整。
- **Learning Rate:** 初始设定为 $1e-4$ 至 $5e-4$, 使用Adam优化器进行训练。
- **迭代次数:** 总训练迭代次数设定为200,000至500,000次。
- **Loss Function:** 使用渲染损失和正则化损失优化模型。

6.1 评价指标

- **PSNR:** 在测试集上计算生成图像与真实图像之间的峰值信噪比。
- **其他指标:** 可根据需要选择使用的评估指标。

7 实验流程

1. 数据准备： 使用COLMAP导出相机参数和场景点云，划分数数据集为训练集和测试集。
2. 模型训练： 在训练集上训练NeRF模型，调整网络结构、损失函数和优化器进行优化。
3. 模型评估： 在测试集上生成新视图，计算评价指标（如PSNR）以评估模型性能。
4. 结果分析和报告： 使用Tensorboard可视化训练过程中的损失曲线

8 github链接

repo的public github链接：

<https://github.com/Ethereal-Redolent/hw/tree/master>