*课程大作业报告*

基于Nerf的足球运动员虚拟分身三维构建

徐阳，常烁晨

摘 要: 在数字化时代，随着科技的进步，元宇宙技术逐渐成为现实与虚拟世界之间的桥梁，元宇宙为体育领域带来了革新，特别是在虚拟足球形象构建方面展现出巨大潜力。本项目提出的足球运动员虚拟分身三维构建系统框架包括数据的采集、预处理、模型重建、个性化处理等环节，确保了数据处理和模型建立的整合性与效率。项目利用了NeRF技术，通过Kiri Engine软件捕捉足球运动员的二维图像并处理数据，结合Nerf Studio框架进行高精度的三维形象重建。本项目的方法不仅确保了虚拟形象的逼真度，而且通过Roop软件实现了形象的个性化定制，进一步提升了用户的沉浸感。这项工作在优化数据处理流程和提高模型建立便利性方面取得了创新，为足球运动员虚拟分身的三维构建提供了新的视角。

Three-Dimensional Construction of Football Players' Virtual Avatars Based on NeRF

Yang Xu, Shuochen Chang

**Abstract**: In the digital age, with the advancement of technology, Metaverse has gradually become the bridge between the real and virtual worlds, bringing innovation to the sports field, especially showing great potential in the construction of virtual football images. This project proposes a three-dimensional construction system framework for the virtual avatar of football players, which includes data collection, preprocessing, model reconstruction, and personalized processing, ensuring the integration and efficiency of data processing and model building. The project utilizes NeRF technology, capturing the two-dimensional images of football players with Kiri Engine software and processing data, combined with the Nerf Studio framework for high-precision three-dimensional image reconstruction. Our method not only ensures the realism of the virtual image but also achieves personalized customization through Roop software, further enhancing the user's immersive experience. This work has innovated in optimizing the data processing flow and improving the convenience of model building, offering a new perspective for the three-dimensional construction of football players' virtual avatars.

**Key word：**NeRF, Scene Representation, Metaverse, 3D Reconstruction

# **简介与意义/Introduction**

## 项目意义和依据/Significance

在当前的技术发展趋势中，元宇宙已成为连接现实与虚拟世界的重要桥梁。特别是在体育领域，如足球，它为球迷提供了一个全新的、互动性强的观赏和参与方式。通过创建足球运动员的虚拟形象，本项目不仅推进了元宇宙技术在体育领域的应用，还为球迷们提供了更加身临其境的体验。此外，该技术在运动员训练、比赛策略分析等方面也具有巨大的应用潜力，能够帮助教练和运动员更好地理解和提高比赛表现。

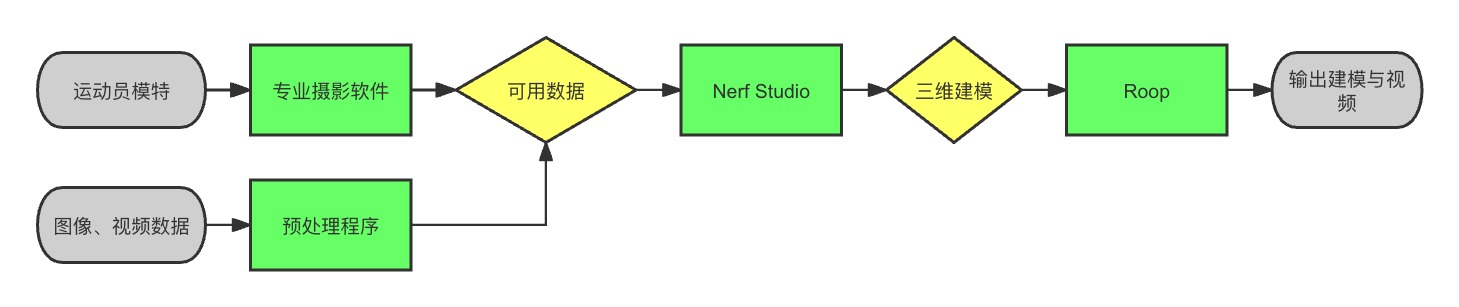
尽管元宇宙在体育领域具有广泛的应用前景，但在实现足球运动员虚拟形象的过程中，仍面临着诸多技术挑战。例如，如何精确捕捉并模拟运动员的运动特征、表情和动作是一个关键难题。此外，为了实现高度逼真和实时的虚拟体验，需要大量的数据和高效的建模渲染技术。同时，如何给用户提供个性化的体验，也是技术上的挑战。

针对这些挑战，本项目采用了NeRF[1] 技术来建立足球运动员的虚拟形象。这一技术能够通过深度学习算法高效地处理大量复杂的数据，从而创建出高度逼真的三维模型。本项目使用专业摄影软件如Kiri Engine 捕捉足球先锋的动作图片数据，使用开源框架Nerf Studio进行了精准高效的三维模型重建，并尝试使用roop软件进行换脸提供个性化的虚拟形象。我们重点关注了运动员动作的捕捉与三维模型重建，确保虚拟形象能够真实反映运动员的特点和风格。

## 本方法/系统框架/Article Structure

本项目的系统框架如图1所示，使用运动员模特构建虚拟形象，首先使用专业软件捕获足球运动员动作图像数据，或使用程序对一般的图像视频数据进行预处理，得到模型可用的带有相机位置与角度的图像数据。然后将数据用于Nerf Studio，使用了Neural Radiance Fields (NeRF)技术来创建具有高度真实感的三维运动员模型。NeRF能够生成复杂光线传输和细节次的场景，为用户提供逼真的三维体验。使用三维建模得到演示视频，并使用Roop等换脸技术个性化地为用户生成虚拟运动员形象的演示视频。

整个框架围绕数据处理、模型建立和个性化用户体验来构建，确保了从捕捉现实世界的动作到在元宇宙中呈现虚拟形象的整个流程的高度集成和效率。这种系统设计使得运动员的虚拟形象不仅具有高度的逼真感，而且能够以独特性和可适应性满足用户的需求。

图1 系统框架

# **相关工作/Related Works**

虚拟形象构建领域有许多有卓越的研究与技术，本项目主要调查了NeRF、换脸技术、虚拟试衣技术等前沿领域，并选择使用NeRF来构建项目，以下是对调查的各个领域的相关工作介绍。

## NeRF

NeRF（Neural Radiance Fields）是一种基于深度学习的3D重建技术，它通过训练一个神经网络来表示场景的连续体积密度和颜色信息。这种方法可以从一组稀疏的、有限角度的图像中重建出复杂的3D场景，生成新的视角图像。NeRF的关键创新在于它可以处理光线穿过场景时的复杂交互，如遮挡和光照变化，从而生成高质量、逼真的3D渲染图像。该技术在计算机视觉和图形学领域具有重要的应用前景。本项目所采用的人物形象三维构建，所基于的核心技术原理就是NeRF技术。

## 换脸技术

换脸技术，基于深度学习和生成对抗网络（GANs），是数字图像处理领域的一项重要进展。该技术依赖于如StyleGAN等高级算法，有效生成几乎无法被人眼识别的逼真人脸图像。其核心机制是将源图像的面部特征精确地转移到目标图像上，同时完整保留目标图像的表情和动作特征。通过对大量人脸数据的深入学习，这些模型不仅提高了图像的真实感，还能精确捕捉和再现面部表情和动作，使换脸后的成果在视觉上与原始面部极为相似。此技术在电影制作、游戏开发、安全验证等多个领域均显示出其广泛的应用潜力，并在数字娱乐、虚拟现实和身份验证等方面展示了其创新能力。

DeepFaceLab[2] 是最成功的换脸技术项目之一。DeepFaceLab是一个开源深度学习框架，致力于面部交换技术。自2018年开源以来，其主要贡献在于提供了一种先进的框架，引入了高效组件用于制作高质量的面部交换视频。该框架设计易于使用、性能高效、可扩展，支持对流程进行修改以满足特定研究需求。其可扩展性支持处理大规模数据集，实现电影级质量结果。DeepFaceLab的流程包括提取、训练和转换阶段，采用先进算法进行面部检测、对齐、分割和融合。这种综合方法及持续更新使其在娱乐和研究领域（特别是深伪视频检测）成为重要工具。该项技术可以为“运动人物形象三维构建”项目提供重要支持。DeepFaceLab通过先进的深度学习技术实现了高质量的面部交换，这一技术可以扩展到三维空间，辅助实现2D运动员照片向3D模型的转换。对于本项目的最终输出阶段， DeepFaceLab的虚拟换脸技术可以应用于三维运动员模型，实现将任意2D普通照片转换为3D运动员形象的任务要求。其可以提高模型的真实感和逼真度，拓宽虚拟换脸项目的应用范围，例如在虚拟现实、游戏设计或动画电影制作中的应用。通过这种跨领域的融合，项目探索了深度学习在视觉艺术和娱乐产业中的潜力。

## 虚拟试衣技术

近年来，虚拟试衣技术的兴起，极大促进了服装电商市场的蓬勃发展。虚拟试衣技术结合了增强现实、3D建模和深度学习算法的最新进展。虚拟试衣模型借助深度学习算法进行训练，根据大量输入数据，学习并优化服装的适配和显示效果，而增强现实技术和3D建模技术允许用户在使用虚拟试衣技术时，在虚拟环境中看到自己穿着不同服装的样子，提高了试衣的互动性和沉浸感，精确再现用户的身体特征和服装的细节，确保服装在虚拟试穿时的贴合度和外观真实性。

举例来说，Flow-Style-VTON[3] 是虚拟试衣技术领域的知名开源深度学习项目。该项目采用了一种基于风格的全局外观流方法，通过深度学习技术有效地模拟了服装在不同人体上的逼真外观。该项目支持直接使用未经处理的原始图像，简化了虚拟试衣的过程，同时提高了试衣效果的真实性和自然度。虚拟试衣技术可以应用到3D建模领域，实现三维人物形象着装的丰富性，无需做更多预处理，就可以完成一个3D形象的运动服装试穿。虚拟试衣技术在简化实现3D形象构建流程的同时，扩展了虚拟替代技术的应用范围。如今虚拟试衣不仅适用于专业的运动员模型，也适用于广泛的服装设计和数字时尚场景。在3D建模领域有着的广阔应用潜力。

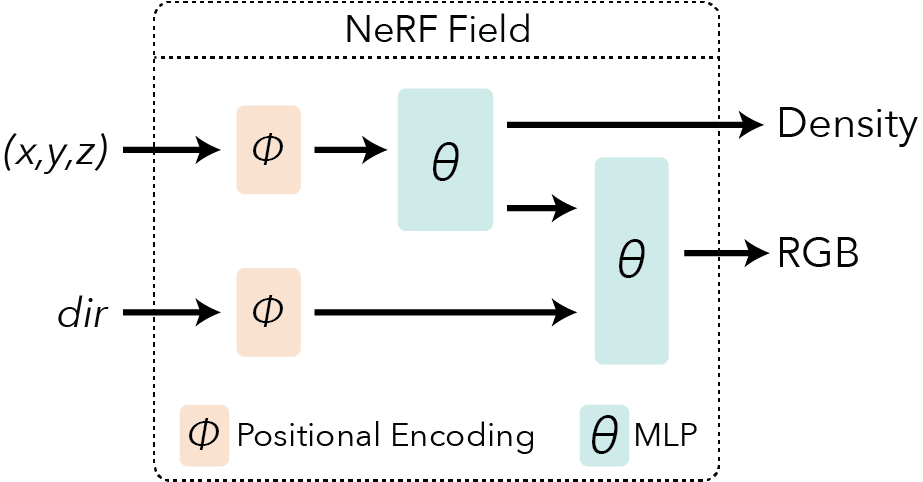
# **研究内容与方法（或算法）/Contnts and Methods(or Algorithm)**

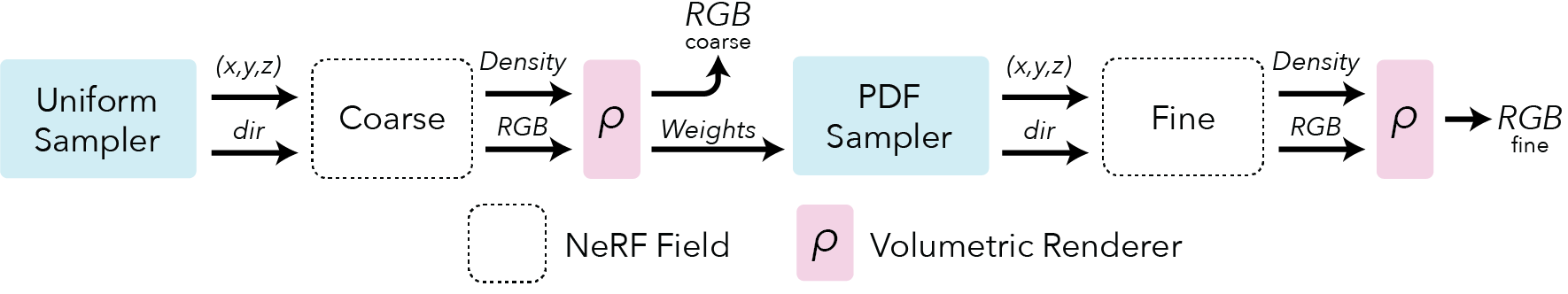
## NeRF（Neural Radiance Fields）算法

神经辐射场（Neural Radiance Fields，后简称NeRF）技术是一种基于深度学习的方法，用于从稀疏的二维图像重建场景的三维表示。 NeRF 模型可以学习场景几何形状、相机姿势和物体的反射特性，使其能够从新颖的视角渲染场景的真实感视图。[1] NeRF于 2020 年首次推出，因其在计算机图形和内容创建方面的潜在应用而受到广泛关注，其在计算机绘图和动画、医学成象、虚拟现实、卫星图像和规划等等领域均有所应用。

在NeRF模型中，将场景编码为由全连接深度神经网络（DNN）优化的体积函数。通过沿相机光线采样许多点，NeRF的全连接深度神经网络输入为相机的空间位置和以欧拉角表示的观察方向，输出为该空间位置处体积密度和与视图相关的发射辐射亮度，其原理由图2所示。NeRF通过查询沿相机光线的5D坐标来合成视图，并使用经典的体积渲染技术将输出颜色和密度投影到图像中。NeRF的整体流程图如图3所示。该网络可以预测 2D 图像中任意点的光强度（或辐射度），从而从不同角度生成新颖的 3D 视图。

然而，最初的 NeRF 模型存在训练和渲染速度慢、适应不同场景不灵活等局限性​​​​。本项目使用了开源项目NeRF Studio的Nerfacto模型，该模型是许多有效的已发表方法的组合，有着优秀的渲染精度和逼真度。

图2 NeRF中的场表示

图3 NeRF Pipeline 总览

## NeRF Studio

NeRF Studio是由伯克利人工智能研究中心（BAIR）的（KAIR）实验室于2022年推出的开源项目，其提供了简单的API，可以简化创建、训练和测试NeRF的端到端流程。该库通过模块化每个组件来支持 NeRF 的更可解释的实现。本项目使用了NeRF Studio提供的nerfacto模型，实现了2D到3D的虚拟形象建模过程，并探讨了不同参数对模型效果的影响。

### 环境搭建

本项目环境搭建使用的操作系统为Linux。

1）首先确保设备系统安装有CUDA的NVIDIA显卡，推荐使用Anaconda进行环境管理，安装Anaconda后，在终端输入如下命令创建虚拟环境：

conda create --name nerfstudio -y python=3.8

conda activate nerfstudio

python -m pip install --upgrade pip

2）接下来安装NeRF Studio依赖项，在已经激活conda虚拟环境的终端输入命令中安装PyTorch 2.0.1和CUDA 11.8：

pip install torch==2.0.1+cu118 torchvision==\

0.15.2+cu118 --extra-index-url <https://download.pytorch.org/whl/cu118>

3）要构建必要的CUDA扩展，还需要安装cuda-toolkit。建议使用conda安装，在终端中输入：

conda install -c "nvidia/label/cuda-11.8.0" cuda-toolkit

4）安装完以上环境后，还需安装tiny-cuda-nn，tiny-cuda-nn是NVIDIA针对NeRF的一系列工作开发的一个轻量级的高效框架，包含一个MLP和多分辨率哈希网格以及各种编码、损失和优化器的支持。在终端中输入：

pip install ninja git+https://github.com/NVlabs/tiny-cuda-nn/#subdirectory=bindings/torch

若网络不通畅，也可以选择在Github上下载项目后进行本地编译，具体可参考<https://github.com/nvlabs/tiny-cuda-nn>。需要注意的是，本地编译要求GCC/G++ 8或更高版本以及CMake-v3.21或更高版本。

5）安装Nerf Studio

安装完上述依赖项，在终端输入：

pip install nerfstudio

即可完成Nerf Studio安装。

### 数据获取

安装完Nerf Studio后，可使用ns-download-data nerfstudio --capture-name=poster下载Nerf Studio的测试数据。

若要使用自定义的训练数据，推荐使用Polycam、Kiri Engine、Record 3D等专业3D相机软件获取可直接使用的数据。以下以Kiri Engine为例，介绍足球运动员姿势图片数据集的获取。

在谷歌商店下载并安装Kiri Engine后，打开软件，点击下方选项栏最后的个人一栏，先点击设置（Settings）将语言切换为中文，并打开开发者模式（Developer Mode）。然后点击下方中心的加号准备录制，选择开发者模式下的“相机位姿”拍照模式开始拍摄，被拍摄的模特摆出足球运动员先锋的跑步姿势，用于之后足球运动员三维形象的构建。操作流程如图4所示

随后可在主页上方的“开发（Dev）”栏中找到拍摄的素材数据，其包括了拍摄的图片以及拍摄时对应的摄像机的坐标和拍摄角度，下载数据导入本地文件备用。

图4 数据拍摄流程图

### 训练

Nerf Studio的提供了一组预先实现的NeRF方法，如NeRF、 Instant-NGP[4] 、 Nerfacto、Instruct-NeRF2NeRF[5] 、 K-Planes[6] 等等。本项目中，使用Nerfacto进行训练。Nerfacto模型是Nerf Studio自己编写的模型，该模型是许多有效的方法的组合。

Nerf Studio通过在终端输入命令来控制各项任务。如使用ns-train训练数据，使用ns-viewer使用查看器，使用ns-export导出三维模型点云或mesh。以下详细介绍各项命令：

ns-train用于训练数据，ns-train后第一个参数为训练模型，如ns-train nerfacto，可通过在终端输入ns-train nerfacto –help查看可设置的各项参数。常用参数为--data {data\_path}、--load\_dir {checkpotin\_path}，分别用于指定数据路径和checkpoint路径（如果要载入checkpoint）。

ns-viewer用于实时查看训练结果，以及载入训练完成的模型，需要设置--load-config参数指定载入模型的yml文件。

## Roop

Roop是一个开源的轻量级视频/图片换脸工具，无需从头开始搭建深度学习模型，可以直接进行GUI的调用。该项目在Windows、Macos、Linux等操作系统以及不同的CPU、GPU硬件架构上均有适配以及对应的加速优化命令。

本实验在M1pro芯片的Macbook pro上运行，Windows、Linux等的运行方式均类似。

### 环境搭建

本项目在Anaconda下进行环境管理，首先确保实验主机安装了Anaconda、git等基本工

具。首先运行以下命令创建工作目录、安装视频处理插件：

cd ~

mkdir roop

cd roop

brew install ffmpeg

进入工作目录后输入以下指令：

conda create -n roop -y python=3.10

conda activate roop

python -m pip install --upgrade pip

随后将代码文件打包到本地并安装对应python包：

git clone https://github.com/s0md3v/roop

pip install -r requirements.txt

为了进行GPU加速，还需要安装依赖项

pip uninstall onnxruntime onnxruntime-silicon

pip install onnxruntime-silicon==1.13.1

执行项目，使用以下指令进行GPU加速：

python run.py --execution-provider coreml

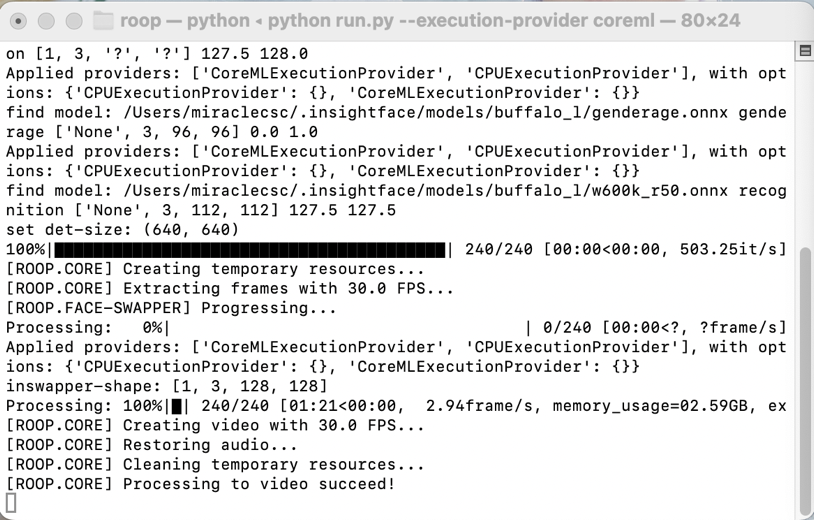


图5 roop使用过程

在终端输入命令后，将会打开GUI界面，如图5，上传人脸照片以及需要进行换脸的视频文件，对GUI进行操作，点击 Start即可实现roop的图片/视频换脸操作。

# **实验结果与分析/Experiment Results and Analysis**

本项目完成了足球运动员的三维虚拟形象建模。在报告的实验结果分析部分，将展示此项目对3D运动员形象的构建结果。本项目选择了足球运动员中“前锋”作为建模结果，前锋在足球场上的职责是进攻得分，本项目模拟出高速带球奔跑的前锋形象。

本项目的实验结果主要包括两部分：模型的3D点云&mesh，以及渲染后的模型展示视频。

## 点云&mesh

本项目将预处理过后的视频格式作为数据进行输入，经过训练后可以将三维模型的点云导出。在本项目中使用的NeRF Studio所提供的nerfacto模型有三种规格选择，分别是nerfacto、nerfacto-big、nerfacto-huge。更大的模型会有更加精细化的3D形象构建，相应的，其训练所需的计算资源（GPU及显存）也更高。

### nerfacto

最小规模的模型 nerfacto 在单张RTX4090训练过程中大约需要6G显存。完成3D建模后导出的点云&mesh如图6所示。



图6 nerfacto点云&mesh模型

### nerfacto-big

中等规模的模型 nerfacto-big 在单张RTX4090训练过程中大约需要12G显存。完成3D建模后导出的点云&mesh如图7所示，可以看到相比于基础版的nerfacto模型，mesh的精细度有相当的提升。

图7 nerfacto-big点云&mesh模型



### nerfacto-huge

最大规模的模型 nerfacto-huge 在单张RTX4090训练过程中大约需要24G显存。完成3D建模后导出的点云&mesh如图8所示，可以看到huge模型具有最高的完成度，模型的精细程度最好。



图8 nerfacto-huge点云&mesh模型

## 3D模型展示

我们此处展示的是使用nerfacto-huge训练后的三维建模结果。本项目旨在实现任意的普通2D输入到3D形象的构建，因此本项目通过支持换脸技术，使得对于任意的2D普通形象的输入（如周杰伦的照片），都可以得到对应的3D足球前锋形象。

### 任意选取普通2D人脸图片作为输入，如图9所示。对渲染的视频进行换脸。

图9 普通2D图片（周杰伦）

### 输出结果如图10所示。



图10 将2D普通照片转化为三维模型结果

# **特色与创新/ Distinctive or Innovation**

本项目基于NeRF技术，使用开源框架Nerf Studio并结合roop，将2D普通照片转化成3D运动员形象，实现了普通二维图片到三维运动员虚拟形象的完整流程。

首先，本项目通过Kiri Engine获取自定义的数据集。在自定义的数据集上训练得到一个符合基本要求的三维足球运动员形象模型。相比于普通的三维重建技术，本项目重点关注个性化的特定形象三维重建，借助GUI界面的roop工具、视频切片脚本等技术，实现了真实通过输入图片建模的的三维虚拟足球运动员形象。

此外，本项目针对不同实现方法、不同模型以及不同训练参数，实现了对三维模型输出效果的对比。在项目报告中，通过训练不同规模的模型，导出对应的点云&mesh，直观展现了不同模型对二维输入的三维运动员建模效果。

References:

1. Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020). NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. In European Conference on Computer Vision (ECCV).
2. Petrov, I., Gao, D., Chervoniy, N., Liu, K., Marangonda, S., Umé, C., ... & Zhang, W. (2020). DeepFaceLab: Integrated, flexible and extensible face-swapping framework. Retrieved from <https://arxiv.org/abs/2005.05535>
3. He, S., Song, Y.-Z., & Xiang, T. Style-Based Global Appearance Flow for Virtual Try-On. Retrieved from <https://arxiv.org/pdf/2204.01046.pdf>
4. Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. ACM Transactions on Graphics, 41(4), 102:1-102:15. <https://doi.org/10.1145/3528223.3530127>
5. Haque, A., Tancik, M., Efros, A., Holynski, A., & Kanazawa, A. (2023). Instruct-NeRF2NeRF: Editing 3D Scenes with Instructions. In Proceedings of the IEEE/CVF International Conference on Computer Vision.
6. Fridovich-Keil, S., Meanti, G., Warburg, F. R., Recht, B., & Kanazawa, A. (2023). K-Planes: Explicit Radiance Fields in Space, Time, and Appearance. In CVPR.

时间安排与分工统计表

|  |  |  |  |
| --- | --- | --- | --- |
| **组员信息**（含组长） | | | |
| 学生姓名 | 徐阳 | 学 号 | 521021910363 |
| 项目分工 | 方案制定、试验研究、环境搭建、数据处理、二维素材提供、模型训练、总结报告撰写、报告排版、答辩PPT制作 | | |
| 学生姓名 | 常烁晨 | 学 号 | 521021910369 |
| 项目分工 | 相关工作整理、实验与分析、环境搭建、三维素材、模型训练、对比建模、总结报告撰写、项目答辩介绍 | | |
| **时间安排/**  **Schedule** | 2023.12.8 初步探讨选题，接触NeRF框架广泛了解前沿技术进展，如视频换脸项目DeepFaceLab、虚拟换衣Virtual-Try-On，尝试搭建环境。  2023.12.16 尝试各种项目，选择NeRF Studio，搭建NeRF Studio环境，学习NeRF Studio项目操作流程。  2023.12.19 数据准备，采用Kiri Engine获取符合训练建模要求的个性化数据。  2023.12.21 训练基础nerfacto模型，在ns-viewer中导出点云模型、建模演示视频，并导入blender查看。  2023.12.22 训练nerfacto-big、nerfacto-huge模型，并在ns-viewer中导出点云和mesh，对比分析不同规格模型的生成mesh。  2023.12.31 实现roop换脸，完成示例视频的个性化面部替换重建。  2024.1.5 完成总结报告以及项目汇报PPT。 | | |