

# 团 体 标 准

T/INFOCA 4-2024

## 基于神经辐射场（NeRF）的三维场景建模过程指南

Guideline for three-dimensional scene modeling process based on neural radiation fields

（征求意见稿）

（本稿完成日期：2024 年 3 月 25 日）

××××-××-××发布

××××-××-××实施

中关村现代信息消费应用产业技术联盟 发布



# 目 次

|                    |     |
|--------------------|-----|
| 目 次                | I   |
| 前 言                | II  |
| 引 言                | III |
| 1 范围               | 1   |
| 2 规范性引用文件          | 1   |
| 3 术语和定义            | 1   |
| 4 缩略语              | 1   |
| 5 基于神经辐射场的三维场景建模过程 | 2   |
| 5.1 概述             | 2   |
| 5.2 过程图            | 2   |
| 5.3 过程描述           | 2   |
| 6 追溯和检查方法          | 4   |
| 6.1 追溯方法           | 4   |
| 6.2 检查             | 4   |

## 前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第1部分:标准化文件的结构和起草规则》给出的规定起草。

本文件的某些内容可能涉及专利。本文件的发布机构不承担识别这些专利的责任。

本文件由中关村现代信息消费应用产业技术联盟提出并归口。

本文件起草单位:中国传媒大学、北京大学、灵然创智(天津)动画科技发展有限公司、哈尔滨工业大学、中国科学院大学、广东南方新媒体股份有限公司

本文件主要起草人:叶龙、胡飞、钟微、方力、张启、王苔社、熊瑞勤、张炜、于浩、花侯鹏、范晓鹏、王兴涛、张新峰、安泓宇、李欢洋、王兵、罗泽文、郭永康。

## 引 言

随着科技的不断进步，计算机图形学、虚拟现实等领域对于高度逼真的三维场景建模的需求愈发迫切。

基于神经辐射场的三维场景建模技术是近年来涌现的一种创新方法，它通过深度学习的方式，使用神经网络对场景中每个点的辐射强度和密度进行建模，从而实现高质量的场景渲染。由于该技术的复杂性和多样性，缺乏统一的标准导致了在不同研究和应用中存在着差异，阻碍了该技术的广泛应用和交流。

为了推动神经辐射场技术的实际应用，降低技术应用门槛，提高三维场景建模的效率和准确性，推动相关领域的技术创新，特制定此标准。



# 基于神经辐射场（NeRF）的三维场景建模过程指南

## 1 范围

本文件规定了基于神经辐射场的三维场景建模技术过程。

本文件适用于虚拟现实中基于神经辐射场的三维场景建模技术。

## 2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

YD/T 4310-2023 移动互联网环境下虚拟现实业务术语

## 3 术语和定义

下列术语和定义适用于本文件。

### 3.1

**虚拟现实** virtual reality

通过计算机模拟产生三维虚拟环境，进行视觉等感官的模拟，利用头戴虚拟显示器和耳机等设备，让用户获得沉浸感，同时通过各类带有传感器的交互设备，感知用户状态，实现交互功能，从而构建与真实环境相近的虚拟场景。

[源自：YD/T 4310-2023，3.1.1]

### 3.2

**神经辐射场** neural radiance fields

基于深度学习的三维场景建模方法，通过训练神经网络来估计场景中每个点的颜色和密度信息实现高度真实感的图像渲染。

## 4 缩略语

下列缩略语适用于本文件。

Adam：自适应矩估计算法（Adaptive Moment Estimation）

NeRF：神经辐射场（Neural Radiance Fields）

## 5 基于神经辐射场的三维场景建模过程

### 5.1 概述

神经辐射场(NeRF)技术基于深度学习的原理,通过使用神经网络模型来捕捉场景中的辐射信息。该技术利用前馈神经网络来表示场景中每个点的辐射颜色和密度,从而实现了对复杂几何形状和光照效果的精确建模。

### 5.2 过程图

图 1 给出了基于神经辐射场 (NeRF) 的三维场景建模的过程。基于神经辐射场的三维场景建模的过程由以下 4 个子过程组成:

- a) 数据采集;
- b) 神经网络建模;
- c) 神经网络训练;
- d) 场景渲染。

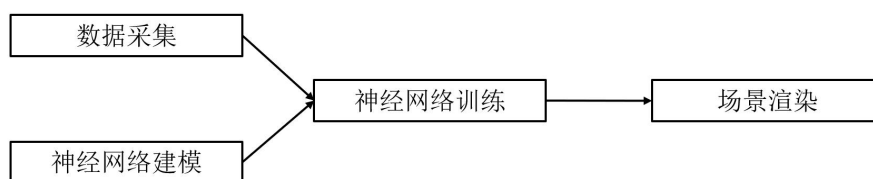


图 1 基于神经辐射场的三维场景建模过程

### 5.3 过程描述

#### 5.3.1 三维建模

数据采集是基于神经辐射场(NeRF)的三维场景建模的第一个过程,涉及以下关键子步骤:

##### 5.3.1.1 相机参数设置

- a) 焦距调整: 设置摄像机的焦距以捕捉场景细节,并根据场景尺度调整焦距,以保持适当的景深。
- b) 光圈大小调整: 调整光圈大小,控制进光量,平衡图像的亮度和深度信息获取。
- c) 快门速度选择: 选择合适的快门速度,防止运动模糊,确保图像清晰度。
- d) 感光度设定: 设置感光度,权衡图像亮度与噪声水平,确保不同光照条件下采集高质量的图像。

5.3.1.2 使用高分辨率摄像机从多个角度和方向拍摄场景,以获取全面的视角和覆盖不同光照条件的信息。

5.3.1.3 进行相机标定,获取准确的内外参数,确保图像在后续处理中的精确对准。



5.3.1.4 进行颜色校正，纠正摄像机传感器的颜色偏差，保持图像色彩一致性。

5.3.1.5 将采集的数据分为训练集和验证集，用于神经网络的有效学习和验证。

## 5.2.2 神经网络建模

神经网络建模是整个基于神经辐射场的三维场景建模的第二个过程。通过神经网络建模从空间坐标到三维场景颜色和密度的函数映射。神经网络建模的过程如下：

5.2.2.1 定义网络的图像输入层，确保能够接受预处理后的图像数据。通常，RGB 图像可以作为输入。

如果使用深度信息，定义深度图的输入层，以接收预处理后的深度图数据。

5.2.2.2 定义神经网络的隐藏层数量，根据场景复杂性和计算资源进行选择。

5.2.2.3 定义每个隐藏层中的神经元数量。每个隐藏层中的神经元数量需根据问题复杂性进行调整，建议根据实验结果进行优化。

5.2.2.4 选择适当的非线性激活函数，以引入非线性特性，建议根据实验结果进行优化。

5.2.2.5 定义输出层，确保能够输出每个采样点的颜色和密度信息。

5.2.2.6 定义初始学习率，一般设置为一个小的正数，如 0.001，建议根据实验结果进行优化。

5.2.2.7 宜定义学习率的衰减策略，以确保训练的稳定性，建议根据实验结果进行优化。

5.2.2.8 定义神经网络训练时的批次大小，宜选择合适的批次大小以提高训练效率，如 32 或 64，建议根据实验结果进行优化。

5.2.2.9 定义用于训练的损失函数，建议根据实验结果进行优化

5.2.2.10 选择适当的优化算法，以最小化损失函数，建议根据实验结果进行优化。

5.2.2.11 采用适当的权重初始化策略，以加速训练过程，建议使用 Adam 等优化器，可根据实验结果进行优化。

## 5.2.3 神经网络训练

神经网络训练是整个基于神经辐射场的三维场景建模的第三个过程。利用数据采集的数据训练神经网络建模模型，并保存模型参数。神经网络训练的过程如下：

5.2.3.1 定义训练的迭代次数，根据训练集大小和模型收敛情况进行选择。

5.2.3.2 定义早停策略，以在验证集上的性能不再提升时停止训练。

5.2.3.3 定义保存模型的策略，包括保存整个模型或仅保存参数。

5.2.3.4 定义模型恢复策略，以在需要时加载已经训练好的模型。

5.2.3.5 在训练集上根据 5.2.2 以及 5.2.3.1-5.2.3.4 的设置，在高性能显卡上训练神经网络模型。在训练过程中，根据实验结果调整 5.2.2 中设置的参数值，并定时保存训练模型。

## 5.2.4 场景渲染

场景渲染是整个基于神经辐射场的三维场景建模的第四个过程。从视点出发经过每个像素的射线，使用训练好的神经网络对每个采样点进行推断，得到颜色和密度的估计值。根据推断得到的颜色和密度，进行必要的修正。在每个像素上通过对射线上的颜色和密度进行积分，并映射到可视化的颜色范围，得到渲染图像。场景渲染的过程如下：

- 5.2.4.1 定义视点的位置，通常为摄像机的位置，确保视点能够捕捉到场景的关键信息。确定射线的方向，通常为从视点出发经过每个像素的射线，确保能够覆盖整个场景。
- 5.2.4.2 在射线上进行采样，以获取射线上的采样点，确保细致地捕捉场景的细节。
- 5.2.4.3 使用训练好的神经网络对每个采样点进行推断，得到颜色和密度的估计值。
- 5.2.4.4 根据推断得到的颜色和密度，进行必要的修正
- 5.2.4.5 在每个像素上通过对射线上的颜色和密度进行积分，得到最终的图像像素值。
- 5.2.4.6 将神经网络输出的颜色值映射到可视化的颜色范围，以获得真实彩色图像。

## 6 追溯和检查方法

### 6.1 追溯方法

采用“一项目一标识”的原则，为基于神经辐射场的三维场景建模的每个项目流程设立唯一的标识码。这一标识码可在项目立项时生成，并贯穿整个制作周期。

应对图 1 所示制作过程的所有阶段进行标识，构建每个环节的标识码，形成全过程的追溯链条。

### 6.2 检查

制定详细的制作记录，记录文件应在相应阶段得到合理保管，文件留存时间不少于两年。