

摘要

如何生成高质量的三维物体是计算机视觉领域研究的重点。近年来提出的神经三维表示具备通过二维的介质生成三维表示的能力，大大降低了三维资产的生成难度和成本。如何在输入缺失的情况下生成具有高质量颜色纹理的神经辐射场是重要的研究课题。在输入图片为单色的情况下，生成的神经辐射场的几何信息已经确定，需要恢复颜色纹理信息；在输入信息仅为文本信息的情况下，生成的神经辐射场的几何和颜色纹理信息都需要模型进行恢复和预测。本研究根据输入信息的缺失程度不同，分别研究神经辐射场的着色问题和基于文本的神经辐射场生成任务。

本文研究神经三维表示的颜色可控生成。在进行广泛的文献调研和大量实验的前提下，针对上述研究课题，本文分别提出了一个单色神经辐射场的着色模型和一个基于文本输入的三维可控生成模型，探究了二维模型在神经表示下的三维生成模型中的应用方式。本文的贡献可以分为如下几个部分：

1. 颜色质量提升：提出了一个从单色图像进行神经辐射场着色的模型，通过分析图像中物体的形状和光照信息，在新的颜色空间中为物体表面建立了亮度和密度表示。在此基础上利用图像着色模块恢复颜色，并提出了多个颜色可控生成的子模块，完成了颜色可控生成的进一步研究和探索。实验证明，提出的模型可以创建具有可行一致性和颜色鲜艳度的三维表示；
2. 基于文本的三维可控生成：提出了一种新的三维可控生成流程。在几何构建过程中，采用了逐步训练多种三维表示的方式提升模型的几何质量，并利用具有三维一致性的预训练模型提供指引，在表面蒙皮恢复阶段，提出了一个能够利用正常的隐式分类指引权重大小生成高质量蒙皮的分数蒸馏函数。在大量实验的验证下，提出的模型能够很好地利用扩散模型蕴含的信息，生成几何精细、蒙皮生动的三维模型。

综上所述，本文从三维资产的颜色质量提升和基于文本的可控生成两个关键子课题出发，对现有神经表示模型进行了研究和改进，提升了三维资产的生成效果，推动了以神经辐射场为代表的非传统图形学管线在三维资产质量提升和生成任务中的发展。

关键词：神经三维表示，着色，基于文本的可控生成