Title: Camera generation model system based on NeRF technology

Introduction

"In today's digital age, 3D content creation and rendering technologies are developing rapidly, providing us with unprecedented opportunities for virtual experiences. Among them, Neural Radiation Field technology (NeRF) has demonstrated its ability to reconstruct high-quality 3D from sparse views Powerful capabilities in scenarios. However, applying these advanced technologies to mobile or computing power-constrained devices remains a challenge."

NeRF technology overview

"NeRF uses deep learning to build a continuous 3D scene representation, learning the volumetric density and color information of the scene through an optimization process. This method can reconstruct realistic 3D scenes from a set of sparse and unstructured 2D images. and synthesize new images from unseen perspectives."

Project Objectives

"This project aims to develop a NeRF-based camera generation model system that can not only run efficiently on devices with limited computing power, but also generate features with diverse illumination and High-quality images from the perspective. We hope that the input photos and output 3D models can be synchronized, which can help users interact better "

Approach:

Network pruning: "We will use advanced network pruning technology to optimize the NeRF model, reduce its demand for computing resources, and enable it to run smoothly on mobile devices."

Variational Auto encoder (VAE): "By integrating VAE technology, our system will be able to generate views under different environmental conditions, thus enhancing the user experience."

Application scenarios

"The system has broad application potential in virtual reality, game development, online education and other fields. For example, it can be used to create interactive virtual museum exhibitions or provide dynamically generated environments for games."

Project advantages and challenges

"Our project, by combining NeRF with VAE technology, not only solves the challenge of deploying high-quality 3D content generation models on constrained devices, but also improves the diversity and fidelity of generated images. However, further optimization of the model to increase generation speed and Efficiency remains our main challenge."

1. **PSNR (Peak Signal-to-Noise Ratio)**：PSNR is a common indicator to measure the quality of image reconstruction. The higher the value, the smaller the difference between the reconstructed image and the original image, and the better the quality.
2. **SSIM (Structural Similarity Index Measure)**：SSIM is a measure of the similarity of two images, taking into account factors such as brightness, contrast, and structure. The value ranges from 0 to 1. The closer the value is to 1, the more similar the images are.
3. **LPIPS (Learned Perceptual Image Patch Similarity)**：LPIPS is an image similarity evaluation index based on deep learning, which is used to measure the perceptual difference between images. The lower the value, the smaller the perceptual difference.

Tum Dataset, pruning method detail

Pruning method: Structure pruning or sparse pruning

Transfer learning

Training nn in big dataset, then make pruning (try sparse pruning first)

Then apply transfer learning when we take new input