

WorldBuilder

Abstract

Recent developments in the field of Computer Vision have enabled improvements in 2D and 3D graphics processing. These advancements have sparked the transition from 2D to 3D development. Processes such as NeRF, started the automation from 2D images to 3D meshes, which are used across multiple industries. This paper proposes state-of-the-art architecture improvements the original NeRF work.

Intro

The NeRF paper, authored in 2020, uses only an 8 layer deep fully-connected neural network without any convolutional layers. Since then, there have been other iterations of NeRF such as PixelNeRF which utilizes a ResNet34 backbone pretrained on ImageNet for their experiments. Additionally, PixelNeRF features were preprocessed and upsampled prior to being fed into the model to fit the 512 pixel size.

While both the original NeRF and PixelNeRF produce adequate results, the results from the techniques can likely be improved using modern-day architectures. This paper documents an ablation study using modern deep learning architectures to improve the quality of the 3d images space generated from the original NeRF process. We supplement the 8 layer deep neural network with a suite of other depths, sizes, and activation function techniques.

To compare the performance of our proposed architectures vs the current techniques, we will analyze the performance on the widely used metrics of PSNR (Peak Signal to Noise Ratio), SSIN (Structural Similarity Index measure), and LPIPS (Learned Perceptual Image Patch Similarity). For both PSNR and SSIN, a higher value indicates better model performance. in LPIPS, a lower value indicates better performance.

Deliverables

- ☐ NeRF Architecture Training Pipeline
- ☐ 5 NeRF Architecture Variations
- ☐ Photo/Video showing the qualitative comparisons
- ☐ Final Report in IEEE Format
- ☐ Link to Git Repo

Authors

- Nicholas Akin nma1810@mavs.uta.edu
- Angel Solis ais8610@mavs.uta.edu

References

- NeRF Paper: <https://arxiv.org/abs/2003.08934>
- One-2-3-45++ <https://arxiv.org/abs/2311.07885>
- PixelNeRF: <https://arxiv.org/pdf/2012.02190>
- Nerf Data: <https://drive.google.com/drive/folders/1cK3UDIJqKAAM7zyrRYVFJ0BRMgrwhh4>
- LIDAR RORAIMA Parime <https://www.kaggle.com/datasets/rogriofmeireles/lidar-roraima-parime-research>
- From Images to 3d Shapes FI3s <https://www.kaggle.com/datasets/lehomme/from-images-to-3d-shapesfi3s>
- Tanks and Temple <https://www.kaggle.com/datasets/jinnywjy/tanks-and-temple-m60-colmap-preprocessed>
- Prosopo: A real face dataset for 3D reconstruction <https://www.kaggle.com/datasets/cantonioupao/prosopo-a-face-dataset-for-3d-reconstruction>