FS-NeRF: Fast Sparse Input Neural Radiance Field⚡

Group23: 吳中赫、林晉暘、嚴士函

I. NeRF with Sparse Inputs

- NeRF [1] typically requires hundreds of images per scene.
- Produces severe distortions when trained with few images.
- Cause: Under-constrained volume rendering equations.

NeRF with dense input views

NeRF with sparse input views

II. Prior Work on Few-Shot NeRF

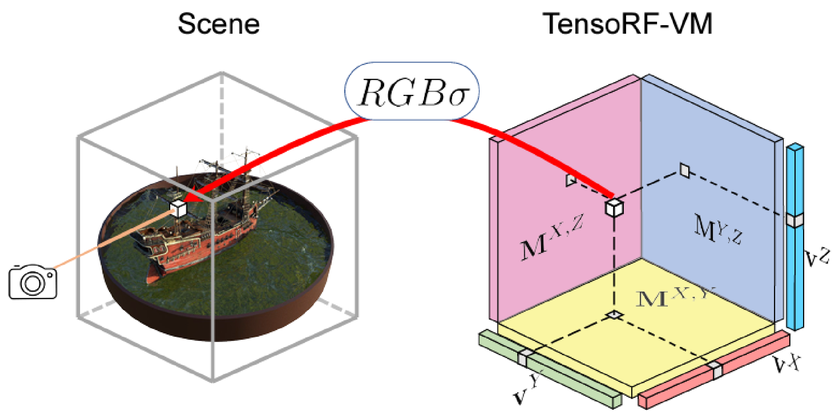
ViP-NeRF[2]:

- Introduce visibility regularization to train the NeRF with sparse input views.
- Estimate the dense visibility prior reliably using plane sweep volumes (PSV).
- Reformulate the NeRF MLP to output visibility thereby significantly reducing the training time
- The prior is dense and reliable, but it is slow to train the model.

-We need a fast training NeRF architecture.

III. Voxel Grid Representation

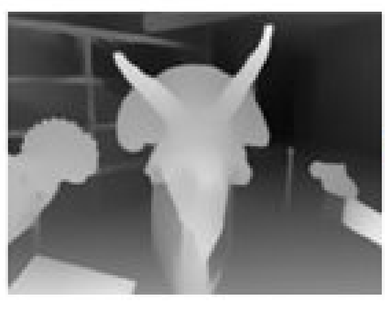
- Unlike ViP-NeRF use pure MLPs to represent the scene, we consider the full volume field as a 4D tensor, which is inspired by TensoRF[3].
- This can achieve fast training and smaller model size tham ViP-NeRF.



IV. Monocular Depth Regularization

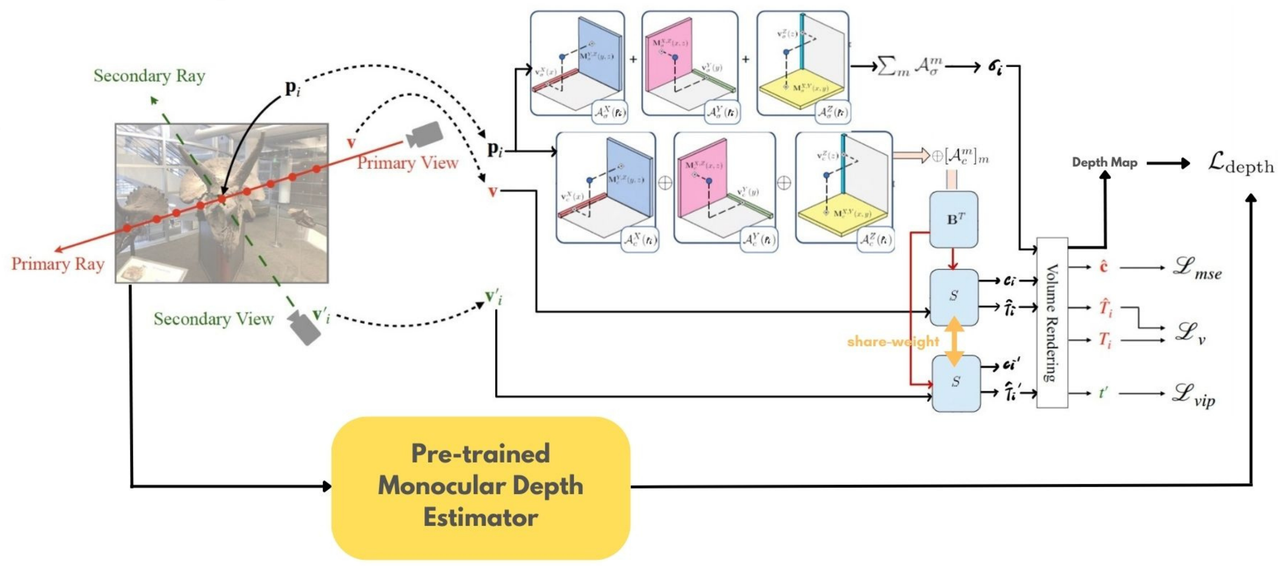
- Besides visibility regulariztion, we further add monocular depth as a supervision of the NeRF.
- Simply calculate the loss between depth map estmate by a pretrained model with the depth map we predicted from NeRF.

RGB



Monocular depth

V. Pipeline



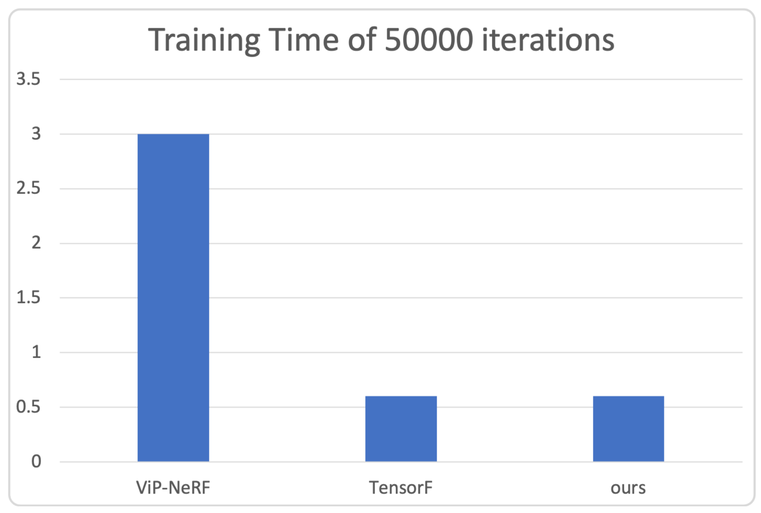
VI. Results

Given only 2 input views from LLFF dataset:

Input view | Ours novel view | TensoRF novel view

Training time

experiment on RTX4090



VII. Conclusions

- Dramatically reduce the training time by replacing the MLP in ViP-NeRF with Voxels Grids.
- Further add monocular depth as a supervision of the NeRF.
- Our work achieve state-of-art quality of NeRF with sparse input views.

VIII. Reference

[1] Mildenhall et al., “Representing Scenes as Neural Radiance Fields for View Synthesis”, ECCV 2020.
[2] Nagabhushan Somraj et al., “ViP-NeRF: Visibility Prior for Sparse Input Neural Radiance Fields”
[3] Anpei Chen et al., “TensoRF: Tensorial Radiance Fields ”, ECCV 2022.