3D Reconstruction Framework for Development of Robots and IoT for Precision Agriculture

Brian Park, Shivam Kumar Panda, Andrew Choi, Ph.D., Khalid M. Jawed, Ph.D., Jungseock Joo, PhD., Sriram Narasimhan, PhD.
Structures-Computer Interaction Laboratory, Mechanical and Aerospace Engineering Department, UC Los Angeles, CA 90095

Introduction

- In agricultural robotics, there is no complete testing framework for robots and sensors used in precision agriculture
- It is expensive to test our physical robot in the outside world, especially in its earlier stages of development
- Therefore, we require a virtual testing environment that includes a 3D, high-resolution rendering of crops, weeds, and soil to simulate our robot's sensors and analyze its behavior
- In order to optimize the quality of these 3D renderings, we will develop an active learning model that is trained to discover the view angles of the object that require the most additional information.

Objectives

1. Optimize our current Unity sampling pipeline to construct higher quality NeRFs.
2. Run COLMAP Ablation Experiments to verify that our sampling groundtruth method produces results in equal quality with COLMAP.

Methodology

Objective 1: Unity Sampling Pipeline Improvements

Originally, our Unity Sampling pipeline rotated around the simulated 3D Unity object in a single, circular ring. While this was effective, it did not capture a variety of view angles, and the camera's movement path was very unrealistic.

To remedy this, we changed the Unity camera movement path to sample more images and make two circular rings. Additionally, throughout the sampling phase, we introduce positional noise, which ensures that the resulting camera poses look more natural and not completely uniform.

We also added more natural shadows from the Unity artificial light. We added walls with high fidelity textures in order to ensure that our sampled images do not contain infinite space. From previous experiments, we have found that the presence of infinite space causes Instant-NGP to fail.

Objective 2: COLMAP Ablation Experiments

Our goal with the COLMAP Ablation Experiments is to ensure that the resulting NeRFs of COLMAP and our groundtruth method are identical.

In order to generate a NeRF using Instant-NGP, we need a folder of images and a corresponding transforms.json file, which contains information about the camera's physical parameters and each frame's transformation matrix.

Therefore, we first run our Unity program to obtain the sampled images. We can run those images through COLMAP to generate the transforms.json file. For our groundtruth method, we can run the sampled images through the Python script get_transforms.py.

Once we have both transforms.json files ready, we can run the Instant-NGP script and compare results.

Results: Unity Sampling Pipeline

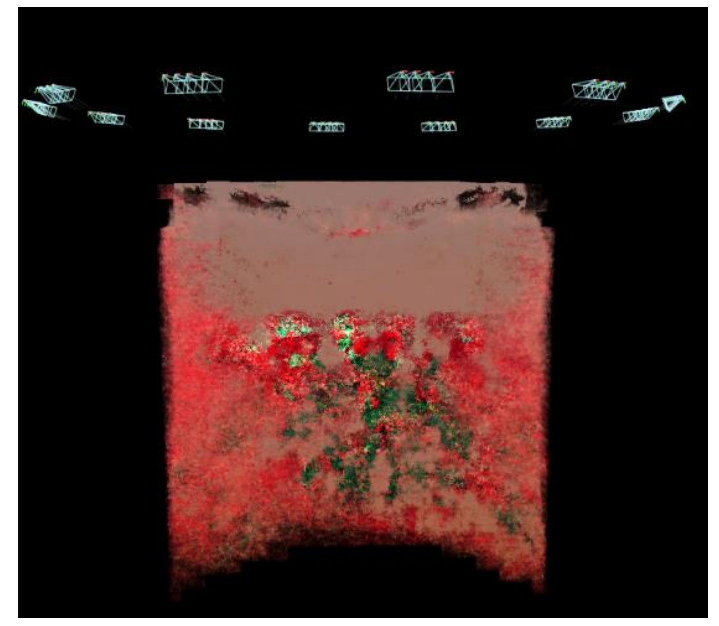

Figure 1: Original Camera Sampling Movement

Figure 2: New Camera Sampling Movement

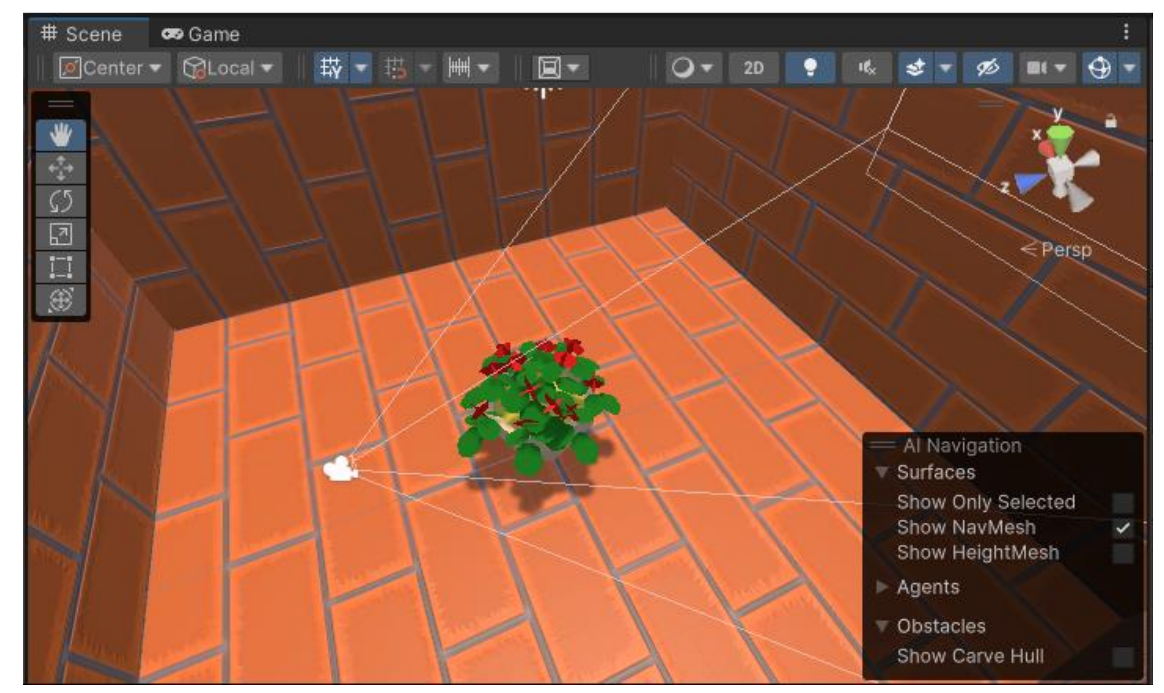

Figure 3: New Unity Environment

Results: COLMAP Ablation Experiments

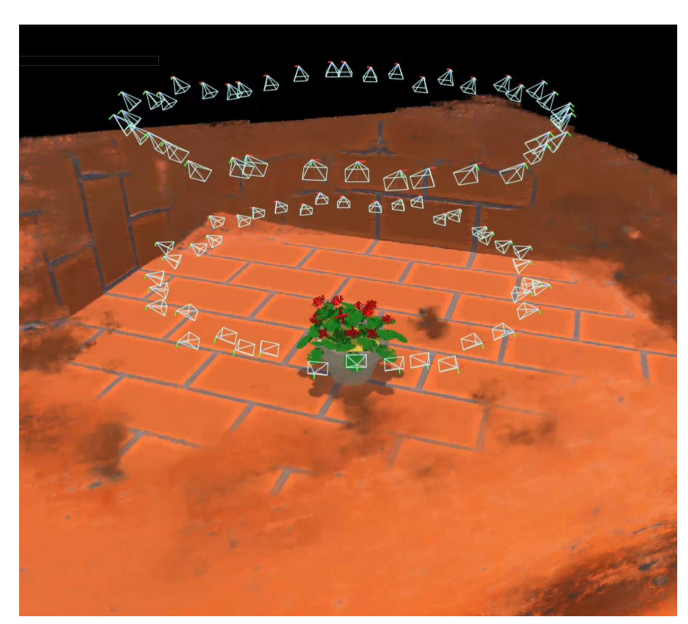

Figure 4: COLMAP NeRF (80 images)

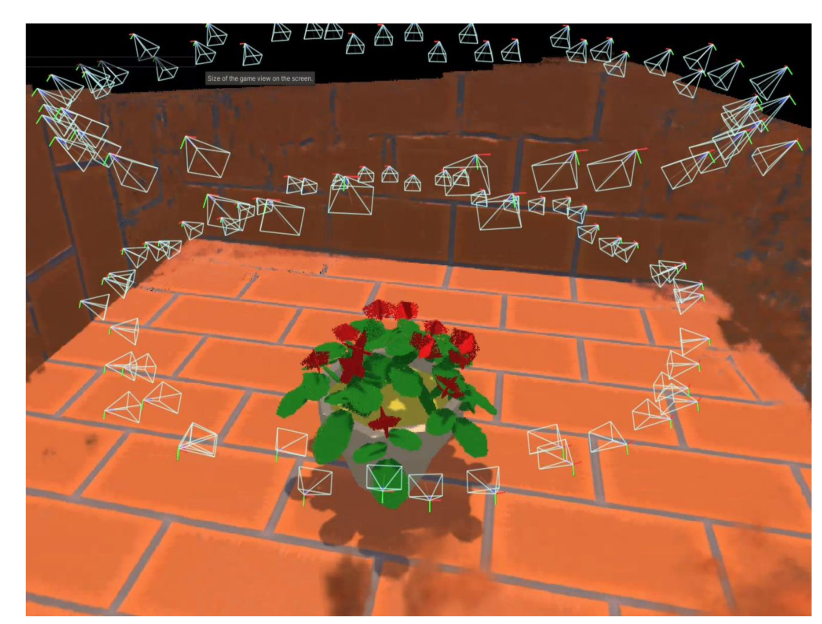

Figure 5: COLMAP NeRF (80 images)

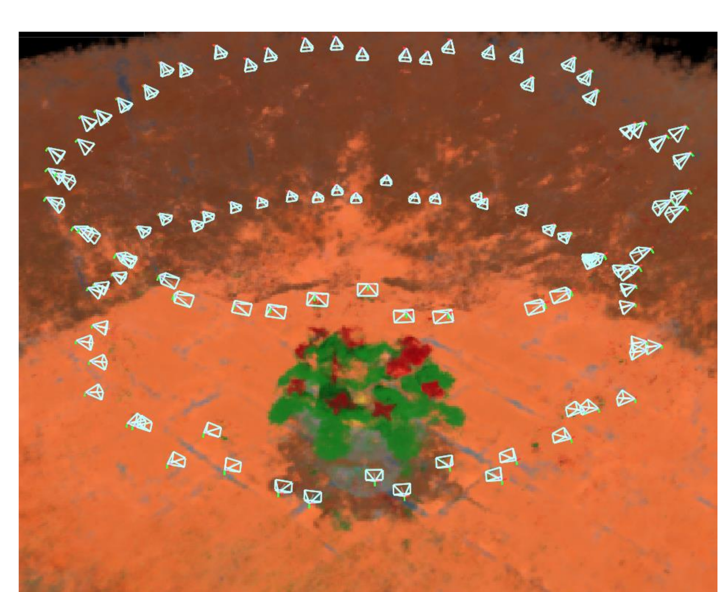

Figure 6: GT NeRF with Unity camera parameters (100 images)

Figure 7: GT NeRF with COLMAP camera parameters (100 images)

Conclusion

Objective 1: Unity Sampling Pipeline

Figure 1 shows the resulting NeRF and its camera poses using the original camera sampling movement. Figure 2 shows the resulting NeRF and its camera poses using the new camera sampling movement.

It is clear that the addition of high fidelity walls, more densely sampled images, positional noise, and multiple rings has significant increased the quality of our NeRFs. This aligns with our expectations, since the more various view angles we have, the more information we are giving Instant-NGP.

Objective 2: COLMAP Ablation Experiments

As shown in Figure 6, our GT method produces a lower quality NeRF, compared to Figure 4 and Figure 5.

Instant-NGP requires feature extraction of the camera's extrinsic matrix and the camera's intrinsic matrix. After investigation, it was found that our GT method was calculating the camera's intrinsic matrix incorrectly. Therefore, we used the intrinsic matrix that COLMAP calculated and incorporated it into our GT transforms.json.

The resulting Figure 7 shows that our GT method is able to produce a NeRF with the same quality as the NeRF that COLMAP generates. Since the camera's intrinsic matrix values should not change through iterations, we can use hard-coded estimate values for the GT method.

Future Works

1. Now that we verified that our sampling pipeline produces accurate results, we will develop a Python module that integrates the sampling pipeline with the active learning model. This Python module will be controlled by the active learning model. It will be able to load in objects from the ShapeNet dataset, sample images from the given viewpoints, and train an Instant-NGP model.

Literature Cited

Pan, Xuran., Lai, Zihang., Song, Shiji., Huang, Gao., "ActiveNERF: Learning Where To See With Uncertainty Estimation.", EECV 2022.
Mueller, Thomas., Evans, Alex., Scied Christopher., Keller, Alexander., "Instant Neural Graphics Primitives with a Multiresolution Hash Encoding"., SISGRAPH 2022.

Acknowledgements

A special thanks to Professor Khalid Jawed, Andrew Choi, Shivam Kumar Panda, and the Structures-Computer Interaction Laboratory at the University of California, Los Angeles.

This project was sponsored by the NSF REU program (Award No. 1925360).

NSF

UCLA

STRUCTURES-COMPUTER INTERACTION LABORATORY