3D Reconstruction Using Multi-View Imagery

Dongwook Lee Joel Markapudi  
lee.dongw@northeastern.edu markapudi.j@northeastern.edu

**1. Overview of Project Idea**

This project aims to address the challenge of creating accurate 3D reconstructions from 2D images using the ETH3D dataset. The dataset provides a wide range of multi-view images and corresponding 3D ground truth, making it an ideal testing ground for advanced feature extraction and depth estimation techniques. Our objective is to generate high-quality 3D models by utilizing recent methods like multi-view stereo (MVS).

We are particularly focused on improving the precision of 3D reconstructions in outdoor environments, which present unique obstacles such as occlusions, variable lighting, and irregular structures. By refining existing methods and introducing new optimizations, we aim to explore what can be achieved in 3D modeling from 2D images. The demonstrated success of techniques like MVS and Poisson surface reconstruction has inspired us to explore and deepen our understanding of solving reconstruction problems.

This project also opens up possibilities for a wide range of real-world applications. Accurate 3D reconstructions are vital in fields like autonomous driving, where detailed environmental maps are essential, and in augmented reality, for immersive experiences. Beyond these, our work has potential benefits in areas like robotics, urban planning, scene modeling, and even healthcare, where 3D models can support decision-making and analysis.

**2. Literature Survey**

From the literature we've explored, many key algorithms stand out for experimentation. RapidVol[3] and Trackerless 3D Freehand Ultrasound[4] introduce rapid reconstruction methods, offering faster volumetric results for large datasets. The Mesh-Guided Multi-View Stereo[2] employs pyramid scaling for depth refinement, improving occlusion handling. BlendedMVS[5] focuses on maintaining lighting and texture through blended image generation. The LEGO Models paper[1] introduces space carving and visual hull construction, offering alternatives for specific scenes. Some of these approaches may enhance parts of our methodology. They guide us in improving depth estimation, model consistency, and overall reconstructions. We also found that many techniques for symmetric cuboids or medical scans do not suit feature-rich outdoor or scenic data.

**3. Dataset**

We will be utilizing the **ETH3D High-Resolution Multi-View Dataset,** which includes high-quality images, calibration data, and 3D ground truth such as point clouds and depth maps. It provides undistorted images that have undergone lens correction, as well as distorted and RAW images. The dataset includes occlusion files that mark areas blocked from view, aiding depth estimation. It also has calibration data, such as intrinsic (focal length, principal point) and extrinsic (camera orientation and position) parameters.

**4.** **Plan of Activities**

Initially, We plan on preprocessing and implementation steps feature extraction and depth estimation techniques. By the mid-point, we aim to complete multi-view stereo (MVS) and surface reconstruction. For the final phase, we will focus on evaluation and validation against ground truth data, using qualitative inspection and quantitative metrics. We then finalize the results and compile the report. We aim to work collaboratively, with both team members contributing to all aspects of the project, rather than splitting tasks. We'll share progress in real time, following an agile methodology, rapid iterations, and continuous integration.

**References**

1. **Reconstructing Creative LEGO Models**
2. **Mesh-Guided Multi-View Stereo with Pyramid Architecture.** [ <https://ieeexplore.ieee.org/document/9156553> ]  
   Y. Wang, T. Guan, Z. Chen, Y. Luo, K. Luo and L. Ju, "Mesh-Guided Multi-View Stereo With Pyramid Architecture," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 2036-2045, doi: 10.1109/CVPR42600.2020.00211.
3. **RapidVol: Rapid Reconstruction of 3D Ultrasound Volumes from Sensorless 2D Scans**
4. **Trackerless 3D Freehand Ultrasound Reconstruction: A Review**
5. **BlendedMVS: A Large-scale Dataset for Generalized Multi-View Stereo Networks**  [ <https://arxiv.org/abs/1911.10127> ]  
   BlendedMVS: A Large-scale Dataset for Generalized Multi-View Stereo NetworksYao, Y., Luo, Z., Li, S., Zhang, J., Ren, Y., Zhou, L., Fang, T., & Quan, L. (2019). BlendedMVS: A Large-scale Dataset for Generalized Multi-view Stereo Networks. ArXiv. https://arxiv.org/abs/1911.10127