EXPLORING TYCHO CRATER: 3D RECONSTRUCTION WITH NEURAL RADIANCE FIELDS FROM SPARSE VIEWS.

Suwan Lee, Seokju Lee (slee@kentech.ac.kr), Korea Institute of Energy Technology (KENTECH), South Korea.

Introduction: Recently, learning-based neural rendering techniques using *artificial intelligence* have achieved remarkable advancements in various computer vision and graphics tasks such as view synthesis, 3D reconstruction, pose estimation, re-lighting, *etc*. In this study, we perform an experiment of 3D reconstruction based on neural rendering[1] for lunar terrain imagery.

As our target area, we have selected the Tycho crater (43.37°S, 348.68°E), which already possesses high-resolution 3D terrain data obtained through LRO LOLA and NAC observations. Notably, the central peak of the crater rises approximately 2 km above the crater floor, spanning about 15 km in width. This region is relatively young in lunar standards, estimated to be around 110 million years old. Consequently, it has maintained its sharp and steep appearance with minimal erosion, making it an ideal candidate for detecting visually prominent features and incorporating them into various matching algorithms for 3D reconstruction.

Objective: To the best of our knowledge, this study marks the pioneering application of neural rendering techniques in reconstructing 3D lunar terrain, with the goal of validating the effectiveness of data-driven learning-based algorithms. Specifically, the objective is to compare the reconstruction performance across various quantities of training data under sparse viewpoint scenarios.

Dataset: Our experiment aims to assess the degradation in the performance of neural networks under different sparse view scenarios on the same lunar terrain. To implement these specific experimental conditions, we created a photorealistic dataset utilizing the LROC QuickMap [1]. Initially, we visualize the target data[2] [2] in the 3D interface and export images from randomly dense viewpoints. We capture a total of 455 images, after which we perform data preprocessing by applying structure-from-motion and bundle adjustment algorithms using COLMAP [3]. Figure 1 shows our region of interest, illustrating the optimized camera pose results obtained from the collected images. Once the complete dataset is constructed, we randomly sample adjacent sparse views (*e.g.*, 2, 3, …, 50 views) and feed them as training data for the neural rendering network.

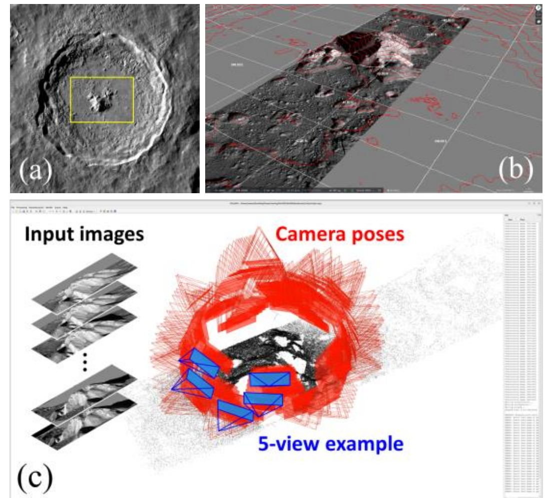

Figure 1. (a) A LROC WAC image of Tycho Crater. Yellow box is location of the central peak, which is our modeling target. (b) A visualization of LROC NAC ROI Mosaics on QuickMap 3D interface. (c) Dataset preprocessing using COLMAP. Sparsely sampled viewpoints for training are represented in blue as an example, while all viewpoints are depicted in red.

Neural Rendering for 3D Reconstruction: The quote, *"What I cannot create, I do not understand"* by Richard Feynman underscores the connection between the act of creation and the comprehension of complex concept. In the realm of 3D computer graphics, neural rendering achieves an understanding of the 3D scene by training on feature representations from multi-view RGB images. It synthesizes realistic novel-view images with high-fidelity rendering, leveraging both neural networks and traditional computer graphics methods.

The pioneer deep learning framework of novel view synthesis [4,5] improve the feature representation of neural network to generate the novel view with the geometrical pixel relations. Volumetric scene representation is applied to view synthesis [6], leveraging the differentiable density fields to exploit geometric constraints. Neural Radiance Fields (NeRF) [7] introduce view-dependent radiance fields with positional encodings and hierarchical ray-marching sampling, rendering high-fidelity generated images.

---

[1] The motivation behind employing learning-based rendering techniques from the field of graphics, as opposed to traditional methods such as *multi-view stereo* and *structure-from-motion* in computer vision, is to facilitate the development of more intricate photometric and optical modeling of the lunar surface in forthcoming research.

[2] Data name: NAC_ROI_TYCHOFLRLOA_E430S3486.

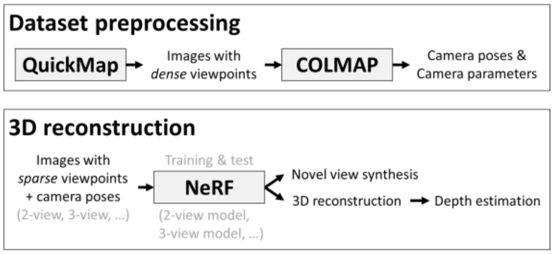

Figure 2. Overall pipeline of our experiments: dataset preprocessing and 3D reconstruction tasks.

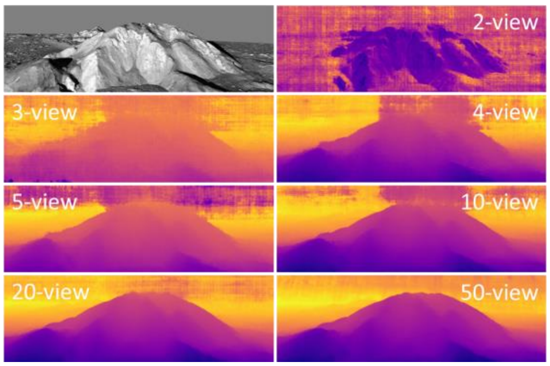

Figure 3. Comparison of depth estimation results from models trained with different numbers of training viewpoints (visualized using the 'plasma' colormap in the Python package).

The rendering weights obtained from the volumetric representation and sampling z-value can be calculated to depth as summarized in Figure 2. Intuitively, regions with higher rendering weight correspond to object surface area, while lower rendering weights regions correspond to vast area. Finally, we rescale the obtained depth values to the original scale of optimized poses from COLMAP using least-squares. Note that this rescaling is performed on the training samples to ensure a fair comparison.

Results and Discussion: We validate the performance of 3D reconstruction with depth estimation by rendering the depth map at the same view direction using models trained with different numbers of viewpoints (2-, 3-, 4-, 5-, 10-, 20-, 50-view). Note that models trained with a larger number of viewpoints include the viewpoints learned by models with fewer viewpoints.

Figure 3 illustrates that as the number of viewpoints decreases, rendering weights become broader in their representation, leading to a more uniform distribution. This indicates that the model struggles to capture sufficient radiance variations. As a result, the model fails to accurately discern the overall shape and boundaries of the target scene. However, we have also confirmed that performance can be quickly restored with a relatively small increase in the number of viewpoints (≥ 5-view).

Future Works: Sparse view scenarios pose an inherent challenge in remote sensing. In our upcoming research, we aim to introduce neural rendering techniques capable of achieving enhanced precision in 3D reconstruction within these limited conditions. Based on observations from the Korea Pathfinder Lunar Orbiter (KPLO)'s Lunar Terrain Imager (LUTI) and Wide-angle Polarimetric Camera (PolCam), we expect the sophisticated design of a learning-based optical modeling for the lunar surface.

Acknowledgments: We would like to extend our sincere gratitude to the Korea Aerospace Research Institute (KARI) for their invaluable advice and comments throughout this research endeavor. This research is supported by the Korea Astronomy and Space Science Institute (KASI).

References:

[1] https://quickmap.lroc.asu.edu

[2] S. M. Klem, et al., "Controlled LROC Narrow Angle Camera High Resolution Mosaics." 45th Annual Lunar and Planetary Science Conference, 2014.

[3] J. L. Schonberger, et al., "Structure-from-Motion Revisited." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[4] T. Zhou, et al., "View Synthesis by Appearance Flow." European Conference on Computer Vision (ECCV), 2016.

[5] J. Flynn, et al., "DeepStereo: Learning to Predict New Views from the World's Imagery." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[6] V. Sitzmann, et al., "DeepVoxels: Learning Persistent 3D Feature Embeddings." IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

[7] B. Mildenhall, et al., "NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis." European Conference on Computer Vision (ECCV), 2020.