IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10641392

THE APPLICATION OF NEURAL RADIANCE FIELDS (NERF) IN GENERATING DIGITAL SURFACE MODELS FROM UAV IMAGERY

D. Abate[1], K. Themistocleous[1], D. Hadjimitsis[1]

[1]Eratosthenes Centre of Excellence, Limassol, Cyprus

ABSTRACT

Neural Radiance Fields (NeRFs) is emerging as an innovative approach for generating in the reality-based 3D modelling techniques by employing an artificial neural network that optimizes a volumetric scene function based on a collection of input views [1, 2, 3]. This approach offers unprecedented prospects for multiscale 3D modelling and analysis.

This study delves into the application of NeRF technology for processing aerial-born imagery [4] to evaluate its effectiveness for the creation of Digital Surface Models (DSM) compared with traditional photogrammetric techniques.

Index Terms— NeRF, Photogrammetry, Digital Surface Model, UAV

1. INTRODUCTION

NeRF technology has been recently extensively utilized by the scientific community for a diverse range of applications across various domains. In the field of heritage conservation and archaeology, researchers have employed NERF technology for the detailed 3D reconstruction of cultural heritage sites and artifacts [5, 6].

In medical science, NERF technology has shown promising results in creating detailed, personalized semantic facial models from monocular videos [7].

The technology has also made significant strides in handling complex visual challenges such as reconstructing transparent objects [8].

For aerial topographical surveys, NERF has been adapted to work efficiently with drones, enabling large-scale, high-fidelity 3D scene reconstruction [9].

Furthermore, NERF's application extends to environmental monitoring [10], where it contributes to understanding and documenting environmental changes, aiding in conservation efforts. Additionally, in the realm of virtual reality, NERF technology has been instrumental in creating immersive 3D virtual environments [11], significantly enhancing the user experience in virtual reality applications.

Through these varied applications, NERF technology demonstrates its versatility and potential to revolutionize numerous fields by providing high-precision, efficient, and detailed 3D modeling capabilities.

This paper focuses on evaluating the geometric accuracy of a model derived from NeRF in comparison to a conventional photogrammetric workflow.

NeRF attempts to understand how light travels through the scene and how it is reflected off objects. It guesses how much light is coming from different parts of the scene and how it changes as it is observed from different angles.

As the network gets a better understanding on how the light moves around in the images, NeRF starts to build a detailed 3D model of the scene.

The dataset employed in this study originates from an aerial survey conducted using a Real Time Kinematic (RTK) Unmanned Aerial Vehicle (UAV) over a rural area.

According to the flight altitude, the camera optics and the sensor size, the expected Ground Sampling Distance (GSD) has been calculated in 1 cm.

The selected case study area is primarily characterized by its flat morphology, with the notable exception of a slope located on one of its sides. This landscape is further distinguished by the presence of scattered bushes and trees, which add to its distinct topographical features.

This comparison aims to determine the effectiveness and precision of NeRF technology in capturing and reconstructing rural landscapes, particularly when compared to established photogrammetric methods.

2. NEURAL RADIANCE FIELDS

Neural Radiance Fields (NeRF) is a machine learning framework that has made significant strides in synthesizing highly realistic 3D scenes from 2D images. At its core, NeRF utilizes a fully connected deep neural network to model the volumetric scene function.

The neural network receives a unified 5D vector encompassing spatial coordinates (x, y, z) and viewing angles (θ, f) as input, and it generates the corresponding volume density (s) and direction-dependent emitted radiance (RGB) for each point and direction.

Starting from the camera poses, this process generates a continuous, high-resolution volumetric scene representation.

By intelligently sampling along camera rays, NeRF can predict the color of light traveling along them, effectively producing novel view synthesis with high precision. The approach is distinguished by its capacity to capture intricate details and complex lighting conditions, rendering images that can be virtually explored from any perspective.

Leveraging Neural Radiance Fields, it is then possible to extract 3D point coordinates, offering a cutting-edge approach to reconstructing intricate spatial geometries from a series of 2D images.

3. METHODOLOGY

The methodology employed for generating Digital Surface Models (DSM) from aerial imagery is concisely outlined in Figure 1 and elaborated upon in the subsequent sections.

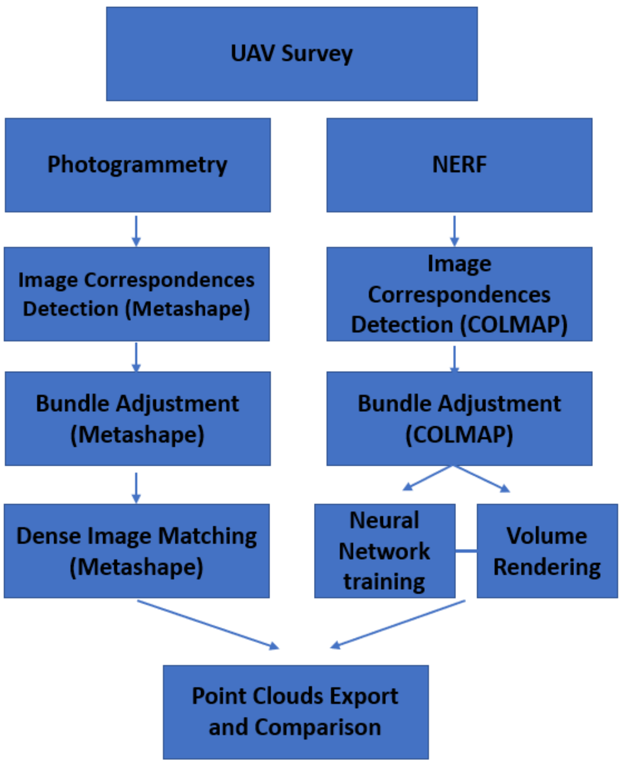

Figure 1. Methodological Workflow

Initially, the image dataset was processed using Agisoft Metashape [13], where the traditional photogrammetric workflow was employed. This workflow is composed of three main stages: (i) detection of image correspondences, where the software identifies similar points across multiple images to establish a relationship between them; (ii) bundle adjustment, a process that refines the camera positions and 3D point coordinates to minimize reprojection errors and enhance the accuracy of the model; and (iii) dense image matching (Figure 2), which involves creating a dense point cloud by closely matching numerous points across the images [12].

Figure 2. Photogrammetric dense point cloud

Similarly, the same dataset underwent processing using NeRF technology, following a distinct but equally systematic workflow. Nerfstudio was used [14] (Figure 3).

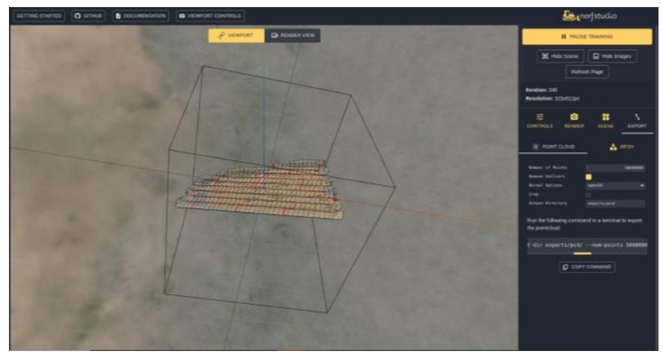

Figure 3. Nerfstudio environment

The NeRF processing can be outlined in several key steps: (i) detection of image correspondences and bundle adjustment using COLMAP [15]; (ii) neural network training, during which a deep learning model learns to represent the 3D scene as a continuous volumetric field, using the input images; and (iii) volume rendering, where the trained NeRF model synthesizes novel views of the scene by rendering the continuous volumetric scene representation (Figure 4).

Figure 4. NeRF dense point cloud

4. DATA ANALYSIS

The two resultant dense point clouds were finally exported and then processed using the 'Cloud-to-Cloud Distance' plugin available in CloudCompare. This plugin facilitates the computation of spatial differences between the two-point clouds by measuring the distance from each point in one cloud to the nearest point in the other cloud. This procedure was used to assess and quantify the discrepancies between the point clouds. [16].

In the analysis, it was observed that the model generated NeRF technology exhibited an apparent barrel distortion. This distortion manifested as a bulging effect, particularly evident in the central region of the reconstructed DSM, where it appeared to curve outward in a range between 30 cm and 1 meter (Figures 5 and 6).

The observed distortion in the NeRF model is attributable to the results of the bundle adjustment process in COLMAP, leading to an unstable camera network. This instability is potentially exacerbated by the flatness of the terrain, which might not provide sufficient variation in features for accurate camera positioning. In contrast, the use of data in Agisoft Metashape, which takes in consideration the RTK corrections data embedded in the images Exif files, helps avoid such distortion. thus, mitigating the issues observed in the NeRF model.

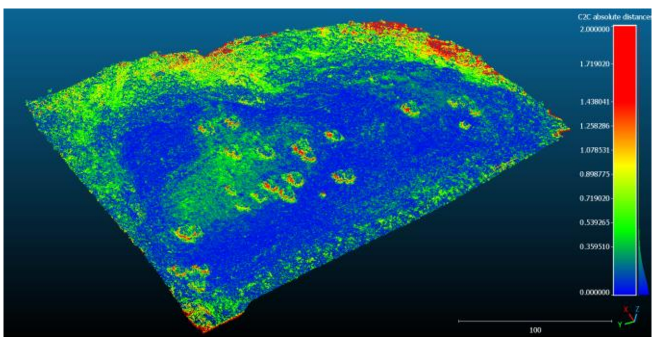

Figure 5. NeRF model Cloud-to-Cloud Distance (1)

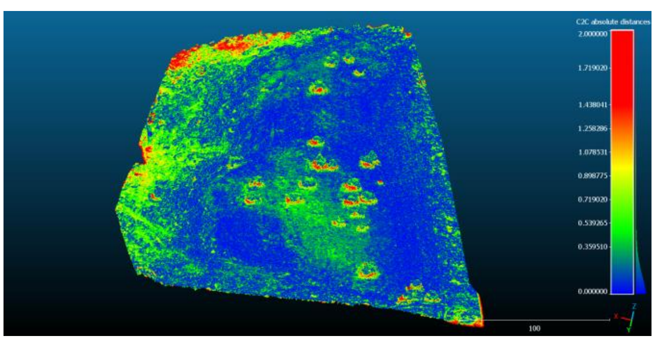

Figure 6. NeRF model Cloud-to-Cloud Distance (2)

This observed effect was further corroborated by employing a reference plane and calculating the distances between each point cloud and the plane. This method provided additional confirmation and a more quantifiable measure of the effect, enhancing the reliability of the preliminary assumptions. (Figure 7 and 8).

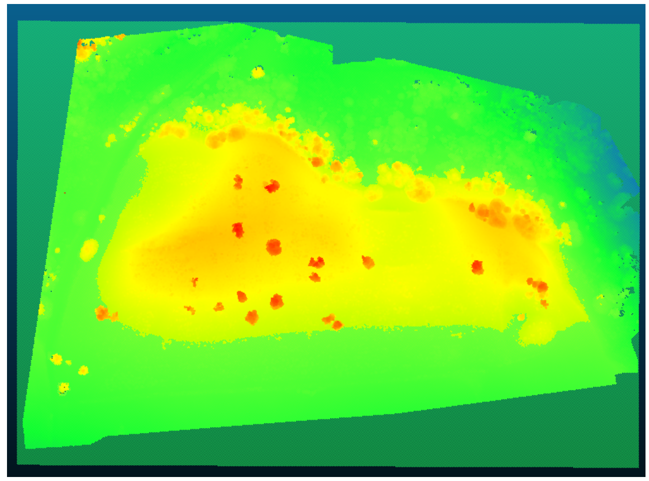
Figure 7. Best fitting plane Photogrammetric model

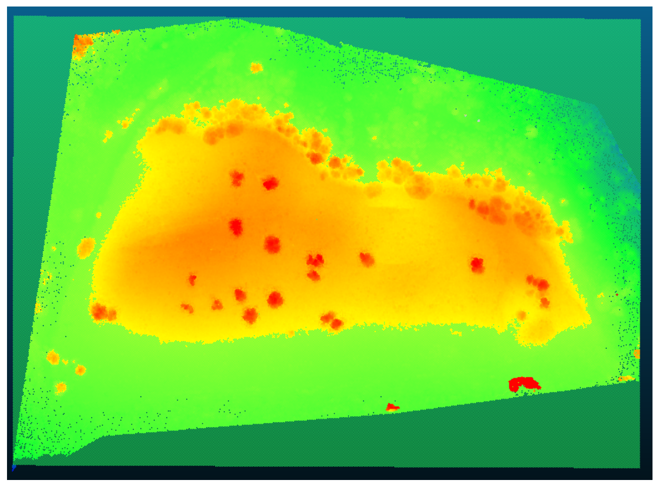
Figure 8. Best fitting plane NeRF model

Despite the distortion, it is noteworthy that NeRF demonstrated a remarkable capability in reconstructing a significant amount of detail, particularly for complex features such as trees and low vegetation, where traditional photogrammetric methods often struggle. Such detailed reconstruction is challenging for standard photogrammetry due to the complex geometry and variable lighting conditions typically found in natural environments. Given the robustness of the photogrammetric camera network, the ability of NeRF to accurately render these detailed features underscores its potential as a powerful tool for detailed and realistic 3D modeling, which can be used in several research domain (i.e., agriculture and forestry) (Figure 9 and Figure 10).

Figure 9. Photogrammetric point cloud, vegetation

Figure 10. NeRF point cloud, vegetation

5. CONCLUSIONS

This paper presented an evaluation study comparing the output of traditional photogrammetric techniques with the Neural Radiance Fields (NeRF) approach for 3D reconstruction and generation of Digital Surface Models. The study utilized a dataset from an aerial survey of a rural environment, processed through both methods. Although the NeRF model displayed significant barrel distortion, this is. Due to the image alignment step conducted in COLMAP. Despite this, the NeRF model impressively captured intricate details, excelling particularly in complex natural features like trees, an area where traditional photogrammetry often encounters difficulties.

Integrating NeRF with traditional photogrammetric methods could leverage the strengths of both methods, potentially leading to more accurate and detailed reconstructions. This hybrid approach could provide a powerful solution for diverse applications in 3D modeling and environmental surveying.

6. REFERENCES

[1] R. Martin-Brualla, N., Radwan, M.S. Sajjadi, J.T. Barron, A. Dosovitskiy, D. Duckworth, 2021. Nerf in the wild: Neural radiance fields for unconstrained photo collections. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 7210–7219.

[2] B. Mildenhall, P.P Srinivasan, M. Tancik, J.T/ Barron, R. Ramamoorthi, R. Ng, 2021. Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM, 65(1), 99–106.

[3] K. Gao, Y. Gao, H. He, D. Lu, L. Xu, J. Li, 2022. Nerf: Neural radiance field in 3d vision, a comprehensive review.

[4] R. Marí, G. Facciolo, and T. Ehret, 2022. Sat-nerf: Learning multi-view satellite photogrammetry with transient objects and shadow modeling using rpc cameras. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1311-1321).

[5] F. Condorelli, F. Rinaudo, F. Salvadore, S. Tagliaventi, 2021. a Comparison Between 3d Reconstruction Using Nerf Neural Networks and Mvs Algorithms on Cultural Heritage Images. The Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., 43, 565–570.

[6] E. Balloni, L. Gorgoglione, M. Paolanti, A. Mancini, and R. Pierdicca, 2023. Few Shot Photogrametry: A comparison between Nerf and Mvs-Sfm For The Documentation Of Cultural Heritage, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-M-2-2023, 155–162.

[7] X. Gao, C. Zhong, J. Xiang, Y. Hong, Y. Guo, and J. Zhang, 2022. Reconstructing personalized semantic facial nerf models from monocular video. ACM Transactions on Graphics (TOG), 41(6), pp.1-12.

[8] J. Ichnowski, Y. Avigal, J. Kerr, and K. Goldberg, 2021. Dex-NeRF: Using a neural radiance field to grasp transparent objects. arXiv preprint arXiv:2110.14217.

[9] Z. Jia, B. Wang, and C. Chen, 2023. Drone-NeRF: Efficient NeRF Based 3D Scene Reconstruction for Large-Scale Drone Survey. arXiv preprint arXiv:2308.15733.

[10] R. Singh, A. Sharma, and M. Vatsa, 2022. Forest-NeRF: Neural Radiance Fields for 3D Environmental Monitoring. Journal of Environmental Management, 300, 113816.

[11] H. Lee, J. Park, and S. Lee, 2021. VR-NeRF: Creating Immersive 3D Virtual Worlds Using Neural Radiance Fields. IEEE Transactions on Visualization and Computer Graphics, 28(5), pp.2060-2070.

[12] Remondino F, Spera MG, Nocerino E, Menna F, Nex F. State of the art in high density image matching. Photogram Rec. 2014;29(146):144–66.

[13] https://www.agisoft.com/

[14] https://docs.nerf.studio/

[15] https://colmap.github.io/index.html

[16] https://www.danielgm.net/cc/