EXPLORING PHOTOGRAMMETRY AND NEURAL RADIANCE FIELDS FOR PRESERVING BAGUIO STATUES AND MONUMENTS

Yohann C. Gayao
College of Information in Technology and Computer Science
University of the Cordilleras
Governor Pack Road Upper Session
ycg4409@students.uc-bcf.edu.ph

Bill R. Malitao
College of Information in Technology and Computer Science
University of the Cordilleras
Governor Pack Road Upper Session
brm1653@students.uc-bcf.edu.ph

Christella Marie R. Ocampo
College of Information in Technology and Computer Science
University of the Cordilleras
Governor Pack Road Upper Session
cro4664@students.uc-bcf.edu.ph

ABSTRACT

The Cordillera Administrative Region in the Philippines has a rich cultural heritage, including cultural statues and monuments. However, these artifacts are facing threats due to deterioration and inadequate preservation measures. This study aims to address this concern by comparing Photogrammetry and NeRF; two advanced digitization techniques, to accurately preserve digitally these cultural artifacts. However, more data is yet to be found regarding the comparative efficiency of these techniques in handling detailed objects like cultural artifacts. Therefore, this research aims to conduct a quantitative comparative analysis of Photogrammetry and NeRF in terms of accuracy, processing time, and resource requirements. The accuracy of the generated 3D models has been assessed by comparing them to the physical artifacts using established evaluation metrics. Processing time and resource requirements, such as computational power, have been measured as well to evaluate the efficiency of each technique. By comparing the performance of Photogrammetry and NeRF, this study aims to identify the most efficient approach for the digital preservation of CAR's cultural heritage. The researchers' hypothesis posits that while Photogrammetry might excel in ideal settings, its cost intensity could be a bottleneck. Conversely, NeRF's versatility allows it to perform well when faced with limitations while remaining cost-effective, making it a good alternative. By identifying the optimal approach for digitizing the Cordillera's cultural treasures, the researchers pave the way for developing effective preservation strategies.

Categories and Subject Descriptors

[Digital Preservation]: 3D Model Reconstruction - *geometric accuracy, ground truth, plane fitting assessment, cloud-to-cloud comparison,*
[Artificial Intelligence]:

General Terms

Machine Learning, Deep Learning,

Keywords

Photogrammetry; Neural Radiance Fields; Splatfacto; Instant-NGP; Nerfacto

1. INTRODUCTION

One of the rising concerns of our society today is cultural preservation. Cultural preservation pertains to the restoration and conservation of artifacts and historical sites. These include paintings, sculptures, documents, and locations with historical significance. Cultural preservation is a challenging duty as there are numerous issues; financial constraints, climate change, and natural disasters that can cause irreparable damage to cultural treasures. Cultural heritage sites are legacies that must be preserved.

As it is known, Numerous conservation efforts include restoring historical buildings, passing on an ancient craft, or recording traditional tales. However, These techniques are antiquated and are not impervious to the passage of time as the artifacts and landmarks are still subjected to natural elements. Factors such as deterioration, climate change, and the challenges of limited resources require innovative solutions to protect these heritage sites continuously [1].

Storing them digitally through modern techniques is starting to become a reliable addition to cultural preservation efforts [2]. Numerous cultural institutions have also begun to digitize their collections to make them more accessible and protected. Traditional methods are proven inadequate when utilized alone as they do not address the main issue, which is the inevitability of any physical artifact's eventual deterioration. AI algorithms play a crucial role in identifying damage and formulating targeted restoration plans for deteriorated artifacts and structures [1].

Applying both traditional and digital methods together is preferable to address deterioration concerns.

By leveraging principles of image analysis and computational geometry, photogrammetry plays a vital role in transforming visual data into valuable three-dimensional representations, providing crucial insights [3]. Photogrammetry encompasses methods of image measurement and interpretation to derive the shape and location of an object based on the dataset of gathered photographs [4]. As of now, innovations in the field of photogrammetry have made it much more flexible in its application and present new opportunities in the representation of monuments [5].

Neural Radiance Fields, on the other hand, represent a powerful framework that excels in capturing complex lighting effects, intricate details, and realistic textures, making it a pivotal technology in the realm of computer-generated imagery [6]. However, each scene must be optimized individually, with no knowledge shared between the images. Unnecessarily time-intensive and in the limit of single or extremely sparse views, it is unable to make use of any prior knowledge to speed up the process of model creation [7]

accommodating dynamic elements. The neural network's ability to capture complex, non-linear relationships in the data contributes to generating visually compelling 3D reconstructions, showcasing the symbiotic relationship between photogrammetry, neural radiance fields, and the transformative potential of machine learning in computer vision applications.

However, the efficiency and processing speed of the methods are yet to be tested and compared. Certain metrics such as the accuracy and image clarity in regards to exposure to different kinds of lighting have not been tested on cultural statues and monuments and in the Cordilleran region specifically.

The researchers aim to find out how the emerging capabilities of Neural Radiance Fields compare to the established strengths of photogrammetry. Through meticulous data gathering and comparison of NeRF and photogrammetry 3D models, relevant data regarding their respective outputs have been compared to one another in terms of efficiency, speed, and overall image quality.

The results have been utilized to contribute meaningfully to the ongoing dialogue surrounding digital preservation, fostering advancements in the field and supporting informed

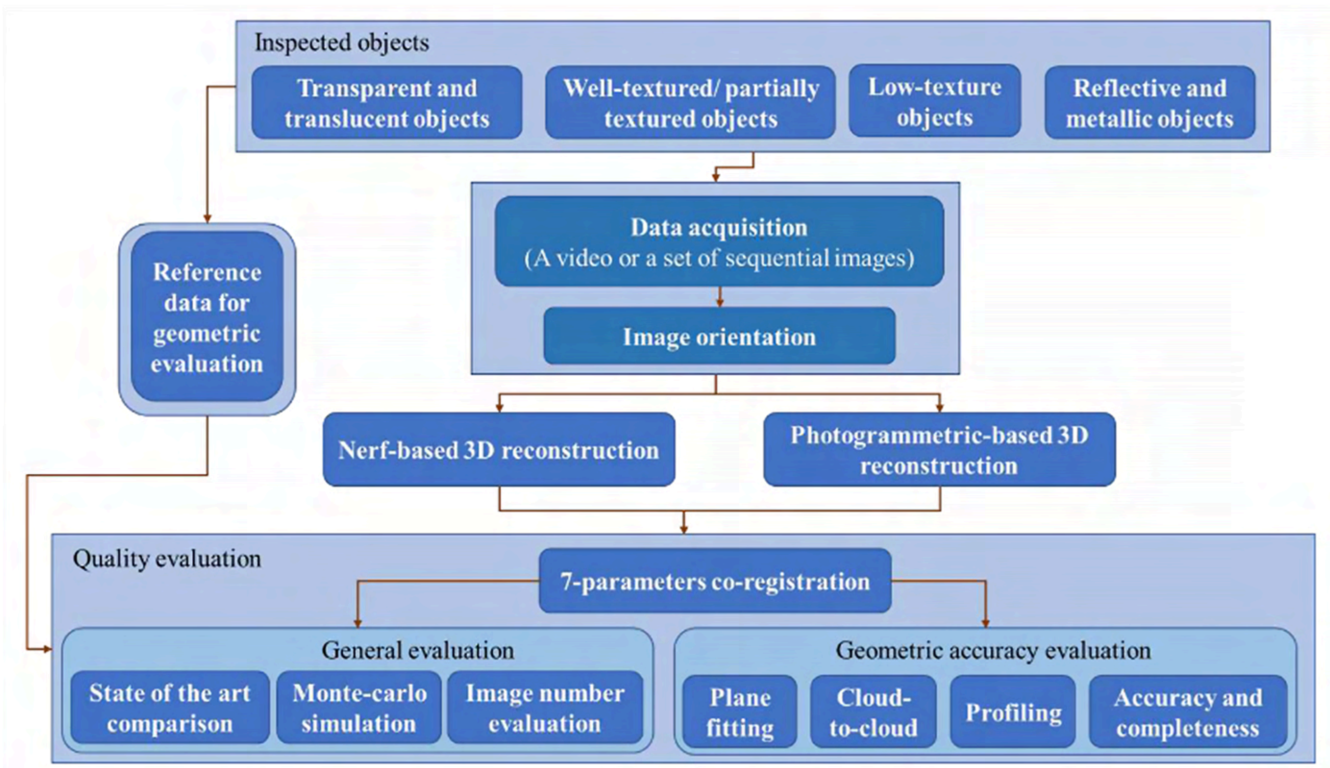

Figure 1. Overview of the proposed procedure to assess the performance of NeRF-based 3D reconstruction with respect to conventional photogrammetry [4]. Excerpt from A Critical Analysis of NeRF-based 3D Reconstruction. Remote Sensing.

The utilization of machine learning within Neural Radiance Fields is instrumental in achieving their remarkable capabilities. These networks are trained on extensive datasets containing diverse scenes and lighting conditions, enabling them to generalize and infer intricate 3D structures from 2D images. During the training process, the neural networks learn to predict the radiance or color of a point in 3D space based on its position and appearance in the input images. This learned representation allows NERF to generate novel views of a scene with unprecedented realism, even

decision-making for the conservation of cultural artifacts in urban settings.

2. METHODOLOGY

This study employs a methodical and experimental quantitative design to explore digital preservation methods for statues and monuments in Baguio City. The process begins with a well-defined research question, focusing on evaluating the current methods or environmental impact on deterioration. Following a thorough literature review, a suitable study design is chosen. Methodical studies involve interviews and observations, while experimental setups manipulate variables and utilize control groups to test preservation techniques. Representative statues are then selected for sampling.

Figure 1 employs a three-phased methodological framework. Initially, meticulous textural assessment of the objects is conducted to establish a benchmark for subsequent geometric evaluation. This data serves as a crucial reference point for the later stages. Subsequently, a set of sequential images is acquired and utilized as datasets for both NeRF-based and photogrammetric-based 3D reconstruction techniques. Finally, a comprehensive quality evaluation is performed, encompassing general and geometric accuracy assessments. Notably, the geometric accuracy evaluation leverages reference data as ground truth to ensure the validity and reliability of comparisons.

3. RESULTS AND DISCUSSIONS

Our study explored photogrammetry and Neural Radiance Fields for preserving Baguio statues and monuments. The researchers hypothesized that photogrammetry performs better in an ideal setting, but its cost intensity is an issue, and that NeRF is a viable alternative. Our findings partially supported this hypothesis. Models of NeRF showed similar quality to the models of photogrammetry while using fewer images.

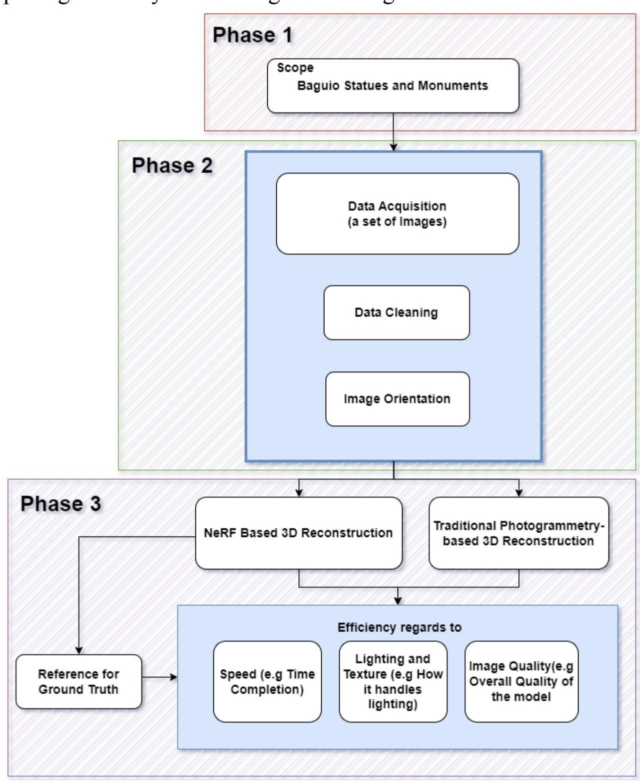

Figure 2. Model to compare NeRF to Traditional Photogrammetry in regards to efficiency, speed, lighting, and image quality

Figure 2 was constructed in reference to Figure 1. Displaying a set of phases and procedures to be conducted to properly assess and compare photogrammetry and Neural Radiance Fields digital preservation techniques. The model also employs a three-phased methodological framework. Respectively, documentation, data acquisition, and 3D model reconstruction.

Figure 3. Statue Images taken to be used as Datasets

Phase 1, in Figure 3, the statues and monuments in Baguio were found and documented. Each statue or monument is classified based on the object's texture/material used (e.g. metal, cement, wood, glass, or mixed). Detailed documentation included notes on material composition, and surrounding lighting conditions at multiple times of the day. A table containing the statues' primary data can be observed below.

Name	Lighting	Material	Surface Finish
Crucifix Statue	Harsh Directional Lighting	Painted	Smooth
Manuel L. Quezon Statue	Diffused Lighting	Painted Concrete	Rough
Centennial Marker	Harsh Directional Lighting	Marble	Smooth
Burnham Bust	Harsh Directional Lighting	Painted Concrete/Marble	Smooth
Emilio Aguinaldo Monument	Diffused Lighting	Concrete	Rough

Figure 4. Table of Classification of the Statues

It is shown in Figure 4 that three of the statues were documented with harsh directional lighting due to sunlight, while the other two were documented during a calmer clime. The materials of the statues were predominantly concrete, or painted concrete, while one statue was made from stone. Regarding the surface finish, on the other hand, only 2 statues were smooth while the rest were rough. These listed attributes were taken into consideration in the evaluation of the outputted 3D models.

For data acquisition, the researchers made sure to employ certain image capture guidelines. Particularly, image overlapping, taking more photos than required, effective use of the frame, good lighting, and avoiding blind spots. These guidelines helped ensure that the quality and formatting of the images were uniform.

However, since the researchers were only able to use phones for data collecting, consequently, a few issues were encountered. Specifically, limited storage life and battery capacity, lower resolution, and lack of configuration options for the camera.

Furthermore, the researchers meticulously documented the hardware specifications, particularly those relevant to photogrammetry and NeRF dataset processing. The chosen hardware plays a critical role in evaluating the comparative performance between these two techniques, as both are inherently dependent on hardware capabilities and input data. Due to limitations imposed by a mid-tier GPU with only 8GB of VRAM, the researchers implemented a pre-processing step for the photo dataset to reduce VRAM consumption. This pre-processing also facilitates a more efficient rendering process within the photogrammetry and NeRF software, focusing computational resources solely on the object of interest.

Figure 5. Creation of the Dataset for the Statues

Then, in Phase 2, Figure 5, the researchers proceeded with the creation of datasets to push the boundaries of 3D reconstruction using Neural Radiance Fields (NeRF). The researchers chose cultural artifacts like statues and monuments, considering their size, intricate shapes, and diverse materials. To make things even more in-depth, the statues and monuments were photographed under different lighting conditions.

Figure 6. Images masked for use

As seen in Figure 6, the images for the dataset have undergone data cleaning, and unnecessary details such as their backgrounds have been covered. The researchers manually masked the individual images through the use of Adobe Photoshop 2024 and its mask layering feature. This process was used to hide parts of the images without deletion. The images were then exported as PNGs to maintain the mask layering.

Figure 7. The Photogrammetric Reconstruction Process of the Statues

After gathering and cleaning images for the dataset, the researchers proceeded with Phase 3. In Figures 7 and 8, a comparative reconstruction of the statues and monuments is shown. In this phase, the researchers used the constructed datasets to apply Nerfacto which is an implementation of NerfStudio, Alongside this process, photogrammetric models were reconstructed using Agisoft Metashape. The evaluation focused on reconstruction speed, accuracy (both large-scale shape fidelity and detail preservation), and the influence of lighting variations.

Figure 8. NeRF Reconstruction Process of the Statues

This phase delves into the granular analysis and optimization of the chosen NeRF-based reconstruction workflow. A battery of metrics, encompassing error calculations (e.g. STD, Mean_E, RMSE, MAE), cloud-to-cloud comparisons, and performance profiling have been employed to rigorously quantify the strengths and weaknesses of the used NeRF approach.

Metrics comparing the results of the Neural Radiance Fields method against ground truth photographs guided our refinement process. Our analysis also compared results from advanced mesh optimization features within Agisoft Metashape.

Statue / Monument	**Photogrammetry Time(Seconds)**	**Nerfacto Time (Seconds)**
Emilio Aguinaldo Statue	342	1036
Burnham Statue	769	1215
Manuel L, Quezon Bust	347	1022
Centennial Marker	330	990
Crucifix Statue	327	1064
MEAN	423	1065.4

Figure 10. Table of Total Time Processed

In Figure 10, Nerfacto took a considerably longer amount of time compared to its more traditional counterpart, photogrammetry. With a mean of “1065.4”, larger than twice the photogrammetry’s mean of “423”. The researchers hypothesize that NeRF’s utilization of machine learning has caused this difference in processing time.

Through proper surveying, and the compilation of relevant images with varying viewpoints, lighting, and time of the day taken; the researchers have managed to create a dataset that satisfies the requirements of NeRF and photogrammetry, allowing the researchers to properly utilize variants of both digital preservation techniques to recreate the statues to an adequate degree. Through the utilization of both the models that were produced by NeRF and photogrammetry techniques, the researchers performed a comparative reconstruction where the models were judged with set metrics and compared to the ground truth image of the statues. The metrics were measured via numerous error calculations, to determine its efficiency in terms of speed, lighting, and image quality.

Figure 9. Comparative Images between Captured Image, Photogrammetry, and NeRF

Notably, photogrammetry required a significantly larger image set, suggesting its inherent labor-intensiveness [8] and the diminishing final product accuracy when analyzing an object from afar without multiple control ground points [9]. While NeRF achieved model generation with fewer images, its quality suffered accordingly, highlighting its dependence on a substantial dataset as well for accuracy and fidelity. Also, NeRF models its input using infinitesimally small 3D points along a ray, which causes aliasing when rendering views of varying resolutions [10].

In Phase 3, Figure 5, the researchers were able to maximize the capabilities of both NeRF and photogrammetry and their techniques within hardware constraints. This involved reprocessing incomplete or inaccurate models until obtaining the optimal outcome.

While quantitative results are pending, the study's findings temporarily position NeRF as a viable alternative to photogrammetry, albeit not a direct replacement. Both techniques demonstrate unique strengths and weaknesses, rendering their suitability context-dependent.

This research study centered on a comparative analysis of the efficiency of Neural Radiance Fields-based 3D reconstruction techniques versus established photogrammetry methods for the digital preservation of cultural heritage objects. Specifically, for the scope, the focus is on statues and monuments located within Baguio City. Only Nerfacto, an implementation of NeRFStudio, and established photogrammetry methods have been evaluated to pinpoint the most appropriate approach for accurately capturing these heritage assets. The researchers employed quantitative metrics. These include error calculation, cloud-to-cloud comparison, and profiling in assessing geometric accuracy and fidelity. This study prioritized object materials (stone, metal, wood), and varying lighting conditions, alongside analysis of factors like rendering time and image texture quality.

4. CONCLUSIONS AND RECOMMENDATIONS

This study lays the groundwork for further exploration of NeRF and photogrammetry in digital preservation. To advance the understanding and utilization of these techniques, several avenues for future research can be pursued. The digital landscape is actively advancing, and both NeRF and photogrammetry are no exceptions. It is crucial to delve into recent and versatile variants of these methodologies. Evaluating their strengths and limitations in the context of cultural heritage preservation can offer valuable insights. Another way is by developing robust comparative frameworks that go beyond basic error calculations, for it is essential. Metrics should encompass factors such as geometric accuracy, material authenticity, textural fidelity, and visual realism. Ground truth data obtained from high-precision scanning technologies should be used for comparative evaluation, ensuring the reliability and objectivity of comparisons. Establishing standardized benchmarks and data sets would enable researchers to share and compare their findings more effectively.

Both techniques show ample potential in fields other than cultural preservation. 3D model reconstruction in architecture, gaming, filmmaking, and even forensics where a crime scene can be reconstructed in 3D to offer a different and perhaps more holistic point-of-view. These examples are only a few of the many directions that could be taken when utilizing photogrammetry and NeRF, highlighting their versatility even more.

Lastly, while NeRF and photogrammetry are two distinct techniques, exploring the possibility of integration or the creation of a hybrid may prove beneficial to the field of study.

5. ACKNOWLEDGEMENTS

To the panel members, for sharing their suggestions and remarks that have greatly helped us to develop the project. To our adviser Miss Lovely Jen Reformado, Ma'am Thelma Domingo-Palaoag, and the members of CITCS, for enthusiastically viewing and testing the project and providing feedback to the researchers. More importantly, the researchers are grateful for their friends and family who gave them wholehearted guidance, moral support, and encouragement that kept them going.

6. REFERENCES

[1] Akyol, Gamze & Avci, Ali Berkay. (2023). AI APPLICATIONS IN CULTURAL HERITAGE PRESERVATION: TECHNOLOGICAL ADVANCEMENTS FOR THE CONSERVATION. 10.6084/m9.figshare.24077862.

[2] Murtiyoso, A., & Grussenmeyer, P. (2023). INITIAL ASSESSMENT ON THE USE OF STATE-OF-THE-ART NERF NEURAL NETWORK 3D RECONSTRUCTION FOR HERITAGE DOCUMENTATION. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-M-2–2023, 1113–1118.

[3] Magnani M, Douglass M, Schroder W, Reeves J, Braun DR. The Digital Revolution to Come: Photogrammetry in Archaeological Practice. American Antiquity. 2020;85(4):737-760. doi:10.1017/aaq.2020.59

[4] Burdziakowski, P. (2020). Increasing the Geometrical and Interpretation Quality of Unmanned Aerial Vehicle Photogrammetry Products using Super-Resolution Algorithms. Remote Sensing, 12(5), 810. MDPI AG. Retrieved from http://dx.doi.org/10.3390/rs12050810

[5] Guidi, G., & Russo, S. (2019). Photogrammetry for architectural and archaeological documentation. In Digital heritage. Cham: Springer International Publishing.

[6] Remondino, F., Karami, A., Yan, Z., Mazzacca, G., Rigon, S., & Qin, R. (2023). A Critical Analysis of NeRF-Based 3D Reconstruction. Remote Sensing, 15(14), 3585. MDPI AG. Retrieved from http://dx.doi.org/10.3390/rs15143585

[7] Yu, A. W., Park, J., Peng, S., & Ma, C. (2021). PixelNeRF: Neural Radiance Fields from One or Few Images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). https://openaccess.thecvf.com/content/CVPR2021/html/Yu_pixelNeRF_Neural_Radiance_Fields_From_One_or_Few_Images_CVPR_2021_paper.html

[8] C. Zhang, H. Zhou, S. Shah, and H. Yujiang, "Computer vision-assisted photogrammetry and one-image 3D modeling in marine mammals," ResearchGate, Feb. 2023, [Online]. Available:https://www.researchgate.net/publication/370202518_Computer_vision-assisted_photogrammetry_and_one-image_3D_modeling_in_marine_mammals

[9] Chudley, T. R., Christoffersen, P., Doyle, S. H., Abellan, A., & Snooke, N. (2019). High-accuracy UAV photogrammetry of ice sheet dynamics with no ground control. The Cryosphere, 13(3), 955–968. https://doi.org/10.5194/tc-13-955-2019

[10] J. T. Barron, B. Mildenhall, D. Verbin, P. P. Srinivasan and P. Hedman, "Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields," 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 2022, pp. 5460-5469, doi: 10.1109/CVPR52688.2022.00539.