NERF-DRIVEN ALGORITHMS APPLIED ON 360 IMAGES FOR 3D MODELLING OF HERITAGE ASSET IN VIRTUAL ENVIRONMENT

F. Condorelli [1]*, A. Luigini [1]

[1] Free University of Bozen, Italy - (francesca.condorelli, alessandro.luigini)@unibz.it

Commission II

KEY WORDS: NeRF, Neural Networks, 360 images, 3D Modelling, Heritage Education

ABSTRACT

This study aims to fill a gap in existing research by focusing on the application of Neural Radiance Fields (NeRF) algorithms to cultural heritage case studies using equirectangular image data captured with 360° cameras. The main objective is to evaluate and compare the performance of various NeRF algorithms applied to equirectangular images, shedding light on their suitability for cultural heritage preservation. the experiments were carried out on the case study of the library of the seminar in Brixen (Italy). By evaluating the effectiveness of NeRF in combination with conventional photogrammetric methods, the research highlights NeRF's competence in capturing complex details and addressing the challenges encountered in fast and expeditious 3D reconstruction of heritage. The positive results manifest in precise reconstructions, affirming the potential of NeRF in promoting the accuracy and fidelity of 3D models. Despite the computational demands, the study supports further exploration of NeRF-based algorithms, highlighting advantages and some limitations.

1. INTRODUCTION

The context to which this research refers is that of the valorisation of South Tyrol's cultural heritage through the development of interactive videogames for heritage education. The digitization of cultural heritage not only safeguards historical assets but also enables their widespread accessibility and educational use. In this context, the generation of efficient and accurate 3D models is required. Collecting data from the real scene can be done by various methods, such as photogrammetry or laser scanning, but artificial intelligence-based algorithms for creating 3D models, especially for visualisation, are becoming increasingly popular to solve problems such as the quantity and quality of primary data used to create the final model. Among those, Neural Radiance Fields (NeRF) (Mildenhall et al., 2020) is a powerful 3D modelling technique that has gained popularity in recent years. Neural Radiance Fields are a new method for training a neural network to generate and optimise a volumetric representation of a scene from a set of source images. Although its creators focused on NeRF as a means of synthesising new views of a scene, the volumetric representation can be converted into a 3D mesh, making it an alternative to photogrammetry for 3D scanning. The challenges of creating accurate 3D models for heritage environments necessitate advanced techniques and NeRF emerges as a cutting-edge solution due to its ability to reconstruct complex 3D scenes. For this reason, this research focuses on experimenting with technologies that enable the creation of 3D models using artificial intelligence such as NeRF algorithms, evaluating their possible advantages over more traditional methods such as reconstruction using MVS algorithms.
This study addresses a gap in existing research by focusing on the application of NeRF algorithms to cultural heritage case studies utilizing equirectangular image data captured with 360 cameras. While previous studies have extensively examined the efficacy of NeRF algorithms in diverse cultural heritage contexts, this investigation uniquely explores the untapped potential of equirectangular image data. The absence of prior tests with this specific data type underscores the novelty of this research. The primary objective is to evaluate and compare the performance of various NeRF algorithms when applied to equirectangular imagery, shedding light on their suitability for cultural heritage preservation.
The EARTH LAB of Free University of Bozen (Italy) has been working in the field of digitization of heritage for years with the aim of realizing educational paths of art education, both in schools and museums with a particular focus on the valorization of South Tyrol heritage, especially of the city of Brixen. The project involves the digitization of architectural heritage of the historic center of Brixen for the documentation and valorization of the heritage with the creation of educational paths in the form of applied games. A survey campaign started in 2020, using the technologies of 3D laser scanning, terrestrial and aerial photogrammetry, for the creation of immersive environments.
We concentrated on the situation that is the rapid 3D reconstruction of a virtual environment using 360 videos taken with an Ista360 camera. The case study chosen is the library of the seminary in Brixen, an important heritage of the city, built in 1772.
This building was reconstructed with the NeRF method and compared with photogrammetry, highlighting the advantages and disadvantages of using this technique to obtain 3D models.

2. STATE OF THE ART: THE USE OF NERF IN THE CONTEXT OF CULTURAL HERITAGE

The application of photogrammetry encounters difficulties in the case of views with ambiguous observations, such as large areas of homogeneous colors, repetitive texture patterns or significant color variations (Croce et al., 2023).
In these cases, conventional methods can lead to inaccurate reconstructions with noisy or missing surfaces. NeRFs, however, demonstrate a greater ability to handle these complexities, providing more accurate and detailed renderings. In particular,

* Corresponding author

Figure 1. 360 image of the library in Seminary in Brixen.

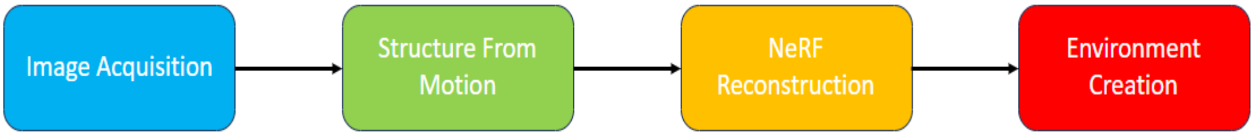

Figure 2. Proposed workflow.

they could be a solution to the challenge of reconstructing metallic, translucent and/or transparent surfaces, objects with homogeneous textures, arrangements involving occlusions, vegetation and complex details. In the field of cultural heritage, Neural Radiance Fields offer a cost-effective solution for rendering objects such as sculptures, archaeological remains, sites and paintings. For this reason, 3D reconstruction in the cultural heritage domain has seen significant progress (Condorelli et al., 2021; Martin-Brualla et al., 2021; Llull et al, 2023; Palestini et al., 2024). Researchers are increasingly using NeRF-based techniques to acquire and recreate detailed 3D models of cultural heritage sites. The advantages mainly concern the fact that NeRF algorithms excel at producing highly realistic renderings of scenes, making them particularly suitable for applications in the cultural heritage sector. The ability to capture complex details in any lighting conditions contributes to the faithful representation of historical artifacts, buildings and landscapes. Despite this progress, challenges remain, including scalability for large-scale heritage sites, managing dynamic scenes, and further improving the efficiency of reconstruction processes. Moreover, NeRF-based techniques often demand high computational power. Finally there may be room for improvement in enhancing the resolution of the final point cloud (Murtiyoso et al., 2023).
Ongoing research is dedicated to addressing these challenges and pushing the boundaries of NeRF applications in the cultural heritage sector. Following this trend, this research involves experimenting with different NeRF algorithms on equirectangular images which have the advantage of giving the possibility of rapid surveying. The basis of NeRF is, as is known, photogrammetry and the application of SfM algorithms, since these algorithms use camera poses as a starting point, this study does not focus on a metric comparison.

3. CASE STUDY AND METHODOLOGY: THE APPLICATION OF NERF ON THE ANCIENT LIBRARY OF SEMINARY IN BRIXEN

This research addresses a challenging scenario, focusing on 3D reconstruction of environments using 360-degree videos captured with an Insta360 camera. The use of the 360 camera has the advantage of obtaining a rapid survey, which is especially necessary in cases of creating environments for video games for educational purposes, which is the main focus of the EARTH lab's research.
The chosen case study is part of the cultural heritage complex of Brixen which is being digitized. The late Baroque library of the Major Seminary of Brixen (Figure 1) represents a significant part of the city's heritage. Built in 1772, the library has a typical Rococo rectangular vault, characterized by a double height with a balcony on all four sides. The ceiling, supported by columns that symbolize writing and tradition, is decorated with frescoes by the Austrian painter Franz Anton Zeiller, known for his works in the Brixen cathedral. These frescoes symbolize the six theological disciplines and originally dictated the thematic organization of the books below.

We selected Nerfstudio (Tancik et al., 2023) as a good method to implement NeRF because it allows the processing of videos in equirectangular format, such as those captured by 360-degree cameras, facilitating the reconstruction of indoor environments such as the library. In fact, Nerfstudio is a library of modular components for implementing NeRF that allows the combination of different pre-implemented methods. Using Nerfstudio you it is possible to train a neural network on a set of custom images and export the 3D data that generates point cloud and meshes. Among the available models, we chose the Nerfacto model. The Nerfacto model is designed for processing real data in fixed scenes and has

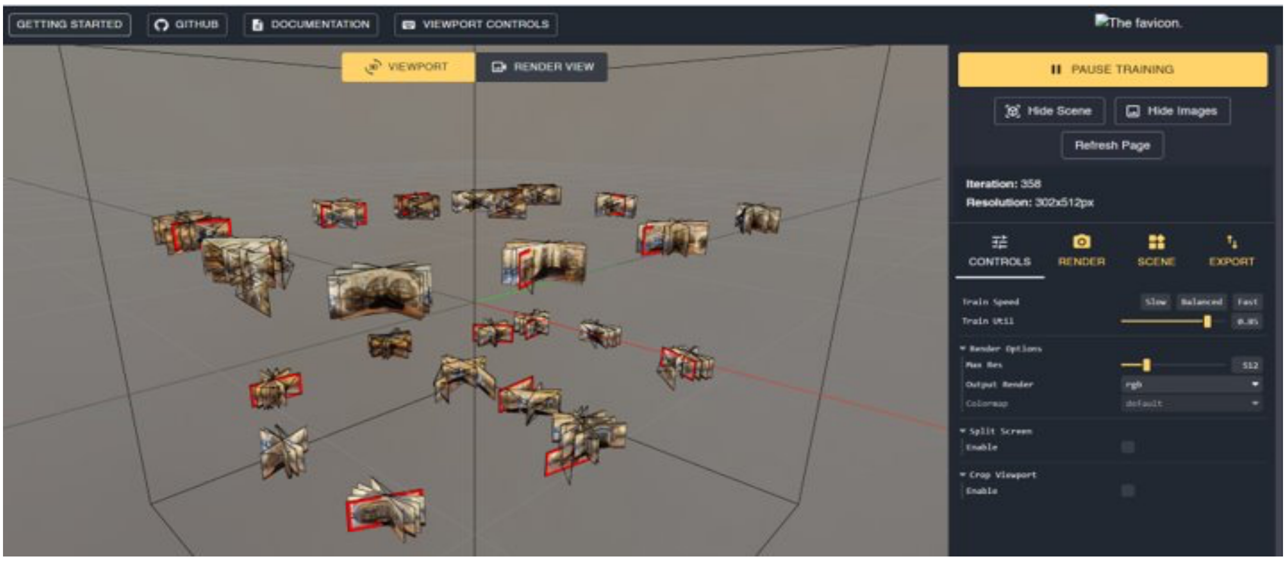

Figure 3. Training of the model in Nerfstudio.

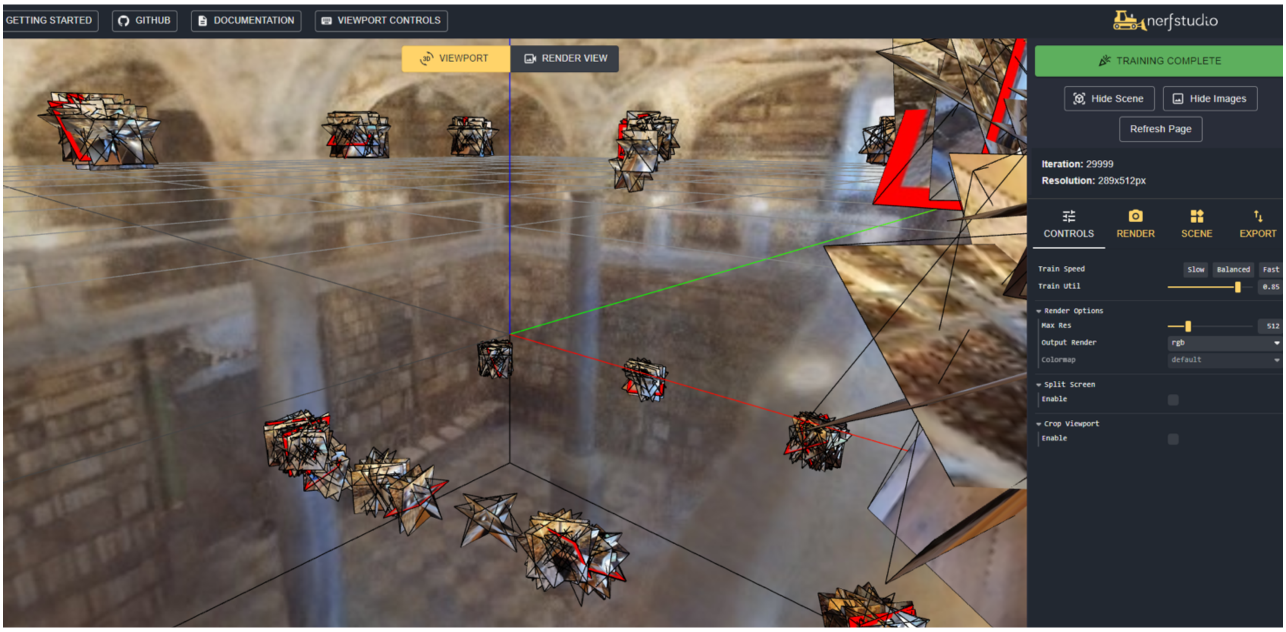

Figure 4. Implementation of Nerfacto in Nerfstudio.

been integrated into Nerfstudio as a mix of methods taken from various publications. The advantages of this method consist in a refinement of camera poses and errors due to incorrect acquisition. It also improves the quality of reconstruction by focusing on important regions.

The workflow we propose follows an effective process to generate 3D reconstructions suitable for cultural heritage. The integration of cutting-edge features resulting from the scientific community's progress on NeRFs contributes to the robustness of the methodology.

This is the workflow we follow, also shown in Figure 2.

1. The first step in creating high-quality NeRF is to carefully capture a scene: typically, a dense collection of photos must be acquired from which 3D geometry and color can be derived. In our case, using a Sony Insta360 camera, we acquired the images via a 360-video made of equirectangular images. To obtain the best possible reconstruction quality, each surface must be viewed from different directions. The more information a model has about an object's surface, the better it can discover the object's shape and its interaction with light.

2. Using a Structure-from-Motion (SfM) pipeline, the captured photos are subject to camera parameter estimation. This includes determining the position, orientation and lens properties for each camera, forming the basis for generating a point cloud representing the geometry of the scene. These parameters associate e Schönberger ach pixel with a point and a direction in 3D space and represent the basis in the NeRF reconstruction process. The COLMAP pipeline (Schönberger et al., 2016) was used as SfM algorithms. This software does not directly support equirectangular image processing. The source images were subjected to a pre-resize process to prepare them for the next phase.

3. Starting from the camera poses calculated with the previous step, the Nerfacto NeRF model is trained from scratch in each captured position. The training process can be followed by viewing the Nerfstudio viewver (Figure 3) where parameters relating to the speed and resolution of the 3D reconstruction process can also be set. At the end of the training the result is a scene synthesized in an accurate and realistic manner. It is then possible to obtain the 3D model which can be exported both as a point cloud and as a mesh (Figure 4).

4. RESULTS

The results of the processing with Nerfacto (Figure 5) were positive, the final reconstruction is in fact quite detailed, avoiding extremely noisy areas, as often happens in processing of this type. As expected, the use of distance-proportional sampling and the sampling proposal contributed to an optimal distribution of samples, ensuring dense coverage for nearby objects and adequate sampling for distant objects. The processing speed was not excessive considering that the quality of the final rendering was concentrated on the most relevant regions of the scene.

To evaluate the accuracy and fidelity with respect to reality, a comparison of the point clouds obtained with Nerfacto with those of the photogrammetry in CloudCompare software was carried out.

The precision in the reconstruction was certainly respected by Nerfacto, as can be seen from Figure 6.

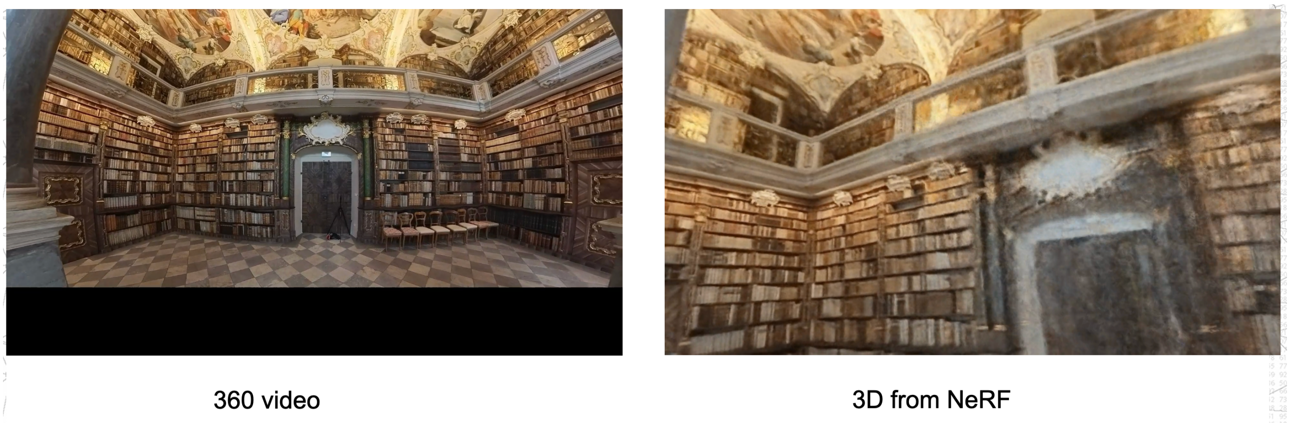

Figure 5. Results of the 3D reconstruction with Nerfacto compared with the initial 360 video.

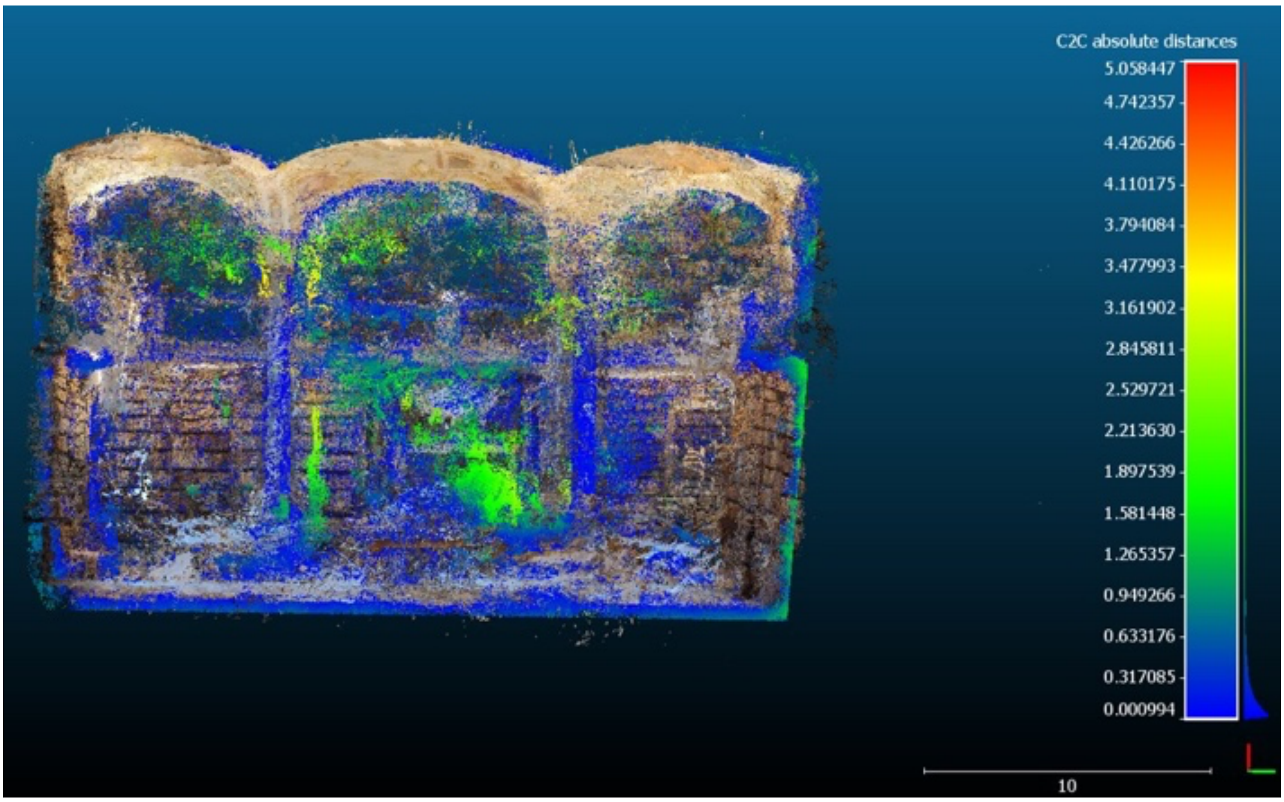

Figure 6. Cloud-to-cloud comparison between the reconstruction from NeRF and the reconstruction with SfM.

The same dataset was also experimented with in another framework born in the wake of Nerfstudio namely SDFStudio (Yu et al., 2022), which is also modular for neural reconstruction of implicit surfaces. The potential lies in being able to experiment with different NeRF algorithms using the same dataset preparation and the same type of setup. SDFStudio includes several algorithms for surface reconstruction such as UniSurf, VolSDF (Yariv et al., 2021), NeuS (Wang et al., 2021) and Neus-facto (Sun et al., 2023). In particular the latter was used but for the present case together with Neuralangelo (Li et al., 2023). NeuS-facto is inspired by Nerfacto where the proposal network from Mip-NeRF360 (Barron et al., 2022) is used for sampling points along the ray. The authors applied this idea to NeuS to speed up the sampling process and reduce the number of samples for each ray. In addition to having computational times that are extraordinarily much longer than Nerfacto, the results were not adequate applied to 360 images. This first test shows that Nerfstudio is more appropriate for equirectangular image type data.

Our findings demonstrate the feasibility and effectiveness of NeRF-based techniques in achieving high-quality 3D models from a set of 360-degree images. One of the advantages includes the ability to synthesize new views of a scene, reducing the effort required for capturing the necessary photographs. These models can be used to enrich cultural heritage education and enhance the visual fidelity of video game environments. The results of our study provide valuable insights into the potential of NeRF-based algorithms in the 3D modeling process for virtual worlds, offering a promising avenue for creating engaging videogame environments in the future.

5. DISCUSSION AND CONCLUSIONS

This study provides valuable insights into the field of computer vision and heritage documentation, improving our understanding of the capabilities and limitations of NeRF algorithms in different imaging scenarios. Continued research and innovation in this field promises to further improve the digitization and preservation of cultural heritage.

While traditional methods remain relevant, the integration of AI-based algorithms, particularly NeRF, signals a paradigm shift towards more efficient, dynamic and accessible approaches to the digitization of cultural heritage.

Despite the computational power requirements and the need to improve the resolution of the final point cloud, NeRF proves advantageous for addressing difficult scenarios that can occur with traditional methods such as photogrammetry.

The research successfully demonstrates the application of NeRF, particularly by Nerfstudio, for the 3D reconstruction of cultural heritage environments. The use of 360-degree images enables efficient processing of large indoor environments. The Nerfacto model proves efficient in processing real-world data and provides both point cloud and mesh representations.

Although Nerfacto presents some improvements over classical photogrammetry in the 3D reconstruction of static scenes, it is important to remember that its implementation requires an in-depth understanding of its components and parameters, making it less accessible for less experienced users. Furthermore, Nerfacto is designed to work with real data, and its effectiveness can vary greatly depending on the quality and variety of input data. In scenarios where the collected data has limitations or errors, the reliability of the model may be compromised. This was also seen by experimenting with the Neus-facto and Nauralangelo algorithms which, for the case in question of images taken with a 360 camera, the results were not satisfactory. While Nerfacto offers improvements, these may not be so radical as to completely justify adopting the model in all situations. In some contexts, classical photogrammetry may still represent a more practical and efficient solution.

Ultimately, Nerfacto certainly offers improvements and represents a promising prospect, but its adoption should be carefully evaluated based on the specific needs of the project and available expertise.

Certainly the integration of NeRF into the digitalization process opens new horizons for creating engaging educational experiences, increasing cultural awareness and preserving historic sites for future generations. Future work should focus on testing new NeRF-based algorithms on challenging cultural heritage cases by varying the type of input data.

ACKNOWLEDGEMENTS

The authors would like to thank the project 'Digital Bozen and Brixen. Documentation, valorization and educational process for cultural heritage' in which this research is being conducted.

REFERENCES

Barron, J., Mildenhall, B., Verbin, D., Srinivasa, P., Hedman, P., 2022: Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields, In: CVPR, 2022.

Condorelli, F., Rinaudo, F., Salvadore, F., & Tagliaventi, S., 2021: A comparison between 3D reconstruction using NeRF neural networks and MVS algorithms on cultural heritage images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 43(B2-2021), 565–570.

Croce, V., Caroti, G., Luca, L. De, & Piemonte, A., 2023: Neural radiance fields (NeRF): review and potential applications to digital cultural heritage. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLVIII-M-2-2023, 453–460.

Li, Z., Muller, T., Evans, A.,Taylor, R.,Unberath, M., Liu, M., Lin, H., 2023: Neuralangelo: High-Fidelity Neural Surface Reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition, 2023.

Llull, C., Baloian, N., Bustos, B., Kupczik, K., Sipiran, I., Baloian; A., 2023: Evaluation of 3D Reconstruction for Cultural Heritage Applications. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, 2023, pp. 1642-1651.

Martin-Brualla, R., Radwan, N., Sajjadi, M. S. M., Barron, J. T., Dosovitskiy, A., & Duckworth, D., 2021: NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 7206–7215.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R., 2020: NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. *Lecture Notes in Computer Science* (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12346 LNCS, 405–421.

Müller, T., Evans, A., Schied, C., & Keller, A., 2022: Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics*, 41(4).

Murtiyoso, A., Markiewicz, J., Karwel, A. K., and Kot, P., 2023: Investigation on the use of NeRF for heritage 3D dense reconstruction for interior spaces, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*., XLVIII-1/W3-2023, 115–121, https://doi.org/10.5194/isprs-archives-XLVIII-1-W3-2023-115-2023, 2023.

Matthew Tancik, Ethan Weber, Evonne Ng, Ruilong Li, Brent Yi, Terrance Wang, Alexander Kristoffersen, Jake Austin, Kamyar Salahi, Abhik Ahuja, David Mcallister, Justin Kerr, and Angjoo Kanazawa, 2023: Nerfstudio: A Modular Framework for Neural Radiance Field Development. In *ACM SIGGRAPH 2023 Conference Proceedings (SIGGRAPH '23).* Association for Computing Machinery, New York, NY, USA, Article 72, 1–12. https://doi.org/10.1145/3588432.3591516.

Oechsle, M., Peng, S., Geiger, A., 2021: UNISURF: Unifying Neural Implicit Surfaces and Radiance Fields for Multi-View Reconstruction. In. *International Conference on Computer Vision (ICCV)*, 2021.

Palestini, C., Meschini, A., Perticarini, M., Basso, A., 2024: Neural Networks as an Alternative to Photogrammetry. Using Instant NeRF and Volumetric Rendering. In: Giordano, A., Russo, M., Spallone, R. (eds) *Beyond Digital Representation. Digital Innovations in Architecture, Engineering and Construction.* Springer, Cham. https://doi.org/10.1007/978-3-031-36155-5_30

Schönberger, J. L., Frahm, J. M., 2016: Structure-from-motion revisited. In: I*EEE Conference on Computer Vision and Pattern Recognition*, CVPR 2016, Vol. 2016, 4104-4113, IEEE Computer Society.

Sun, J., Zhang, Z., Chu, T., Li, G., Zhao, Z., Xing, W., 2023: Joint Optimization of Triangle Mesh, Material, and Light from Neural Fields with Neural Radiance Cache, arXiv:2305.16800.

Yariv, L., Gu, J., Kasten, Y., Lipman, Y., 2021: Volume rendering of neural implicit surfaces. In: *Thirty-Fifth Conference on Neural Information Processing Systems*, 2021.

Yu, Z., Chen, A., Antic, B., Peng, S., Bhattacharyya, A., Niemeyer, M., Tang, S., Sattler, T., Geiger, A., 2022: SDFStudio: A Unified Framework for Surface Reconstruction, https://github.com/autonomousvision/sdfstudio.

Wang, P., Liu, L., Liu, Y., Theobalt, C., Komura, Taku, K., Wenping, W., 2021: NeuS: Learning Neural Implicit Surfaces by Volume Rendering for Multi-view Reconstruction, arXiv:2106.10689.