Neural Radiance Fields for Landscape Architecture

Maximilian Schob[1], Jörg Rekittke[2]

[1]Norwegian University of Life Sciences (NMBU)/Norway · maximilian.schob@nmbu.no
[2]Norwegian University of Life Sciences (NMBU)/Norway

Abstract: In this paper, we examine potential applications of Neural Radiance Fields (NeRF) in the field of landscape architecture. NeRF is a state-of-the-art method for novel view synthesis and volumetric scene reconstruction based on real-world training data. Our paper addresses NeRF and its derived models with a focus on the use and application of Instant-NGP, a method developed by researchers from the technology company NVIDIA. We discuss experimental applications of NeRF based on the case study of the post-disaster landscape of Ahr Valley, Germany, affected by a 100-year flood in 2021. In particular, we are interested in the benefits of NeRF in comparison to other landscape modeling methods, such as Structure-from-Motion (SfM) or Multi-View-Stereo (MVS), which use similar data as input.

This study shows that the application of NeRF technology can be a promising alternative for capturing and visualizing landscape scenes. The study focuses especially on tasks and situations where the larger spatial context – the landscape – is of interest and importance. The technological aspects of how NeRF models work are relevant, but our main focus is on their potential implications for the field of landscape architecture. Technical development and research in the scientific field of computer vision are accelerating rapidly. As users, rather than developers, of digital tools, we believe that NeRF technology requires professional validation through real-world landscape projects.

Keywords: Neural Radiance Field, NeRF, Novel View Synthesis, Instant-NGP (Instant Neural Graphics Primitives), Ahr Valley Flood 2021

Fig. 1: Neural Radiance Fields NeRF model based on data extracted from a drone flight southwest of Ahrweiler, Germany (NeRF model: SCHOB 2022; Drone imagery: REKITTKE 2022)

Journal of Digital Landscape Architecture, 8-2023, pp. 428-442. © Wichmann Verlag, VDE VERLAG GMBH · Berlin · Offenbach. ISBN 978-3-87907-740-3, ISSN 2367-4253, e-ISSN 2511-624X, doi:10.14627/537740046.

1 Introduction

Upon reviewing a corresponding article on computer science, we developed a more comprehensive understanding of what can be generated with a Neural Radiance Field (NeRF) and, consequently, the relevance and potentials of this technology and method in future developments of spatial design in general and more specifically, in landscape architecture. The NeRF approach comes from the computer science field of self-learning systems – as "neural" refers to "self-learning" – and we recognize the task of introducing this method to digital landscape architecture as an urgency, which our contribution is centrally dedicated to. In this context, we set out to generate on-site images – starting with the university campus and ending with a significant flooded area after a disaster – that we could use for our related experiments with NeRF technology.

2 NeRF – Neural Radiance Fields

Neural Radiance Fields (NeRF) is a method for novel view synthesis and volumetric scene reconstruction based on real-world training data. NeRF was introduced by (MILDENHALL et al. 2020) and, since then, has gained traction in computer vision and related fields (GAO et al. 2022, TEWARI et al. 2022). "In its basic form, a NeRF model represents three-dimensional scenes as a radiance field approximated by a neural network. The radiance field describes color and volume density for every point and for every viewing direction in the scene" (GAO et al. 2022). The original approach by Mildenhall et al. "represents a scene using a fully-connected (non-convolutional) deep network, whose input is a single continuous 5D coordinate (spatial location (x; y; z) and viewing direction (θ; ϕ)) and whose output is the volume density and view-dependent emitted radiance at that spatial location" (MILDENHALL et al. 2020). The NeRF model uses a set of two-dimensional RGB images, and their camera poses to create synthetic three-dimensional scenes. These scenes can be rendered into new images or video animations of photo-realistic quality (GAO et al. 2022). NeRF models can also be exported as simple mesh models. Emerging from the field of computer vision, the primary focus of the NeRF method is to produce visual representations of a scene, surface, or object. Unlike other methods and sensors in remote sensing and environmental modeling, NeRF does not originate from a surveying or measurement context. NeRF uses internal coordinate systems instead of geographic reference systems, which were not a priority in its development. In the field of landscape architecture, the fact that the NeRF model is not connected to a real-world coordinate system that potentially links data to a ground-truth reference might be unusual at first. The rooting in scanning and surveying that led to the rise of lidar point cloud models in the field is being replaced with a kind of ground truth of images. The article "Ground truth to fake geographies: machine vision and learning in visual practices" by GIL-FOURNIER & PARIKKA (2021) is of importance to this discussion and, particularly where the authors argue that "ground truth has shifted from a reference to the physical, geographical ground to the surface of the images."

A NeRF model is not limited to generating a radiation field but can also be used to generate point-based radiation fields (XU et al. 2022) or voxel-based models (YU et al. 2021). Since its publication in 2020, the paper 'NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis' by MILDENHALL et al. (2020) has inspired many researchers to advance, adjust, and refine their methods (GAO et al. 2022, TEWARI et al. 2022). Especially the processing speed metric has been a major threshold in making the method available to a wider array of users, as it is directly connected to the complexity of scenes and the hardware necessary for their generation. Comparing the element of speed between the newer and older models, we can observe that the former outperforms the latter by several orders of magnitude. While processing a specific scene takes twelve hours in 2020 (MILDENHALL et al. 2020), the same scene takes only about five seconds in the middle of 2022 (MÜLLER et al. 2022).

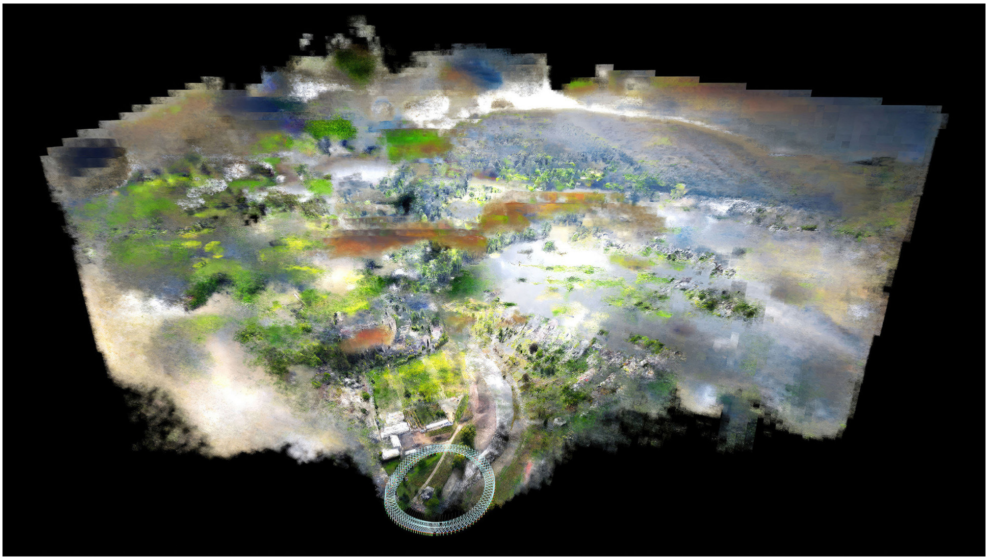

Fig. 2: Visualization of the technical-spatial extent of a zoomed-out NeRF model, viewed from above. The light blue object circle at the bottom center of the image shows the camera position and direction of the drone images used as training data. An inward-oriented circle (camera orbit) was flown (NeRF model: SCHOB 2022; Drone imagery: REKITTKE 2022).

GAO et al. (2022) provide an overview of existing literature grouped based on their focus on applications such as three-dimensional reconstruction, image processing, or urban applications. Especially interesting is a model for large-scale scene reconstruction (TANCIK et al. 2022) that lets us envision potentially global scale models. Significant improvements have been made to NeRF models, creating a wide range of applications, including "urban mapping / modelling / photogrammetry, image editing / labelling, image processing, and 3D reconstruction and view synthesis of human avatars and urban environments" (GAO et al. 2022). Recent advances in NeRF model performance have also made this technology more accessible to professionals in related fields outside of computer vision. More specifically, professionals from fields involved in digital visualization and aesthetics, such as landscape architecture, will be encouraged to test and develop their own models in relation to their specific tasks and topics.

2.1 NeRF Aesthetics

Customary techniques for reconstructing three-dimensional landscape scenes, such as point clouds or vectors derived from photogrammetry, result in models whose aesthetics are detached from their physical context – the surrounding landscape. Where the lidar scanner rays end for a point cloud model, a black hole opens as the model background. Such models represent their own digital aesthetics and stand in stark contrast to the realism of photography and film. From an aesthetic point of view, there was always a significant difference between the navigable three-dimensional model and the modeled real scene. Lidar scans, or photogrammetric scene reconstructions, seem to be functionally limited by their rootedness in technical correctness and dimensional accuracy. It is difficult to implement diverse, complex, and associative topical links in these models. In contrast, NeRF technology translates the ability of photography or film to capture the full context of a scene, including the background, into a complex three-dimensional model with an identifiable background (Fig. 1 and 2). A NeRF model generates a detailed reconstruction of variables, which are key in registering a scene. In addition to position and color, other variables transferred to the model include light intensity, darkness, and transparency. The ability to incorporate these elements presents an unprecedented three-dimensional realism (Fig. 3). Point cloud models, with high point densities and an even distribution of points, appear sparse up close and denser from a distance, where this higher density does not correspond to increased information. Combining point cloud models with different resolutions or resampled data can thus be useful for creating large models. Christophe Girot describes such combined models through the term he coined as *cloudism* (GIROT 2020). As we have established the possibilities of the NeRF method, we propose to counteract the newfangled term *cloudism*, again with a classic term – realism. Landscape in the form of a model is still most accurately understood – by laypersons and experts alike – when it corresponds to the common appearance of the surrounding landscape.

Fig. 3: Comparison of training data in the form of drone images (left) and the resulting NeRF model (right). The slightly lower image definition of the NeRF is noticeable only on close inspection. However, the landscape as such is clearly identifiable (NeRF model: SCHOB 2022; Drone imagery: REKITTKE 2022).

2.2 Instant-NGP

For the generation of our experimental NeRF outcomes, we used Instant-NGP (Instant Neural Graphics Primitives), an open-source software framework developed by NVIDIA. The software framework processes Neural Graphics Primitives that, in addition to NeRF, can also be used for Gigapixel images, neural Signed Distance Functions (SDF), and Neural Radiance

Caching (NRC). Our paper is limited in scope to NeRF models. Instant-NGP is proving to be one of the most popular and regularly updated NeRF generation solutions. It trains a NeRF in seconds using multi-resolution hash encoding. The coordinates are hashed and used as an index into a stack of multi-resolution data arrays, drastically reducing the number of parameters per model. The NeRF model is constrained by a unit cube bounding box set at a coordinate space of $[0,1]^3$. The model has the highest resolution around a central point positioned at the center of the unit cube, at [0.5, 0.5, 0.5].

3 Case Study Ahr River Valley

For our initial experimentation with the application of NeRF technology, we focused on the case of the Ahr Valley in Germany in the aftermath of the 2021 flood disaster. We obtained the related fieldwork data through camera tours and UAV flights on-site. In the summer of 2021, between July 12th and July 15th, the Ahr River Valley experienced a 100-year flood as a result of pronounced heavy regional rainfall events in connection with a low-pressure system. In addition, the soils in the affected regions of Rhineland-Palatinate and South Westphalia could hardly absorb any additional water (GERMAN WEATHER SERVICE 2021). After the flood, which took the lives of many people and caused extreme destruction, the hasty reconstruction activity did not necessarily lead to sustainable design and building.

Fig. 4: NeRF model based on data extracted from a drone flight southwest of Ahrweiler, Germany. The frame circle in cyan shows the individual frames and the viewing direction of the drone flight (NeRF model: SCHOB 2022; Drone imagery: REKITTKE 2022).

“The moment after a natural disaster is a window of time that can be used to adapt-to-climate (change), but this opportunity is in many cases demonstrably wasted. […] After a disaster, amnesia leads people to forget about what primarily should be designed and built" (REKITTKE & NINSALAM 2022). Without a thorough analysis of a disaster, economically and ecologically sensible decisions become unlikely. There is a danger in conducting post-disaster analysis and the subsequent planning and design relying solely on documents like maps or legal texts, which operate on a high level of abstraction. It is imperative to incorporate what the people themselves have seen (REKITTKE & NINSALAM 2022). We are interested in NeRF technology for this particular reason, as it opens up new possibilities for visual realism and the ability to integrate different temporal layers into a single landscape model. Disasters reveal snapshots of many aspects that should have been taken into account during planning phases and are overshadowed once the developments are carried through a short time later. In this pursuit, we aspire to preserve the memories of a flood disaster by creating appropriate landscape models. Like in an autopsy, the aim is to fill the common gap between reality, recollection, and forward planning with evidence that is supposed to trigger cogitation (REKITTKE & NINSALAM 2022).

3.1 Data Collection and Processing

For testing NeRF models in the context of post-flood Ahr River landscapes, we created an extensive dataset consisting of 106 video and image samples using UAV-mounted cameras (Fig. 4) and ground-based handheld smartphone devices. Our data were collected in two separate sessions. The first was during the flood event in July 2021– sporadic and *ad hoc*. The second was in September 2022, in the course of systematic fieldwork. Our aim was to create cases using one of the most common NeRF methods available. All NeRF models were trained locally using Instant-NGP. The GitHub repository (GITHUB / instant-ngp 2022) provides documentation on software and hardware requirements, installation, pre-processing, training, and rendering NeRF, as well as exporting. Another document on GitHub provides additional advice on the process (GITHUB / nerf_dataset_tips 2022). We created a separate NeRF model for each set of input data, following a list of six sequential processing steps: 1) data acquisition, 2) data pre-processing and frame extraction, 3) pose estimation, 4) NeRF model training, 5) video export, and 6) post-processing.

1) Data acquisition for each site was carried out using lightweight, field study-ready collection devices: a DJI Mini drone and an iPhone 11 Pro. For all data acquisition, we used the highest possible resolution of the devices. The drone videos were shot at 2.7K resolution (2720x1530), 23,97 fps, in MPEG-4 format. With the smartphone camera, we shot videos at Full HD resolution (1920x1080), 59,94 fps, in MPEG-4 format, and pictures at 12MP resolution (4032x3024), in JPEG format. Shots in the wide-angle camera mode (0.5x, 13 mm equivalent focal length, 120° field of view) were particularly effective. In total, we collected 96 videos and 10 image sequences, a raw data package of 35 Gigabytes.

2) For the video shots from the field, we extracted a set of sequential frames ranging from 50 to over 400 images. The image sequences were filtered to the same amount. NeRF models based on the method of Müller et al. (2022) do not infinitely increase resolution or quality with a more extensive set of input images. Mildenhall et al. (2020) and Müller et al. (2022) use tens to hundreds of images to train NeRF models. We followed the recommendations presented on the GitHub forum (GITHUB / nerf_dataset_tips 2022). Furthermore, we tested the application of digital image enhancement techniques such as sharpening, noise reduction, and super-resolution to improve matching image detection.

3) We used the COLMAP pipeline that was part of the Instant-NGP codebase to process an estimate of the camera poses for each image set. The resulting JSON file containing the camera parameters for each image was saved in a folder along with the original images in the format TRANSFORMS.JSON.

4) We trained our NeRF models using the interactive GUI (Graphical User Interface) that was included in the codebase. The GUI offers a variety of different tools for training, visualization, and export, as well as allowing the user to interactively move through a scene while the model is being rendered in real time. Training begins by launching the GUI from an Anaconda prompt, and within seconds, the model evolves from blurry noise to a clear representation of a scene. Once the training reaches a satisfactory level, the training progress can be saved as a JSON file – called *snapshot* – which can be used to reload the NeRF model or to create an animation. The GUI facilitates interactive creation and saves a camera path along a set of key frames that can be used to export a video animation.

5) The codebase allows exporting of flythrough video animations of the NeRF model using the previously generated training progress and camera path. The export is handled outside the GUI in an Anaconda code prompt using Python bindings. We exported a range of video animations up to 4K resolution at 30 fps.

6) The exported videos can be easily edited using common video and image editing tools. Since both the input and output of the NeRF model are image-based, digital editing and processing pipelines such as image sharpening, noise removal, or frame interpolation can be applied before and after NeRF modeling. The user has full control over both pre- and post-NeRF model media, as would be the case with photography, photogrammetry, or map-making.

4 NeRF for Landscape Architecture

Although we have been working with NeRF technology for a limited time and therefore did not yet utilize its full potential, we can already identify and highlight some of its specific strengths. We offer a selection of tested applications for NeRF technology, generally for landscape architecture and, more specifically, in the context of our case study. In addition to the enormous technological advances that have determined the rise of NeRF technologies in recent years, the method has to be tested in relation to issues concerning landscape architecture.

4.1 Multi-Resolution Models

For NeRF models, it applies that their resolution is not developed in relation to the model but to the depth of information captured from the input images. “The multiresolution aspect of the hash encoding covers the full range from a coarse resolution Nmin that is guaranteed to be collision-free to the finest resolution Nmax that the task requires. Thereby, it guarantees that all scales at which meaningful learning could take place are included, regardless of sparsity” (MÜLLER et al. 2022). In the context of the Ahr Valley after the flood, we are able to create a lightweight but complex three-dimensional model that enables capturing environments at multiple scales: from the rocks in the Ahr riverbed to the flowing water, from the riverbanks and adjacent vegetation to patterns of the urban fabric, and from the mountains in the background to the clouds in the sky above the valley. The main benefit we see in our case-based NeRF models is that they feature a high spatial depth, capturing the sky, clouds, and even distant landscape features such as mountains, valleys, and urban areas. In the case of the Ahr Valley, the NeRF model consolidates all flood-relevant factors to be discussed simultaneously in one model: the change in sediment flows in the river, the altered course of the river, the destruction of urban settlements and agriculture in the Ahr valley near the floodplain, the topography of the valley where water has accumulated downstream. More advanced NeRF models can present the rich contextual depth of the mapping in the form of a navigable three-dimensional model. Xiangli et al. (2022) outline the nature of such prospective models by expanding the notion of rendering scenes at multiple resolutions by "modeling different scenes at multiple scales with drastically varying views on multiple data sources" (XIANGLI et al. 2022).

4.2 Object Focus versus Open World Scene?

Based on our current research and studies, we observe a certain level of contradiction regarding the great landscape potential of NeRF models and their technical nature. NeRF is designed to feature a central point in the model, from which the resolution gradually decreases towards the edges of the bounding cube. This raises the question of whether NeRF models are inherently object-oriented (single object) and how this might impact the modeling of non-point-centric open-world scenes, such as landscapes. Is our positive assessment of the NeRF landscape model accurate, or will the triumph of the NeRF models primarily extend to object models, for example, in the context of architectural projects? Instant-NGP NeRF models are constrained by a maximum resolution bounding box at its center. But in our case study, we fed this "machine" exclusively with data from landscape photography and landscape videos, thereby obtaining effective landscape models. We suggest that future explorations are "to be continued."

4.3 Comparison to Photogrammetry and Point Cloud Models

There are partial overlaps in data collection and processing methods between NeRF models and photogrammetric processing. Both methods use two-dimensional raster images as input data and share further similarities in the initial processing of this input data. A wide range of NeRF methods, including the one of Müller et al. (2022), use COLMAP, a package for SfM, to extract camera poses. Models derived from lidar scanning or photogrammetric modeling still offer higher geometric accuracy than NeRF models (LEHTOLA et al. 2022). For our own comparison of the different methods, we created a set of 406 frames from a selected drone flight, which we used as input for corresponding NeRF and photogrammetric models. The NeRF model was generated with Instant-NGP, and the photogrammetry model was processed as a dense point cloud in Agisoft Metashape. The point cloud was exported as a file in LAS format with 12 million points and a file size of 326 MB.

The output quality of a photogrammetry process is assessed based on the accuracy of where the resulting points are positioned with respect to a ground truth reality. Passive sensing data, such as intensity return data or true color imagery captured by other sensors, may only support subsequent analysis or enhance visualization. In many ways, the NeRF model sits somewhere between the perception of space and the perception of textures, materials, light, and color. There are various metrics can be used to assess the quality of a NeRF model. Many fundamental advantages of photogrammetry, such as a relation to a "real world" Coordinate Reference System (CRS) and transformations of the model with respect to this CRS, are not yet realized in the NeRF system we used. Nonetheless, this does not preclude the possibility of implementing them in future applications.

A NeRF differs from all the traditional three-dimensional scene formats commonly used in the field, some of which include vectors or meshes, grids, and point clouds. Each format can have distinct ways of formulating a representation, which can yield a unique set of advantages and disadvantages depending on the format used. It can therefore be difficult to judge a NeRF in relation to the qualities of these formats. Some of these metrics may be embedded in the process or acquisition technology used for data generation to determine the potential for accurately representing a scene from the start. Different metrics apply to data obtained from photogrammetry or lidar scanning. Such comparisons exist in various areas that are tangentially linked to landscape architecture, for example, in heritage preservation. Data obtained through lidar scanning is among the most accurate geometric data available in modern scanning techniques but lacks the ability to accurately capture the texture and diagnostic color information (DOSTAL & YAMAFUNE 2018). Perhaps the most fascinating outcome of the method is not the NeRF model itself but the images and videos that are generated from the model (Fig. 5 and 6). MILDENHALL et al. (2020) state that the results of view synthesis are best viewed as videos.

Fig. 5: Comparison of a NeRF model (left) and a model derived from photogrammetry (right) of the same area. The point cloud model is black in the background, while the NeRF model captures the depth of the landscape at several levels of resolution (NeRF model: SCHOB 2022; Drone imagery: REKITTKE 2022).

Fig. 6: Comparison of a NeRF model (left) and a model derived from photogrammetry (right) of the same area. Even in the close-up scene, the NeRF model provides higher visual resolution than the point cloud model (NeRF model: SCHOB 2022; Drone imagery: REKITTKE 2022).

4.4 Material Noise

Metrics such as reflection or change in transparency and color as a function of position are not considered part of the physical enterprise of the object or scene in the current literature on photogrammetrically derived models. Rather, they are viewed as factors that distort or corrupt the signal in ways that may need to be eliminated in order to derive a high-quality model. In scenes produced by photogrammetry, removing water-surface reflection effects presents a challenge (PARTAMA et al. 2018). Materials featuring difficult optical properties – including but not limited to absorptivity, reflectivity, scattering, challenging texture and complex shape or geometry – still pose challenges in photogrammetry (NICOLAE et al. 2014). The distinction between signal and noise is pronounced in the literature on photogrammetry and, more generally, in remote sensing and earth observation. The NeRF model, on the other hand, allows data otherwise defined as noise to be used to visualize unstable materialities, surfaces, and objects. In our case study, for example, we had the means to illustrate the unstable nature of the Ahr River – with its changing water levels up to extreme flooding conditions.

4.5 Multi-Source NeRF

A multi-source model is based on input data generated through multiple acquisitions for the same or a similar area. This method can be used in situations where only a sparse set of input data is available to increase the resolution of the NeRF model by adding additional input data for angles or features not previously captured. For our case study, we trained a NeRF model of the Kalvarienberg Monastery and the surrounding area in the town of Ahrweiler with several of our drone acquisitions and eventually improved the model's resolution. In addition, our model captures the changing light conditions between diffused and direct sunlight caused by different cloud conditions during the recording period. The various parameters – geometry, color, lighting, texture, and translucency – captured by a NeRF model may be acquired independently. Each parameter can be derived from a separate set of input data. The final NeRF synthesis model allows navigating between these parameters in relation to the position and direction of a particular view. In working with a 3D landscape model, this synthesis is a novelty that opens up considerable potential.

4.6 Multi-Temporal NeRF

It sounds almost unattainable within the limitations and resources of our current time, but a simultaneous coupling of movement through time, and movement through space, is fundamentally possible with NeRF technology. This option is yet to be defined and therefore, we propose to use "Multi-temporal NeRF models" when referring to the visualization of changing layers of time in the course of changing positions. Multi-temporal NeRF models use multiple sets of images captured at different times and utilize them as the input to produce novel views that interpolate between the images. The model synthesizes the input data and allows it to move through time while moving through space. A Multi-temporal NeRF model makes it possible to capture movements, visualize ongoing processes, and depict all kinds of patterns of change. For example, the growth or the changing state of the health of vegetation can be documented in this way. The intensity of the changes captured by the model can be related to the temporal extent of the capture and the intensity of the change in the underlying object or study surface. Multi-temporality is a common concept and method in geosciences, in which remote sensing observations collected at different times are combined into a single multi-temporal image or model. Multi-temporal analyses enable the detection and visualization of changes in spatial patterns over time. The concept has taken root in areas such as architecture and landscape architecture to understand changes in the built environment. The landscape architect, in particular, can think of numerous possible uses. The depiction of seasonal changes in the city, landscape, and vegetation are only a few of them. We find this method to be the most effective to this date for purposes of representing the "before" and "after" conditions of a site.

In 2021, it was already demonstrated that NeRF models could be trained with unstructured collections of photographs taken at different times, from different angles, and under different lighting conditions. The model registers the static geometries of the scene but interpolates between color and illumination in dependence on the view position (MARTIN-BRUALLA et al. 2021). It is possible for a NeRF to process a sequential set of images of the same scene at different times of the day, times of the year, and so on. The associated different lighting conditions of these different images, which show a time difference, allow the generation of an outstanding level of multi-temporality in a single NeRF result. The resulting model allows the user to literally move through time as they move through space – made possible by adjusting the different radiation fields between the time-shifted images.

Fig. 7: Multi-temporal NeRF, 2019 (left) and 2022 (right) captured in a single model (NeRF model: SCHOB 2022; Drone photo material (AWVISION 2022), for 2019 and 2022)

In our case study, the Multi-temporal NeRF shifts the digital model from a state of representation to a state of simulation of the underlying flood event. Our NeRF model captures and interpolates situations found in two self-contained trajectories. The model combines two drone flights – one in 2019 and the other in 2022 – over the Ahr Valley municipality of Rech, Germany. The acquisition from 2019 shows a historic bridge over the Ahr River, connecting the two halves of the village. The second acquisition captures the same location in the autumn of 2022, after the flood event destroyed parts of the bridge, swept away several buildings south of the bridge, and visibly changed the course of the river (Fig. 7). Both drone flights have different trajectories and viewpoints, allowing us to create a model relating one set of views to the 2019 acquisition and the other set of views to the 2022 acquisition. In the resulting NeRF model, we can navigate through the internal digital coordinate system and observe the changing states in quasi-real-time.

4.6.1 Flooding as a Multi-Temporal NeRF Application

Due to the presence of temporal components – such as large amounts of water that flows in and out of a particular site – flood zones are generally considered to be suitable for the application of multi-temporal NeRF. Instant-NGP's interactive GUI also provides a set of visual debugging tools that can be used to uncover the internal structure of input and output neurons. The parameters of these tools are recorded as part of a keyframe animation. For our case study dealing with speculative scenarios of a flood-ravaged valley, we used the partial activation of neurons as a rising green light field to simulate the rising waters of the flood through visuals (Fig. 8). Through this experimentation, we found a simple yet very powerful tool for digital flood simulation with full visibility of all model elements. Supporting this hypothesis (LI et al. 2022) have shown that NeRF models can be used to simulate ultra-complex climate events.

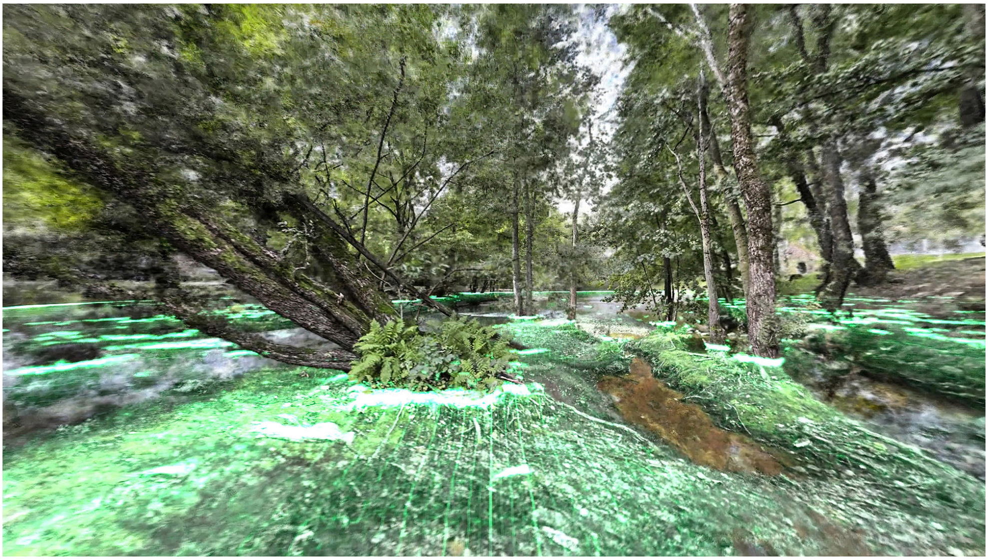

Fig. 8: Flood simulation using a partial activation of neurons in the NeRF model (NeRF model and photo material: SCHOB 2022)

5 Conclusion

This paper presents a baseline study on various approaches and methods of working with NeRF models in landscape architecture. The aim is to inspire and encourage researchers in related fields to develop in-depth studies on the applications of NeRF. Our interest in NeRF is fundamentally rooted in the idea that landscape is anything but static, which is, at the same time, where we find significant potential in this approach. NeRF models can be useful for “wandering” through changing light conditions or addressing moving objects, such as water, clouds, birds, cars, trains, and others. Different materials can be evaluated through direct comparison. Reflections and light, as well as structure, can be included in their changing appearance. Changed terrain, for example, differing terrain heights in the course of a construction project, can be evaluated. In addition, the same scenes and objects could be recorded under different lighting conditions in order to enable a critical evaluation. For example, early in the morning, at noon, in the evening, or in cloudy weather. The NeRF interpolates between the input images, allowing for seamless switching between different states within the resulting model. It is an important task to think of a landscape model not as a static set of coordinates, i. e. point clouds, raster, or vector data, but as a set of parameters that are constantly changing. As in a landscape as such, the model changes depending on the viewer's position – an unstable model. The fact that NeRF crosses the border between modeling and simulation suits the instability and openness of the landscape subject. We are dealing with a technology that is still very new and largely untested but whose potential seems enormous. Most computer graphics algorithms and techniques developed over more than half a century assume meshes or point clouds as three-dimensional scene representations for rendering and editing. Neural rendering, on the other hand, is such a recent field that the term was first used in 2018. For this reason, there is an inevitable gap between the available methods that can work with classic three-dimensional representations and those that can be applied to neural representations (TEWARI et al. 2022). This is definitely true for the field of landscape architecture, and we look forward to the developments and publications that will qualify NeRF models for landscape architecture in the years to come.

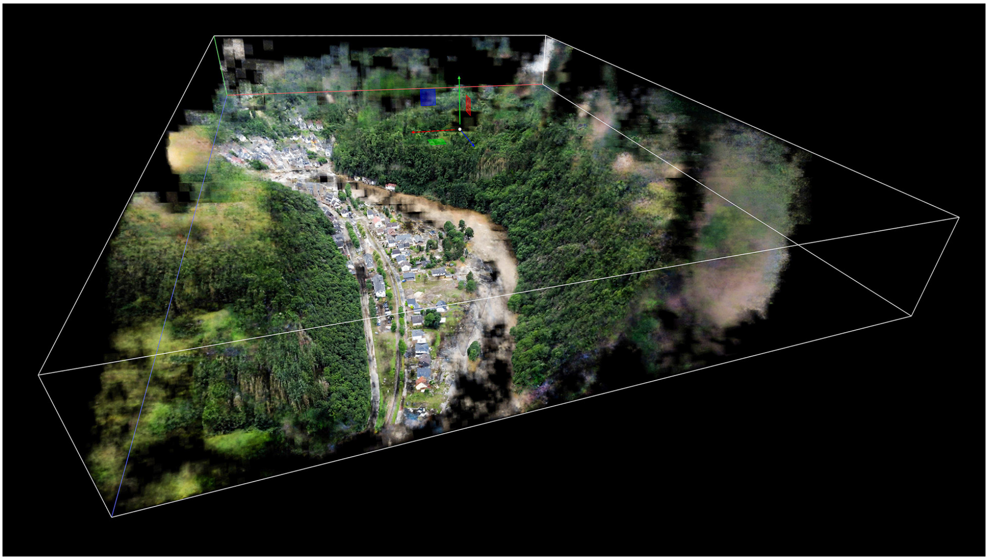

Fig. 9: NeRF model limited to a bounding box that serves as section planes (NeRF model: SCHOB 2022; Drone photo material: OSTERBURG 2021)

In 2019, Christophe Girot described how the point cloud model overcomes the separation between model, architectural drawing, renderings, or other visualizations. "In the "cloudist" approach, there exists no separation between a model, a section and a plan: they all stem from the same cloud of design information. Separate renderings or visualizations become quite unnecessary, since the views generated are directly derived from the model, with their own singular aesthetic" (GIROT 2019). NeRF models remove the threshold between different forms of representation (Fig. 9). The NeRF method offers qualities similar to point clouds but significantly reduces the separation or the visual contrast between the model and reality.

References

AWVISION (2022), Rech an der Ahr (2019/2022) nach der Flut. youtube.com. https://www.youtube.com/watch?v=vQhgOmIte_o (08.08.2022).

DOSTAL, C. & YAMAFUNE, K. (2018), Photogrammetric texture mapping: A method for increasing the Fidelity of 3D models of cultural heritage materials. Journal of Archaeological Science: Reports, 18, 430-436. https://doi.org/10.1016/j.jasrep.2018.01.024.

GAO, K., GAO, Y., HE, H., LU, D., XU, L. & LI, J. (2022), NeRF: Neural Radiance Field in 3D Vision, A Comprehensive Review (arXiv:2210.00379). arXiv. http://arxiv.org/abs/2210.00379 (30.11.2022).

GERMAN WEATHER SERVICE (2021), Hydro-climatological classification of heavy and permanent precipitation in parts of Germany in connection with the low-pressure area "Bernd" from July 12th to 19th, 2021 (by: JUNGHANEL, T. et al.). https://www.dwd.de/DE/leistungen/besondereereignisse/niederschlag/20210721_bericht_starkniederschlaege_tief_bernd.html (24.10.2021).

GIL-FOURNIER, A. & PARIKKA, J. (2021), Ground truth to fake geographies: Machine vision and learning in visual practices. AI & Society, 36 (4), 1253-1262. https://doi.org/10.1007/s00146-020-01062-3.

GIROT, C. (2020), Cloudism. In: MIHYE, AN, M., HOVESTADT, L. & BÜHLMANN, V. (Eds.), Architecture and Naturing Affairs. Birkhäuser, Berlin/Boston, 96-101. https://doi.org/10.1515/9783035622164-013.

GIROT, C. (2019), Cloudism. In: BRAAE, E. & STEINER, H. (Eds.), Routledge research companion to landscape architecture. Routledge.

GITHUB/instant-ngp (2022), NVIDIA Research Projects. https://github.com/NVlabs/instant-ngp (Original work published 2022) (18.03.2023).

GITHUB/nerf_dataset_tips (2022), NVIDIA Research Projects. https://github.com/NVlabs/instant-ngp/blob/ea776524ce891b19ad740bbff81a51f146493a67/docs/nerf_dataset_tips.md (Original work published 2022) (18.03.2023).

LEHTOLA, V. V., KOEVA, M., ELBERINK, S. O., RAPOSO, P., VIRTANEN, J.-P., VAHDATIKHAKI, F. & BORSCI, S. (2022), Digital twin of a city: Review of technology serving city needs. International Journal of Applied Earth Observation and Geoinformation, 114, 102915. https://doi.org/10.1016/j.jag.2022.102915.

LI, Y., LIN, Z.-H., FORSYTH, D., HUANG, J.-B. & WANG, S. (2022), ClimateNeRF: Physically-based Neural Rendering for Extreme Climate Synthesis (arXiv:2211.13226). arXiv. http://arxiv.org/abs/2211.13226 (28.12.2022).

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 (arXiv:2008.02268). arXiv. http://arxiv.org/abs/2008.02268 (05.11.2022).

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 (arXiv:2003.08934). arXiv. http://arxiv.org/abs/2003.08934 (28.11.2022).

MÜLLER, T., EVANS, A., SCHIED, C. & KELLER, A. (2022), Instant neural graphics primitives with a multiresolution hash encoding. ACM Transactions on Graphics, 41 (4), 1-15. https://doi.org/10.1145/3528223.3530127 (21.12.2022).

NICOLAE, C., NOCERINO, E., MENNA, F. & REMONDINO, F. (2014), Photogrammetry applied to Problematic artefacts. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5, 451-456. https://doi.org/10.5194/isprsarchives-XL-5-451-2014.

OSTERBURG, S. (2021), Ahrbrück 15.07.2021. youtube.com. https://www.youtube.com/watch?v=Sb5jLrXnxM0 (27.01.2023).

PARTAMA, I. G. Y., KANNO, A., UEDA, M., AKAMATSU, Y., INUI, R., SEKINE, M., YAMAMOTO, K., IMAI, T. & HIGUCHI, T. (2018), Removal of water-surface reflection effects with a temporal minimum filter for UAV-based shallow-water photogrammetry: Removal of Water-Surface Reflection Effects for UAV-Photogrammetry. Earth Surface Processes and Landforms, 43 (12), 2673-2682. https://doi.org/10.1002/esp.4399.

REKITTKE, J. & NINSALAM, Y. (2022), ///flood.landscape.autopsy: Digital Post-Disaster Retrospection for Sustainable Design Decision Making. Journal of Digital Landscape Architecture, 7-2022, 365-378.

TANCIK, M., CASSER, V., YAN, X., PRADHAN, S., MILDENHALL, B. P., SRINIVASAN, P., BARRON, J. T. & KRETZSCHMAR, H. (2022), Block-NeRF: Scalable Large Scene Neural View Synthesis. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 8238-8248. https://doi.org/10.1109/CVPR52688.2022.00807.

TEWARI, A., THIES, J., MILDENHALL, B., SRINIVASAN, P., TRETSCHK, E., YIFAN, W., LASSNER, C., SITZMANN, V., MARTIN-BRUALLA, R., LOMBARDI, S., SIMON, T., THEOBALT, C., NIEßNER, M., BARRON, J. T., WETZSTEIN, G., ZOLLHÖFER, M. & GOLYANIK, V. (2022), Advances in Neural Rendering. Computer Graphics Forum, 41 (2), 703-735. https://doi.org/10.1111/cgf.14507.

YU, A., FRIDOVICH-KEIL, S., TANCIK, M., CHEN, Q., RECHT, B. & KANAZAWA, A. (2021), Plenoxels: Radiance Fields without Neural Networks (arXiv:2112.05131). arXiv. http://arxiv.org/abs/2112.05131 (20.12.2022).

XIANGLI, Y., XU, L., PAN, X., ZHAO, N., RAO, A., THEOBALT, C., DAI, B. & LIN, D. (2022), BungeeNeRF: Progressive Neural Radiance Field for Extreme Multi-scale Scene Rendering. In: AVIDAN, S., BROSTOW, G., CISSÉ, M., FARINELLA, G. M. & HASSNER, T. (Eds.), Computer Vision – ECCV 2022 (Vol. 13692, 106-122). Springer Nature, Cham, Switzerland. https://doi.org/10.1007/978-3-031-19824-3_7.

XU, Q., XU, Z., PHILIP, J., BI, S., SHU, Z., SUNKAVALLI, K. & NEUMANN, U. (2022), Point-NeRF: Point-based Neural Radiance Fields (arXiv:2201.08845). arXiv. http://arxiv.org/abs/2201.08845 (20.12.2022).