

Neural Implicit Representation in Modern Computer Graphics

Goldi Soni

Assistant Professor

Amity University
Raipur, Chhattisgarh

gsoni@rpr.amity.edu

K. Jaianjan

B.Tech CSE

Amity University
Raipur, Chhattisgarh

k.jaianjan@s.amity.edu

Kshitij Puri Goswami

B.Tech CSE

Amity University
Raipur, Chhattisgarh

kshitij.goswami@s.amity.edu

Abstract

Recently, neural implicit representations gained popularity in computer graphics as an effective approach to modelling and rendering of various 3D scenes. Contrary to typical representations like polygon meshes, volumetric voxels or point clouds, neural implicit representations define scene geometry, its visual appearance and properties using continuous function approximated with a neural network. In this paper, we consider some of the most popular neural implicit representations used nowadays, which include Neural Radiance Fields (NeRF), Deep Signed Distance Functions (DeepSDF), SIREN networks, Gaussian splatting, Instant Neural Graphics Primitives, and NeRF extensions. All of these neural implicit methods enable novel view synthesis and efficient 3D reconstruction from images. The analysis will include an exploration of what particular technique allows better quality, reduced rendering time, better scalability or dealing with such real-life challenges as dynamic scenes, reflective surfaces or rendering large outdoor environments. At the same time, the paper will consider other related ideas like hybrid representations in Plenoxels, using tensor decompositions or neural implicit representations combined with voxel grids. By comparing different scientific papers dedicated to neural implicit representations, it is possible to identify their advantages and disadvantages as well as the ways to overcome possible difficulties in the future.

Keywords

Neural Implicit Representation, NeRF (Neural Radiance Fields), 3D Reconstruction, View Synthesis, Volumetric Rendering

1.Introduction

The field of computer graphics has seen a lot of development in recent years, resulting in sophisticated virtual worlds that are applicable for use in video games, virtual reality, animation, and even scientific visualization. Conventional scene representations include explicit models such as polygon meshes, voxelization, and point clouds. While such techniques are common and effective, they suffer from some issues connected with their memory usage, scalability, and ability to model high-resolution geometry and lighting.

This is when neural implicit representations come into play. The novel paradigm implies training neural networks on a continuous function representing 3D scenes. Rather than describing explicit geometry, implicit neural techniques enable one to encode all scene features – shapes, colors, and lighting – as neural network weights, which makes it possible to obtain arbitrarily high resolution models with complex details.

One of the most important techniques in this field is called Neural Radiance Fields (NeRF), in which an object is represented using a continuous function that creates realistic pictures of the scene from any viewpoint. Since NeRF's creation, there have been many developments and innovations in this technology to improve its efficiency, performance, and effectiveness. Some notable technologies in this area include DeepSDF, SIREN, Instant Neural Graphics Primitives (Instant-NGP), Gaussian Splatting, and Mip-NeRF.

The purpose of this literature review is to look into some of the main breakthroughs made in the field of neural implicit representations of modern computer graphics. The analysis will cover different techniques and ideas introduced recently and explain their goals and importance.

2.Objective

The purpose of this review article is to discuss recent developments in the domain of neural implicit representations in computer graphics. Specifically, it is intended to discuss the modeling, reconstruction, and rendering of 3D scenes with the help of neural networks. This paper will evaluate various techniques that use neural implicit representations, including NeRF, DeepSDF, SIREN, Gaussian Splatting, and Instant Neural Graphics Primitives. In addition, this paper aims to discuss the merits and drawbacks of each technique and their applicability to certain tasks such as view synthesis and 3D scene reconstruction. Further, this paper aims to examine the relative efficiency of neural implicit representations when used with neural networks and identify various challenges associated with the utilization of the same.

3.Literature Review

The current review paper explores 30 research articles on neural implicit representations in contemporary computer graphics. Some of the methods considered in the paper include Neural Radiance Fields (NeRF), DeepSDF, SIREN, Gaussian Splats, Instant Neural Graphics Primitives, as well as other recent neural rendering methods. It should be noted that most of the reviewed research papers prove the effectiveness of such neural methods in terms of rendering quality, 3D reconstruction accuracy, and view synthesis as opposed to more conventional techniques based on meshes and voxels. Nevertheless, it is worth mentioning that neural techniques tend to be very time-consuming and data-intensive, which is another challenge the research addresses.

The contribution of NeRF lies in the introduction of a novel technique for synthesizing photorealistic images using a neural network model of a 3D scene given only a small number of input images. The motivation behind NeRF was to address the limitations of conventional 3D representations by modeling the scenes as continuous functions, which could represent the geometric shape and view-dependent properties of the scene. In the NeRF model, an MLP network takes the 3D coordinates and viewing direction as input and generates the output as color and density at each point. Volumetric rendering is used to generate rays through the pixels of an image and sample the scene to obtain the final output image. The training of NeRF involved the comparison of generated images with ground truth images. [1]

The DeepSDF framework suggests a new approach to encoding 3D shapes via neural networks based on continuous signed distance functions. The idea behind the paper was to design a novel and versatile method of representing shapes, which could surpass the disadvantages inherent in the traditional voxel grid and mesh methods. It includes a neural network with an encoder that accepts 3D points and a latent vector of shapes as inputs and computes the corresponding signed distance to the surface of the object. This neural net is trained with points sampled in the vicinity of various shapes and the respective values of distances to their surfaces. Auto-decoder architecture is used to optimize latent vectors for each shape individually. [2]

SIREN solves the issue of the inability of implicit neural representations in representing high-frequency components and their derivatives. It was aimed at developing an architecture of a neural network that would enable accurate representation of complex signals and solving differential equations. To achieve that, SIREN utilized periodic sine activation functions for all layers of the network instead of the typical ones such as ReLU. It is possible due to the periodicity of the sine function that allows a more accurate representation of the high-frequency components and their derivatives. In addition, a specific way of setting weights helps train the network without instability. Experiments proved that SIREN could successfully represent images, sounds, and even three-dimensional objects with high precision and solve physical equations like Poisson and wave equations. [3]

The goal of this study was to address the limitation of the inefficient rendering and training times associated with NeRF-based techniques. The researchers suggested a new representation for scenes with millions of 3D

Gaussians, rather than using neural networks. Every Gaussian contains parameters such as its location, size, color, and opacity. The system first starts by creating a sparse point cloud and then fine-tunes the Gaussian parameters via gradient descent. For the visualization, the model used a quick rasterization step that projects the Gaussians into an image space and blends them effectively. An adaptive density control technique helps add or remove Gaussians to increase the resolution when necessary. This method allowed rendering at a rate higher than 100 FPS while preserving good visual quality. This paper won the Best Paper award at SIGGRAPH 2023. [4]

The Mip-NeRF model was invented as a solution to the problem of aliasing that existed in the original NeRF algorithm for scenes rendered at various scales. The goal here was to design a scale-invariant scene representation with consistent image quality at any scale, i.e., zooming in/out. Unlike the NeRF approach, where one sampled points from along a ray, Mip-NeRF used conical frustrums for each pixel and performed positional encoding within this space. By doing so, Mip-NeRF managed to estimate the size of each point and prevent undersampling or oversampling during rendering. Since Mip-NeRF employed integrated positional encoding, it was innately anti-aliased. [5]

The Instant-NGP paper centers around the goal of significantly speeding up neural graphics primitives such as NeRF and signed distance functions. The authors aimed at bringing down the training time of these models from hours to seconds without compromising on the quality of the result. This is done by employing an encoding mechanism that creates feature vectors from spatial positions in hash tables. These features are then passed through a small neural network that gives output such as color or density values based on inputs provided to it. This multiresolution feature of Instant-NGP helped it achieve real-time training as well as neural rendering of these models, which was not possible before. [6]

In Occupancy Networks, the idea was to model 3D shapes as a continuous function predicting the presence or absence of a point in the geometry within a shape. The intention was to solve the problems faced by voxel representations due to limited memory resources. An architecture based on a neural network which takes inputs in the form of three-dimensional coordinate values along with the latent code generated through the encoder network using inputs like images and point clouds was proposed. The output is the occupancy probability value indicating the occupancy of that particular point. The surface will be formed by the decision boundary. [7]

The objective of Neural Volumes is to train on the representation of dynamic scene volumes directly from 2D images without the need for direct supervision with 3D volumes. The technique involves training a neural network to output a 3D voxel grid that encodes colors and densities within the volume. Through differentiable volumetric rendering, this volume is projected onto images. Training proceeds using an image reconstruction loss alone, allowing for efficient training of the network. Temporal encoding helps the model to encode dynamic scenes as well. The technique makes it possible for the network to learn a renderable representation of dynamic objects such as people. [8]

Plenoxels attempts to determine whether neural networks are needed in radiance field rendering by using a sparse voxel grid in place of neural networks to store density information and spherical harmonic coefficients indicating view-dependent colors within each voxel. The parameters contained in the voxels are optimized through gradient descent with respect to a rendering loss function. In addition, the use of sparse voxel grids saves memory space, while spherical harmonics are capable of capturing the directional appearance of the object. Ray marching is used to render the scene, and it is almost identical to the process used in NeRF except that there is no neural network inference involved. [9]

NeRF-W is an extension of the original NeRF algorithm that takes into account the changes due to dynamic light and the appearance of transient objects in the photos of real-life scenes. In other words, the goal was to extend neural radiance fields for realistic situations when photographs are taken under different conditions at

different points in time. The method involves separating the scene into two parts – the static scene and transient objects, while introducing appearance embeddings that allow modeling lighting and weather changes. At the training stage, the network manages to differentiate between constant and transient objects. [10]

Point-NeRF introduces a novel representation for the neural radiance field using a point-based approach, aiming to make the process more efficient. It was intended to minimize the amount of computations that were used when querying an empty space in the case of conventional NeRF approaches. A point cloud was generated where a feature vector was stored for each point. Only neighboring points, and not the whole space, were considered during ray intersection. Features were then passed through a tiny neural network, which estimated colors and densities. Using the geometry information provided by a point cloud proved to be highly beneficial. [11]

The Mip-NeRF 360 model is an improvement on the Mip-NeRF model that addresses the problem of modeling and rendering outdoors unbounded scenes. It aims at modeling unbounded backgrounds and geometries such as landscapes and cityscapes. Mip-NeRF 360 applies the technique of spatial contraction whereby infinite space is contracted into a finite space representation. In addition, it applies proposal networks that are used to predict the most appropriate sample points along rays to render a more efficient image. Combining this technique with the anti-aliasing technique of Mip-NeRF results in high-quality images from both nearby and faraway areas. [12]

The paper proposes an efficient radiance field representation scheme based on tensor factorization called TensorRF. The purpose of this study was to minimize the computational complexity and memory consumption in comparison with classic NeRF models. Rather than employing a vast neural network architecture, the radiance field is decomposed into several tensors containing all the necessary information about the scene. While these tensors contain spatial data, a smaller neural decoder transforms them into color information. The main advantage of the proposed model is that the number of parameters is reduced dramatically, which leads to faster training and rendering. [13]

Neural Radiance Caching is an effort to speed up global illumination rendering for path tracing. The idea behind it was to reuse previously calculated lighting information through neural networks to cut down costs on computations. The technique relies on training the neural network during rendering in order to make predictions about the outgoing radiance for each point of the scene depending on spatial coordinates and direction. Rather than calculating costly light transport for every ray, the network calculates results by sampling the cache. As more samples appear, the network improves its predictions, and the algorithm becomes much faster without losing visual fidelity. [14]

The NSVF approach involves combining sparse voxels with radiance fields using deep learning to enhance efficiency. In particular, this work sought to ensure that there was no unnecessary computation done when rendering points located in an unoccupied area in space. This involved breaking down the scene into a sparse voxel octree with each voxel having a latent feature vector. During query, the coordinate values for the queried point alongside the features of the neighboring voxels were then fed into a neural network to compute the density and color values. NSVF performed better than NeRF in terms of training and rendering times. [15]

In this paper, we analyze how difficult it is for neural networks to learn functions with high frequencies. The aim of the paper was to address the problem of spectral bias, which made neural networks learn lower frequency features first. The solution involved using Fourier feature mappings of input coordinates before applying them to the neural network. It allowed the transformation of the coordinate representation to multiple sine and cosine functions that enable capturing finer structures. The approach helped to increase the performance of implicit neural representations, like NeRF. Fourier feature encoding is currently a popular practice for neural networks. [16]

The idea behind KiloNeRF was to make use of a huge number of small neural networks to speed up NeRF rendering. The goal was to replace the single big NeRF network with a great number of small specialized

networks each taking care of its own area. Every network learned scene data for its particular spatial cell. At the time of rendering, only appropriate neural networks were considered for the points on a ray. Due to the fact that neural networks were small, they could fit into the cache of a graphics processing unit easily. KiloNeRF was trained through distillation of a whole NeRF network. [17]

The goal of Ref-NeRF is to enhance the realism of reflection properties in reflective surfaces. The aim was to improve on the previous NeRF by modeling the reflection of surfaces based on their views. In order to predict the reflection direction, Ref-NeRF first predicts the surface normals. Appearance of the materials is predicted through the use of spherical harmonics and the roughness parameter that regulates the sharpness of the reflections. This representation ensures that the network can better simulate mirror-like reflection effects. [18]

The DVGO algorithm suggests optimizing voxel grids directly without the need to use neural networks for the reconstruction of the radiance field. It aimed to achieve extremely fast training without compromising rendering quality. The idea is to store colors and densities in a dense voxel grid directly and optimize them using gradient descent based on the loss incurred when rendering. The algorithm uses progressive refinement and regularization to enhance the process. Such a simple yet powerful method helps reconstruct scenes within minutes, not hours. DVGO showed that basic optimization strategies paired with GPU technology can rival neural network methods in radiance field reconstruction. [19]

The goal of NeX is to generate novel views in real-time without extensive computations on neural networks. NeX learns a set of basis images for modeling the appearance and combines these according to the viewing directions. Novel views are rendered by performing lookups and blending of textures using common graphics hardware. This approach does not require neural computations for rendering, resulting in an extremely fast process. NeX managed to render frames at more than 1000 fps on commercial graphics cards. [20]

SRN aims at learning the representation of 3D scenes based on the guidance of 2D images alone. The idea was to design neural networks capable of comprehending the geometry of the scene without the need for any direct 3D information. This is achieved using an encoder responsible for deriving the latent vector from input images and a decoder that converts 3D positions and view vectors into colors. Training is facilitated through differentiable rendering by matching the generated images with real-world images. The network indirectly learns the intrinsic 3D structure of the scene. [21]

Pi-GAN presents an approach where the generative adversarial networks (GANs) generate images which are realistic, of high resolution, and consistent in 3D. This work was carried out with an aim to develop realistic objects that would be consistent when viewed from different perspectives. Pi-GAN works based on implicit neural representations and employs periodic activation functions like those used in SIRENs. Here, the generator generates a 3D representation, while volume rendering produces the final image. The discriminator judges whether the image is realistic and consistent from a multi-perspective viewpoint. Pi-GAN trained on 2D image datasets to produce 3D object priors. [22]

The NeRF++ model addresses the shortcomings of the base NeRF model by extending its application in large outdoor environments with distant backgrounds. The goal of the model was to solve the problem of unbounded spaces, where both objects and backgrounds were infinite. NeRF++ decomposes the modeling of scenes into foreground and background parts. The former utilizes standard Cartesian coordinates, while the latter uses an inverted sphere coordinate system. This makes it possible to represent distant regions efficiently without consuming unnecessary computational resources of the network. [23]

The present research presents a level-of-detail mechanism to be used with neural implicit surfaces that allows real-time rendering. The goal was to make neural representations usable for graphics applications that require real-time processing, including games and virtual reality. Neural representations are organized in an octree format, where every node stores a tiny neural network describing the local geometry. When it comes to rendering, levels of detail that best fit the current visual requirements are chosen according to the error estimates in the screen space. More distant locations will be described by simple networks, while closer ones – more complicated. [24]

The goal of Urban Radiance Fields is to create an approach that would help to model large city scenes with neural radiance fields. The need for such an approach arose due to difficulties associated with such aspects as large scene scale, lighting variations, and the presence of dynamic elements in the scene. To address this problem, the authors used the idea of semantic segmentation and created multiple neural networks corresponding to each semantic element. Appearance encodings are used to model lighting variations between images. LiDAR information can also be considered. [25]

The Block-NeRF model solves the problem of scaling NeRF up to represent very large scenes, for instance, city blocks. In this approach, the scene is divided into overlapping blocks and trained as individual NeRF models. During the rendering process, the corresponding blocks are merged to form the output image. Appearance embeddings are utilized to ensure coherence among the blocks that have been acquired under varying lightings. Large scenes can be trained incrementally using the proposed technique. [26]

Accelerated training of neural implicit representations through meta-learning is suggested in this research work. The intention here was to decrease the amount of time needed for the training of a model in new scenes. In this regard, gradient-based meta-learning is used to train initial network weights that can converge quickly for new tasks. While meta-training, a set of optimal initial weights is obtained for quick fine-tuning in new scenes. This prior helps in generalizing the knowledge about scene structures. [27]

CodeNeRF is an extension of NeRF which learns category-specific representations such that there would be separation between the shape and appearance of the objects. The goal was to provide a way for controlled object generation and manipulation in a category. This is achieved by using two latent codes which represent geometry and appearance. Both the codes together with the spatial coordinates are used as inputs to a neural network that predicts the color and density. By blending these codes in different ways, novel object variations can be created. [28]

The method NeRF-VAE combines neural radiance fields and a variational autoencoder in order to develop generative 3D scene models. The goal of this project is to build a probability distribution for scenes, which will allow them to generate new scenes. The encoder transforms the images into latent codes that correspond to scenes. In turn, the generator renders images by using a NeRF-based approach with the help of latent scene codes. The probabilistic model NeRF-VAE enables the sampling of scenes according to their distribution. In addition, NeRF-VAE can decompose scenes into different objects. [29]

Instant-NGP attempts to speed up significantly both the learning phase and rendering of NeRF, which usually takes several hours. The central goal of the research was making NeRF applicable to real-time purposes while preserving its quality during scene rendering. In order to achieve this goal, Instant-NGP applies a multiresolution hash encoder, which maps 3D coordinates to their feature representations kept in hash tables of different resolutions. After encoding, the resulting representation is passed through a smaller neural network, predicting color and density for each 3D point in a scene. This approach allowed researchers to simplify the structure of neural networks used and to perform computations more efficiently without losing the ability to accurately capture the details of the scene. The system is also optimized through the use of highly efficient GPU implementation and CUDA kernels. The proposed approach decreased NeRF learning time down to mere seconds and provided support for real-time scene rendering. [30]

As seen from the analysis provided above, there is definitely a move away from classical and explicit representations, namely, meshes, voxels, and point clouds towards neural implicit representation in the field of computer graphics. These classical techniques were hampered by poor memory efficiency, scalability, and difficulty with reproducing small geometric details and lighting effects. Neural implicit representations, in turn, provide continuous scene representation, automatic geometry and texture learning, and improved rendering and 3D scene reconstruction capabilities.

Nonetheless, there are some limitations and challenges to overcome in this field. The major problem is that most current neural implicit approaches rely on large amounts of data, are computationally expensive and time-consuming, which makes these techniques unsuitable for real-time use cases and application in low-power

devices. Handling complex scenarios, including dynamics, large scales, and generalization in different environments remain a challenge too.

4. Comparison of past published research works

The below table provides an overview of the advancements in the field of implicit representation via neural networks for computer graphics applications between 2019 and 2023. During this time frame, research efforts have evolved from basic neural network approaches like DeepSDF and Neural Radiance Fields (NeRF) to sophisticated and optimized solutions such as Instant Neural Graphics Primitives, tensor representations, and Gaussian Splats.

Sl No.	Title	Year of Publication	Proposed Objective	Methodology	Result
1	NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis	2020	For producing photorealistic novel views of complicated 3D environments using only a few photos.	Uses neural networks (Multilayer Perceptron or MLP) to generate a continuous scene representation that returns color and density for each point in 3D space and viewing directions with a volumetric rendering technique.	Created highly realistic view synthesis and introduced neural radiance fields as a novel representation in 3D scenes.
2	DeepSDF: Learning Continuous Signed Distance Functions for Shape Representation	2019	For constructing a small but continuous neural network representation to represent intricate 3D forms.	Employs neural networks to learn the signed distance function (SDF), which maps 3D coordinates and latent shape codes to distances to achieve continuous shape representations.	Constructed 3D models with very few parameters than existing approaches based on voxel representation.
3	SIREN: Implicit Neural Representations with Periodic Activation Functions	2020	For enhancing neural implicit representations to allow capturing high-frequency information and derivative information about signals.	Includes sine activations within the neural networks, along with a specific initialization strategy to encode complex signals.	Accurately represented small-scale structures in pictures, sound, and 3D objects by solving differential equations.
4	3D Gaussian Splatting for Real-Time Radiance Field Rendering	2023	For enabling real-time ray tracing of radiance fields without sacrificing visual quality.	Describes scenes by optimizing millions of 3D Gaussian elements in the scene, each with properties like position, color, and opacity, and then performs rasterization rendering.	Allowed rendering more than 100 frames per second and received SIGGRAPH 2023 Best Paper Award.
5	Instant Neural Graphics Primitives (Instant-NGP)	2022	For decreasing the training and ray tracing time of neural	Uses multi-resolution hash encoding with a tiny neural network and GPU	Accelerated NeRF training from hours to mere seconds without

SI No.	Title	Year of Publication	Proposed Objective	Methodology	Result
			graphics primitives like NeRF.	optimizations for efficient scene representation.	compromising on the quality of rendering.

Table 1:- Comparison of past published research works (2019 to 2023)

5. Conclusion

Neural implicit representations have been found to transform modern computer graphics by implementing 3D scenes with neural networks. Contrary to the previous techniques which relied on mesh, voxel, or point cloud geometry, neural implicit representations have managed to capture geometry, colors, and light as continuous functions. Methods such as NeRF, DeepSDF, SIREN, Instant Neural Graphics Primitives, and Gaussian Splatting have demonstrated exceptional results in novel view synthesis, 3D reconstruction, and photorealistic rendering despite working with minimal data.

Recently, efforts have been directed towards enhancing the efficiency, scalability, and speed of rendering through multiresolution hash encoding, tensor decomposition, sparse voxelization, and Gaussian encoding techniques. Neural rendering is gaining momentum due to these developments in rendering speed and efficiency. They make it possible to implement real-time applications like virtual reality, gaming, and visualization.

Nevertheless, there are still several challenges in terms of computation, slow training times, and difficulties when dealing with dynamic scenes and large environments. It is recommended to focus on designing efficient neural network models, scalable solutions, and hybrid frameworks combining traditional computer graphics and machine learning.

In summary, neural implicit representations stand out as revolutionary techniques for flexible, compact, and realistic modeling of 3D scenes for future graphics engines.

6. Future Scope

The future trends in neural implicit representation modeling include the development of approaches to making the process of learning more efficient by decreasing training times and costs while retaining the quality of rendered images. Scientists will strive to develop scalable algorithms that would allow them to render large-scale environments, e.g., city-level scenes or outdoor scenes. The other important area of research includes the development of real-time capabilities of neural rendering to facilitate the creation of applications in the areas of VR/AR or gaming. The integration of neural implicit representations into traditional computer graphics techniques is also one of the potential directions. It is essential to address the challenges associated with the inclusion of the dynamics into the scene in the context of neural rendering. It is important to improve the data efficiency of models, which means that it should become possible to obtain highly realistic reconstructions based on a relatively small number of training images. Finally, scientists will try to apply their findings to solving real-world problems, including such areas as autonomous driving, digital twins, robotics, and scientific visualization.

References

1. 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*. In Proceedings of the European Conference on Computer Vision (ECCV).
2. Park, J. J., Florence, P., Straub, J., Newcombe, R., & Lovegrove, S. (2019). *DeepSDF: Learning continuous signed distance functions for shape representation*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
3. Sitzmann, V., Martel, J. N. P., Bergman, A. W., Lindell, D. B., & Wetzstein, G. (2020). *Implicit neural representations with periodic activation functions*. In Advances in Neural Information Processing Systems (NeurIPS).
4. Kerbl, B., Kopanas, G., Leimkühler, T., & Drettakis, G. (2023). *3D Gaussian splatting for real-time radiance field rendering*. ACM SIGGRAPH Conference Proceedings.
5. Müller, T., Evans, A., Schied, C., & Keller, A. (2022). *Instant neural graphics primitives with a multiresolution hash encoding*. ACM SIGGRAPH Conference Proceedings.
6. Barron, J. T., Mildenhall, B., Tancik, M., Hedman, P., Martin-Brualla, R., & Srinivasan, P. P. (2021). *Mip-NeRF: A multiscale representation for anti-aliasing neural radiance fields*. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).
7. Mescheder, L., Oechsle, M., Niemeyer, M., Nowozin, S., & Geiger, A. (2019). *Occupancy networks: Learning 3D reconstruction in function space*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
8. Lombardi, S., Simon, T., Saragih, J., Schwartz, G., Lehrmann, A., & Sheikh, Y. (2019). *Neural volumes: Learning dynamic renderable volumes from images*. ACM Transactions on Graphics.
9. Fridovich-Keil, S., Yu, A., Tancik, M., Chen, Q., Recht, B., & Kanazawa, A. (2022). *Plenoxels: Radiance fields without neural networks*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
10. 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*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
11. Xu, Q., Xu, Z., Philip, J., Bi, S., Shu, Z., Sunkavalli, K., & Neumann, U. (2022). *Point-NeRF: Point-based neural radiance fields*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
12. Barron, J. T., Mildenhall, B., Verbin, D., Srinivasan, P. P., & Hedman, P. (2022). *Mip-NeRF 360: Unbounded anti-aliased neural radiance fields*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
13. Chen, A., Xu, Z., Geiger, A., Yu, J., & Su, H. (2022). *TensorRF: Tensorial radiance fields*. In European Conference on Computer Vision (ECCV).
14. Müller, T., Rousselle, F., Novák, J., & Keller, A. (2021). *Neural radiance caching for real-time global illumination*. ACM Transactions on Graphics.
15. Liu, L., Gu, J., Lin, K. Z., Chua, T. S., & Theobalt, C. (2020). *Neural sparse voxel fields*. In Advances in Neural Information Processing Systems (NeurIPS).
16. Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., & Ng, R. (2020). *Fourier features let networks learn high frequency functions in low dimensional domains*. In Advances in Neural Information Processing Systems (NeurIPS).
17. Reiser, C., Peng, S., Liao, Y., & Geiger, A. (2021). *KiloNeRF: Speeding up neural radiance fields with thousands of tiny MLPs*. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).
18. Verbin, D., Hedman, P., Mildenhall, B., Zickler, T., Barron, J. T., & Srinivasan, P. P. (2022). *Ref-NeRF: Structured view-dependent appearance for neural radiance fields*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
19. Sun, C., Sun, M., & Chen, H. T. (2022). *Direct voxel grid optimization: Super-fast convergence for radiance fields reconstruction*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).

20. Wizaradwongsa, S., Phongthawee, P., Yenphraphai, J., & Suwajanakorn, S. (2021). *NeX: Real-time view synthesis with neural basis expansion*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
21. Sitzmann, V., Zollhöfer, M., & Wetzstein, G. (2019). *Scene representation networks: Continuous 3D-structure-aware neural scene representations*. In Advances in Neural Information Processing Systems (NeurIPS).
22. Chan, E. R., Monteiro, M., Kellnhofer, P., Wu, J., & Wetzstein, G. (2021). *Pi-GAN: Periodic implicit generative adversarial networks for 3D-aware image synthesis*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
23. Zhang, K., Riegler, G., Snively, N., & Koltun, V. (2020). *NeRF++: Analyzing and improving neural radiance fields*. arXiv preprint arXiv:2010.07492.
24. Takikawa, T., Litalien, J., Yin, K., Kreis, K., Loop, C., Nowrouzezahrai, D., Jacobson, A., McGuire, M., & Fidler, S. (2021). *Neural geometric level of detail: Real-time rendering with implicit 3D shapes*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
25. Rematas, K., Liu, A., Srinivasan, P. P., Barron, J. T., Tagliasacchi, A., Funkhouser, T., & Ferrari, V. (2022). *Urban radiance fields*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
26. Tancik, M., Casser, V., Yan, X., Pradhan, S., Mildenhall, B., Srinivasan, P. P., Barron, J. T., & Kretzschmar, H. (2022). *Block-NeRF: Scalable large scene neural view synthesis*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
27. Tancik, M., Mildenhall, B., Wang, T., Schmidt, D., Srinivasan, P. P., Barron, J. T., & Ng, R. (2021). *Learned initializations for optimizing coordinate-based neural representations*. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
28. Jang, W., & Agapito, L. (2021). *CodeNeRF: Disentangled neural radiance fields for object categories*. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV).
29. Kosiorek, A. R., Strathmann, H., Zoran, D., Moreno, P., Schneider, R., Mokrá, S., & Rezende, D. J. (2021). *NeRF-VAE: A geometry-aware 3D scene generative model*. In International Conference on Machine Learning (ICML).
30. Müller, T., Evans, A., Schied, C., & Keller, A. (2022). *Instant-NGP: Instant neural graphics primitives*. ACM SIGGRAPH.