

# A REVIEW OF NEURAL RADIANCE FIELDS AND GAUSSIAN SPLATTING TECHNIQUES FOR SCENE REPRESENTATION AND RENDERING

Samarth Joshi<sup>1</sup>, Punit Kumar<sup>2</sup>

<sup>1</sup>Student, Artificial Intelligence and Data Science Poornima Institute of Engineering and Technology Jaipur, India  
0009-0000-8445-1113

<sup>2</sup>Assistant Professor, Artificial Intelligence & Data Science Poornima Institute of Engineering and Technology  
Jaipur, India.

[punit.kumawat@poornima.org](mailto:punit.kumawat@poornima.org)

DOI: <https://www.doi.org/10.58257/IJPREMS37523>

## ABSTRACT

This paper reviews advances of the representation and rendering of 3D scenes emphasizing methodologies proposed by Neural Radiance Fields (NeRF) and Gaussian splatting methods. We review earlier works that established foundational methodologies: first, NeRF's neural-based volumetric rendering and subsequent work on Gaussian splatting for real-time applications. We also look further at segmentation optimization introduced with FlashSplat. General features, weaknesses, as well as discussion on the research gaps of these approaches will guide further exploration towards scalable, efficient, and dynamic rendering solutions.

**Keywords**—Gaussian splatting, 3D scenes, volumetric rendering, NeRF weaknesses, dynamic rendering solutions.

## 1. INTRODUCTION

Representation and rendering of scenes has been two major challenges in computer vision and computer graphics, extending into broad applications including virtual reality, augmented reality, gaming, and photorealistic content creation. Traditional techniques-triangle meshes and voxel-based-have been employed to achieve 3D modeling and rendering for many years. The problem with these methods is that they are usually manually designed, extensively computationally demanding, and also inaccurate, especially for complex scenes. This has transformed the face of deep learning, which in return has brought forth new images to represent scenes and render data with models.

Neural Radiance Fields represent a new frontier in the field, substituting neural networks for encoding scenes as volumetric radiance fields. This 2020 concept was known as NeRF, first introduced by Mildenhall et al. NeRFs showed exceptional skill in synthesizing photorealistic novel views of static 3D scenes. NeRF optimized a design for a neural network which would correlate spatial coordinates and viewing directions with radiance and density values. Doing so achieved results previously impossible through conventional graphics pipelines. However, their success came at the cost of extreme computational inefficiency. Training and inference times for NeRF take unacceptable lengths of time for real-time applications or wide-scale deployments. Moreover, its dependence on fixed representations constrains its capacity to adjust to dynamic or interactive environments, thereby highlighting the necessity for more adaptable alternatives.

Because of these limitations, improvements in efficiency and the applicability to a wide variety of situations have been proposed by alternative techniques. Gaussian splatting is one notable development, proposed by Müller et al. This technique differs from the paradigm of neural networks, since it represents scenes as sets of 3-D Gaussian distributions. Requiring only the mathematical simplicity and low-cost computations of Gaussian primitives, real-time performance is achieved without visual sacrifice. Gaussian splatting therefore opens the door to interactivity, with low-latency rendering at the core. But, similar to NeRF, however, this comes with significant challenges in scaling up to dynamic environments and memory usage optimization in more complex scenes.

The second development is FlashSplat, which takes the concept of Gaussian splatting one step further through the integration of an efficient segmentation requirement in 3D environments. This allows for more accurate segmentation by aligning masks in 2D with Gaussian primitives in 3D, improving the efficiency in rendering as well as expounding the capabilities of Gaussian splatting by adding element concepts such as object isolation, scene editing, and mixed reality integration. Despite its contributions, FlashSplat's reliance on 2D data highlights the ongoing challenge of bridging 2D and 3D representation techniques effectively.

This work aims to integrate the contributions of NeRF, Gaussian splatting, and FlashSplat by discussing their respective methods, respective relative advantages, and limitations. We describe how this body of works collectively accelerated the advances in the representation and rendering of 3D scenes, while pointing out research questions that

have to be tackled to reach fully scalable, dynamic, and efficient solutions. Through an integrated analysis of these methods from an integrated perspective, we hope to provide a holistic insight into the current developments and open up future directions in this fast-changing field.

## 2. OVERVIEW OF RECENT LITERATURE

### NeRF: Environments as Neural Radiance Fields

The work introduced by Mildenhall et al. makes a paradigm shift in the synthesis of novel views for static 3D scenes with the help of a fully connected neural network that encodes volumetric data as a continuous radiance field mapping spatial coordinates  $(x,y,z)$  and viewing directions  $(\theta,\phi)$  into density and RGB color values making NeRF to produce high fidelity renders with good details and physical lighting effects.

The training process contains the optimization of network parameters by using a large dataset of 2D images captured from various viewpoints. The major advancement NeRF has made involves positional encoding: it captures high-frequency details in a fine-grained and neat way by mapping input coordinates onto a higher-dimensional space. Such an approach actually relaxes constraints that neural networks face when trying to represent complex spatial variations. The several limitations this model has make the performance of NeRF hindered. However, typically hours or days are needed to train NeRF for a single scene because of the aforementioned reasons of a large number of ray samples required for optimization. Moreover, inference is slow since rendering one frame involves tracing thousands of rays through the radiance field and performing neural network evaluations for each sample.

Additionally, the static nature of NeRF's representation renders it inappropriate for dynamic or interactive applications, thereby limiting its applicability in contexts such as real-time virtual environments or extensive reconstructions.

### 3D Gaussian Splatting for Real-Time Radiance Field Rendering

While NeRF relies on the neural network, 3D Gaussian splatting, proposed by Müller et al., represents scenes as an aggregation of distributions in space. As for each of the Gaussians, position, orientation, size, and colour, enable to efficiently model complex geometries, this enables the modelling of complex geometries and visual attributes of scenes. The rendering procedure now reads: project these Gaussians onto the image plane through a differentiable splatting methodology that integrates their contributions.

Gaussian splatting exhibits superior computational efficiency, allowing the technique to run in real time on current GPUs. This efficiency is largely due to a reduced number of neural network evaluations, whose place is taken over by computationally efficient mathematical operations. Additionally, this approach allows for accelerated optimization by directly fine-tuning Gaussian primitive parameters rather than requiring training of a neural network.

It demonstrates extreme flexibility: for instance, capabilities for real-time scene editing, interactive navigation, and high-quality rendering. However, it relies on fixed-number Gaussian primitives, which can create potential memory constraints in large or highly detailed scenes. Also, like NeRF, Gaussian splatting assumes a static presentation of the scene, hence reducing its applicability in dynamically changing environments.

### FlashSplat: Optimal Segmentation for Gaussian Splatting

FlashSplat extends the Gaussian splatting framework to satisfy the desires of good and even accurate segmentation in the realm of 3D rendering. This method provides a new approach towards the segmentation of 3D Gaussians with the aid of 2D masks. Segmentation is modeled as a linear programming problem, thereby relating 2D image inputs to 3D Gaussians for the best possible results.

An important benefit of the FlashSplat approach is computational efficiency. It combines the ability to opportunistically include segmentation in the framework of Gaussian splatting without additional demands on resources typically associated with post-processing or independent segmentation models that are otherwise affecting practical applications concerning object isolation, interactive editing, and manipulation of mixed reality scenes.

In flashsplat, occlusions and noisy inputs are handled very robustly, so the model is very suitable for real-world applications. However, its dependency on a 2D mask limits its accuracy in volumetric segmentation tasks, especially when using 3D data that is sparse or incomplete. In future extensions, integration of 3D volumetric priors can be useful to further improve the accuracy and robustness achieved by the model.

## 3. ANALYSIS

These techniques studied, NeRF, 3D Gaussian splatting, and FlashSplat, represent great contributions to the field of three-dimensional scene representations and rendering. Each technique had significant challenges concerning propositions of a rather novel approach but exposed their very own shortcomings, pointing out potential further ways for research. NeRF presented a critical state-of-the-art in the community-demonstrating neural representations could achieve photorealism and beyond. Positional encoding and volumetric integration revolutionized the architecture of

novel view synthesis, but reliance on computationally expensive neural networks in both training and rendering limits its feasibility. The long training period and also not suitable for real-time applications that limit the application of NeRF in settings such as gaming and virtual reality. Furthermore, since NeRF produces static representations, it does not fit dynamic environments and this calls for models that adapt to changing situations.

Conversely, the utilization of 3D Gaussian splatting markedly diminishes computational demands by substituting the neural representation with Gaussian primitives. This approach illustrates that it is feasible to attain high-quality renderings without the substantial resource requirements typically associated with a neural network. By facilitating real-time performance, Gaussian splatting has created new avenues for interactive applications, encompassing virtual tours, augmented reality (AR), and scenarios involving live rendering. Nevertheless, this technique presents certain challenges, especially in terms of memory efficiency and scalability. Representing complex scenes by a small number of Gaussian primitives leads to artifacts or low resolution in fine details. Also, similar to NeRF, Gaussian splatting relies on the static scene structure, and dynamic scene modeling is another open research problem.

FlashSplat generalizes Gaussian splatting as a way to get efficient segmentation with the help of 2D masks and uses linear programming for optimal alignment between both 2D and 3D data to make the Gaussian splatting more flexible and useful. These capabilities to separate objects and efficiently manipulate scenes have particularly made FlashSplat extremely valuable for developing applications in mixed reality and for interactive scene editing.

However, the reliance of FlashSplat on 2D inputs limits it in volumetric segmentation since boundaries have to be defined sharply in 3D. Hence, the algorithm could be further improved by including some kind of volumetric priors or using machine learning techniques for better robustness in segmentation.

A prevalent constraint among these methodologies is their inherent static characteristic, which limits their utility in dynamic or extensive environments. Subsequent investigations should prioritize the creation of flexible frameworks that can effectively represent scenarios that change over time. Furthermore, enhancing memory efficiency and scalability will be essential for managing highly intricate scenes without sacrificing performance.

It turns out that the neural representations and geometric primitives of NeRF and Gaussian splatting do have a valid hybrid methodology. It seems it has the opportunity to take the benefits in the implementation that each provides and combine the scalability and efficiency of the approach in Gaussian splatting with the expressive ability of the neural networks. Other imminent upgrades in hardware, such as dedicated GPUs or neural accelerators, may also ease some of the computational drawbacks models such as NeRF are likely to inherit. In a nutshell, NeRF-to-Gaussian splatting and ultimately FlashSplat signal the move towards much more practical and efficient rendering solutions. The challenge remains open, however-about how to come up with an even more balanced solution that takes care of all quality, efficiency, scalability, and adaptability. The approaches developed here are all indicative of a need for further innovation in unlocking 3D scene representation and rendering technologies.

#### 4. CONCLUSION

A review of some of the quite recent advances in 3D scene representation and rendering has pronounced the high advancement made in that particular domain using methods such as NeRF, 3D Gaussian splatting, and FlashSplat. All these methods have contributed to the evolution of scene modeling through novel techniques for rendering high-quality images, reducing computation overhead, and enhancing adaptability in scenes.

However, they open up many areas that need more research and development work, particularly in regard to scalability and dynamic scene management and performance in real-time. NeRF is indeed a pioneering approach in neural scene representation that has greatly moved the existing bar for photorealistic rendering. It has brilliantly provided novel views with much detail and accuracy in a very efficient manner and has thus been thought of as an indispensable tool for various applications pertaining to computer vision and graphics. The main disadvantage of NeRF is that it has very heavy computational requirements, rendering it inapplicable for real-time applications or large dynamic environments. As such, the challenge moving forward is to preserve the high quality of NeRF rendering with even more efficient ways of reducing its computational cost-for example, through model optimization, compression, or parallel processing strategies.

The 3D Gaussian splatting is an important advancement in this area. As scene depiction now makes use of Gaussian distributions, computation can be improved significantly in order to reach for real-time performance without loss of rendering quality. Thus, these results will open up new interactive applications in the fields of virtual and augmented reality. However, like NeRF, such Gaussian splatting also presents specific challenges in terms of the scene representation and limited memory to store a number of Gaussian primitives. This method is difficult to handle dynamic changes of the scene and definitely not applicable to interactive real-world environments.

Work in this direction should then proceed on to improving the flexibility of Gaussian splatting by bringing in dynamic object tracking or adaptive scene representations that make for interactive, ever-changing environments. FlashSplat then extends the Gaussian splatting approach by adding optimal segmentation of 3D scenes with a 2D mask. It unlocks tremendous opportunities in mixed reality, scene editing, as well as object recognition applications due to its ability to precisely segment and manipulate objects within a scene. FlashSplat's unique contribution is the integration of linear programming for optimal segmentation that computationally efficiently aligns 2D input with 3D Gaussian primitives.

Still, as mentioned in the review, FlashSplat depends on 2D segmentation, which would not tolerate as challenging volumetric segmentation tasks. With that said, one assumes that integrating 3D segmentation methods, and even leveraging models of machine learning for-automatic segmentation, could enhance the method even more and make it even more robust.

Considering such contributions, it is very apparent that though NeRF, Gaussian splatting, and FlashSplat made immense headways in the 3D scene representation capability, this field is actually at its pending developing step. Real-time rendering of very minute and photorealistic 3D scenes while maintaining scale and efficiency is a great challenge. Perhaps the most promising direction for further work is hybrid models that combine the strengths of neural networks with geometric representations like Gaussian splatting. Hybrid systems hold the promise of utilizing the significant expressive power of neural networks to capture detailed features while geometric primitives serve to improve scalability and speed. Additional positive influences include continued hardware technology advancement, like the development of dedicated graphics processing units and neural processing units, which can reduce some of the computational limitations currently preventing real-time rendering capabilities. Coupling these hardware acceleration capabilities with advanced rendering algorithms will enable the next generation of scene representation technologies to be applied in a much more extensive spectrum than that provided now, which ranges from real-time games to complex simulations and virtual environments. In conclusion, the 3D scene representation field is revolutionizing rapidly with NeRF, Gaussian splatting, and FlashSplat. While these techniques have pointed towards rendering high-quality efficient and interactive scenes, some challenges still persist in terms of scalability, dynamic adaptation toward the changes in the scene, and finally achieving real-time performance. Further research will thus address all these limitations by injecting hybrids, improving memory efficiency, and incorporating dynamic scene modelling techniques into it. With continuing innovations both in algorithms and hardware, the dream of real-time photorealistic rendering of difficult 3D scenes appears now closer than ever, so that new possibilities of wide-ranging application may now become possible.

## 5. REFERENCES

- [1] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, Ren Ng. arXiv:2003.08934
- [2] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, George Drettakis. arXiv:2308.04079
- [3] Qiuhong Shen, Xingyi Yang, Xinchao Wang. arXiv:2409.08270
- [4] Haibo Yang, Yang Chen, Yingwei Pan, Ting Yao, Zhineng Chen, Chong-Wah Ngo, Tao Mei arXiv:2409.07452
- [5] Zeyu Cai, Duotun Wang, Yixun Liang, Zhijing Shao, Ying-Cong Chen, Xiaohang Zhan, Zeyu Wang arXiv:2409.05099
- [6] Flynn, J., Broxton, M., Debevec, P., DuVall, M., Fyffe, G., Overbeck, R., Snavely, N., Tucker, R CVPR 2019 (Deep View: view synthesis with learned gradient descent.)
- [7] Henzler, P., Mitra, N.J., Ritschel, T.CVPR 2020 (Learning a neural 3d texture space from 2d exemplars. )
- [8] Kara-Ali Aliev, Artem Sevastopolsky, Maria Kolos, Dmitry Ulyanov, and Victor Lempitsky. ECCV 2020 (Neural Point-Based Graphics. In Computer Vision )
- [9] Jonathan T. Barron, Ben Mildenhall, Dor Verbin, Pratul P. Srinivasan, and Peter Hedman. CVPR 2022 (Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. )
- [10] Markus Gross and Hanspeter (Eds) Pfister. Elsevier 2011 (Point-based graphics.)
- [11] Jeff P. Grossman and William J. Dally. Rendering Techniques 1998 (Point Sample Rendering)
- [12] Peter Hedman, Pratul P. Srinivasan, Ben Mildenhall, Jonathan T. Barron, and Paul Debevec. ICCV 2021 ( Baking Neural Radiance Fields for Real-Time View Synthesis.)