Proceedings of CONF-SEML 2025 Symposium: Intefrating AI into Software Engineering
DOI: 10.54254/2755-2721/156/2025.CH23277

A Comprehensive Review of Advancements and Evaluation Frameworks in AIGC for 3D Content Creation: Focusing on 3D Gaussian Splatting

Ran Xie

Hefei University of Technology, Hefei, China
237479085@qq.com

Abstract: In recent years, AIGC has made significant progress in various fields, particularly in the generation of 2D images. However, in the 3D domain, traditional 2D model training methods have not achieved ideal results due to the lack of sufficient high-quality datasets. To address this issue, methods such as utilizing 2D diffusion models as priors and emerging 3D Gaussian Splatting have demonstrated the significant potential of AIGC in the 3D field. With the rapid development of AIGC 3D, the related evaluation metrics have not yet been unified. Although quantitative evaluation of 3D content is challenging, establishing a standardized evaluation system is crucial for future research. This article summarizes the technological progress and evaluation systems in the AIGC 3D field, focusing particularly on the potential applications of 3D Gaussian point set technology, and discusses future development directions and the challenges they face.

Keywords: AIGC, 3D Gaussian Splatting, Evaluation metrics

1. Introduction

Artificial Intelligence Generated Content (AIGC) has witnessed substantial advancements and developments in recent years. An increasing number of enterprises and research areas have leveraged AIGC to tackle practical problems that were previously difficult to solve, achieving remarkable results. For instance, ChatGPT [1] which are generative models developed by OpenAI, have demonstrated huge success on texts generation field.

With the research of more advanced architectures and technologies, other generative content, especially on the image generation, has also made tremendous progress. In the 2D domain, particularly with the research of diffusion models, the speed and quality of 2D content generation have seen significant improvements.

Unlike the 2D, 3D field initially did not have a rich and high-quality dataset like the 2D field, the 3D models trained using the traditional 2D model training method did not perform well. To solve this problem, DreamFusion first proposed a method that use a 2D diffusion model as a prior to optimize a Neural Radiance Fields (NeRF) by employing the score distillation sampling (SDS) loss. Based on this method, with the goal of further improving the quality and geometric accuracy of 3D contents, many new models have been proposed and significant improvements have been achieved.

At the same time, newly developed technologies, especially the 3D Gaussian Splatting technology, can significantly shorten the model training speed while maintaining quality compared to

DOI: 10.54254/2755-2721/156/2025.CH23277

NeRF-based models, potentially becoming a new research direction for AIGC 3D. To address the limitations of single-model 3D generation, multi-model collaboration methods are also being steadily advanced. Models optimized for specific problems in specific scenarios are also gradually emerging. In summary, AIGC 3D is developing rapidly and attracting more and more researchers to delve deeper into it.

According to the research, surveys related to AIGC 3D have not emphasized the evaluation metrics for this field, Although it's difficult to quantify and evaluate three-dimensional content, with the rapid development of AIGC 3D, it is crucial to establish a standard evaluation system. At the same time, there is no detailed explanation for the application of 3D Gaussian Splatting technology, while research has proven that 3D Gaussian Splatting technology has great potential in the field of AIGC 3D and should also be given equal attention, helping practitioners better navigate the expanding research frontier.

In this survey, we make the following contributions. First, to the best of the authors' knowledge, we provide a comprehensive summary of the existing mainstream evaluation frameworks for AIGC in 3D. Second, we review the technological advancements in AIGC for 3D content creation, with a particular focus on the application of 3D Gaussian Splatting technology. Finally, we discuss several promising directions for future development, along with the associated challenges that need to be addressed.

The paper is organized as follows. Section 2 introduces the evaluation metrics for AIGC in 3D. Section 3 reviews the technological advancements in the AIGC 3D domain, focusing on 3D Gaussian Splatting. Section 4 examines the applications of AIGC in 3D content creation. Section 5 discusses the challenges and future development directions for AIGC in 3D. The final section concludes the survey.

2. 3D AIGC evaluation

A comprehensive and accurate evaluation system is essential for identifying issues within models and facilitating the adoption of appropriate solutions. Evaluation metrics play a crucial role in this process. In this section, we discuss several common evaluation methods used for AIGC in 3D content creation.

2.1. Qualitative comparison

Qualitative comparison is a fundamental and efficient approach to assess a model's performance. Typically, models are trained on the same datasets alongside conventional models in the field, with their results compared. In the current 3D generation domain, the focus is primarily on improving the quality and geometric accuracy of 3D content. Consequently, training results are generally evaluated based on these two key aspects. Qualitative comparisons allow for direct visual inspections of the generated outputs, providing insight into the model's strengths and weaknesses.

2.2. Quantitative comparison

Quantitative comparison, in contrast, involves measuring specific differences between a model and other models in predefined metrics. This approach offers a more precise way to demonstrate the advantages of a model. However, given that 3D content is significantly more complex than 2D data in terms of volume and structure, there is currently no comprehensive and standardized method for evaluating 3D content.

Nonetheless, some studies have demonstrated that certain evaluation techniques developed for 2D content, such as those used in Text-to-Image models, can be adapted for 3D evaluation. For instance, CLIP R-Precision has been utilized in models like DreamFusion, Dream3D [2], and RichDreamer [3]

DOI: 10.54254/2755-2721/156/2025.CH23277

to assess the alignment between generated images and text, offering a quantitative measure of the model's ability to generate content that corresponds accurately to a given text prompt. Similarly, CityDreamer [4] uses Fréchet Inception Distance (FID) to evaluate the quality of generated images, further highlighting the potential for applying 2D evaluation metrics to 3D content creation.

2.3. User study

User studies are a widely used evaluation method in 3D generation, as the ultimate goal of these models is to serve the needs of users. Collecting direct feedback from users provides valuable insights into the practical performance of the models.

User studies are typically conducted in two stages. First, offline testing is performed, where a group of participants is recruited to evaluate content generated by different models within the same scenario. Second, online surveys or questionnaires are distributed to reach a broader audience. The feedback from all participants is then aggregated to summarize user preferences and assess the strengths and weaknesses of the models. This method offers an essential user-centric perspective on model performance.

2.4. Ablation study

While the aforementioned evaluation methods are designed to compare the performance of a model against others, an ablation study focuses on understanding the impact of specific changes or improvements within a model. This type of study is crucial for model development, as it allows researchers to determine whether newly introduced methods or modifications genuinely enhance model performance.

An ablation study is akin to a controlled variable experiment. By systematically removing certain components or features that were designed to improve the model, researchers can observe changes in performance and evaluate whether these modifications are effective. This process not only helps to validate the proposed improvements but also highlights areas of the model that may need further refinement. Through ablation studies, researchers can identify both the strengths and limitations of their models, providing valuable insights for future work and guiding further advancements in the field.

3. 3D content generation methods based on 2D guidance

Given that current 3D datasets suffer from limitations in richness and quality, leveraging abundant 2D data to guide 3D content generation has become a widely adopted strategy. This section is divided into three parts: an introduction to the key technologies involved, a review of the research based on the NeRF + 2D diffusion method, and finally elaborate on the current research related to the application of 3D Gaussian Splatting to 3D content generation.

3.1. Overview of key technologies

Diffusion models [5]: Diffusion models refer to a class of generative techniques based on the Denoising Diffusion Probabilistic Model (DDPM) framework. DDPM trains a model to perform the reverse diffusion process - starting from a noisy signal and applying iterative denoising steps to recover the original data distribution. By learning to denoise across different noise levels, the model can generate new samples by starting with random noise and applying the reverse diffusion process.

NeRF: NeRF is a neural rendering technique that has gained significant attention for novel view synthesis in 3D. NeRF consists of two core components: a volumetric ray tracer and a multi-layer perceptron (MLP). While NeRF has been successfully used as a global representation in AIGC-3D

DOI: 10.54254/2755-2721/156/2025.CH23277

applications, it can be computationally expensive, particularly in rendering complex scenes in real-time.

3D Gaussian Splatting [6]: 3D Gaussian Splatting represents 3D scenes using a sparse set of weighted Gaussian distributions positioned in 3D space. This technique captures complex scene structures implicitly and offers a more efficient approach to novel view synthesis than traditional methods. By modeling surface elements or points as Gaussian blobs, it enables compact and rich scene representations.

3.2. Diffusion-based generative method

DreamFusion [7]: Pioneering the NeRF + Diffusion Approach

DreamFusion was one of the first methods to combine Score Distillation Sampling (SDS) with 2D diffusion models to optimize NeRF, yielding state-of-the-art results in 3D content generation. This approach helped to address the challenge of generating high-quality 3D content while maintaining geometric accuracy and texture consistency. However, DreamFusion highlighted some key issues, such as avoiding the emergence of multi-face objects (Janus problem) and improving the generation speed.

Magic3D [9]: Coarse-to-Fine Optimization

To overcome these challenges, Magic3D introduced a two-stage, coarse-to-fine optimization framework, which improves both the speed and quality of 3D content generation. This method demonstrated that optimizing the generation process in stages leads to better results in terms of detail and fidelity.

Fantasia3D [11]: Decoupling Geometry and Appearance

In contrast, Fantasia3D suggested that coupling geometry and appearance through volume rendering may not be optimal for high-quality 3D content. They proposed disentangling the modeling and learning of geometry and appearance for better results. For geometry, they used a hybrid scene representation and encoded surface normals as input for the diffusion model. For appearance, they introduced spatially varying bidirectional reflectance distribution function (BRDF) to better model surface materials, which is essential for photorealistic rendering.

3.3. 3DGS-based generative method

DreamGaussian [12] is the first attempt to apply 3D Gaussian Splatting technology to the generation of 3D contents. It incorporates Score Distillation Sampling into 3DGS and companions mesh extraction and texture refinement in UV space to enhance rendering quality. GSGEN [13] recognizes that due to the expressive display of 3DGS, the generation of geometric content can be directly rectified with a point cloud prior. Utilizing this feature can significantly alleviate the Janus problem, and GSGEN also employs a density-based densification to further optimize the generation quality. Similarly, GaussianDreamer [14] shares the same basic concept with GSGEN, but GaussianDreamer applies noisy point growing and color perturbation to enhance the initialized Gaussians, reducing the generation time to less than 15 minutes.

4. Applications

4.1. 3D assets generation

With the development of digital technology, Digital 3D assets have become indispensable, especially with the introduction of concepts such as digital twins and the metaverse, making Digital 3D assets prevalent across various industries, including education, gaming, architecture, animation, virtual and

DOI: 10.54254/2755-2721/156/2025.CH23277

augmented reality, among others. By leveraging AIGC technology, the time spent on modeling 3D content can be significantly reduced, thereby enhancing work efficiency.

4.2. 3D scene generation

In fields such as film and television production, cultural heritage conservation, and others, it is often necessary to convert 2D image data captured by cameras into high-precision 3D models. With the help of 3D AIGC technology, the system can automatically generate 3D scenes that are highly consistent with the image content by inputting a small number of 2D images. This technology can provide efficient and accurate solutions for complex scene reconstruction that traditional methods find difficult to achieve, especially when dealing with scenes that are structurally complex and rich in detail, significantly improving reconstruction accuracy and reducing manual intervention.

4.3. Real object modeling

Although text-to-image generation has made rapid progress, creating realistic 3D objects remains a significant challenge because it requires considering the specific shapes of the surfaces to be rendered. Fully automatic 3D content generation is still limited by the manual labor required to design textures. Therefore, if the texture design process could be automatically achieved through text, it would greatly simplify the workflow of 3D modeling, reduce the workload of designers, and enhance the flexibility and efficiency of the modeling process.

5. Challenge

Although significant progress has been made in AIGC 3D in recent years, there are still some unresolved issues that can significantly impact the efficiency and quality of 3D content generation. This section will elaborate on these challenges and provide some directions for future development.

5.1. Evaluation metrics

Despite the availability of numerous evaluation metrics that can assess models from different perspectives, the evaluation of 3D content in terms of quality and geometry still largely relies on human rating. Due to the high data complexity of 3D content, quantifying its quality is challenging, but doing so could bring significant development to the field. There are also studies that suggest methods for quantifying 2D content can be applied to 3D content as well. Furthermore, research has already developed testbench [15] based on existing evaluation metrics, and providing evaluation solutions for AIGC 3D is also a research direction.

5.2. Generation

The challenges for generating 3D content mainly focus on the trade-off between speed and quality and addressing the Janus problem. To improve generation speed, utilizing 3D Gaussian Splatting might become a solution direction; for the Janus problem, most research suggests it is caused by over-reliance on 2D diffusion priors. The expressive representation of 3DGS itself can effectively alleviate this issue; additionally, creating a coarse 3D prior-guided for 2D diffusion prior to guide generation is also a solution.

References

[1] *“Chatgpt: Optimizing language models for dialogue” Nov. 2022.*

[2] *Xu, Jiale, et al. "Dream3d: Zero-shot text-to-3d synthesis using 3d shape prior and text-to-image diffusion models." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.*

DOI: 10.54254/2755-2721/156/2025.CH23277

[3] *Qiu, Lingteng, et al. "Richdreamer: A generalizable normal-depth diffusion model for detail richness in text-to-3d." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.*

[4] *Xie, Haozhe, et al. "Citydreamer: Compositional generative model of unbounded 3d cities." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.*

[5] *Rombach, Robin, et al. "High-resolution image synthesis with latent diffusion models." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.*

[6] *Kerbl, Bernhard, et al. "3d gaussian splatting for real-time radiance field rendering." ACM Trans. Graph. 42.4 (2023): 139-1.*

[7] *Poole, Ben, et al. "Dreamfusion: Text-to-3d using 2d diffusion." arXiv preprint arXiv:2209.14988 (2022).*

[8] *Lin, Chen-Hsuan, et al. "Magic3d: High-resolution text-to-3d content creation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.*

[9] *Chen, Rui, et al. "Fantasia3d: Disentangling geometry and appearance for high-quality text-to-3d content creation." Proceedings of the IEEE/CVF international conference on computer vision. 2023.*

[10] *Wang, Zhengyi, et al. "Prolificdreamer: High-fidelity and diverse text-to-3d generation with variational score distillation." Advances in Neural Information Processing Systems 36 (2024).*

[11] *Liang, Yixun, et al. "Luciddreamer: Towards high-fidelity text-to-3d generation via interval score matching." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.*

[12] *Tang, Jiaxiang, et al. "Dreamgaussian: Generative gaussian splatting for efficient 3d content creation." arXiv preprint arXiv:2309.16653 (2023).*

[13] *Chen, Zilong, et al. "Text-to-3d using gaussian splatting." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.*

[14] *Yi, Taoran, et al. "Gaussiandreamer: Fast generation from text to 3d gaussians by bridging 2d and 3d diffusion models." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.*

[15] *He, Yuze, et al. "T^3 Bench: Benchmarking Current Progress in Text-to-3D Generation." arXiv preprint arXiv:2310.02977 (2023).*