УДК 004.896 + 007.52

A REVIEW OF 3D GAUSSIAN SPLATTING METHOD AND ITS APPLICATIONS IN ROBOTICS

Mohrat M. (ITMO University)

Scientific supervisor – professor, doctor of technical sciences, Kolyubin S. A. (ITMO University)

Introduction. Novel-view synthesis, an important task in computer vision, has traditionally relied on Neural Radiance Field (NeRF) methods, which optimize Multilayer Perceptron (MLP) models. While NeRF approaches offer continuous representations conducive to optimization, their reliance on MLPs leads to computationally intensive rendering processes. In contrast, 3D-GS, introduced in [1], offers a compelling alternative with higher rendering speeds. By harnessing the parameters of 3D Gaussians, 3D-GS efficiently models complex scenes, mitigating the computational burden associated with traditional NeRF methods. Its differentiable rendering algorithms and explicit scene representations not only enable real-time rendering but also afford enhanced levels of editability and control.

Main part. 3D Gaussian Splatting is a rasterization technique for real-time 3D reconstruction and rendering of images taken from multiple points of view. Unlike NeRFs, 3D Gaussian splatting is considered as a point-based rendering method, where the 3D space is defined as a set of 3D Gaussians that represent both appearance and geometry, where each Gaussian has its position, rotation, scale, color, and opacity parameters. These parameters can be optimized during training, and each 3D Gaussian is projected "splatted" onto the 2D image plane by rasterization.

To have a 3D representation of a scene based on 3D-GS, several steps should take place, starting from the initialization of Gaussians with 3D points from Structure from Motion (SfM) method or SLAM-generated 3D maps, then a differentiable Gaussian rasterization to create and image that will be fed into optimization process using Stochastic Gradient Descent. Despite the fact that Gaussians' parameters are optimized through a machine learning approach, rendering does not require any heavy processing or any MLP queries as it's executed by a tile-based rasterizer.

A trained scene could be represented with up to several millions of 3D Gaussians. Thus, a main problem with this approach is the need to store the aforementioned parameters for each Gaussians. Therefore, the problem boils down to inefficient memory utilization. In turn, the authors of many works managed to create a compact representation based on 3D-GS, while retaining the ability to create high-quality representations. For instance, authors of [2] proposed a more compacted representation of view-dependent color by employing a grid-based neural field rather than Spherical Harmonics SH function and proposed a learnable masking strategy that reduces the number of Gaussians.

Approach of 3D-GS has a variety of applications that range between different domains such as robotics and autonomous driving. For instance, in SLAM, research demonstrates the use of 3D Gaussians to enhance dense SLAM with a single RGB-D camera, improving rendering speed and map accuracy [3]. Additionally, one work studied the possibility of scenes' decomposition into static and dynamic components [4]. Another work showed that 3D Gaussians can be further augmented with linguistic embedding for enhanced scene understanding [5]. These applications showcase the transformative potential of 3D GS across different sectors.

Conclusion. To sum up, this work reviews the method of 3D Gaussian Splatting that presents a promising alternative to traditional NeRF methods for novel-view synthesis. Despite facing challenges such as memory inefficiency, this work demonstrates how researchers have devised compact representations to maintain high-quality outputs and fast rendering speed. The versatility of 3D-GS is evident in its applications across various domains. This work showcases three key applications for robotics, which shows the potential of 3D-GS in reshaping the representations and interaction with 3D space.

References.

1. B. Kerbl, G. Kopanas, T. Leimkuhler, and G. Drettakis, “3D gaussian splatting for real-time radiance field rendering,” ACM Transactions on Graphics, vol. 42, July 2023.
2. J. C. Lee, D. Rho, X. Sun, J. H. Ko, and E. Park, “Compact 3D gaussian representation for radiance field,” 2023.
3. N. Keetha, J. Karhade, K. M. Jatavallabhula, G. Yang, S. Scherer, D. Ramanan, and J. Luiten, “Splatam: Splat, track map 3D gaussians for dense rgb-d slam,” arXiv, 2023.
4. Y. Yan, H. Lin, C. Zhou, W. Wang, H. Sun, K. Zhan, X. Lang, X. Zhou, and S. Peng, “Street gaussians for modeling dynamic urban scenes,” 2024.
5. M. Qin, W. Li, J. Zhou, H. Wang, and H. Pfister, “Langsplat: 3D language gaussian splatting,” 2023.