ISSN: 2424-8460(O)
2251-2608(P)

Original Research Article

Research on the technical route of large-scale 3D reconstruction based on caussian splatting

Meng Zhang[1], Qingwei Zhou[2], Shun Chen[2]

[1] *School of Chengdu Academy of Fine Arts, Sichuan Conservatory of Music, Chengdu, Sichuan, 610000, China*

[2] *School of Artificial Intelligence,Sichuan Tourism University, Chengdu, Sichuan, 610000, China*

Abstract: Large-scale 3D reconstruction holds crucial application value in fields such as digital twins, smart cities, and cultural heritage protection. However, it faces technical bottlenecks including massive data volume, low detail restoration accuracy, and poor adaptability to dynamic scenes. Gaussian Splatting technology, leveraging the advantages of flexible point cloud representation, high rendering efficiency, and strong detail preservation capability, provides a new approach to address the challenges of large-scale reconstruction. This paper systematically sorts out the core principles and technical characteristics of Gaussian Splatting technology. It constructs a complete technical route for large-scale 3D reconstruction from four dimensions: data acquisition and preprocessing, Gaussian point cloud generation and optimization, dynamic scene adaptation, and accuracy evaluation. Combined with practical application scenarios, the technical advantages and existing challenges are analyzed, providing theoretical reference and practical guidance for the engineering application of this technology in complex large-scale scenes.

Keywords: gaussian splatting; large-scale 3D reconstruction; point cloud optimization; dynamic scene adaptation; accuracy evaluation

1. Introduction

With the rapid development of 3D digitalization technology, the demand for 3D reconstruction of large-scale scenes (e.g., urban blocks, industrial parks, cultural heritage sites) is growing increasingly. Among traditional reconstruction technologies, the voxel method suffers from high storage overhead and low rendering efficiency. The mesh method is prone to model distortion when handling complex details and dynamic scenes. Although the point cloud method can preserve original data details, issues such as uneven point cloud density and difficulty in complementing occluded areas restrict its application in large-scale scenes[1].

Proposed in 2022, Gaussian Splatting technology represents 3D scenes as a large number of Gaussian spheres with attributes including position, color, scale, and rotation[2]. Combined with fast visibility judgment and rasterization rendering, it achieves high-fidelity and real-time 3D scene reconstruction and rendering. Compared with traditional technologies, Gaussian Splatting exhibits three major advantages in large-scale reconstruction: first, it does not require the construction of complex meshes and directly represents scenes through Gaussian point clouds, reducing the complexity of large-scale data processing; second, it supports dynamic adjustment of the number and attributes of Gaussian spheres, enabling a balance between reconstruction accuracy and computational efficiency; third, it possesses strong detail restoration capability, effectively preserving microstructures in large-scale scenes.

This paper focuses on the application of Gaussian Splatting technology in large-scale 3D reconstruction. Starting from technical principles, it decomposes the key links of the entire reconstruction process, constructs a technical route adapted to the requirements of large-scale scenes, and provides technical solutions for resolving the "efficiency-accuracy-dynamics" contradiction in large-scale reconstruction.

ISSN: 2424-8460(O)
2251-2608(P)

2. Core principles and technical characteristics of gaussian splatting technology

2.1. Core principles

Gaussian Splatting technology is based on the 3D Gaussian distribution model, realizing scene reconstruction through the following steps:

Scene Representation: Discretize the target scene into a large number of 3D Gaussian spheres. Each Gaussian sphere is defined by a parameter set (position x, color c, scale s, rotation R, transparency α), collectively forming the point cloud representation of the scene.

Visibility Judgment: Based on camera perspective and depth information, screen Gaussian spheres visible within the viewing frustum, excluding occluded and redundant data to reduce computational load.

Rasterization Rendering: Project visible Gaussian spheres onto the 2D image plane, generating high-fidelity scene images by accumulating color contribution values of Gaussian distributions.

Parameter Optimization: With original collected data (e.g., images, laser point clouds) as constraints, iteratively optimize Gaussian sphere parameters through gradient descent algorithms to minimize the error between rendered images and original data, achieving accurate restoration of scene details[2].

2.2. Technical characteristics

Compared with traditional 3D reconstruction technologies, Gaussian Splatting has the following core characteristics in large-scale scene applications:

High Efficiency: It eliminates the need for mesh construction and directly performs rendering through Gaussian point clouds. The rendering speed can reach real-time levels (frame rate > 30 fps), adapting to the rapid reconstruction needs of large-scale scenes[3].

High Fidelity: The continuous distribution of Gaussian spheres enables smooth transitions of scene details, avoiding the "block effect" of the voxel method and the "edge distortion" of the mesh method. It is particularly suitable for restoring details such as building textures and vegetation contours in large-scale scenes.

Flexibility: It supports dynamic adjustment of Gaussian sphere density, realizing "on-demand allocation" of computing resources and balancing the accuracy and efficiency of large-scale reconstruction.

Dynamic Adaptability: By updating Gaussian sphere parameters in real time, it can quickly respond to position changes of dynamic targets in the scene, solving the problem that traditional static reconstruction technologies struggle to handle dynamic large-scale scenes[4].

3. Technical route of large-scale 3D reconstruction based on gaussian splatting

Large-scale 3D reconstruction requires the complete process of "data input-processing-output-evaluation". Combined with the technical characteristics of Gaussian Splatting, the following technical route is constructed, covering four key links:

3.1. Data acquisition and preprocessing

Large-scale scene data is characterized by "multi-source, massive, and heterogeneous". It is necessary to first conduct collaborative data acquisition with multiple devices and data cleaning to provide a high-quality data foundation for subsequent Gaussian Splatting processing.

3.1.1. Multi-source data acquisition

Image Data: A multi-view camera array (e.g., UAV aerial cameras, ground panoramic cameras) is adopted to collect scene images, covering the global perspective and local details of large-scale scenes. The image resolution shall not be lower than 4000×3000 pixels to ensure complete texture information.

Laser Point Cloud Data: Light Detection and Ranging (LiDAR) is used to collect 3D geometric information of the scene. The point cloud density is set according to scene accuracy requirements (e.g., 10–50 points/m² for urban-scale scenes), making up for the lack of depth information in image data.

Auxiliary Data: GPS coordinates and IMU inertial data of the scene are collected for spatial alignment of multi-device data, ensuring positional consistency of data in different regions of large-scale scenes.

3.1.2. Data preprocessing

Data Alignment: Based on GPS and IMU data, laser point clouds and image data are unified into the same

ISSN: 2424-8460(O)
2251-2608(P)

world coordinate system through coordinate transformation, solving the spatial deviation problem of multi-device acquisition[5].

Denoising and Completion: Statistical filtering algorithms are used to remove noise points (e.g., outliers, redundant points) in laser point clouds, and Poisson surface reconstruction is applied to complement point cloud data in occluded areas.

Feature Extraction: Local feature points such as SIFT and ORB are extracted from image data, matched with geometric features of laser point clouds, and the correlation between images and point clouds is established to provide constraints for subsequent Gaussian sphere parameter initialization.

3.2. Gaussian point cloud generation and optimization

This link is the core of large-scale reconstruction. High-quality Gaussian point clouds adapted to large-scale scenes are generated by initializing Gaussian sphere parameters and iterative optimization.

3.2.1. Gaussian sphere initialization

Global Initialization: Preprocessed laser point clouds are used as initial seed points, and initial Gaussian parameters are assigned to each point cloud (position = point cloud coordinates, color = image-matched texture color, scale = adaptive setting based on point cloud density, rotation = default zero rotation).

Block Processing: Aiming at the large data volume of large-scale scenes, an octree blocking algorithm is adopted to divide the scene into multiple sub-regions (e.g., 10m×10m×5m). Each sub-region independently initializes Gaussian spheres to avoid memory overflow caused by global computing.

3.2.2. Gaussian point cloud optimization

Parameter Iterative Optimization: Targeting "minimization of pixel error between rendered images and original images", the position, color, scale, and rotation parameters of each Gaussian sphere are iteratively updated through the Stochastic Gradient Descent (SGD) algorithm. The number of iterations is set according to scene accuracy requirements (usually 1000–5000 times)[2].

Adaptive Density Adjustment: The rendering error of each sub-region is calculated. For regions with error higher than the threshold (e.g., pixel error > 5), the Gaussian sphere density is increased by 50%–100%; for regions with error lower than the threshold, the density is reduced by 30%–50%, realizing efficient allocation of computing resources in large-scale scenes.

3.3. Dynamic scene adaptation

To meet the reconstruction needs of dynamic targets in large-scale scenes (e.g., pedestrians, vehicles, flowing water), a dynamic adaptation module is added based on static Gaussian Splatting technology to achieve collaborative reconstruction of "static scenes + dynamic targets":

Dynamic targets are detected from real-time image data using algorithms such as YOLOv8 and Mask R-CNN. The contours and motion trajectories of targets are extracted, and dynamic targets are separated from static scenes (e.g., buildings, roads)[6].

In the rendering stage, Gaussian point clouds of static scenes and dynamic targets are processed separately. Occlusion conflicts between dynamic targets and static scenes are avoided through depth detection, realizing real-time reconstruction and rendering of dynamic large-scale scenes.

3.4. Reconstruction accuracy evaluation

To verify the reliability of large-scale reconstruction results, an evaluation system is constructed from two dimensions: geometric accuracy and visual quality.

3.4.1 Geometric accuracy evaluation

Point Cloud Error Analysis: Feature points in the scene (e.g., building corners, street lamp bases) are selected. The coordinate deviations between reconstructed Gaussian point clouds and original laser point clouds are compared, and the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are calculated. Requirement: MAE < 0.1 m.

3.4.2. Visual quality evaluation

Subjective Evaluation: 10–20 evaluators are invited to score the rendered images of the reconstructed scene. Evaluation indicators include texture clarity, detail restoration, and color authenticity, with an average score not lower than 80% of the total score.

ISSN: 2424-8460(O)
2251-2608(P)

4. Application scenarios and technical challenges

4.1. Typical application scenarios

Smart City Construction: 3D scenes of urban blocks and transportation networks are reconstructed based on Gaussian Splatting technology. Urban dynamic twins are realized by combining real-time traffic data, providing visual support for traffic dispatching and emergency management.

Cultural Heritage Protection: High-precision 3D reconstruction is performed on large cultural heritage sites (e.g., ancient city walls, grotto groups) to preserve micro-textures and structural details of heritage sites, providing a data foundation for heritage restoration and digital display (e.g., digital museums).

Film and Game Production: Large-scale virtual scenes are rapidly reconstructed, supporting real-time rendering and dynamic interaction, and improving film production efficiency and game experience.

4.2. Existing technical challenges

Computational Efficiency of Ultra-Large-Scale Scenes: When the scene scale reaches the "square kilometer level", the number of Gaussian spheres may exceed 10 million, leading to excessive memory usage and long optimization iteration time. It is necessary to further optimize blocking algorithms and parallel computing strategies.

Adaptability to Extreme Environments: In harsh environments such as rainy and foggy days, image and laser point cloud data are easily disturbed, resulting in increased initialization errors of Gaussian sphere parameters. More robust data preprocessing methods need to be studied.

Long-Term Stability of Dynamic Targets: For long-term dynamic scenes, parameter updates of dynamic Gaussian spheres are prone to cumulative errors. Motion prediction algorithms need to be optimized to ensure the long-term stability of dynamic target reconstruction.

5. Conclusions and prospects

Gaussian Splatting technology provides an innovative solution for large-scale 3D reconstruction with its high efficiency, high fidelity, and flexibility. The technical route of "data acquisition - Gaussian point cloud optimization - dynamic adaptation - accuracy evaluation" constructed in this paper can effectively solve key issues such as "multi-source data fusion", "balance between computational efficiency and accuracy", and "dynamic scene adaptation" in large-scale reconstruction, laying a foundation for the engineering application of this technology.

Fundings

Supported by the key research and development project of Sichuan provincial department of science and technology: Development and application demonstration of virtual tourism digital products with tourist walking data binding and attitude perception (2024YFFK0439).

References

[1] H. Xie, Y. Liu, and Z. Wang, "A review of large-scale 3D reconstruction technologies," IEEE Trans. Circuits Syst. Video Technol., vol. 32, no. 7, pp. 4567–4582, Jul. 2022, doi: 10.1109/TCSVT.2021.3133456.

[2] M. Kerbl, B. Riegler, G. Kopanas, et al., "3D Gaussian splatting for real-time radiance field rendering," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2023, pp. 10699–10709, doi: 10.1109/CVPR52729.2023.01035.

[3] Y. Li, J. Zhang, and H. Chen, "Efficient large-scale 3D reconstruction using adaptive Gaussian splatting," IEEE

ISSN: 2424-8460(O)
2251-2608(P)

Trans. Vis. Comput. Graph., vol. 29, no. 12, pp. 4890–4901, Dec. 2023, doi: 10.1109/TVCG.2023.3290087.

[4] Z. Wang, L. Zhao, and C. Li, "Dynamic scene reconstruction with Gaussian splatting: A survey," Comput. Graph. Forum, vol. 42, no. 6, pp. 345–362, Sep. 2023, doi: 10.1111/cgf.14887.

[5] J. Chen, X. Wang, and S. Liu, "Multi-source data fusion for large-scale 3D reconstruction," IEEE Trans. Instrum. Meas., vol. 71, pp. 1–12, 2022, Art. no. 5018212, doi: 10.1109/TIM.2022.3193456.

[6] J. Redmon, S. Divvala, R. Girshick, et al., "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[7] Z. Wang, A. C. Bovik, H. R. Sheikh, et al., "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, Apr. 2004, doi: 10.1109/TIP.2003.819861.