ISSN: 2710-0588

3D Scene Reconstruction of Architectural Engineering Based on 3D Gaussian Splatting

Haotian Wang, Feng Ding, Yuan Liu

Xi'an Technological University School of Mechatronic Engineering, Xi'an 710016, China

Abstract

In order to solve the problems of high time cost, poor rendering quality and low efficiency of manual maintenance in the field of engineering and architecture, a novel method combining 3D Gaussian algorithm for real-world reconstruction was proposed. The method uses UAV aerial photography for data collection, and uses point cloud data for reconstruction, which realizes the real-time rendering of the radiation field, which provides a significant acceleration for scene optimization and novel view synthesis. It greatly improves the efficiency and accuracy in the field of construction engineering.

Keywords

3D reconstruction, UAV, radiation field, 3D gaussian.

1. Introduction

At present, there are various methods for establishing 3D object models, covering various technologies from manual modeling to automatic modeling. The most common modeling technique is to create objects by constructing a mesh composed of polygons, which is suitable for creating complex objects composed of simple shape combinations. By modeling through images or videos, this method utilizes computer vision technology to reconstruct three-dimensional structures from multiple images taken from different perspectives. Through steps such as feature matching, geometric calibration, and dense matching, the three-dimensional coordinates of each point in the scene are calculated; The point cloud based 3D reconstruction method generates a 3D model of an object [1] from the point cloud perspective by scanning the object. Point cloud based 3D reconstruction is a technique that utilizes computer vision and graphics to reconstruct the three-dimensional shape and structure of an object or scene from a set of discrete 3D point data. The main processes include data registration [2], point cloud data preprocessing [3], segmentation [4], triangulation [5-6], and mesh rendering.

Debevec et al [7]. proposed a method of representing the initial model using parametric geometry. Using modeling software for reconstruction has obvious drawbacks, such as insufficient restoration of the reconstruction effect and small reconstruction scale. Chen et al. [8] developed a multi view 3D reconstruction system based on the Windows environment, which can input images from various perspectives to obtain a 3D model. This image-based 3D reconstruction mainly includes sparse point cloud reconstruction and dense point cloud reconstruction. Sparse point cloud reconstruction (Structure From Motion, SFM) mainly includes two parts: feature point detection and matching, and SFM reconstruction. On this basis, Lowe [9-10] proposed the SIFT (Scale Invariant Feature Transform) algorithm, which solves the problem that the initial detection operator is easily affected by factors such as lighting, scale, and rotation. In Multi View Stereo (MVS) reconstruction of dense point clouds, the calculation of depth maps is particularly important. The calculation method mainly involves calculating geometric parameters to obtain image depth information, such as Plan Sweeping proposed by Gallup et al. [11], Patch Match proposed by Bleyer et al. [12], and Deep MVS proposed by Huang et al. [13]

ISSN: 2710-0588

The development of traditional geometric modeling techniques is relatively mature. However, if this modeling technique is consistently used in the field of construction engineering, not only does it require a long cycle and numerous modeling processes, but most importantly, there are still differences between the modeling effect and the real scene in many areas such as water surfaces, mirrors, or areas with unclear lighting. Therefore, image-based and point cloud based 3D reconstruction has become a hot topic in the field of 3D reconstruction. This article mainly introduces the application of 3D Gaussian Splatting optimization algorithm in the field of construction engineering. 3D Gaussian Splatting is a 3D reconstruction method based on point cloud data. [14] By mapping point cloud data to a grid in 3D space and using Gaussian functions for point cloud information difference and interpolation, the original point cloud data can be reconstructed.

3D Gaussian Splatting Reconstruction and Process

1.1. 3D Gaussian Splatting reconstruction concept

3D Gaussian Splatting is a probability based 3D reconstruction method. The core idea is to treat each 3D point as a Gaussian distribution, in order to better handle the sparsity and noise of point cloud data during the reconstruction process. Compared to traditional 3D reconstruction methods such as triangular mesh reconstruction, 3D Gaussian Splatting can reconstruct complex geometric structures and surface details more smoothly and naturally. 3D Gaussian Splatting, like NeRF, is mainly used for new view synthesis. Its feature is the use of rasterization rendering instead of NeRF's (volumetric rendering along a ray), which uses multiple 3D Gaussian spheres to represent the scene. It can ensure high quality while training fast inference speed.

In 3D Gaussian Splatting, each 3D point cloud data point is represented as a 3D Gaussian distribution, which defines the position and diffusion degree of the point in 3D space. This representation method can effectively describe the neighborhood characteristics and local geometric shapes of points. By converting point cloud data into Gaussian distribution, sparsity and noise of the data can be better handled. The parameters of Gaussian distribution, such as mean and covariance matrix, can be estimated through statistical methods. During the reconstruction process, by appropriately rendering the Gaussian distribution, a realistic 3D model can be generated. During the rendering process, lighting models and shading techniques are used to simulate real-world light reflection and shadow effects.

1.2. 3D Gaussian Splatting Real Scene Reconstruction Process

1) Point cloud initialization: Obtain an initialized sparse point cloud (sampling points) through SFM. We obtain the initial point cloud through colmap, initialize Gaussian spheres based on these point clouds to generate a 3D Gaussian ellipsoid set, place a Gaussian sphere at each point cloud position, set the center point position as the point cloud position, and randomly initialize others.

2) Projection: Based on the camera's internal and external parameters (image pose), the Gaussian sphere is projected onto the image, and 99% of all visible Gaussian spheres are projected onto the image. The method we use here is to project a 3D Gaussian onto a 2D pixel plane, using the following formula:

$$G(x) = e^{-\frac{1}{2}(x)^T \Sigma^{-1}(x)}$$

$$\sum = RSS^T R^T$$

$$\sum' = \mathrm{J}W \sum W^T J^T$$

x is a vector of random variables representing a sample point.

ISSN: 2710-0588

Σ is a covariance matrix of $d \times d$, representing the covariance relationship between random variables.

R is represented by quaternions with 4 parameters, S is a diagonal matrix with 3 parameters, so the covariance has a total of 7 parameters.

3) Rasterization rendering: Use α blending to perform rasterization rendering (Differentiated Tile Rasterizer) in the projection overlap area, which is a deterministic function and does not require learning. The process of mixing these Gaussian spheres is differentiable, and the formula used is:

$$C = \sum_{i \in N} c_i \alpha_i \prod_{j=1}^{i-1} (1-\alpha_j)$$

4) Loss comparison iteration: Update the 59 dimensional coefficients of each Gaussian sphere, and the loss function is:

$$L = (1-\lambda)L_1 + \lambda L_{D-SSIM}$$

5) Gradient feedback: Updating the properties of each Gaussian sphere is an optimization problem, and 3D Gaussian will clone and split the 3D Gaussian sphere based on the gradient. During the learning process, Gaussian spheres with large gradients have problems of insufficient and excessive reconstruction. The Gaussian spheres in the areas with insufficient reconstruction have small variances, so cloning is necessary. In the areas with excessive reconstruction, the Gaussian spheres have large variances, so splitting is necessary, as shown in Figure 1. After a fixed number of iterations, a removal operation is performed to remove almost transparent (transparency close to 0) Gaussian spheres and Gaussian spheres with excessive variances:

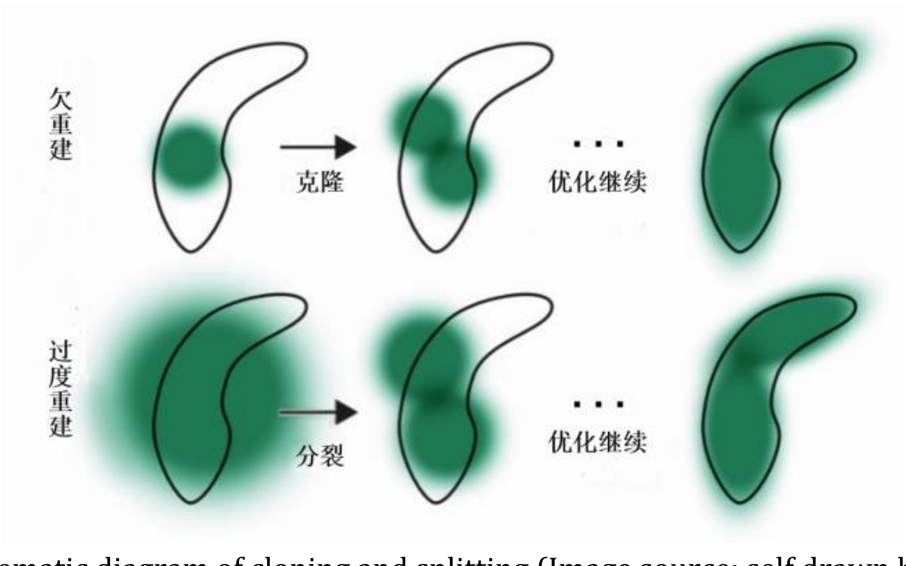


Figure 1. Schematic diagram of cloning and splitting (Image source: self drawn by the author)

By integrating the above process into a complete optimization process, the 3D Gaussian algorithm optimization rendering process can be summarized as shown in Figure 2.

ISSN: 2710-0588

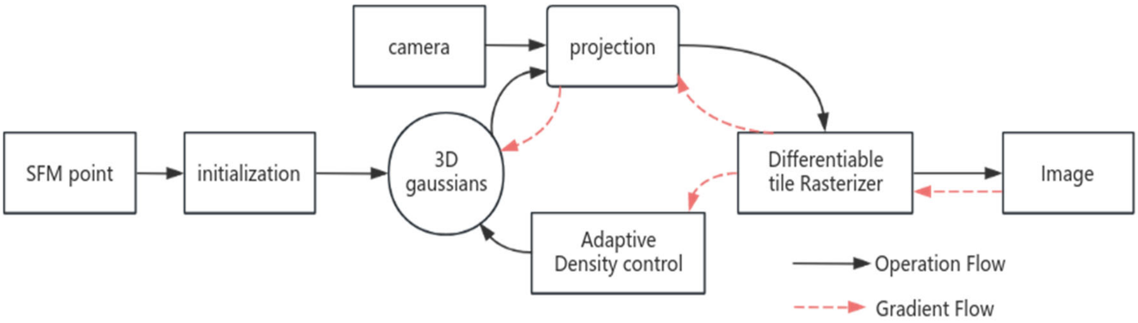


Figure 2. 3D Gaussian Algorithm Optimization Rendering Process
(Image source: self drawn by the author)

2. Experimental and Analysis of Realistic Reconstruction of Engineering Buildings

In order to compare the advantages and characteristics of 3D Gaussian Splatting reconstruction, we selected buildings from some engineering projects for reconstruction and compared them with traditional 3D reconstruction.

The raw data of images is crucial for the results of 3D reconstruction. Traditional architectural photogrammetry techniques generally collect data through optical photogrammetry satellites or drone aerial photography. The images captured by drones [15] are clearer, with controllable range and clearer details, making them more advantageous in the reconstruction of engineering buildings. This article uses DJIdrones to take photos of the reconstructed buildings and obtain measurement data. And use both 3D Gaussian 3D reconstruction technology and traditional 3D reconstruction technology to reconstruct the building. The traditional reconstruction technique uses the oblique photography software Context Capture Center Master (CCMaster), which is a widely used professional building reconstruction software in the industry.

2.1. Experimental preparation

The current remote sensing technology often fails to meet the reconstruction needs. Due to its low cost, high resolution, and strong flexibility, drone technology has gradually become a powerful supplement to satellite remote sensing. This article uses drones equipped with high-resolution cameras to cruise and capture images of buildings from the air, collecting raw data from different perspectives. As shown in Figure 3, a portion of the 464 original images collected using a drone.

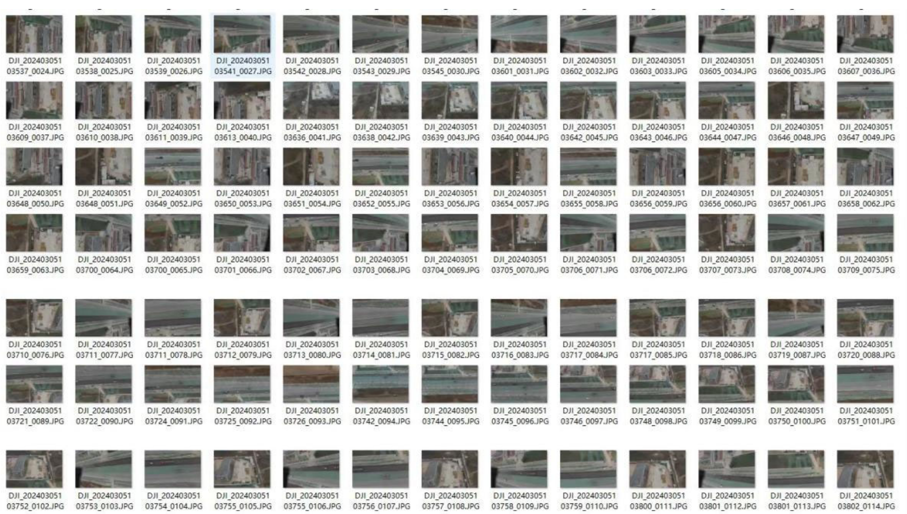


Figure 3. Image captured by drone (Image source: self taken by the author)

ISSN: 2710-0588

All experiments in this study were conducted under the same environmental configuration. CCMaster directly imported images for grid reconstruction, while in the 3D Gaussian method, due to the fact that the original image data cannot be directly used by the 3D Gaussian algorithm, data preprocessing is required to estimate the internal and external parameters of the camera. This step is completed using Colmap software. Mainly through the camera positioning function, the camera's pose is recognized and estimated, that is, the position and direction of the camera in the world coordinate system. Then, based on the initial feature point cloud and camera pose, sparse 3D reconstruction results are generated to obtain point cloud data. It is also possible to obtain dense point clouds through sparse point clouds, as well as the point cloud data required for 3D Gaussian algorithms. The sparse point cloud data used in this experiment is used for 3D reconstruction.

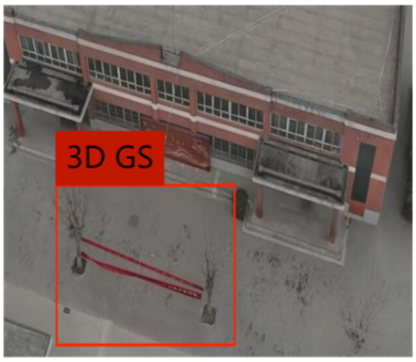
2.2. Comparison of experimental results

In order to visually compare the performance of CCMaster and 3D Gaussian reconstruction, this article conducted a comparative analysis from three aspects: the reconstruction effect, reconstruction time, and reconstruction indicators of the model.

2.2.1. Comparative analysis of reconstruction effects

464 original images of the buildings that need to be reconstructed were taken as raw data in the experiment. Two reconstruction methods were used to reconstruct them, and the shooting angles and non shooting angles of the images were compared. We selected a portion of the reconstructed images as reference.

As shown in Figure 4, it can be seen from the banner in front of the building that the 3D Gaussian 3D reconstruction in Figure 4a can clearly reconstruct the overall color and shape of the banner, while the traditional 3D reconstruction technique in Figure 4b produces poor results.



a 3D Gaussian reconstruction effect b CCMaster reconstruction effect

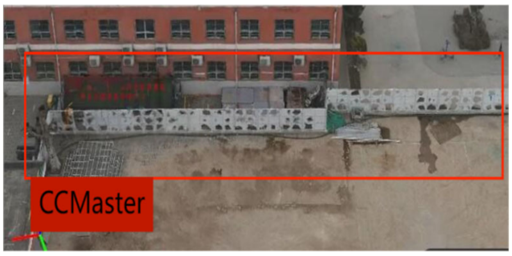
c Original image

Figure 4. Comparison between 3D Gaussian and traditional techniques (Image source: self drawn by the author)

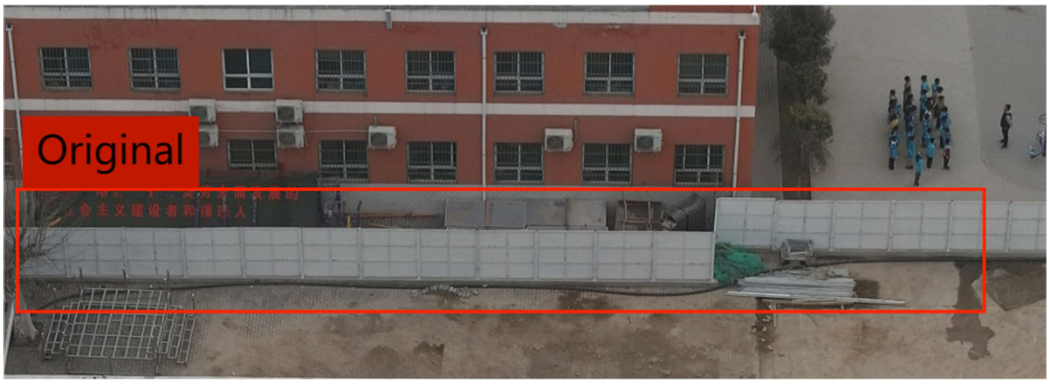
ISSN: 2710-0588

The wall reconstructed by 3D Gaussian in Figure 5 is intact, while the wall reconstructed by traditional techniques on the right has multiple holes and is severely damaged.

a 3D Gaussian reconstruction effect



b CCMaster reconstruction effect



c Original image

Figure 5. Comparison between 3D Gaussian and traditional techniques (Image source: self drawn by the author)

In Figure 6, there are many details such as the branches and leaves of trees, doors and windows, and more. The Gaussian reconstruction image is shown on the left side.

Figure 6. Comparison between 3D Gaussian and traditional techniques (Image source: self drawn by the author)

ISSN: 2710-0588

2.2.2. Comparison of reconstruction time

The time required for reconstruction is significant, which has always been one of the key issues in the field of 3D reconstruction technology. In traditional reconstruction methods, obtaining camera pose, depth information, etc. often requires a lot of time, which also leads to a high time cost for the entire reconstruction. And 3D Gaussian can greatly shorten the reconstruction time. Due to the fact that the 3D Gaussian reconstruction results depend on the quality of image training and are affected by the performance and operation of the calculator, the computer we use cannot fully represent the time required for reconstruction. However, the comparison results are still significant, as shown in Table 1 for different iteration times.

Table 1. Time for different iterations of 3D Gaussian and traditional 3D reconstruction techniques

number of iterations/time	5000	30000	50000
CCMaster/h	4.1	20.4	31
3D Gaussian/h	2.2	6.5	10.6

The reconstruction time is directly proportional to the number of training steps, and the more steps are trained, the more time is required. In practical applications, this will seriously affect the experimental progress, so it is necessary to choose an appropriate number of iterations.

The reconstruction of this scene was trained a total of 50000 iterations and took 10.6 hours. Figure 7 shows the convergence of reconstruction loss during the training process from 5000 to 50000 steps.

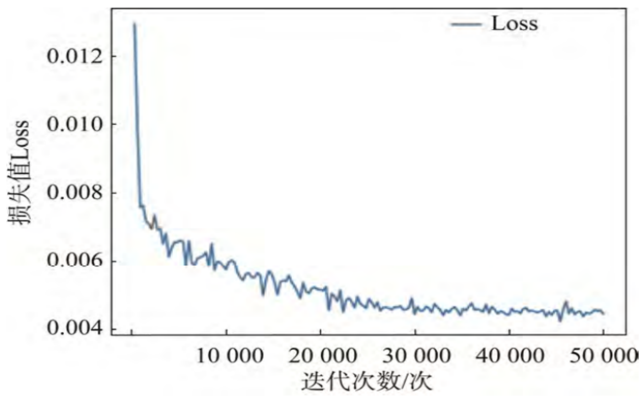


Figure 7. Convergence of training loss for 3D Gaussian reconstruction (Image source: self drawn by the author)

From the figure, it can be seen that the clear boundary in the reconstruction process is from 30000 iterations, and the most important optimization process is from 5000 to 30000 iterations. The optimization is more obvious in this process, and there is basically no difference in the optimization from 30000 to 50000 iterations. At this point, the 3D Gaussian reconstruction result is almost the final reconstruction result. In this experiment, the final reconstruction time was taken as the time consumed by 50000 iterations. In specific reconstruction tasks, the convergence time of the model reconstruction varies due to different reconstruction models, so a fixed number of iterations is often set to complete the reconstruction. However, compared to the time required by traditional reconstruction techniques, the consumption time of 3D Gaussian reconstruction is quite considerable.

ISSN: 2710-0588

2.2.3. Evaluation

PSNR (Peak Signal to Noise Ratio) is a commonly used indicator for evaluating image quality, mainly used to measure the quality difference between the original image and the compressed or reconstructed image.

The range of PSNR values is from 0 to infinity, measured in dB. The higher the PSNR value, the higher the quality of the reconstructed image. However, it should be noted that PSNR is only a rough measure of image quality, and sometimes two images with high PSNR may appear significantly different to the human eye; On the contrary, two images with very low PSNR may not appear significantly different to the human eye. The PSNR after 3D Gaussian reconstruction and traditional 3D reconstruction are shown in Table 2.

Table 2. Comparison of evaluation indicators for 3D reconstruction

number of iterations/PSNR	5000	30000	50000
CCMaster/dB	19.6	21.1	21.5
3D Gaussian/dB	22.4	25.6	27.8

Through a series of comparisons, we can conclude that the reconstruction effect of 3D Gaussian is far superior to traditional 3D reconstruction.

2.3. Display of 3D Gaussian reconstruction results

The 3D Gaussian reconstruction results are shown in Figure 8, which displays the reconstruction of buildings from four different perspectives.

Figure 8. a Front rendering

Figure 8. b Back rendering

Figure 8. c Right side rendering

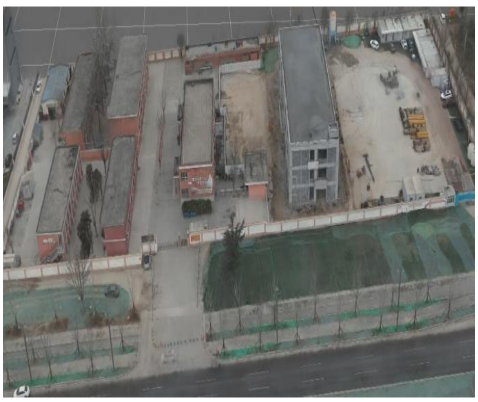


Figure 8. D Left side rendering

ISSN: 2710-0588

As shown in Figure 8, various landmark buildings, vegetation, vehicles, and people can be fully reconstructed; But some artifacts or unclear reconstructions can be found in the surrounding environment. This is because during the aerial photography process of the drone, the target building may be obstructed by surrounding buildings or not captured, which affects the reconstruction effect of its bottom.

During the reconstruction process, the captured images are affected by external factors such as lighting, weather, and angles, and the equipment needs to be continuously adjusted according to different reconstruction scenarios and reconstruction requirements. Due to the fact that 3D Gaussian data collection relies on unmanned aerial vehicles to capture images, the quality requirements for captured images are very high, especially sensitive to factors such as lighting and shadows. Therefore, ensuring good shooting equipment and a good shooting environment is a prerequisite for ensuring the quality of 3D Gaussian reconstruction. Secondly, during the reconstruction process, computer equipment, network conditions, and other factors can also affect the reconstruction speed, so it is recommended to use better equipment as much as possible in the experimental process.

3. Conclusion

By comparing and analyzing with traditional multi view stereo matching reconstruction techniques, this article has drawn the following conclusions:

(1) The Gaussian model can effectively process point cloud data in construction projects, avoiding various situations such as model missing, blurring, and a large number of artifacts in traditional 3D modeling during the reconstruction process.

(2) Compared with traditional reconstruction algorithms, the 3D Gaussian Splatting method has advantages in computational complexity, can save a lot of reconstruction time and improve efficiency, and is suitable for processing large-scale point cloud data in construction projects.

(3) In the process of engineering construction reconstruction, there are many factors that affect the reconstruction effect. Algorithm parameters and point cloud data structures can be adjusted according to different reconstruction scenarios to make the 3D Gaussian Splatting algorithm more suitable for engineering environments.

Acknowledgments

Xi'an Science and Technology Plan Projects (No.23GXFW0033, No.23GXFW0032)

Author

Wang Haotian (1999) Male Master's Student Research Direction: Architectural 3D Reconstruction Corresponding Author: Professor Ding Feng Male Research Direction: Fault Diagnosis 136790921@qq.com

References

[1] Ji.H.Y,He.Y.M,Zhang.W.J,Wu.H, Research on urban 3D reconstruction technology based on unmanned aerial vehicle oblique photogrammetry point cloud data [J] Automation and Instrumentation,2024, (09): 330-333+338.

[2] Mohd Isa S N, Abdul Shukor S A, Rahim N A, et al. A review of data structure and filtering in handling 3D big point cloud data for building preservation [C] ‖ 2018 IEEE Conference on Systems, Process and Control (ICSPC) , 2018: 141-146.

[3] Wang.M,Yan.K,Luo.X,etc, Research and Application of Preprocessing Technology Based on 3D Point Cloud Data [J] Power Big Data,2024,27(09):61-68.DOI:10.19317/j.cnki.1008-083x.2024.09.007.

ISSN: 2710-0588

[4] SUN Jinhu. Research on key technologies of point cloud segmentation and fusion [D] . Nanjing: Nanjing University of Aeronautics and Astronautics, 2013.

[5] Zhang.W.L, Research on Grid Technology of 3D Point Cloud Data [D] Changchun: Changchun University of Technoogy University of Technology, 2023. DOI:10.27805/d.cnki.gccgy.2023.000696.

[6] Tan.C, Neural network-based triangular mesh simplification algorithm [D] Hefei: University of Science and Technology of China,2022.DOI:10.27517/d.cnki.gzkju.2022.002141.

[7] Debevec P E, Taylor C J, Malik J. Modeling and rendering architecture from photographs: a hybrid geometry-and image-based approach [C] ‖Proceedings of the 23rd annual Conference on Computer Graphics and Interactive Techniques, 1996: 11-20.

[8] Chen X, Wu Q F, Wang S Z. Research on 3D reconstruction based on multiple views [C] ‖2018 13th International Conference on Computer Science＆Education (ICCSE), 2018: 1-5.

[9] Lowe D G. Object recognition from local scale-invariant features [C] ‖Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999. DOI: 10. 1109 /ICCV. 1999. 790410.

[10] Lowe D G. Distinctive image features from scale-invariant key points [J] . International Journal of Computer Vision, 2004, 60(2) : 91-110.

[11] Gallup D, Frahm J M, Mordohai P, et al. Real-time planesweeping stereo with multiple sweeping directions [C] ‖ IEEE Conference on Computer Vision and Pattern Ｒecognition, 2007. DOI: 10. 1109 /CVPＲ. 2007. 383245.

[12] Bleyer M, Rhemann C, Rother C. PatchMatch stereostereo matching with slanted support windows [C] ‖ British Machine Vision Conference, 2011.

[13] Huang P H, Matzen K, Kopf J, et al. DeepMVS: learning multi-view stereopsis [C] ‖ IEEE Conference on Computer Vision and Pattern Ｒecognition, 2018: 2821-2830.

[14] Fridovich-Keil and Yu, Matthew Tancik, Qinhong Chen, Benjamin Recht, and Angjoo Kanazawa. 2022. Plenoxels: Radiance Fields without Neural Networks. In CVPR.

[15] Mao.B.Y,Zhang.W,Liang.Z.Q,Peng.Q.Y,Wang.Y,Wang.Y.X, The application process, principles, and examples of drone photogrammetry and 3D reconstruction technology in the field of architecture [J] Huazhong Architecture, 2024, 42 (09): 53-59.