LiDAR-Guided Illumination-Aware 3D Gaussian Splatting for Cultural Heritage

Xiao Liu[1], Xinyi Li[1], Wan Li[2], Tao Liu[1], Wei Sun[1], Sheng Zhang[3*]

[1] Wuhan Geomatics Institute, Wuhan, China – (liuxiao2436@gmail.com, islixy@whu.edu.cn, t.liu@sina.com, gnss.wei@gmail.com)
[2] Hubei Surveying and Mapping Quality Supervision and Inspection Station, Wuhan, China – w.li@sina.com
[3] Langfang Natural Resources Comprehensive Survey Center, CGS, Langfang 065000, China – zhangsheng_cgs@sina.com

Keywords: 3D Gaussian Splatting; Cultural Heritage; 3D Reconstruction; LiDAR; Illumination-aware.

Abstract

To address the issues of geometric distortion and loss of details in 3D modeling for complex cultural heritage scenes, this paper proposes an improved 3D Gaussian Splatting (3DGS) reconstruction method that integrates LiDAR and illumination-awareness. First, high-precision 3D coordinates from LiDAR point clouds are utilized to guide the initialization of Gaussian Primitives, establishing a precise geometric foundation and effectively overcoming deformation on weakly textured surfaces. Second, an illumination-aware network is constructed to dynamically adjust appearance parameters by combining global illumination from images with LiDAR reflectance intensity. This decouples complex lighting from material properties, accurately reproducing the unique textures of artifacts. Finally, a multi-dimensional joint loss function incorporating photometric, scale, and appearance smoothness constraints is introduced to collaboratively optimize scene geometry, appearance, and camera poses. Experimental results on indoor and outdoor cultural heritage preservation scenarios demonstrate that the proposed method significantly outperforms various comparative algorithms in terms of both visual fidelity and geometric accuracy. The quantitative and qualitative evaluations confirm that our approach effectively eliminates geometric distortions and recovers fine texture details, providing an efficient and reliable technical solution for the digital preservation of cultural heritage.

1. Introduction

Cultural heritage is a valuable asset that embodies the historical and cultural memory of a nation, and its precise, complete, and long-term digital preservation has become a major contemporary challenge. Reality-based 3D modeling, as a core technical approach, aims to create high-fidelity digital archives (Chabuk and Al-Amiri, 2022; Llull et al., 2023). However, existing technologies still face severe challenges when dealing with cultural heritage scenes characterized by complex structures, diverse materials, and variable lighting conditions.

Traditional 3D reconstruction methods primarily rely on Photogrammetry and Light Detection and Ranging (LiDAR). Although photogrammetry is flexible, it is susceptible to lighting and viewing angles, often resulting in texture stretching and blurred details (Rangelov et al., 2024a). LiDAR technology can ensure geometric accuracy, but its limited color information and poor texture realism often necessitate complex fusion with image data (Gomes et al., 2018; Laurent et al., 2025). Therefore, achieving a balance among efficiency, geometric accuracy, and texture detail has long been a core challenge in the 3D modeling of cultural heritage scenes (Clini et al., 2024).

In recent years, while Neural Radiance Fields (NeRF) can achieve highly realistic novel view synthesis (Mazzacca et al., 2023) , their reliance on dense network queries leads to time-consuming training and rendering processes, making it difficult to meet the real-time and high-efficiency demands of cultural heritage projects. Against this backdrop, the 3D Gaussian Splatting (3DGS) technique has emerged. As an explicit, rasterization-based representation method, 3DGS utilizes millions of 3D Gaussian Primitives to represent a scene (Basso et al., 2024), achieving unprecedented real-time rendering speeds and high-quality reconstruction results, and has demonstrated enormous potential in various fields (Dahaghin et al., 2024; Jamil and Brennan, 2025).

Despite the outstanding performance of 3DGS, its direct application to complex cultural heritage scenes still has limitations (Rangelov et al., 2024b; Wang et al., 2024): Firstly, the accuracy of point clouds initialized solely by Structure from Motion (SfM) is limited. This can easily lead to geometric deformation, especially when dealing with weakly textured or smooth surfaces (such as porcelain or stone tablets), due to the poor quality of the initial point cloud (Jiang et al., 2024). Secondly, the simplistic lighting model in general 3DGS makes it difficult to realistically represent the complex illumination (e.g., mixed light sources, dynamic shadows) and special materials (e.g., metallic luster, glazed highlights) found in cultural heritage scenes (Lim et al., 2024).

To address the aforementioned challenges, this paper proposes an improved 3DGS reconstruction method tailored for cultural heritage scenes. The method aims to solve the modeling difficulties in complex cultural heritage scenarios by deeply integrating the precise geometric information from LiDAR point clouds with an advanced illumination-aware network.

2. Methodology

2.1 Method Overview

The workflow of the improved 3D Gaussian Splatting reconstruction method proposed in this paper is illustrated in Figure 1, and it primarily consists of the following steps: LiDAR-guided Gaussian initialization, construction of an illumination-aware network, and joint loss optimization. The specific implementation steps are as follows: First, a multi-view video stream and LiDAR point cloud data of the cultural artifact are acquired and registered via timestamps to establish their spatial correspondence. Next, the 3D coordinates from the LiDAR point cloud are utilized, in conjunction with the pixel color values from the camera images, to complete the initialization of the Gaussian Primitives. Then, the Spherical Harmonics (SH) coefficients of the Gaussian Primitives are dynamically adjusted through an Illumination-aware Network, a

* Corresponding author.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-561-2026

Modulation Network, and a Bias Network. Subsequently, a multi-dimensional loss function is constructed to collaboratively optimize the parameters of the Gaussian Primitives, the camera poses, and the weights of the Illumination-aware, Modulation, and Bias Networks. Finally, a high-precision 3D Gaussian scene representation of the target cultural artifact is output.

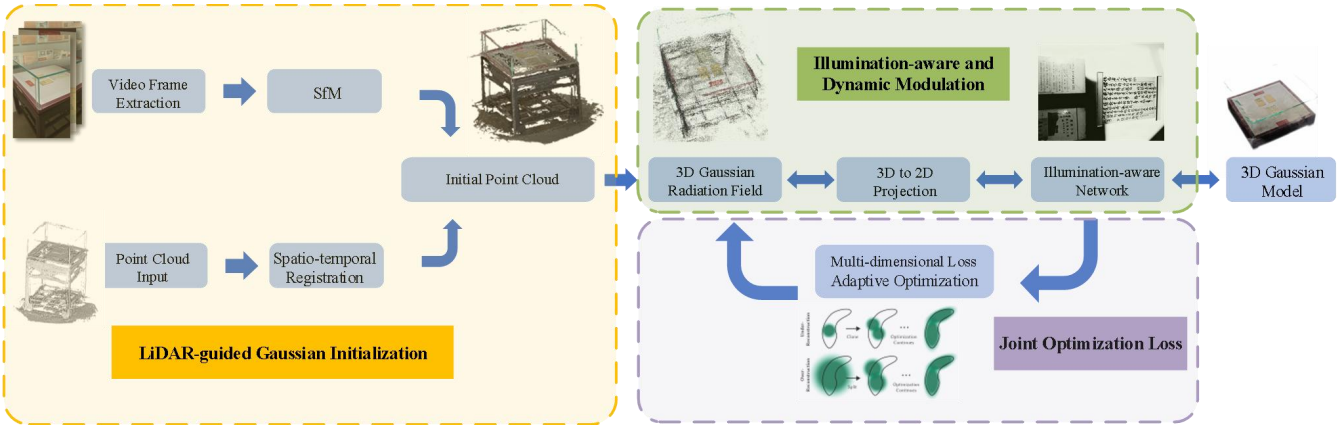


Figure 1. 3D Reconstruction Method for Cultural Heritage Scenes Based on Improved 3D Gaussian Splatting

2.2 LiDAR-guided Gaussian Initialization

First, for the time-synchronized multi-view images and LiDAR point cloud, Structure from Motion (SfM) technology is employed to recover the intrinsic and extrinsic camera parameters, after which a spatial correspondence between the point cloud and pixels is established via projection. For each 3D point in the LiDAR point cloud, its corresponding 2D pixel coordinates on the camera image are obtained using the following formula:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \left[R/t \right] \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{1}$$

In the formula, z_c represents the depth in the camera coordinate system, (u, v) are the pixel coordinates, K is the intrinsic matrix, R is the rotation matrix, tis the translation vector, (X_w, Y_w, Z_w) are the world coordinates.

The aforementioned operation identifies the corresponding pixel location in the camera image for each LiDAR point, thereby establishing the spatial correspondence between the 3D point cloud and the 2D camera images. Subsequently, the aligned multi-modal data is utilized to initialize the Gaussian primitives. The high-precision 3D coordinates from the LiDAR point cloud are directly assigned to each Gaussian primitive as its initial position parameters. Leveraging the spatial correspondence established in the preceding steps, the RGB color value of each 3D point is extracted from its corresponding pixel in the camera image and converted into 0th-order Spherical Harmonics (SH) coefficients, which serve as the initial appearance parameters for the Gaussian primitive. All higher-order (greater than 0th-order) SH coefficients are initialized to zero. The initial scale and opacity parameters of the Gaussian primitives are uniformly initialized to a small, preset isotropic value.

2.3 Illumination Perception and Dynamic Modulation

This paper constructs a neural network model to dynamically adjust the appearance of Gaussian primitives, with its architecture shown in Figure 2. The model dynamically adjusts the appearance of Gaussian primitives based on ambient illumination and is primarily composed of three parts: an Illumination-aware Network LightNet( ), a Modulation Network ModNet( ), and a Bias Network BiasNet( ). ModNet adjusts the intensity and contrast of the appearance through multiplicative modulation to simulate variations in lighting intensity, while BiasNet adjusts the base tone of the appearance via an additive bias to simulate the color or color cast of the ambient light.

First, the current view image is input into a pre-trained lightweight Convolutional Neural Network (CNN), specifically MobileNetV2 pre-trained on the ImageNet dataset, to extract global illumination information. Concurrently, the laser reflectance intensity information corresponding to each point is directly read from the LiDAR data. Then, the global illumination information and the reflectance intensity information are concatenated and input into the Illumination-aware Network LightNet( ) to generate a refined illumination environment estimate $L_{refined}$. The Illumination-aware Network is a Multi-Layer Perceptron(MLP) containing three fully connected layers, with the number of neurons in each layer set to 256, 128, and 64, respectively; each layer is followed by a ReLU activation function. Finally, $L_{refined}$ is input into the Modulation Network ModNet( ) and the Bias Network BiasNet( ), and the Spherical Harmonic (SH) coefficients are adjusted according to the following formula:

$$SH_i' = SH_i \odot ModNet(L_{refined}) + BiasNet(L_{refined}) \tag{2}$$

In the formula, SH_i' are the adjusted SH coefficients, SH_i are the original SH coefficients, $\odot$ represents the Hadamard product, $L_{refined}$ is the refined illumination environment estimate, ModNet() is the modulation network, and BiasNet() is the bias network. Herein, the Modulation Network (ModNet) is composed of a cross-attention mechanism and a conditional MLP, while the Bias Network BiasNet( ) consists of an MLP branch and a physical feature branch, the outputs of which are ultimately fused.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-561-2026 | © Author(s) 2026. CC BY 4.0 License.

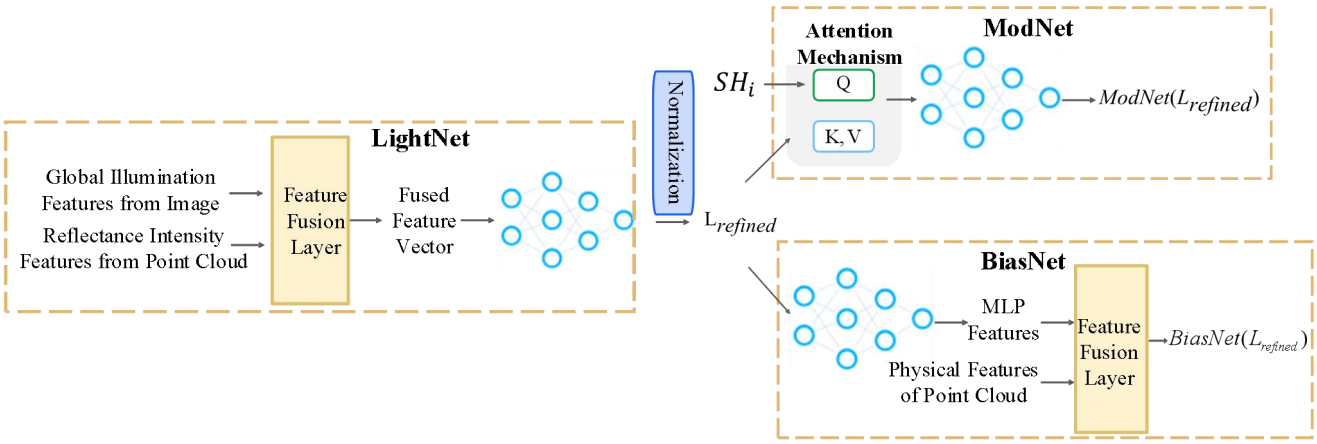


Figure 2. Architecture of the Illumination-aware Network

2.4 Joint Optimization Loss

To achieve the simultaneous optimization of scene geometry, appearance, and camera poses, this paper constructs a multi-dimensional composite loss function. By minimizing this loss function, it is possible to concurrently optimize all parameters of the Gaussian primitives (position, scale, opacity, SH coefficients), the camera poses, and all network weights of the illumination-adaptive appearance network. This optimization process is, in itself, the reconstruction process.

The formula for the multi-dimensional loss function is as follows:

$$L_{total} = L_{render} + \lambda_1 L_{scale} + \lambda_2 L_{smooth} \tag{3}$$

In the formula, λ_1 and λ_2 are hyperparameters used to balance the respective loss terms. The definitions of each loss term are as follows:

(1) Photometric Reconstruction Loss (L_{render}): This represents the discrepancy between the rendered image and the ground-truth observed image.

$$L_{render} = (1-\alpha)\cdot \| I_{render} - I_{gt} \|_1 + \alpha \cdot L_{D-SSIM}(I_{render}, I_{gt}) \tag{4}$$

In this formula, I_{render} is the image rendered using the adjusted Spherical Harmonic coefficients SH_i', I_{gt} is the corresponding ground-truth image, $\|\bullet\|_1$ represents the L1 norm, L_{D-SSIM} is the structural dissimilarity loss, and α is a hyperparameter for balancing the two terms, typically set to 0.2.

(2) Scale Regularization Loss (L_{scale}): This is used to constrain the scale of Gaussian primitives in high-frequency detail regions on the cultural heritage surface, preventing excessive blurring.

$$L_{scale} = \sum_{i \in \Omega_{high}} max(0, \| s_i \|_2 - s_{th}) \tag{5}$$

In this formula, Ω_{high} is the set of high-frequency regions, s_i is the scale parameter of the i-th Gaussian primitive, s_{th} is a preset upper-limit threshold for the scale, and $\|\bullet\|_2$ denotes the L2 norm.

(3) Appearance Smoothness Loss (L_{smooth}): This serves to maintain a smooth transition in appearance between spatially adjacent Gaussian primitives, thereby eliminating color noise artifacts.

$$L_{smooth} = \sum_i \sum_{j \in N(i)} \| SH_i' - SH_j' \|_2^2 \, N(i) \tag{6}$$

In this formula, $N(i)$ is the set of neighboring Gaussian primitives j that are spatially adjacent to the i-th Gaussian primitive, SH_i' and SH_j' are the adjusted Spherical Harmonic coefficients for two adjacent primitives i and j in the neighborhood, respectively, and $\|\bullet\|_2$ represents the square of the L2 norm.

Finally, the Adam optimizer is used to iteratively optimize the total loss L_{total}.

3. Experiment and Analysis

3.1 Experimental Preparation

Experimental data was collected on-site using a Canon EOS R5 panoramic camera array and a Faro Focus Premium handheld laser scanner. The scanner was configured with a high-resolution scanning mode, providing an angular resolution of 0.018° and an average point density of approximately 3mm at a distance of 10m. It features a ranging accuracy of ±1mm, providing high-fidelity geometric priors for the reconstruction.

Indoor cultural heritage scenes: To ensure sample diversity and representativeness, we selected four objects that cover the typical challenging cases in heritage digitization: a "Celadon-glazed Equestrian Porcelain Figurine" (complex topology), a "Famille Rose Phoenix-shaped Handles Vase" (highly specular surface), a "Blue-glazed porcelain ewer" (texture-less smooth surface), and the "Lingquan Mountain Feng Shui inscribed stele" (high-frequency text details). A total of 920 high-resolution (8192×5464 pixels) photographs were collected, yielding a point cloud of approximately 30 million points.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-561-2026 | © Author(s) 2026. CC BY 4.0 License.

Outdoor ancient architecture scene: To validate the method's generalization capabilities in large-scale, uncontrolled environments, we selected the "Xingyin Pavilion," a prominent historical building. This site was specifically chosen as a multi-faceted trial because it simultaneously presents three critical challenges: massive data scale (230 million points), intricate geometric structures (complex dougong brackets and eaves), and drastic natural lighting variations (strong sunlight and cast shadows). Validating on this scene effectively demonstrates the method's scalability and robustness to outdoor conditions.

The experimental hardware was a server equipped with an Intel Xeon Gold 6248R CPU and an NVIDIA RTX 4090 GPU. To determine the hyperparameters for the loss function, an optimization experiment was conducted using PSNR, SSIM, and LPIPS as metrics, which resulted in the final determination of λ_1=0.5, λ_2=0.1 (Table 1).

λ_1	λ_2	PSNR/db↑[(1)]	SSIM↑	LPIPS↓
0	0	31.25	0.901	0.128
0.2	0	31.89	0.917	0.115
0.5	0	32.41	0.926	0.105
0.5	0.05	32.98	0.935	0.094
0.5	0.1	33.42	0.938	0.089
0.5	0.2	33.05	0.936	0.092
0.8	0.1	32.76	0.931	0.098

Table 1. Scaling Coefficients and Evaluation Metrics

3.2 Experimental Analysis

3.2.1 Ablation Study: To systematically evaluate the effectiveness of the core modules proposed in this paper, we used the original 3DGS method as a baseline and progressively introduced the LiDAR-guided Gaussian initialization, the illumination-aware network, and the joint loss optimization module. We conducted a comprehensive analysis covering both reconstruction quality and computational efficiency.

As can be seen from Table 2, compared to the baseline model, the gradual introduction of LiDAR initialization, the illumination-aware network, and the joint optimization loss leads to significant improvements in all model metrics. When all three improvement strategies were applied, PSNR increased by 17.2%, SSIM increased by 11.4%, and LPIPS decreased by 51.9%.

Method	PSNR/db↑	SSIM↑	LPIPS↓
A. Original 3DGS (Baseline)	28.5	0.842	0.185
B. A + Point Cloud Gaussian Initialization	30.1	0.881	0.152
C. B + Illumination-aware Network	31.8	0.912	0.121
D. C + Joint Loss Optimization	33.4	0.938	0.089

Table 2. Ablation Study Results

Figure 3 shows the performance of the original 3DGS and the fully integrated improved method on the three metrics as the number of iterations increases during the training process. The PSNR and SSIM of the improved method are significantly higher than those of the original 3DGS, with a more pronounced improvement in the early iteration stages. This indicates that the proposed method has a stronger reconstruction capability in terms of image quality and structural information preservation, and can converge faster to a better result. Throughout the entire iteration process, the LPIPS of the improved method is consistently lower than that of the original 3DGS, and it drops rapidly in the early stages before finally stabilizing. This demonstrates that the improved method is superior to the original 3DGS in perceptual quality, producing images that are more consistent with human visual perception.

To assess the processing cost and model complexity, we analyzed the computational overhead between Original 3DGS and Ours. We compared the Training Time, Peak VRAM Usage, Number of Gaussians (Model Vertices), and Rendering Speed (Table 3).

Method	Training Time (min)	No. of Gaussians (Millions)	Peak VRAM Usage (GB)	Rendering Speed (FPS)
Original 3DGS (Baseline)	35	4.8	4.2	145
Ours (Full Method)	45	5.2	6.1	128

Table 3. Efficiency Comparison

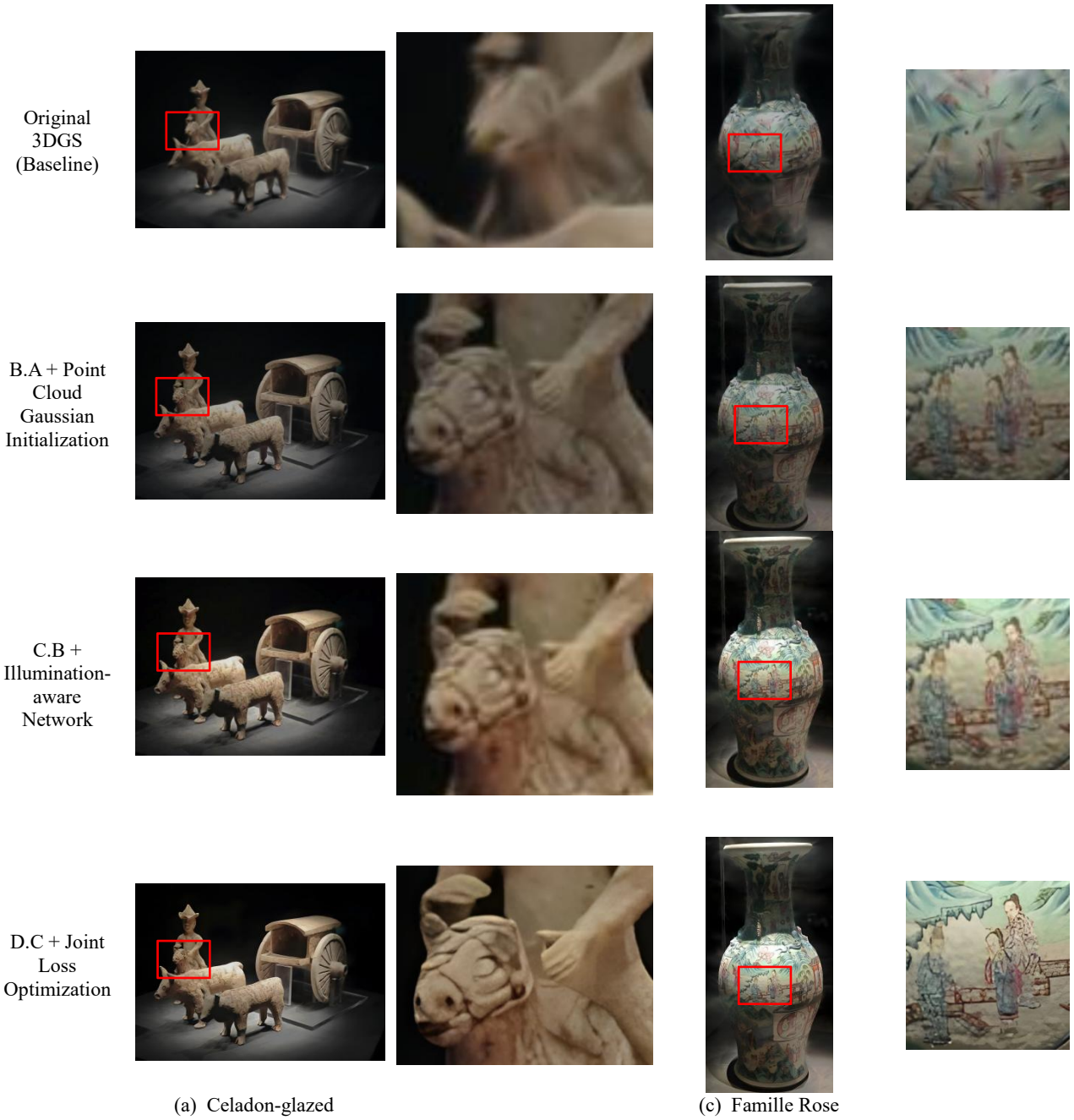
The final model consists of approximately 5.2 million Gaussian primitives (vertices), representing a slight increase over the baseline due to scale regularization allowing for adaptive densification in detail-rich regions, which directly contributes to higher geometric fidelity. In terms of training cost, the training time increased by approximately 28% (from 35 to 45 mins), which is considered a reasonable trade-off for significant quality gains. While the peak VRAM usage increased by 1.9 GB, it remains well within the capacity of standard high-end GPUs like the RTX 3090 or 4090. Crucially, the rendering speed remains at 128 FPS, far exceeding the real-time threshold of 60 FPS. Furthermore, the model demonstrates strong scalability for large-scale cultural heritage scenes, as VRAM usage exhibits a linear relationship with scene complexity, confirming the lightweight design of the illumination network.

Figure 4 visually demonstrates the improvements in visual effects brought about by the gradual introduction of each module in the ablation study. Specifically, the reconstruction result of the baseline method (Figure 4a) exhibits significant geometric noise and floating artifacts. After introducing the LiDAR prior for initialization (Figure 4b), the model's macroscopic geometric structure is effectively corrected, becoming clean and regular, but the texture details remain blurry. On this basis, the application of the illumination-aware network (Figure 4c) significantly enhances the model's sense of lighting. For example, the glaze on the "Famille Rose Phoenix-shaped Handles Vase" shows realistic highlights and color, shedding the original "plastic-like" feel.

Finally, the introduction of the joint loss optimization (Figure 4d) becomes key to recovering high-frequency details. Both the complex structural contours of the horse's head on the "Celadon-glazed Equestrian Porcelain Figurine" and the fine features such as the facial features and hair of the lady figure on the body of the "Famille Rose Phoenix-shaped Handles Vase" are clearly reconstructed, achieving the best visual effect.

(1) ↑ indicates that a higher value is better, while ↓ indicates that a lower value is better.

Figure 3. Comparison curves of PSNR, SSIM, and LPIPS indicators between 3DGS and ours



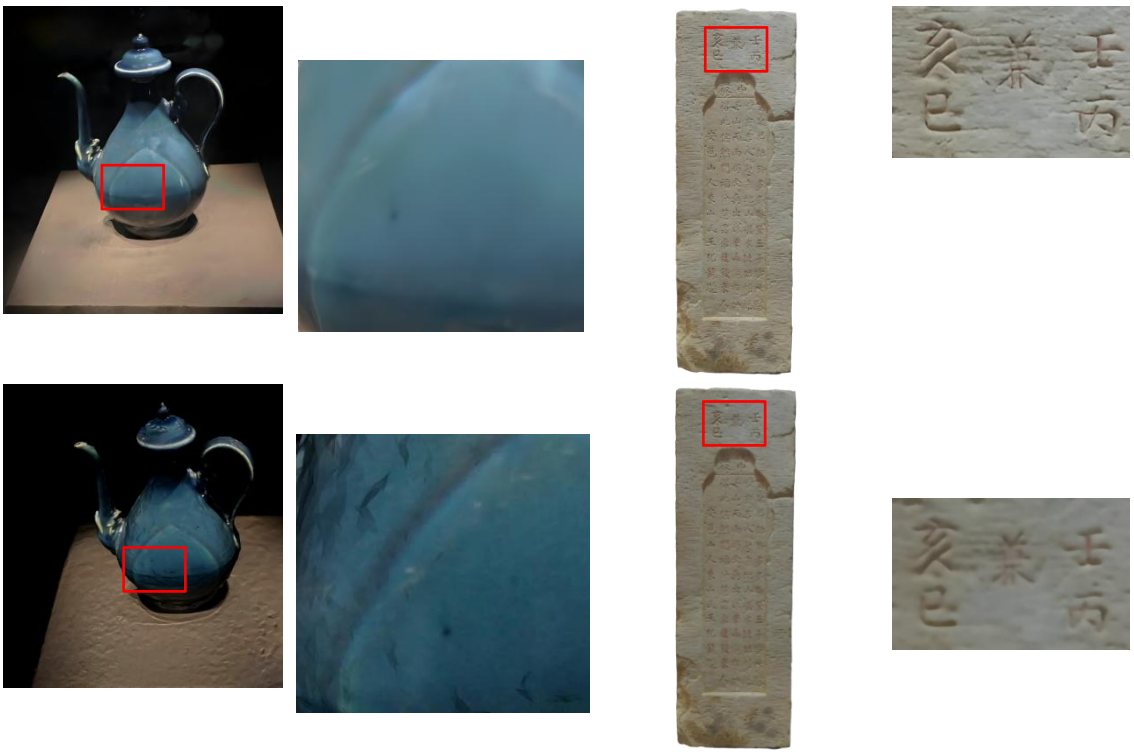
(a) Celadon-glazed Equestrian Porcelain Figurine (b) Detailed view (c) Famille Rose Phoenix-shaped Handles Vase (d) Detailed view

Figure 4. Comparison of visual effects

https://doi.org/10.5194/isprs-annals-XI-2-2026-561-2026

3.2.2 Comparative Experiment: To further validate the comprehensive performance of the proposed method, this paper conducts an in-depth comparative analysis of its visual effects against the widely used traditional oblique photogrammetry 3D reconstruction method in several typical scenarios.

As shown in Figure 5, in the task of indoor cultural relic reconstruction, for objects with continuous smooth curved surfaces such as the blue-glazed chicken-heart ewer, the traditional oblique photogrammetry method, being based on a Mesh surface representation, produces reconstruction results with blocky artifacts and texture distortion, making it difficult to realistically reproduce the warm and smooth texture of the porcelain glaze. In contrast, the method proposed in this paper can generate a highly continuous and smooth geometric surface, accurately restoring the specular highlights and delicate texture of the glaze. For objects rich in high-frequency details, such as the Fengshui Inscribed Stele, the proposed method successfully preserves the clear and sharp edges of the inscription's strokes, whereas the traditional method loses text details due to its smoothing effect.



(a) Blue-glazed Porcelain Ewer (b) Detailed view (c) Fengshui Inscribed Stele (d) Detailed view

Figure 5. Comparison of indoor cultural relic reconstruction effects

Further extending the comparison to more structurally complex large-scale outdoor scenes, Figure 6 shows the reconstruction results on the ancient building "Xingyin Pavilion". On the building facade and complex dougong structures (Figure 6a), the method of this paper achieves clear and sharp edges and precise geometric restoration, while the comparative method produces blurring and "melting" effects, leading to a significant loss of structural details. When dealing with the stone lion sculpture at the entrance (Figure 6b), the method of this paper completely preserves its rich surface details, whereas the comparative method shows obvious holes and artifacts in the recessed areas. When processing the eaves tile-end structures (Figure 6c), the method of this paper is still able to distinguish and reconstruct independent components, while the comparative method adheres these structures into an unrecognizable blurry form.

Oblique photogrammetry method

(a) Architectural facade details

(b) Sculpture details

(c) Tile-end structure

Figure 6. Comparison of outdoor scene reconstruction effects

3.2.3 Robustness Analysis: To evaluate the applicability of the proposed method under varying conditions, we analyzed its robustness across different LiDAR qualities and lighting environments. While our primary experiments utilized high-precision terrestrial laser scanning (Faro Focus), the method demonstrates resilience to lower point densities. The LiDAR data is primarily used for Gaussian initialization and scale regularization. In regions with sparse point clouds, the joint optimization driven by photometric consistency allows the Gaussians to densify and adapt, filling in geometric gaps while maintaining the macroscopic structure provided by the LiDAR skeleton. This implies the method's potential compatibility with lower-cost mobile LiDAR sensors, provided the initial registration is accurate.

Based on our experiments, we recommend a LiDAR point density of at least 1cm at the object surface to ensure optimal initialization quality, though the method remains stable at lower resolutions due to the joint optimization constraints. The "Xingyin Pavilion" experiment (outdoor scene) specifically challenges the method with uncontrolled natural illumination, including strong sunlight and cast shadows. As shown in Figure 6, our Illumination-aware Network successfully decouples these environmental factors from the material properties. Unlike traditional photogrammetry, which often bakes shadows into the texture map, our method utilizes LiDAR intensity as a lighting-invariant feature to guide the BiasNet, ensuring consistent appearance recovery even under dynamic lighting conditions.

4. Conclusions and Discussion

This paper addresses the critical challenges of geometric distortion and texture fidelity in cultural heritage digitization by proposing an improved 3D Gaussian Splatting framework that deeply integrates LiDAR priors with illumination perception. While our work aligns with the recent trend of LiDAR-enhanced 3DGS, which focuses on geometric fidelity, we specifically address the photometric challenges inherent in cultural heritage. The core contributions and findings of this study are summarized as follows:

Positioning within LiDAR-enhanced 3DGS Trends: The integration of LiDAR into 3D Gaussian Splatting has emerged as a rapidly evolving paradigm. Recent studies have successfully demonstrated that LiDAR priors can significantly improve SLAM tracking stability and mitigate geometric collapse in large-scale mapping. Our method aligns with this cutting-edge trend by adopting LiDAR-guided initialization to ensure geometric fidelity. However, a critical distinction lies in the ultimate objective: while most existing variants, such as LiGSM (Shen et al., 2025), prioritize geometric consistency and pose estimation, our work specifically addresses the photometric challenges inherent to cultural heritage, such as complex specular highlights and dynamic shading. By introducing an Illumination-aware Network atop a high-precision LiDAR-geometric foundation, we achieve high-fidelity texture reconstruction while maintaining accurate geometry, thereby filling a critical gap in digital heritage preservation.

Mechanisms of Geometric Anchoring and Photometric Decoupling: The robust geometric restoration achieved by our framework can be primarily attributed to the deterministic spatial anchors provided by the high-precision LiDAR point clouds. This geometric skeleton, coupled with our scale regularization constraint L_{scale}, effectively guides the adaptive densification of Gaussian primitives, thereby mitigating the implicit search failures and "floating" artifacts inherent to purely vision-based Structure-from-Motion (SfM) initializations. Furthermore, regarding texture fidelity, our approach differs from conventional 3DGS pipelines that primarily rely on traditional Spherical Harmonics (SH). High-order SH often leads to "baked-in" lighting artifacts when handling glazed or metallic surfaces. In contrast, our Illumination-aware Network fuses global image illumination with LiDAR reflectance intensity. This cross-modal mechanism effectively decouples environmental lighting from intrinsic material properties, mitigating the lighting artifacts found in existing methods. These findings underscore three critical conclusions: (1) explicit LiDAR priors play a crucial role in stabilizing 3DGS geometry in textureless heritage regions; (2) incorporating physical reflectance attributes effectively resolves the illumination entanglement in neural rendering; and (3) point-based explicit representations offer a promising paradigm for high-frequency detail preservation compared to conventional mesh-based pipelines.

Robustness, Limitations, and Future Directions: Quantitative evaluations confirm that our method achieves a favorable balance between geometric accuracy and visual quality, outperforming comparative algorithms in PSNR, SSIM, and LPIPS metrics. The proposed strategy demonstrates strong robustness to data quality: the scale regularization term enables detail recovery even with lower-density point clouds, provided the macroscopic geometric skeleton is preserved. This finding suggests potential for future compatibility with lower-cost mobile LiDAR sensors, reducing the hardware threshold for field deployment. However, the current reliance on high-precision registration remains a constraint. Future work will focus on exploring lightweight network architectures to further reduce computational overhead, broadening the applicability of

this technology to diverse cultural heritage documentation scenarios.

References

Basso, A., Condorelli, F., Giordano, A., Morena, S., Perticarini, M., 2024: Evolution of rendering based on radiance fields. The Palermo case study for a comparison between NeRF and Gaussian Splatting. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-2/W4-2024, 57-64. doi.org/10.5194/isprs-archives-xlviii-2-w4-2024-57-2024.

Chabuk, M. A., Al-Amiri, S. A., 2022: Virtual Reconstruction in 3D Documentation of Built Cultural Heritage. *Journal of STEPS for Humanities and Social Sciences*, 1(3), 0-0. doi.org/10.55384/2790-4237.1118.

Clini, P., Nespeca, R., Angeloni, R., Coppetta, L., 2024: 3D representation of Architectural Heritage: a comparative analysis of NeRF, Gaussian Splatting, and SfM-MVS reconstructions using low-cost sensors. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-2/W8-2024, 93-99. doi.org/10.5194/isprs-archives-xlviii-2-w8-2024-93-2024.

Dahaghin, M., Morín Castillo, M. M., Riahidehkordi, K., Toso, M., Del Bue, A., 2024: Gaussian Heritage: 3D Digitization of Cultural Heritage with Integrated Object Segmentation. *arXiv (Cornell University)*. doi.org/10.48550/arxiv.2409.19039.

Gomes, L., Silva, L., Bellon, O. R. P., 2018: Exploring RGB-D Cameras for 3D Reconstruction of Cultural Heritage. *ACM Journal on Computing and Cultural Heritage*, 11(4), 1-24. doi.org/10.1145/3230674.

Jamil, O., Brennan, A., 2025: Immersive heritage through Gaussian Splatting: a new visual aesthetic for reality capture. *Frontiers in Computer Science*, 7, 0-0. doi.org/10.3389/fcomp.2025.1515609.

Jiang, C., Gao, R. L., Shao, K.R., Wang, Y., Xiong, R., Zhang, Y., 2024: LI-GS: Gaussian Splatting with LiDAR Incorporated for Accurate Large-Scale Reconstruction. *arXiv (Cornell University)*. doi.org/10.48550/arxiv.2409.12899.

Laurent, A., Coupry, B., Brument, B., Mélou, J., Quéau, Y., Fritz, C., Durou, J.-D., 2025: Combining geometric and photometric 3D reconstruction techniques for cultural heritage. *Journal of Cultural Heritage*, 73, 43-51. doi.org/10.1016/j.culher.2025.02.013.

Lim, H., Chang, H. K., Choi, J., Yeum, C. M., 2024: LiDAR-3DGS: LiDAR Reinforced 3D Gaussian Splatting for Multimodal Radiance Field Rendering. *arXiv (Cornell University)*. doi.org/10.48550/arxiv.2409.16296.

Llull, C., Baloian, N., Bustos, B., Kupczik, K., Sipiran, I., Baloian, A., 2023: Evaluation of 3D Reconstruction for Cultural Heritage Applications. *Proc. IEEE Int. Conf. Comput. Vis. Workshops (ICCVW)*, 1634-1643. doi.org/10.1109/iccvw60793.2023.00179.

Mazzacca, G., Karami, A., Rigon, S., Farella, E. M., Trybała, P., Remondino, F., 2023: Nerf for heritage 3D reconstruction. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-M-2-2023, 1051-1058. doi.org/10.5194/isprs-archives-xlviii-m-2-2023-1051-2023.

Rangelov, D., Waanders, S., Waanders, K., van Keulen, M., Miltchev, R., 2024a: Impact of Camera Settings on 3D Reconstruction Quality: Insights from NeRF and Gaussian Splatting. *Sensors (Basel)*, 24(23), 7594. doi.org/10.3390/s24237594.

Rangelov, D., Waanders, S., Waanders, K., van Keulen, M., Miltchev, R., 2024b: Impact of Data Capture Methods on 3D Reconstruction with Gaussian Splatting. *Preprints*. doi.org/10.20944/preprints202412.1959.v1.

Shen, J., Yu, H., Wu, J., Yang, W., Xia, G.-S., 2025: LiDAR-enhanced 3D Gaussian Splatting Mapping. arXiv (Cornell University). doi.org/10.48550/arXiv.2503.05425

Wang, R., Hua, C.-H., Shingys, T., Niu, M., Yang, Q., Gao, L., Zheng, Y., Yang, J., Wang, Q., 2024: Enhancement of 3D Gaussian Splatting using Raw Mesh for Photorealistic Recreation of Architectures. *arXiv (Cornell University)*. doi.org/10.48550/arxiv.2407.15435.