Photogrammetry and 3D Gaussian Splatting for Cultural Heritage. Pros, Cons and Main Differences

Photogrammetry and 3D Gaussian Splatting for Cultural Heritage. Pros, Cons and Main Differences

Xinchen Li [1], Alessio Martino [1], Filiberto Chiabrando [1], Xiang Li [1]

[1] Department of Architecture and Design (DAD), Politecnico di Torino, Viale Pier Andrea Mattioli, 39, 10125 Torino – (xinchen.li, alessio.martino, filiberto.chiabrando, xiang_li)@polito.it

Keywords: 3D Reconstruction, Photogrammetry, 3D Gaussian Splatting, Cultural Heritage, 3D Surveying

Abstract

This paper presents a comparative analysis of traditional photogrammetric methods and 3D Gaussian Splatting (3DGS) technology in the digitisation of Cultural Heritage (CH). Two representative datasets, differing in scale and image acquisition conditions, were selected to systematically evaluate the performance of both methods in terms of visual quality, geometric accuracy, computational efficiency and stability. The results indicate that 3DGS significantly outperforms traditional photogrammetry methods in terms of rendering quality and real-time visualisation capabilities, generating more realistic and immersive visual effects. However, its geometric accuracy is generally slightly lower than that of traditional methods, a difference that is particularly pronounced in small-scale datasets or under low-resolution input conditions. Among the various implementation methods, Postshot and LichtFeld Studio demonstrated higher stability and robustness, whilst the original GraphDeco method exhibited greater sensitivity to data scale and parameter settings. Photogrammetry offers reliability in high-precision geometric reconstruction, whilst 3DGS demonstrates significant potential for complementing this with a high-fidelity visual experience. The research findings try to provide practical guidance for selecting 3D reconstruction methods across different cultural heritage application scenarios.

1. Introduction

The digital documentation and 3D reconstruction of cultural heritage assets are of great significance for their preservation, dissemination and preventive conservation. Currently, a variety of mature methods have been established for 3D mapping based on active or passive sensors, such as LiDAR (Light Detection and Ranging) technology using active sensors, and photogrammetry methods utilising passive imagery.

Furthermore, SLAM (Simultaneous Localisation and Mapping) technology, which integrates data from multiple sensors, has also been widely adopted. Among these, photogrammetric workflows based on Structure-from-Motion (SfM) and Multi-View Stereo (MVS) have long been established as industry standards. Through feature matching and multi-view geometric reconstruction, these methods can generate dense point clouds and mesh models containing colour information, achieving a high-precision representation of the geometric form of cultural heritage assets (Aicardi et al., 2016).

During the information integration stage, the point clouds or mesh models generated by photogrammetry are typically used to construct Historical Building Information Models (HBIMs), thereby extending the process from geometric reconstruction to semantic representation. However, the HBIM modelling process remains heavily reliant on manual or semi-automated operations, resulting in low efficiency. In recent years, with the advancement of deep learning, researchers have begun to explore the integration of semantic segmentation and structural recognition into the cultural heritage reconstruction workflow.

For example, (Pierdicca et al., 2020) automatically identified architectural components such as bracket sets, beams and columns, providing structured semantic information for HBIM.

Furthermore, in data-constrained scenarios, single-view or sparse-view reconstruction methods have been employed to infer three-dimensional structures from limited or even a single historical image, offering new possibilities for the reconstruction of damaged or lost heritage (Li et al., 2025).

Although traditional photogrammetric techniques demonstrate high reliability in terms of geometric accuracy, certain limitations persist in practical applications. For instance, in areas with weak texture such as large flat surfaces, highly reflective materials like glass or metal, or low-light environments as the interior of a tomb, issues regarding noise or incomplete reconstruction often arise.

Furthermore, as the results are primarily geometry-centric, they struggle to adequately capture the lighting, materials and visual atmosphere of cultural heritage assets. Additionally, traditional workflows have limited capabilities in real-time rendering, making it difficult to meet the demands of interactive displays and immersive applications.

With the rapid development of artificial intelligence, the technical paradigm of 3D reconstruction is gradually shifting from geometry-based modelling methods to learning-based representation methods. Neural rendering, as a key area combining computer graphics and machine learning, achieves new-view synthesis and high-fidelity visual reconstruction by learning scene representations from images (Gao et al., 2026).

Early methods employed multi-layer perceptrons (MLPs) as implicit representations, whereas the recently proposed 3D Gaussian Splatting (3DGS) utilises an explicit radiance field representation, encoding scene information as a set of optimisable three-dimensional Gaussian distributions and achieving efficient real-time rendering through differentiable, tile-based rasterisation. Compared to traditional methods, 3DGS offers significant advantages in terms of visual realism, lighting expression and rendering efficiency, and is better able to capture the material properties and spatial atmosphere of cultural heritage (Chen and Wang, 2024).

Furthermore, generative reconstruction methods (Generative 3D Modelling) are increasingly being applied in the field of cultural heritage. Technologies based on diffusion models and generative artificial intelligence (AIGC) enable inferential reconstruction based on historical documents or surviving information in the absence of original data, providing a new technical pathway for the digital restoration of damaged or lost heritage.

In this context, this paper aims to analyse the applicability and stability of the 3D Gaussian Splatting method across two cultural heritage scenarios, comparing it with traditional photogrammetry methods. It focuses on examining the differences between the two approaches in terms of visual representation, geometric accuracy and application characteristics, assessing the potential of 3DGS in the field of cultural heritage digitisation.

2. Research Progress

Image-based 3D scene reconstruction converts multi-view images or videos into 3D digital models that can be used for analysis, processing and interaction. Early methods primarily relied on SfM and MVS techniques, utilising feature matching and multi-view geometry to recover 3D structures. However, such methods still have certain limitations in areas with weak texture, under complex lighting conditions, and in the synthesis of new views.

With the advancement of deep learning, Neural Radiance Fields (NeRF) have emerged as a significant breakthrough in the field of 3D reconstruction. NeRF utilises multi-layer perceptrons (MLPs) to learn the mapping relationship between spatial coordinates and colour and volume density, implicitly encoding the 3D scene within the neural network and synthesising images from new viewpoints based on volumetric rendering principles.

Compared to traditional point cloud or mesh representations, NeRF is capable of constructing continuous scene functions, thereby achieving high-fidelity visual reconstruction results.

However, this method typically relies on substantial computational resources, incurring high costs during both training and rendering phases, which limits its real-time capabilities and large-scale applications. To address these issues, 3DGS has been proposed as an explicit radiance field representation method.

This approach represents the scene as a set of three-dimensional Gaussian distributions, each containing attributes such as position, scale, orientation, colour and opacity, and is optimised by minimising the error between the rendered image and the actual observation. Combined with differentiable rasterization techniques, 3DGS significantly enhances rendering efficiency whilst maintaining high visual quality, achieving near-real-time performance.

Compared to MLP-based implicit representation methods, 3DGS offers distinct advantages in computational efficiency and visual performance, and has been progressively applied in fields such as virtual reality, digital twins and interactive media (Chen and Wang, 2024)

In recent years, Neural Radiance Fields and 3D Gaussian Splatting have demonstrated immense potential in the field of cultural heritage digitisation, with related research evolving from single-model reconstruction towards diversified applications such as high-precision documentation, intelligent restoration, and immersive dissemination. In terms of digital documentation, NeRF and 3DGS have significantly improved the efficiency and visual realism of cultural heritage modelling.
To address the issue of view-dependent reflections arising from glass display cases, glazed ceramics and polished metals, frameworks such as RAS-NeRF effectively eliminated 'ghosting' and blurring in reconstructions by introducing geometric constraints and adaptive frequency modulation mechanisms, enhancing the ability to perceive complex materials (Wang et al., 2025). Utilising the 3DGS framework, (Li et al., 2025) achieved high-quality geometric detail and texture reconstruction using sparse image data in remote archaeological sites or resource-constrained environments.

As conservation requirements increase, the integration of generative AI with traditional monitoring technologies has emerged as a new trend. (Yu et al., 2025) proposed a complementary approach combining laser scanning with 3DGS, utilising LiDAR to ensure millimetre-level structural accuracy of architectural heritage whilst leveraging 3DGS's strengths in real-time, high-fidelity texture rendering to construct digital archives that are both precise and vivid. Regarding the reconstruction of damaged artefacts, (Stoean et al., 2024) investigated the integration of NeRF rendering models with diffusion models such as Stable Diffusion (SD). By utilising semantic priors to intelligently infer and complete missing geometric structures and textures, this approach has opened up new avenues for the digital restoration of archaeological artefacts.

In the field of heritage dissemination, the achievements of NeRF and 3DGS are being rapidly integrated into modern digital narrative systems. To address performance bottlenecks in the visualisation of large-scale cultural heritage data, research is shifting towards lightweight management based on Gaussian representations, enabling smooth interaction within real-time rendering environments such as Unreal Engine (Baradaran Rahimi et al., 2025; Basso et al., 2024). By providing highly realistic virtual models, these technologies empower virtual museum tours, online education and cross-media game narratives.

Through augmented reality (AR) and interactive scenarios, they not only broaden the public's access to cultural heritage but also significantly enhance the immersive experience and public engagement in heritage dissemination. However, it is important to note that the raw output of 3D Gaussian Splatting is not yet fully mature for rigorous metric documentation. Unlike traditional photogrammetry, 3DGS lacks a continuous surface topology, making it currently less suitable for precise 3D metric representations or the extraction of high-resolution 2D architectural drawings.

3. Materials and Methods

This study evaluates a 3D reconstruction technique for cultural heritage based on neural radiance fields by comparing the reconstruction results of 3DGS with those of traditional photogrammetric methods, to explore the effectiveness and potential advantages of 3DGS in the digitisation of cultural heritage. (Figure 1)

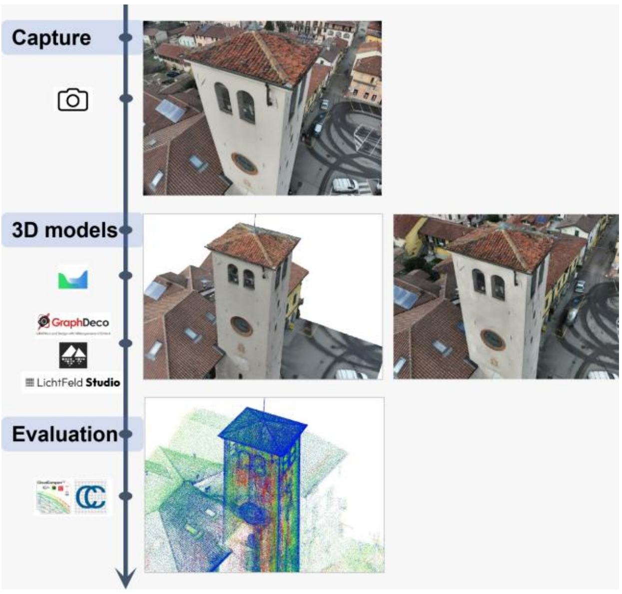

Figure 1. Workflow of the proposed methodology for photogrammetry and 3DGS comparison.

3.1 Data Acquisition and Pre-processing

To ensure a fair and representative comparison, whilst considering the impact of heritage sites of varying scales and complexities on reconstruction methods, this study selected two datasets of cultural heritage sites with distinct scale characteristics for experimentation.

All data were collected using UAV imagery captured by a DJI Mavic 3M, with the raw images having a resolution of 5280 × 3956 pixels. During data acquisition, differentiated flight path strategies were designed to account for the geometric characteristics of different heritage sites, ensuring both the completeness of image coverage and the ability to capture fine details.

Specifically, for the Torre dell'Orologio dataset in Villanova d'Asti, in which a 26 m tall clock tower built in the 15th century has been acquired employing a combination of close-range vertical ascent trajectories (flying at approximately 7.60 meters of distance from the tower) and horizontal circling trajectories was employed to balance the capture of tower details with comprehensive coverage of the overall structure and surrounding environment, resulting in a total of 849 images. In contrast, the Torre Chianca dataset, which represents a 16th-century watchtower, commissioned by Emperor Charles V as part of a strategic coastal defense network designed to protect the Salento peninsula from Turkish and Saracen incursions. Has been digitised by following two circling flight paths with a radius of about 20 meters at different altitudes (approximately 22 and 30 meters) to capture multi-scale information of the tower itself and its surroundings, resulting in a total of 94 images. (Figure 2).

During subsequent processing, to ensure comparability between photogrammetry and 3DGS methods, all input images underwent uniform resolution control and pre-processing. Following extensive testing, ImageMagick was employed to downsample the images: the Torre dell'Orologio dataset was downsampled by a factor of 4 (resize4, 1320x989 px), whilst the Torre Chianca dataset was downsampled by a factor of 2 (resize2, 2640x1978 px). This pre-processing step is a fundamental practice in modern 3D reconstruction pipelines: processing full-resolution, multi-megapixel photographs is heavily demanding on hardware and often contains redundant high-frequency data that does not strictly contribute to the macroscopic 3D structure. Downsampling drastically accelerates model convergence while maintaining sufficient detail. This configuration strikes a balance between computational efficiency and reconstruction quality and was maintained consistently throughout all reconstruction workflows.

Figure 2. Torre dell'Orologio and Torre Chianca datasets.

3.2 3D Reconstruction

For 3D reconstruction, this study employs two typical paradigms to model and conduct a comparative analysis of cultural heritage objects: namely, a photogrammetric method combining traditional SfM with MVS, and the 3DGS method based on radiative field representation.

3.2.1 Photogrammetric Reconstruction: For what concerns the photogrammetric reconstruction, Agisoft Metashape 2.2.1 was used to implement a standard SfM and MVS workflow. The traditional approach was followed: first, a correspondence between the multi-view images was established through image feature detection and matching, and the camera's internal and external orientation parameters, as well as its pose, were estimated. Subsequently, a 3D mesh model is constructed using a multi-view stereoscopic reconstruction method based on depth maps.

Since before the process a uniform downsampling pre-processing of the input imagery was performed via

ImageMagick, the highest quality parameters provided by the software are employed at every key stage of the photogrammetric workflow to avoid further resolution compression within the software and to provide a relatively reliable reference model for subsequent geometric accuracy assessments.

3.2.2 3D Gaussian Splatting: Regarding 3DGS reconstruction, this study is based on the methodological framework proposed by Kerbl et al., 2023, which achieves high-quality real-time rendering within a short training period through explicit 3D Gaussian representations, interleaved optimisation and density control mechanisms, and a visibility-oriented fast rendering algorithm. To evaluate the performance differences between various implementation strategies, this paper compares three 3DGS algorithms: the original open-source GraphDeco implementation, LichtFeld Studio, and the commercial software Jawset Postshot v1.0.110.

In the GraphDeco workflow, COLMAP 3.8 (Schönberger and Frahm, 2016) is first utilised for camera parameter estimation and sparse point cloud reconstruction, with the results converted into the format required by 3DGS. Concurrently, GNSS-based georeferencing is introduced to ensure model scale consistency. The training process is conducted on an NVIDIA A100 GPU within the Google Colab platform, using default parameters and running 30,000 iterations, ultimately generating a 3D Gaussian model and its configuration file in .ply format.

In the workflows of the Postshot and LichtFeld, both first pre-process the input images using RealityScan to obtain camera parameters and pose information in COLMAP format, as well as a sparse point cloud to use as a base for the 3DGS training, which Agisoft Metashape does not allow directly to export. Subsequently, model construction is completed using default training parameters within their respective environments, and a .ply file in a standardised format is output for subsequent analysis.

3.3 Model Evaluation

During the model evaluation phase, this study conducted a comprehensive analysis of the applicability of the 3DGS method in cultural heritage scenarios, focusing on two aspects: geometric accuracy and visual rendering quality.

Firstly, regarding the assessment of geometric accuracy, a 3D mesh model generated using photogrammetry (Agisoft Metashape) was used as a reference benchmark to perform a quantitative analysis of the geometric deviations in the 3DGS model. Specifically, CloudCompare v2.13.2 was employed to pre-process the 3DGS point cloud, including cropping, Statistical Outlier Removal (SOR) and scalar value filtering, to extract the main subject area with clear structural features.

Subsequently, homogeneous feature points were manually selected using the point-to-point alignment method to achieve initial registration between the 3DGS point cloud and the reference mesh model, since no control points were measured for the scope of this research. On this basis, the Cloud-to-Mesh Distance (C2M) method was utilised to calculate geometric deviations, and the mean distance, standard deviation, and a heatmap of the spatial distribution of errors were generated, thereby systematically evaluating the geometric accuracy of the 3DGS model.
Regarding visual quality assessment, this study combines both subjective and objective methods to conduct a comparative analysis of the rendering effects of different 3DGS models. On the one hand, manual visual inspection is used to conduct a qualitative assessment of the model's overall realism, texture clarity and detail representation. On the other hand, multiple image quality metrics were introduced for quantitative analysis, including Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM) and Learned Perceptual Image Patch Similarity (LPIPS) (Zhang et al., 2018), to comprehensively measure the visual fidelity of the models in the texture reconstruction of cultural heritage (Croce et al., 2024).

4. Results and Discussion

4.1 Photogrammetric Results Analysis

Regarding photogrammetric reconstruction results, differences between the two datasets in terms of acquisition scale and the number of images affected both model quality and processing efficiency (Table 1). The Torre dell'Orologio dataset (849 images), acquired at a low altitude and on a closer range, achieved a higher ground resolution (8.21 mm/pixel) and a denser distribution of tie points (310,090), providing a clear advantage in terms of detail representation. However, the time required for registration and depth map calculation for this dataset increased significantly (11 min 12 s and 10 min 48 s respectively), resulting in a relatively high computational cost.

In contrast, the Torre Chianca dataset (94 images) was acquired at a higher flight altitude (30.8 m) and has a lower ground resolution (1.66 cm/pixel), but its reconstruction process is more efficient (39 s for registration and approximately 5 min for depth map generation), demonstrating better computational efficiency even when handling high-resolution images (5280 × 3956 pixels) and processing them using the High accuracy parameter.

It is worth noting that although the high-resolution dataset is finer at the input scale, its final number of mesh faces (1,791,053 faces) is actually slightly lower than that of the low-resolution dataset (2,064,729 faces). This indicates that mesh complexity is not only influenced by image resolution but is also closely related to factors such as scene coverage, viewpoint redundancy and reconstruction parameters. Furthermore, the reprojection errors for both datasets remained at the sub-pixel level (approximately 0.3–0.6 pixels), indicating that the photogrammetric workflow exhibits good geometric stability at both scales. Overall, high-density image acquisition helps to enhance detail representation but significantly increases computational costs, whereas smaller-scale datasets demonstrate a better balance between efficiency and accuracy; this difference provides an important reference for subsequent comparisons with the 3DGS method. (Figure 3)

Features	Torre Chianca	Torre dell'Orologio
Image pre-processing	resize 2	resize 4
Number of images	94	849
Average flying distance	20 m	7.60 m
Ground Sample resolution	1.66 cm/pix	8.21 mm/pix
Coverage area	0.0213 km^2	958m^2
Tie Points	58,376 of 69,898	310,090 of 430,423
Alignment time	39 s	11 m 12 s
Depth maps processing time	5 m 4 s	10 m 48 s
Model faces	2,064,729	1,791,053

Table 1. Photogrammetric reconstruction statistics based on different dataset.

Figure 3. Photogrammetric model results.

4.2 3DGS Models Evaluation

Visual Quality Assessment: As shown in Table 2, there are marked differences between the various 3DGS implementation methods in terms of training efficiency, initialisation and rendering quality. Specifically, both Postshot and LichtFeld Studio are capable of generating clear and stable rendering results on both datasets, with a high degree of consistency in building edges, textural details and overall structure. In contrast, the GraphDeco method performs significantly worse on both datasets, producing output that is generally blurred and lacking in detail. (Figure 4)

In terms of the training process and time costs, the GraphDeco method utilised the default 30,000 iterations for training on both datasets. Training on the Torre dell'Orologio dataset took approximately 33 minutes, whilst training on the Torre Chianca dataset took approximately 1 hour and 3 minutes, demonstrating significant data dependency and instability. At the same time, there were differences in convergence quality across the two datasets, with poorer model optimisation results observed on the smaller dataset. In contrast, Postshot's training times on the two datasets were approximately 1 hour 8 minutes and 55 minutes respectively, with approximately 30k and 43k iterations, indicating a relatively stable training process overall.

LichtFeld's training times were approximately 41 minutes and 49 minutes respectively, demonstrating consistent computational efficiency across datasets of different sizes. In terms of initialisation, both Postshot and LichtFeld are trained using sparse point clouds generated by RealityScan; compared to GraphDeco, which relies directly on COLMAP output, their initial conditions are more uniform. Regarding training parameter settings, GraphDeco uses 'resolution=1', i.e. the resolution of the original input image, without additional downsampling, whilst Jawset Postshot and LichtFeld Studio both employ default adaptive strategies, including control of the number of Gaussians and adjustment of the number of training steps.

Dataset	Torre Chianca (94)		
Method	GraphDeco	Jawset Postshot	LichtFeld Studio
Image Input	resize2 (res=1)	resize2	resize2
Iteration	30k	~30k	30k
Training Time	1h 03m	1h 08 m	41m 12s
Initialization	~90k pts (COLMAP)	~450k pts (RealityScan)	RealityScan
Rendering Quality	Blurry, unstable	Sharp, stable	Sharp, stable

Dataset	Torre dell'Orologio (849)		
Method	GraphDeco	Jawset Postshot	LichtFeld Studio
Image Input	resize4 (res=1)	resize4	resize4
Iteration	30k	~43k	~85k
Training Time	33m 11s	55m 12s	49m 42s
Initialization	~600k pts (COLMAP)	~450k pts (RealityScan)	RealityScan
Rendering Quality	Blurry, unstable	High quality	High quality

Table 2. Comparison of 3DGS methods in the two datasets

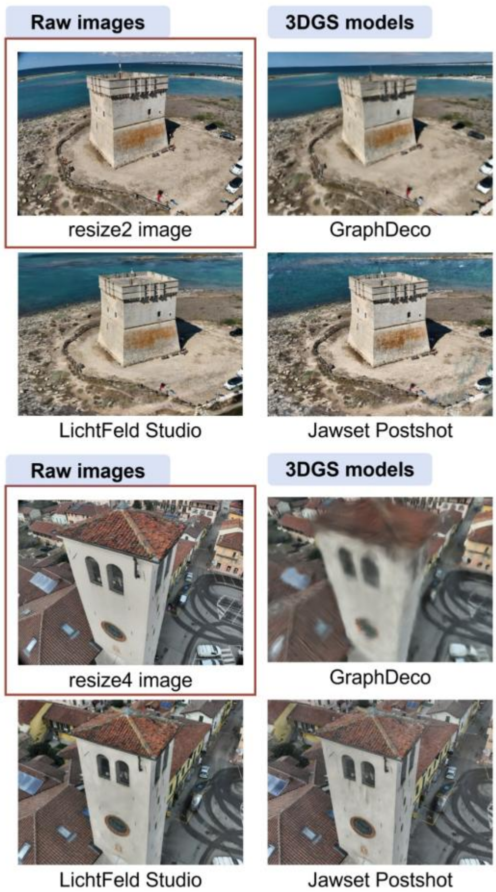

Figure 4. Visual comparison of different 3DGS methods.

The results indicate that although GraphDeco, as the original implementation, offers greater flexibility, it is more sensitive to input resolution and data scale in practical applications, whereas

Postshot and LichtFeld demonstrate greater robustness in terms of parameter automation and training stability.

To further validate the stability and applicability of the GraphDeco method across different data scales and parameter settings, this paper conducts supplementary experiments on the input image scale and training resolution parameters.

Quantitative results (Table 3) indicate that, on the small-scale Torre Chianca dataset, different parameter combinations have a limited impact on model performance; SSIM consistently remains between 0.45 and 0.51, PSNR is approximately 20.9–21.7, and LPIPS is approximately 0.55–0.59. Overall performance is relatively consistent but remains at a low level.

The corresponding rendering results also reveal that the model generally suffers from visual blurring, with difficulty in restoring fine details. In contrast, on the large-scale Torre dell’Orologio dataset, the ‘noresize’ approach combined with a moderate training resolution yields the best results (SSIM ≈ 0.91– 0.92, PSNR ≈ 29.6– 29.9, LPIPS ≈ 0.20– 0.22). These results further indicate that the GraphDeco method is highly dependent on data scale and input resolution; its stability and generalisation capabilities are limited on small-scale datasets, whereas superior rendering quality can only be achieved under conditions of large-scale, high-resolution inputs (Figure 5). Overall, in the context of cultural heritage digitisation, the actual performance of the 3DGS method depends not only on the algorithm itself, but also to a large extent on the specific implementation strategy and data pre-processing workflow.

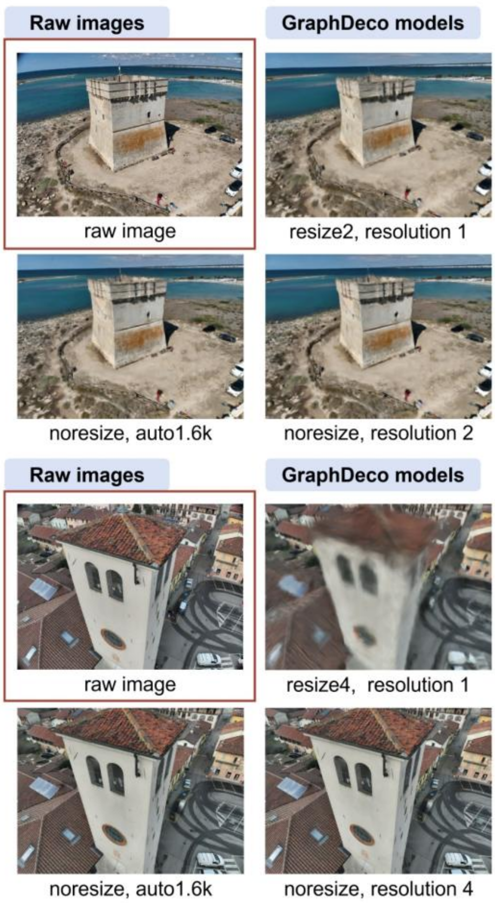

Figure 5. GraphDeco rendering quality effect.

Dataset	Image Input	Resolution	SSIM	PSNR	LPIPS
Torre Chianca	resize 2	resolution 1	0.51	21.47	0.56
	noresize	auto (1.6k)	0.51	21.74	0.55
		resolution 2	0.45	20.87	0.59
		resolution 1	out of memory		
Torre dell'Orologio	resize 4	resolution 1	0.62	19.81	0.51
	noresize	auto (1.6k)	0.91	29.59	0.22
		resolution 4	0.92	29.87	0.20
		resolution 2	out of memory		

Table 3. Different 3DGS models visual performance.

4.2.1 Geometric Accuracy Comparison with Photogrammetry Models: In the geometric accuracy assessment, a mesh model generated using photogrammetry was used as a reference. CloudCompare was employed to calculate the C2M distance between the 3DGS point cloud and the reference model, thereby enabling a quantitative analysis of the geometric consistency of the different methods (Table 4)

Dataset	Torre Chianca			
N° of Images	94			
Image Input	resize 2			noresize
Resolution	resolution 1	/	/	resolution 2
Methods	GraphDeco	Postshot	LichtFeld	GraphDeco
N° of point	178,225	1,509,127	789,289	90,316
Mean distance [m]	0.49	0.2	0.14	1.6
STD deviation [m]	0.51	0.33	0.1	1.02

Dataset	Torre dell’Orologio			
N° of Images	849			
Image Input	resize 4			noresize
Resolution	resolution 1	/	/	resolution 4
Methods	GraphDeco	Postshot	LichtFeld	GraphDeco
N° of point	71,884	1,129,966	510,354	1,262,633
Mean distance [m]	0.5	0.17	0.1	0.13
STD deviation [m]	0.36	0.31	0.14	0.14

Table 4. Different 3DGS models geometric accuracy for the two datasets.

The results indicate that there are significant differences in geometric accuracy among the various 3DGS implementation methods. In the Torre Chianca dataset, GraphDeco achieved an average distance of 0.49 m (STD=0.51 m) under the resize2 and resolution=1 conditions, which is significantly higher than that of Postshot (Mean=0.20 m, STD=0.33 m) and LichtFeld Studio (Mean=0.14 m, STD=0.10 m). Under the noresize, resolution=2 conditions, GraphDeco’s error further increased (Mean=1.60 m, STD=1.02 m), demonstrating lower geometric stability. In contrast, both Postshot and LichtFeld Studio exhibited smaller error ranges and more uniform distributions on this dataset. (Figure 6)

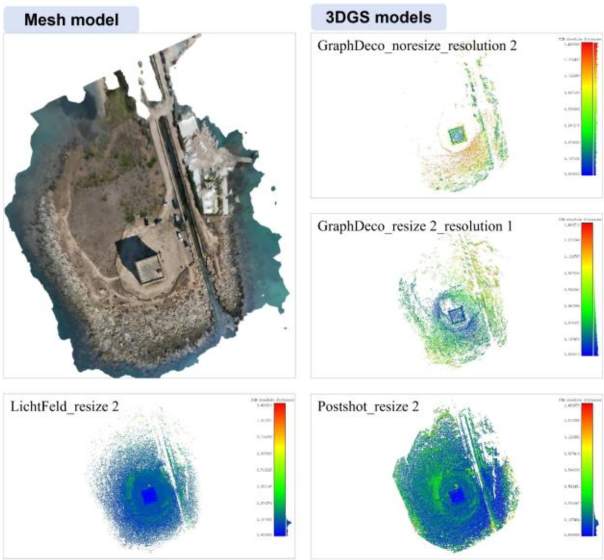

Figure 6. C2M distance comparison between photogrammetry and 3DGS models on Torre Chianca dataset.

In the Torre dell'Orologio dataset, the overall error of all methods was significantly reduced. GraphDeco achieved an average distance of 0.50 m (STD=0.36 m) under the conditions of resize4 and resolution=1, whilst this was reduced to 0.13 m (STD=0.14 m) under the conditions of noresize and resolution=4, indicating an improvement in geometric accuracy under high-resolution input conditions.

However, even under optimal settings, its accuracy remains slightly lower than that of LichtFeld Studio (Mean = 0.10 m, STD = 0.14 m). Overall, LichtFeld Studio achieved the best or near-best geometric accuracy in both datasets, with a more concentrated error distribution and a smaller standard deviation, demonstrating superior stability. (Figure 7)

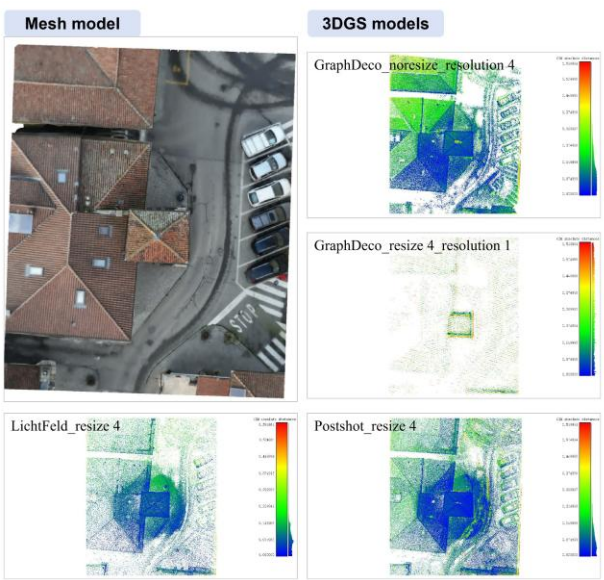

Figure 7. C2M distance comparison between photogrammetry and 3DGS models on Torre dell'Orologio dataset.

A closer examination of the spatial error distribution reveals that the geometric errors of the GraphDeco model exhibit a clear dependence on parameters. Within the same methodological framework, models generated at different resolution settings show significant differences in both the morphology and numerical range of error distributions, with marked variations in local structural displacement and point cloud dispersion.

This indicates that data scale primarily influences the overall level of error, whilst parameter settings largely determine the stability of the geometric structure and the quality of the reconstruction. In contrast, the point clouds generated by Postshot and LichtFeld exhibit more consistent error distribution characteristics under varying data conditions; errors are primarily concentrated along edges and in non-structured regions, whilst the overall structure maintains a high degree of consistency. Overall, the 3DGS method still lags behind traditional photogrammetry results in terms of geometric reconstruction accuracy. Therefore, while it is currently not viable for survey-grade 3D reconstruction or precise architectural documentation, it is highly successful as a visualization tool. Its output is more than sufficient for applications prioritizing visual realism over metric exactness, such as extended reality (XR) applications and digital heritage dissemination, with variations observed between different implementations of the Gaussian model.

5. Conclusions

This study focuses on the 3D reconstruction of cultural heritage sites, conducting a comparative analysis of traditional photogrammetry methods and the 3D Gaussian Splatting method. It evaluates the performance of these two approaches, with particular emphasis on visual quality, geometric accuracy, computational efficiency and method stability. The experimental results indicate that each method has its own strengths in different aspects and is suitable for different application requirements. (Table 5)

Aspect	Photogrammetry (SfM+MVS)	3D Gaussian Splatting
Visual Quality	Moderate, texture dependent	High, realistic rendering
Geometric Accuracy	High, reliable	Moderate
Model Representation	Mesh and point cloud	Gaussian point
Rendering	Offline	Real-time
Processing Time	Moderate (data dependent)	Moderate (training and fast rendering)
Stability	High	Sensitive to data and parameters
Data Requirement	Flexible	Sensitive to data scale
Workflow Control	High (transparent pipeline)	Varies (open-source vs commercial)
Ease of Use	Moderate (manual tuning required)	High (commercial tools), lower (open-source)
Application Suitability	Surveying, documentation	Visualization, XR, interactive applications

Table 5. Comparison between photogrammetry and 3DGS in cultural heritage applications

In terms of visual presentation, the 3DGS method offers distinct advantages in rendering realism and detail, capable of generating more continuous and natural textural effects, particularly in scenes with complex lighting and surface details.

By contrast, whilst mesh models generated by photogrammetry possess high structural integrity, they are relatively limited in terms of textural representation. In terms of geometric accuracy, photogrammetry continues to demonstrate greater stability and precision; the mesh models it generates serve as reliable geometric references. Whilst 3DGS models are generally of a usable standard in most cases, their accuracy remains subject to the scale of the data and the parameter settings.

In terms of computational efficiency and workflow, the two methods exhibit distinct characteristics. The photogrammetry workflow is relatively mature and produces predictable results, but typically requires longer processing times, with computational costs being particularly high under high-resolution and large-scale data conditions. The 3DGS method offers high efficiency during the training and rendering stages and supports real-time rendering; however, its performance is relatively sensitive to the quality of input data and parameter settings. In particular, GraphDeco exhibits lower stability with small-scale datasets, whilst Jawset Postshot and LichtFeld Studio enhance robustness through automated workflows, albeit at the cost of reduced process controllability and interpretability to some extent.

Overall, photogrammetry and 3DGS methods are not mutually exclusive; rather, they represent two complementary technical approaches. Photogrammetry is better suited to scenarios requiring high geometric accuracy, such as surveying and heritage documentation, whilst 3DGS holds greater potential for visual presentation, interactive applications and real-time rendering.

Therefore, in the context of cultural heritage digitisation, methods should be selected based on specific requirements to implement appropriate technical application strategies. Looking forward, a highly promising avenue for future research is the evaluation of hybrid 'mesh-to-Gaussian' pipelines. By leveraging an accurate photogrammetric mesh as the foundational geometry for Gaussian splatting, it may be possible to bridge the gap between these technologies, guaranteeing both rigorous survey-grade spatial accuracy and exceptional real-time visual quality.

References

Aicardi, I., Chiabrando, F., Grasso, N., Lingua, A.M., Noardo, F., Spanò, A., 2016. Uav photogrammetry with oblique images: first analysis on data acquisition and processing. https://doi.org/10.5194/isprsarchives-XLI-B1-835-2016

Baradaran Rahimi, F., Demers, C.M.H., Karimi Dastjerdi, M.R., Lalonde, J.F., 2025. Agile digitization for historic architecture using 360° capture, deep learning, and virtual reality. Autom. Constr. 171, 105986. https://doi.org/10.1016/J.AUTCON.2025.105986

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. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLVIII-2-W4-2024, 57–64. https://doi.org/10.5194/isprs-Archives-XLVIII-2-W4-2024-57-2024

Chen, G., Wang, W., 2024. A Survey on 3D Gaussian Splatting. 2025. A Survey on 3D Gaussian Splatting 1, 35. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Croce, V., Billi, D., Caroti, G., Piemonte, A., De Luca, L., Véron, P., 2024. Comparative Assessment of Neural Radiance Fields and Photogrammetry in Digital Heritage: Impact of Varying Image Conditions on 3D Reconstruction. Remote Sensing 2024, Vol. 16, Page 301 16, 301. https://doi.org/10.3390/RS16020301

Gao, K., Gao, Y., He, H., Lu, D., Xu, L., Li, J., 2026. NeRF: Neural Radiance Field in 3D Vision: A Comprehensive Review (Updated Post-Gaussian Splatting).

Kerbl, B., Kopanas, G., Leimkuehler, T., Drettakis, G., 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Trans. Graph. 42, 14. https://doi.org/10.1145/3592433

Li, C., Sui, D., Li, J., Li, Yacong, Li, Yang, Guo, M., 2025. Frequency and tensorial neural radiance fields for digital preservation of ancient buildings. npj Heritage Science 2025 13:1 13, 158-. https://doi.org/10.1038/s40494-025-01695-x

Pierdicca, R., Paolanti, M., Matrone, F., Martini, M., Morbidoni, C., Malinverni, E.S., Frontoni, E., Lingua, A.M., 2020. Point Cloud Semantic Segmentation Using a Deep Learning Framework for Cultural Heritage. Remote Sensing 2020, Vol. 12, Page 1005 12, 1005. https://doi.org/10.3390/rs12061005

Schonberger, J.L., Frahm, J.M., 2016. Structure-from-Motion Revisited. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2016-December, 4104–4113. https://doi.org/10.1109/CVPR.2016.445

Stoean, R., Bacanin, N., Stoean, C., Ionescu, L., 2024. Bridging the past and present: AI-driven 3D restoration of degraded artefacts for museum digital display. J. Cult. Herit. 69, 18–26. https://doi.org/10.1016/j.culher.2024.07.008

Wang, Y., Sui, D., Li, Yacong, Li, Yang, Guo, M., 2025. RAS-NeRF for novel view synthesis of complex reflective artifacts. npj Heritage Science 2025 13:1 13, 161-. https://doi.org/10.1038/s40494-025-01727-6

Yu, Y., Verbree, E., van Oosterom, P., Pottgiesser, U., Peng, Y., Poux, F., 2025. From comparison to integration: A workflow evaluation of 3D Gaussian splatting and LiDAR point cloud for modern architectural heritage. Autom. Constr. 180, 106509. https://doi.org/10.1016/j.autcon.2025.106509

Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O., 2018. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 586–595. https://doi.org/10.1109/CVPR.2018.00068