Combining Photogrammetry and Gaussian Splatting

Fabio Remondino[1], Elisa Mariarosaria Farella[1], Gianluca Bertolasi[1], Simone Rigon[1], Rongjun Qin[2]

[1] 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy
Email: (remondino, elifarella, gbertolasi, srigon)@fbk.eu

[2] Department of Civil, Environmental and Geodetic Engineering; Department of Electrica and Computer Engineering, The Ohio State University, Columbus, Ohio, USA - Email: qin.324@osu.edu

Keywords: Photogrammetry; Gaussian Splatting; fusion; non-collaborative surfaces

Abstract

Among the image-based methods, photogrammetry is a consolidated 3D reconstruction technique able to provide highly accurate metric products, widely exploited in many domains. Photogrammetry is, however, conditioned by the characteristics of the captured scene, with good performance in well-textured areas and limits when non-collaborative surfaces, such as reflective or transparent, are present. In such cases, the photogrammetric reconstruction is often affected by noise, incomplete geometry and artifacts, reducing its final reconstruction quality. In recent years, different AI-based reconstruction methods have emerged as alternative (or complementary) 3D reconstruction and rendering solutions. In particular, 3D Gaussian Splatting (GS) has demonstrated impressive capabilities in rendering photorealistic scenes in challenging situations with high visual fidelity. However, its application in large-scale scenarios or when highly accurate 3D metric products are required is still limited, due to the high computational resources needed and the intrinsic optimization of GS methods for photometric rendering quality. To address these bottlenecks, this work proposes a hybrid reconstruction pipeline, leveraging the strengths and benefits of each technique. The method exploits the accurate geometry of photogrammetry in well-textured regions and the GS capabilities to improve completeness and visual aspect in areas featuring non-collaborative surfaces. A fusion strategy is proposed to combine the two results into a single 3D model, presenting examples from two aerial and one terrestrial dataset.

1. Introduction

Image-based 3D reconstruction has become an indispensable and requested task for urban planning (Agugiaro et al., 2020), autonomous driving (Yan et al., 2025a), heritage 3D documentation (Farella et al., 2022), digital twin creation (Jeddoub et al., 2023), etc. Advances in photogrammetry, laser scanning and, more recently, neural scene representations (Mildenhall et al., 2020; Tewari et al., 2022; Kerbl et al., 2023) have supported these purposes and intensely improved the realism and geometric fidelity of 3D scenarios (Chen et al., 2025). Conventional photogrammetry workflows, based on feature extraction, bundle adjustment, multi-view stereo (MVS) and surface meshing, are today widely adopted due to their geometric accuracy, efficiency, reliability and compatibility with established professional pipelines. The recent emergence of deep learning has fundamentally changed 3D processing workflows (Stathopoulou and Remondino, 2023; Liu et al., 2023; Akhavi et al., 2025; Wang et al., 2026). Modern AI-based approaches treat multi-view reconstruction as an end-to-end optimization problem, based on implicit representations, leveraging neural networks to learn robust feature representations, correspondence estimation and 3D data estimation directly from data (Wang et al, 2025; Yang et al., 2025). Additionally, differentiable rendering pipelines allow to optimize 3D predictions directly from image-level supervision, reducing the dependence on ground-truth geometry (Gao and Qi, 2024). However, image-based methods often struggle in complex scenarios featuring low-texture areas, reflective objects, transparent and semi-transparent materials, repetitive patterns, thin structures or fine-scale appearance variations. The optical behaviour of most of these surfaces violates the basic assumptions of most optical and vision-based methods, which typically rely on Lambertian reflectance models. As a result, significant gaps in both geometry and appearance are present, notwithstanding noise, over-smoothing or artifacts in visually intricate regions (Figure 1b). But the landscape of image-based 3D reconstruction is rapidly transforming, boosted by the latest developments in neural scene representations. Among them, Neural Radiance Fields - NeRF (Mildenhall et al., 2020) and 3D Gaussian Splatting - GS (Kerbl et al., 2023; Fei et al., 2024; Chen et al., 2025) recently emerged as particularly influential methods. These methods do not use explicit geometry (points or meshes) but learns how the object looks from any viewpoint. NeRF is an implicit scene representation and synthesises novel views of complex scenes by optimising an underlying continuous volumetric scene function from a set of oriented input images using a deep, fully connected neural network (often referred to as a multilayer perceptron, or MLP). The neural network is queried at multiple points along each camera ray, with colour and density values integrated to render the final pixel value (Remondino et al., 2023; Liao et al., 2026). On the other hand, GS has emerged as a highly efficient alternative to implicit radiance fields, enabling novel view synthesis with almost real-time rendering capabilities, high visual fidelity and scene editing capabilities (Dalal et al., 2024). Following the literature, 3D GS has rapidly replaced NeRF in many applications, mainly because it dramatically improves speed and usability while keeping comparable visual quality. By representing a scene as a set of anisotropic gaussians with optimizable color, opacity, and spatial covariance, GS methods (Huang et al., 2024; Jiang et al., 2024, Xie et al., 2024) provide a continuous volumetric description that excels at reproducing fine appearance details and view-dependent effects (Figure 1c).

1.1 Aim of the work

While GS models offer impressive visual realism, they generally lack the explicit metric guarantees and geometric quality that are central to photogrammetric applications. Moreover, many areas of a scene are generally well reconstructed with conventional photogrammetric methods (MVS), whereas non-collaborative surfaces generally produce noisy point clouds or large gaps. Therefore, such contrast presents a timely research opportunity: can the complementary strengths of GS and conventional photogrammetry be combined into a unified 3D reconstruction framework? Photogrammetry provides reliable camera geometry, scalable multi-view consistency and a metrically accurate triangulation principle for well-textured areas, while GS contributes with detailed radiance modelling, high rendering efficiency and dense scene representation on areas with thin

a) b) c)
Figure 1. A thin metallic structure (a) reconstructed with a conventional photogrammetric MVS method (b) and GS (c).

	GS	NeRF
Scene Representation	Explicit: the scene is modelled as a collection of millions of individual 3D Gaussians (ellipsoids)	Implicit: the scene is modelled as a continuous function learned by a neural network
Rendering Process	Rasterisation: the final pixel colour is determined by directly projecting or "splatting" the Gaussians onto the 2D image plane	Volume rendering: a neural network is queried at multiple points along each camera ray, with colour and density values integrated to render the final pixel colour.
Training Time	Fast: not really a training, more an optimisation of Gaussian parameters	Slow: very computationally intensive, can take hours or even days
Rendering Speed	Almost real-time: it could achieve very high frame rates (90+ FPS)	Slow: the original NeRF was very slow, taking seconds per frame
Memory Usage	High: 3DGS requires a lot of VRAM to store all the parameters for millions of Gaussians	Low: trained models are relatively small neural networks, which are very memory efficient
Core Technology	Differentiable rasterisation & gradient-based optimisation	Neural networks & differentiable volume rendering

Table 1: Main characteristics and differences between NeRF and GS.

structures or not cooperative surfaces. Smartly integrating the two paradigms has the potential to produce 3D assets that are both geometrically robust and visually compelling, bridging a gap existing in actual image-based 3D modelling. The aim of this work is to present the developments and experiences of such fusion methodology.

2. Gaussian Splatting

Though NeRF dominated the research landscape until 2023, the community’s attention shifted in mid-2023 toward Gaussian Splatting methodologies, marking a further paradigm shift in 3D reconstruction applications. Unlike NeRF (Table 1), which defines a scene as a continuous function optimized across space, or traditional graphics that uses a discrete mesh, GS uses a large collection of explicit, anisotropic 3D Gaussian primitives (also called splats) that collectively approximate the scene's geometry and appearance. These primitives are differentiable ellipsoids defined by a set of parameters for each individual Gaussian in the scene, including: the 3D point of the Gaussian (x, y, z); shape, orientation and scale of the ellipsoid embedded in a covariance matrix; the opacity governing the transparency of the Gaussian (0 for empty space, 1 for solid surfaces); some coefficients (Spherical Harmonics - SH) used to model view-dependent colours and light reflections across the surface of the Gaussian. The number of SH coefficients depends on the degree (or order) of the SH used to represent the colour and view-dependency of each Gaussian. Standard implementations use a maximum degree of 3, implying 3 coefficients (degree 0) to represent the base colour and 45 coefficients (degrees 1, 2 and 3) to represent the view-dependent light reflection effects, such as shading, reflections, and directional lighting. These SH coefficients allow the Gaussian to change its colour and lighting depending on the viewing angle. Higher SH degrees allow for more complex reflections but significantly increase the learning approach and final file size. The learning and rendering approaches are performed through an initialization and optimization stage, designed for rapid convergence and real-time performance. The former leverages the sparse point cloud of the photogrammetric orientation step, whereas the latter optimizes the parameters of every Gaussian in the scene through a differentiable pipeline using an image loss function which compares the rendered view to the actual input image.

Pan et al. (2025) proposed a novel initialization pipeline able to achieve high fidelity reconstruction from dense image sequences without relying on the sparse point clouds of the image orientation phase. Ji and Yao (2025) eliminated the need for known camera poses, introducing a hierarchical training strategy that trains and merges multiple 3D Gaussians into a single, unified 3DGS model representing the entire scene.

3DGS methods are generally tailored for small scenarios, with fast training convergence, high rendering efficiency and competitive visual fidelity (Xu et al., 2024; Yu et al., 2024). City-scale approaches are also demonstrating impressive visual results (Lin et al., 2024; Liu et al., 2024; Xie et al., 2025), although scaling up to large urban datasets raises issues related to memory usage. Indeed, practical deployment demands efficient training and compression strategies: large-scale GS reconstructions can require tens of millions of Gaussians, consuming over 30 GB of GPU memory and reducing rendering speeds below real-time (Yan et al., 2026).

Different studies have proposed solutions to improve rendering and reconstruction quality (Lu et al., 2024; Peng et al., 2024; Steiner et al., 2025) and efficiency (Fan et al., 2024; Lee et al., 2024; Hahlbohm et al., 2026). In Liu et al. (2025), the Gaussian kernels for geometry and low-order Spherical Harmonics (SH) for color encoding are replaced with Deformable Beta Splatting (DBS), a deformable and compact approach that enhances both geometry and colour representation.

3DGS, as a primitive- point-based representation, remains often incompatible with mesh-based pipelines, e.g. in AR/VR environments or for game engines. Therefore, a mesh-based reconstruction approach that jointly optimizes geometry and appearance through differentiable rendering (Held et al., 2026) or a differentiable renderer that directly optimizes triangles via end-to-end gradients (Held et al., 2025a) is proposed. Guédon et al. (2025) suggested a fully differentiable procedure that constructs the mesh - including both vertex locations and connectivity - at every iteration directly from the parameters of the Gaussians. 3D smooth convexes are also proposed as primitives for modelling geometrically-meaningful radiance fields from multi-view images (Held et al., 2025b).

3. Integrating photogrammetry and 3DGS

Integrating photogrammetry and Gaussian Splatting is currently one of the most promising directions for achieving complete, high-quality 3D scene reconstruction in an efficient way. The aim is to preserve the accurate geometry of the photogrammetric dense reconstruction in regions where conventional MVS techniques perform reliably, while replacing areas where these methods fail, and Gaussian Splatting provides improved results in terms of completeness of the geometry and appearance.

The proposed integration methodology (Figures 2 and 3) consists of the following steps:

- Image orientation: the available images are oriented in order to recover precise camera poses and a sparse point cloud of the scene (Figure 2b). This step defines the common reference system for both photogrammetric and Gaussian Splatting processes.
- Dense point cloud generation (Figure 2c): through MVS methods, a dense 3D point-based representation of the scene is generated. The dense output includes accurate geometry in well-textured regions with collaborative surfaces, as well as potentially incomplete and noisy geometry in reflective, transparent or low-texture areas.
- Identify regions of interest (Figure 2d): areas where GS should be applied to improve the quality/completeness of the digital scene are masked using two alternative methods:
  - image segmentation method: based on Multimodal Large Language Models - MLLMs, e.g., SAM2 (Ravi et al., 2024), SAM3 (Carion et al., 2025) or Sa2Va (Yuan et al., 2025), specific areas are identified with dedicated queries and prompts to guide their segmentation (Figure 4).
  - 3D classification method: based on supervised approaches, areas not well reconstructed with conventional MVS methods are isolated in the 3D dense point cloud; classes are project-based defined (Figure 5).
- GS applied to the masked regions: different methods are tested, including 2D Gaussian Splat (Huang at al., 2024), BlockGaussian (Wu et al., 2025) and Jawset PostShot[1], to generate a 3D Gaussian primitives-based representation of the areas, encoding the 3D position (X,Y,Z), normals (Nx, Ny, Nz), radiance coefficients represented through Spherical Harmonics and splat parameters (opacity, scale, rotation).
- Post-processing and cleaning of GS results to remove unwanted splats.
- Conversion of photogrammetric points: the photogrammetric dense 3D point-based representation is converted into Gaussian primitives to enable joint visualization and integration. The conversion includes:
  - computation of the 3 colour component (SH degree 0) from the RGB values of the MVS points: the RGB values are first normalized in the range [0,1] (Eq.1):

$$colour_j = \frac{RGB_j}{255} \qquad \text{(Eq.1)}$$

  and mapped to the principal component of the SH (Eq.2):

$$f_{dc,j} = \frac{\left(colour_j - \frac{1}{2}\right)}{C_0}, \quad j \in \{R, G, B\} \qquad \text{(Eq.2)}$$

  where $C_0 \approx 0.2821$ is the 0-degree SH basic function, used to normalize the principal colour component. This guarantees that photogrammetric colours are compatible with Gaussian radiance representation and prevents excessive and unrealistic brightness.
  - computation of higher-degree components: the other 45 SH coefficients are all initialized to zero. Photogrammetric points, however, are not part of the original GS training and do not have radiance information. Therefore, these coefficients cannot be reliably estimated and are initialized to zero during the integration, providing a neutral contribution without introducing artifacts.
- Opacity (α) setting: for each Gaussian, it is defined as a function of the global dense point cloud density (Eq.3), ensuring that even sparser regions remain relatively opaque:

$$\alpha = \log\frac{1}{\mathrm{d}} \qquad \text{with } \alpha_{min} \le \alpha \le \alpha_{max}$$
$$d = \frac{1}{N}\sum_{i=1}^{n} \| p_i - p_{i,NN} \| \qquad \text{(Eq.3)}$$

  where d is the average nearest-neighbour distance, defined as the Euclidean distance between each point p_i and its closest neighbour $p_{i,NN}$.
- Scale (s) setting: it is set according to the global average nearest-neighbour distance, reflecting the prevalent point spacing (Eq.4).

$$s = \log(d) \qquad \text{(Eq.4)}$$

- Rotations settings: they are set as identity quaternions, since photogrammetric points do not provide reliable local orientation information.
- Export and visualization: the merged dataset is exported as a (binary) PLY file compatible with Gaussian Splatting viewers. Large datasets can be compressed using strategies such as SOG (Morgenstern et al., 2025) to reduce storage requirements while preserving visual quality.

Figure 2. Example of combination pipeline with a heritage ornament (medallion) rich of green and orange glassy and shiny gems (a). Recovered camera poses (96 views) and sparse point cloud (b). Photogrammetric dense cloud with some holes on the reflective gems (c). Isolated gems where 3DGS should be applied (d). Final fused 3D result (e).

[1] https://www.jawset.com/

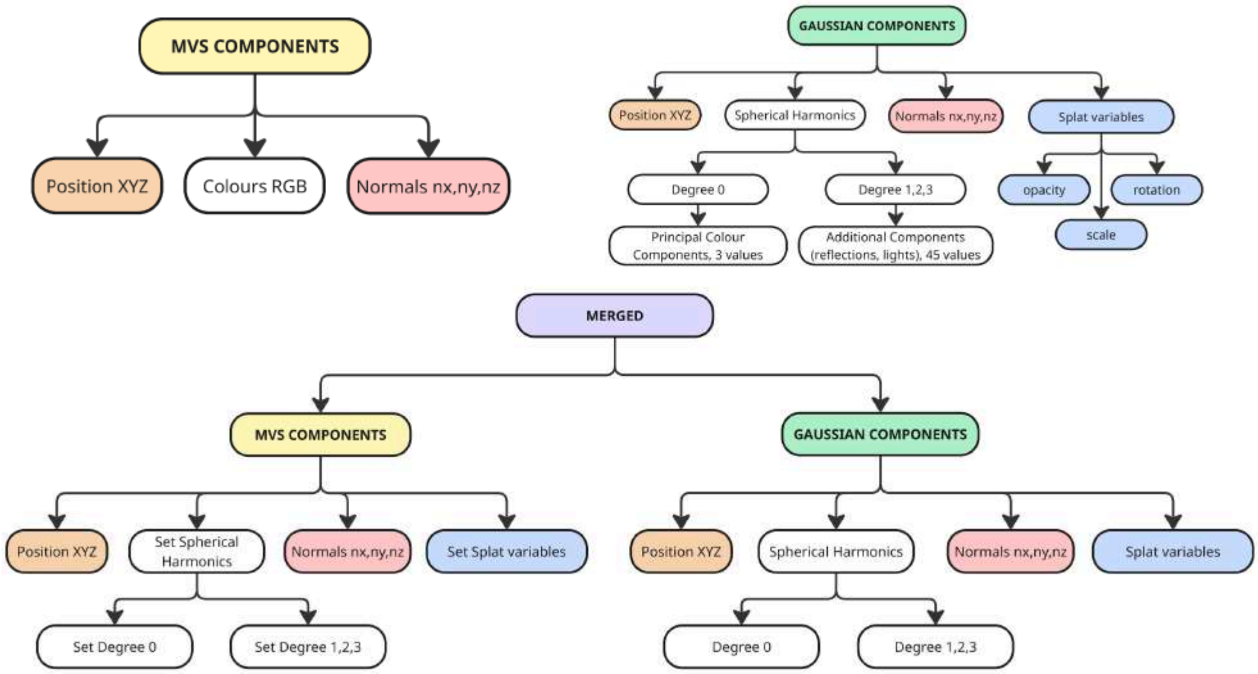

Figure 3. Photogrammetric (MVS) dense point cloud and GS components (top) and merged result with its characteristic (bottom).

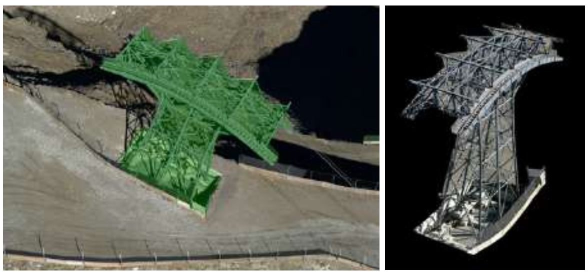

Figure 4. Examples of a MMLM result and derived mask for the lift pylon to constraint the 3DGS application.

Figure 5. Example of point clouds classification with Random Forest for separating 3DGS and dense photogrammetric components for fusion purposes.

4. Datasets and results

The proposed integration methodology is evaluated on different-scale datasets, and in particular:

1) Aerial-scale: two UAV datasets (Figure 6) provided by AVT Airborne Sensing[2] are used:

- Lift pylon: it consists of 67 nadir and oblique images acquired with a Share 203S Pro multi-view camera and with 1.3 cm average GSD; the scene contains a cable-car pylon with metallic shining elements and a well-textured surrounding environment.
- Electricity switchgear area: it is surveyed with a DJI ZenMuse P1 camera and consists of 212 nadir and oblique images acquired on two different flying heights and with 1.5 cm average GSD; the electricity structures are thin and shining, posing problems to conventional 3D reconstruction methods. For this area, a point cloud acquired with a Riegl miniVUX3-UAV is available and used as ground truth.

2) Artefact-scale: a medallion image block, composed of 96 images with a 2 mm GSD (Figure 2b), rich of glassy and shiny gems.

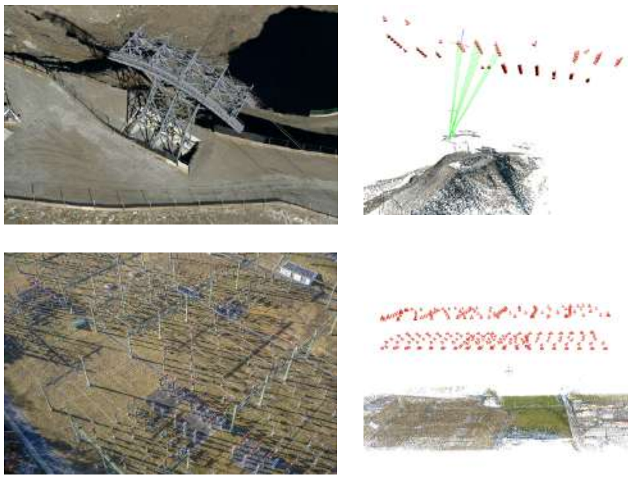

Figure 6. The lift pylon (top) and the electricity switchgear area (bottom) with the recovery camera network and sparse point cloud for the UAV blocks.

The image orientation results for the two aerial blocks are reported in Figure 6, while Figure 7 shows the reconstruction and integration visuals. These results demonstrate that the structure and material of the lift pylon and the electricity switchgear area hampered the generation of a complete dense point cloud with a conventional MVS method, while the GS properly rendered all metallic elements. At the same time, the gems of the medallion could not be entirely reconstructed with a MVS approach, whereas the GS visual results show completeness in the visualization.

[2] https://avt-as.eu/en/

Figure 7. Dense reconstruction with classical MVS (left), 3DGS results (middle) and merged photogrammetric and GS 3D data (right).

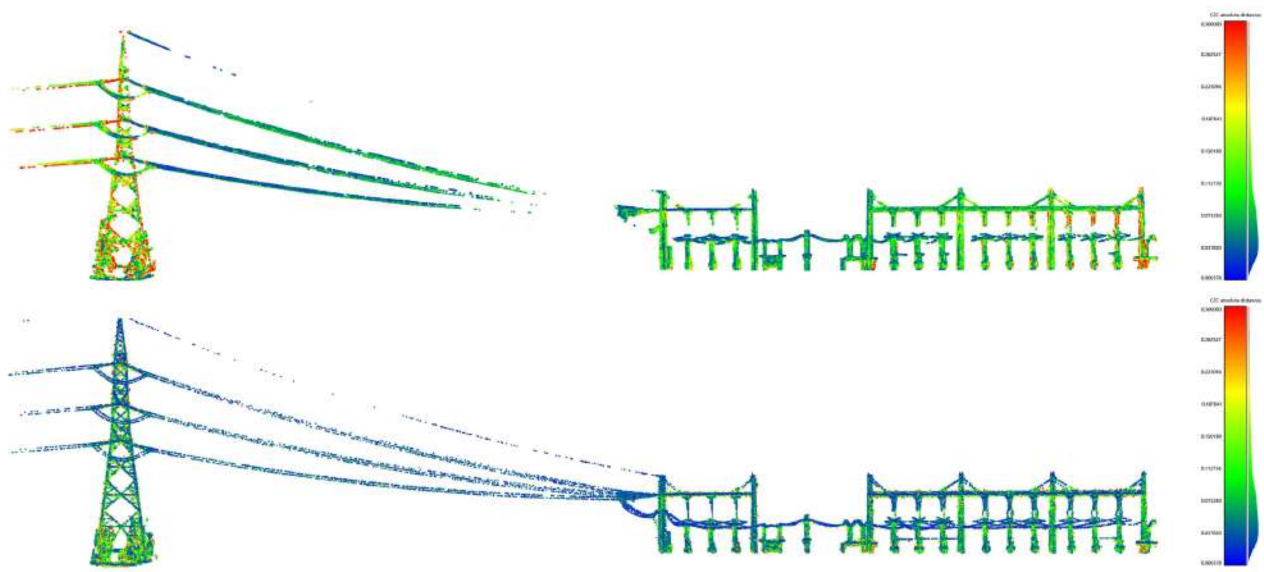

Figure 8. Visuals of the cloud-to-cloud distances against LiDAR ground truth for the electricity switchgear scene. Photogrammetry (top) and merged (bottom). Distance range [0-0.3m].

Dataset	Mean	St. Dev.
Photogrammetry	0.094	0.065
Merged	0.057	0.044

Table 2. Metric assessment (LiDAR ground truth) for the electricity switchgear area [m]. Cloud-to-cloud distance using the LiDAR dataset as ground truth.

A metric assessment and a comparison of the photogrammetric dense reconstruction and the proposed combination for the electricity switchgear area is presented in Table 2 and Figure 8.
The improved visual appearance for challenging regions reconstructed with GS methods and used in the combined datasets is also assessed using pixel-based and perceptual texture metrics. Some of the original input images of the three datasets are compared with rendered views of GS models and MVS reconstructions. Tables 3 and 4 demonstrate that GS representation outperforms the photogrammetric datasets in all metrics in the case of non-collaborative surfaces.

	MSE↓	PSNR (dB)↑	SSIM↑	MS-SSIM↑	LPIPS↓	ΔE↓
Pylon	0.011	19.55	0.872	0.888	0.120	3.142
Switchgear	0.001	29.98	0.981	0.988	0.024	1.045
Gems	0.0002	35.39	0.986	0.987	0.013	0.468

Table 3. Photogrammetric datasets - rendered texture assessment.

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-1363-2026

	MSE↓	PSNR (dB)↑	SSIM↑	MS-SSIM↑	LPIPS↓	ΔE↓
Pylon	0.006	21.59	0.896	0.914	0.081	2.754
Switchgear	0.0006	31.73	0.988	0.991	0.021	0.902
Gems	0.0001	38.31	0.991	0.991	0.006	0.257

Table 4. Merged datasets - rendered texture assessment.

Finally, cross-sections on the MVS and GS results for the medallion are shown in Figure 9. The increased completeness on the gem's surfaces is evident, but also the higher noise generated by GS in the combination.

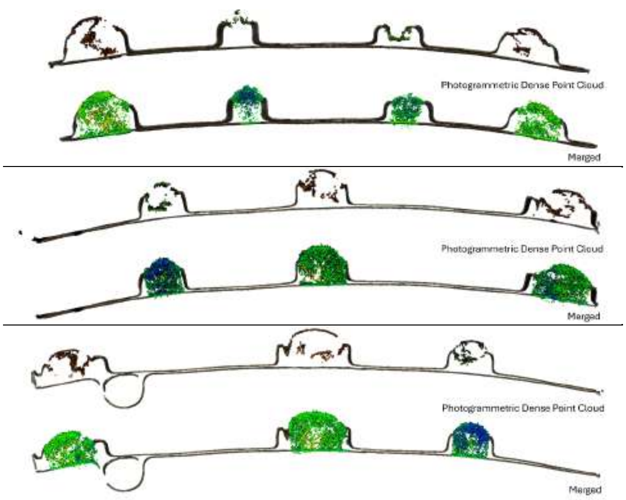

Figure 9. Visual comparison: cross-sections on the medallion dataset.

5. Online visualization for metric purposes

3DGS has revolutionized real-time rendering with its state-of-the-art novel view synthesis, but its utility for accurate geometric measurement remains under-utilized. Current point measurement methods still rely on demanding stereoscopic workstations or direct picking on often-incomplete and inaccurate 3D point clouds or meshes. As 3DGS renders exact source views and smoothly interpolates in-between views, a user can intuitively pick congruent points across different views while operating 3DGS models. By triangulating these congruent points, one can precisely generate 3D point measurements. This approach mimics traditional stereoscopic measurement but is significantly less demanding: it requires neither a stereo workstation nor specialized operator stereoscopic capability. Furthermore, it enables multi-view intersection (more than two views) for higher measurement accuracy. This idea, presented in Deng and Qin (2026), is implemented into a web-based application. Interacting with a 3DGS model, users can manually identify points of interest across multiple real-time rendered views. The tool then triangulates these picked points based on the rendered view poses and provides 3D coordinates. This approach leverages the superior visual quality of GS rendering to offer a robust alternative to traditional stereo-measurement. Notably, supporting multi-ray triangulation, it generates an error ellipsoid for each measured point to quantify precision. Figure 10 shows the lift pylon scene visualized in the online tool.

6. Conclusions

The paper presents an integration methodology to merge photogrammetric MVS dense point clouds with 3DGS results into a unique, measurable 3D asset. The proposed methodology wants to exploit multi-view consistency and metrically accurate triangulation results on well-textured areas together with detailed radiance modelling and dense scene representation on regions with less collaborative surfaces or tiny structures.
The evaluation on different-scale datasets, featuring challenging objects and materials for conventional photogrammetric reconstructions, proved that the proposed fusion approach improves completeness and visual fidelity. Moreover, GS-rendered views enable interactive 3D measurements in regions where dense point clouds are incomplete, increasing the overall quality of the surveying product.
Despite the promising results, some limitations of the current implementation should be highlighted. 3DGS primarily improves visual fidelity, but further investigations are needed to ensure a sufficient metric accuracy in regions without dense photogrammetric coverage. Converting photogrammetric points to 3DGS format and semi-automatic masking can introduce minor artifacts in the fused models. This is mainly due to the needed approximations in assigning to the MVS cloud the GS properties that are not encoded in the original data, but that are required for compatibility with GS rendering. Finally, the method lacks object-level understanding or geometric constraints to regularize surfaces, which may affect consistency in representing complex and heterogeneous scenes.
As future work, we will investigate a semantic-based integration of the two methods, using class-aware priors as well as a hybrid integration based on geometric-aware consistency (from depth maps) and surface regularization.

References

Agugiaro, G., González, F. G. G., Cavallo, R., 2020. The city of tomorrow from…the data of today. *ISPRS International Journal of Geo-Information*, 9(9), 554.

Akhavi Zadegan, A., Vivet, D., Hadachi, A., 2025. Challenges and advancements in image-based 3D reconstruction of largescale urban environments: a review of deep learning and classical methods. *Frontiers in Computer Science*, 7, 1467103.

Carion, N., Gustafson, L., Hu, Y.T., Debnath, S., Hu, R., Suris, D., Ryali, C., Alwala, K.V., Khedr, H., Huang, A. and Lei, J., 2025. Sam 3: Segment Anything with concepts. *Proc. ICLR*.

Chen, W., Li, Z., Guo, J., Zheng, C., & Tian, S., 2025. Trends and techniques in 3D reconstruction and rendering: A survey with emphasis on Gaussian Splatting. *Sensors*, *25*(12), 3626.

Dalal, A., Hagen, D., Robbersmyr, K.G., Knausgård, K.M., 2024. Gaussian Splatting: 3D reconstruction and novel view synthesis: a review. *IEEE Access*, 12, pp. 96797-96820.

Deng, D., Qin, R., 2026. Accurate point measurement in 3DGS – A new alternative to traditional stereoscopic-view based measurements. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* ISPRS Congress 2026, in press.

Fan, Z., Wang, K., Wen, K., Zhu, Z., Xu, D., & Wang, Z., 2024. LightGaussian: unbounded 3D gaussian compression with 15x reduction and 200+ fps. *Advances in neural information processing systems*, 37, 140138-140158.

Farella, E.M., Morelli, L., Rigon, S., Grilli, E., Remondino, F., 2022. Analysing key steps of the photogrammetric pipeline for museum artefacts 3D digitisation. *Sustainability*, 14, 5740
.

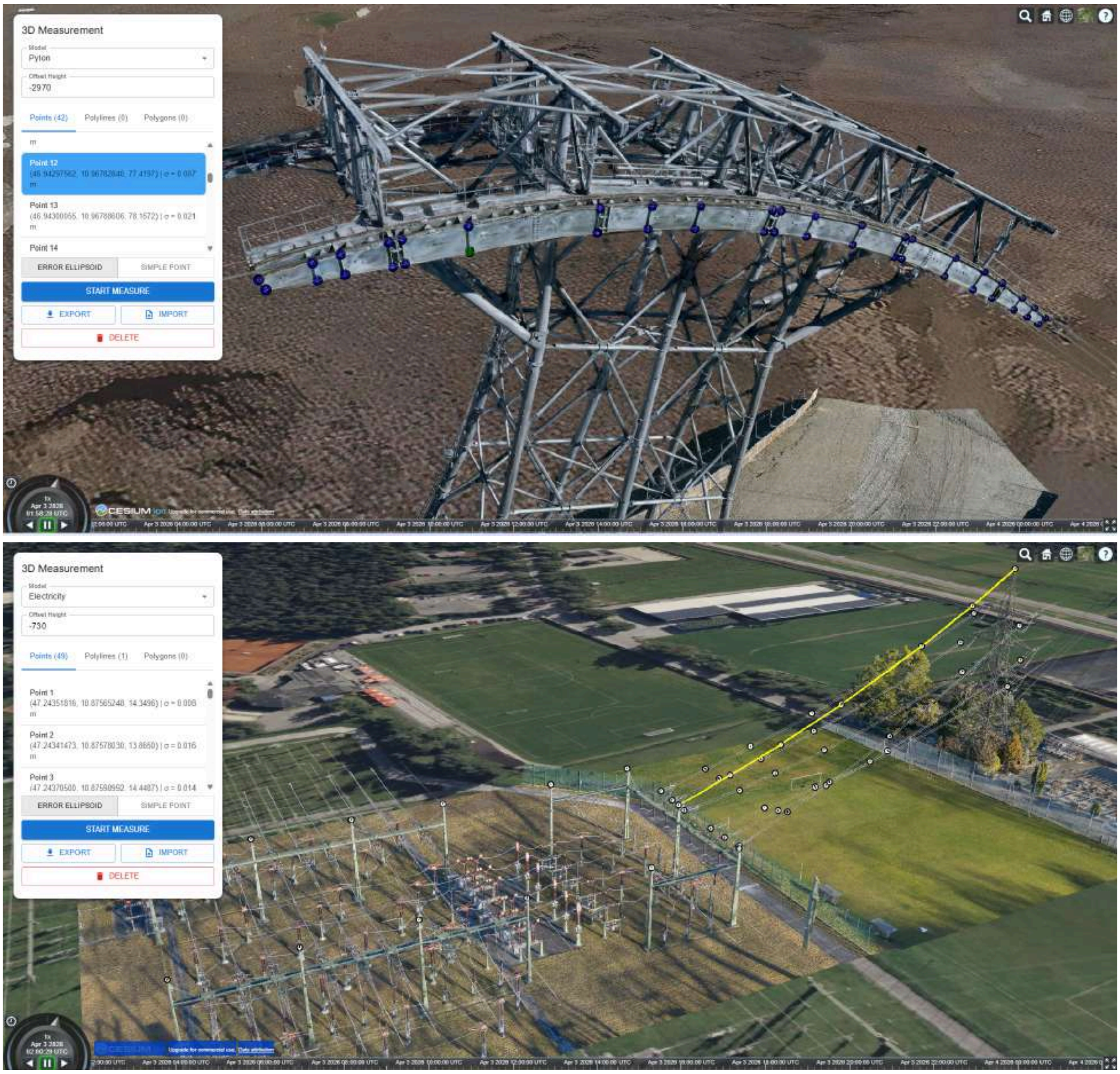

Figure 10. Point measurements on the merged 3D data for the lift pylon (top) and electricity switchgear (bottom) scenes using multi-ray triangulation in the OSU 3DGS measurement tool[3]. Error blue ellipsoids and measured features are enlarged for visualization purposes.

Fei, B., Xu, J., Zhang, R., Zhou, Q., Yang, W., & He, Y., 2024. 3D gaussian splatting as new era: A survey. *IEEE Transactions on Visualization and Computer Graphics*.

Gao, R., Qi, Y., 2024. A Brief review on differentiable rendering: recent advances and challenges. *Electronics*, 13(17), 3546.

Guédon, A., Gomez, D., Maruani, N., Gong, B., Drettakis, G., Ovsjanikov, M., 2025. MILo: Mesh-In-the-Loop Gaussian Splatting for detailed and efficient surface reconstruction. *ACM TOG*, Vol. 44(6), pp. 1-15.

Hahlbohm, F., Franke, L., Eisemann, M., Magnor, M., 2026. Faster-GS: analyzing and improving gaussian splatting optimization. Proc. *CVPR*.

Held, J., Vandeghen, R., Deliege, A., Hamdi, A., Cioppa, A., Giancola, S., Vedaldi, A., Ghanem, B., Tagliasacchi, A., Van Droogenbroeck, M., 2025a. Triangle Splatting for real-time radiance field rendering. Proc. *3DV*.

Held, J., Vandeghen, R., Hamdi, A., Deliege, A., Cioppa, A., Giancola, S., Vedaldi, A., Ghanem, B., Van Droogenbroeck, M., 2025b. 3D Convex Splatting: radiance field rendering with 3D smooth convexes. Proc. *CVPR*.

Held, J., Son, S., Vandeghen, R., Rebain, D., Gadelha, M., Zhou, Y., Cioppa, A., Lin, M., Van Droogenbroeck, M., Tagliasacchi, A., 2026. MeshSplatting: differentiable rendering with opaque meshes. Proc. *CVPR*.

[3] https://gdaosu.github.io/3dgs_measurement_tool/

Huang, B., Yu, Z., Chen, A., Geiger, A., & Gao, S., 2024. 2D gaussian splatting for geometrically accurate radiance fields. Proc. *ACM SIGGRAPH*, pp. 1-11.

Jeddoub, I, Nys, G.A, Hajji, R., Billen, R., 2023. Digital Twins for cities: Analyzing the gap between concepts and current implementations with a specific focus on data integration. *Int. J. of Applied Earth Observation and Geoinformation. Vol. 122*.

Ji, B., Yao, A., 2025. SfM-Free 3D Gaussian Splatting via hierarchical training. Proc. *CVPR*, 21654-21663.

Jiang, Y., Tu, J., Liu, Y., Gao, X., Long, X., Wang, W., & Ma, Y., 2024. Gaussianshader: 3D gaussian splatting with shading functions for reflective surfaces. Proc. *CVPR*, pp. 5322-5332.

Kerbl, B., Kopanas, G., Leimkuhler, T., Drettakis, G., 2023. 3D Gaussian Splatting for real-time radiance field rendering. Proc. *SIGGRAPH*, 42(4).

Lee, J. C., Rho, D., Sun, X., Ko, J. H., Park, E., 2024. Compact 3D gaussian representation for radiance field. Proc. *CVPR*, pp. 21719-21728.

Liao, Y., Di, Y., Zhou, H., Zhu, K., Lu, M., Duan, Q., Liu, J., 2026. A Survey on Neural Radiance Fields. *ACM Computing Surveys*, Vol. 58(2), No. 41, pp. 1 - 33.

Lin, J., Li, Z., Tang, X., Liu, J., Liu, S., Liu, J., Lu, Y., Wu, X., Xu, S., Yan, Y. et al., 2024. Vastgaussian: Vast 3D gaussians for large scene reconstruction. Proc. *CVPR*, 5166–5175.

Liu, R., Sun, D., Chen, M., Wang, Y., Feng, A., 2025. Deformable Beta Splatting. Proc. *SIGGRAPH*, pp. 1-11.

Liu, Y., Luo, C., Fan, L., Wang, N., Peng, J., Zhang, Z., 2024. Citygaussian: Real-time high-quality large-scale scene rendering with gaussians. Proc. *ECCV*, 265-282.

Liu, J., Gao, J., Ji, S., Zeng, C., Zhang, S., Gong, J., 2023. Deep learning based multi-view stereo matching and 3D scene reconstruction from oblique aerial images. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 204, pp. 42-60.

Lu, T., Yu, M., Xu, L., Xiangli, Y., Wang, L., Lin, D., Dai, B., 2024. Scaffold-GS: structured 3D gaussians for view-adaptive rendering. Proc. *CVPR*, pp. 20654-20664.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., and Ng, R., 2020. NeRF: representing scenes as neural radiance fields for view synthesis. *Commun. ACM* 65..

Morgenstern, W., Barthel, F., Hilsmann, A., Eisert, P., 2025. Compact 3D scene representation via self-organizing gaussian grids. Proc. *ECCV*, pp. 18-34.

Pan, W., Zhang, X., Zhai, H., Xiang, X., Jiang, H., Zhang, G., 2025. Liberated-GS: 3D Gaussian Splatting independent from SfM point clouds. Proc. *CVPR*, pp. 26675-26685.

Peng, C., Tang, Y., Zhou, Y., Wang, N., Liu, X., Li, D., & Chellappa, R., 2024. BAGS: blur agnostic gaussian splatting through multi-scale kernel modeling. Proc. *ECCV*, pp. 293-310.

Ravi, N., Gauber, V., Hu, Y.-T., Hu, R., Ryali, C, …, Feichtenhofer, C., 2024. SAM 2: Segment Anything in images and videos. *ArXiv:2408.00714.*

Remondino, F., Karami, A., Yan, Z., Mazzacca, G., Rigon, S., Qin, R., 2023. A critical analysis of NeRF-Based 3D reconstruction. *Remote Sensing*, 15, 3585.

Son, S., Gadelha, M., Zhou, Y., Fisher, M., Xu, Z., Qiao, Y.-L., Lin, M.C., Zhou, Y., 2025. DMesh++: An Efficient Differentiable Mesh for Complex Shapes. Proc. *ICCV*, pp. 26590-26599.

Stathopoulou, E.K., Remondino, F., 2023. A survey of conventional and learning-based methods for multi-view stereo. *The Photogrammetric Record*, Vol. 38(183), pp. 374-407.

Steiner, M., Köhler, T., Radl, L., Windisch, F., Schmalstieg, D., Steinberger, M., 2025. AAA-Gaussians: anti-aliased and artifact-free 3D gaussian rendering. Proc. *ICCV*, pp. 27650-27659.

Tewari, A., Thies, J., Mildenhall, B., Srinivasan, P., Tretschk, E., Yifan, W., Lassner, C., …, Golyanik, V., 2022. Advances in neural rendering. *Computer Graphics Forum, SIGGRAPH courses*, Wiley Online Library.

Wang, J., Chen, M., Karaev, N., Vedaldi, A., Rupprecht, C., Novotny, D., 2025. VGGT: Visual geometry grounded transformer. Proc. *CVPR*, 5294-5306.

Wang, F., Zhu, Q., Chang, D., Gao, Q., Han, J., Zhang, T., Hartley, R., Pollefeys, M., 2026. Learning: A survey. *TPAMI.*

Xie, T., Zong, Z., Qiu, y., Li, X., Feng, Y., Yang, Y., Jiang, C., 2024. PhysGaussian: Physics-Integrated 3D Gaussians for generative dynamics. Proc. *CVPR*.

Xie, H., Chen, Z., Hong, F., Liu, Z., 2025. Generative Gaussian Splatting for unbounded 3D city generation. Proc. *CVPR*.

Xu, J., Mei, Y., Patel, V., 2024. Wild-GS: Real-time novel view synthesis from unconstrained photo collections. Proc. *NIPS*, 37.

Yan, Z., Li, L., Shao, Y., Chen, S., Wu, Z., Hwang, J.-N., Zhao, H., Remondino, F., 2026. 3dsceneeditor: Controllable 3D scene editing with Gaussian Splatting. Proc. *IEEE/CVF WACV*, pp. 1852-1863.

Yan, Z., Dong, W., Shao, Y., Lu, Y., Haiyang, L., Liu, J., Wang, H., Wang, Z., Wang, Y., Remondino, F., Ma, Y., 2025a. Renderworld: World model with self-supervised 3D label. Proc. *IEEE ICRA*.

Yan, Z., Yin, M., Shao, Y., and Remondino, F., 2025b. Evaluating 3D Gaussian Splatting for urban scene reconstruction. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-1/W6-2025, 251-258.

Yang, J., Sax, A., Liang, K. J., Henaff, M., Tang, H., Cao, A., Chai, J., Meier, F., Feiszli, M., 2025. Fast3r: Towards 3D reconstruction of 1000+ images in one forward pass. Proc. *CVPR*, 21924-21935

Yu, Z., Sattler, T., Geiger, A., 2024. Gaussian opacity fields: Efficient adaptive surface reconstruction in unbounded scenes. *ACM TOG*, 43(6), pp. 1-13.

Yuan, H., Li, X., Zhang, T., Huang, Z., Xu, S., Ji, S., Tong, Y., Qi, L., Feng, J., Yang, M.-H., 2025. Sa2va: Marrying SAM2 with llava for dense grounded understanding of images and videos. *arXiv preprint arXiv:2501.04001.*