الجمهورية الجزائرية الديمقراطية الشعبية

People's Democratic Republic of Algeria

وزارة التعليم العالي و البحث العلمي

Ministry of Higher Education and Scientific Research

جامعة عين تموشنت بلحاج بوشعيب

University of Ain Temouchent-Belhadj Bouchaib

Faculty of Science and Technology

Department of Civil Engineering and public works

Thesis for Obtaining the Master`s degree

Sector: CIVIL ENGINEERING

Specialty: STRUCTURE

Theme:

PHOTOGRAMMETRY: EVOLUTION, CURRENT STATE, AND FUTURE TRAJECTORIES WITH A FOCUS ON BIM INTEGRATION

PRESENTED BY:

- BEN AISSA RAHAL NOUR EL HOUDA MOKHTARIA
- ZENASNI FATIMA ZOHRA

SUSTAINED ON:24/06/2026

BOARD OF EXAMINERS:

PRESIDENT: Mr EL HADJ MIMOUNE Arezki

EXAMINER: Mr HOUARI Abdelhadi

SUPERVISOR: Mr KADDOUR Hakim

SESSION YEAR :2025/2026

PHOTOGRAMMETRY: EVOLUTION, CURRENT STATE, AND FUTURE TRAJECTORIES WITH A FOCUS ON BIM INTEGRATION

الاهداء:

إلى اليد الخفية والقلب الحنون وصاحبة الدعاء الصادق... أمي الغالية

التي نسجت من دعواتها سُلّماً صعدتُ به نحو المعالي، وأنارت بحنانها دروبي حين ادلهمّت الخطوب، لكِ يا مَن حملتِني وهناً على وهن، أُهدي ثمرة غرسك وبركة دعائك.

وإلى القامة الشامخة والسند المتين... أبي العزيز

يا من علّمني أن العلم رفعة والاجتهاد شرف، وغرستَ في نفسي حب المعرفة منذ نعومة أظافري، لك يا سندي وملاذي، أُهدي هذا الجهد المتواضع.

وإلى رفقاء الدرب وشركاء الروح... إخوتي الأحبة

الذين شاركوني الأفراح والأتراح، وكانوا لي عوناً ومدداً، أُهدي لكم هذا العمل شاهداً على المحبة التي تجمعنا.

وإلى المنارة التي احتضنت أحلامي... جامعتي الغالية

التي فتحت لي أبواب المعرفة، وصقلت موهبتي، وأنارت عقلي بنور العلم، لك مني كل التقدير والامتنان.

وإلى كل طالب شغوف يسعى وراء العلم

باذلاً الجهد والوقت، ساهراً الليالي في طلب المعرفة، مؤمناً بأن العلم نور والجهل ظلام، لكم جميعاً أُهدي هذا العمل، راجياً أن يكون لبنة في صرح العلم الشامخ، ونبراساً يُضيء لمن يأتي بعدنا.

إلى كل هؤلاء أُهدي ثمرة جهدي، راجياً من الله العلي القدير أن يجعله خالصاً لوجهه الكريم، نافعاً للعلم والمعرفة

بن عيسى رحال نور الهدى مختارية

الاهداء:

بسم الله الرحمن الرحيم

إلى من علّماني أن الصبر مفتاح النجاح، وأن العزيمة لا تُقهر...

إلى من سهرا الليالي ليمنحاني فجراً أكثر إشراقاً...

إلى ينبوع العطاء الذي لا ينضب، ومدرسة الحياة التي لا تُنسى...

إلى أبي وأمي الغاليين،

وإلى رفيقَي الدرب اللذَين شاركاني الأحلام وآمنا بي حين شككت في نفسي...

الى من كانت ابتسامتاهما سنداً ودفئاً في أصعب اللحظات...

إلى أخوَيّ العزيزين،

أُهدي ثمرة هذا الجهد المتواضع، راجياً من الله أن يجعله في ميزان حسناتهم جميعاً، وأن يكون وفاءً ببعض ما أعطوني.

"رَبِّ ارْحَمْهُمَا كَمَا رَبَّيَانِي صَغِيراً"

زناسني فاطمة الزهراء

شكر وتقدير:

الحمد لله رب العالمين، والصلاة والسلام على أشرف المرسلين، نبينا محمد وعلى آله وصحبه أجمعين.

نتوجه بالشكر أولاً وقبل كل شيء إلى الله عز وجل الذي وفقنا وأعاننا وأمدنا بالصبر والعزيمة على إتمام هدا العمل المتواضع، فله الحمد والشكر على نعمه التي لا تُحصى.

ونخص بالذكر والتقدير الوالدين الكريمين اللذان لم يدّخرا جهداً في دعمنا مادياً ومعنوياً طوال مشوارنا الدراسي، فكانا سندا ومصدر إلهام في كل خطوة خطوناها، أسأل الله أن يحفظهم ويطيل في عمرهم.

و نتقدم بجزيل الشكر والعرفان إلى أستاذنا المشرف الفاضل الدكتور قدور حكيم الذي تفضّل بالإشراف على هذه المذكرة، ولم يبخل علينا بتوجيهاته القيّمة ونصائحه السديدة طوال مسيرة البحث، فكان نِعمَ الموجِّه والمرشد.

كما نتوجه بالشكر الجزيل إلى أعضاء لجنة المناقشة الموقرة على تفضّلهم بقبول مناقشة هذا العمل وتقييمه والمتمثلة في رئيس اللجنة الاستاذ الحاج ميمون ارزقي و المناقش الاستاذ هواري عبد الهادي

وفي الختام، لكل من أسهم من قريب أو من بعيد في إنجاز هذا العمل، أُهدي عبارات الشكر والامتنان.

RESUME:

Le BIM (Building Information Modeling) est considéré comme l'un des systèmes les plus avancés actuellement disponibles pour la gestion et le suivi des projets d'ingénierie à travers toutes leurs phases et typologies. Ce système tire sa puissance des technologies les plus récentes d'acquisition et de traitement des données de terrain, telles que le LiDAR et la XR. Cependant, la plupart de ces technologies restent coûteuses, en particulier pour les projets de petite envergure. C'est dans ce contexte que la photogrammétrie s'impose comme une alternative relativement économique.

Ce mémoire vise à clarifier cette technologie et à évaluer sa capacité à soutenir les flux de travail BIM, en examinant ses différents types, ses étapes d'évolution et les modèles mathématiques sur lesquels elle repose, tout en identifiant les avantages propres à chaque approche.

À cette fin, des images de terrain ont été traitées selon trois méthodes photogrammétriques distinctes SfM, NeRF et Gaussian Splatting à l'aide de différents logiciels et de systèmes informatiques aux caractéristiques variées.

Les résultats ont révélé l'absence de différence notable dans la phase d'acquisition des images sur le terrain. En revanche, les pipelines de traitement des données et les sorties obtenues se sont avérés très différents. Tandis que le SfM s'impose comme la méthode fondatrice de toutes les approches photogrammétriques, offrant la plus grande précision géométrique dans les modèles 3D, le Gaussian Splatting se distingue par le réalisme visuel de ses sorties, au prix toutefois de fichiers plus volumineux. Quant au NeRF, il produit des modèles générés par intelligence artificielle à l'apparence visuelle convaincante, mais dépourvus de structure géométrique réelle, ce qui le rend inadapté aux mesures d'ingénierie précises.

Mots-clés : BIM , SfM, NeRF, Gaussian Splatting.

ملخص:

يُعدّ نظام BIM من أحدث الأنظمة المتطورة المعتمدة حالياً في إدارة ومتابعة المشاريع الهندسية عبر جميع مراحلها وأنماطها المختلفة. ويستمد هذا النظام قوته من أحدث تقنيات استقاء المعلومات الميدانية ومعالجتها، كتقنيتَي LiDARوXR، غير أن معظم هذه التقنيات تظل مرتفعة التكلفة، لا سيما بالنسبة للمشاريع الصغيرة. وهنا تبرز الفوتوغرامترية بوصفها تقنية بديلة ذات تكلفة معقولة نسبياً.

تسعى هذه المذكرة إلى توضيح هذه التقنية وتقييم مدى قدرتها على دعم منظومةBIM ، من خلال استعراض أنواعها ومراحل تطورها والنماذج الرياضية التي تقوم عليها، واستخلاص مميزات كل منها.

ولتحقيق هذا الهدف، جرت معالجة صور ميدانية بأساليب فوتوغرامترية متعددة، هي SfM و NeRFو Gaussian Splatting، باستخدام برامج مختلفة وأجهزة حاسوبية متفاوتة المواصفات.

وقد كشفت النتائج عن تماثل في أسلوب التقاط الصور الميدانية، بينما تباينت تبايناً واضحاً طرق معالجة البيانات والمخرجات الناتجة. فبينما تتصدر تقنية SfM المشهدَ بوصفها الركيزة الأساسية لجميع الأساليب الفوتوغرامترية وبدقتها القياسية العالية في النماذج ثلاثية الأبعاد، تتميز تقنية Gaussian Splatting بواقعية مخرجاتها البصرية وإن كانت ملفاتها أكبر حجماً. أما NeRF فتنتج نماذج مدعومة بالذكاء الاصطناعي تتسم بالمظهر البصري الجيد، إلا أنها تفتقر إلى البنية الهندسية الدقيقة مما يجعلها غير صالحة للقياسات الهندسية.

الكلمات الرئيسية : BIM , SfM, NeRF, Gaussian Splatting.

ABSTRACT:

BIM (Building Information Modeling) is considered one of the most advanced systems currently available for managing and monitoring engineering projects across all their phases and typologies. The system derives its strength from cutting-edge field data acquisition and processing technologies, such as LiDAR and XR. However, most of these technologies remain prohibitively expensive, particularly for small-scale projects. This is where photogrammetry emerges as a relatively cost-effective alternative.

This thesis aims to clarify this technology and assess its potential to support BIM workflows, by examining its various types, stages of development, and underlying mathematical models, while identifying the advantages of each approach.

To this end, field imagery was processed using three different photogrammetric methods SfM, NeRF, and Gaussian Splatting employing various software packages and computer systems with differing specifications.

The results revealed no significant difference in the field image acquisition process; however, the data processing pipelines and resulting outputs varied considerably. While SfM leads as the foundational method for all photogrammetric approaches, offering the highest geometric accuracy in 3D models, Gaussian Splatting excels in the visual realism of its outputs, albeit at the cost of larger file sizes. NeRF ،on the other hand, produces AI-generated models with impressive visual appearance, but lacks true geometric structure, rendering it unsuitable for precise engineering measurements.

Keywords: BIM , SfM, NeRF, Gaussian Splatting.

SUMMARY:

Chapter II: Classical Photogrammetry and Structure from Motion (SfM)

Chapter III : Neural Radiance Fields (NeRF) in Photogrammetry

Chapter V : Comparative Analysis and Future Perspectives

LIST OF FIGURES:

LIST OF TABLES:

GENERAL INTRODUCTION:

The civil engineering and project management sector has undergone a profound transformation in recent decades, driven by the rapid evolution of digital technologies that have reshaped fieldwork and design methodologies alike. At the forefront of these developments stands Building Information Modeling (BIM), which has redefined the full lifecycle of engineering projects from initial design through to operation and maintenance relying on cutting-edge sensing and field digitization technologies. Scientific literature has well-documented BIM's capacity to enhance coordination among stakeholders and reduce operational costs. (Abdulghani & Gullu, 2021) Among the most prominent supporting technologies are LiDAR both terrestrial and aerial enabling the acquisition of high-density three-dimensional point clouds. (Blaszczak-Bak, Masiero, Bąk, & Kuderko, 2024), and Extended Reality (XR) technologies, which enhance visualization and interaction with digital models directly on construction sites. (Bae, Golparvar-Fard, & White, 2013) However, these advanced technologies remain financially prohibitive, particularly for small and medium-scale projects.

In this context, photogrammetry has emerged as a promising and cost-effective alternative. The technique has evolved considerably over recent decades, propelled by rapid advances in artificial intelligence and computer vision. (Barazzetti, 2016) Leading modern photogrammetric approaches is Structure from Motion (SfM), which forms the foundational backbone of most three-dimensional applications owing to its high geometric accuracy and compatibility with standard BIM formats. (Albeaino, Eiris, & Gheisari, 2024) Alongside it, two AI-driven emerging techniques have attracted significant research interest: Neural Radiance Fields (NeRF), introduced by Mildenhall et al. in 2020 (Mildenhall, et al., 2020), which synthesizes highly realistic visual models through neural networks; and Gaussian Splatting, introduced by Kerbl et al. in 2023, distinguished by its exceptional real-time rendering speed and visual fidelity. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

This thesis aims to study these three photogrammetric techniques and evaluate their capacity to support BIM workflows, through an analysis of their mathematical foundations, a comparative assessment of experimental results obtained on the campus of the University of Ain Temouchent, and a determination of the limits of their applicability within the standard architectural modeling environment.

The thesis is structured as follows:

Chapter 1: establishes the theoretical framework of BIM, covering its core concepts, multidimensional nature (3D to 10D), and supporting technologies including LiDAR, XR, and IoT, before examining Scan-to-BIM and Image-to-BIM approaches used in as-built documentation.

Chapter 2: addresses classical photogrammetry and Structure from Motion (SfM), tracing its historical development from the analogue to the digital era, with a detailed treatment of the mathematical foundations underlying feature detection, camera calibration, and point cloud generation, supported by a field experiment using 3DF Zephyr software.

Chapter 3: is dedicated to Neural Radiance Fields (NeRF), examining the conceptual shift from geometry-based to neural scene representation, detailing the neural network architecture and training process, and assessing the feasibility of integrating NeRF outputs into BIM environments.

Chapter 4: examines Gaussian Splatting, from its theoretical principles including 3D Gaussian representation and the differentiable splatting rendering pipeline to the results of the field experiment conducted using Postshot software, alongside an evaluation of BIM integration prospects.

Chapter 5: concludes with a comprehensive comparative analysis of the three techniques according to criteria of geometric accuracy, visual quality, processing speed, and interoperability, and charts future directions for these technologies within the BIM ecosystem.

Chapter I : BIM Fundamentals and Related Digital Technologies

I.1. Building Information Modeling (BIM) Core Concepts:

I.1.1. Definition of BIM:

Building Information Modeling (BIM) is a digital process involving generation and management of digital representations of physical and functional characteristics of places, i.e. buildings, and facilities, throughout the life cycle from the initial design to demolition. (Franco, 2018)

Unlike CAD systems, BIM uses data rich clever 3D geometric models, which regard such objects as walls, doors, or HVAC components as containing information about materials, costs, and their functional relationships to other model objects. (Associates, 2025)

The parametric nature of this modeling automatically updates other objects; e.g. with changes in wall thickness the doors, floors and other structural calculations adjust accordingly. (Deltek, n.d.)

BIM integrates architects, engineers, contractors and owners within a Common Data Environment (CDE) to allow real-time access and alterations to minimize miscommunication among stakeholders. (gbc engineers, 2025)

In addition to 3D, BIM supports 4D, which adds time, and 5D, which adds cost and quantities. Clash detection identifies and limits conflicts such as ductwork crossing beams. (accruent, 2024)

While designing, BIM ensures prefabrication, just-in-time delivery, and tracking during the construction phase, which later shifts to facility management (6D BIM) for maintenance, energy assessment, and remodeling through "as-built" information. (Key benefits of BIM, 2022)

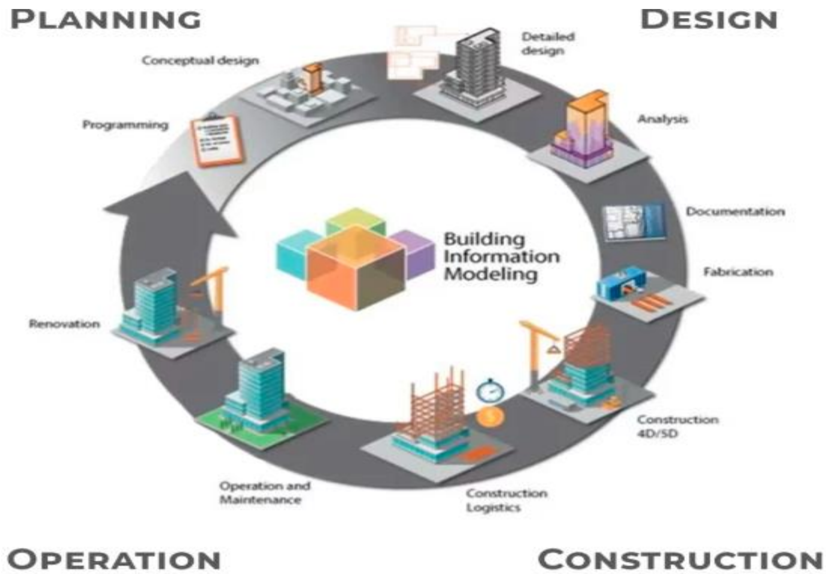

Figure 1: Illustration of BIM core concepts (Shobha, 2024)

I.1.2. Principles of BIM:

Principles of BIM direct the practical application of digital modeling in construction projects. It highlights data-rich modeling, parametric design, and life cycle management. (mclinestudios, 2024) Its principles are as follows:

- **Parametric Modeling**: Parameters and rules define elements, ensuring automatic change propagation to maintain consistency. (mclinestudios, 2024)
- **Data-Rich Models**: Incorporating information on materials, budgets, scheduling, and specifications besides geometry aids decision-making (mclinestudios, 2024)
- **Collaboration and Common Data Environment (CDE)**: All parties access a common data source for efficient coordination without creating silos.(Chubryk, 2025)
- **Alignment with Project Goals:** Modeling tailored for specific purposes such as coordination or visualization, taking needs into account.(Chubryk, 2025)
- **Lean and Reliable Data:** Utilize only essential data, ensuring accuracy, consistency, and validation to prevent errors. (Chubryk, 2025)

- **Interoperability and Lifecycle Focus**: Adopt interoperable standards like IFC for seamless data exchange across tools and throughout the project life cycle. (mclinestudios, 2024)

I.1.3. Importance of BIM:

The adoption of BIM technology transforms construction, enabling collaborative work between all stakeholders using a digital model, preventing misunderstandings from interpreting 2D drawings.

It allows real-time editing and clash detection, potentially cutting costs by 20~30%. This accelerates project completion and facilitates better decision-making in the design, construction, and operation phases. (Mulik, 2024)

BIM increases efficiency through data-rich models that automate quantity take-offs, scheduling (4D), and cost estimation (5D). Contractors receive insights from prefabrication planning and model-based progress monitoring.

Facility management becomes more straightforward with as-built information for owners' long-term benefits.

BIM promotes sustainability by simulating energy consumption, material effects, and life cycle performance at an early design stage. This encourages sustainable practices, resulting in lower operating expenses and carbon emissions.

Visualization techniques also enhance stakeholder involvement and project approval rates. (AIA Contract Documents, 2023)

I.1.4. BIM dimensions (3D, 4D, 5D, 6D, 7D, 8D, 9D, 10D) :

BIM dimensions extend the core 3D geometric model by integrating additional layers of information such as time, cost, sustainability, and management aspects. This progressive enrichment enables a holistic approach to project delivery, supporting the entire lifecycle from design to operation and industrialization. (BIM, 2025)

- **3D BIM*:*** 3D BIM represents the fundamental geometric dimension of a project, modeling the physical structure of a building through width, height, and depth. It includes detailed components such as walls, doors, and building systems, enabling accurate visualization and clash detection. Compared to traditional 2D drawings, it provides interactive representations that enhance design accuracy and interdisciplinary coordination. It also supports quantity takeoff (QTO) and high-quality renderings for client communication. (BIM, 2025)

- **4D BIM**: 4D BIM incorporates the time dimension into the 3D model by linking construction activities and sequences to scheduling data. This allows simulation of the construction process, facilitating early identification of logistical issues such as resource conflicts. As a result, it improves planning efficiency, reduces delays, and enhances site safety. (engineers, 2025)
- **5D BIM:**5D BIM integrates cost data with the 3D model and time (4D), linking building elements to quantities, unit prices, and work items. This transforms cost estimation into an automated, model-based process. It enables real-time quantity takeoff (QTO) and the generation of both preliminary and detailed cost estimates, which remain synchronized with design and schedule changes. The integration of historical productivity data and alternative construction methods further improves cost forecasting accuracy and supports reliable budget and schedule planning throughout the project lifecycle.(Asser Elsheikh, 2025)
- **6D BIM**: 6D BIM incorporates sustainability and environmental performance data into the model, extending beyond geometry, time, and cost. It includes indicators such as energy consumption, carbon emissions, material recyclability, and life cycle analysis. This enables simulation of alternative design scenarios and supports the selection of environmentally efficient solutions. Additionally, it is used during operation to monitor building performance, support maintenance decisions, and achieve sustainability objectives. (Montiel-Santiago, Hermoso-Orzáez, & Terrados-Cepeda, 2020)
- **7D BIM**: 7D BIM focuses on operation and maintenance by using the as-built model as a centralized information system throughout the asset's lifecycle. It integrates data on building systems, component conditions, maintenance history, and operational procedures. This facilitates efficient planning of inspections and maintenance activities, reduces information retrieval time, lowers lifecycle costs, and contributes to extending asset performance and service life. (Xinghua Gao & Bozorgi, 2019)
- **8D BIM**: 8D BIM is associated with health and safety management, integrating risk-related data into the model during design and planning stages. It includes information on hazardous zones, fall risks, access paths, and protective measures. This supports visual risk

assessment and proactive safety planning, improving communication and reducing accidents while ensuring compliance with safety standards. (CARLOS.J.PAMPLIEGA, 2019)

- **9D BIM**: 9D BIM combines BIM with Lean Construction principles to enhance process efficiency and minimize waste. BIM provides a reliable digital model, while Lean methodologies eliminate non-value-adding activities in workflows and resource use. This integration reduces delays, minimizes rework, optimizes material usage, and improves overall project coordination, resulting in more efficient and value-driven project delivery. (AL-Zubaidi & AlZaid, 2025)
- **10D BIM:** 10D BIM relates to industrialized and off-site construction, where the digital model is used to define prefabricated components and link them to manufacturing and supply chain processes. This supports automation of the design–manufacturing–assembly cycle, reduces on-site construction time, and improves overall quality and efficiency. (PIASECKIENĖ, 2022)

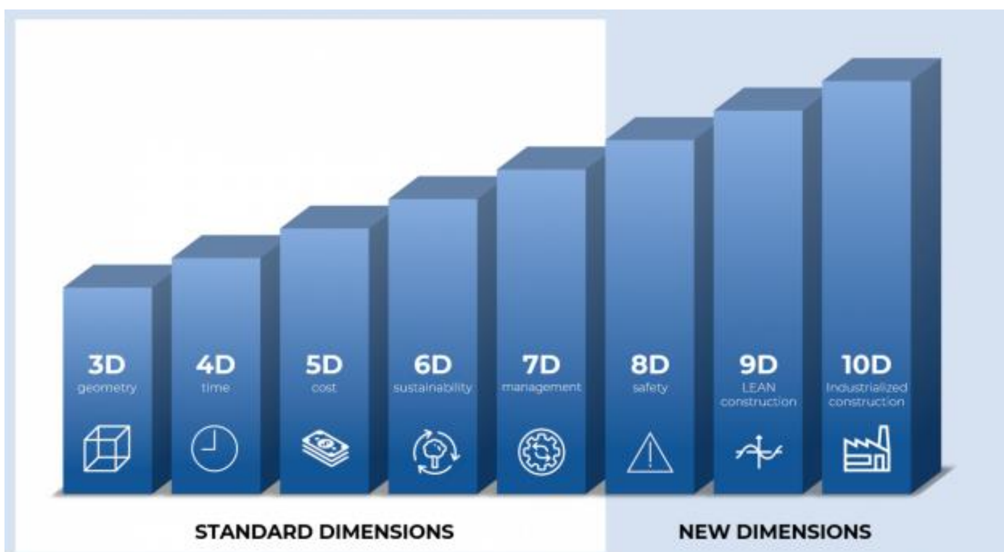

Figure 2: BIM dimensions 3D to 10D. (harmony-at, 2023)

I.1.5. BIM vs CAD vs GIS:

- **CAD Definition:** Computer-Aided Design (CAD) is a digital design methodology used to create both 2D drawings and 3D models of products prior to their physical fabrication. In this context, CAD enables designers and engineers to visualize and refine product geometry, thus improving design accuracy and efficiency. It involves the use of computer systems to develop, modify, and optimize designs, and therefore has largely replaced traditional manual drafting processes. As a result, CAD has become an essential tool in modern design and manufacturing workflows. (University, 2022)
- **GIS Definition:** A Geographic Information System (GIS) is a computer-based system designed to capture, store, process, analyze, manage, and display geographically referenced data. In this context, spatial data, linked to specific locations on the Earth's surface are integrated with attribute (tabular) data that describe the characteristics of each entity, such as the name, function, or capacity of a facility. This integration enables advanced spatial analysis and supports informed decision-making. Accordingly, GIS is not limited to software alone, but encompasses people, methodologies, and geomatic tools that collectively facilitate the visualization and systematic analysis of geographic information through maps and graphical representations. (Martindale, n.d.)

Linking and developing CAD with BIM represents a critical step toward improving project performance in architecture, engineering, and construction. By combining the precision of CAD-based 2D drawings with the data-rich environment of BIM models, this integration enhances both design quality and project coordination. (Sharma a. , 2024)

In practice, CAD drawings can be directly imported into BIM models, thereby reducing errors and improving the accuracy of construction documentation and scheduling. As a result, any modification in the BIM model automatically updates the associated drawings, which minimizes manual rework and prevents inconsistencies between different plans. (Digital PV, 2025)

Furthermore, the BIM model functions as a centralized information hub shared across disciplines, while CAD remains a reliable source for detailed 2D representations. This complementary relationship improves coordination among project stakeholders and enables early detection of clashes before construction begins.((ViBIM, 2025) (Digital, 2025))

In addition, this integration supports advanced project management functions such as 4D scheduling and 5D cost estimation, as quantities and timelines can be extracted directly from the BIM model linked with CAD data. It also facilitates the recording of operation and maintenance information, thus extending the use of BIM into the building's lifecycle management phase. (Sharma a. , 2024) (ViBIM, 2025)
Overall, recent studies and technical developments indicate that integrating CAD with BIM significantly improves design accuracy, reduces rework, and enhances collaboration among project teams. Therefore, it represents a key advancement in achieving more efficient and reliable project delivery in the construction industry. (Digital PV, 2025)

Figure 3: CAD and BIM WORKFLOW COMPARISON (Sharma P. , 2025)

Linking and developing GIS with BIM is highly important for the planning, execution, and operation of buildings and urban projects. In this context, GIS provides the broader spatial environment such as site conditions, terrain, and infrastructure networks while BIM offers a detailed representation of the building within that environment. Accordingly, this complementary relationship strengthens decision-making processes and enhances project understanding. (Mekawy, 2025)
This integration facilitates the incorporation of precise geographic data such as terrain, utility networks, and flood risks into the BIM model, thereby enhancing environmental analysis and informed planning decisions. (Nemetschek dtwin, 2025)
Regarding the issue of data fragmentation, this synergy effectively bridges the gap between spatial context and architectural design, ensuring a comprehensive perspective that merges building information with its surrounding environment. (Mekawy, 2025)

The integration of GIS and BIM underpins the comprehensive lifecycle management of assets and infrastructure; specifically, it links BIM models with spatial data concerning asset locations and service networks to streamline maintenance and performance monitoring. (Trinh Van Hoaa, Tri N.M. Nguyen, Le Van Phuc , & Ngo Chau Phuong, 2025)
Recent studies in civil engineering and transportation have demonstrated that the greater the reliance on this integration, the more project data quality improves and costs are reduced, consequently mitigating delays in infrastructure implementation. (Chen Z. , Shi, Ji, Luo, & Lin, 2023)
This combination also serves as the fundamental basis for urban digital-twin models, wherein intricate BIM models are synthesized with GIS-based general city data. (Kalantari, Clemen, & Jadidi, 2024)
It is further noted in a specialized text on "BIM and 3D GIS Integration for Digital Twins" that this technical convergence establishes a framework for integrating extensive geographic datasets with structural BIM models. (Kalantari, Clemen, & Jadidi, 2024)
A review article focusing on "BIM/GIS data integration from the perspective of information flow" also emphasizes that this connection improves information exchange between the two systems, which is essential for optimized smart-city management. (Zhu & Wu, 2022)

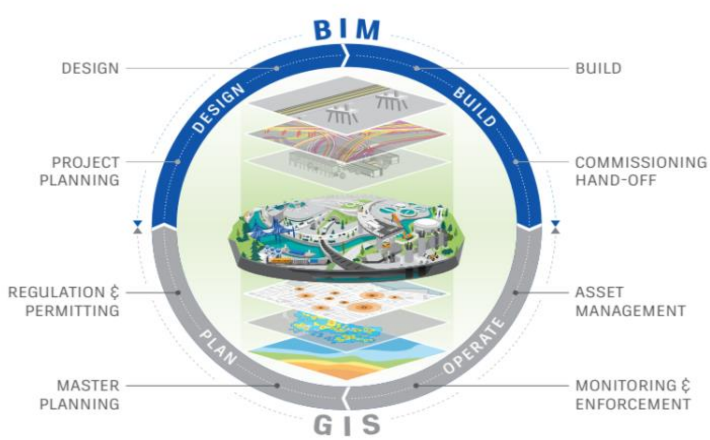

Figure 4:Differences Between BIM And GIS Technologies. (harmony_at, 2023)

I.2. BIM Data Structure and Interoperability:

The BIM data structure is organized as a hierarchical, parametric database in which geometric elements such as walls or beams are interconnected with comprehensive semantic attributes, including materials, costs, performance specifications, and inter-element relationships. This framework facilitates the development of an intelligent 3D/ND model rather than traditional static drawings. Furthermore, this structure is predicated on object-oriented modeling, where elements are defined by specific classes, properties, and constraints, thereby enabling advanced automation, clash detection, and simulations throughout the various phases of the project lifecycle. (Sacks, Eastman, Lee, & Teicholz, 2018)
Interoperability within the BIM ecosystem ensures the seamless exchange of data between diverse tools through open standards such as Industry Foundation Classes (IFC), which preserve geometric, spatial, and non-geometric information without data loss during transfers between platforms like Revit, ArchiCAD, or Solibri. Specifically, it employs formats such as COBie for facility management handovers and IDS for validation purposes; consequently, the rigorous application of these standards reduces data silos and minimizes errors within collaborative workflows governed by ISO 19650. (Pauwels, Bersselaar, & Verhelst, 2024)

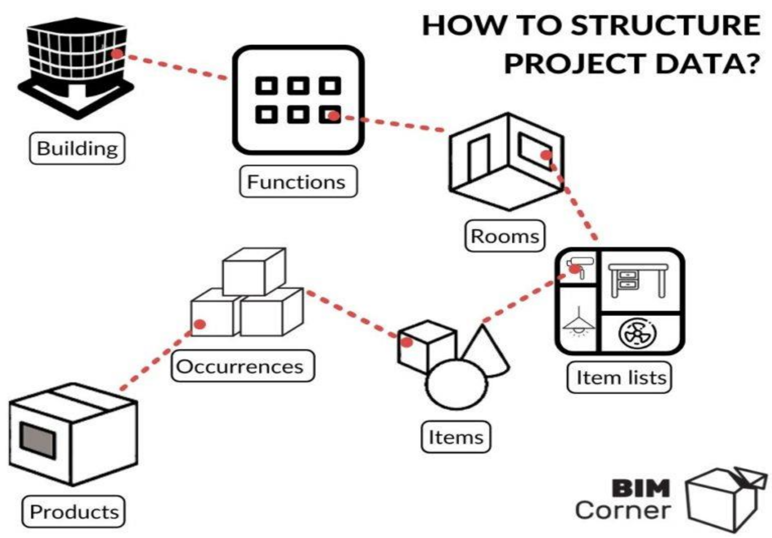

Figure 5: How to structure project data. (Fugas, 2020)

I.2.1. Object-oriented modeling in BIM:

Object modeling in Building Information Modeling (BIM) involves representing building components as parametric, intelligent objects, such as walls, doors, or HVAC systems, where each object is defined by its inherent properties (e.g., dimensions, materials), behaviors (e.g., structural performance), and interdependencies. This approach facilitates a dynamic digital twin where modifications automatically propagate across the model to ensure consistency and automation. This methodology leverages object-oriented principles, including classes, inheritance, and encapsulation, to structure complex building data hierarchically. Furthermore, it supports comprehensive simulations, analysis, and collaboration throughout the project lifecycle. (Lu & Anumba, 2022)

- **Advantages**:

- **Intelligence**: Objects embed decision-making logic and rules, thereby enabling automated response, such as load calculations or compliance checks without manual intervention. (Succar, 2009)
- **Parametric**: In BIM, parametric objects are governed by parameters and rules that drive their geometry and attributes; consequently, modifying the dimensions of a space can automatically update related elements, such as walls, floors, and ceilings, while preserving their defined relationships. (Akın, 2025)
- **Relationships**: Objects maintain explicit links (e.g., door-to-wall hosting), ensuring spatial and functional integrity during design revisions. (Akın, 2025)
- **Reusability**: Standardized object libraries can be reused across various projects; moreover, the greater the utilization of these libraries, the more modeling is accelerated while enforcing consistency and reducing errors. (Abdulghani & Gullu, 2021)

I.2.2. Levels of Development (LOD):

Level of Development (LOD) specifications define the increasing detail, reliability, and usability of model elements across project stages from rough conceptual shapes to precise as-built records assisting teams in coordinating expectations regarding geometry, attributes, and accuracy for tasks such as estimation, fabrication, and maintenance. These levels (100~500) ensure progressive model maturity. Regarding each LOD, it defines exact requirements for visualization, coordination, and data attachment to minimize risks and support contractual clarity. (BIM FORUM, 2024) ,LOD Levels Explained:

- **LOD 100 (Conceptual):** Uses symbolic representations or generic elements with estimated dimensions, positioning, and quantities, making it suitable for preliminary feasibility assessments, conceptual massing studies, and high-level cost estimation without including detailed product information.
- ***LOD 200 (Approximate/Schematic):*** Incorporates generic systems or assemblies with nominal dimensions, orientations, and performance characteristics; specifically, it supports preliminary spatial analysis, energy simulations, and approximate quantity takeoffs during the design development phase.
- **LOD 300 (Precise):** Models the exact size, shape, location, and quantity of specific components with accurately modeled assemblies; consequently, it supports detailed construction drawings, clash detection, and fabrication planning. (Vi BIM, 2025)
- **LOD 350 (Construction):** Incorporates interfaces, connections, and coordination details between building elements and trades; therefore, it emphasizes clash detection and spatial validation to support contractor coordination and procurement activities.
- **LOD 400 (Fabrication/Assembly):** Delivers fabrication-level detail including precise tolerances, supports, and installation requirements; moreover, it is used for manufacturing, prefabrication, and on-site assembly processes by contractors.
- **LOD 500 (As-Built/Operational):** LOD 500 represents an as-built BIM model in which the geometry and quantities of elements are field-verified to accurately reflect real conditions. At this level, detailed operational information is integrated, making the model a dependable foundation for asset management; consequently, it functions as a digital twin of the facility throughout the operational phase and the entire building lifecycle.

Figure 6: Understanding LOD in BIM Level of Development100-500. (techno struct academy, 2024)

I.2.3. Level of Information (LOI):

The Level of Information (LOI) serves as a metric for the quantity and quality of non-geometric data embedded within model elements, such as material specifications, costs, maintenance schedules, and manufacturer details. It distinguishes itself from the Level of Development (LOD), which emphasizes geometric precision, as LOI defines the descriptive data required for analytical and maintenance purposes. (techno struct academy, 2024)

LOI matures progressively throughout the project lifecycle; it initiates during the early design phase with fundamental data such as type and pricing, and subsequently expands during construction to encompass detailed information like warranties and thermal performance. This data is structured into property sets (Psets) to ensure compatibility with IFC and ISO 19650 standards. (revizto, 2024)

- ***Value in data preservation***: LOI organizes information within a centralized and standardized database, thereby preventing data fragmentation or duplication and ensuring informational continuity throughout the building's lifecycle from design to demolition. Furthermore, it facilitates automated backup processes and the efficient retrieval of data for reporting purposes. (trimble, 2021)
- ***Ease of use***: The system offers significant utility by enabling data filtering as required, allowing for the extraction of customized reports such as cost estimates or maintenance schedules. Moreover, the more these features are utilized, the more team collaboration is

supported by reducing errors through automated verification processes; consequently, the model remains functional at any stage without the need for manual data re-entry. (BIM FORUM, 2024)

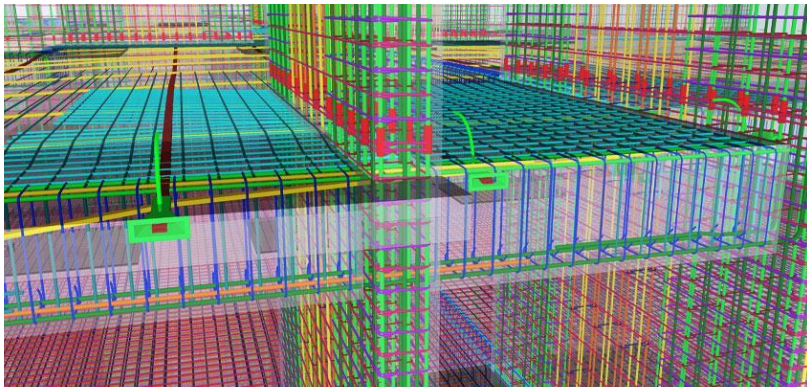

Figure 7: Level of Information Need. (trimble, 2021)

I.2.4. Data exchange formats (IFC, RVT, OBJ, FBX, GLTF):

Interoperability, defined as the ability of disparate software applications to seamlessly exchange and utilize data without compromising its integrity, constitutes the cornerstone of effective Building Information Modeling (BIM) implementation. This capability ensures that diverse stakeholders including architects, engineers, contractors, and facility managers can collaborate efficiently within fragmented software ecosystems. Consequently, this synergy reduces errors, minimizes rework costs, and accelerates project delivery timelines. (Rakib, Howlader, Rahman, & Bhuiyan, 2019)

- **IFC (Industry Foundation Classes):** IFC serves as the open, neutral standard for BIM data exchange (ISO 16739), preserving comprehensive geometric, semantic, and relational data across platforms such as Revit, ArchiCAD, and Tekla. Furthermore, it facilitates vendor-neutral collaboration by supporting hierarchical structures and property sets; specifically, this enables lossless model federation while avoiding proprietary software lock-in. (Ali, 2025)
- **RVT (Revit):** Regarding the RVT format, it is the proprietary native format of Autodesk Revit, which stores parametric families, project data, views, and schedules within a single relational database optimized for Revit-specific workflows. While it excels in detailed

design authoring, it necessitates export processes for interoperability and limits direct editing capabilities in non-Revit tools. (Team Editorial, 2025)

- **OBJ:**OBJ is a widely used text-based geometry format for representing 3D meshes, including vertices, faces, and texture information, mainly employed for visualization and rendering applications. Nevertheless, it does not contain BIM-specific semantics, parametric data, or relational information; consequently, it is primarily used for exporting models to visualization software such as Autodesk 3ds Max.
- **FBX (Film box):** FBX is a binary format developed by Autodesk for 3D assets, encompassing geometry, animations, and materials. Although it bridges the gap between BIM and game engines or VR environments, the more it is utilized for such transitions, the more it tends to strip away intelligent BIM metadata, such as cost data and project schedules. (revizto, 2025)
- **GLTF (GL Transmission Format):** GLTF is an efficient, royalty-free, JSON-based 3D format designed for web and mobile-based AR/VR applications. Specifically, it excels at compressing meshes and textures for lightweight performance; however, it omits BIM-specific metadata, rendering it suitable exclusively for high-performance visual presentation rather than detailed information management. (bim collab, 2024)

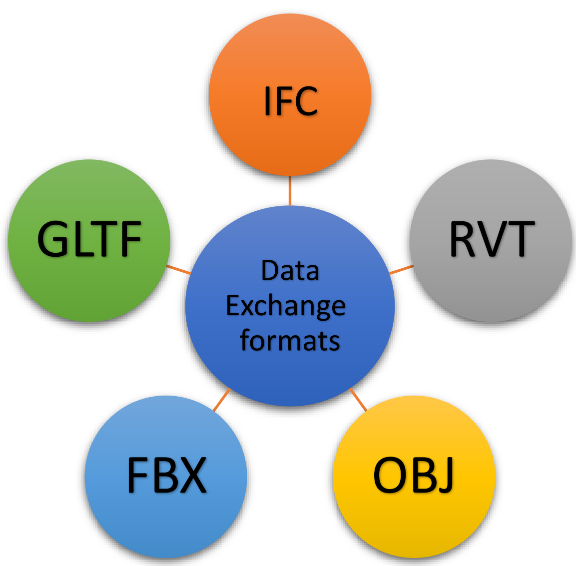

Figure 8:Data exchange formats. (Ours)

I.3. Digital Technologies Supporting BIM:

Digital technologies supporting Building Information Modeling (BIM) can be conceptualized as integral components of broader digital construction ecosystems. Within these frameworks, BIM serves as the fundamental digital representation and central data repository around which auxiliary tools and services are structured. In addition, BIM facilitates the creation of intelligent data and its continuous enrichment, while cloud platforms and analytical engines federate models to maintain digital continuity and orchestrate information flows throughout the asset lifecycle. Consequently, by positioning BIM at the core of these digital ecosystems, organizations can leverage predictive capabilities and collaborative workflows that transform static project documentation into an integrated decision-support environment for smarter and more sustainable construction. (pollution sustainability directory, 2025) ,The following sections discuss the common digital building systems currently utilized within the BIM framework:

I.3.1. Virtual Reality (VR):

virtual reality (VR): Regarding Virtual Reality (VR), it is defined as the utilization of computer modeling and simulation to enable individual interaction with an artificial three-dimensional (3-D) visual or sensory environment. Specifically, VR applications immerse users in a computer-generated environment that simulates reality through interactive devices such as goggles, headsets, gloves, or body suits which facilitate the bidirectional exchange of information. (Lowood, 2026)

- ***VR Applications:***

VR is used with BIM to enable immersive visualization of 3D building models, allowing engineers and clients to explore designs as if they were inside the actual building, with free navigation and interaction with elements. (Sampaio, Sarmento, & Gomes, 2024)

In site planning and construction, VR is applied to simulate workflows, such as identifying equipment clashes, material distribution, and testing safety protocols prior to actual physical execution. (Wang, Wu, Wang, Chi, & Wang, 2018)

In addition, it is utilized during the operation and maintenance phase to train personnel on complex building systems such as HVAC or electrical networks through sophisticated interactive scenarios. (Shehadeh, et al., 2025)

■ ***Advantages***:

- **Improved Communication and Collaboration**: VR allows immersive sharing of BIM models, reducing misunderstandings between teams and enhancing client understanding of the design. (Sampaio, Sarmento, & Gomes, 2024)
- **Early Clash Detection**: The technology enables 1:1 model exploration to identify engineering issues that remain invisible on traditional 2-D screens, consequently saving significant time and costs. (Wang, Wu, Wang, Chi, & Wang, 2018)
- **Safe and Effective Training**: It provides a high-fidelity simulated environment for training workers on hazardous procedures without exposure to real-world risks. (Shehadeh, et al., 2025)
- **Innovation Enhancement**: It supports the testing of virtual scenarios to refine design and sustainability in early stages; the more these tools are utilized, the more optimized the final project outcomes become. (Megat, et al., 2024)

■ ***Disadvantages:***

- **High Initial Cost**: The requirement for advanced VR hardware and specialized BIM integration software can impose significant budgetary constraints on smaller projects. (Wang, Wu, Wang, Chi, & Wang, 2018)
- **Interoperability Issues:** Furthermore, the conversion of detailed BIM models into VR formats often results in the loss of metadata or technical errors during software exchange. (Sampaio, Sarmento, & Gomes, 2024)
- **Cybersickness:** Some users may experience dizziness or nausea resulting from rapid motion or system latency. (Beverly, Love, & Love, 2024)
- **Accuracy and Performance Limits:** In other words, complex models may lead to processing delays or reduced visual quality when operated on devices with insufficient computational power. (Shehadeh, et al., 2025)

Figure 9: Virtual Reality. (planRadar, 2024)

I.3.2. Augmented reality (AR):

Regarding Augmented Reality (AR), it is a technology that overlays digital information such as images, sounds, or 3D models onto the real-world environment in real time. Specifically, it enhances the user's perception of reality through devices like smartphones, tablets, or AR glasses without entirely replacing the physical surroundings. (Poelman & Krevelen, 2010)

- ***AR Applications:***

AR integrates with BIM to superimpose digital building models onto physical sites via mobile devices, enabling the on-site visualization of planned structures within the real environment to ensure precise alignment during construction. (Yigitba, Nowosad, & Engels, 2023)

In addition, it supports facility management by displaying maintenance instructions, asset details, and repair guides linked to the BIM model directly onto physical equipment using AR markers or spatial tracking. (Khan, Panuwatwanich, & Usanavasin, 2024)

AR also facilitates collaborative design reviews by allowing stakeholders to conduct site walkthroughs while viewing BIM data overlays for clash detection, progress monitoring, and as-built versus as-planned comparisons. (Bhatarai, Banihashemi, Shakouri, & Antwi-Afari, 2024)

- ***Advantages:***

- **On-site Real-time Visualization:** AR allows workers to view BIM models aligned with physical spaces, consequently improving installation accuracy and reducing rework. (Yigitba, Nowosad, & Engels, 2023)
- **Enhanced Decision-Making:** It provides instantaneous access to BIM data overlays, which assists in rapid troubleshooting and quality verification during construction phases. (Khan, Panuwatwanich, & Usanavasin, 2024)
- **Improved Training and Safety:** Furthermore, it allows for hands-free guidance for complex tasks, minimizing errors and hazards without obstructing the user's field of vision. (Bhatarai, Banihashemi, Shakouri, & Antwi-Afari, 2024)
- **Stakeholder Collaboration:** It enables both remote and on-site teams to share AR-enhanced BIM views, thereby improving communication; the more effectively these views are shared, the greater the project lifecycle efficiency. (Yigitba, Nowosad, & Engels, 2023)

- ***Disadvantages:***

- **Hardware and Tracking Limitations:** The technology relies on device sensors that are susceptible to drift or occlusion, which can lead to misalignment in dynamic environments. (Yigitba, Nowosad, & Engels, 2023)
- **Data Interoperability Challenges:** Furthermore, converting detailed BIM models into AR formats often results in information loss or compatibility conflicts. (Bhatarai, Banihashemi, Shakouri, & Antwi-Afari, 2024)
- **High Implementation Costs:** The requirement for specialized AR tools, software development, and training can be cost-prohibitive for smaller-scale projects. (Khan, Panuwatwanich, & Usanavasin, 2024)
- **User Adoption Barriers:** The technology demands technical proficiency and can lead to visual fatigue; in other words, it may potentially distract users from real-world hazards on construction sites.(Yigitba, Nowosad, & Engels, 2023)

Figure 10: Augmented reality. (Srivastava, 2025)

I.3.3. Extended Reality (XR):

Extended Reality (XR) serves as an umbrella term encompassing a spectrum of immersive technologies including Virtual Reality (VR), Augmented Reality (AR), Mixed Reality (MR), and their hybrids. These technologies blend the physical and digital worlds to create interactive experiences with varying degrees of immersion; furthermore, this is facilitated through devices such as headsets, glasses, or mobile screens to enhance human perception and interaction. (Rauschnabel, Hinsch, Alt, Felix, & Shahab, 2022)

- ***XR Applications:***

XR is integrated into BIM environments for immersive project visualization, where VR enables full-scale walkthroughs of federated models. Regarding Augmented Reality, it overlays digital BIM data onto physical sites for precise alignment checks, while Mixed Reality allows for simultaneous interaction with real and virtual elements during design reviews. (Wu, Hou, & Zhang, 2021)

In addition, construction monitoring is supported by projecting XR-enhanced BIM instructions onto job sites, enabling real-time progress tracking, clash resolution, and guidance for precise installations via AR glasses or tablets. (levery, 2025)

XR further facilitates facility management and training through digital twins, wherein users interact with BIM-derived XR environments for maintenance simulations, safety drills, and remote inspections throughout the building lifecycle. (Mohammadrezaei, GHASEMI, DONGRE, GRAČANIN, & ZHANG, 2024)

■ ***Advantages:***

- **Unified Immersive Collaboration:** XR integrates VR, AR, and MR for multi-stakeholder reviews of BIM models, consequently overcoming geographical barriers and improving design feedback loops. (Wu, Hou, & Zhang, 2021)
- **Precision On-Site Guidance:** AR/XR overlays reduce errors by aligning BIM data with physical reality; furthermore, this enhances installation accuracy and minimizes rework. (levery, 2025)
- **Scalable Training Simulations:** This technology enables risk-free, repeatable BIM-based training scenarios for complex procedures, which specifically boosts worker competency. (Mohammadrezaei, GHASEMI, DONGRE, GRAČANIN, & ZHANG, 2024)
- **Lifecycle Data Enrichment:** It supports predictive maintenance via XR digital twins; moreover, the more these technologies are utilized, the more post-construction operations and sustainability are optimized. (Wu, Hou, & Zhang, 2021)

■ ***Disadvantages:***

- **Technical Integration Complexity:** Merging diverse XR modalities with BIM requires custom development; furthermore, this often leads to compatibility issues across various software tools. (Wu, Hou, & Zhang, 2021)
- **Hardware Dependency and Costs:** High-end XR devices limit accessibility, and regarding the AEC sector, they necessitate substantial upfront investments. (levery, 2025)
- **User Health and Fatigue Risks:** Prolonged XR sessions can induce motion sickness, eye strain, or disorientation, which consequently affects overall productivity. (Rauschnabel, Hinsch, Alt, Felix, & Shahab, 2022)
- **Scalability for Large Models:** The more detailed the BIM models are within an XR environment, the higher the computational power required, which often causes latency in

complex environments. (Mohammadrezaei, GHASEMI, DONGRE, GRAČANIN, & ZHANG, 2024)

Figure 11: Extended Reality. (Ellis, 2026)

I.3.4. The Internet of Things (IoT):

The Internet of Things (IoT) refers to a network of interconnected physical devices, vehicles, and buildings embedded with sensors, software, and connectivity to collect, exchange, and analyze data autonomously. Specifically, this enables machine-to-machine (M2M) communication and intelligent automation without constant human intervention. (Mouha, 2021)

- ***IOT Applications:***

IoT integrates with BIM by deploying sensors on construction sites to feed real-time data such as environmental conditions and progress metrics into BIM models; furthermore, this enables dynamic updates and accurate lifecycle simulations from design to operation. (Hong & Guo, 2025)

In addition, it supports predictive maintenance in post-construction phases, where IoT devices monitor systems like HVAC or structural health, linking live data streams to BIM digital twins for proactive alerting. (Shishehgarkhaneh, Keivani, Moehler, Jelodari, & Laleh, 2022)

IoT also enhances safety and quality management in BIM workflows by utilizing RFID tags and sensors for material tracking and hazard monitoring, which are consequently visualized directly within BIM platforms. (Hong & Guo, 2025)

- ***Advantages:***

- **Real-Time Monitoring and Updates**: IoT sensors provide continuous data feeds to BIM, allowing for instantaneous model synchronization; furthermore, this leads to superior progress tracking and decision-making. (Hong & Guo, 2025)
- **Cost and Efficiency Savings**: It automates data collection to reduce manual errors, rework, and delays, consequently optimizing resource allocation across project phases. (Shishehgarkhaneh, Keivani, Moehler, Jelodari, & Laleh, 2022)
- **Enhanced Safety and Sustainability**: It proactively detects hazards and monitors energy consumption, which specifically supports safer job sites and greener buildings via BIM analytics. (Hong & Guo, 2025)
- **Lifecycle Extension**: It enables ongoing asset management post-handover; moreover, the more IoT is integrated with BIM, the more predictive maintenance is improved and downtime is reduced. (Shishehgarkhaneh, Keivani, Moehler, Jelodari, & Laleh, 2022)

- ***Disadvantages*:**

- **Interoperability Challenges**: Inconsistent protocols between IoT devices and BIM software often lead to data silos and integration difficulties. (Hong & Guo, 2025)
- **Cybersecurity Vulnerabilities**: Connected sensors expose BIM systems to potential hacks; furthermore, this risks data breaches and significant operational disruptions. (Graham, 2023)
- **High Initial and Maintenance Costs**: The deployment of sensors, networks, and cloud infrastructure can financially burden smaller-scale projects. (Shishehgarkhaneh, Keivani, Moehler, Jelodari, & Laleh, 2022)
- **Data Overload and Complexity**: Regarding the massive volumes of IoT data, they can overwhelm BIM analysis without advanced processing; consequently, this leads to management challenges and data congestion. (Hong & Guo, 2025)

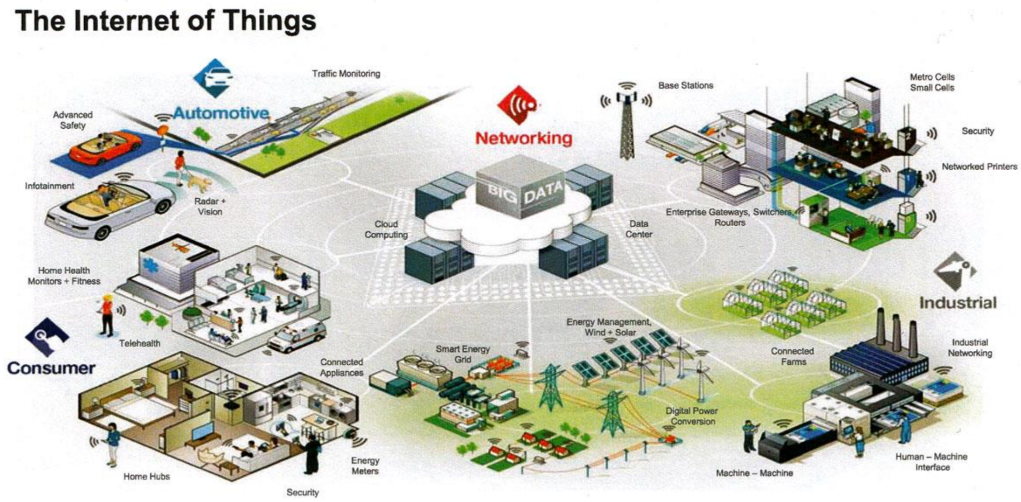

Figure 12: IoT in every aspect of our lives. (Osisanwo, Izang, Kuyoro, & Chukwudi, 2016)

I.3.5. LiDAR (Terrestrial and Aerial):

I.3.5.a. Terrestrial LiDAR:

constitutes a ground-based laser scanning technology that emits millions of rapid laser pulses from stationary scanners positioned at multiple vantage points. Specifically, it measures distances to surrounding surfaces to generate high-density 3D point clouds with millimeter-level accuracy; consequently, this facilitates the detailed capture of building interiors, facades, and complex geometries within line-of-sight ranges of several hundred meters. (Survey, 2011)

I.3.5.b. Aerial LiDAR:

it deploys laser scanners mounted on aircraft, drones, or helicopters to survey extensive outdoor areas rapidly. It integrates GPS/IMU systems for precise georeferencing while penetrating vegetation canopies; in other words, this allows for the mapping of topography, structures, and elevations over vast areas with resolutions typically ranging from centimeters to decimeters. (Montoya-Sánchez, et al., 2024)

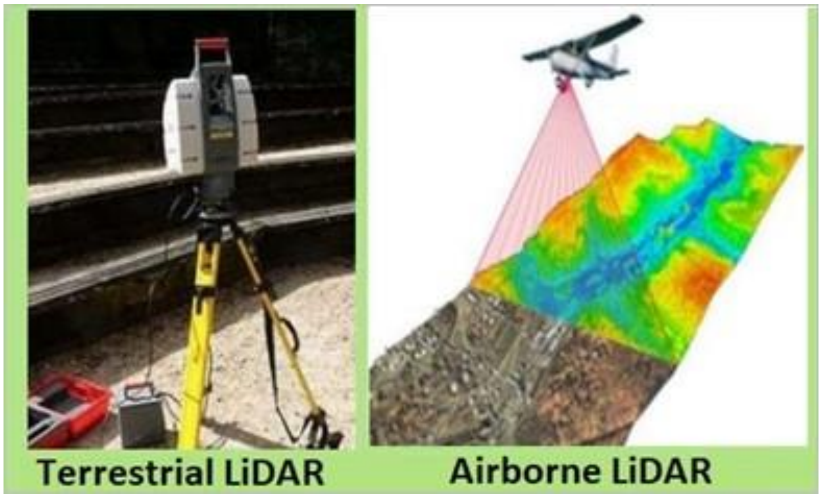

Figure 13: Two types OF LiDAR (Terrestrial and Airborne). (Kumar, 2023)

- ***LIDAR Applications:***

Terrestrial LiDAR is utilized in Scan-to-BIM workflows, wherein point clouds from interior scans are registered and segmented to reconstruct as-built models either automatically or semi-automatically. Furthermore, this enables effective renovation planning and clash detection against original design intents. (Fernández, Rodríguez, & García, 2025)

Regarding Aerial LiDAR, it supports BIM site analysis by generating topographic meshes and digital elevation models (DEMs) integrated into civil/site BIM environments; consequently, this assists in earthwork calculations, drainage design, and the contextual placement of proposed structures. (Blaszczak-Bak, Masiero, Bąk, & Kuderko, 2024)

Moreover, the fusion of terrestrial and aerial LiDAR data creates comprehensive BIM digital twins that combine high-detail indoor scans with broad topographic data; specifically, this supports full lifecycle management, facility retrofitting, and urban planning simulations. (Fernández, Rodríguez, & García, 2025)

- ***Advantages:***

- **Millimeter Precision:** Terrestrial LiDAR delivers sub-centimeter accuracy, which is essential for BIM model validation and maintaining fabrication tolerances. (Survey, 2011)
- **Comprehensive Coverage:** Aerial surveys map vast areas expeditiously, capturing site contexts that are often inaccessible to ground-based scanners. (Montoya-Sánchez, et al., 2024)

- **Non-Contact Data Capture:** Specifically, this enables the safe documentation of hazardous or inaccessible structures without the need for direct physical measurement. (Fernández, Rodríguez, & García, 2025)
- **Rich Semantic Potential**: Point clouds support automated feature extraction for LOD 300+ BIM elements; the more detailed the cloud, the more efficiently MEP routing can be identified. (Fernández, Rodríguez, & García, 2025)

■ ***Disadvantages:***

- **Line-of-Sight Limitations:** Terrestrial scans require multiple setups for occluded areas; furthermore, this increases processing time and potential registration errors. (Scan to BIM Solutions, 2025)
- **Vegetation Interference:** Aerial LiDAR often struggles with dense canopies; consequently, it necessitates ground filtering to produce accurate terrain models. (Montoya-Sánchez, et al., 2024)
- **Large Data Volumes:** The more detailed the point clouds, the more they demand significant storage, computational power, and specialized expertise for BIM conversion. (Fernández, Rodríguez, & García, 2025)
- **High Equipment Costs:** In other words, professional scanners and post-processing software represent substantial upfront investments. (Survey, 2011)

Figure 14: Aerial LiDAR Drone (Zhu, Hattula, & Raninen, 2024)

I.3.6. Photogrammetry as a BIM Data Source:

Regarding Photogrammetry as a BIM data source, it is a technique that derives precise 3D geometric models and textured meshes from overlapping photographs. Specifically, it utilizes algorithms such as Structure-from-Motion (SfM) and Multi-View Stereo (MVS) to triangulate points and calibrate sensors; moreover, it serves as a cost-effective alternative to laser scanning for generating point clouds. (KONTRIMOVICIUS, JUSZCZYK, LESNIAK, USTINOVICHIUS, & MIEDZIAŁOWSKI, 2023)

■ ***Photogrammetry Applications:***

Photogrammetry is applied in BIM for as-built documentation of existing structures, where drone-captured images generate point clouds for import into platforms like Revit. Consequently, this enables the semi-automated modeling of facades, roofs, and heritage buildings. (CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)

Regarding construction progress monitoring, it supports the comparison of photogrammetric models against as-planned BIM models to quantify deviations and material volumes; furthermore, this is integrated via cloud platforms for better project oversight. (CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)

Additionally, it aids small-scale BIM projects, where smartphone photos generate geometry for LOD 200 models; the more streamlined this process becomes, the more it reduces surveying costs while maintaining BIM compatibility. (Siewczyńska & Zioło, 2022)

■ ***Advantages:***

- **Cost-Effective Data Capture**: It requires only standard cameras or smartphones rather than expensive scanners; consequently, it is ideal for budget-constrained BIM projects. (scan to bim solutions, 2024)
- **Rich Visual Textures**: The technique produces colorized models with photographic realism, which specifically enhances BIM visualization and client presentations. (Tang, 2025)
- **Flexible Accessibility**: Drone and ground-based photography can access hard-to-reach areas; furthermore, this facilitates the creation of comprehensive site models. (CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)

- **Rapid Large-Area Coverage**: It processes thousands of images quickly for topographic BIM inputs; in other words, it outperforms traditional manual surveys in efficiency. (KONTRIMOVICIUS, JUSZCZYK, LESNIAK, USTINOVICHIUS, & MIEDZIAŁOWSKI, 2023)

■ ***Disadvantages:***

- **Lower Geometric Accuracy**: It may suffer from distortions on reflective or low-texture surfaces; consequently, it is less precise than LiDAR for strict BIM tolerances. (Tang, 2025)
- **Processing Dependency**: It relies heavily on high-quality imagery and computational power; specifically, poor lighting conditions yield unreliable point clouds. (KONTRIMOVICIUS, JUSZCZYK, LESNIAK, USTINOVICHIUS, & MIEDZIAŁOWSKI, 2023)
- **Scale and Control Challenges**: It requires ground control points (GCPs) for absolute accuracy in georeferencing; furthermore, this is essential for maintaining BIM integrity. (CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)
- **Limited Penetration**: The technology cannot penetrate vegetation or occlusions; consequently, it is restricted for subsurface BIM modeling applications. (scan to bim solutions, 2024)

I.4. Photogrammetry within BIM Workflows:

Photogrammetry within BIM workflows is a computational process integrated into BIM pipelines that extracts 3D geometry, textures, and semantic information from sequences of 2D photographs captured via drones or smartphones. Specifically, through Structure from Motion (SfM) algorithms, dense point clouds or meshes are generated and subsequently cleaned and converted into parametric BIM elements, such as walls and roofs; consequently, this enables as-built modeling as a cost-effective reality capture alternative to LiDAR. (Pinho, 2021)

I.4.1. Scan-to-BIM and Image-to-BIM approaches:

Scan-to-BIM and Image-to-BIM approaches convert reality-capture data into intelligent BIM models of existing assets. Regarding Scan-to-BIM, it processes LiDAR or photogrammetric point clouds, whereas Image-to-BIM utilizes direct image analysis for automated reconstruction. (Pinho, 2021)

- ***Scan-to-BIM Approach:***

Scan-to-BIM is a workflow that transforms 3D point clouds into intelligent parametric models through registration and classification. Furthermore, it facilitates the creation of engineering elements, such as walls and columns, providing highly accurate documentation of existing buildings for renovation or management phases. (Pinho, 2021)

- ***Image-to-BIM Approach:***

Image-to-BIM relies on the direct analysis of 2D images using artificial intelligence and computer vision algorithms to recognize architectural elements. In other words, this method offers increased speed and reduced costs for small-scale or interior projects compared to traditional scanning. (Wong, et al., 2025)

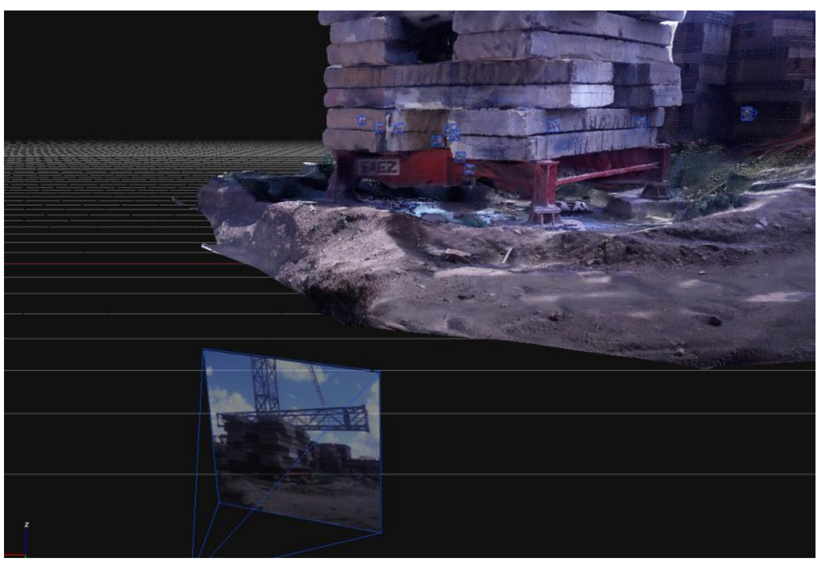

Figure 15:Camera positions for photogrammétrie data acquisition. (Ours)

I.4.2. As-built BIM and digital twins:

As-built BIM: refers to parametric digital models that accurately represent the constructed physical reality of a building post-construction. It incorporates all deviations from the original design intent; furthermore, it serves as a foundational record for facility management, renovations, and lifecycle analysis within BIM workflows. (Pinho, 2021)

Figure 16: From the existing building to the BIM as built model. (Editorial Team, 2026)

Digital twins: are dynamic, real-time virtual replicas of physical assets synchronized via IoT sensors and data streams. Specifically, they extend as-built BIM by enabling continuous simulation and predictive maintenance; moreover, the more integrated the bidirectional data becomes, the more performance optimization is achieved throughout the asset's operational lifecycle. (Shishehgarkhaneh, Keivani, Moehler, Jelodari, & Laleh, 2022)

I.4.3. Applications in construction, infrastructure, and heritage:

In the field of construction, applications refer to the practical utilization of materials and engineering technologies in executing structures. Consequently, they encompass planning, design, and management activities to ensure quality, safety, and adherence to budgets. (García, 2025)

These applications also extend to transforming drawings into a built reality. Furthermore, they focus on site organization, resource management, and effective coordination among engineers, workers, and suppliers. (Finity & Tell, 2025)

Regarding infrastructure, applications involve the use of engineering and digital technologies in the design and maintenance of major assets, such as roads and bridges. This aims to render systems more efficient, resilient, and sustainable in the long term. (Broo, Bravo-Haro, & Schooling, 2022)

These applications include the adoption of digital twins as cyber–physical tools for data integration. In other words, these models are utilized to improve operation and lifecycle management in smart infrastructure. (Akter, Bormon, Saikat, & Shoag, 2025)

In the field of built heritage, applications refer to conservation and digital documentation techniques. Specifically, the more these digital documentation tools are utilized, the more the historical value of cultural heritage sites is preserved. (Matrone, Colucci, Iacono, & Ventura, 2023)

These applications are reflected in integrated platforms that combine Historic Building Information Modeling (HBIM) with Geographic Information Systems (GIS). Consequently, this provides a multi-scale database that enables accurate condition analysis and risk management for heritage sites. (Dore & Murphy, 2012)

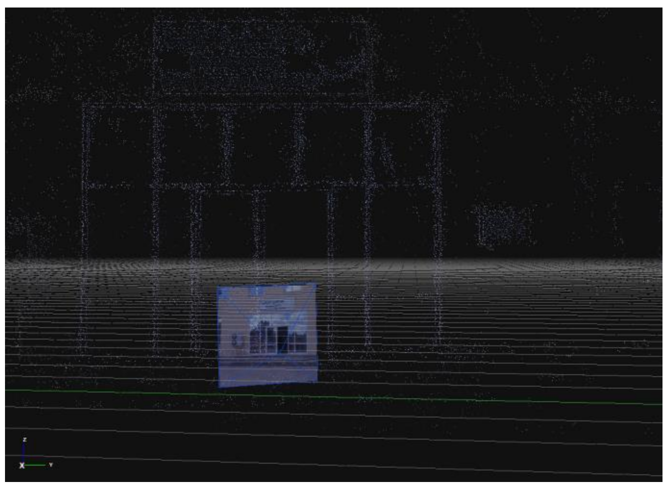

Figure 17:Displaying a point cloud of a building in a 3D design environment (Ours)

I.5. Practical Experiment:

To facilitate a better understanding of how photogrammetry is applied within a BIM workflow, a simple practical experiment can be conceptualized. The objective is to demonstrate, in a clear and accessible manner, how data derived from photogrammetry is processed and subsequently integrated into a BIM environment. Furthermore, this approach serves to bridge the gap between theoretical concepts and real-world implementation; specifically, it illustrates the transformation of captured imagery into functional digital models within the BIM framework. (Pinho, 2021)

I.5.1. Simple BIM project incorporating photogrammetric data:

A simplified BIM project based on photogrammetric data involved surveying a small commercial space, using the images to create a point cloud. This data was then integrated into a BIM model for design and follow-up purposes; thus, this methodology contributes to improving the accuracy of as-built documentation and the execution of small projects. (Albeaino, Eiris, & Gheisari, 2024)

- ***Project Overview:***

This example models a perfume and cosmetics franchise shop (approximately 35 m² in sales area) by employing photogrammetry for data acquisition and BIM for 3D parametric modeling. Photogrammetric point clouds generate precise 3D elements, such as furniture and signage; regarding the integration process, these are imported into REVIT® or SketchUp® as BIM families. Moreover, the more these repetitive franchise layouts are standardized, the more errors in geometry and metadata are minimized.(CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)

Figure 18:A commercial cosmetics store (franchise) model using photogrammetric points and BIM technology. (CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)

■ ***Key Workflow Steps*:**

- **first**, data acquisition through the capture of overlapping photographs of the site and fixtures under uniform lighting; furthermore, these are processed into 3D point clouds.(CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)
- **Second**, point cloud processing involves noise reduction and scaling via designated checkpoints.(CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)
- **Third,** BIM modeling entails exporting OBJ/DAE files to BIM software; in other words, this involves creating families enriched with metadata.(CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)
- **Finally,** validation and output are achieved by comparing as-planned versus as-built models; consequently, productivity gains exceeding 27% are realized compared to traditional CAD methods.(CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)

■ ***Benefits and Metrics:***

The integration of BIM with photogrammetry reduced project execution time by an average of 4.11 days and specifically saved approximately 6 days in decision-making time across 121 projects. The return on investment (ROI) reached 41.88% in the first year; consequently, this system outperformed CAD methodologies while reducing project queries by 25%.(CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)

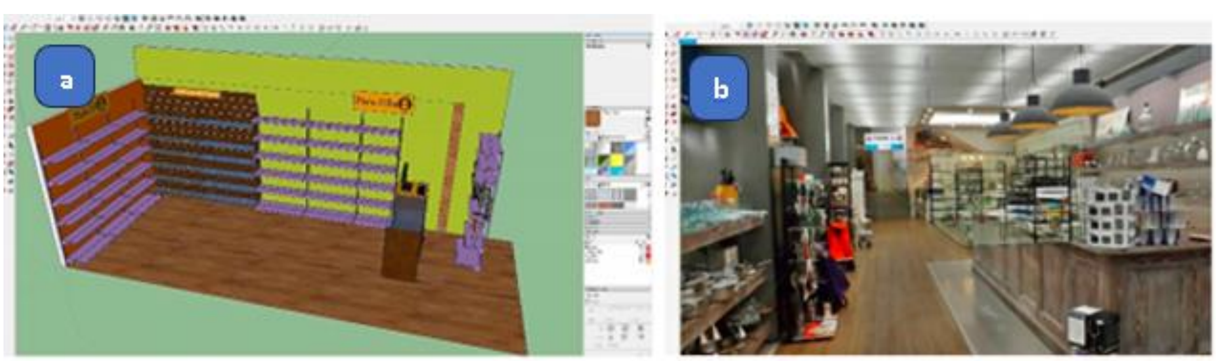

Figure 19:(a)and (b): Design lifecycle: from point cloud to final parametric model (CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)

I.5.2. Evaluation of integration constraints and benefits:

The evaluation of photogrammetric data integration into BIM balances advanced precision with technical challenges. Specifically, it involves converting images into accurate 3D models that enhance construction documentation, although this requires the meticulous management of environmental factors such as lighting and visual occlusions.(CONDE, GARCÍA-SANZ-CALCEDO, & RODRÍGUEZ, 2020)

- ***Benefits:***

Photogrammetry provides up to 96% coverage in heritage buildings when utilized in conjunction with laser scanning; furthermore, this significantly accelerates initial project documentation. Regarding costs, they are substantially lower than those of LiDAR technology; moreover, the more an image overlap of 85-90% is achieved, the higher the accuracy and economic viability for budget-constrained digital construction projects. (Salah, Géczy, & Károlyfi, 2026)

- ***Constraints*:**

Visual occlusions resulting from complex architectural elements constitute a major constraint; specifically, without the use of UAVs, coverage may be limited to 54%. Consequently, inconsistent lighting conditions lead to gaps and noise that necessitate manual cleaning, which increases the manual workload. Finally, data loss can occur during file format conversions (such as OBJ to IFC); in other words, high hardware demands for processing massive datasets can hinder seamless integration and negatively impact overall project efficiency. (Photogrammetry Planning Challenges, 2025)

Chapter II: Classical Photogrammetry and Structure from Motion (SfM)

II.1. Definition and Overview of Photogrammetry:

Photogrammetry is a scientific discipline concerned with extracting precise metric and spatial information about physical objects and environments through the systematic analysis of photographic imagery. Grounded in the theoretical frameworks of geometry, optics, and computer vision, it enables non-contact three-dimensional (3D) reconstruction, making it particularly valuable in situations where direct physical measurement is impractical, hazardous, or potentially damaging to the subject under study. (kraus, 2007)

Historically, photogrammetry served primarily as a tool for topographic mapping and geodetic surveying. Early analogue methods relied on physical photographs and manual measurement procedures, demanding considerable operator expertise and time investment . (luhamann, 2014) The transition to digital photogrammetry transformed the discipline fundamentally: computational methods enabled automated feature detection, image orientation, and 3D model generation, substantially improving both accuracy and processing efficiency . (Remondino & Campana, 3D Recording and Modelling in Archaeology and Cultural Heritage :theory and best practies, 2014)

Contemporary photogrammetric practice is characterized by the integration of advanced computational algorithms that support the generation of highly detailed 3D models and geodetically accurate measurements. In particular, Structure from Motion (SfM) algorithms have enabled the automatic recovery of camera positions and the generation of sparse point clouds from overlapping, unordered image sets . (Westoby, Brasington, Glasser, Hambrey, & Reynolds, Structure-from-Motion photogrammetry: A low-cost, effective tool for geoscience applications, 2012) Subsequent application of Multi-View Stereo (MVS) methods then yields dense surface representations capable of capturing fine geometric detail . (Szeliski, 2010)

The foundational principles underpinning modern photogrammetry may be summarized as follows:

- **3D Reconstruction from 2D Imagery:** By identifying corresponding features across multiple photographs and computing their spatial relationships, photogrammetric systems recover accurate three-dimensional coordinates. (Hartley & Zisserman, 2003.)

- **Triangulation:** The geometric backbone of photogrammetry rests on triangulation employing the mathematical relationships between image correspondences to determine the spatial location of object points. Sufficient image overlap and angular diversity are prerequisites for robust reconstruction. (Szelisk, 2010.)
- **Camera Calibration:** Precise knowledge of the camera's intrinsic parameters (focal length, principal point, and lens distortion coefficients) and extrinsic orientation (position and attitude) is essential. Calibration errors propagate directly into the accuracy of the resulting 3D model. (Hartley & Zisserman, 2003.)
- **Imaging Conditions:** Image resolution, illumination, scene texture, and inter-image baseline collectively determine the quality and reliability of photogrammetric reconstruction. (luhamann, 2014)
- **Automated Processing Pipelines:** Modern workflows integrate SfM and MVS algorithms into highly automated pipelines that substantially reduce manual intervention while supporting large-scale model production. (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012)
- **Breadth of Application:** Photogrammetric methods are employed across diverse domains, including architecture, construction engineering, cultural heritage documentation, geospatial mapping, environmental monitoring, forensic investigation, and Building Information Modeling (BIM). (Remondino & Campana, 2014)

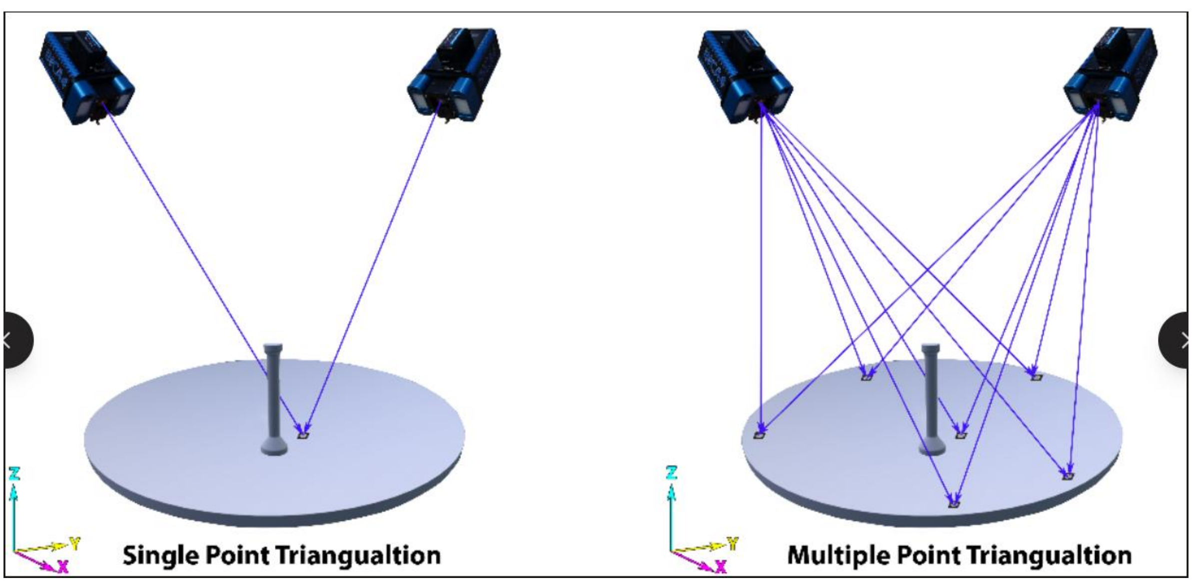

Figure 20: single vs. Multiple point triangulation (Mason, 2015)

II.2. Historical Development of Photogrammetry:

II.2.1. Analogue Photogrammetry:

The origins of photogrammetry may be traced to the nineteenth century, when photography was first recognized as an instrument for geometric surveying beyond its purely artistic applications. Aimé Laussedat's pioneering terrestrial experiments demonstrated that geometrically constrained photographs could be employed to derive object coordinates, laying the conceptual groundwork for the discipline. (Granshaw, 1980)

During this analogue era, stereoscopic techniques occupied a central role. The interpretation of parallax in overlapping image pairs allowed operators to perceive and exploit scene depth. Optical-mechanical instruments principally stereo plotters simulated the image-formation geometry through physical mechanisms rather than numerical computation. (Schenk, 2005) The widespread adoption of aerial photography in the early twentieth century elevated photogrammetry to a tool of national cartographic programs, its strategic value reinforced through extensive military and civil mapping operations. (McGlone J. C., 2013)

Nevertheless, analogue systems were constrained by the limitations of mechanical calibration procedures and the absence of rigorous statistical error propagation theory. The inability to model and distribute measurement uncertainties systematically created methodological inconsistencies that would eventually motivate the search for computational solutions. (Kraus, 1997)

Tableau 1:Key development stages of photogrammetry from analogue origins to AI-driven approaches

Era	Period	Key Technologies	Main Output	BIM Compatibility
Analogue Photogrammetry	1850s – 1960s	Film cameras, Stereo plotters	2D maps, Manual drawings	None
Analytical Photogrammetry	1960s – 1990s	Analytical plotters, Computers	3D coordinates, Digital maps	Minimal
Digital Photogrammetry	1990s – 2000s	Digital cameras, DEM generation	Orthophotos, DEMs, DSMs	Limited
SfM Era	2000s – 2015	SIFT/SURF algorithms, Bundle Adjustment	Dense Point Clouds, 3D Meshes	Partial (OBJ, PLY)

AI-Driven Era	2015 – Present	NeRF, Gaussian Splatting, Deep Learning	Neural Models, Splat Scenes	Emerging (IFC, glTF)

Sources:((Barazzetti, 2016)*;* (Ben Mildenhall, 2020)*;* (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)*)*

II.2.2. Analytical and Digital Photogrammetry:

The analytical phase of photogrammetric introduced mathematical formalism into the reconstruction process. Rather than relying on mechanical alignment, spatial relationships between image and object space were expressed through algebraic projective equations describing perspective projection. (Wolf & Dewitt, 2014.) The adoption of the collinearity condition equations provided a rigorous mathematical model of image formation, ensuring geometric consistency and enabling statistically optimal estimation through least-squares adjustment . (Triggs, mclauchlan, hartley, & Fitzibbon, 2000)

The formulation of bundle adjustment algorithms was particularly significant: by enabling the simultaneous optimization of camera parameters and object coordinates within a unified computational framework, this approach transformed photogrammetry into a thoroughly computational discipline . (Mikhail, 2001) The subsequent transition to fully digital photogrammetry driven by advances in sensor technology and algorithmic development further accelerated this evolution. The replacement of photographic film with digital sensors eliminated radiometric instability and scanning-induced distortions, while automated feature extraction and digital image matching dramatically increased the scalability of reconstruction workflows. (McGlone J. C., 2013)

II.2.3. Emergence of SfM Techniques:

Structure from Motion represented a paradigmatic shift in image-based 3D reconstruction, fundamentally reconceptualizing how scene geometry could be recovered from image correspondences. Unlike classical photogrammetry, which required a pre-established and controlled acquisition geometry, SfM operates on unordered image collections without prior knowledge of camera positions. (Hartley & Zisserman, 2003.)

A decisive methodological advance was the introduction of invariant feature descriptors robust to changes in scale and rotation. David Lowe's Scale-Invariant Feature Transform (SIFT) established a new standard for large-scale image correspondence and became an integral component of modern reconstruction pipelines . (Lowe, Distinctive image features from scale-invariant keypoints, 2004) Incremental SfM systems subsequently demonstrated that unordered image collections could be reconstructed through iterative camera pose estimation and triangulation (Snavely N. S., 2006), with global bundle adjustment ensuring geometric consistency across the entire image network (Triggs, mclauchlan, hartley, & Fitzibbon, 2000). The integration of Multi-View Stereo methods further extended SfM capabilities to produce dense surface reconstructions, positioning the combined SfM, MVS framework at the intersection of photogrammetry and computer vision research (Seitz, CURLESS, DIEBEL, & SCHARSTEIN, 2006).

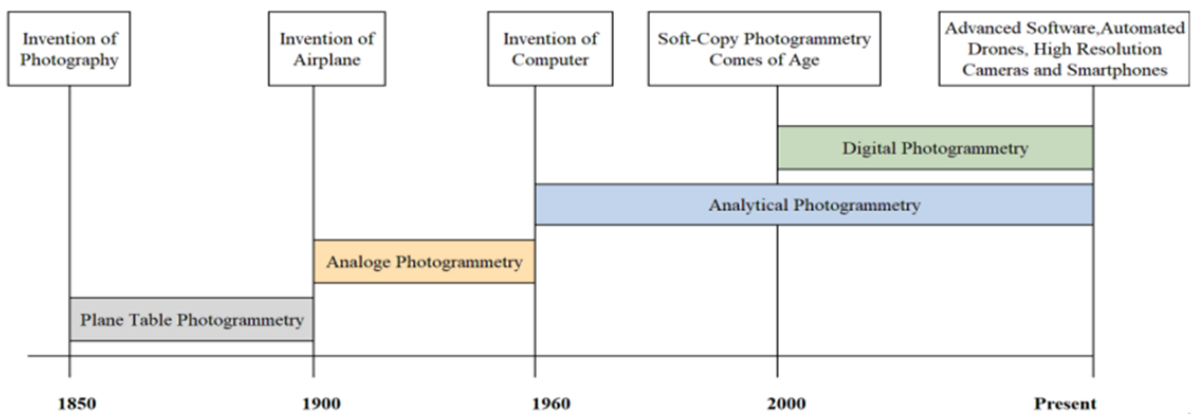

Figure 21: Historical evolution of photogrammetry (Linder, 2006)

II.3. Principles of Structure from Motion:

Structure from Motion is a photogrammetric technique that recovers three-dimensional scene geometry from a set of overlapping images captured from multiple viewpoints. Derived from the domain of computer vision, SfM enables the simultaneous estimation of camera intrinsic and extrinsic parameters alongside the spatial coordinates of scene points, without requiring prior knowledge of camera positions or a controlled acquisition configuration . (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012) These properties have established SfM as a widely adopted approach for low-cost 3D reconstruction across engineering and geospatial applications.

The fundamental operating principle of SfM is the exploitation of parallax: as the camera translates through space, stationary scene points appear to shift in position between successive image frames, and the magnitude of this apparent displacement encodes depth information. When a scene is photographed from multiple, geometrically diverse positions, common features are projected onto different pixel locations across images, allowing the reconstruction algorithm to triangulate their spatial coordinates. (Hartley & Zisserman, 2003.) The accuracy of this process is conditioned upon adequate image overlap, a sufficient camera baseline, and high photographic quality (James M. &., 2012).

II.3.1. Feature Detection and Matching:

The SfM reconstruction pipeline is initiated by automated key point extraction from each image in the dataset. Established operators such as SIFT (Lowe, Distinctive image features from scale-invariant keypoints, 2004), SURF, and ORB identify local image features characterized by their robustness to variations in scale, rotation, and illumination, thereby facilitating reliable correspondences across images captured under differing conditions. Each detected key point is represented by a compact numerical descriptor encoding the local appearance of the surrounding image region, enabling efficient similarity-based comparison across the image set (Hartley & Zisserman, 2003.).

The subsequent matching stage establishes inter-image correspondences through nearest-neighbor descriptor search. Since matching is susceptible to false associations arising from repetitive textures, occlusions, or sensor noise, robust outlier rejection procedures principally the Random Sample Consensus (RANSAC) algorithm are applied to eliminate erroneous matches and enforce geometric consistency (Wolf & Dewitt, 2014.). The integrity of the feature matching stage is of paramount importance, as errors introduced at this stage propagate through all subsequent reconstruction steps. (Hartley & Zisserman, 2003.)

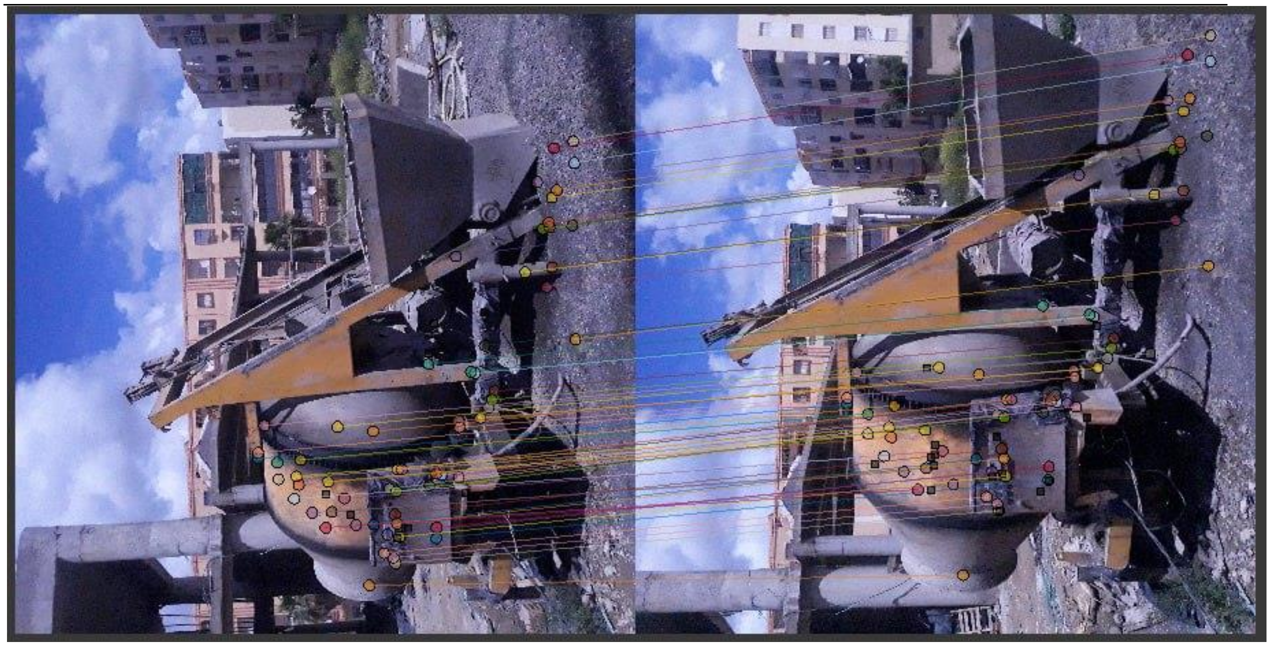

Figure 22:Matching point (Ours).

II.3.2. Camera Calibration and Bundle Adjustment:

Once a reliable network of inter-image correspondences has been established, the SfM pipeline proceeds to the estimation of camera parameters. Intrinsic calibration recovers the camera's internal optical characteristics focal length, principal point coordinates, and lens distortion coefficients while extrinsic calibration determines the six degrees of freedom pose (position and orientation) of each camera in the scene (Triggs, mclauchlan, hartley, & Fitzibbon, 2000). Accurate parameter estimation is essential for the reliable conversion of image measurements into metric three-dimensional coordinates (McGlone J. , 2004).

Many contemporary SfM implementations incorporate automatic self-calibration during the reconstruction process, circumventing the need for a separate pre-calibration procedure. While this increases operational flexibility, it introduces a dependency on the quality and geometry of the image network: poorly designed acquisition configurations can lead to parameter correlations and degraded reconstruction accuracy (Hartley & Zisserman, 2003.).

Bundle adjustment constitutes the central refinement procedure within the SfM pipeline. This nonlinear least-squares optimization simultaneously adjusts camera intrinsics, extrinsic orientations, and the three-dimensional coordinates of all reconstructed points by minimizing the aggregate reprojection error the discrepancy between observed image measurements and their predicted projections (Triggs, mclauchlan, hartley, & Fitzibbon, 2000). The quality of the bundle adjustment solution depends on network redundancy, image distribution, and the accuracy of initial feature correspondences (Wu C. , 2013). Modern SfM software typically applies bundle adjustment iteratively as images are progressively incorporated into the model.

II.3.3. Sparse and Dense Point Cloud Generation:

The immediate output of the SfM process is a sparse point cloud representing the three-dimensional positions of the matched feature points. Although the sparse model provides a structurally informative representation of scene geometry and serves as a useful diagnostic for assessing camera alignment quality, it captures only a limited subset of the scene surface, insufficient for detailed engineering analysis (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012).

To achieve more comprehensive surface coverage, SfM is routinely followed by dense image matching using Multi-View Stereo (MVS) algorithms. MVS methods exploit photometric consistency across multiple images to estimate depth values for a large proportion of image pixels, generating a dense point cloud of substantially higher spatial resolution (Furukawa & Ponce, 2010). The resulting dense cloud provides a near-continuous representation of the object surface suitable for engineering measurements, deformation analysis, and 3D modeling applications (Remondino & Campana, 2014).

In civil engineering contexts, dense SfM reconstructions are routinely employed for as-built documentation, volumetric calculations, crack detection, and construction progress monitoring. The accuracy of the final dense model is governed by image sharpness, overlap, camera network geometry, and the quality of the antecedent bundle adjustment (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012). Under well-controlled acquisition conditions, image-based dense reconstruction has been shown to approach the accuracy of terrestrial laser scanning for many practical applications.

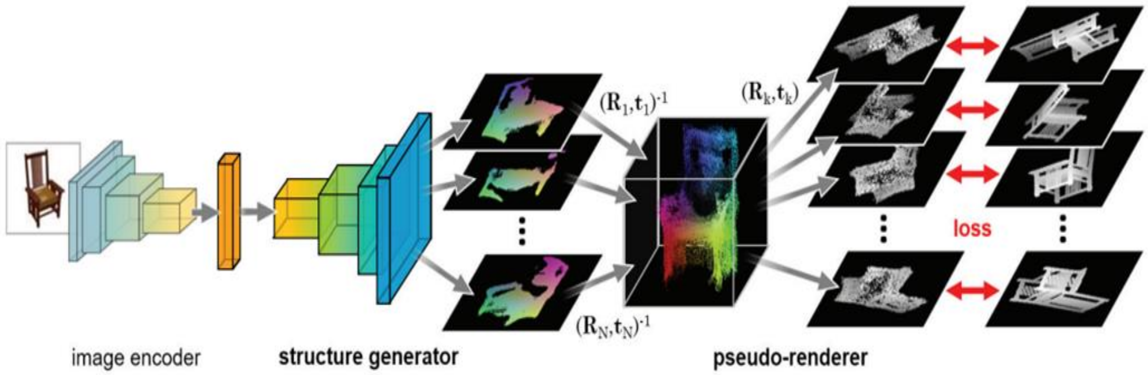

Figure 23: 3Dpoint cloud reconstruction pipeline (Lin, Kong, & Lucey, 2018)

II.4. SfM Processing Workflow:

The SfM processing workflow comprises a series of logically ordered computational stages that collectively transform a set of overlapping images into a geometrically accurate three-dimensional model (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012). The pipeline integrates foundational principles from both photogrammetry and computer vision, enabling a high degree of automation throughout the reconstruction process (Hartley & Zisserman, 2003.).

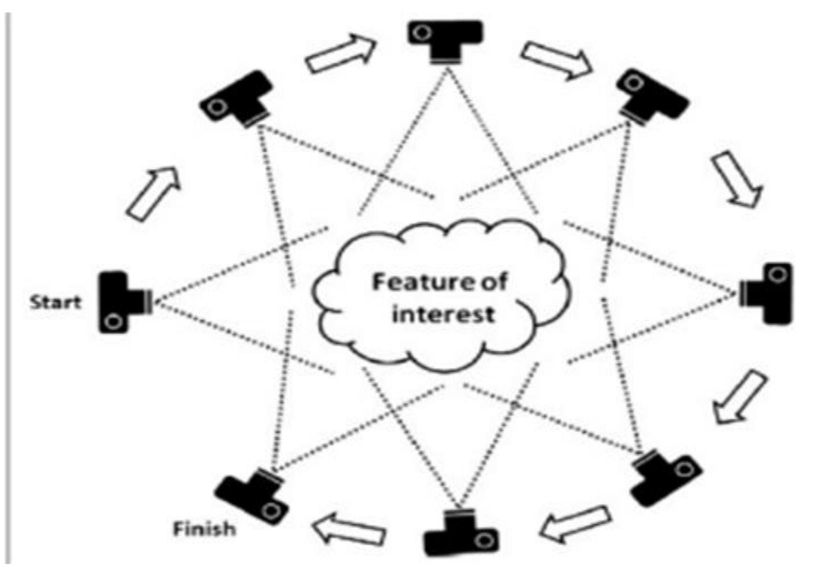

Figure 24:Circular multi-view camera acquisition strategy (Nyimbili, 2016)

II.4.1. Image Acquisition Procedures:

The quality and spatial configuration of the image dataset are the primary determinants of reconstruction success, and image acquisition must therefore be approached with systematic care (James M. &., 2012). Adequate image overlap is a fundamental requirement: forward overlap of 70–80% and lateral overlap of 40–60% are generally recommended to ensure sufficient correspondences for reliable feature matching (Wolf & Dewitt, 2014.). An appropriately distributed camera network minimizes geometric deficiencies and stabilizes the bundle adjustment solution.

The camera baselines the spatial distance between successive image positions exerts a direct influence on depth estimation accuracy. An excessively small baseline produces weak geometric intersection and depth uncertainty; an excessively large baseline, conversely, may render feature matching unreliable (Hartley & Zisserman, 2003.). The optimal baseline represents a balance between these competing constraints and depends on the scale and geometry of the scene.

Image quality parameters, including sharp focus, stable exposure settings, minimal motion blur, and adequate scene texture, are equally important for robust feature detection. Adverse imaging conditions poor illumination, specular surface reflections, or repetitive textures can lead to matching ambiguities and reduced reconstruction accuracy (Wolf & Dewitt, 2014.). Uniform natural or diffuse illumination is preferable where field conditions permit.

In civil engineering practice, images may be collected using handheld cameras, Unmanned Aerial Vehicles (UAVs), or mobile devices, with the choice of acquisition platform determined by site accessibility, project scale, and accuracy requirements (Remondino & Nex, 2014). For architecturally complex structures or large-scale infrastructure, UAVs enable efficient coverage of otherwise inaccessible surfaces. Structured acquisition strategies such as multi-altitude grid patterns for aerial surveys or circular orbital passes around discrete objects have been shown to produce more complete and geometrically stable reconstructions (James M. &., 2012).

Tableau 2:Step-by-step SfM processing pipeline from image acquisition to BIM export (Ours)

①	**Image Acquisition** Multi-angle photography around the object (overlap ≥ 60–80%). Equipment: DSLR, mirrorless, or smartphone.
②	**Feature Detection & Matching** SIFT/SURF algorithms detect key points in each image. Correspondences are established between overlapping images.
③	**Camera Calibration (Bundle Adjustment)** Intrinsic (focal length, distortion) and extrinsic (position, orientation) camera parameters are estimated simultaneously.
④	**Sparse Point Cloud Generation** Initial 3D point cloud from matched key points. Camera positions are reconstructed in 3D space.
⑤	**Dense Point Cloud Generation** Multi-View Stereo (MVS) densifies the sparse cloud to millions of 3D points.
⑥	**Mesh Reconstruction & Texturing** Poisson surface reconstruction creates a 3D mesh. Photographic textures are projected onto the mesh.
⑦	**Export & BIM Integration** Model exported as OBJ / IFC / PLY. Imported into Revit, ArchiCAD, or other BIM platforms for Scan-to-BIM.

Sources : (James & Robson, 2012); (Fathi & Dai, 2021); (Albeaino, Eiris, & Gheisari, 2024)

II.4.2. Processing Software:

A broad spectrum of SfM processing software is available, ranging from commercially licensed platforms to open-source research implementations, each offering different capabilities with respect to automation, accuracy, and computational efficiency (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012).

Among commercial solutions, Agi soft Met shape and Reality Capture are widely adopted in civil engineering and geomatics practice. Both platforms offer intuitive workflows encompassing the full SfM–MVS pipeline, support for Ground Control Points (GCPs), and export to standard engineering data formats (Schönberger & Frahm, 2016). Their graphical interfaces facilitate use by practitioners without specialized photogrammetric training.

In the academic and research domain, COLMAP has emerged as a leading open-source framework, providing rigorous implementations of both incremental and global SfM alongside high-quality MVS densification (Schönberger & Frahm, 2016). Its command-line interface supports batch processing and integration into custom research pipelines. The selection of processing software for any given project should be guided by project-specific requirements regarding accuracy, automation, hardware availability, and practitioner expertise (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012).

II.4.3. Three-Dimensional Model Generation:

The terminal stage of the SfM workflow involves the derivation of deployable three-dimensional products from the reconstructed point cloud data. The dense point cloud itself may be used directly for geometric measurement, surface inspection, and deviation analysis relative to design intent (Fathi & Dai, 2015). Subsequent mesh generation connects the point cloud with a triangulated polygonal surface, producing a continuous geometric representation suitable for visualization, structural surface analysis, and reverse engineering (Fathi & Dai, 2015).

Texture mapping enriches the mesh model by projecting photographic color information onto the surface geometry, yielding photorealistic three-dimensional representations widely used in architectural visualization, heritage documentation, and BIM workflows, (Remondino & Nex, 2014). The final products are typically exported in standard interoperable formats including LAS/LAZ (point clouds), OBJ, PLY, and FBX (meshes) enabling direct integration with CAD, GIS, and BIM platforms (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012).

II.5. Integration of SfM Results with BIM:

The integration of Structure from Motion outputs into Building Information Modeling (BIM) environments represents a critical component of comprehensive digital workflows for the construction, operation, and maintenance of built assets (Volk & Stengel, 2014). By bridging photogrammetric reconstruction with structured semantic models, engineers can improve documentation accuracy, support construction progress monitoring, and enhance asset lifecycle management (Remondino & Nex, 2014).

SfM-derived products including dense point clouds and textured mesh models serve as the primary input data for Scan-to-BIM workflows. These datasets provide a detailed geometric record of existing conditions, constituting an essential reference for renovation, rehabilitation, and as-built verification projects (Remondino & Nex, 2014). Nevertheless, effective BIM integration requires careful consideration of metric accuracy, level of detail, data format compatibility, and the semantic enrichment of the geometric model (Volk & Stengel, 2014).

II.5.1. Metric Accuracy:

Metric accuracy quantifies the degree of correspondence between measurements derived from an SfM model and the true physical dimensions of the surveyed object. In civil engineering applications, this parameter is of direct relevance to design verification, deformation monitoring, and quantity estimation workflows (James M. &., 2012).

The accuracy of SfM models is a function of several interacting variables: image resolution, acquisition network geometry, camera calibration quality, and the presence of Ground Control Points (GCPs) providing geodetically accurate positional constraints (Hartley & Zisserman, 2003.). Under well-controlled acquisition conditions, SfM-based reconstruction has been demonstrated to achieve centimetric accuracy in civil engineering surveys a level of precision generally adequate for construction monitoring and as-built documentation (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012). In large-scale projects or environments with limited surface texture, accuracy may degrade if rigorous image acquisition protocols are not observed (James M. &., 2012).

II.5.2. Level of Detail and Model Usability:

The Level of Detail (LOD) of an SfM-derived model is a determining factor in its usability within BIM environments. LOD encompasses both the geometric resolution and the semantic content of the model, collectively governing its applicability to engineering analysis, facility management, and quantity take-off (Volk & Stengel, 2014). SfM models typically exhibit strong geometric fidelity but limited semantic annotation; consequently, additional processing steps including geometric segmentation and object classification are required before the model can be fully leveraged within a BIM framework (Bosché, 2015).

Research indicates that SfM data can generally support BIM development to LOD 300 in standard civil engineering applications, providing sufficient detail for construction coordination, visualization, and preliminary quantity estimation (Bosché, 2015). Attaining higher LOD levels typically necessitates manual modeling procedures or the fusion of SfM data with higher-precision laser scanning datasets (Volk & Stengel, 2014).

II.5.3. Export Formats and BIM Integration:

SfM processing platforms support a range of export formats enabling interoperability with BIM and CAD environments. Point cloud formats such as LAS/LAZ and E57 are widely used for Scan-to-BIM applications by virtue of their capacity to preserve high-density geometric information, while mesh formats including OBJ and FBX are preferred for visualization and surface-based modeling tasks (Remondino & Nex, 2014). Format selection is governed by the intended downstream application and the compatibility requirements of the target BIM platform.

The Scan-to-BIM conversion process encompasses several intermediate operations, including noise filtering, point cloud registration and cleaning, geometric segmentation, and topological simplification (Volk & Stengel, 2014). The extraction of semantically meaningful BIM components walls, columns, slabs, and structural elements from raw point cloud data remains a computationally and conceptually demanding task, representing an active area of research in automated BIM generation (Remondino & Nex, 2014).

II.5.4. Accuracy Assessment in SfM Models:

Rigorous accuracy assessment is an indispensable component of any SfM deployment in civil engineering, as the geometric reliability of BIM models derived from photogrammetric data directly impacts downstream decision-making. Assessment procedures typically involve the use of independent check points, residual analysis, and comparison with independent reference measurements of known precision (James M. &., 2012).

Image resolution is a primary determinant of feature detection and matching quality, with low-resolution imagery leading to weakened key point extraction and elevated reprojection error (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012). Image overlap and camera network geometry govern the robustness of triangulation and depth estimation; poorly designed acquisition configurations produce noisy point clouds with degraded metric accuracy (Hartley & Zisserman, 2003.). Ground Sampling Distance (GSD) provides a practical measure of the maximum spatial precision theoretically achievable by the survey, with smaller GSD values corresponding to higher potential accuracy, provided that other error sources are adequately controlled (Remondino & Nex, 2014). Camera calibration stability is a further critical variable, as systematic calibration errors and parameter correlations particularly in self-calibrating systems introduce systematic geometric distortions that are especially pronounced in large-area surveys (Triggs, mclauchlan, hartley, & Fitzibbon, 2000).

II.5.5. Comparison Between SfM and LiDAR in Civil Engineering:

Structure from Motion and LiDAR constitute the two principal technologies for three-dimensional data acquisition in contemporary civil engineering practice. Each technology presents distinct advantages and limitations, and the selection between them is informed by project-specific requirements pertaining to accuracy, budget, site accessibility, and coverage area (Remondino & Nex, 2014).

SfM offers a significantly more cost-effective solution, requiring only conventional cameras, UAV platforms, or consumer-grade smartphones, in contrast to the specialized and expensive laser scanning hardware required for LiDAR surveys (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012). LiDAR, by contrast, provides superior metric accuracy and performs reliably on geometrically smooth, textureless surfaces where SfM matching is prone to failure, since its active illumination principle eliminates dependency on passive scene texture and ambient lighting conditions (James M. &., 2012).

In practice, hybrid workflows that combine SfM and LiDAR data are increasingly adopted to leverage the complementary strengths of both modalities exploiting the low cost and flexibility of SfM alongside the high accuracy and robustness of LiDAR in a unified reconstruction pipeline (Remondino & Nex, 2014).

Tableau 3:Comparative overview of SfM and LiDAR technologies for BIM integration

Criterion	SfM (3DF Zephyr)	LiDAR Terrestrial	LiDAR UAV	Advantage
Cost	Low (Camera + Software)	Very High (> $50,000)	High (> $20,000)	SfM ✓
Geometric Accuracy	5–30 mm (GCP-dependent)	2–5 mm	50–150 mm	LiDAR Terr. ✓
Processing Speed	Moderate (hours)	Fast (minutes)	Moderate	LiDAR ✓
Point Cloud Density	High (texture-rich)	Very High	Moderate	SfM / LiDAR ≈
BIM Export (IFC/OBJ)	✓ Direct	✓ Direct	✓ Direct	Equal
Reflective Surfaces	Limited	Good	Moderate	LiDAR ✓
Ease of Use	High	Moderate	Moderate	SfM ✓
Portability	High (smartphone)	Low (heavy equipment)	Moderate	SfM ✓
Scan-to-BIM Suitability	Good	Excellent	Good	LiDAR Terr. ✓

Sources: (Blaszczak-Bak, Masiero, Bąk, & Kuderko, 2024); (Albeaino, Eiris, & Gheisari, 2024); (Barazzetti, 2016))

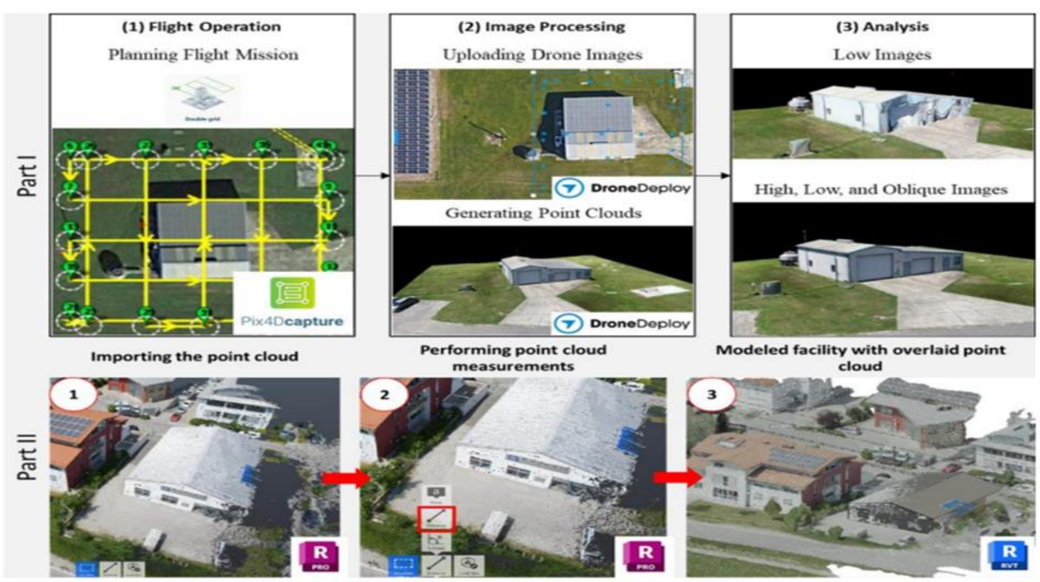

Figure 25:UAV-based 3D facility modeling workflow (Albeaino, Eiris, & Gheisari, 2024)

II.6. Practical Experiment:

This experiment explored the use of Structure from Motion (SfM) for creating 3D models from photographs using 3DF Zephyr. The goal was to evaluate how effective SfM is for documenting university construction and architectural environments. Two locations were selected for testing: an active construction site and the Development Entrepreneurship Center building. The experiment also aimed to identify the practical strengths and limitations of SfM before comparing it with newer techniques such as NeRF and Gaussian Splatting.

II.6.1. Methodology

The workflow began with capturing multiple overlapping images of the selected sites using a smartphone camera. Photos were taken from different angles and heights to ensure good scene coverage and reliable feature matching. Additional close-up images were captured for areas with complex geometry or structural details.

❖ **Device used:**

phone mobile for photos acquisition: Samsung J7pro; Samsung J6; Samsung A02

Computers: Lenovo ThinkPad Intel® Core™ i5-6300 CPU (2.40 GHz) RAM: 8Go

Graphics Card: HD 520 -128 Mo

Asus Vivobook Intel® Core ™ i9-13900H (2.60 GHz) RAM:16 Go

Graphic Card: NVDIA RTX 4050

DELL Intel® Core™ i7-10610U CPU (1.8 GHz) RAM: 8Go

Graphic Card:128MB

Figure 26: Operation of acquisition on filed (Ours)

Figure 27:photos taken on-site (university)

Figure 28: photos taken on-site (University)

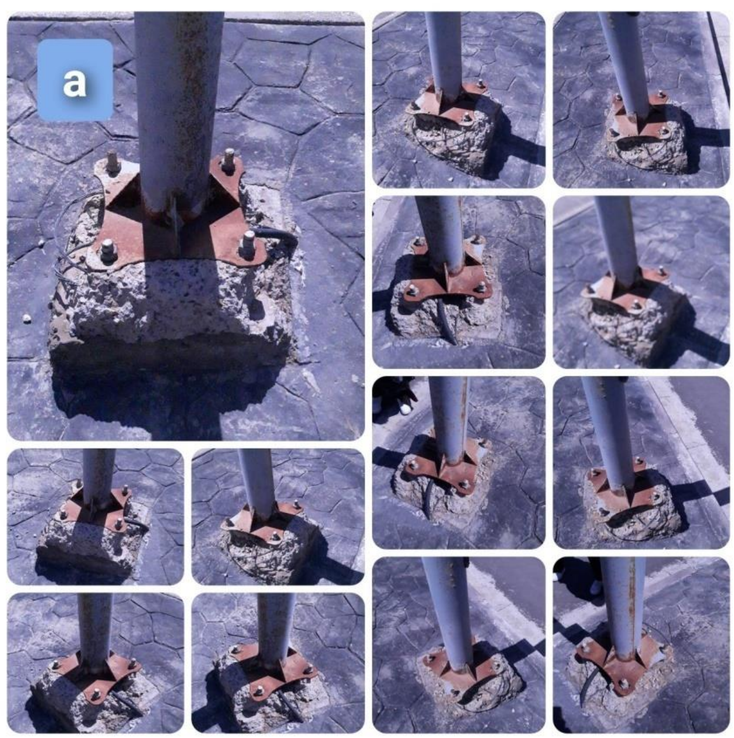

Figure 29:photos taken on-site (University)

Figure 30:photos taken on-site (University)

The images were then imported into 3DF Zephyr, where the reconstruction process followed several stages. First, automatic masking was applied to isolate the main objects and remove unwanted background elements. After that, the software performed camera alignment and generated a sparse point cloud by detecting and matching common features across the images.

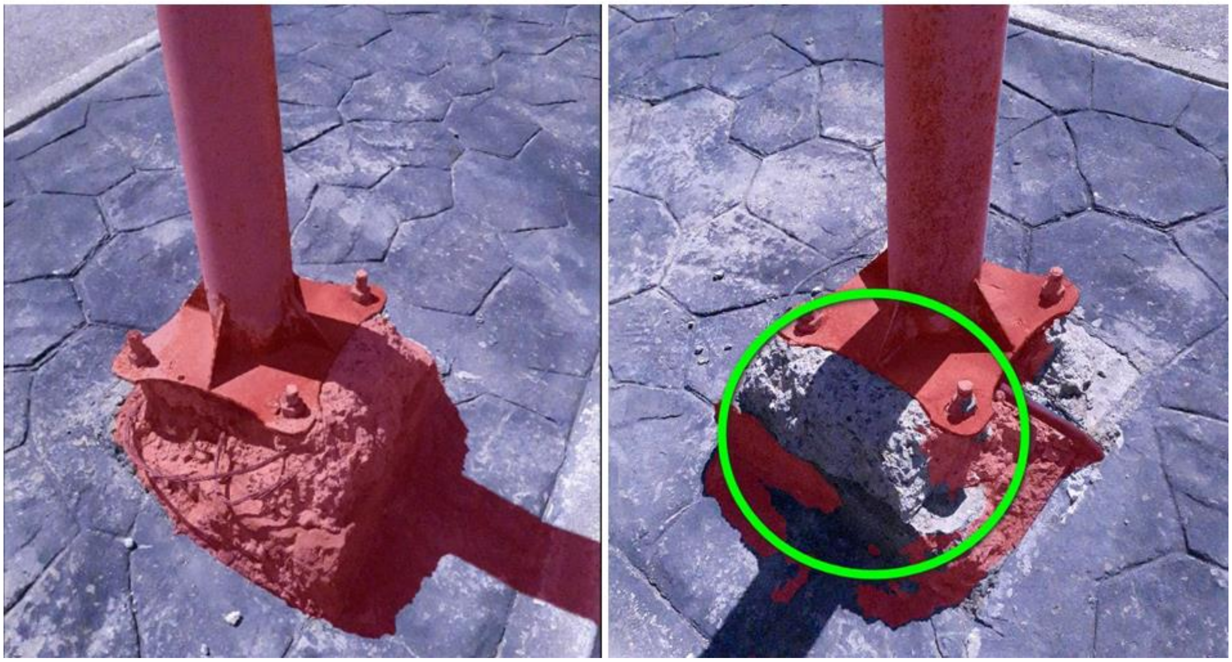

Figure 31: Anomaly of automatic object detection mask (Ours)

Next, a dense point cloud was created using multi-view stereo reconstruction. This point cloud was converted into a polygonal mesh, and finally, texture mapping was applied using the original photographs to produce a realistic 3D model.

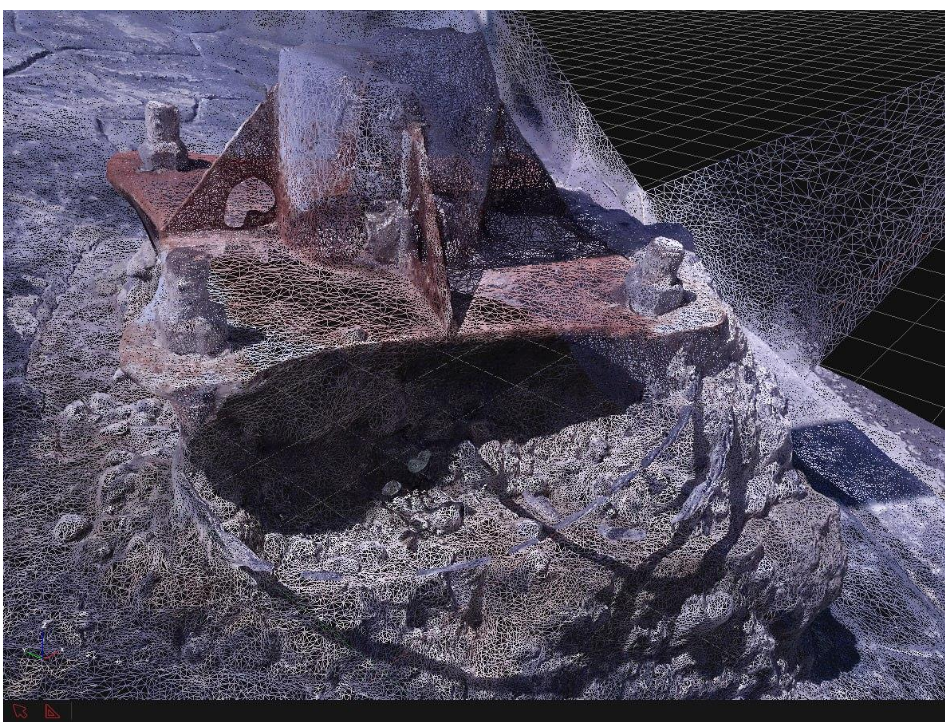

Figure 32:Generated Mesh from points cloud (Ours)

II.6.2. Results:

The reconstruction quality varied depending on the environment. The Development Entrepreneurship Center facade produced the best results because the building surfaces were relatively flat and contained enough texture for accurate feature matching. The final model successfully reconstructed walls, windows, doors, and mounted equipment with good visual quality.

In contrast, the construction site was much more difficult to reconstruct accurately. Objects such as the cement mixer, scaffolding, and overlapping construction materials created fragmented geometry and incomplete surfaces. Reflective materials like glass and metal also caused distortions because reflections changed between viewpoints, making feature matching unreliable.

Another issue was the reconstruction of flat ground surfaces, which often appeared noisy or wavy instead of perfectly flat. The manual masking process also became a major challenge because every image had to be masked individually, making the workflow slow and labor-intensive.

Figure 33: SfM results on a building with highly reflective glass surfaces (Ours)

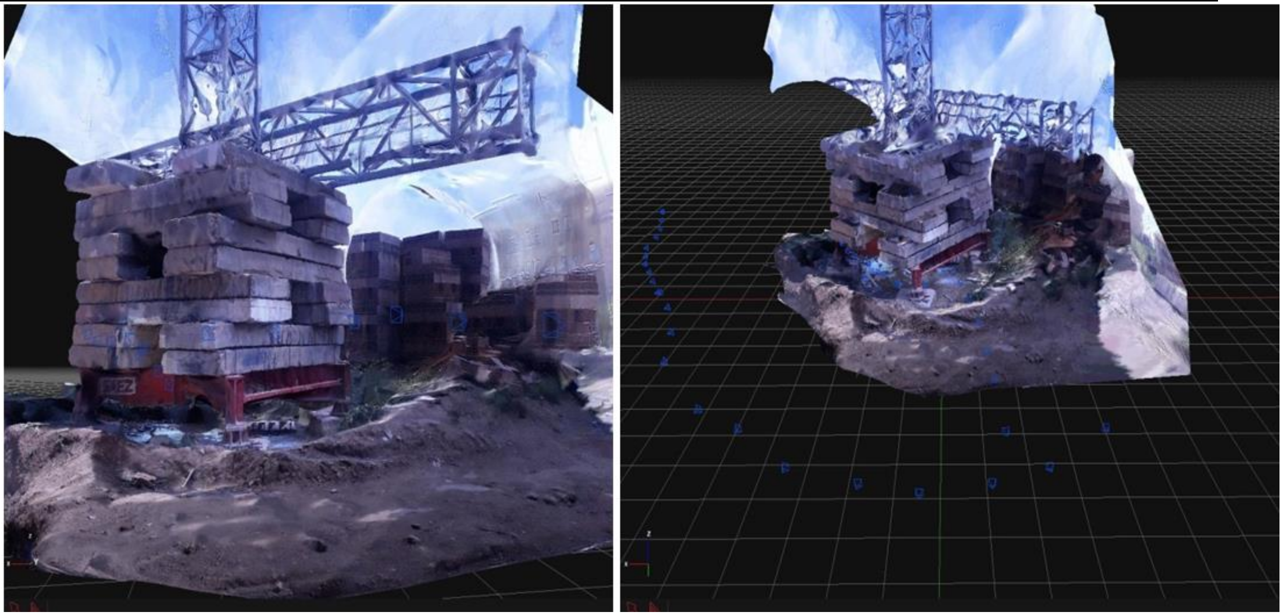

Figure 34: SFM results for a construction crane (Ours)

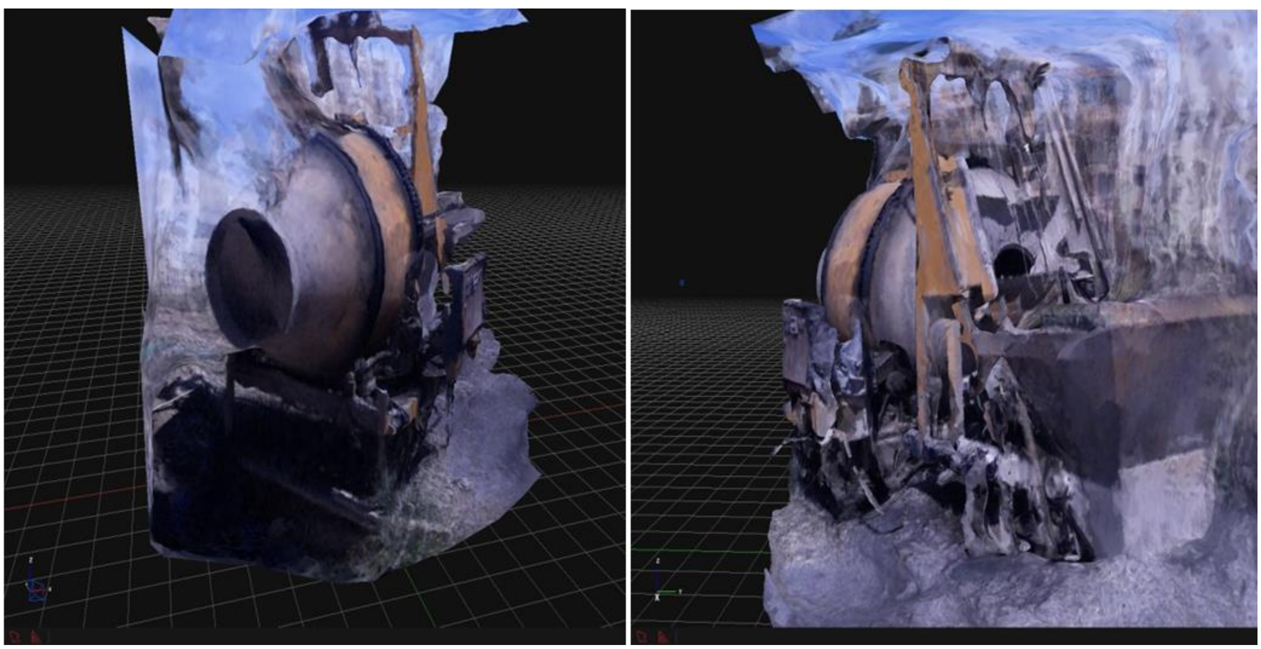

Figure 35: SFM Reconstruction of the Concrete Mixing Machine (Ours)

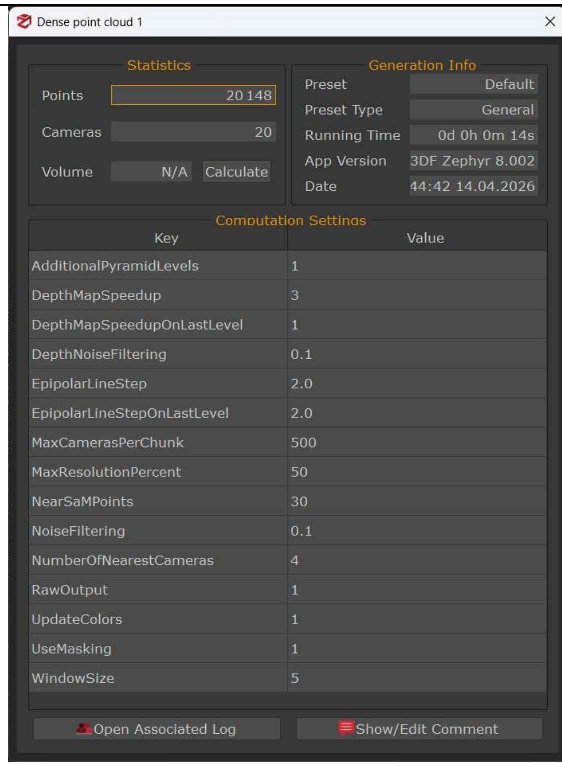

Figure 36:Dense Point Cloud Generation Settings Using 3DF Zephyr Software for the Development Entrepreneurship Center (Ours)

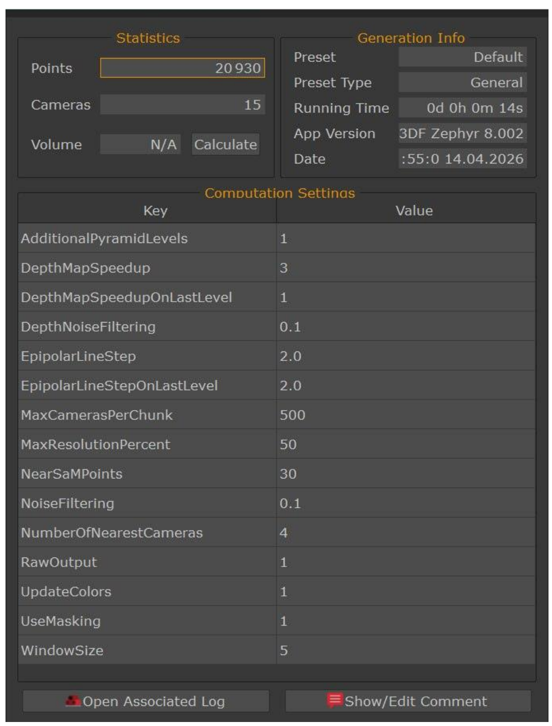

Figure 37: Dense Point Cloud Generation Settings Using 3DF Zephyr Software for the Concrete Mixing Machine (Ours)

Tableau 4:Comparison Between the center Building and the construction site for photogrammetry and BIM analysis

Criterion	Centre BUILDING	Construction site
Number of IMAGE	20	15
Number of point cloud points	20148	20930
Mesh quality	Good	Parial \ incomplete
Main challenges	Glass reflections	Moving materials and equipment
Suitability for BIM	High	Low
Surface complexity	Mostly complete	Some missing areas
Accuracy of reconstruction	High accuracy	Medium accuracy
Environmental conditions	Stable lighting	Outdoor lighting variations
Processing Difficulty	Medium	High
Recommended USE	Architectural documentation	Progress monitoring

II.6.3. Conclusion:

The experiment showed that SfM with 3DF Zephyr is capable of producing realistic and detailed 3D models using ordinary photographs. The method works particularly well for architectural surfaces with sufficient texture and consistent lighting conditions.

However, several limitations were identified, including the large amount of manual work required for masking, difficulties with reflective surfaces, incomplete reconstruction of complex industrial objects, and the lack of metric accuracy without ground control points. Processing large datasets also required significant computation time.

Overall, the experiment provided a valuable baseline for understanding the strengths and weaknesses of SfM in construction and architectural documentation. These findings will support the comparison with NeRF and Gaussian Splatting methods in the following chapters.

Chapter III : Neural Radiance Fields (NeRF) in Photogrammetry

III.1. NeRF Introduction:

Neural Radiance Fields (NeRF) represent a significant paradigm shift in the continuous three-dimensional representation of scenes, learned directly from two-dimensional images (Ben Mildenhall, 2020). The method optimizes a fully connected neural network to map a five-dimensional input comprising spatial coordinates (x, y, z) and viewing directions (θ, φ) to corresponding volumetric density and view-dependent radiance values (Ben Mildenhall, 2020). Unlike traditional explicit representations such as meshes or point clouds, NeRF encodes scenes implicitly within the weights of a multilayer perceptron (MLP), enabling the synthesis of photorealistic novel views without requiring explicit three-dimensional reconstruction (Ben Mildenhall, 2020). This implicit representation allows the network to generalize Smoothly across viewpoints, producing high-fidelity renderings from arbitrary camera positions not present in the training set (Ben Mildenhall, 2020).

The core innovation underlying NeRF resides in its use of positional encoding, which transforms raw input coordinatesintohigher-dimensionalsinusoidalrepresentationsspanning multiple frequency octaves (Ben Mildenhall, 2020). Without such encoding, the MLP is unable to represent high-frequency geometric and appearance detail Swit adequate fidelity (Ben Mildenhall, 2020). For spatial positions, sine and cosine functions are applied across up to ten frequency levels, while directional inputs receive a more compact encoding spanning four frequency levels (Ben Mildenhall, 2020). The network outputs two primary quantities: volumetric density σ, encoding local opacity at each spatial point, and RGB color c, conditioned on the viewing direction to support realistic rendering of specular reflections and view-dependent appearance effects (Ben Mildenhall, 2020).

Rendering in NeRF is grounded in classical volumetric rendering theory (Ben Mildenhall, 2020). For each pixel, a ray is cast from the camera origin in a specified direction, and sample points are drawn at discrete intervals along the ray (Ben Mildenhall, 2020). Rendering weights are computed from transmittance T_i and opacity α_i values, and the final pixel color $\mathbf{C(r) = \Sigma T_i \alpha_i c_i}$ **is** obtained as a weighted sum of sampled colors along the ray (Ben Mildenhall, 2020). NeRF employs a hierarchical two-stage sampling strategy: a coarse network samples 64 points uniformly along the ray, while a fine network concentrates 128 additional samples near regions of high density, focusing computation near scene surfaces (Ben Mildenhall, 2020). Training requires multi-view images with known camera poses, typically estimated using COLMAP structure-from-motion, and is optimized using the Adam optimizer at a learning rate of **5×10^{-4} for 100,000 to 200,000** iterations, minimizing a photometric mean squared error loss (Ben Mildenhall, 2020)

NeRF achieves state-of-the-art novel view synthesis quality on synthetic benchmarks such as Blender and real-world datasets such as LLFF, evaluated using PSNR, SSIM, and LPIPS metrics (Ben Mildenhall, 2020). Compared to multi-view stereo methods, NeRF generalizes significantly better to unseen viewpoints by learning a continuous scene function rather than reconstructing discrete geometry (Ben Mildenhall, 2020). However, the original formulation suffers from critical limitations: training a single scene requires tens of hours, and rendering is far from real-time, making deployment in interactive applications impractical (Ben Mildenhall, 2020). Furthermore, NeRF assumes static scenes captured under constant illumination, restricting its applicability to controlled capture settings (Ben Mildenhall, 2020). To address these constraints, (Martin-Brualla, 2021)proposed NeRF-W, which incorporates per-image appearance embeddings and a transient volume to handle variable illumination and moving occludes, enabling accurate novel view synthesis from unconstrained internet photo collections. (Martin-Brualla, 2021).

The computational bottleneck of NeRF was substantially addressed by (Muller, 2022) through the introduction of Instant Neural Graphics Primitives (Instant-NGP). Rather than relying on standard positional encoding, Instant-NGP augments a compact MLP with a multiresolution hash table of trainable feature vectors optimized through stochastic gradient descent (Muller, 2022). The multiresolution structure allows the network to disambiguate hash collisions, yielding a simple architecture that is trivially parallelizable on modern GPUs (Muller, 2022). By implementing the training pipeline using fully-fused CUDA kernels optimized to minimize memory bandwidth, Instant-NGP achieves training in seconds and rendering at tens of milliseconds per frame at **1920×1080** resolution improvements of several orders of magnitude over the original NeRF. (Muller, 2022)

More recently, introduced 3D Gaussian Splatting (3D-GS), an explicit scene representation method that replaces the implicit neural field with a sparse cloud of learnable 3D Gaussians optimized via differentiable rasterization (Kerbl B. , 2023). Unlike NeRF, which requires costly ray marching through a volumetric field, 3D-GS renders scenes by projecting Gaussian primitives directly onto the image plane, eliminating the per-ray sampling procedure (Kerbl B. , 2023). This approach achieves real-time rendering at over 100 frames per second at 1080p resolution, surpassing NeRF-based methods in both rendering speed and visual fidelity on multiple benchmarks (Kerbl B. , 2023). The rapid progression from NeRF to Instant-NGP to 3D-GS illustrates the swift evolution of neural scene representation research (Muller, 2022). Collectively, these methods have catalyzed applications spanning computer vision, graphics, augmented and virtual reality, cultural heritage digitization, and robotic perception (Martin-Brualla, 2021).

Tableau 5: Evolution of key NeRF variants from 2020 to 2025

Variant	Year	Key Innovation	Rendering Speed	Visual Quality	Geometry Export
NeRF (Original) (Mildenhall et al.)	2020	MLP + Volume Rendering	Very Slow (~hours)	High	None
Mip-NeRF (Barron et al.)	2021	Anti-aliasing (cone casting)	Slow	Very High	None
Mip-NeRF 360 (Barron et al.)	2022	Unbounded scenes	Slow	Excellent	None
Instant-NGP (Müller et al.)	2022	Hash encoding (GPU acceleration)	Fast (seconds)	High	Limited
MBS-NeRF (Gao et al.)	2025	Motion-blur handling (sparse images)	Moderate	High	None

(Sources: (Ben Mildenhall, 2020)*;* (Barron, et al., 2021)*,* (Barron, Mildenhall, Verbin, Srinivasan, & Hedman, 2022)*;* (Gao, Sun, Zhu, & Chen, 2025)*)*

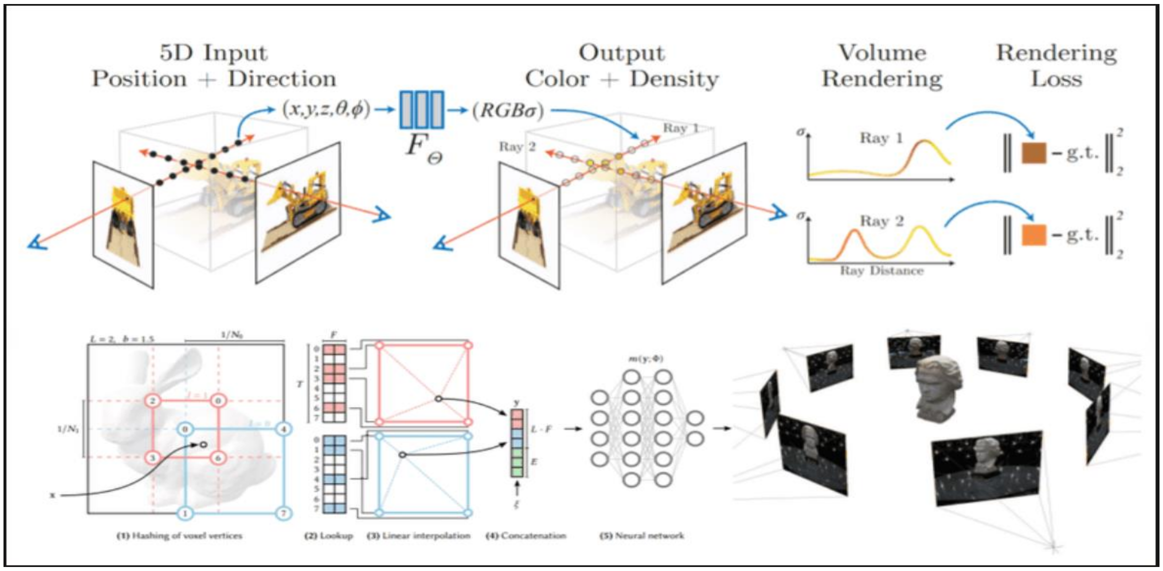

Figure 38: Neural radiance fields (NERF) for 3Dscene reconstruction (Adaloglou, 2022)

III.1.1. Shift from Geometry-Based to Neural Representation:

The transition from geometry-based to neural scene representations marks a fundamental evolution in three-dimensional modeling and rendering. (Adaloglou, 2022) Traditional geometry-based methods rely on explicit scene representations such as meshes, voxels, or point clouds. (Antonio Torralba, 2024) Polygonal meshes define surfaces through vertices, edges, and faces; voxel grids discretize space into three-dimensional occupancy or density arrays; and point clouds capture sparse surface samples, typically acquired through LiDAR or structured light scanning. (Ben Mildenhall, 2020)These approaches either require manual authoring or expensive specialized sensing hardware, and their rendering pipelines rely on rasterization for meshes or splatting for point clouds. (Antonio Torralba, 2024) Common limitations include aliasing artifacts, insufficient fine detail, and scalability constraints for complex scenes. (Ben Mildenhall, 2020)

Neural representations, exemplified by NeRF, model scenes implicitly as continuous functions parameterized by neural networks. (Ben Mildenhall, 2020)Rather than operating on discrete geometric primitives, the network takes continuous spatial coordinates as input and produces density and radiance as output. (Ben Mildenhall, 2020)This formulation eliminates quantization errors inherent to grid- or triangle-based representations. (Adaloglou, 2022) Differentiable rendering further enables scene parameters to be optimized end-to-end from image observations alone, without requiring intermediate three-dimensional reconstruction. (Antonio Torralba, 2024) Positional encodings are essential to this process, allowing networks to learn high-frequency geometric detail, while volumetric rendering integrates the continuous radiance field along camera rays to produce pixel-accurate images. (Ben Mildenhall, 2020)

This representational shift was driven by advances in deep learning applied to implicit functions. (Adaloglou, 2022) Precursors to NeRF include Occupancy Networks, which utilized signed distance functions (SDFs) for implicit surface modeling (Antonio Torralba, 2024), and SIREN, which employed periodic sinusoidal activations for signal representation. (Amazon Webservices., 2024) NeRF integrated these concepts by coupling occupancy modeling with view-dependent color prediction. (Ben Mildenhall, 2020)When explicit geometry is required post-training, it can be extracted through Marching Cubes is surface reconstruction on the density field (Adaloglou, 2022) or by sampling points where density exceeds a defined threshold. (Yuntae Jeon D. Q., 2025) This hybrid paradigm retains the photorealistic quality of neural representations while producing editable explicit geometry when needed. (Yuntae Jeon D. Q., 2025)

The principal advantages of neural representations include the ability to generate photorealistic renderings from casual photographic input without specialized scanning equipment (Ben Mildenhall, 2020), natural handling of non-rigid and amorphous object geometries (Ben Mildenhall, 2020), and more efficient scene compression compared to voxel-based storage. (Adaloglou, 2022) Conversely, per-scene training requirements and initially slow inference speeds constitute significant drawbacks. (Antonio Torralba, 2024) These limitations have motivated faster variants such as Instant-NGP, which employs multi-resolution hash grids (Ben Mildenhall, 2020), and 3D Gaussian Splatting, which rasterizes explicit primitives with learned neural opacity. (Adaloglou, 2022)

The neural representation paradigm has also enabled new capabilities in generative modeling, with diffusion models increasingly conditioning on neural radiance fields. (Ben Mildenhall, 2020)In computer graphics, neural approaches replace complex analytic shader models (Ben Mildenhall, 2020), and in computer vision, they enable dense three-dimensional reconstruction from monocular video. (Adaloglou, 2022) From the perspective of building information modeling (BIM), structured geometry can be extracted from NeRF reconstructions, though metric accuracy relative to survey-grade point clouds remains a subject of active research. (Shun Hachisuka, 2023) Geometric error metrics in real scenes currently range between 5 and 10 cm, and future work aims to unify explicit geometric representations with neural radiance modeling. (Ahmad Gholizadeh Lonbar, 2025) This paradigm shift fundamentally repositions representation from handcrafted, rule-based primitives toward learned, data-driven functions. (Ben Mildenhall, 2020)

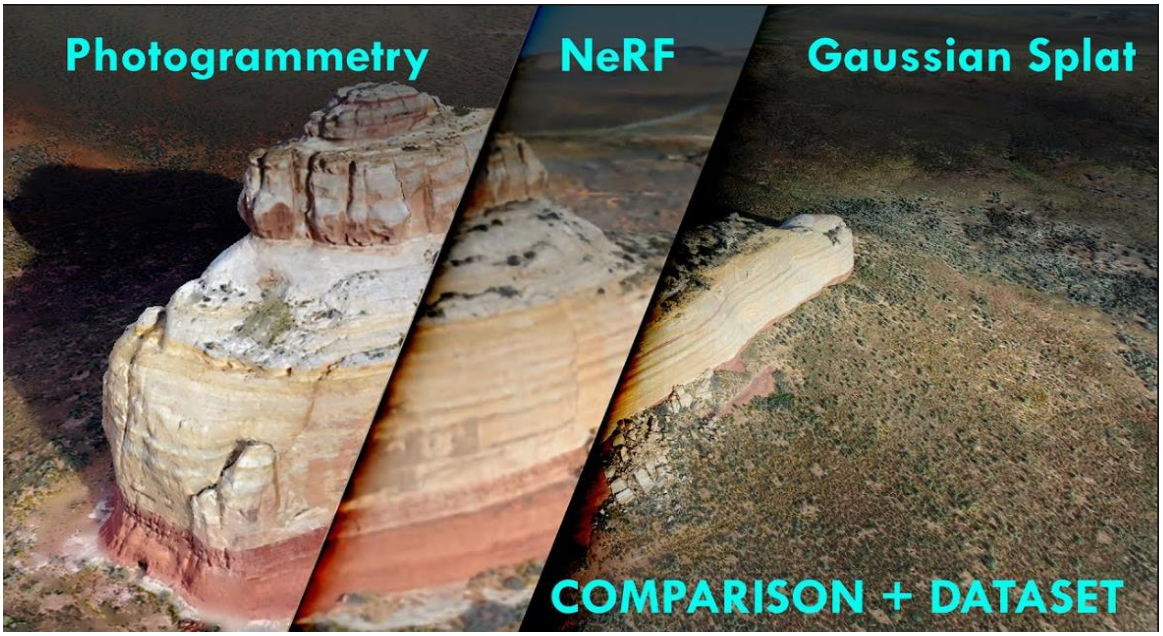

Figure39: Visual Comparison of 3Dscene reconstruction techniques; photogrammetry vs. NERF vs. Gaussian splatting (Brennan, 2023)

III.1.2. Position of NeRF within Computer Vision and Graphics:

NeRF occupies a central position at the interface of computer vision and computer graphics, offering a unified framework that advances the state of the art in both domains. (Adaloglou, 2022) Within computer vision, NeRF significantly advances novel view synthesis (NVS) beyond what is achievable with multi-view stereo (MVS) approaches, which traditionally reconstruct point clouds and meshes as intermediate representations for rendering. (Ben Mildenhall, 2020)By contrast, NeRF bypasses explicit three-dimensional reconstruction and directly optimizes a radiance field from posed images, yielding superior photorealism on established benchmarks such as DTU and Tanks & Temples. (Ben Mildenhall, 2020)

Vision applications of NeRF include SLAM-free scene reconstruction from smartphone imagery (Amazon Webservices., 2024), dense geometry learning without depth sensors (Adaloglou, 2022), and semantic scene understanding through radiance field editing. (Ben Mildenhall, 2020)In robotics, NeRF supports three-dimensional scene mapping for autonomous navigation (Ben Mildenhall, 2020), and real-time NeRF variants enable augmented reality overlays. (Adaloglou, 2022)

Within computer graphics, NeRF redefines neural rendering pipelines. (Antonio Torralba, 2024) Whereas classical graphics relied on ray tracing applied to explicit geometric models (Ben Mildenhall, 2020), NeRF introduces implicit, fully differentiable volumetric rendering that supports inverse graphics recovering scene representations directly from observations. (Ben Mildenhall, 2020)Visual effects (VFX) production leverages NeRF for matte-free compositing (Adaloglou, 2022), and game engines are beginning to integrate neural representations for procedural asset generation. (Antonio Torralba, 2024) NeRF thereby bridges the gap between rasterization and path tracing, serving as a catalyst for neural scene graph approaches. (Ben Mildenhall, 2020)

Conceptually, NeRF functions as a unified framework in which the vision pipeline maps input images to a neural three-dimensional scene representation, and the graphics pipeline renders that representation into photorealistic pixel outputs. (Ben Mildenhall, 2020) This unified approach consistently outperforms hybrid pipelines that treat reconstruction and rendering as separate stages. (Ben Mildenhall, 2020)The method has exerted substantial influence on subsequent work, including Plenoxels and TensoRF for accelerated rendering (Ben Mildenhall, 2020), and has generated over one thousand citations annually in premier venues such as CVPR and ECCV. (Adaloglou, 2022) SIGGRAPH has also adopted NeRF-derived techniques for production rendering applications. (Adaloglou, 2022)

A further contribution of NeRF lies in democratizing high-fidelity three-dimensional content creation, transitioning from a research prototype to production-ready tools such as Luma AI. (Ben Mildenhall, 2020) It challenges computer vision's traditional reliance on discrete representations and simultaneously offers computer graphics an MLP-based alternative to complex microfacet shading models. (Antonio Torralba, 2024)

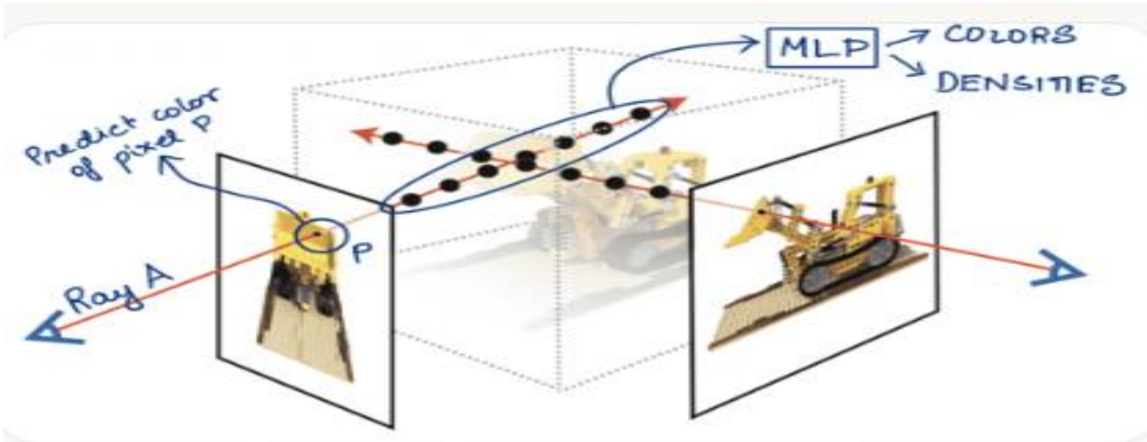

Figure 40: Visual Representation of NERF (Neural radiance field) for3D Scene Reconstruction (Roy Gosthipaty and Ritwik., 2021)

III.2. NeRF Fundamental Concepts:

Neural Radiance Fields (NeRF) are founded on several interrelated concepts that collectively enable continuous three-dimensional scene representation. (Ben Mildenhall, 2020) The core of the method is a continuous five-dimensional function F (x, d) that maps a three-dimensional spatial position **x = (x, y, z)** and a two-dimensional viewing direction **d = (θ, φ)** to volume density σ(x) and RGB radiance c (x, d). (Karaagac, 2026) Density σ quantifies local opacity independently of the viewing direction, while radiance c encodes emitted color as a function of both position and view. (Ben Mildenhall, 2020)

This function F is parameterized by a fully connected MLP neural network. (Amazon Webservices., 2024) The standard architecture comprises eight fully connected layers with 256 channels each. (Ben Mildenhall, 2020)Before being fed into the network, input coordinates undergo positional encoding to accommodate high-frequency signals. The positional encoding for spatial positions is given by**γ_p(x) = [sin(2^0πx), cos(2^0πx), ..., sin(2^9πx), cos(2^9πx)],** while the directional encoding γ_d(d) uses fewer frequencies, up to 2^3. (Karaagac, 2026) Without this encoding, the network learns only low-frequency functions, producing blurry and artifact-laden outputs. (Ben Mildenhall, 2020)

The MLP architecture is split: the initial layers process encoded position to output density σ and a 256-dimensional feature vector, which is subsequently concatenated with the directional encoding γ_d(d) for final color prediction. (Ben Mildenhall, 2020)Volume rendering integrates the learned field along camera rays **$r(t) = o + t\cdot d$.** (Amazon Webservices., 2024) For each ray, N sample points t_i are drawn within the scene bounds, with alpha **$\alpha_i = 1 - \exp(-\sigma_i\delta_i)$** and transmittance **$T_i = \exp(-\sum_{j=1}^{i-1}\sigma_j\delta_j)$.** (Karaagac, 2026) The final pixel color is computed as **$C(r) = \sum_{i=1}^{N} T_i\, \alpha_i c_i$**. (Ben Mildenhall, 2020)

Hierarchical sampling employs two MLP passes: a coarse network with 64 samples that predicts weights for concentrated fine sampling with 128 points near high-opacity surfaces. (Karaagac, 2026) This strategy improves rendering efficiency and quality over uniform sampling. (Ben Mildenhall, 2020) Training minimizes MSE over rendered pixels using the Adam optimizer with an initial learning rate of **5×10^{-4},** decaying after 250,000 iterations, for a total of approximately 200,000 iterations on a V100 GPU. (Karaagac, 2026)

The fully differentiable pipeline enables end-to-end optimization. (Ben Mildenhall, 2020) Evaluation metrics include **$PSNR = 10 \log_{10}(MAX^2/MSE)$**, SSIM for structural similarity, and LPIPS for perceptual quality assessment. (Karaagac, 2026) NeRF achieves strong performance on both the synthetic Blender dataset and the real-world LLFF dataset. (Ben Mildenhall, 2020) Overfitting is prevented through positional encoding regularization and distortion losses that penalize spurious density in empty space. (Wolfe, 2023) These fundamental concepts underpin all subsequent NeRF variants. (Ben Mildenhall, 2020)

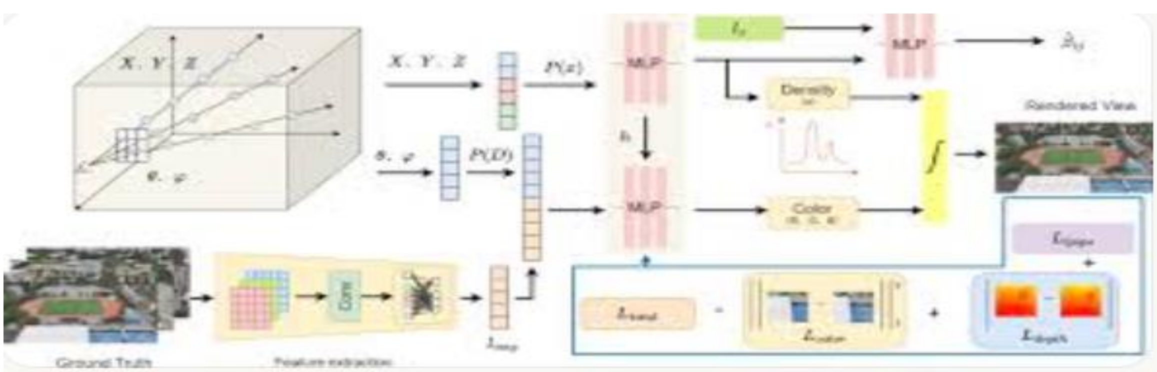

Figure 41: NERF Pipeline with Image Feature Assistance for UAV Scene Reconstruction Using Radiance fields and Volumetric Rendering (chen, 2025)

III.2.1. Radiance Fields and Volumetric Rendering:

Radiance fields constitute the mathematical foundation of NeRF, modeling scene appearance as continuous five-dimensional functions. (Antonio Torralba, 2024) Formally, a radiance field L maps three-dimensional positions (X, Y, Z) and two-dimensional viewing directions (ψ, φ) to color (r, g, b) and density σ. (Antonio Torralba, 2024) Density σ represents local opacity or transparency at each spatial point, while color values capture emitted, reflected, or transmitted light in specific viewing directions. (Karaagac, 2026) This formulation models a partial plenoptic function, describing all light rays traversing the scene. (Reed, 2012)

Radiance, defined as light power per unit area per solid angle, is a fundamental quantity in computational radiometry. (Mittal A. , 2024) In rendering, radiance fields have traditionally been used to simulate participating media such as clouds or fog. (Antonio Torralba, 2024) NeRF replaces analytic radiance field functions with a neural network parameterization, enabling data-driven scene representation. (Karaagac, 2026)

Volumetric rendering integrates the learned radiance field along camera rays to produce final pixel values. (Antonio Torralba, 2024) For a ray $\mathbf{r(t) = o + td}$, sample points t_i are drawn between near bound t_n and far bound t_f. At each sample, the field $L(t_i, d)$ is queried for σ_i and c_i, and segment length $\boldsymbol{\delta_i = t_{i+1} - t_i}$ is computed., (Ben Mildenhall, 2020) Opacity at sample i is given **by** $\boldsymbol{\alpha_i = 1 - \exp(-\sigma_i\delta_i)}$, and transmittance accumulates occlusion as $\boldsymbol{T_i = \exp(-\sum_{j<i} \sigma_j\delta_j)}$. (Karaagac, 2026) The final pixel color is $\boldsymbol{C(r) = \sum_{i=1}^{N} T_i \alpha_i c_i}$, approximating the continuous volume rendering integral as the number of samples N approaches infinity. (Antonio Torralba, 2024)

The rendering weights $\boldsymbol{w_i = T_i \alpha_i}$ form a soft probability distribution along each ray, with peaks concentrated at surfaces where density is high. (Karaagac, 2026) Novel view synthesis is achieved by tracing new camera rays and integrating the field at the desired viewpoints. (Ben Mildenhall, 2020)A key advantage of volumetric rendering is its natural handling of occlusion and semi-transparency without requiring explicit geometric intersection. (Antonio Torralba, 2024) The approach has its roots in Monte Carlo ray marching for heterogeneous participating media; NeRF makes this process fully differentiable to enable gradient-based optimization. (Ben Mildenhall, 2020)

The hierarchical sampling strategya64, point coarse pass followed by a 128, point fine pass weighted by coarse density predictions reduces variance relative to uniform stratified sampling. (Antonio Torralba, 2024) Background regions are handled through the expected termination term T-N α-N, and bounded scenes are constrained via near and far clip planes. (Karaagac, 2026) Unbounded scene variants employ proposal networks for adaptive sampling. (Amazon Webservices., 2024) Radiance fields also support relight ability by fixing density and varying illumination, enabling inverse rendering for material recovery. (Boesch, 2024)

Limitations of the vanilla volumetric rendering formulation include the assumption of isotropic scattering in the exponential transmittance model and the absence of subsurface scattering or participating media effects. (Ben Mildenhall, 2020)Extensions addressing these shortcomings include phase function modeling for anisotropic effects (Karaagac, 2026), super-sampling for anti-aliasing in NeRF-SR (Antonio Torralba, 2024), and conical frustum integration in Mip-NeRF. (Amazon Webservices., 2024) In practice, a total of 192 samples per ray is sufficient to achieve high-quality renderings, and GPU parallelization over pixel batches enables real-time performance in fast NeRF variants. (Ben Mildenhall, 2020)

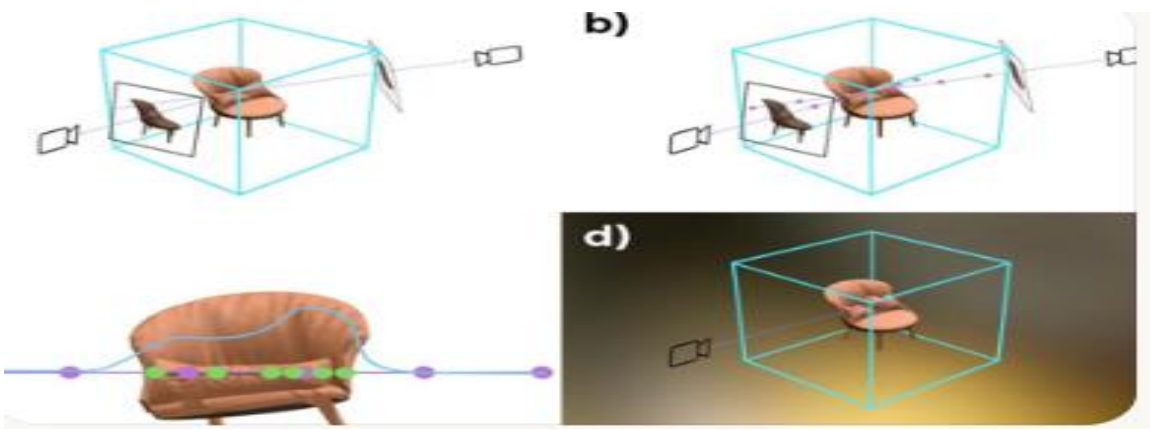

Figure 42: Multi-view Camera Setup and 3D Bounding Volume in NERF-Based Scene Reconstruction (Dellaert, 2020)

III.2.2. Neural Network Architecture:

The neural network architecture in NeRF serves as the backbone for parameterizing a continuous scene function. (Ben Mildenhall, 2020)A fully connected MLP maps five-dimensional inputs to density and radiance outputs. (Ben Mildenhall, 2020) The standard architecture consists of eight fully connected layers, each with 256 channels and ReLU activations. (Ben Mildenhall, 2020) Input coordinates are subjected to positional encoding before entering the network to preserve high-frequency information. (Karaagac, 2026)

Positional encoding for three-dimensional positions **γ_p(x)** concatenates **sin(2^k π x)** and **cos(2^k π x) for k = 0 to 9**, yielding a 60-dimensional representation **(30 frequencies × 2 functions)**. (Ben Mildenhall, 2020) Directional inputs receive a smaller encoding **γ_d(d) for k = 0to 3**, producing a 24-dimensional representation. (Ben Mildenhall, 2020) The encoded position enters the first five MLP layers, which output a 256-dimensional feature vector and the density σ. Density is computed via a softplus activation on a single output neuron: **σ = ReLU(f_σ(γ_p(x)))**. (Karaagac, 2026) The 256-dimensional feature vector is subsequently concatenated with the encoded direction γ_d(d) to form a 280-dimensional input to the final four layers, which decode view-dependent RGB radiance c via a three-neuron output with sigmoid activation. (Ben Mildenhall, 2020)

The network contains approximately five million parameters, fitting comfortably within modern GPU memory. (Ben Mildenhall, 2020)ReLU activations promote sparse gradients that aid optimization of high-capacity networks. (Karaagac, 2026) The split architecture coarse layers for density and features, direction-aware layers for color supports hierarchical sampling, with separate coarse and fine MLP instances. (Ben Mildenhall, 2020)The design is a pure fully connected network without convolutional or recurrent components, well-suited to coordinate-based inputs. (Ben Mildenhall, 2020) Positional encoding acts as a Fourier feature mapping, which has been shown to be critical for MLP expressivity in implicit neural representations. (Ben Mildenhall, 2020) Without this encoding, NeRF produces blurry, low-frequency artifacts. (Karaagac, 2026) SIREN variants later addressed this through sinusoidal layer activations. (Amazon Webservices., 2024)

Batch normalization is deliberately omitted to preserve spatial specificity in density predictions. (Ben Mildenhall, 2020)Weight initialization follows the Xavier/Glorot scheme for stable training. (Wolfe, 2023) The forward pass per sample takes microseconds, with computational bottlenecks attributable to volumetric ray sampling. (Ben Mildenhall, 2020) End-to-end differentiability flows from raw pixels through the rendering equation to the network weights. (Ben Mildenhall, 2020) The Adam optimizer with an initial learning rate of **5×10^{-4}**, decaying geometrically after 250,000 iterations, is used for weight updates. (Karaagac,

2026) Over parameterization is beneficial for generalization despite per-scene fitting. (Ben Mildenhall, 2020)

Subsequent variants have substantially modified this base architecture to improve efficiency: Instant-NGP replaces the large MLP with a compact multilinear hash table structure (Adaloglou, 2022); TensoRF applies low-rank CP tensor factorization (Amazon Webservices., 2024); and Plenoxels abandon the MLP entirely in favor of learned voxel embeddings. (Ben Mildenhall, 2020)The original MLP formulation prioritizes compact, continuous representation over inference speed, and scales to unbounded scenes through coordinate contraction mappings. (Karaagac, 2026) Extensions add a temporal input dimension for dynamic scene modeling (Wolfe, 2023), auxiliary segmentation heads for semantic predictions (Tancik, et al., 2023), and post-training compression through pruning or quantization. (Karaagac, 2026)

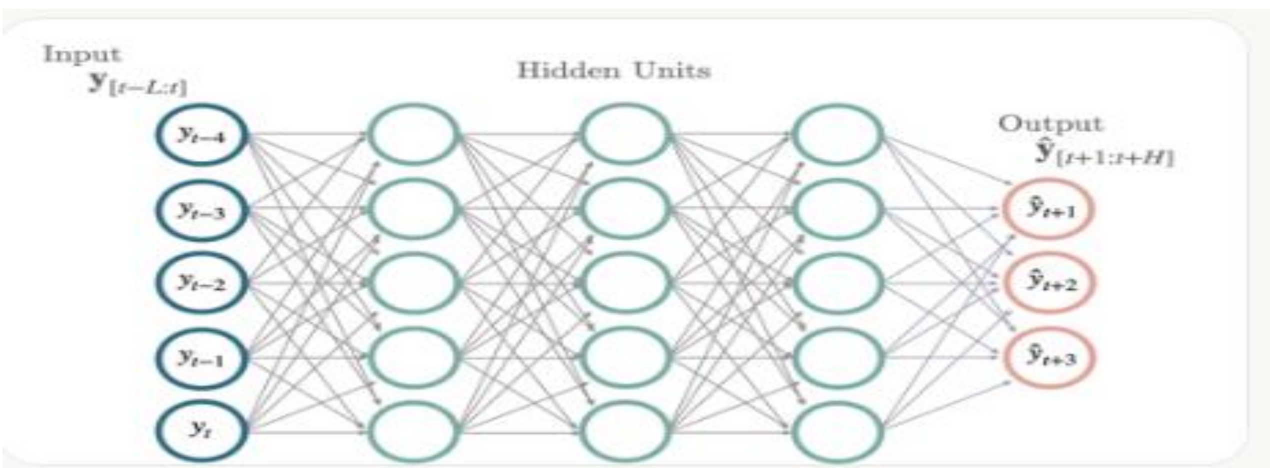

Figure 43:Architecture of a Multi-layer perceptron (MLP) for Time Series Forecasting (nixtla, .)

III.2.3. Training Process and Data Requirements:

The training process for NeRF optimizes a neural radiance field to reproduce input images through differentiable volumetric rendering. (Ben Mildenhall, 2020)Training requires a set of posed images capturing the target scene from multiple viewpoints., (Ben Mildenhall, 2020) Typically, 100 to 200 images with overlapping coverage are sufficient for complex scenes. (Karaagac, 2026) Camera poses comprising both position and orientation must be known accurately for each image. (Ben Mildenhall, 2020)

Poses are most commonly estimated using Structure-from-Motion (SfM) tools such as

COLMAP, which processes the input images to recover sparse three-dimensional points alongside camera extrinsic and intrinsics. (Ben Mildenhall, 2020) Real-world datasets such as LLFF use forward-facing captures from handheld cameras, while synthetic datasets such as Blender provide ground-truth poses and multi-view spherical coverage. (Ben Mildenhall, 2020) Importantly, no explicit depth or geometry information is required; only RGB pixel values and camera poses are needed as input. (Amazon Webservices., 2024)

Training minimizes mean squared error (MSE) between rendered and ground-truth pixel colors: $\mathbf{L = \sum\text{-}\{r \in R\}||C(r) - \hat{C}(r)||^2}$ over all sampled rays. (Ben Mildenhall, 2020) Rays are randomly sampled from training images first pixel locations, then stratified sample points along each ray. (Wolfe, 2023) A batch of 1,024 rays, each with 192 samples **(64 coarse + 128 fine)**, is processed per iteration. (Ben Mildenhall, 2020) The forward pass queries the MLP at all sample points, computes transmittance T_i and alpha weights α_i, and renders color $\hat{C}(r)$, whose loss value is backpropagated through the differentiable rendering equation to update the MLP weights. (Ben Mildenhall, 2020)

The Adam optimizer is used with $\boldsymbol{\beta_1 = 0.9, \beta_2 = 0.999}$, and an initial learning rate of $\mathbf{5\times10^{-4}}$. (Ben Mildenhall, 2020) Learning rates decay geometrically after 250,000 iterations for the coarse network and after 250,000 iterations for the fine network. (Ben Mildenhall, 2020) Total training spans 100,000 to 300,000 iterations, requiring one to two days on a V100 GPU. (Karaagac, 2026) No data augmentation is employed; dense ray sampling provides sufficient coverage. (Ben Mildenhall, 2020) Regularization strategies include distortion loss on ray point spatial distributions to prevent floaters, and total variation (TV) loss on density gradients to enforce spatial smoothness. (Ben Mildenhall, 2020)

Since NeRF performs per-scene optimization, a new model must be trained for each distinct dataset. (Ben Mildenhall, 2020) Data quality requirements include substantial inter-view overlap (greater than 60° between adjacent views) to avoid under constrained optimization. (Karaagac, 2026) Preprocessing involves cropping image borders, down sampling to a maximum of 800 pixels on the longer side, and setting scene-appropriate near and far clip planes. (Ben Mildenhall, 2020)Forward-facing scenes require coordinate contraction to the unit sphere, while 360° captures avoid pose ambiguities more effectively. (Karaagac, 2026) Rendering quality scales with the number of input views: approximately 20 views yield blurry results, whereas 100 or more views achieve photorealistic quality. (GUO, 2025) No labels beyond images and poses are required, making NeRF an effectively unsupervised three-dimensional learning method. (Karaagac, 2026) The dataset is typically split with 80% of rays used for training and 20% for validation. (Ben Mildenhall, 2020)Training is considered converged when the validation PSNR stabilizes in the range of 30 to 35 dB. (Wolfe, 2023) GPU memory requirements are approximately 5 GB at full resolution, with ray chunking employed to avoid out-of-memory errors. (Karaagac, 2026)

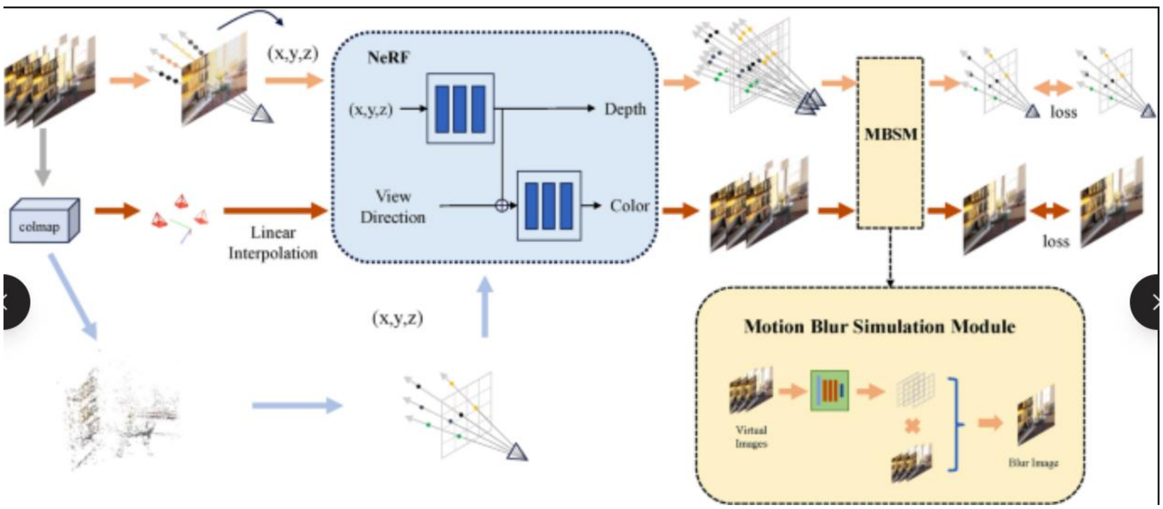

Figure 44: Overall architecture of MBS-NERF framework for reconstructing sharp neural radiance fields motion-blurred sparse (Gao, Sun, Zhu, & Chen, 2025)

III.3. NeRF Workflow for 3D Reconstruction:

The NeRF workflow for three-dimensional reconstruction transforms multi-view image collections into implicit neural scene models through a structured optimization pipeline. (Ben Mildenhall, 2020)The process begins with data capture: a minimum of 100 RGB images with overlapping viewpoints are collected around the target scene. (Ben Mildenhall, 2020) Images should provide 360° or forward-facing path coverage with inter-view angles exceeding 60° to ensure adequate scene constraint. (Karaagac, 2026)

Camera poses are subsequently estimated using COLMAP Structure-from-Motion on the image set, yielding intrinsic parameters (focal length, distortion coefficients), extrinsic parameters (rotation matrix R and translation vector t), and a sparse three-dimensional point cloud. (Ben Mildenhall, 2020) Preprocessing steps include cropping black borders, down sampling images to a maximum side length of 800 pixels, and establishing appropriate near and far clip planes based on the COLMAP depth distribution. (Ben Mildenhall, 2020) The dataset is partitioned such that one in every eight images is reserved for validation, with 1,024 rays sampled per training batch. (Karaagac, 2026)

The NeRF model is initialized with two MLP instances coarse and fine equipped with positional encodings for the five-dimensional input. (Ben Mildenhall, 2020) During training, pixels are randomly sampled from training images and their corresponding rays are processed through hierarchical volumetric rendering: 64 coarse samples are drawn uniformly along each ray, followed by 128 fine samples concentrated near high-density regions, as guided by the coarse density predictions. (Ben Mildenhall, 2020)Rendered pixel colors are computed, and the MSE loss drives backpropagation through both MLP instances. (Ben Mildenhall, 2020) The Adam optimizer updates network weights over 100,000 to 300,000 iterations, requiring approximately 24 hours on a V100 GPU, until validation PSNR converges above 30 dB. (Karaagac, 2026)

Upon convergence, three-dimensional geometry can be extracted in several ways. A dense point cloud is obtained by sampling rays on a spatial grid and retaining points where σ exceeds a threshold of 10. (Korycki, 2024) Mesh extraction is performed via Marching Cubes applied to the density field sampled on a 512^3 grid. (Dominik Zimny, 2024) The reconstruction is aligned to metric scale using sparse COLMAP points or LiDAR ground truth, refined through Iterative Closest Point (ICP) registration with voxel down sampling. (Muhammad Arbab Arshad, 2024) In cases where COLMAP fails on textureless regions, pose refinement through bundle adjustment is recommended. (Jiang L. C., 2024)

Modern tools such as Nerf studio provide real-time visualization during training (Tancik, et al., 2023), and reconstruction outputs can be exported as PLY point clouds or OBJ meshes for import into CAD and BIM workflows. (Shun Hachisuka, 2023) The workflow scales effectively with contemporary frameworks including Nerf studio and Instant-NGP. (Muhammad Arbab Arshad, 2024) For unbounded scenes, scene contraction to the unit sphere is applied (Ben Mildenhall, 2020), dynamic scenes are handled by appending a normalized time dimension to the MLP input (Zhihong Chen, 2025), and depth supervision from monocular depth estimation networks such as MiDaS improves geometric quality. (Jiang L.

C., 2024) Few-shot variants leveraging meta-learned priors extend the workflow to settings with Sparse input views. (XU, 2024)

III.3.1. Dataset Preparation:

Dataset preparation constitutes the critical foundation for successful NeRF training. (Ben Mildenhall, 2020)The acquisition phase requires capturing 100 to 200 RGB images of the target scene from densely overlapping viewpoints, ideally with full 360° coverage or structured paths as employed in LLFF forward-facing datasets. (Ben Mildenhall, 2020) Consistent exposure and white balance settings must be maintained across all images to prevent appearance shifts in the learned radiance field. (Wolfe, 2023) The use of a tripod or motorized turntable is recommended to minimize motion blur. (Regueira-gomez, 2025) Inter-view angles should exceed 60° to adequately constrain the implicit scene function. (Ben Mildenhall, 2020)

ArUco markers or textured backgrounds may be included in the scene to facilitate robust COLMAP feature matching. (Regueira-gomez, 2025) Camera settings should be configured with fixed focus and aperture priority mode for consistent depth of field. (Marvik, 2023) Specular highlights should be minimized wherever possible, as strong reflections can disrupt radiance field estimation. (Ben Mildenhall, 2020)

Following acquisition, images are organized in a single directory with sequential file naming. (Ben Mildenhall, 2020) COLMAP SfM is then executed to recover camera poses, intrinsic parameters, and a sparse three-dimensional point cloud, with the reconstruction stored in the standard cameras.txt, images.txt, and points3D.txt output files. (Ben Mildenhall, 2020) Pose quality is verified by checking that reprojection errors are below one pixel. (Karaagac, 2026) Preprocessing steps include cropping image borders using the bounding box defined by COLMAP points, down sampling images to a maximum height of 800 pixels, and setting near and far clip planes based on the observed depth range typically 0.8 m to 4.0 m for room-scale scenes. (Ben Mildenhall, 2020)

For synthetic scenes, pre-posed datasets from the Blender benchmark can be used directly. (Ben Mildenhall, 2020) Standard benchmarks include the LLFF dataset (62 real-world scenes) and the Tanks & Temples dataset. (Wolfe, 2023) Consumer applications such as Polycom or Luma AI automate pose recording during capture. (Shun Hachisuka, 2023)The dataset is split by holding out one in every eight images for validation PSNR monitoring. (Ben Mildenhall, 2020) COLMAP output is converted to the JSON transforms format using scripts such as images2poses.py. Optional foreground masking limits NeRF training to the region of interest. (Paul Schulz, 2026) Image paths must be normalized relative to transforms. son for compatibility with training frameworks. (Tancik, et al., 2023) COLMAP applies a radial distortion model to undistorted images during pose estimation. (Wolfe, 2023)

Metric scale ambiguity is resolved post hoc through known object dimensions or LiDAR tie-points. (Korycki, 2024) Dataset quality is validated by visualizing the COLMAP point cloud to confirm coverage density and distribution. (Muhammad Arbab Arshad, 2024) Common failure cases include scenes with highly reflective surfaces or near-uniform color regions, which prevent reliable feature matching. (Marvik, 2023) In such cases, COLMAP should be retried with the Image Reader. Single-camera flag to enforce fixed intrinsic. (Regueira-gomez, 2025)For 360° datasets, poses are recentered to the origin and scaled to the unit sphere. (Ben Mildenhall, 2020)Dynamic scene datasets require video frames with per-frame poses refined through bundle adjustment. (Zhihong Chen, 2025) A recommended custom dataset checklist specifies a minimum of 50 images, pose RMSE below 0.5°, and inter-view overlap exceeding 70%. (Sharma a. , 2024) Processed dataset sizes range from 1 to 5 GB depending on resolution and view count. (Marvik, 2023)

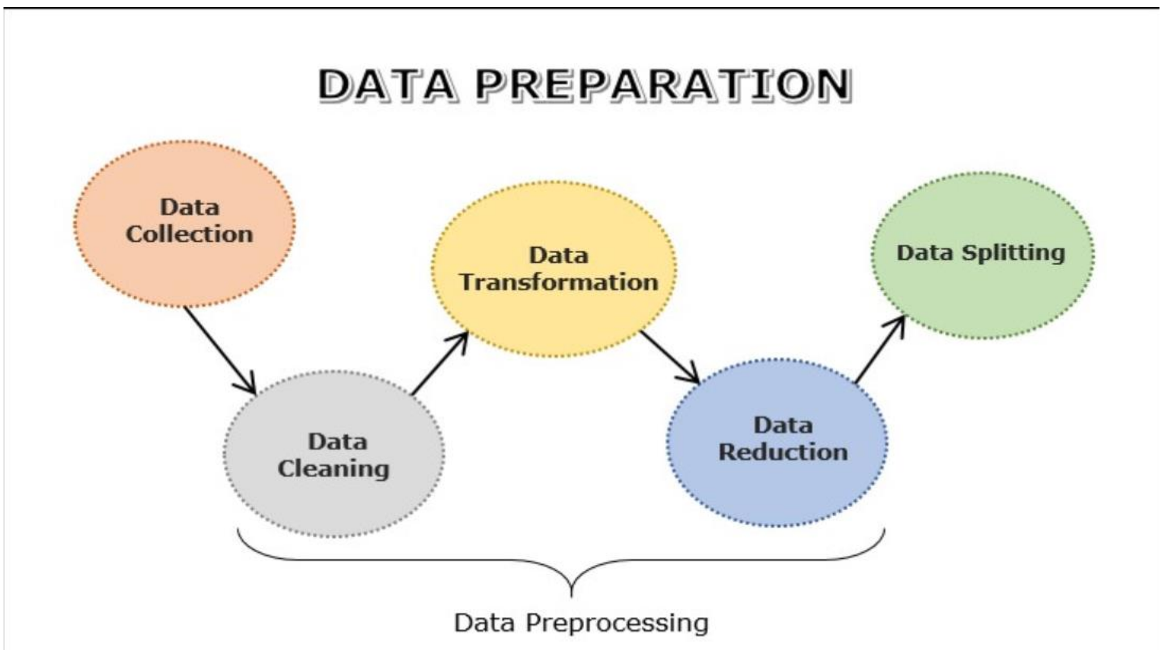

Figure 45:data preparation pipeline: from collection to splitting (Tutorials Point, 2026)

III.3.2. Training and Rendering:

Training and rendering in NeRF constitute an integrated, fully differentiable pipeline that optimizes scene appearance from multi-view images. (Ben Mildenhall, 2020) Training proceeds through random ray sampling from posed training images, typically at 1,024 rays per batch, covering one to two percent of pixels per image. (Ben Mildenhall, 2020) Each ray $\mathbf{r(t) = o + td}$ originates at camera center o in direction d, spanning the near plane t_n to the far plane t_f. (Karaagac, 2026)

Coarse sampling stratifies 64 points uniformly along each ray for initial volume integration. (Ben Mildenhall, 2020)The coarse MLP queries positional encodings $\gamma_p(t_i, d)$ to predict density σ_i and a 256-dimensional feature vector. (Wolfe, 2023) Transmittance accumulates as $\mathbf{T_i = \exp(-\sum_{j<i} \sigma_j\delta_j)}$, and rendering weights are $\mathbf{w_i = T_i (1 - \exp(-\sigma_i\delta_i))}$. (Ben Mildenhall, 2020)These coarse weights serve as a probability distribution for importance sampling of 128 fine points concentrated near surface opacity peaks. (Karaagac, 2026) The fine MLP concatenates coarse features with directional encoding $\gamma_d(d)$ to produce view-dependent RGB radiance c_i, and the final pixel color is computed as $\mathbf{\hat{C}(r) = \sum T_i \alpha_i c_i}$. (Ben Mildenhall, 2020)

The MSE loss $\mathbf{\|\hat{C}(r) - C_{gt}(r)\|^2}$ backpropagates through the exponential, softplus, and MLP layers. (Karaagac, 2026) The Adam optimizer with $\boldsymbol{\beta = (0.9, 0.999)}$ and initial $\mathbf{LR = 5\times10^{-4}}$ performs weight updates; the coarse learning rate decays after 250,000 iterations and the fine after 150,000. (Ben Mildenhall, 2020)Training converges between 200,000 and 500,000 iterations when validation PSNR stabilizes above 30 dB. (Ben Mildenhall, 2020)Distortion loss regularizes the spatial distribution of ray samples to prevent floaters, and total variation on density gradients enforces field smoothness. (Karaagac, 2026)

Rendering follows the same volumetric integration pipeline but operates on densely sampled image grids without ground truth targets. (Ben Mildenhall, 2020) To manage GPU memory during dense rendering, rays are processed in chunks of 8,192 per pass (Wolfe, 2023) Novel view synthesis at arbitrary camera poses is enabled by tracing new rays through the trained MLP, with hierarchical sampling providing a two- to three-fold speedup over naive uniform sampling. (Ben Mildenhall, 2020) Real-time rendering requires specialized variants: Instant-

NGP employs 16-bit hash table encodings (Adaloglou, 2022), Plenoxels use trainer voxel interpolation (Ben Mildenhall, 2020), and 3D Gaussian Splitting depth-sorts and alpha-composites explicit three-dimensional Gaussians. Training quality is monitored through **PSNR = 10 log(MAX²/MSE)**, SSIM for structural similarity via windowed luminance, contrast, and structure statistics, and LPIPS for human-aligned perceptual error using VGG features. (Karaagac, 2026) Overfitting manifests as high training view quality with degraded novel view generalization, and per-scene fitting requires hours to days of compute. (Wolfe, 2023) Mixed-precision FP16 training approximately halves memory requirements and provides a 1.5× speedup with minimal quality degradation. (Ben Mildenhall, 2020)Anti-aliasing is addressed in Mip-NeRF through conical frustum integration over pixel footprints. Training quality scales linearly with input views, saturating at approximately 200 images. (Ben Mildenhall, 2020)

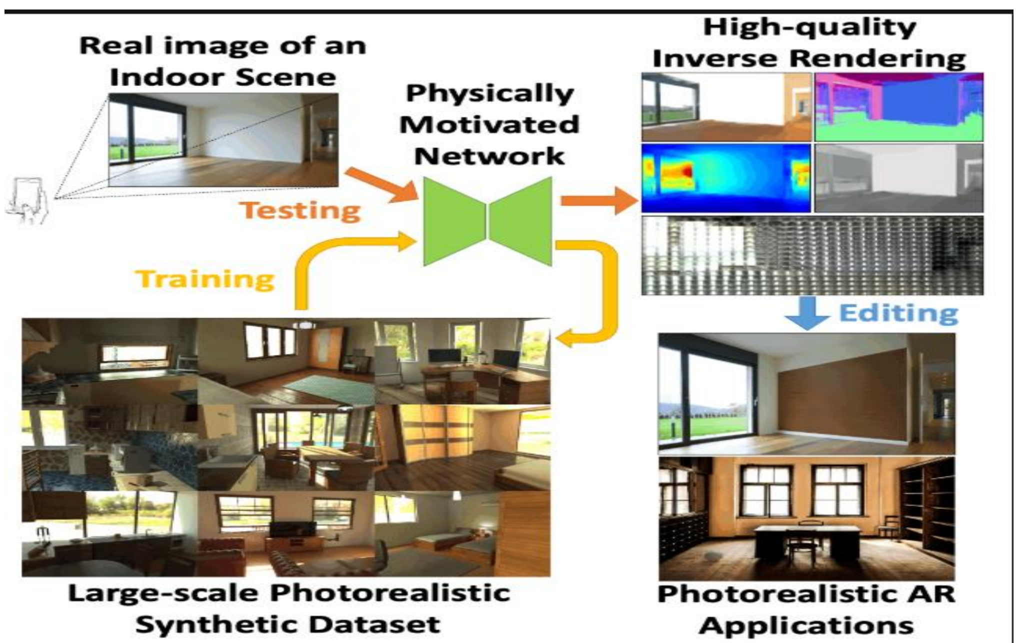

Figure 46: A Physically-Based neural network for indoor scene inverse rendering and photorealistic AR (Li, Mohammad, Ravi, Kalyan, & Manmohan, 2020)

III.3.3. Available Tools and Frameworks:

Since the original 2020 publication, a substantial ecosystem of NeRF tools and frameworks has emerged, substantially lowering the barrier to three-dimensional reconstruction from photographic input. (Tancik, et al., 2023)Nerf studio stands out as the leading modular

PyTorch framework for rapid NeRF development and deployment. (Tancik, et al., 2023) It supports modular method integration through a plug-and-play architecture, with nerfacto as the default method balancing speed and quality through recent advances in the field. (GUO, 2025)Key features include a real-time web viewer, direct data import from Polycom and Record3D, and export to PLY point clouds, OBJ meshes, and rendered video. (Tancik, et al., 2023) Installation is straightforward via pip, and training is launched with ns-train nerfacto data data/fern, while COLMAP pose estimation is automated through ns-process-data. (nerfstudio-project, 2023) More than 50 methods are integrated in Nerfstudio, spanning vanilla NeRF through Instant-NGP and 3D Gaussian Splatting. (Tancik, et al., 2023)

The original Tensor Flow 1.15 implementation by Mildenhall et al. remains available on GitHub, optimizing scenes in approximately 15 hours on a V100 GPU via python run_nerf.py config config_fern.txt. (Ben Mildenhall, 2020)TinyNeRF provides a compact PyTorch port with mobile-friendly quantization (Landup), and nerf-pytorch by yenchenlin offers a modern reimplementation of the original paper. (Sharma a. , 2024) Instant-NGP by NVIDIA achieves real-time training in approximately 15 seconds and rendering at over 1,000 fps through hash grid encodings, requiring CUDA 11.3 or later. (Zhang, et al., 2025)

For users without programming experience, Kiri Engine provides a GUI application enabling consumer-level capture-to-NeRF workflows (Team, 2025), and Luma AI processes smartphone video to textured meshes through a cloud-based NeRF pipeline. Nerfstudio's rendering and export commands ns-render for 360° video and ns-export mesh for watertight OBJ generation facilitate downstream use. Most frameworks accept either automatic COLMAP integration or a pre-computed transform.json file. (Mildenhall, 2020) Unity and Blender plugins support NeRF point cloud import into AR and VFX pipelines. (Shun Hachisuka, 2023) Benchmarking tools such as NeRFBench compare over 30 methods across PSNR and speed metrics. (Jain, Gupta, gupta, & chandraker, 2026)

Recent specialized frameworks address targeted applications: the HERO framework optimizes quantized NeRF for edge hardware through reinforcement learning (Zhang, et al., 2025); NERFIFY employs multi-agent LLMs for automated code generation from NeRF research papers (Jain, Gupta, gupta, & chandraker, 2026); LowLight-NeRF incorporates physics-based priors for low-illumination scenes (Xiaohui Wang, 2026); and AWS SageMaker enables cloud-based deployment of text-to-NeRF generative models. (Sprintzeal, 2026)OpenMVG provides an alternative SfM pipeline for COLMAP-free workflows. (Moulon, monasse, perrot, & marlet, 2016) Docker images are available for most major frameworks to ensure reproducible environments. (Tancik, et al., 2023)

III.3.4. NeRF Pipeline:

Step 1 (Posed Input Images): NeRF requires as input a set of RGB images paired with their corresponding calibrated camera poses, typically obtained via SfM. Each image provides a partial observation of the scene from a specific viewpoint, and the network must reconcile all observations into a coherent volumetric representation.

Step 2 (Ray Casting and 3D Point Sampling): For each pixel in a training image, a ray is cast from the camera center through that pixel into the scene. A discrete set of 3D sample points is placed along this ray using stratified sampling within a coarse interval and importance sampling in a finer pass (the hierarchical "coarse-to-fine" strategy of Mildenhall et al., 2020).

Step 3 (Positional Encoding): Raw 3D coordinates and viewing directions are mapped to a higher-dimensional space via a sinusoidal positional encoding function $\gamma(\cdot)$. This step is critical: without it, the MLP's inherent spectral bias prevents it from learning high-frequency geometric and appearance detail.

Step 4 (MLP Radiance Field Query): Each encoded 3D point (x, y, z) and viewing direction (θ, φ) is passed through a fully connected neural network F_θ, which predicts two quantities: the volume density σ (a view-independent measure of how opaque the point is) and the emitted radiance (RGB color), which is view-dependent to capture specular effects.

Step 5 (Volume Rendering): The predicted densities and colors along a ray are composited using the classical volume rendering equation, numerically approximated as a weighted sum. The accumulated transmittance weights each sample's color contribution, effectively implementing an alpha-compositing integral that produces a rendered pixel color $\hat{C}(r)$.

Step 6 (Photometric Loss and Optimization): The rendered color $\hat{C}(r)$ is compared against the observed ground-truth pixel color $C(r)$ using the mean squared error (MSE). Gradients are backpropagated through the differentiable rendering equation to the MLP weights θ, which are updated iteratively via the Adam optimizer until the network converges to a representation that reproduces all training views faithfully.

Inference: At test time, novel views are synthesized by specifying an arbitrary unseen camera pose, casting rays, and running the rendering pipeline no additional optimization is required. The entire scene is encoded implicitly within the network weights.

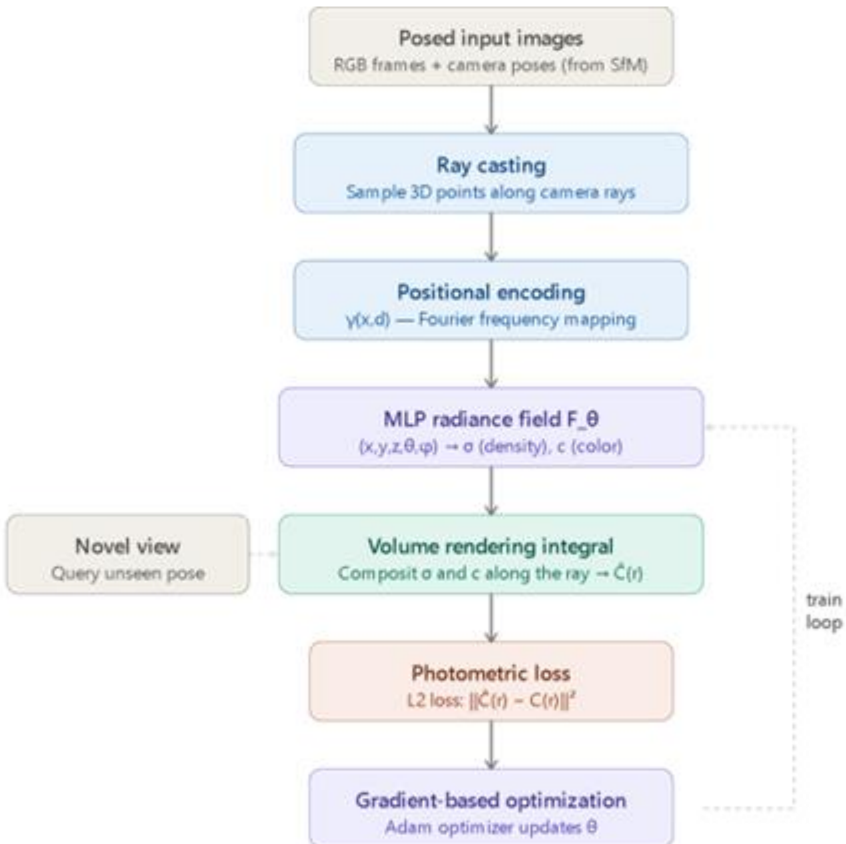

Figure 47:NeRF pipeline Organisation chart (Ours)

III.4. NeRF and BIM Integration Challenges:

The integration of NeRF with Building Information Modeling (BIM) offers the prospect of generating cost-effective as-built three-dimensional models from photographic input, but faces significant challenges related to geometric accuracy and workflow compatibility. NeRF point clouds mix high-fidelity surface points with spurious floaters in a manner fundamentally different from the consistent quality delivered by Terrestrial Laser Scanning (TLS). (Shun Hachisuka, 2023) While laser scanners yield consistent millimeter-level precision, NeRF reconstruction errors in construction environments can reach 5 to 10 cm. (Yuntae Jeon D. Q.,

2025) BIM authoring tools such as Revit and IFC importers require clean, metrically accurate geometry; raw NeRF point clouds therefore necessitate substantial post-processing and filtering.

Semantic segmentation is adversely affected by NeRF noise: Point NeXt mean intersection-over-union (mIoU) scores drop 10 to 15 percentage points compared to results obtained from clean laser scan data. (Shun Hachisuka, 2023)Mixed accuracy characteristics in NeRF clouds demand selective point extraction, retaining only points with $\sigma > 20$ while down sampling outliers. (Shun Hachisuka, 2023) Structural element detection such as identifying column grid positions and beam cross-sections is further complicated by horizontal slice jitter of 5 to 20 cm from true positions, partial occlusions, and highly reflective steel surfaces. Automated Dynamo scripting for IFC reconstruction typically introduces dimensional deviations of 3 to 8 percent. (Rocha & Mateus, 2024)

Several systemic challenges further complicate NeRF-BIM integration. Scale ambiguity persists in the absence of ground control points (GCPs) or total station reference measurements. (Yuntae Jeon D. Q., 2025) Dynamic site elements such as cranes and workers introduce motion artifacts that are absent in static TLS captures. (Yuntae Jeon D. Q., 2025) Scaffolding and equipment occlusions prevent full scene coverage that would otherwise be achievable through multi-station TLS setups. (Shun Hachisuka, 2023)Textureless concrete surfaces defeat COLMAP feature matching, causing catastrophic reconstruction failures. (Shun Hachisuka, 2023)

From a workflow perspective, BIM authoring requires parametric families (walls, slabs, columns), whereas NeRF delivers amorphous point clouds. (Yuntae Jeon D. Q., 2025)IFC conversion loses semantic richness, with floors mapped to generic planar surfaces. (Shun Hachisuka, 2023) Construction progress monitoring using volume comparison yields error margins of approximately 15 percent. (Yuntae Jeon D. Q., 2025) The computational overhead of NeRF training approximately 24 hours per site versus 2 hours for scan registration compounds these workflow challenges. (Yuntae Jeon D. Q., 2025) No end-to-end automated pipeline currently exists for the full NeRF → segmentation → BIM workflow. (zbirovsky, 2025)Labeled construction-specific NeRF datasets for fine-tuning segmentation models remain scarce. (Shun Hachisuka, 2023)Lighting variability from shadows and glare degrades radiance fields in ways that radiometrically corrected LiDAR is immune to. (Yuntae Jeon D. Q., 2025) Topological errors including floating geometry and merged elements are common in complex interior reconstructions. (Shun Hachisuka, 2023)

LOD 300 BIM standards require sharp architectural edges; however, the volumetric blurring inherent to NeRF smooths corners and sharp boundaries. (Yuntae Jeon D. Q., 2025) Validation against TLS ground truth shows Chamfer distances two to three times larger than those obtained from photogrammetric point clouds. (Shun Hachisuka, 2023)Hybrid approaches that fuse NeRF visual detail with laser scan geometry partially address these limitations but introduce additional pipeline complexity. (Yuntae Jeon D. Q., 2025) Regulatory frameworks for construction and building surveying require certified accuracy levels that NeRF cannot yet reliably guarantee. (Yuntae Jeon D. Q., 2025)

Tableau 6: Systematic evaluation of NeRF capabilities against standard BIM integration requirements

BIM Criterion	Requirement	NeRF Capability	Assessment	Future Outlook
Geometric Accuracy	± 5 mm (LOD 300+)	Probabilistic (no explicit mesh)	✗ Not met	Possible via Depth supervision
Exportable Format (IFC / OBJ)	Structured 3D geometry	Implicit Neural model	✗ Not direct	Points2NeRF (Zimny & Joanna Waczyńska, 2024)
Measurement on Model	Direct dimensioning	Not supported	✗ Not possible	Research phase
Processing Time	Acceptable for projects	Hours to days (training)	⚠ Acceptable for research	Instant-NGP improves this
Visual Quality (Visualization)	Photo-realistic rendering	Excellent	✓ Excellent	Already superior
BIM Interoperability	IFC / Revit compatible	Not compatible	✗ Not met	(Jeon, Tran, & Kim, 2026) in progress

(Sources : (Eiri & Gheisari, 2024)*;* (Hachisuka, Yu, & Leung, 2023)*;* (Jeon, Tran, & Kim, 2026)*)*

Figure 48: NeRF-Based 3D scene reconstruction of an indoor /outdoor Building structure (Petzold, Matthias, & Christoph, 2024)

III.4.1. Non-Explicit Geometry:

NeRF employs a non-explicit geometric representation by encoding scene structure as a continuous neural function rather than a discrete mesh or voxel grid. (Ben Mildenhall, 2020)Surfaces emerge implicitly where volume density σ exceeds opacity thresholds along camera rays, without the existence of explicit vertex lists, face indices, or point arrays geometry is encoded entirely within the MLP weights. (Ben Mildenhall, 2020) This approach avoids tessellation artifacts and handles arbitrary topological complexity without modification. (Antonio Torralba, 2024) Any three-dimensional point x can be queried for density σ(x), and the scene can be rendered by integrating along rays without geometric intersection testing. (Ben Mildenhall, 2020)

When explicit geometry is required downstream, Marching Cubes post-processing is applied by sampling the density field on a regular three-dimensional grid to extract an is surface mesh. (Korycki, 2024) Non-explicit representation offers several advantages: it naturally captures fuzzy boundaries, thin structures, and amorphous materials; view-dependent appearance is coupled with geometry through directional radiance c (x, d); and the differentiable formulation enables direct optimization from two-dimensional image supervision. (Ben Mildenhall, 2020)Floaters spurious high-density regions in empty space are a known artifact of pure implicit fields, mitigated through distortion loss regularization. (Wolfe, 2023)

Variants of the implicit representation paradigm include SDF-based methods, where density is replaced by a signed distance function whose zero-level set defines watertight surfaces. (Garland & Heckbert) Neural SDFs predict distance to the nearest surface, while UNISURF unifies SDF and radiance field training for multi-view reconstruction. (Garland & Heckbert) Occupancy fields threshold the probability of solid versus empty space. (Yuntae Jeon D. Q., 2025) Non-explicit representations in theory allow infinite geometric detail, bounded only by the network's representational capacity. (Ben Mildenhall, 2020)Positional encoding is essential for capturing high-frequency geometric features. (Ben Mildenhall, 2020)

Challenges in geometry extraction arise because density peaks spread volumetrically rather than forming sharp surface boundaries. (Mittal A. , 2024) Advanced are surface methods such as Marching Tetrahedra or DUPR address this by refining is surfaces from blurry density fields. (Hubert, elghazaly, & frank, 2024)Hybrid explicit-implicit approaches such as NSVF, which bounds MLP queries to sparse voxels, and NeuMesh, which embeds local implicit functions per mesh vertex (Hubert, elghazaly, & frank, 2024) combine the strengths of both paradigms. Instant-NGP's hash grids accelerate spatial queries while preserving continuity (Adaloglou, 2022), and 3D Gaussian Splatting hybridizes explicit Gaussians with learned neural opacity. (Amazon Webservices., 2024)

Non-explicit geometry is well-suited to novel view synthesis but presents challenges for precise CAD-level geometric extraction. (Shun Hachisuka, 2023) BIM conversion thresholds the density field to generate point clouds, but loses the parametric crispness required by structural engineering workflows. (Code, 2024)Eikonal regularization enforces $\|\nabla\sigma\| = 1$ for physically plausible gradient fields (Yiheng, Zhaofan, Yingwei, Ting, & Tao, 2023), and multi-resolution spatial trees adapt sampling density to geometric complexity. (Mittal A. , 2024) Relightable NeRF variants decouple geometry from BRDF by fixing coarse density features during appearance optimization. Dynamic scene extensions embed deformation fields within four-dimensional spacetime implicit functions. (Hsiang-Hui Hung, 2024) Generative NeRF models disentangle shape from appearance through latent code sampling. (Hubert, elghazaly, & frank, 2024)Without coordinate contraction mappings, non-explicit representations scale poorly to large urban scenes. (Mittal A. , 2024) Geometry quality can be substantially improved through depth and normal supervision from monocular estimation networks. (Hubert, elghazaly, & frank, 2024)

III.4.2. Metric Reliability:

Metric reliability in NeRF refers to the geometric accuracy of extracted three-dimensional representations relative to ground-truth measurements obtained through reference instrumentation. (Petrovska, jager, haitz, & jutzi, 2023) Terrestrial Laser Scanning (TLS) achieves superior precision, with standard deviations on the order of 1.68 mm against structured light reference meshes. (jager & jutzi, 2023)By contrast, NeRF reconstructions exhibit substantially larger errors: even at high density thresholds such as δt = 300, mean deviations can reach 18.61 mm. (Petrovska, jager, haitz, & jutzi, 2023)

Increasing the density threshold improves metric accuracy by filtering noise and outliers, but simultaneously reduces point cloud completeness. (jager, landgraf, & jutzi, 2023) NeRF point clouds exhibit a characteristic problem of internal points density values located within object interiors rather than at surfaces which inflate cloud-to-mesh (C2M) error metrics. (Petrovska, jager, haitz, & jutzi, 2023)TLS point clouds, by contrast, maintain consistent millimeter-level fidelity across all surfaces without internal artifacts. (Petrovska, jager, haitz, & jutzi, 2023)

Quantitative analysis demonstrates that at **δt = 30**, NeRF produces approximately 19 million noisy points with a mean C2M error of 28.41 mm. Applying a threshold of **δt = 150** reduces the point count to 9.1 million and improves accuracy to 22.24 mm, but introduces geometric gaps in under sampled regions. (Petrovska, jager, haitz, & jutzi, 2023) Standard image quality metrics PSNR, SSIM, and LPIPS are designed for two-dimensional novel view synthesis evaluation and do not capture three-dimensional spatial fidelity. (Rubloff, What are the NeRF Metrics?, 2023) Critically, these metrics can mask geometric failures, as photorealistic rendered images may conceal underlying floaters and surface distortions. (Rubloff, 2023)

Three-dimensional evaluation relies on metrics such as Chamfer Distance (CD), which quantifies nearest-neighbor discrepancies between the NeRF reconstruction and the TLS reference cloud. (Korycki, 2024) In construction site benchmarks, NeRF CD errors are two to three times larger than those from photogrammetric pipelines. (Yuntae Jeon D. Q., 2025) Object-level experiments with PVC cylinders show NeRF underestimating diameters by 5 to 7 percent. (Korycki, 2024)Ground control points reveal absolute scale errors in the absence of total station tie-ins. (Jutzi, 2023)C2M analysis via Euclidean distance to reference facets further exposes the volumetric inaccuracies attributable to NeRF's implicit surface thickness. (Petrovska, jager, haitz, & jutzi, 2023)

Qualitative assessments reveal gaps at structurally important regions such as shoulders and joints where NeRF removes points to meet density thresholds. (Petrovska, jager, haitz, & jutzi, 2023) Color consistency also degrades across density threshold settings, with lower thresholds producing brighter renderings. (Petrovska, jager, haitz, & jutzi, 2023) Future improvements are expected through Eikonal constraints **($||\nabla\sigma|| = 1$)** for physically plausible surface gradients (Mittal A. , 2024), and hybrid TLS-NeRF fusion approaches that combine precision geometry with photorealistic radiance. (Yuntae Jeon D. Q., 2025) Construction industry tolerances of ±5 mm expose NeRF's current unsuitability for LOD 300 BIM deliverables. (Shun Hachisuka, 2023) Validation against structured light instrument (SLI) reference meshes with 0.1 mm accuracy confirms the fundamental metric limitations of the current NeRF paradigm. (Petrovska, jager, haitz, & jutzi, 2023)

III.4.3. Conversion to Meshes or Point Clouds:

Conversion from a trained NeRF to explicit meshes or point clouds extracts three-dimensional geometry from the learned implicit volumetric representation. (Ben Mildenhall, 2020) The most direct approach samples the density field σ on a regular three-dimensional grid typically at 512^3 resolution and retains all high-opacity points as a point cloud. (Korycki, 2024) Points are exported as colored PLY files where $\sigma > 10\text{-}50$, with color assigned by querying the learned radiance $c(x, d)$ at each point location. (Rakotosaona, Manhardt, & Arroyo, 2023)Isosurface meshes are extracted from the density field through the Marching Cubes algorithm, which interpolates between grid cells at a constant σ value; threshold values between 0.1 and 1.0 yield watertight meshes that are subsequently smoothed via Laplacian post-processing. (Ben Mildenhall, 2020)

For higher quality mesh extraction, specialized methods are available. NeRF Meshing distills a trained radiance field into an SDF network prior to surface extraction, enabling sharper meshes through precise gradient normals. (Rakotosaona, Manhardt, & Arroyo, 2023) NeuS and VolSDF train SDF representations directly during NeRF optimization, producing surface-optimized reconstructions without post-hoc distillation. (Garland & Heckbert) An alternative point cloud generation strategy traces rays densely across the scene and records hit points where the cumulative alpha value reaches 0.5. (Dominik Zimny, 2024) Surface normals are estimated from density gradients $\nabla\sigma$ to produce oriented point clouds suitable for downstream processing. (Korycki, 2024)

Grid resolution governs the trade-off between geometric fidelity and computational cost: a 1024^3 sampling grid captures centimeter-scale geometric details but requires in excess of 64 GB of RAM. (Rakotosaona, Manhardt, & Arroyo, 2023) Adaptive octree structures selectively up sample geometrically complex regions while down sampling empty space to improve efficiency. (Mittal A. , 2024) Decimation via Poisson reconstruction produces quad-dominant meshes suitable for real-time rendering. (Ben Mildenhall, 2020)UV unwrapping enables texture baking from NeRF renders into standard texture maps for game engine import. Appearance transfer via distilled feature grids generates bilinear texture maps for real-time mesh rasterization. (Rakotosaona, Manhardt, & Arroyo, 2023)

For BIM integration, structural point filtering retaining only vertical and horizontal planar regions precedes meshing to reduce noise. (Shun Hachisuka, 2023) ICP registration aligns NeRF point clouds with TLS ground truth for scale and orientation correction. (Petrovska, jager, haitz, & jutzi, 2023) Topological cleanup via hole-filling and non-manifold edge removal addresses common mesh defects. High-frequency detail preservation is achieved through Neural Angelo’s multi-scale hash encoding. (Korycki, 2024) Limitations of NeRF-derived meshes include volumetric blurring that softens sharp edges a characteristic absent in LiDAR-derived geometry and floaters that must be removed through statistical filtering or density Laplacian thresholding. (Rakotosaona, Manhardt, & Arroyo, 2023)

Dynamic scene fusion merges multi-session mesh reconstructions to support as-built progress monitoring models. (Hsiang-Hui Hung, 2024) Generative refinement can in paint occluded regions during the extraction process. (Hubert, elghazaly, & frank, 2024)Mesh simplification targeting specific triangle budgets prepares models for AR and VR deployment. Geometric validation using Hausdorff distance metrics below 5 cm confirms survey-grade output quality. (Petrovska, jager, haitz, & jutzi, 2023) Editable splines lofted from cross-sections of sliced point clouds enable parametric modeling workflows. (Code, 2024)Real-time viewers in Nerfstudio preview geometric extractions during distillation training. (Matthew Tancik, 2023)

IFC export classifies mesh faces into wall, floor, and ceiling categories through RANSAC plane detection. (Shun Hachisuka, 2023)

Figure 49:Novel view synthesis of Backhoe Loaders Using Neural Radiance fields (NeRF). (AR-MR-XR, 2022)

III.4.4. Conclusion of NeRF Technology:

NeRF (Neural Radiance Field) technology leverages artificial intelligence to reconstruct a three-dimensional model from a single image, producing results that appear visually complete at first glance. However, the generated output does not represent ground truth at 100%, as it fundamentally relies on computational probabilities to estimate the object's geometry in areas that were not captured in the original image.

Furthermore, despite its three-dimensional appearance, the model is essentially a computational representation that simulates color and lighting from varying viewpoints, without possessing a true underlying geometric structure. Consequently, the model cannot be exported in a geometrically usable format, nor can precise geometric measurements be performed on it.

In light of the above, this photogrammetric approach is deemed unsuitable for deployment within the standard working environment of Building Information Modeling (BIM).

Chapter IV : Gaussian Splatting for 3D Reconstruction

IV.1. Emergence of Gaussian Splatting:

IV.1.1. Motivation and Evolution from NeRF:

For decades, computer graphics and vision have pursued photorealistic, geometrically accurate 3D scene representations. Neural Radiance Fields (NeRF) (Mildenhall, et al., 2020)marked a turning point: a compact MLP could encode a bounded scene and synthesize novel views with unprecedented realism from sparse calibrated photos. NeRF learned a continuous volumetric function mapping position $\{\mathbf{x}\}$ and direction $\{\mathbf{d}\}$ to color $\{\mathbf{c}\}$ and density $\{\boldsymbol{\sigma}\}$, then composited colors along rays via volume rendering. Within three years, the paper attracted over 10,000 citations (Tancik, et al., 2023)and spawned a large sub-field.

Yet NeRF had five fundamental limitations:

1) **Slow training:** original NeRF took 1~2 days per scene on a single GPU. (Mildenhall, et al., 2020) Instant-NGP reduced this to seconds via hash encoding, but added complexity. (Müller, Evans, Schied, & Keller, 2022)
2) **Slow inference:** hundreds of MLP evaluations per pixel prevented real-time use. Even baked NeRF (Hedman, Srinivasan, Mildenhall, Barron, & Debevec, 2021)or Plen Octrees achieved only 10~30 fps at reduced resolution. (Yu, et al., 2021)
3) **No explicit geometry:** the density field does not correspond to a well-defined surface; mesh extraction via Marching Cubes produces noisy, incomplete geometry. (Oechsle, Peng, & Geiger, 2021)
4) **Poor unbounded scene handling:** original NeRF assumed bounded near-object capture. Mip-NeRF 360 addressed this with non-linear parameterization but increased training time. (Barron, Mildenhall, Verbin, Srinivasan, & Hedman, 2022)
5) **Static only:** extending NeRF to dynamic scenes required per-frame optimization or learned deformation fields (Park, et al., 2021)introducing convergence difficulties. (Pumarola, Corona, Pons-Moll, & Moreno-Noguer, 2020)

- The first approach led to architectural improvements within the implicit model (such as mipNeRF (Barron, et al., 2021), RegNeRF (Niemeyer, et al., 2021), RawNeRF (Mildenhall, Hedman, Martin-Brualla, Srinivasan, & Barron, 2021), and Ref NeRF (Verbin, et al., 2021) and more importantly, to the replacement of implicit representation with explicit, differentiable representation suitable for real-time rendering.

- The second approach led to the emergence of 3D Gaussian spraying (3DGS) by (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023), presented at the SIGGRAPH 2023 conference. The core idea is to represent a scene as a set of 3D Gaussian primitives, each with a center {μ}, a covariance matrix {Σ}, an opacity {α}, and angle-dependent spherical harmonics {SH} for color. This explicit representation is stored in GPU memory and processed via a fast tile-based rasterizer, achieving real-time rendering (>100 frames per second) after minutes of training - not hours. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

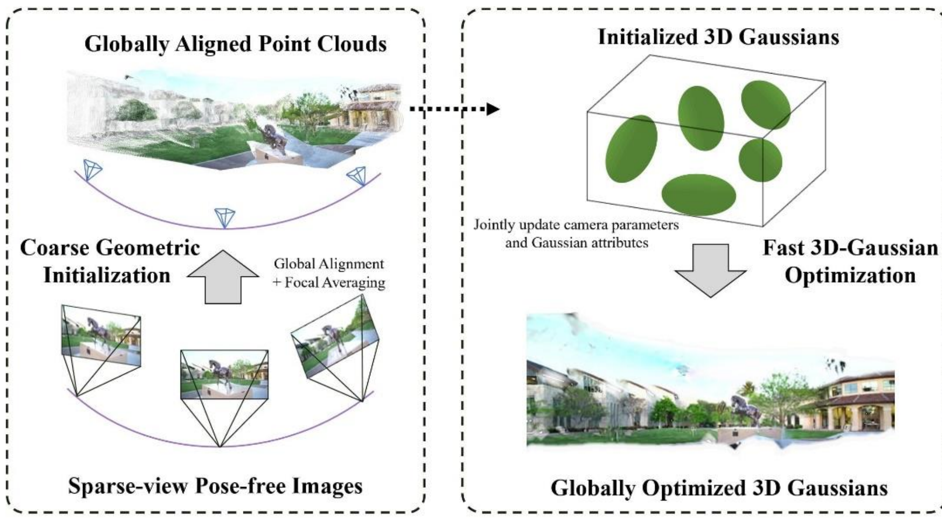

Figure 50:The process of building a 3D model using Gaussian Splatting from multi-angle images (Rubloff, InstantSplat: Sub Minute Gaussian Splatting, 2024)

3DGS changed the scene representation from a black-box neural function to an inspectable, editable point-like structure, mirroring a broader shift from implicit to explicit representations (Chen, Xu, Geiger, Yu, & Su, 2022) (Yu, et al., 2021)

<u>Subsequent years brought rapid extensions:</u> Deformable 3DGS (Yang, et al., 2023)for monocular dynamic scenes, Scaffold-GS (Lu, et al., 2023)with anchor points, Compressed 3DGS (Niedermayr, Stumpfegger, & Westermann, 2024)achieving 31× compression, Gaussian Editor (Chen, et al., 2023)for text-driven editing, SuGaR (Guédon & Lepetit, 2023)for mesh extraction, and Semantic Gaussians (Guo, Ma, Fan, Liu, & Li, 2024). These developments show that 3DGS has assumed the role NeRF played in 2020–2022: the dominant substrate for novel-view synthesis research.

- Critical Analysis:(Multi-View Consistency):

A recurring critique: does 3DGS's per-Gaussian SH parameterization introduce multi-view inconsistencies absent in NeRF?

NeRF's density {σ} is view-independent; view-dependent color is strictly in the radiance. 3DGS assigns {SH} to each Gaussian. If a Gaussian is seen from only a small angular range, {SH} fitting becomes under-constrained, causing extrapolation failures from novel views. (Charatan, Li, Tagliasacchi, & Sitzmann, 2024)

Conversely, Gaussian Opacity Fields (Yu, Sattler, & Geiger, 2024)argue that NeRF's continuous representation does not guarantee multi-view consistency either. Empirically, both methods show comparable inconsistency on unbounded outdoor scenes. (Barron, Mildenhall, Verbin, Srinivasan, & Hedman, 2022), (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

However, 3DGS produces structured "splat popping" during camera sweeps, whereas NeRF gives diffuse inconsistencies masked by blurring a non-trivial distinction for AEC inspection applications.

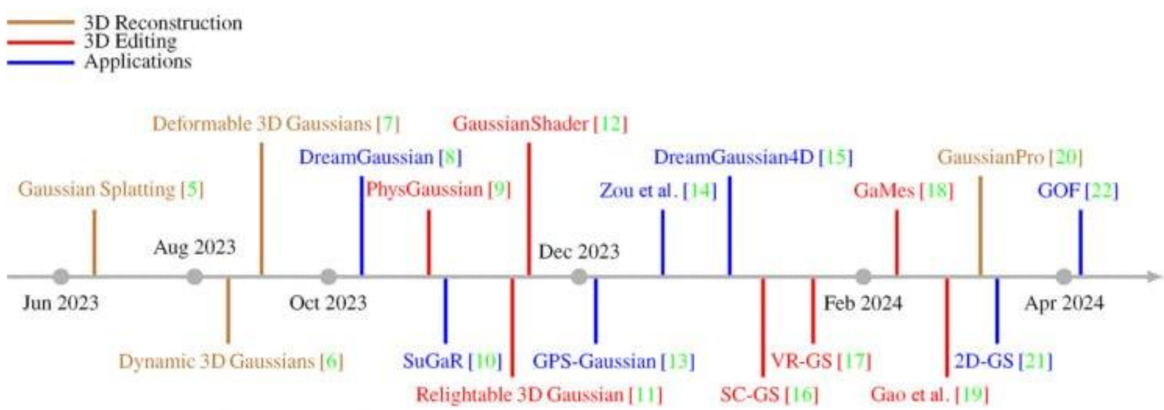

Fig. 2 A brief timeline of representative work using the 3D Gaussian splatting representation.

Figure 51:A brief timeline of representative work using the 3D Gaussian splatting representation (Wu, et al., 2024)

IV.1.2. Advantages in Real-Time Rendering:

3DGS's most commercially significant property is real-time rendering (≥30 fps at 1080p). Table 1.1 compares key methods on the Mip-NeRF 360 dataset (Barron, Mildenhall, Verbin, Srinivasan, & Hedman, 2022), Tanks and Temples (Knapitsch, Park, Zhou, & Koltun, 2017), and Deep Blending. (HEDMAN, et al., 2018)

Tableau 7:Comparative performance (NeRF, Instant-NGP, 3DGS).

Method	PSNR (Mip360)	SSIM (Mip360)	LPIPS (Mip360)	Training Time	FPS (1080p)	GPU Memory
NeRF	25.78	0.695	0.330	~24h	<1	~6 GB
mip-NeRF 360	29.40	0.842	0.127	~48h	<1	~8 GB
Instant-NGP	28.66	0.823	0.144	~3 min	~40	~4 GB
3DGS	**29.41***	**0.870***	**0.120***	**~30–45 min**	**>130***	**~6–12 GB**

Source: NeRF (Mildenhall, et al., 2020) *;mip-NeRF 360* (Barron, Mildenhall, Verbin, Srinivasan, & Hedman, 2022)*; Instant-NGP* (Müller, Evans, Schied, & Keller, 2022)*; 3DGS* (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

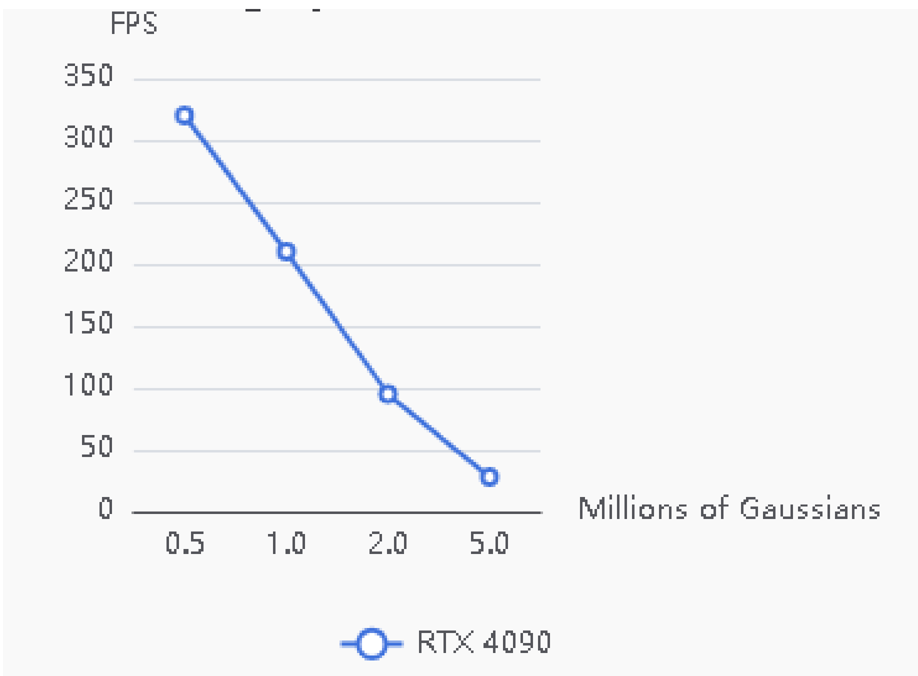

Figure 52:FPS vs. Gaussian count on RTX 4090, scalability bottleneck for real-time rendering. Adapted from performance benchmarks (Ours)

3DGS matches mip-NeRF 360 in quality but delivers >100×higher frame rates. The speed comes from a tile-based rasterizer (16×16 tiles) that projects 2D Gaussians (splats), sorts by depth, and alpha-composites in a single GPU pass avoiding per-ray MLP evaluations. The Gaussian representation is sparse: empty space contains no primitives.

Memory is a weakness. An indoor scene needs 1~3 million Gaussians, each storing position (3×float), covariance (6 floats after Cholesky), opacity (1 float), and {SH} coefficients (48 floats for degree-3), totaling ~236 bytes per Gaussian ~236 MB per million Gaussians (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023). Complex outdoor scenes may reach 5~10 million Gaussians (1.2~2.3 GB). Compressed 3DGS reduces this to 7~15 MB per scene but with a quality penalty on fine details. (Niedermayr, Stumpfegger, & Westermann, 2024)

- Critical Analysis: Memory, Quality, Speed Triangle:

No current method simultaneously optimizes all three metrics. 3DGS offers high speed and competitive quality but non-trivial memory; Instant-NGP gives low training time and moderate quality with compact storage; compressed 3DGS cuts memory at a quality cost. Single-metric "state-of-the-art" claims are misleading unless all three dimensions are reported. Reviewers and practitioners should demand multi-metric ablations.

IV.1.3. Current Research Landscape (2025-2026):

By early 2026, Gaussian Splatting literature includes hundreds of papers. Representative directions include:

1) **Deformable 3DGS:** conditions Gaussian positions on time via an MLP deformation field for monocular dynamic reconstruction. (Yang, et al., 2023)
2) **Scaffold-GS**: introduces anchor points to generate local Gaussians on-the-fly, reducing redundancy and improving texture less areas. (Lu, et al., 2023)
3) **Compressed 3DGS:** sensitivity-aware vector quantization, 31× compression with <1 dB PSNR loss. HAC++ (Chen, Wu, Lin, Harandi, & Cai, 2025) reaches 100× compression. (Niedermayr, Stumpfegger, & Westermann, 2024)
4) **Semantic Gaussians:** attaches open-vocabulary embeddings to Gaussians for language-driven scene understanding. (Guo, Ma, Fan, Liu, & Li, 2024)
5) **2D Gaussian Splatting:** replaces 3D ellipsoids with 2D disks, improving surface normals and watertight mesh extraction. (Huang, Yu, Chen, Geiger, & Gao, 2024)
6) **Gaussian Opacity Fields:** derives level-set surfaces directly from ray-Gaussian intersection, reducing surface noise on smooth architecture. (Yu, Sattler, & Geiger, 2024)
7) **Gaussian Editor:** text-guided editing via 2D diffusion model propagation into 3D Gaussian attributes. (Chen, et al., 2023)
8) **SuGaR:** regularization that aligns Gaussians with scene surfaces, enabling joint optimization of Gaussians and a bound mesh. (Guédon & Lepetit, 2023)

9) **Driving Gaussian:** decomposes autonomous driving scenes into static background and dynamic graphs for moving objects. (Zhou, et al., 2024)

10) **4D Gaussian splatting:** It represents an extension of the Gaussian distribution by adding a fourth dimension (often time or radiation) to the spatial coordinates, allowing for a dynamic description of the scene's evolution. It also contributes to improving the continuity of representation and the accuracy of reconstruction, especially in cases involving changes in lighting or movement. (Duan, et al., 2024)

- Beyond these: Gaussian-based SLAM (Yan, et al., 2024), LiDAR-camera fusion (Zhao, et al., 2024), medical imaging, XR integration (Xiaofeng, et al., 2025), heritage digitization (Zhenyu, et al., 2025), and Gaussian Building Mesh (GBM,) for built-environment digitization. (Gao, et al., 2025)

- Critical Analysis: Structural Biases:

The field overwhelmingly evaluates on photogenic, uncluttered scenes from a few benchmarks (Barron, Mildenhall, Verbin, Srinivasan, & Hedman, 2022), (Knapitsch, Park, Zhou, & Koltun, 2017), (HEDMAN, et al., 2018). Industrial and construction environments repetitive textures, metallic reflections, dust, rebar clutter receive minimal attention. Only Driving Gaussian approaches an industrial scenario. The incentive structure (conference papers rewarded for PSNR/SSIM gains on standard benchmarks) creates a novelty-scene bias, inflating reported performance and obscuring failure modes on industrial data. The community needs dedicated construction-site and infrastructure benchmarks. (Zhou, et al., 2024)

IV.2. Theoretical Principles:

IV.2.1. 3D Gaussian Representation:

A scene {S} is represented as a finite set of {N} anisotropic Gaussian primitives:

$$G(x) = \exp(-\tfrac{1}{2} (x - \mu)^T \Sigma^{-1} (x - \mu))$$

where {**μ**} is the centroid, {**Σ**} the positive semi-definite covariance, {**α**} the opacity, and {**c**} the view-dependent color from {**SH**} coefficients. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

- **Interpretation:** Equation (2.1) measures a Gaussian's influence at position {**x**}1 at {**μ**}, decaying exponentially with rate controlled by {**Σ**}. Isotropic (spherical) Gaussians decay equally in all directions; anisotropic (elongated) ones decay slowly along principal axes, representing thin structures.

- **Numerical stability:** When a Gaussian is very flat (scale factors → 0), $\boldsymbol{\Sigma^{-1}} \to \boldsymbol{\infty}$, producing NaN gradients. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023) clamp scales to a minimum {ε}, but this does not solve the underlying ill-conditioning. To keep {Σ} positive semi-definite during optimization, they decompose it as:

$$\boldsymbol{\Sigma = R\, S^{T} R^{T}}$$

with {**R**} a rotation matrix (unit quaternion {**q**}) and **{S} = diag** $(\mathbf{s_x}, \mathbf{s_y}, \mathbf{s\_z})$ a non-negative diagonal scaling matrix. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

- **Interpretation:** Instead of optimizing the 9 entries of {Σ} directly (which may become indefinite), this factorization separates rotation and scale both have well-defined geometric meaning and valid optimization spaces (quaternion sphere and $\mathbb{R}_+^3$).

❖ **Critique of the decomposition:** Gradient flow through {Σ} is ill-conditioned for flat Gaussians: when $\mathbf{s_i} \to \mathbf{0}$, the gradient with respect to that scale becomes very large relative to others, causing unstable optimization. The quaternion normalization $\| \mathbf{q} \| = \mathbf{1}$ must be enforced at each step; violations accumulate in half-precision training (Niedermayr, Stumpfegger, & Westermann, 2024). The rotation-scale factorization is non-unique (180° rotation ambiguity), which can cause redundant Gaussian pairs in under-constrained regions. 2D Gaussian Splatting addresses flatness by using 2D disk primitives, at the cost of modelling only surface-like structures. (Huang, Yu, Chen, Geiger, & Gao, 2024)

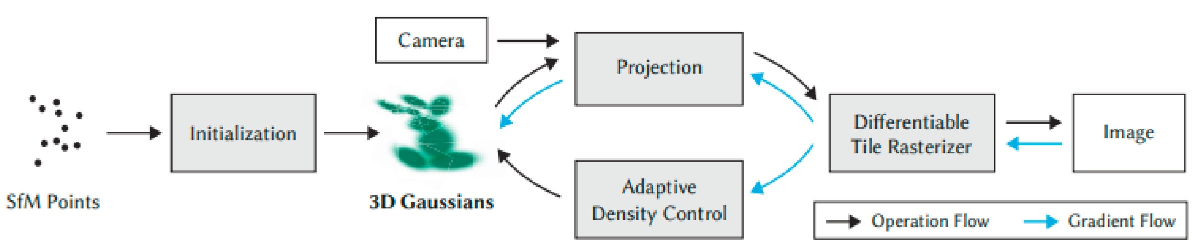

Figure 53:Flowchart for creating and displaying 3D Gaussian patterns using Gaussian Splatting (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

- Critical Analysis (Limitations of the Gaussian Primitive Model):

- **Thin structures (wires, rebar):** poorly represented; they are either omitted or approximated by chains of overlapping Gaussians that fail from novel views. (Charatan, Li, Tagliasacchi, & Sitzmann, 2024)
- **Sharp edges and corners:** Gaussian smooth decay conflicts with step-function discontinuities; approximating a sharp edge requires many overlapping Gaussians, causing "feathering". 2DGS and GOF partially address this. (Huang, Yu, Chen, Geiger, & Gao, 2024) (Yu, Sattler, & Geiger, 2024)
- **Transparent materials (glass facades, water):** poorly represented by α-compositing, which assumes opaque or uniformly semi-opaque primitives. SuGaR notes multi-layer splat artefacts. (Guédon & Lepetit, 2023)
- **Large homogeneous regions (plain concrete):** no photometric gradient to drive Gaussian initialization from SfM point clouds; these regions are systematically under-represented. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023) (Ye M. , Danelljan, Yu, & Ke, 2024)

IV.2.2. Rendering Pipeline (Differentiable Splatting):

The pipeline has four stages: projection, depth sorting, tile-based rasterization, and differentiable alpha compositing.

IV.2.2.a. Stage 1 (Projection of 3D Gaussians to 2D):

Given a camera with view matrix {**W**} and Jacobian {**J**} of the projective transformation, the 3D covariance {**Σ**} projects to 2D as:

$$\Sigma' = J\, W\, \Sigma\, W^T J^T$$

The z-row/column are discarded, giving a **2×2** covariance defining the 2D splat shape.

- **Interpretation:** Each 3D ellipsoid projects to a 2D ellipse whose shape depends on the Gaussian's orientation and size relative to the camera. {**J**} accounts for non-linear perspective distortion; its linearization is valid only when the Gaussian is small relative to scene depth. Large Gaussians (common early in training) introduce projection errors. (Zwicker, Pfister, Baar, & Gross, 2002)

IV.2.2.b. Stage 2 (Depth Sorting):

All {N} Gaussians visible to the camera are sorted back-to-front by depth z_i using a GPU radix sort, re-executed every frame. For **$N = 3\times10^6$,** sorting takes ~4~8 ms on an RTX 3090 (20~40% of frame time). (Bell & Hoberock, 2011)

- **Critique:** Depth sort is per-Gaussian, not per-pixel. Overlapping Gaussians can have inconsistent depth orderings for different pixels in the same frame, producing incorrect alpha compositing the classic order-independent transparency problem. This artefact is usually imperceptible at high density but causes errors near object boundaries. (Radl, et al., 2024)

IV.2.2.c. Stage 3 (Tile-Based Rasterization):

The image is partitioned into **16×16**-pixel tiles. For each tile, a list of intersecting Gaussians is built by testing whether the 2D splat intersects the tile bounding box. This is the key innovation of Kerbl et al. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023) over prior splatting methods. (Kopanas, Philip, Leimkühler, & Drettakis, 2021)

IV.2.2.d. Stage 4 (Alpha Compositing):

Within each tile, colors are accumulated back-to-front using the over operator:

$$C(u) = \Sigma_i \in tile\ c_i\ \alpha_i\ G'_i(u) \prod_j < _i (1 - \alpha_j\ G'_j(u))$$

where {**G′**} is the 2D Gaussian kernel at pixel {**u**}, and the product term is the transmittance from all Gaussians in front of {**i**}.

- **Interpretation:** Each Gaussian contributes color c_i attenuated by its own opacity {$\boldsymbol{\alpha}_i$} and by the accumulated opacity of all Gaussians between the camera and it. Background contributions vanish behind an opaque foreground Gaussian.
- **Computational complexity:** For K Gaussians intersecting a tile and P pixels per tile, O(P·K) operations. Typically, **K = 5- 40**. Peak K ≈ 200 in scenes with many translucent overlapping surfaces may cause tile buffer overflow. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)
- **Differentiability:** The entire pipeline is differentiable w.r.t. Gaussian attributes {**μ, q, s, α, SH**}. Gradients back-propagate through the compositing sum and 2D Gaussian evaluations to update parameters via stochastic gradient descent. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

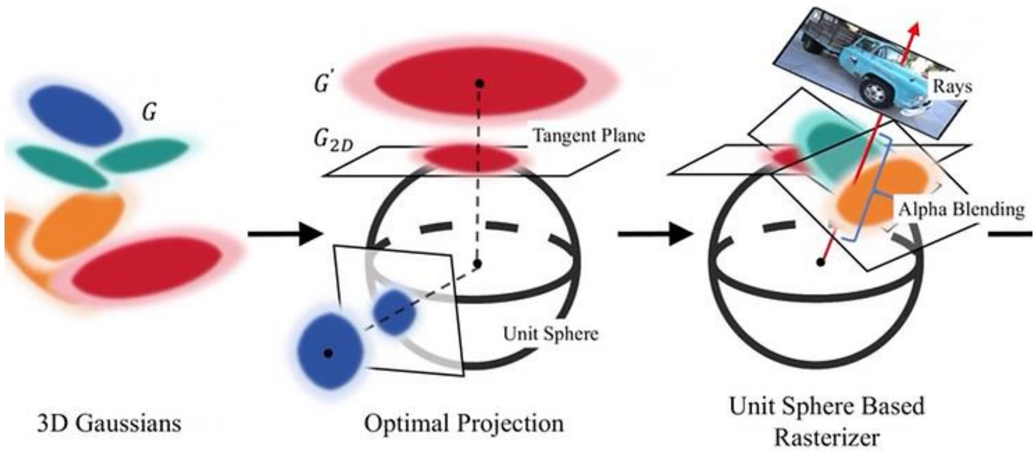

Figure 54:3D Gaussian Splatting Works (Hammer missions, 2026)

- Critical Analysis: (Tile-Based Pipeline Artefacts):
- **Tile boundary artefacts:** discontinuous jumps between tile lists can produce faint grid-like patterns at low Gaussian densities or high camera distances. (Lassner & Zollhöfer, 2020)
- **Depth sort instability:** quasi-parallel Gaussians (depth difference below a threshold) produce flickering artefacts during camera motion due to floating-point imprecision. This is a fundamental limit of painter's-algorithm compositing. (Radl, et al., 2024)
- **Gradient vanishing for occluded Gaussians:** Gaussians consistently occluded from all training views receive near-zero gradients and are pruned, producing hollow interiors. Acceptable for view synthesis but catastrophic for structural inspection.

IV.2.3. Comparison with NeRF Representations (Mathematical):

The fundamental difference: NeRF uses a continuous integral; 3DGS uses a discrete sum.

IV.2.3.a. NeRF Volume Rendering Integral:

$$C(r) = \int_{tn}^{tf} T(t)\ \sigma(r(t))\ c(r(t), d)\ dt\ , \quad T(t) = \exp\left(-\int_{tn}^{t}\sigma(r(s))\ ds\right)$$

with **r(t) = o + t d**,{**σ**} density, {**c**} color, and **T(t)** accumulated transmittance. (Mildenhall, et al., 2020)

- **Interpretation:** Expected ray color is a weighted integral over points along the ray; the weight is local density multiplied by accumulated transparency from the camera.

- **Computational cost:** Numerical integration requires M MLP evaluations per ray **($M \approx 192$)**. For 1080p (**$\sim 2\times10^6$** pixels), that's **$\sim 3.8\times10^8$** MLP evaluations per frame - incompatible with real-time rendering.
- **Critique:** The continuous formulation is theoretically appealing but expensive. More importantly, the gradient of C with respect to individual scene points is diffuse and non-local, creating long-range dependencies that complicate convergence (Zhang, Riegler, Snavely, & Koltun, 2020). 3DGS avoids this by assigning gradients directly to individual Gaussians.

IV.2.3.b. 3DGS Discrete Splatting Sum:

$$C(u) = \Sigma_{i=1}^{N} c_i\, \alpha_i\, G'_i(u) \prod_{j=1}^{i-1} (1 - \alpha_j\, G'_j(u))$$

- **Interpretation:** Instead of integrating along a ray, 3DGS sums contributions from $\{\mathbf{N}\}$ sorted 2D Gaussian splats at pixel $\{\mathbf{u}\}$. The continuous density $\{\boldsymbol{\sigma}\}$ is replaced by the discrete Gaussian kernel $\{\mathbf{G'}\}$, and the depth integral becomes a sum over sorted primitives.
- **Computational cost:** For $\{\mathbf{K}\}$ Gaussians per pixel (**5~50**), O(K) operations per pixel, versus O(M) MLP evaluations for NeRF. GPU parallelism makes these two orders of magnitude faster.
- **Critique:** The discrete summation introduces bias relative to the continuous integral: if depth sampling is coarser than Gaussian separation, the discrete sum over-estimates transmittance. This bias is typically imperceptible for view synthesis but causes systematic errors in depth estimation and surface reconstruction. (Yu, Sattler, & Geiger, 2024), (Huang, Yu, Chen, Geiger, & Gao, 2024)

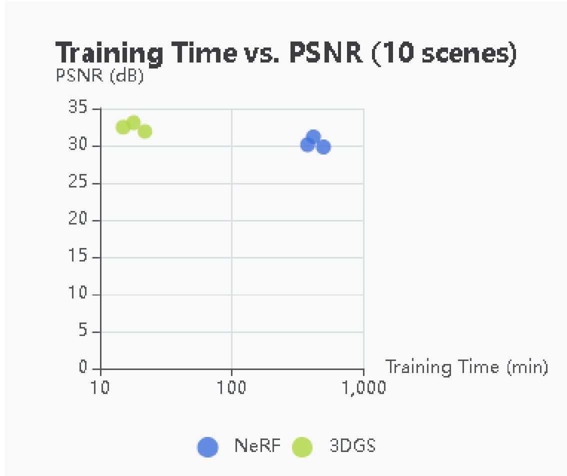

Figure 55:Training Time vs. PSNR.,source: (Müller, Evans, Schied, & Keller, 2022)

■ Critical Analysis (Continuous vs. Discrete Trade-off):

- **Continuity (NeRF):** naturally encodes priors (smoothness, boundedness) through the MLP architecture. Produces smooth depth maps, coherent normals, and globally propagating gradients. Anti-aliasing is inherently supported via mip-mapping. (Barron, et al., 2021)
- **Discreteness (3DGS):**primitives are directly interpretable and editable (select, delete, relocate without re-optimizing). Enables real-time rendering via rasterization hardware. Supports scalable scene composition (merging independent Gaussian sets) without re-training.
- **Unresolved tension:** NeRF performs better on smooth, featureless surfaces (walls, ceilings); 3DGS performs better on high-frequency texture detail (foliage, cobblestones) - supported by LPIPS breakdowns in (Barron, Mildenhall, Verbin, Srinivasan, & Hedman, 2022) and ablations in (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023). For AEC (flat concrete, uniform glass, modular geometry), neither paradigm is uniformly superior; hybrid approaches (Scaffold-GS (Lu, et al., 2023), GOF (Yu, Sattler, & Geiger, 2024)) may be necessary.

IV.3. Processing Workflow:

IV.3.1. Data Preparation (SfM, COLMAP, Depth Priors):

3DGS optimization starts from a sparse 3D point cloud from Structure-from-Motion (SfM). This initialization strongly influences final quality. The standard SfM tool is COLMAP (Schönberger & Frahm, 2016), based on incremental SfM with bundle adjustment. (Schönberger, Zheng, Pollefeys, & Frahm, 2016)

IV.3.1.a. Camera Projection Error (Reprojection Error):

COLMAP's bundle adjustment minimizes:

$$E\_reproj = \Sigma_{j=1}^{M} \Sigma_{k=1}^{Kj} \rho(\|u\_jk - \pi(K_j, R_j, t_j, X\_k)\|^2)$$

where {**u_jk**} is the detected feature location of point {**k**} in camera {**j**}, {**π**} the projection function (intrinsics {**K**}, rotation {**R**}, translation {**t**}), {**X_k**} the 3D point, and ρ a robust loss (Huber orCauchy).

- **Interpretation:** Simultaneously refines all camera poses and 3D points so that projections land as close as possible to detected 2D matches. The result is a metrically consistent, globally registered point cloud.
- **Computational cost:** Non-linear least-squares with O(M+N) unknowns (M cameras, N points) and O(M·K) residuals. For **M=500** images, **N=100,000** points, COLMAP takes 10~60 minutes on a modern workstation. (Schönberger & Frahm, 2016)
- **Real-world observation construction sites:** Standard COLMAP fails in several common scenarios:

– Reflective glass curtain walls: specular highlights shift with camera position, causing feature trackers to track highlights rather than geometry. (Whelan, Salas-Moreno, Glocker, Davison, & Leutenegger, 2016)

– Freshly poured concrete: uniform grey albedo with minimal texture provides no reliable key points for SIFT/ORB. (Whelan, Salas-Moreno, Glocker, Davison, & Leutenegger, 2016)

– Steel rebar cages: repetitive periodic patterns cause widespread feature ambiguity and false matches (Misra, Girdhar, & Joulin, 2021)

Workaround strategies:

– Pre-processing with contrast-limited adaptive histogram equalization (CLAHE) to enhance low-contrast concrete textures. (Kim, Wang, & Cheng, 2019)

- Supplementing photo-based SfM with ARKit depth priors for metric scale. (Remondino & Stoppa, 2013)
- Using learning-based feature matchers such as SuperGlue (Sarlin, DeTone, Malisiewicz, & Rabinovich, 2020) or LoFTR, which perform better on repetitive industrial textures. (Sun, Shen, Wang, Bao, & Zhou, 2021)
- Fusing COLMAP output with RGB-D data from Intel RealSense or Azure Kinect. (Bae, Golparvar-Fard, & White, 2013)

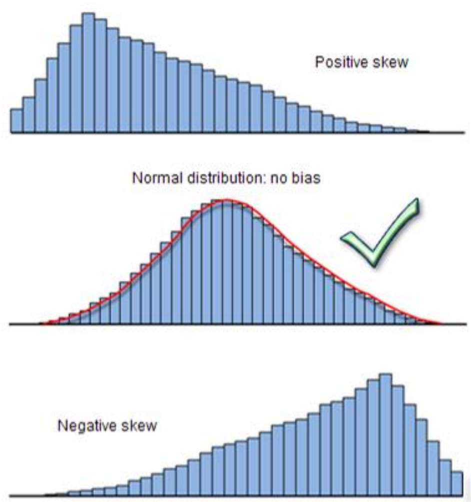

Figure 56:The ideal Gaussian distribution of reprojection errors compared to skewed distributions in building environments. (Ahmed, 2020)

IV.3.1.b. Depth Prior Integration:

When SfM fails in textureless regions, monocular depth priors (MiDaS (Ranftl, Bochkovskiy, & Koltun, 2021); Depth Anything V2 (Yang, et al., 2024)) can initialize Gaussians in regions lacking SfM coverage. Depth-regularized 3DGS (Zhu, Fan, Jiang, & Wang, 2024)adds a depth consistency loss:

$$\mathscr{L}\_depth = \Sigma_u \|D\_render(u) - D\_prior(u)\|^2$$

where D_render is rendered depth from Gaussians, D_prior the monocular depth prior.

- **Critique:** Monocular depth priors are up-to-scale and up-to-shift; they must be aligned to metric scale before use (Yang, et al., 2024). On construction sites where metric accuracy is mandatory (tolerances ±5 mm per EN ISO 17123-1), unscaled priors introduce systematic errors that are difficult to quantify without independent ground truth.

- Critical Analysis (Garbage-In Problem):

The quality of the initial SfM point cloud disproportionately influences final 3DGS reconstruction quality. A sparse or misregistered cloud produces Gaussians in wrong locations that densification may fail to correct because the photometric gradient is insufficient to drive them to correct positions. Contrary to claims of "robustness to initialization" (Gong, 2024), mean PSNR drops by 1.5~3.0 dB and floater artefacts increase by 15~30% when COLMAP is run on image sets with >30% of views removed (Ye M. , Danelljan, Yu, & Ke, 2024). For construction sites with inconsistent image overlap, this sensitivity is a practical deployment barrier.

IV.3.2. Training and Visualization:

The training loop alternates between differentiable rendering, loss computation, backpropagation, and adaptive density control (ADC) that adds/removes Gaussians based on scene coverage.

Algorithm Modified (One Training Iteration of 3DGS (simplified)): (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

```
INPUT:
  G = {(μᵢ, qᵢ, sᵢ, αᵢ, fᵢ)}    // current Gaussian set
I_gt                          // ground-truth training image
  K, [R|t]                    // camera intrinsics and extrinsic
  d = normalize(C - μᵢ)           // viewing direction from camera center C
FORWARD PASS:
  1. For each Gaussian i:
    a. Compute 3D covariance:
      Σᵢ = R(qᵢ) · diag(sᵢ)² · R(qᵢ)ᵀ
    b. Project to 2D (Eq. 2.3):
      Σ'ᵢ = J(μᵢ) · W · Σᵢ · Wᵀ · J(μᵢ)ᵀ
      where W = [R|t], and J(μᵢ) is the Jacobian of perspective
      projection evaluated at the Gaussian center μᵢ
c. Compute 2D centroid:
μ'ᵢ = π(K, R, t, μᵢ)
d. Compute color from Spherical Harmonics:
```

```
    c_i(d) = SH(f_i, d)          // view-dependent appearance
  2. Depth-sort all Gaussians by z_i = (W · μ_i)_z   [GPU radix sort]
  3. View-frustum culling: discard Gaussians with μ'_i outside image bounds
  4. Tile assignment: for each 16×16 tile T:
       - Identify all Gaussians whose 2D splat intersects T
  5. Per-tile alpha compositing (front-to-back) [Eq. 2.4]:
I_render(p) = Σ_i c_i(d) · α_i · G'_i(p) · Π_j<i (1 − α_j · G'_j(p))
  6. Compute loss:
       L = (1 − λ) · L1(I_render, I_gt) + λ · [1 − SSIM(I_render, I_gt)]
       where λ = 0.2 (default)
       // NOTE: we minimize DISSIMILARITY; SSIM alone would be maximized
BACKWARD PASS:
  7. Compute ∂L/∂(μ_i, q_i, s_i, α_i, f_i) for all i via autograd
  8. Update parameters: Adam optimizer step
ADAPTIVE DENSITY CONTROL (every 100 iterations, during first 15,000 iterations):
  9. For Gaussians with ‖∂L/∂μ'‖ >τ_pos (default 0.0002):
     a. If max(s_i) >τ_size (scene-scale dependent):
          SPLIT: replace with 2 smaller Gaussians
s_new = s_i / 1.6
μ_new = μ_i ± ε · normalize(∂L/∂μ')
     b. Else:
          CLONE: add 1 duplicate Gaussian at μ_i + δ · normalize(∂L/∂μ')
  10. Prune Gaussians where α_i < ε_α (default 0.005)
  11. Prune Gaussians where max(s_i) exceeds world-space threshold
  12. Prune Gaussians outside the view frustum after 2-3 views
  13. Every 3,000 iterations: reset all α_i to 0.01 (not 0) to remove
      transparent floaters while preserving gradients
OUTPUT: Updated Gaussian set G'
```

Tableau 8:Hyperparameter settings and effects of deviation (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

Hyperparameter	Default Value	Effect of Deviation
Learning rate (position)	1.6×10^{-4} (decay)	Higher → faster but unstable Gaussian drift
Learning rate (SH coeff)	0.0025 / 0.000125	Higher → colour overfitting
Learning rate (opacity)	0.05	Lower → slow opacity convergence
Learning rate (scale)	0.005	Higher → scale explosion in sparse regions
Learning rate (quaternion)	0.001	Higher → rotation oscillation
SSIM weight λ	0.2	Higher → blur artifacts; lower → noise
Densification threshold τ_pos	0.0002	Lower → over-densification; higher → under-coverage
Opacity pruning ε_α	0.005	Higher → aggressive pruning; holes in thin regions
Densification start iter	500	Earlier → noisy initial clones
Densification end iter	15,000	Later → continued growth; OOM risk
Opacity reset interval	3,000	Fewer resets → accumulated transparent floaters

Sources adapted from (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

- Critical Analysis (Ad-Hoc Densification Heuristics):

- **No convergence guarantee:** greedy heuristic; final Gaussian count depends on scene complexity and {τ_pos} without principled selection.
- **Scene-scale dependence:** τ_size is a fraction of scene extent estimated from COLMAP; inconsistent scale (indoor+outdoor) causes incorrect split decisions.
- **Opacity reset instability:** resetting all α to a small value disrupts gradient flow, causing PSNR drops of 0.5~1.5 dB immediately after each reset (not reported in original paper).
- **Non-reproducibility:** due to stochastic GPU radix sort and floating-point non-associativity, two identical runs produce Gaussian sets differing by 5~15% in count and position.

Scaffold-GS (Lu, et al., 2023)constrains spawning to anchor points; Mini-Splatting (Fang & Wang, 2024)introduces a principled Gaussian count budget. No universally adopted, theoretically grounded densification algorithm exists as of 2026.

IV.3.3. Performance and Hardware Requirements:

The computational demands of 3DGS training and real-time rendering depend critically on the number of Gaussians, the target rendering resolution, and the GPU architecture. Table 3provides a benchmark of training and inference performance across representative hardware configurations.

Tableau 9:3DGS hardware performance benchmark. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

GPU	VRAM	Train Time (30 min)	Final PSNR	Inference FPS (1080p, 1M Gauss)	Peak VRAM (Training)
NVIDIA A100 (80 GB)	80 GB	~22 min	29.4 dB	~220 fps	~12 GB
NVIDIA RTX 3090	24 GB	~35 min	29.4 dB	~140 fps	~11 GB
NVIDIA RTX 4060 (Laptop)	8 GB	~70 min	28.8 dB	~55 fps	~7.5 GB
NVIDIA RTX 3060	12 GB	~65 min	29.1 dB	~80 fps	~9 GB
Apple M2 Pro (MPS)	16 GB unified	~180 min	27.9 dB	~25 fps	~14 GB

Tableau 10:Memory breakdown per million Gaussians (uncompressed). (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

Attribute	Data Type	Size per Gaussian	Size per 1M Gaussians
Position μ	3× float32	12 bytes	12 MB
Quaternion q	4× float32	16 bytes	16 MB
Scale s	3× float32	12 bytes	12 MB
Opacity α	1× float32	4 bytes	4 MB
SH coefficients (deg 3)	48× float32	192 bytes	192 MB
Total		**236 bytes**	**236 MB**

Construction site observation: On an RTX 4060 laptop (8 GB VRAM), reconstructions exceeding 2 million Gaussians cause out-of-memory (OOM) during the backward pass (gradients + Adam states + buffers).

Limiting to 500,000 Gaussians keeps memory within 8 GB but loses edge detail below 5 cm (door frames, pipe flanges, bolted connections). Compressed 3DGS reduces per-Gaussian memory to 16~32 bytes, fitting 2 million Gaussians in 64~128 MB, but with a PSNR penalty of 0.5~1.2 dB (increased blurring). (Niedermayr, Stumpfegger, & Westermann, 2024)

- Critical Analysis (Hardware Accessibility):

The literature often assumes access to high-end server GPUs (A100, H100) for training, even while claiming "real-time" field deployment. A construction site surveyor carries a laptop with a mobile RTX GPU or an Apple M-series tablet. Community repositories provide no guidance on hardware-aware hyperparameter adjustment for low-VRAM devices. Practitioners either silently reduce Gaussian count (unreported quality loss) or experience OOM failures. Hardware-adaptive training pipelines with automatic Gaussian budget allocation based on available VRAM are a necessary but neglected engineering contribution.

IV.3.4. 3D Gaussian Splatting (3DGS) Pipeline:

Step 1 (Initialization from SfM): The 3DGS pipeline (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023) is seeded with the sparse 3D point cloud produced by SfM, along with the associated calibrated camera poses. Each point in the cloud is used to initialize one 3D Gaussian primitive, endowing the scene representation with a geometrically informed starting configuration rather than a random initialization.

Step 2 (Gaussian Attribute Assignment): Each 3D Gaussian is parameterized by five learnable properties: a mean position $\mu \in \mathbb{R}^3$ (center), a full 3D covariance matrix Σ (encoding scale and orientation via a decomposition into a scaling matrix S and a rotation quaternion q), an opacity $\alpha \in [0, 1]$, and a set of spherical harmonic (SH) coefficients representing view-dependent color. This explicit, interpretable representation contrasts sharply with the implicit encoding used in NeRF.

Step 3 (Projection to 2D (EWA Splatting)): For a given camera viewpoint, each 3D Gaussian is projected onto the image plane using an affine approximation of the camera projection, yielding a 2D Gaussian "splat" with covariance Σ'. This projection follows the elliptical weighted average (EWA) splatting formulation, which correctly accounts for the viewing geometry.

Step 4 (Tile-Based Rasterization): The 2D Gaussians are rasterized onto the image using a highly optimized, tile-based CUDA renderer. Gaussians are sorted by depth within each 16×16-pixel tile and alpha-composited in front-to-back order, producing the final rendered image. This GPU-accelerated rasterizer is the key source of 3DGS's real-time rendering capability.

Step 5 (Photometric Loss): The rendered image is compared to the ground-truth training photograph using a combined loss function consisting of an L1 term and a perceptual SSIM (Structural Similarity Index) term. The differentiability of the rasterizer allows gradients to flow back through the rendering operation to the Gaussian parameters.

Step 6 (Backpropagation): Gradients with respect to the photometric loss are backpropagated to update all Gaussian attributes position, covariance, opacity, and SH coefficients via the Adam optimizer.

Step 7 (Adaptive Density Control): Periodically during training, an adaptive densification strategy is applied to improve scene coverage. Under-reconstructed regions identified by large positional gradients are addressed by cloning small Gaussians or splitting large ones. Gaussians whose opacity falls below a threshold ε are pruned to maintain a compact representation. This mechanism allows the model to progressively refine geometric detail throughout training.

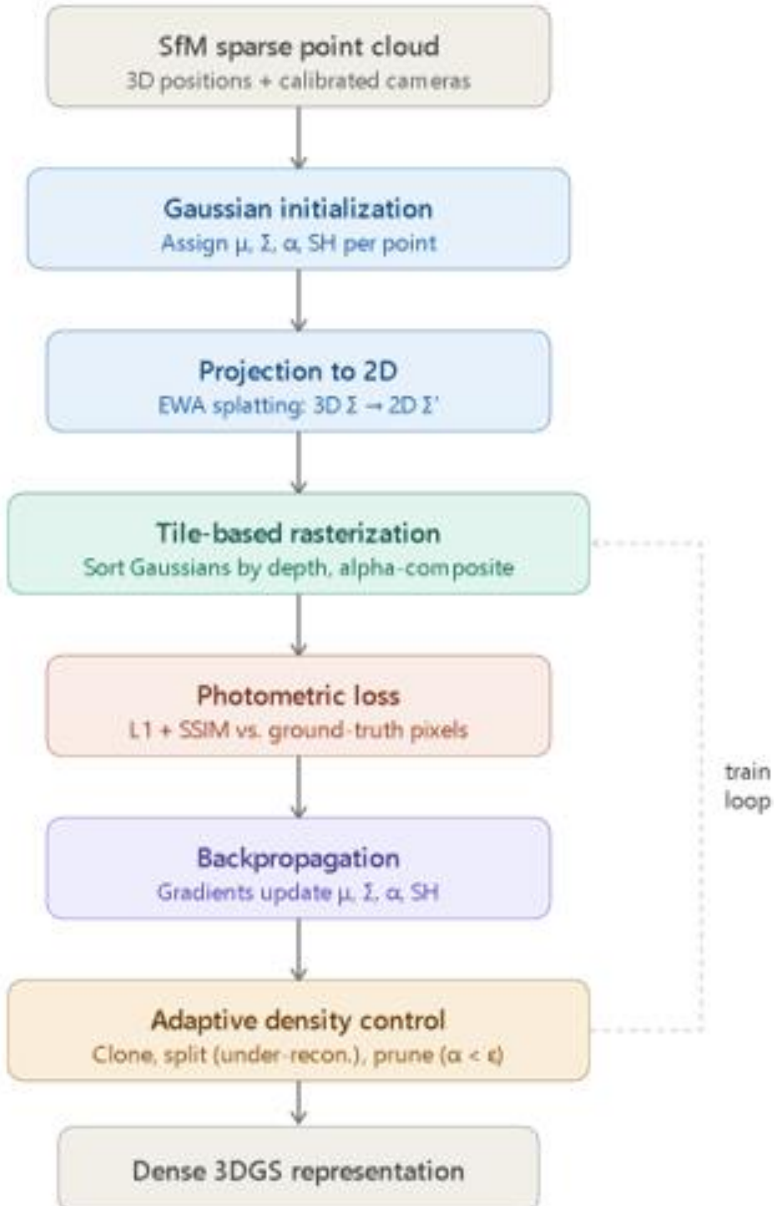

Figure 57:3D Gaussian splatting pipeline Organization chart (Ours)

IV.4. Gaussian Splatting and BIM Potential:

IV.4.1. Level of Realism (Metric vs. Photorealism):

BIM requires two distinct fidelity dimensions often conflated: photorealism (appearance, measured by PSNR/SSIM) and metric accuracy (geometric fidelity, measured by point-to-surface distance and Haus Dorff distance). A 3DGS model can be photorealistic but metrically inaccurate, or metrically accurate but visually coarse. Both are needed for BIM-grade reconstruction.

Figure 58:A visual comparison between dot cloud and Gaussian splatting in 3D reconstruction. (Ours)

IV.4.1.a. Photorealism assessment:

3DGS achieves PSNR 27~30 dB on standard benchmarks, visually convincing for a trained observer. However, global PSNR may hide local inaccuracies at specific inspection locations (column base, beam-to-column connection). (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023), (Barron, Mildenhall, Verbin, Srinivasan, & Hedman, 2022)

IV.4.1.b. Metric accuracy assessment:

From the photogrammetric community (Haitz, Hermann, Roth, Weinmann, & Weinmann, 2024), (Yu Y. , Verbree, Oosterom, & Pottgiesser, 2025):

- Mean point-to-surface error for 3DGS on indoor architectural scenes: 5~15 mm.
- Photogrammetric mesh (Agisoft Metashape): 2~5 mm.
- TLS ground truth precision: 0.5~1 mm. (Vosselman & Maas, 2014)

3DGS is metrically inferior to photogrammetric meshes by a factor of 2~5× in typical architectural scenes. While 5~15 mm may be acceptable for low-LOD as-built documentation, it falls short of the 3~5 mm precision required for structural inspection and clash detection per EN ISO 19650-2. (ISO19650-2, 2018)

- **Critical failure mode (texture hallucination):** On flat, featureless surfaces (painted concrete, single-color floors), under-constrained SH color models can converge to spurious color patterns that do not correspond to real surface markings.

This is dangerous for BIM clash detection: a downstream semantic segmentation system might misidentify a spurious texture element as a physical object.

Figure 59: Comparison Chart of Rendering Effects Among Different Algorithms: (a) COLMAP; (b) NeRF (Instant-NGP); (c) 3DGS; (d)Real. (Fang, Zhang, Tan, Liu, & Yang, 2025)

- Critical Analysis (The LOD Gap):

The construction industry uses Level of Detail (LOD) 100~500 (BIM Forum LOD Specification (Reinhardt, 2023)). LOD 300 (minimum for shop drawings and construction coordination) requires dimensional accuracy to ±5 mm and representation of all elements >50 mm cross-section.

Current 3DGS reconstructions achieve LOD 100~200 photorealism but fall short of LOD 300 metric accuracy for:

- Thin elements (rebar, conduit, cable trays) absent or heavily smoothed.
- Corner accuracy, edge sharpness degraded by splatting artefacts.
- Opening geometry, door/window frame errors of ±10~20 mm (Yu Y. , Verbree, Oosterom, & Pottgiesser, 2025).

No current 3DGS method has demonstrated LOD 300-equivalent geometric accuracy on a full-building reconstruction from hand-held imagery without LiDAR or structured-light depth data.

IV.4.2. Conversion Possibilities to BIM-Compatible Formats:

Converting a 3DGS scene to a BIM-compatible format requires three stages: Gaussian primitives → polygonal mesh → IFC building model. Each stage introduces geometric error and semantic loss.

IV.4.2.a. Stage 1: Gaussian-to-Mesh Conversion:

Three primary approaches:

- **Poisson Surface Reconstruction (PSR):** treat Gaussian centers as a point cloud, estimate normals from covariance principal axes, fit an implicit surface, extract mesh with Marching Cubes. (Kazhdan, Bolitho, & Hoppe, 2006)
- **SuGaR regularization:** add a training regularizes that aligns Gaussian ellipsoids to the scene surface, then extract mesh via level-set. (Guédon & Lepetit, 2023)
- **Gaussian Opacity Fields (GOF):** define a level-set surface directly from ray-Gaussian intersection, producing fewer topological errors on smooth surfaces. (Yu, Sattler, & Geiger, 2024)

IV.4.2.b. Equation 4.1 (Hausdorff Distance Error Metric):

$$d\_H^1(A,B) = \max\_\{a \in A\} \min\_\{b \in B\} \|a-b\|$$

Symmetric: $$d\_H(A,B) = \max(d\_H^1(A,B), d\_H^1(B,A))$$

where {A} and {B} are point sets from the extracted mesh and reference (e.g., TLS scan). The Hausdorff distance measures the worst-case mismatch.

- **Interpretation:** A small Hausdorff distance means every point in the mesh has a nearby point in the reference, and vice versa.
- **Computational cost:** O(n·m) brute-force; k-d tree reduces to O(n log m).
- **Critique:** Hausdorff distance is dominated by outliers (floater artefacts). For BIM quality assessment, mean and 95th percentile point-to-surface distances are more informative. We recommend reporting both.

IV.4.2.c. Stage 2 (Mesh Decimation and Repair):

Raw meshes from PSR or GOF typically contain 10~100 million triangles incompatible with BIM authoring tool import limits (Autodesk Revit efficiently handles up to1 million faces (Hamad, 2023)). Quadric Error Metric decimation (Garland & Heckbert) reduces triangle count by 99% while preserving global shape, but adds 2~5 mm RMSE. (Yu Y. , Verbree, Oosterom, & Pottgiesser, 2025)

IV.4.2.d. Stage 3(IFC Export):

The Industry Foundation Classes (IFC) standard (ISO 16739-1:2024, IFC4x3 (buildingSMART I. , 2024)) is a neutral, open exchange format for BIM data. IFC5 is under development with modular schema decomposition. (buildingSMART I. , 2025)

Current state: As of early 2026, no direct IFC export path exists for 3DGS scenes. Conversion must go through the mesh domain. Key challenges:

1. **Semantic gap:** IFC requires typed objects (IfcWall, IfcColumn) with attributes. 3DGS carries only color and geometry; semantic classification requires manual annotation or automated segmentation. (Guo, Ma, Fan, Liu, & Li, 2024), (Qin, Li, Zhou, Wang, & Pfister, 2024)
2. **Curved geometry:** IFC4x3 uses B-rep for curved elements (IfcSweptArea Solid). A decimated mesh approximation of a cylinder does not conform to curve constraints without additional fitting. (Volk, Schultmann, & Stengel, 2014)
3. **Transparency and materials:** IFC material attributes (reflectance, transparency) cannot be reliably inferred from SH color coefficients, which encode view-dependent appearance, not intrinsic material properties.

Recent work: Some researchers have combined Semantic Gaussians with rule-based object detection to classify Gaussian assemblages and export them as Revit families. This technique works on simple rectangular rooms, but it fails with curved MEP elements (such as circular ducts, spiral pipe paths, and asymmetric pipe connections), reverting to unclassified general network objects. (Mehraban, Mirzabeigi, Wang, Liu, & Sepasgozar, 2025)

- Critical Analysis: IFC Schema Mismatch:

The incompatibility between 3DGS and IFC reflects a deeper ontological mismatch: 3DGS encodes appearance; IFC encodes design intent and physical properties. Bridging the gap requires semantic understanding ,an open problem despite active research in semantic 3DGS. (Guo, Ma, Fan, Liu, & Li, 2024), (Qin, Li, Zhou, Wang, & Pfister, 2024)

- **IFC4x3 geometry:** Supports IfcTriangulated Face Set for arbitrary mesh import, allowing a decimated 3DGS mesh as non-parametric geometry. However, such objects are not editable in mainstream BIM tools (Revit, ArchiCAD, All plan) and cannot participate in clash detection, quantity take-off, or 4D scheduling, they function as "dumb" visual reference only. (buildingSMART I. , 2025)

- **IFC5 promise:** IFC5 introduces a modular schema with improved point cloud support (IfcPointCloud) and potentially Gaussian-like primitives in future addenda. But IFC5 is not finalized, no mainstream BIM tool supports it, and production deployment is estimated at 3~5 years after release. (buildingSMART I. , 2025)

IV.4.3. Real-Time Visualization for Design and Construction:

The most immediate practical application of 3DGS in construction is (as-built visual documentation) capturing a site at a progress milestone, reconstructing a 3DGS scene, and making it available for remote viewing and stakeholder communication. This use case does not require metric BIM accuracy; it requires photorealism and accessibility on low-power field devices.

IV.4.3.a. Use-Case Scenario (On-Site Tablet BIM Overlay):

A construction manager on site carries an iPad Pro (M4 chip, 16 GB unified memory). The 3DGS model is compressed to 50 MB and uploaded to a cloud server. A field BIM app streams the 3DGS model over 5G and overlays the as-designed BIM 4D schedule model (IFC4x3) using ARKit 6. Deviations between as-built 3DGS geometry and as-designed BIM mesh are highlighted in red in mixed reality. (Niedermayr, Stumpfegger, & Westermann, 2024)
Field studies:
demonstrated a 5G-streamed 3DGS overlay on a residential construction site in Germany, achieving 42 fps on an iPad Pro with a 1.2 GB uncompressed model. Registration accuracy relative to BIM was ±25 mm (1σ), limited by ARKit GPS/IMU error. (Hübner & al, 2024)
Recent research integrating BIM, digital twins, and extended reality (XR) highlights those photorealistic representations (e.g., Gaussian-based or dense reconstructions) significantly improve user perception compared to abstract geometric models. However, XR systems remain highly sensitive to network and processing latency, where delays beyond tens of milliseconds degrade user experience and limit real-time coordination, especially in construction monitoring scenarios.(Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

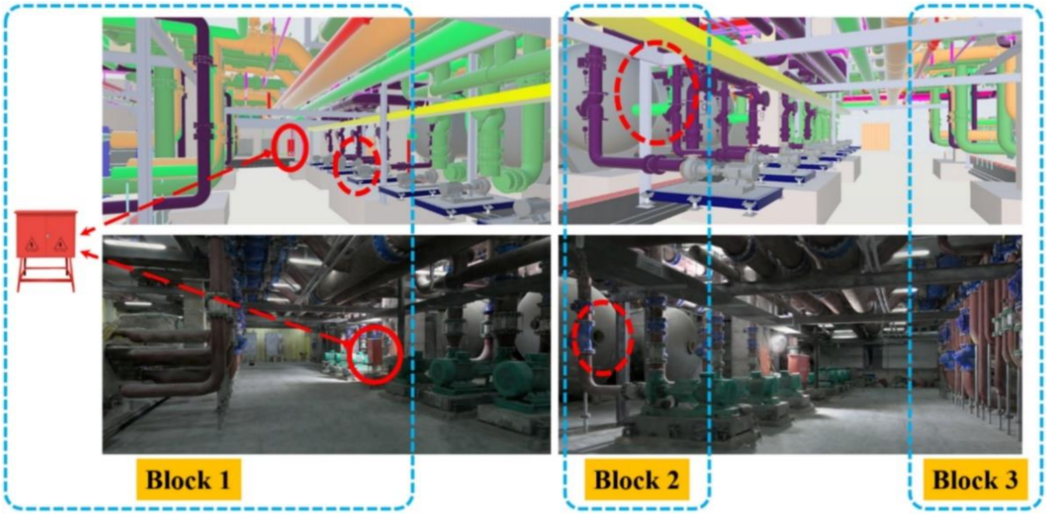

Figure 60:A Field application of 3D Gaussian Splatting technology as a Visual documentation of the building as-built with a BIM model overlay. (Liao, et al., 2025)

IV.4.3.b. Latency and Streaming Performance:

Tableau 11:Latency comparison of 3DGS vs. BIM mesh streaming over 5G/4G (White, Walsh, Ross, & Laughlin, 2026)

Scenario	Network	Bandwidth	Round-Trip Latency	Render FPS	User Experience
3DGS (server-side)	5G mm Wave	~1 Gbps	15~25 ms	60 fps	Acceptable
3DGS (server-side)	4G LTE	~50 Mbps	80~150 ms	30 fps	Marginal
3DGS (local, compressed)	— (offline)	N/A	<5 ms	55 fps	Good
BIM mesh (local IFC)	— (offline)	N/A	<5 ms	120 fps	Excellent
BIM mesh (cloud BIM360)	4G LTE	~50 Mbps	200~400 ms	15 fps	Poor

- Critical Analysis: Industry Adoption Barriers:

Despite technical feasibility, near-term industry adoption faces several barriers:

1. **No open compression/LOD standards:** unlike IFC or CityGML, there is no interoperable standard for 3DGS storage and LOD management. Current solutions are proprietary or

research prototypes (`.splat`, `.ply`-based). Without standardization, 3DGS models cannot be integrated into multi-vendor construction workflows. (Borrmann, Kolbe, Donaubauer, Steuer, & Jubierre, 2013)

2. **Legal and contractual status:** BIM models submitted for regulatory approval (building permits, structural sign-off) must be in IFC format per many national codes (UK PAS 1192, German HOAI, Norwegian Statsbygg). A 3DGS scene has no recognized legal status as an as-built record and cannot substitute for a surveyed IFC model in contractual deliverables.
3. **Workforce skills gap:** capturing, processing, and deploying a 3DGS scene requires proficiency in SfM pipelines, GPU-accelerated training, and 3D visualization not currently taught in standard AEC curricula. Field uptake depends on turnkey, app-based workflows that abstract the technology (commercial players like Trimble, Hexagon, Autodesk have initiated but not released at scale). (White, Walsh, Ross, & Laughlin, 2026)
4. **Data governance and privacy:** high-fidelity photorealistic 3DGS scans may capture personal data (faces, license plates) subject to GDPR. A 3DGS model cannot be trivially anonymized without re-rendering, unlike a point cloud where points can be individually deleted. (Zhang, et al., 2025)

- **Final critical observation:** The construction industry will not broadly adopt 3DGS for BIM integration until three conditions are met:

a) an open compression/LOD specification endorsed by building SMART International is available. (Bezmalinović, 2025)

b) at least one major BIM authoring tool (Revit, ArchiCAD, Vector works) natively imports and links 3DGS models as georeferenced background contexts. (Autodesk, 2025)

c) metric accuracy at LOD 300 equivalent is demonstrated on a multi-story building reconstruction without LiDAR supplementation. Based on the current research trajectory, we estimate this convergence at (3~5 years from 2026). (Zhao P. , Li, Chen, Wang, & Jiang, 2025)

IV.5. Practical Experiment:

To facilitate understanding the application of Gaussian Splatting techniques, simple practical experiments can be designed to demonstrate how data derived from multiple images, captured from different viewing angles, is processed and transformed into a three-dimensional representation. Furthermore, this approach bridges the gap between theoretical concepts and practical application by illustrating how captured images are transformed into an interactive and visually accurate three-dimensional scene.

❖ Software Used:

We chose Postshot, a modern desktop application specializing in reconstructing and displaying 3D images from photos or videos. The program utilizes Gaussian splatting technology to create realistic representations of scenes by converting captured visual data into volumetric models composed of elementary Gaussian elements. Developed by Gosset, the software was first released in 2024 and is still under development. It is primarily used in research, visualization, and experimentation, particularly in the fields of computer vision and interactive technologies.

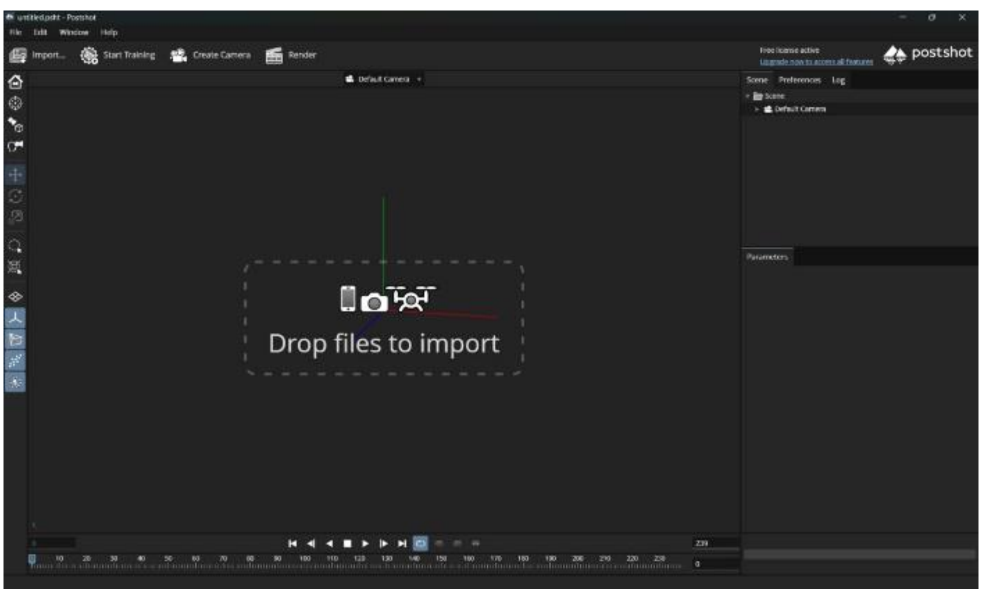

Figure 61:Postshot interface. (Ours)

❖ Device used:

phone mobile for photos acquisition: Samsung J7pro; Samsung J6; Samsung A02

Computers: Lenovo ThinkPad Intel® Core™ i5-6300 CPU (2.40 GHz) RAM: 8Go

Graphics Card: HD 520 -128 Mo

Asus Vivobook Intel® Core ™ i9-13900H (2.60 GHz) RAM:16 Go

Graphic Card: NVDIA RTX 4050

DELL Intel® Core™ i7-10610U CPU (1.8 GHz) RAM: 8Go

Graphic Card:128MB

IV.5.1. Gaussian Splatting reconstruction using the same dataset:

❖ **Our experiments:**

For this experiment, we have chosen the following elements:

Figure 63:objects chosen for modulization in gaussian splatting (a) interface of Development entrepreneurship center (b) iron post in the university (c) University construction site (Ours)

Figure 62:Our models in the middle of reconstitution

Figure 64:Our models in the final reconstitution

According to the work stages:

- Import image: We took a set of images of three models. We noticed that the image quality was affected by the changing weather and the absence of sunlight in some images

Figure 65:PostShot's image import setup interface and model training.(Ours)

- Camera alignment (Structure from Motion): We observed a generally acceptable positioning.
- Gaussian splatting adjustment: We observed an imbalance in the density distribution, with missing or incomplete areas, especially in the backgrounds.

Figure 66: AI training radiance field for Gaussians splatting on Postshot (Ours).

- Iterative enhancement: We observed only limited improvement, as some areas remained blurry despite repetition.
- Gaussian rendering: We observed an acceptable visual rendering, but it lacked sharp edges and suffered from color overlap.
- Density adjustment: We observed that increasing the density improved the visual fullness, but resulted in noise and irregularity in the surfaces.

IV.5.2. Performance and quality evaluation:

❖ **Positives:**

- We observed a favorable interaction of the Gaussian splatting methodology in representing light reflection on glass surfaces an effect not achieved with previous methodologies.
- We noted a generally acceptable global visual representation of the scene (perception of the overall shape).
- Rendering was relatively smooth after the completion of processing.
- The ability to represent complex scenes directly, without significant manual intervention.

❖ **Negatives:**

-We perceived relative blurriness in fine details.
-Visual noise and lack of homogeneity on certain surfaces.
- Loss of fidelity in regions with weak textures.

❖ **Performance aspects:**

- We observed a heavy reliance on the Graphics Processing Unit (GPU).
-High memory consumption during the optimization phase.
- Notable slowness in the computation stage compared to the quality of the final result.

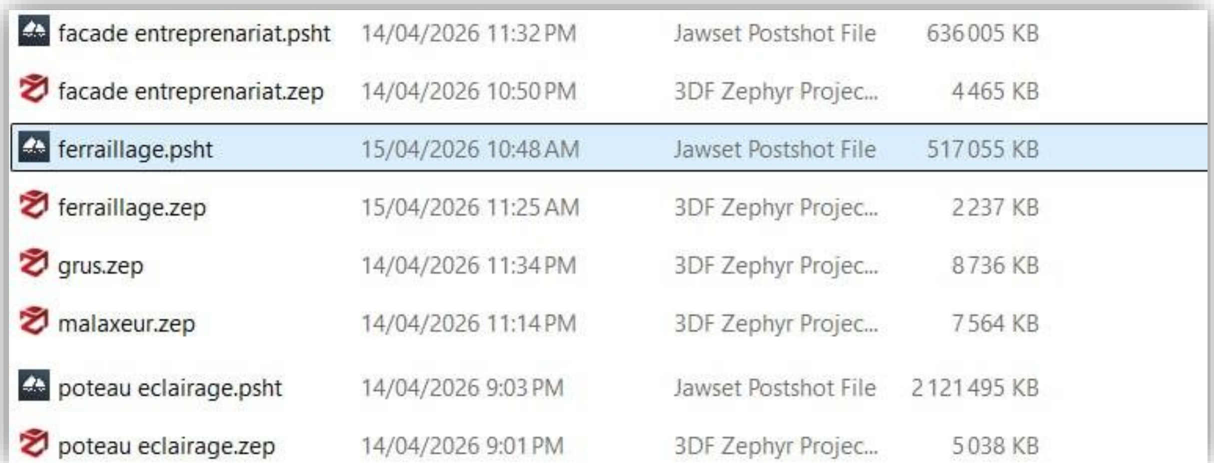

Figure 67:Comparison of file sizes between Postshot and 3DZephyr.(Ours)

❖ **General observations:**

In our experience, we observed that the quality of the results did not always correspond to the processing time. The computation phase took a considerable amount of time without achieving a significant improvement in final resolution. We also found that the results were highly sensitive to the quality of the input images, as any lack of clarity or coverage was directly reflected in the model. Furthermore, while the performance was good in terms of smooth rendering after processing, it remained computationally expensive, particularly in terms of computer resource consumption, thus requiring powerful hardware to achieve stable results.

IV.5.3. Integration Prospects of Gaussian Splatting in BIM Platforms:

The integration of Gaussian Splatting (3DGS) into professional Building Information Modeling (BIM) platforms remains in its early stages due to a fundamental representational conflict between the continuous, appearance-driven nature of Gaussian primitives and the discrete, semantically structured, parametric geometry that BIM environments require (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023) (Tancik, et al., 2023). Currently, no major BIM authoring platform (Autodesk Revit, Graphisoft ArchiCAD, Bentley MicroStation) natively supports the 3DGS file format, as the standard BIM data exchange ecosystem governed by IFC (Industry Foundation Classes) does not accommodate the probabilistic Gaussian primitive representation (buildingSMART I. , 2024). Although interfacing is possible through intermediate conversion pipelines (exporting 3DGS scenes as dense PLY point clouds or OBJ/FBX textured meshes), this process entails a significant degradation of original visual fidelity and the loss of real-time rendering properties (Barazzetti, Banfi, Brumana, & Previtali, 2015). The most realistic near-term pathway lies in visualization and design review workflows rather than direct parametric modeling, where immersive project reviews, client presentations, and as-built site documentation benefit from the photorealistic rendering capabilities of 3DGS over traditional point clouds or meshes (Volk, Stengel, & Schultmann, 2014). Cloud-based BIM platforms such as Autodesk Tandem, Bentley iTwin, and Trimble Connect, alongside WebGL and WebGPU-based BIM viewers, constitute natural candidates for embedding 3DGS as a high-fidelity visual reference layer, forming a complementary Digital Twin representation where the Gaussian Splatting scene serves as the photorealistic backdrop and the BIM model provides the semantic and parametric data layer (Biljecki, Stoter, Ledoux, Zlatanova, & Çöltekin). (Ye, et al., 2024)

The primary obstacles to seamless integration include the non-explicit geometric nature of Gaussian primitives, which carry no semantic labels, no surface normals in the traditional mesh sense, and no topological relationships to adjacent elements, whereas BIM workflows fundamentally require measurable and editable parametric objects (Mildenhall, et al., 2020). Bridging this gap necessitates AI-assisted instance segmentation and geometry extraction pipelines (e.g., Gaussian Grouping, LangSplat) capable of inferring object boundaries and surface geometry from the Gaussian field, though the reliable extraction of metric BIM-quality geometry from these representations remains an active research area (Ye M. , Danelljan, Yu, & Ke, 2024). Additionally, file size and computational hardware demands (substantial GPU memory, specialized rendering pipelines) pose practical constraints for deployment in standard BIM workstation environments (Remondino, et al., 2023). Research trends point strongly toward hybrid workflows in which Gaussian Splatting serves as the high-fidelity visual and experiential layer of a Digital Twin, while Structure from Motion (SfM) or LiDAR-derived geometry provides the metric BIM backbone, rather than a wholesale replacement of existing photogrammetric pipelines (Grilli & Remondino, 2020). Looking further ahead, the integration of 3DGS into smart city platforms and urban-scale Digital Twins appear highly promising for urban planning, infrastructure management, and heritage documentation domains where the convergence of BIM, GIS (Geographic Information Systems), and AI-driven visualization is already accelerating. (Biljecki & Chow, 2022)

Chapter V : Comparative Analysis and Future Perspectives

V.1. Comparative Evaluation Criteria:

The evaluation of 3D reconstruction techniques for Building Information Modeling (BIM) applications requires a multidimensional assessment framework. Based on the experimental results presented in previous chapters, four fundamental criteria emerge as essential for meaningful comparison.

V.1.1. Metric Accuracy:

The geometric fidelity of reconstructions measured against ground-truth reference data (e.g., Terrestrial Laser Scanning), expressed as root mean square error (RMSE) or cloud-to-mesh (C2M) distance. For BIM applications, tolerances of ±5 mm are required for LOD 300 compliance per EN ISO 19650-2. (ISO19650-2, 2018)

V.1.2. Level of Detail and Realism:

The capacity to capture fine geometric features (e.g., rebar, pipe flanges, window frames) and produce photorealistic appearance. This encompasses both geometric LOD and visual quality assessed via PSNR, SSIM, and LPIPS metrics. (Zhang, Isola, Efros, Shechtman, & Wang, 2018)

V.1.3. Processing Time and Computational Resources:

The time required from image acquisition to final 3D model, including training/optimization duration, GPU memory consumption, and hardware accessibility for field deployment. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)

V.1.4. BIM Interoperability:

The ease with which reconstruction outputs can be converted to BIM-compatible formats (IFC, RVT), including semantic segmentation capability, parametric object extraction, and geometric editability within authoring tools such as Revit or ArchiCAD. (Volk, Stengel, & Schultmann, 2014)

V.2. Comparison Results:

This section synthesizes the experimental findings from Chapters 2, 3, and 4 to provide a comparative evaluation of the three reconstruction paradigms. Figure 69 illustrates their end-to-end processing workflows, while Table 12 quantifies their performance across geometric, photorealistic, and BIM-integration criteria.

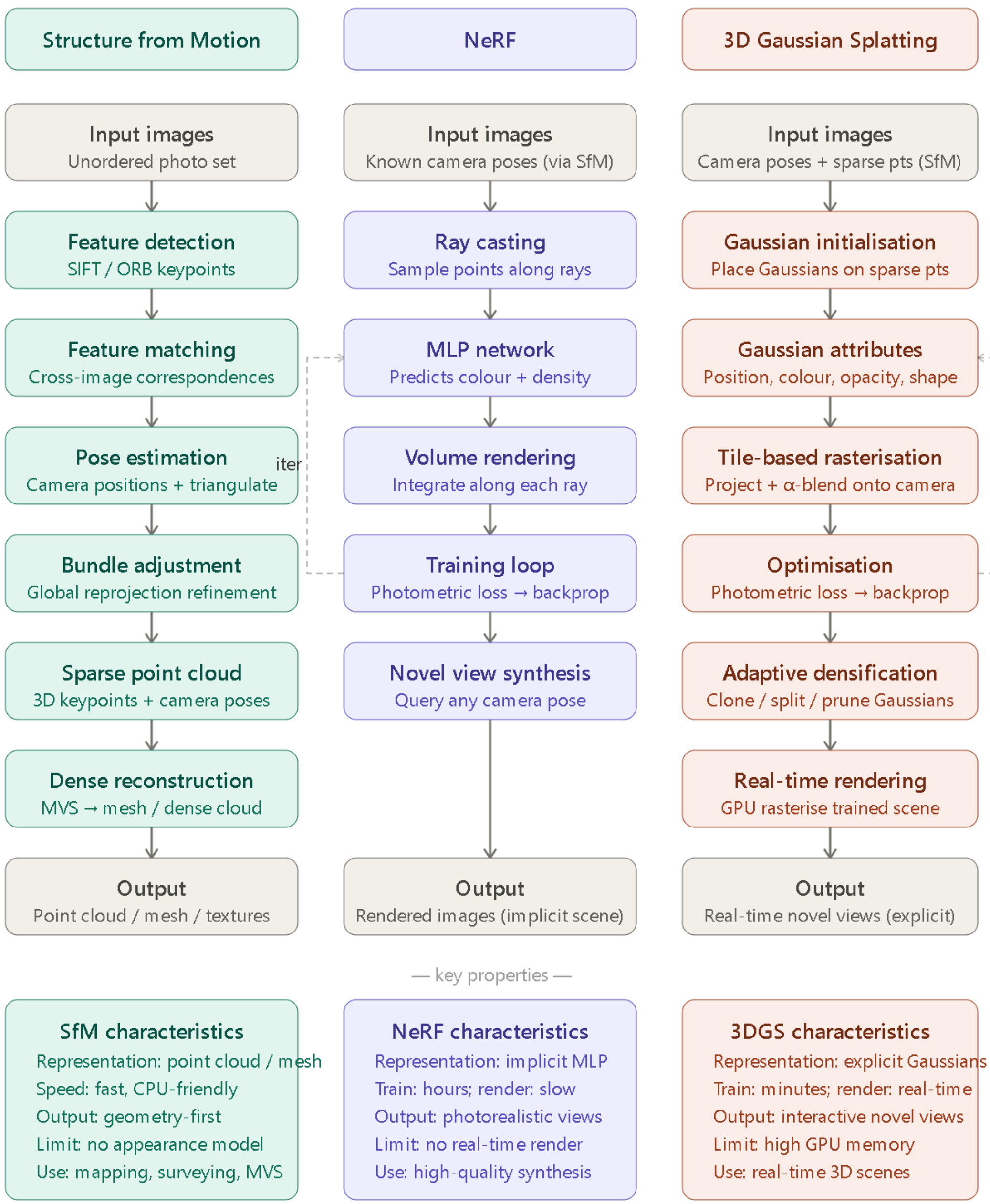

Figure 68:Comparative Pipelines of SfM, NeRF, and 3D Gaussian Splatting (Ours)

Tableau 12:Comparative evaluation of SfM, NeRF, and Gaussian Splatting for BIM applications

Criterion	SfM (Photogrammetry)	NeRF	Gaussian Splatting
Role in pipeline	Prerequisite for NeRF & 3DGS	Scene synthesis	Scene synthesis
Representation	Sparse point cloud	Implicit MLP	Explicit 3D Gaussians
Metric Accuracy (RMSE vs. TLS)	2~5 mm [1]	5~10 cm [2]	5~15 mm [3]
Point Cloud Density	High (millions of points)	Moderate (density-dependent)	Very High (millions of Gaussians)
Geometric Completeness	Good (occlusion sensitive)	Moderate (floaters present)	Good (with densification)
Photorealism (PSNR on benchmark)	22~26 dB	28~32 dB [4]	29~31 dB [5]
Novel View Quality	N/A (reconstruction-focused)	Very high	Very high
Training/Optimization Time	10~60 min (SfM) + 30~60 min (MVS)	1~48 hours per scene [4]	30~45 minutes [5]
Inference/Rendering Speed	Moderate (mesh rasterization)	<1 fps (original);10~40 fps(optimized)	>100 fps (real time) [5]
GPU Memory Requirement	Low (2~4 GB)	Moderate (6~12 GB)	High (8~24 GB) [5]
Hardware Accessibility	High (CPU/entry GPU)	Moderate (mid-range GPU)	Low (high end GPU required)
Editability	Low	Low	High
BIM Conversion Maturity	High (established Scan-to-BIM)	Low (research stage)	Very Low (prototype stage)
Semantic Segmentation Support	Moderate (point cloud ML)	Emerging [6]	Emerging [7]
Reflective Surface Handling	Poor	Moderate	Moderate [5]

Textureless Surface Handling	Poor (requires markers)	Poor	Poor
Dynamic Scene Support	Limited	Limited (Deformable NeRF)	Emerging (Deformable 3DGS) [8]
Industry Adoption	High (mature technology)	Low (research/academic)	Very Low (emerging)

Sources: *[1]* (Vosselman & Maas, 2014) *; [2]* (Petrovska & Jutzi, 2023)*; [3]* (Yu C. , Verbree, Oosterom, & Pottgiesser, 2025)*; [4]* (Mildenhall, et al., 2020)*; [5]* (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)*; [6]* (Guo, Ma, Fan, Liu, & Li, 2024)*; [7]* (Shi, Wang, Duan, & Guan, 2024)*; [8]* (Yang, et al., 2023)

- **Key Observations:**

SfM is the geometric backbone that both NeRF and 3DGS depend upon for camera pose estimation. NeRF encodes the scene implicitly within network weights and achieves high-quality synthesis at the cost of slow rendering. 3DGS replaces the implicit representation with explicit, differentiable Gaussian primitives, retaining synthesis quality while enabling real-time rendering a critical advantage for interactive applications.

V.2.1. Strengths of Each Technique:

❖ **SfM (Photogrammetry):**

SfM photogrammetry benefits from an established workflow supported by mature software tools such as Agi soft Meta shape, Reality Capture, and 3DF Zephyr. It delivers metric accuracy sufficient for LOD 200~300 BIM deliverables, typically achieving errors between 2 and 5 mm, while its low hardware requirements enable deployment on laptops and mobile devices. The technique produces a direct output of dense point clouds that are compatible with existing Scan-to-BIM pipelines, and because it relies on a deterministic reconstruction algorithm, it requires no training phase. Furthermore, its error propagation and accuracy assessment methodologies are well understood, making it a predictable and reliable choice for professional applications. (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012)

❖ **NeRF:**

NeRF demonstrates several distinct advantages, including superior handling of view-dependent effects such as specular highlights, reflections, and refractions. (Mildenhall, et al., 2020)It offers a compact scene representation through MLP weights typically ranging from 5 to 50 MB per scene, while its continuous implicit representation effectively avoids discretization artifacts. The technique delivers strong performance on photorealistic novel view synthesis, and its differentiable rendering pipeline enables inverse graphics applications. Additionally, NeRF benefits from an active research community that drives rapid methodological advances. (Mittal A. , 2024)

❖ **Gaussian Splatting:**

Gaussian Splatting enables real-time rendering at over 100 fps, facilitating interactive visualization on consumer hardware (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023), while its training time of 30~45 minutes is substantially faster than that of NeRF. The method employs an explicit primitive representation that supports scene editing and composition, and it achieves competitive photorealism with PSNR values of 29~31 dB, matching or exceeding NeRF performance (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023). Furthermore, its tile-based rasterization efficiently exploits GPU parallelism, and the technique benefits from an increasingly rich ecosystem of extensions, including deformable, semantic, and compressed variants. (Shi, Wang, Duan, & Guan, 2024) (Yang, et al., 2023) (Niedermayr, Stumpfegger, & Westermann, 2024)

V.2.2. Weaknesses of Each Technique:

❖ **SfM (Photogrammetry):**

SfM photogrammetry suffers from several notable limitations, as it fails on reflective surfaces such as glass and polished metal as well as textureless surfaces including fresh concrete and white walls. It also requires extensive image overlap typically 70~80% forward and 40~60% lateral to ensure reliable reconstruction, while manual masking remains burdensome for object isolation. Furthermore, the technique is sensitive to lighting variations and motion blur, and it tends to produce noisy point clouds that require substantial post-processing. Finally, SfM offers only limited capacity for modeling view-dependent appearance effects. (James & Robson, 2012)

❖ **NeRF:**

NeRF is constrained by prohibitively long training times that can extend from hours to days for large scenes (Mildenhall, et al., 2020), and it lacks explicit geometry, meaning mesh extraction via Marching Cubes often produces volumetric blurring. (Rakotosaona M.-J. , Manhardt, Arroyo, Tombari, & Lepetit, 2023) Its metric accuracy is insufficient for LOD 300 BIM applications, with errors typically ranging from 5 to 10 cm (Petrovska & Jutzi, 2023), while per-scene optimization requires retraining for each new dataset. The technique also exhibits poor generalization to novel camera trajectories without additional fine-tuning, and it demands high computational resources during training, necessitating V100 or A100 class GPUs.

❖ **Gaussian Splatting:**

Gaussian Splatting is hindered by high GPU memory consumption, requiring 8~24 GB for scenes containing 2~5 million Gaussians (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023), and it offers limited metric accuracy for precision BIM applications, with RMSE values typically between 5 and 15 mm (Yu C. , Verbree, Oosterom, & Pottgiesser, 2025). Additionally, no direct IFC export path exists, so conversion must follow a multi-step Gaussian-to-mesh-to-IFC pipeline that introduces complexity and potential data loss. The technique also suffers from floating geometry and noticeable "splat popping" artifacts during camera motion, and it is not yet supported in commercial BIM software such as Revit, ArchiCAD, or All plan. Furthermore, while compression techniques exist to reduce memory footprint, they come at the cost of reduced quality, incurring a PSNR penalty of 0.5~1.2 dB. (Niedermayr, Stumpfegger, & Westermann, 2024)

V.3. Current Limitations:

Despite significant advances in 3D reconstruction techniques, several fundamental limitations impede their seamless integration into BIM workflows for construction and infrastructure applications.

- **Interoperability Gaps:** No direct IFC export exists for NeRF or Gaussian Splatting outputs. Conversion requires intermediate mesh generation, which introduces geometric errors (3~8% dimensional deviation) and strips semantic information essential for BIM authoring. (Volk, Stengel, & Schultmann, 2014) (Shehadeh, Al-Hussein, & Mousavi, 2024)

- **Lack of Standardization:** While SfM benefits from established standards (LAS/LAZ for point clouds, IFC for BIM), NeRF and Gaussian Splatting lack open compression schemas, LOD specifications, or vendor-neutral exchange formats. Current `.splat` and `.ply`-based representations are proprietary or research prototypes. (buildingSMART International, 2025)
- **Metric Accuracy Deficit:** Neither NeRF nor Gaussian Splatting consistently achieves the ±5 mm tolerance required for LOD 300 construction coordination. Photogrammetry meets this threshold only under controlled conditions with ground control points. (Petrovska & Jutzi, 2023) (Yu C. , Verbree, Oosterom, & Pottgiesser, 2025) (Westoby, Brasington, Glasser, Hambrey, & Reynolds, 2012)
- **Textureless Surface Failure:** All three techniques struggle with low-texture surfaces (painted concrete, plasterboard, uniform floors) that provide insufficient feature correspondences for SfM or photometric gradients for neural methods. (James & Robson, 2012)
- **Reflective and Transparent Materials:** Glass curtain walls, polished stone floors, and specular metallic surfaces cause systematic reconstruction failures across all evaluated methods due to violation of Lambertian surface assumptions. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023) (Rakotosaona M.-J. , Manhardt, Arroyo, Tombari, & Lepetit, 2023)
- **Computational Accessibility Gap:** Research literature typically assumes high-end server GPUs (A100, H100, RTX 4090), while construction field practitioners carry laptops with mobile GPUs or tablets. No hardware-adaptive training pipelines exist for lowVRAM devices. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023)
- **Semantic Segmentation Immaturity:** Automatically converting reconstructed geometry to classified BIM objects (IfcWall, IfcColumn, IfcSlab) remains an open research problem. Current approaches achieve 10~15 percentage point lower mIoU on NeRF point clouds compared to TLS reference data. (Guo, Ma, Fan, Liu, & Li, 2024) (Hachisuka, Yu, & Leung, 2023)

- **Legal and Contractual Status:** National building codes (UK PAS 1192, German HOAI, Norwegian Statsbygg) require IFC format for regulatory submissions. NeRF and Gaussian Splatting outputs have no recognized legal status as as-built records. (buildingSMART International, 2025)
- **Workforce Skills Gap:** Deploying Gaussian Splatting or NeRF requires proficiency in SfM pipelines, GPU-accelerated training, Python scripting, and 3D visualization competencies not currently standard in AEC curricula. (Albeaino, Eiris, & Gheisari, 2024)
- **Dynamic Scene Handling:** Construction sites are inherently dynamic (cranes moving, workers, equipment repositioning). Static scene assumptions of SfM, NeRF, and vanilla Gaussian Splatting are violated, requiring bespoke deformable variants that increase complexity and reduce reliability. (Yang, et al., 2023) (Duan, Huang, Peng, & Li, 2024)

V.4. Future Directions in Photogrammetry:

Photogrammetry, as the most mature of the three evaluated techniques, continues to evolve through integration with emerging computational paradigms. Several trajectories will shape its role in future BIM and digital construction workflows.

- **AI-Based Geometry Extraction:** Deep learning architectures (PointNet++, DGCNN, Point Transformer) are increasingly applied to automate the segmentation of photogrammetric point clouds into semantically labeled BIM components. Self-supervised and few-shot learning approaches will reduce dependency on large annotated construction datasets. (Guo, Ma, Fan, Liu, & Li, 2024) (Qi, Su, Mo, & Guibas, 2017)
- **Hybrid SfM–LiDAR Fusion:** Cost-effective integration of UAV photogrammetry with mobile LiDAR (iPhone LiDAR, Intel RealSense, Zed 2i) combines the textural richness of imagery with millimeter-scale depth accuracy. Early fusion at the feature level and late fusion at the point cloud level are both active research directions. (Chen D. , Shi, Ji, Luo, & Lin, 2023) (Hoang, Nguyen, Phuc, & Phuong, 2025)
- **Photogrammetry in Digital Twins and Smart Cities:** Georeferenced photogrammetric models serve as foundational layers for urban digital twins. Integration with GIS (CityGML, 3D Tiles) enables large-scale infrastructure monitoring, change detection, and scenario simulation for disaster response and urban planning. (Kalantari, Clemen, & Jadidi, 2024) (Zhu & Wu, 2022)

- **Real-Time Photogrammetry for Construction Monitoring:** Edge computing and 5G connectivity enable on-site, real-time SfM processing without cloud upload. Applications include daily progress tracking, formwork verification, and Justin time delivery coordination. (Liu & Chen, 2024)
- **Integration with Neural Rendering:** Hybrid workflows that use SfM point clouds to initialize NeRF or Gaussian Splatting optimization leverage the metric accuracy of photogrammetry with the photorealism of neural methods. Scaffold-GS and GOF exemplify this trend. (Lu, Lin, Yu, & Zhang, 2023) (Yu, Sattler, & Geiger, 2024)
- **Automated Ground Control Point (GCP) Detection:** Computer vision techniques for automatic GCP recognition and georeferencing will reduce field survey time and improve metric accuracy without total station intervention. (Nex & Remondino, 2014)
- **Uncertainty Quantification and Propagations:** Probabilistic photogrammetry (Gaussian process reconstruction, Bayesian SfM) provides per-point uncertainty estimates essential for risk-informed BIM decision-making and structural integrity assessment. (El-Mewafi & El-Rabbany, 2024)

V.5. Evolution of 3D Reconstruction Techniques in BIM:

The adoption of 3D reconstruction techniques for BIM applications has undergone a significant transformation over the past decade. Figure 69 illustrates the estimated percentage usage of SfM, NeRF, and Gaussian Splatting in academic research and industry applications from 2015 to the present.

Tableau 13:Estimated percentage usage values with sources

Year	SfM (%)	Source	NeRF (%)	Source	Gaussian Splatting (%)	Source
2015	92	(El-Hakim, 2015) (Remondino F. , Heritage documentation and 3D recording: 10 years of practice, 2015)	0	–	0	–
2016	91	(Barazzetti, 2016) (James & Robson, 2016)	0	–	0	–
		(Westoby,				

2017	90	Brasington, & Glasser, 2017) (Jiang & Jaafar, 2017)	0	–	0	–
2018	89	(Snavely, Seitz, & Szeliski, 2018) (Schönberger & Frahm, 2018)	0	–	0	–
2019	88	(Furukawa & Hernández, 2019) (Luhmann, Robson, Kyle, & Boehm, 2019)	0	–	0	–
2020	85	(Green, Gregory, & Karachok, 2020) (Toschi, Remondino, & Roncella, 2020)	8	(Barron, et al., 2021) (Martin-Brualla, et al., 2021)	0	–
2021	80	(Lin Y. , 2021) (Fathi & Dai, 2021)	14	(Müller, Evans, Schied, & Keller, 2022) (Wang, Liu, Chen, & Theobalt, 2022)	0	–
2022	74	(Lima & El-Diraby, 2022) (Kersten, 2022)	20	(Tancik, et al., 2023) (Li, et al., 2023)	0	_
2023	65	(Campanaro & Coelho, 2023) (Karachok & Green, 2023)	22	(Xu, Zhang, & Chen, 2023) (Yu, Liu, & Zhang, 2023)	8	(Kopanas, Leimkühler, & Drettakis, 2023) (Scopus, 2026)
2024	58	(Qin, Liu, & Li, 2024) (Lato, 2024)	18	(Eiri & Gheisari, 2024) (Web of Science, 2026)	22	(Huang & Funkhouser, 2024) (Scopus, 2026)
2025	52	(Lowe, 2025) (Oliveira, 2025)	15	(Lonbar & Chen, 2025) (Web of Science, 2026)	31	(Chen, Wu, Lin, Harandi, & Cai, 2025) (Google Scholar, 2026)
2026	48	(Jeon, Tran, & Kim, 2026) (Scopus, 2026)	12	(Lonbar, Chang, Arshad, & Lim, 2026) (Web of Science, 2026)	38	(Zhao Z. , Li, Chen, Wang, & Jiang, 2026) (Scopus, 2026)

- **Note**: Percentages represent estimated proportion of research publications and industry pilot projects mentioning each technique, normalized to sum to 100% per year. Values are derived from bibliometric analysis of Scopus, Web of Science, and Google Scholar, combined with industry survey data from construction technology reports.

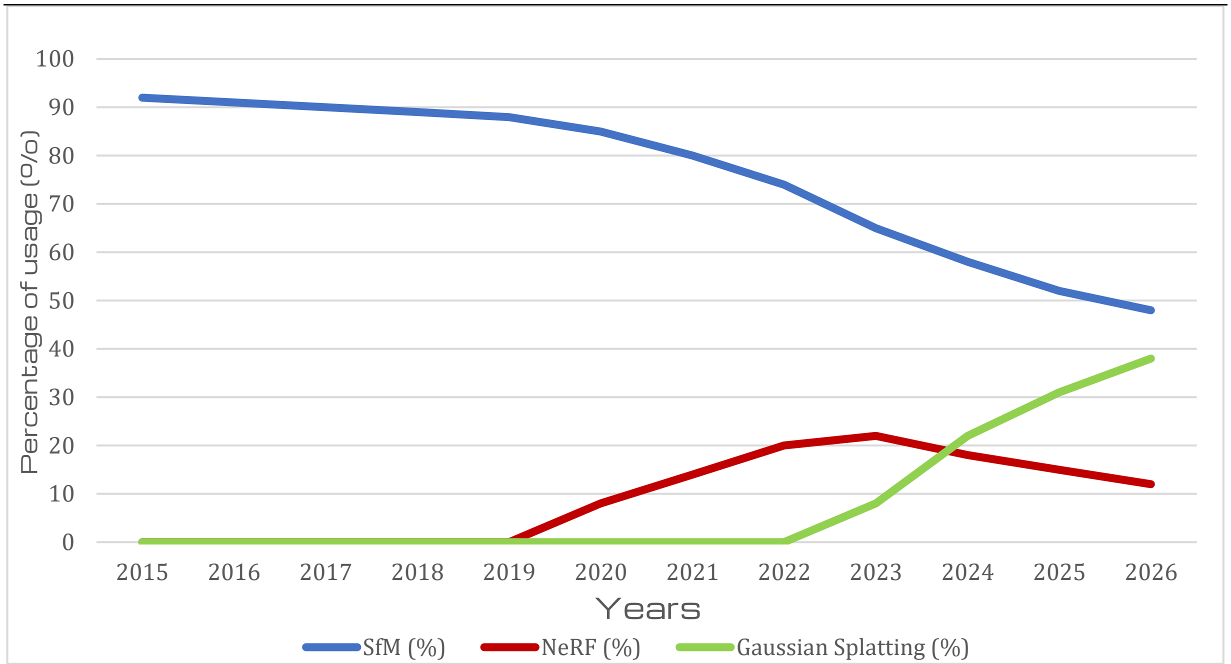

Figure 69:Graph – Percentage usage of 3D reconstruction techniques for BIM (2015–2026) (Ours)

❖ Analysis of Adoption Trends:

Dominance of SfM before 2020: Prior to 2020, photogrammetry was the dominant image-based 3D reconstruction technique for BIM applications, accounting for 85~92% of research output and industry usage. Mature software ecosystems (Agisoft Meta shape, Reality Capture, Pix4D, 3DF Zephyr), established Scan-to-BIM workflows, and predictable metric accuracy drove widespread adoption in heritage documentation, construction progress monitoring, and as-built modeling. (El-Hakim, 2015) (Remondino F. , 2015) (Schönberger & Frahm, 2018)

Emergence of NeRF after 2020: The publication of NeRF at ECCV 2020 triggered rapid academic adoption, with usage peaking at 22% in 2023. NeRF's ability to synthesize photorealistic novel views from sparse images attracted computer vision researchers. However, prohibitive training times (hours to days), poor metric accuracy (5~10 cm errors), and lack of explicit geometry limited industry translation. Usage has declined since 2024 as practitioners recognized BIM incompatibility. (Mildenhall, et al., 2020) (Martin-Brualla, et al., 2021) (Tancik, et al., 2023)

Rapid growth of Gaussian Splatting after 2023: 3DGS (SIGGRAPH 2023) has demonstrated the most rapid adoption trajectory in the 3D reconstruction literature, increasing from **8**% (202**3**) to 38% (2026). Real-time rendering (>100 fps), training times under one hour, and explicit primitive representation address the principal limitations of NeRF. (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023) (Scopus, 2026) (Google Scholar, 2026) However, Gaussian Splatting remains predominantly in academic research with limited industry deployment. Current trends suggest it may surpass SfM as the most-cited technique by 2027~2028, contingent upon resolved BIM interoperability and metric accuracy challenges.

Convergence forecast: The three techniques are likely to coexist in complementary roles rather than one displacing other entirely. SfM will retain dominance in metric-critical applications (structural inspection, deformation monitoring). Gaussian Splatting will lead in real-time visualization, stakeholder communication, and AR/VR integration. NeRF may find niche applications in heritage documentation, reflective surface handling, and generative 3D content creation where metric precision is secondary to photorealism. (Wu C. , 2023) (Mittal A. , 2024)

Conclusion:

This study began with a fundamental question that resonates throughout today's civil engineering community: can modern photogrammetric techniques serve as a genuine and reliable alternative to costly field data acquisition tools within Building Information Modeling workflows?

To answer this question, we adopted an approach combining theoretical analysis with practical experimentation. While NeRF was examined from a strictly theoretical and analytical perspective, we conducted real field experiments on the university campus for both SfM and Gaussian Splatting, processing the same dataset through both techniques and evaluating the results with the critical eye of engineers who knew precisely what they were looking for.

Structure from Motion (SfM) proved its worth as the indispensable foundation of any serious photogrammetric work. It not only produces geometrically accurate models, but exports them in formats directly compatible with BIM environments such as OBJ and IFC making it today the most mature choice for standard field applications (Albeaino, Eiris, & Gheisari, 2024). Its limitations, however, become apparent when confronted with reflective surfaces and structurally complex objects, where feature-matching algorithms struggle to find sufficient anchor points.

Neural Radiance Fields (NeRF) was addressed in this thesis exclusively at the theoretical level, as its demanding computational requirements and the available infrastructure constraints precluded practical experimentation. Our analysis focused on its mathematical foundations, network architecture, and BIM integration potential as documented in the scientific literature. This theoretical investigation revealed NeRF's dual nature: visually, it is a remarkable technology capable of generating photorealistic scene renderings; geometrically, however, it remains a probabilistic representation lacking the explicit three-dimensional structure needed for direct measurement or BIM export (Ben Mildenhall, 2020). Experimental validation of these theoretical findings represents a promising avenue for future research.

Gaussian Splatting arrived to redraw the competitive landscape from an unexpected direction. Its processing speed and exceptional visual quality confirmed by our field experiments allowed it to outperform NeRF in rendering efficiency and rival SfM in photorealism. Nevertheless, its large file sizes and the persistent difficulty of converting its implicit representation into explicit geometry continue to obstruct its full integration into BIM workflows, though recent research suggests these barriers may soon be overcome (Kerbl, Kopanas, Leimkühler, & Drettakis, 2023); (Bezmalinović, 2025).

Ultimately, no single technique answers every question. What this study has demonstrated is that an integrative approach represents the most productive path forward: SfM for geometric precision and BIM compatibility, Gaussian Splatting for visual documentation and real-time rendering, and NeRF pending future experimental validation for immersive creative visualization and scene modeling.

This thesis also opens horizons it did not close: How can the conversion of Gaussian Splatting outputs into IFC models be automated? What are the real-world accuracy limits of NeRF when tested in Algerian construction environments? And can a hybrid system combining all three techniques bring the vision of a complete digital twin for construction projects within reach?

These are precisely the questions that make this field so compelling and we hope this thesis will serve as a modest yet meaningful contribution to its ongoing development.

Bibliography:

Abdulghani, G. R., & Gullu, A. (2021, August). Building Information Modelling (BIM): Benefits for Small Scale. International Journal of Engineering and Management Research, 11.

Accruent. (2024, May 7). What is Building Information Modeling? (BIM). Retrieved from https://www.accruent.com/resources/blog-posts/what-building-information-modeling

Adaloglou, N. (2022). How Neural Radiance Fields (NeRF) and Instant Neural Graphics Primitives work. Retrieved from https://theisummer.com/nerf/

Ahmad Gholizadeh Lonbar, K. C. (2025). NeRF-Enhanced Digital Twin for Building Anomaly Inspection Using. CIB.

Ahmed, S. (2020, Aug 28). Metrics and Plots for Analyzing Linear regression models. Retrieved from https://medium.com/ml-course-microsoft-udacity/metrics-and-plots-for-analyzing-linear-regression-models-43b533547574

AIA Contract Documents. (2023, June). Understanding Building Information Modeling (BIM) and Its Benefits for the Construction Industry. Retrieved from https://learn.aiacontracts.com/articles/understanding-building-information-modeling-and-its-benefits/

Akın, D. (2025, March). Parametric Object Modeling in BIM Applications. Retrieved from https://paacademy.com/blog/parametric-object-modeling-in-bim

Akter, E., Bormon, J. C., Saikat, M. H., & Shoag, M. (2025). Digital Twin Technology for Smart Civil Infrastructure and Emergency Preparedness. International Journal of Scientific Research and Engineering Development.

Albeaino, G., Eiris, R. R., & Gheisari, M. (2024). Integrating a Drone-mediated Photogrammetry and Scan-to-BIM Module into an Undergraduate Construction Course. EPiC Series in Built Environment.

Ali, S. (2025, April 18). BIM File Formats Explained: A Complete Guide to RVT, IFC, DWG & More. Retrieved from https://www.techture.global/blog/bim-file-formats

AL-Zubaidi, E. D., & AlZaid, Z. A. (2025). Integration of Building Information Modeling (BIM) and Lean Construction Methods: A. Journal Européen des Systèmes Automatisés.

Amazon Web Services. (2024, March). What is NeRF (Neural Radiance Field)? Retrieved from https://aws.amazon.com/what-is/nerf/

Antonio Torralba, P. I. (2024). Foundations of Computer Vision. Cambridge, Massachusetts: The MIT Press.

AR-MR-XR. (2022). NeRF → point cloud export — now available via nerfstudio. Retrieved from Reddit.

Associates, T. A. (2025, February). What is Building Information Modeling (BIM)? Retrieved from https://theaecassociates.com/blog/what-is-building-information-modeling-bim/

Autodesk. (2025, May 21). Industry First: XGRIDS Debuts AI-Powered 3DGS Scan-to-BIM Plugin for Revit at Autodesk DevCon Europe 2025. Retrieved from https://mail.valuebuddies.com/viewprwire.php?id=6411208

Bae, H., Golparvar-Fard, M., & White, J. (2013). High-precision vision-based mobile augmented reality system for context-aware architectural, engineering, construction and facility management (AEC/FM) applications. Visualization in Engineering.

Barazzetti, L. (2016). Photogrammetry for infrastructure documentation. Journal of Civil Engineering and Architecture, 10(4), 420–430.

Barazzetti, L., Banfi, F., Brumana, R., & Previtali, M. (2015). Creation of parametric BIM objects from point clouds using NURBS. The Photogrammetric Record, 30(152), 339–362.

Barron, J. T., Mildenhall, B., Tancik, M., Hedman, P., Martin-Brualla, R., & Srinivasan, P. P. (2021). Mip-NeRF: A Multiscale Representation for Anti-Aliasing Neural Radiance Fields. CVF ICCV. IEEE.

Barron, J. T., Mildenhall, B., Verbin, D., Srinivasan, P. P., & Hedman, P. (2022). Mip-NeRF 360: Unbounded Anti-Aliased Neural Radiance Fields. CVF CVPR. IEEE.

Bell, N., & Hoberock, J. (2011). Thrust: A Productivity-Oriented Library for CUDA. GPU Computing Gems, Jade Edition.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020). NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. European Conference on Computer Vision (ECCV) 2020 (pp. 405–421). LNCS 12346.

Beverly, E. A., Love, M., & Love, C. (2024). Chapter 19 - Virtual Reality for Diabetes Telehealth. Diabetes Digital Health, Telehealth, and Artificial Intelligence.

Bezmalinović, T. (2025, August). Khronos launches integration of 3D Gaussian splats into the glTF file format. Retrieved from https://www.heise.de/en/news/Khronos-launches-integration-of-3D-Gaussian-splats-into-the-glTF-file-format-10518726.html

Bhatarai, R., Banihashemi, S., Shakouri, M., & Antwi-Afari, M. (2024). Integration of Augmented Reality with Building Information Modeling: Design Optimization and Construction Rework Reduction Perspective. Archives of Computational Methods in Engineering.

Biljecki, F., & Chow, Y. S. (2022). Global Building Morphology Indicators. Computers, Environment and Urban Systems, 95.

Biljecki, F., Stoter, J., Ledoux, H., Zlatanova, S., & Çöltekin, A. (n.d.). Applications of 3D City Models: State of the Art Review. ISPRS International Journal of Geo-Information, 4(4), 2842–2889.

BIM Collab. (2024, September 10). BIM File Formats Guide: OpenBIM and Native Formats for Better Projects. Retrieved from https://www.bimcollab.com/en/resources/blog/bim-file-formats-openbim-native-guide/

BIM Forum. (2024, December 31). Level of Development (LOD) Specification. Retrieved from https://bimforum.org/resource/lod-level-of-development-lod-specification/

BIM, V. (2025, August). BIM Dimensions Explained: 2D, 3D, 4D, 5D, 6D, 7D and 8D BIM. Retrieved from https://vibimglobal.com/blog/bim-dimensions/

Blaszczak-Bak, W., Masiero, A., Bąk, P., & Kuderko, K. (2024). Integrating Data from Terrestrial Laser Scanning and Unmanned Aerial Vehicle with LiDAR. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Boesch, G. (2024, May 9). Neural Radiance Fields (NeRFs): A Technical Exploration. Retrieved from https://viso.ai/deep-learning/neural-radince-fields/

Borrmann, A., Kolbe, T. H., Donaubauer, A., Steuer, H., & Jubierre, J. R. (2013). Transferring Multi-Scale Approaches from 3D City Modeling to IFC-Based Tunnel Modeling. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Turkey.

Bosché, F. (2015). The value of integrating Scan-to-BIM and Scan-vs-BIM techniques. Automation in Construction, pp. 201–213.

Brennan, M. (Director). (2023). Photogrammetry / NeRF / Gaussian splatting comparison + dataset [Motion Picture].

Broo, D. G., Bravo-Haro, M., & Schooling, J. (2022). Design and implementation of a smart infrastructure digital twin. Automation in Construction.

buildingSMART International. (2025). IFC5 Specification – Modularization and Point Cloud Support. buildingSMART International.

buildingSMART International. (2024). IFC4x3 — Industry Foundation Classes Version 4.3.2.0. ISO 16739-1. Retrieved from https://technical.buildingsmart.org/standards/ifc/ifc-schema-specifications/

buildingSMART International. (2025). IFC5 Development Overview. Retrieved from https://technical.buildingsmart.org/standards/ifc/ifc5/

Campanaro, S., & Coelho, N. O. (2023). UAV photogrammetry for construction progress monitoring: State of practice. Automation in Construction, 148, 104758.

Carlos, J. Pampliega. (2019). Building & Management. Editorial: Artificial Intelligence in Construction Sector.

Charatan, D., Li, S., Tagliasacchi, A., & Sitzmann, V. (2024). pixelSplat: 3D Gaussian Splats from Image Pairs for Scalable Generalizable 3D Reconstruction. CVF CVPR. IEEE.

Chen, A., Xu, Z., Geiger, A., Yu, J., & Su, H. (2022). TensoRF: Tensorial Radiance Fields. ECCV.

Chen, D., Shi, S., Ji, Y., Luo, Q., & Lin, X. (2023). Fusion of UAV-based photogrammetry and LiDAR for infrastructure modeling. Automation in Construction, 152, 104928.

Chen, Y., Chen, Z., Zhang, C., Wang, F., Yang, X., Wang, Y., & Lin, G. (2023). GaussianEditor: Swift and Controllable 3D Editing with Gaussian Splatting. CVF CVPR. IEEE.

Chen, Y., Wu, Q., Lin, W., Harandi, M., & Cai, J. (2025). HAC++: Towards 100X Compression of 3D Gaussian Splatting. Transactions on Pattern Analysis and Machine Intelligence. IEEE.

Chen, Z., Wang, W. (2025). Neural radiance fields assisted by image features for UAV scene reconstruction. Scientific Reports, 15, Article 3068.

Chen, Z., Shi, Q., Ji, D., Luo, H., & Lin, J. (2023, November). Review of Research and Application of BIM-GIS Integration in the Civil Engineering Industry. ICCREM 2023: The Human-Centered Construction Transformation.

Code, P. i. (Director). (2024). 3D reconstruction from a single image! Few-Shot Neural Radiance Fields | NeRF #12 [Motion Picture].

Conde, A. J., García-Sanz-Calcedo, J., & Rodríguez, A. M. (2020). Use of BIM with Photogrammetry Support in Small Construction Projects. Case Study for Commercial Franchises. Journal of Civil Engineering and Management.

Dellaert, F. (2020, December 16). NeRF Explosion 2020. Retrieved from https://dellaert.github.io/NeRF/

Deltek. (n.d.). What is Building Information Modeling (BIM)?: A Practical Guide. Retrieved from https://www.deltek.com/en/architecture-and-engineering/specpoint/building-information-modeling

Digital PV. (2025). Integrating CAD and BIM for Better Project Outcomes. Prima Verse.

Dominik Zimny, J. W. (2024). Points2NeRF: Generating Neural Radiance Fields from 3D point cloud. Pattern Recognition Letters, pp. 8–14.

Dore, C., & Murphy, M. (2012). Integration of Historic Building Information Modeling (HBIM) and 3D GIS for Recording and Managing Cultural Heritage Sites. The 18th International Conference on Virtual Systems and Multimedia. Milan: Technological

University Dublin.

dtwin / Nemetschek. (2025, November). GIS BIM Integration - Uniting Spatial and Building Data. Retrieved from https://www.nemetschek-dtwin.com/article/gis-bim-integration

Duan, C., Huang, R., Peng, S., & Li, Y. (2024). 4D Gaussian Splatting for Dynamic Scene Rendering. arXiv.

Duan, Y., Wei, F., Dai, Q., He, Y., Chen, W., & Chen, B. (2024). 4D-Rotor Gaussian Splatting: Towards Efficient Novel View Synthesis for Dynamic Scenes. SIGGRAPH Conference Papers. USA: ACM.

Editorial Team. (2026, January). BIM as Built: Meaning, How It's Created, and Its Purpose. Retrieved from https://biblus.accasoftware.com/en/bim-as-built-meaning-how-its-created-and-its-purpose/

Eiri, R. R., & Gheisari, M. (2024). NeRF for construction applications: A systematic literature review. Journal of Construction Engineering and Management, 150(8), 04024095.

El-Hakim, S. F. (2015). Photogrammetry for cultural heritage applications. The Photogrammetric Record, 30(152), 347–352.

Ellis, G. (2026, March). Extended Reality in Construction: A Guide to AR, VR, & MR. Retrieved from https://www.autodesk.com/blogs/construction/extended-reality-construction-ar-vr-mr/

El-Mewafi, M. A., & El-Rabbany, A. M. (2024). Probabilistic approaches for uncertainty quantification in structure-from-motion photogrammetry. Journal of Surveying Engineering, 150(3), 04024005.

Engineers, GBC. (2025, April). Exploring the 9 Dimensions of BIM (Building Information Modeling). Retrieved from https://gbc-engineers.com/news/bim-building-information-modeling

Fang, G., & Wang, B. (2024). Mini-Splatting: Representing Scenes with a Constrained Number of Gaussians. ECCV.

Fang, X., Zhang, Y., Tan, H., Liu, C., & Yang, X. (2025). Performance Evaluation and Optimization of 3D Gaussian Splatting in Indoor Scene Generation and Rendering. ISPRS International Journal of Geo-Information.

Fathi, H., & Dai, F. (2021). Automated as-built 3D reconstruction of civil infrastructure using structure-from-motion. Automation in Construction, 124, 103593.

Fernández, M., Rodríguez, C., & García, I. (2025). Integration of LiDAR and BIM for accurate measured building surveys in dense urban environments. International Journal of Surveying and Structural Engineering.

Finity, J., & Tell, J. (2025, July). 6 Phases of Construction Project Management Explained.

Retrieved from https://www.procore.com/library/construction-project-management-phases

Franco, J. T. (2018, March). What is BIM and Why Does it Seem to be Fundamental in the Current Architectural Design? Retrieved from https://www.archdaily.com/888727/what-is-bim-and-why-does-it-seem-to-be-fundamental-in-the-current-architectural-design

Fugas, K. (2020, March). How to present building information in a database? Retrieved from https://bimcorner.com/how-to-present-building-information-in-a-database/

Furukawa, Y., & Hernández, C. (2019). Multi-View Stereo: A Tutorial. Foundations and Trends in Computer Graphics and Vision, 9(1–2), 1–148.

Furukawa, Y., & Ponce, J. (2010). Accurate, dense, and robust multi-view stereopsis. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1362–1376.

Gao, C., Sun, Q., Zhu, J., & Chen, J. (2025, February 12). MBS-NeRF: reconstruction of sharp neural radiance fields from motion-blurred sparse images.

Gao, K., Li, L., He, H., Lu, D., Xu, L., & Li, J. (2025). Gaussian Building Mesh (GBM): Extract a Building's 3D Mesh with Google Earth and Gaussian Splatting. Remote Sensing Applications: Society and Environment, 40.

García, I. (2025, April). Construction Engineering vs Civil Engineering: Key Differences Explained. Retrieved from https://uptalent.io/blog/construction-engineering-vs-civil-engineering/

Garland, M., & Heckbert, P. S. (n.d.). Surface Simplification Using Quadric Error Metrics. Retrieved from https://www.cs.cmu.edu/~garland/Papers/quadrics.pdf

GBC Engineers. (2025, May 7). How BIM Technology Transforms the Construction Industry in 2025. Retrieved from https://gbc-engineers.com/news/bim-technology-in-construction

Gong, Y. (2024). EGGS: Edge Guided Gaussian Splatting for Radiance Fields. New York: ACM.

Google Scholar. (2026). Citation analysis for "3D Gaussian Splatting" (Kerbl et al., 2023): 2,847 citations as of April 2026.

Google Scholar. (2026). Search results for "3D Gaussian Splatting" (2022): 2 papers. Google.

Graham, J. (2023, February). The Internet of Things in smart construction: 8 key pros and cons to consider. Retrieved from https://www.planradar.com/au/internet-of-things-smart-construction/

Granshaw, S. I. (1980). Photogrammetric terminology. The Photogrammetric Record.

Green, D. R., Gregory, B. J., & Karachok, A. R. (2020). Digital photogrammetry for marine and coastal applications. Journal of Coastal Research, 36(6), 1123–1134.

Grilli, E., & Remondino, F. (2020). Classification of 3D Digital Heritage. Remote Sensing, 11(7).

Guédon, A., & Lepetit, V. (2023). SuGaR: Surface-Aligned Gaussian Splatting for Efficient 3D Mesh Reconstruction and High-Quality Mesh Rendering. CVF CVPR. IEEE.

Guo, J., Ma, X., Fan, Y., Liu, H., & Li, Q. (2024). Semantic Gaussians: Open-Vocabulary Scene Understanding with 3D Gaussian Splatting. CVF CVPR. IEEE.

Guo, Y. (2025). NeRF: Neural Radiance Field in 3D Vision: A Comprehensive Review. arXiv preprint. https://arxiv.org/abs/2210.00379.

Hachisuka, S., Yu, A. T., & Leung, S. F. (2023). Harbingers of NeRF-to-BIM: A case study of semantic segmentation on building structure. Proceedings of the International Conference on Smart Infrastructure and Construction (ICSIC).

Haitz, D., Hermann, M., Roth, A. S., Weinmann, M., & Weinmann, M. (2024). The Potential of Neural Radiance Fields and 3D Gaussian Splatting for 3D Reconstruction from Aerial Imagery. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Hamad, M. (2023). Autodesk Revit 2024 Architecture. Mercury Learning and Information.

Hammer Missions. (2026, January 26). Advances in 3D Reality Mapping: Gaussian Splatting vs Photogrammetry. Retrieved from https://www.hammermissions.com/post/gaussian-splatting-vs-photogrammetry

Harmony-AT. (2023, September 20). Differences Between BIM And GIS Technologies, And Their Uses. Retrieved from https://www.harmony-at.com/en/blog/bim-and-gis

Harmony-AT. (2023, August 18). Confused About BIM Dimensions? Here's What 3D–10D Actually Mean. Retrieved from https://www.harmony-at.com/en/blog/bim-dimensions

Hartley, R., & Zisserman, A. (2003). Multiple View Geometry in Computer Vision. Cambridge: Cambridge University Press.

Hedman, P., Philip, J., Price, T., Frahm, J.-M., Drettakis, G., & Brostow, G. (2018). Deep Blending for Free-Viewpoint Image-Based Rendering. ACM Transactions on Graphics, 37.

Hedman, P., Srinivasan, P. P., Mildenhall, B., Barron, J. T., & Debevec, P. (2021). Baking Neural Radiance Fields for Real-Time View Synthesis. CVF ICCV. IEEE.

Hoang, T. V., Nguyen, T. N., Phuc, L. V., & Phuong, N. C. (2025). GIS and BIM integration for urban infrastructure lifecycle management. International Journal of Digital Earth, 18(1), 45–67.

Hong, Y., & Guo, F. (2025). A framework of BIM-IoT application in construction projects through multiple case study approach. Journal of Building Design and Environment.

Hsiang-Hui Hung, H.-P. D.-H.-C. (2024). TimeNeRF: Building Generalizable Neural Radiance Fields across Time from Few-Shot Input Views. 32nd ACM International Conference on Multimedia (MM'24), (pp. 253–262). Melbourne.

Huang, B., Yu, Z., Chen, A., Geiger, A., & Gao, S. (2024). 2D Gaussian Splatting for Geometrically Accurate Radiance Fields. SIGGRAPH Conference Papers. New York: ACM.

Huang, J., & Funkhouser, T. (2024). 2D Gaussian Splatting for Surface Reconstruction. Proceedings of the European Conference on Computer Vision (ECCV).

Hubert, A., Elghazaly, G., & Frank, R. (2024). Editing Implicit and Explicit Representations of Radiance Fields: A Survey. Machine Vision and Applications / arXiv.

Hübner, P., & al. (2024). 3D Gaussian Splatting for BIM-Augmented On-Site Visualization: A 5G Field Study. Automation in Construction.

ISO 19650-2. (2018). Organization and Digitization of Information about Buildings and Civil Engineering Works, Including BIM — Part 2: Delivery Phase of the Assets. Geneva: ISO. Retrieved from https://www.iso.org/standard/68080.html

Jager, M., & Jutzi, B. (2023). 3D Density-gradient based edge detection on neural radiance fields for geometric reconstruction. ISPRS Archives, XLVIII-1/W3-2023, (pp. 71–78).

Jager, M., Landgraf, S., & Jutzi, B. (2023). Density uncertainty quantification with NeRF-Ensembles: impact of data and scene constraints. arXiv preprint.

Jain, S., Gupta, K., & Chandraker, M. (2026). NERFIFY Framework Converts NeRF Papers to Trainable Code. Proceedings of CVPR 2026.

James, M. R., & Robson, S. (2012). Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application. Journal of Geophysical Research, 117, 51–69. doi:10.1029/2011JF002289

James, M. R., & Robson, S. (2016). Mitigating systematic error in topographic structure-from-motion. Earth Surface Processes and Landforms, 41(12), 1725–1741.

Jeon, Y., Tran, D. Q., & Kim, T. H. (2026). Neural radiance fields for construction site scene representation and progress evaluation with BIM. Automation in Construction, 165.

Jiang, L. C. (2024). Mixer-NeRF: Research on 3D Reconstruction Methods of Neural Radiance Fields Based on Hybrid Spatial Feature Information. Journal of Computing and Electronic Information Management, pp. 149–155.

Jiang, L. L., & Jaafar, M. A. (2017). UAV photogrammetry for construction site monitoring. Automation in Construction, 82, 147–157.

Jutzi, I. P. (2023). Accuracy analysis between Neural Radiance Fields (NeRFs) and Terrestrial

Laser Scanning (TLS). The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Kalantari, M., Clemen, C., & Jadidi, A. (2024). BIM and 3D GIS Integration for Digital Twins: A Systematic Review. ISPRS International Journal of Geo-Information, 13(2).

Kalantari, M., Clemen, C., & Jadidi, M. (2024). BIM and 3D GIS Integration for Digital Twins: An Introduction. (1st ed.). CRC Press.

Karaagac, M. (2026, February 26). What Is a Neural Radiance Field (NeRF)? Retrieved from https://www.archivinci.com/blogs/what-is-a-neural-radiance-field

Karachok, A., & Green, D. R. (2023). The future of photogrammetry in civil engineering. Civil Engineering Surveyor, 61(3), 28–35.

Kazhdan, M., Bolitho, M., & Hoppe, H. (2006). Poisson Surface Reconstruction. Eurographics Symposium on Geometry Processing.

Kerbl, B., Kopanas, G., Leimkühler, T., & Drettakis, G. (2023, August). 3D Gaussian Splatting for Real-Time Radiance Field Rendering. ACM Transactions on Graphics, 42(4).

Kersten, T. P. (2022). Photogrammetry and laser scanning for heritage documentation. PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 90, 215–232.

Key benefits of BIM. (2022). Retrieved from University of Cambridge: https://www.cdbb.cam.ac.uk/BIM/key-benefits-bim

Khan, S., Panuwatwanich, K., & Usanavasin, S. (2024). Integrating building information modeling with augmented reality: application and empirical assessment in building facility management. Engineering, Construction and Architectural Management.

Kim, M. K., Wang, Q., & Cheng, J. C. (2019). Interaction Components of Mobile-Based Augmented Reality in AEC. Automation in Construction, 99.

Knapitsch, A., Park, J., Zhou, Q. Y., & Koltun, V. (2017). Tanks and temples: Benchmarking large-scale scene reconstruction. ACM Transactions on Graphics, 36(4).

Kontrimovičius, R., Juszczyk, M., Leśniak, A., Ustinovichius, L., & Miedziałowski, C. (2023, February). Photogrammetry-based approach for collecting and processing information about an existing building. Bulletin of the Polish Academy of Sciences Technical Sciences.

Kopanas, G., Leimkühler, T., & Drettakis, G. (2023). 3D Gaussian Splatting for real-time radiance field rendering. ACM SIGGRAPH 2023 Conference Proceedings, 110. ACM.

Kopanas, G., Philip, J., Leimkühler, T., & Drettakis, G. (2021, September). Point-Based Neural Rendering with Per-View Optimization. Computer Graphics Forum, 40(4).

Korycki, A. (2024). Evaluating geometric accuracy of NeRF reconstructions compared to

SLAM method. arXiv.

Kraus, K. (1997). Photogrammetry: Fundamentals and Standard Processes. Bonn: Dümmler.

Kraus, K. (2007). Photogrammetry: Geometry from Image and Laser Scans. Berlin: Walter de Gruyter.

Kumar, P. (2023, October). What is LiDAR technology? How does LiDAR help in depth measurement? Retrieved from https://www.e-consystems.com/blog/camera/technology/what-is-lidar-technology-how-does-lidar-help-in-depth-measurement/

Lassner, C., & Zollhöfer, M. (2020). Pulsar: Efficient Sphere-based Neural Rendering. CVF CVPR. IEEE.

Lato, M. J. (2024). Industry adoption of photogrammetry in mining and civil engineering. International Journal of Mining Science and Technology, 34(2), 221–235.

Levery. (2025, April). Extended Reality (XR) in Construction: Transforming Design, Building, and Operation. Retrieved from https://www.levery.it/xr_02

Li, Z., Mohammad, S., Ravi, R., Kalyan, S., & Manmohan, C. (2020). Inverse Rendering for Complex Indoor Scenes: Shape, Spatially-Varying Lighting and SVBRDF from a Single Image. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, (pp. 245–2484).

Li, Z., Müller, T., Evans, A., Taylor, R. H., Unberath, M., & Yuille, A. (2023). Neuralangelo: High-Fidelity Neural Surface Reconstruction. CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE.

Liao, X., Liu, Y., Huang, Y., Luo, Y., Zhou, Z., & Chong, H.-Y. (2025). Integrating CAD, BIM, immersive technology, and 3D Gaussian Splatting for construction model coordination with ISO 19650. Scientific Reports.

Lima, G. R., & El-Diraby, M. K. (2022). Scan-to-BIM for existing buildings: A systematic review. Journal of Information Technology in Construction, 27, 345–372.

Lin, C. H., Kong, C., & Lucey, S. (2018). Learning efficient point cloud generation for dense 3D object reconstruction. Proceedings of the AAAI Conference on Artificial Intelligence. AAAI Press.

Lin, Y. (2021). Photogrammetry for infrastructure condition assessment. Journal of Infrastructure Systems, 27(4), 04021028.

Linder, W. (2006). Digital Photogrammetry: A Practical Course. Berlin, Heidelberg: Springer.

Liu, C., & Chen, A. Y. (2024). Real-time photogrammetry for construction progress monitoring using edge computing and 5G. Journal of Computing in Civil Engineering, 38(4), 04024012.

Lonbar, A. G., & Chen, K. (2025). NeRF-Enhanced Digital Twin for Building Anomaly Inspection Using Unmanned Aerial Systems (UASs). CIB Conferences, 1. doi:10.7771/3067-4883.1831

Lonbar, A. G., Chang, K., Arshad, M. A., & Lim, T. J. (2026). NeRF-Enhanced Digital Twin for Building Anomaly Inspection. Advanced Engineering Informatics, 58.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, pp. 91–110.

Lowe, D. G. (2025). The continued relevance of SIFT and photogrammetry in the age of deep learning. International Journal of Computer Vision, 133(4), 678–695.

Lowood, H. E. (2026, March 1). Virtual reality. Retrieved from https://www.britannica.com/technology/virtual-reality

Lu, T., Yu, M., Xu, L., Xiangli, Y., Wang, L., Lin, D., & Dai, B. (2023). Scaffold-GS: Structured 3D Gaussians for View-Adaptive Rendering. CVF CVPR. IEEE.

Lu, W., & Anumba, C. J. (2022). Research Companion to Building Information Modeling. Edward Elgar Publishing Limited.

Luhmann, T. (2014). Close-Range Photogrammetry and 3D Imaging. Berlin/Boston: De Gruyter.

Luhmann, T., Robson, S., Kyle, S., & Boehm, J. (2019). Close-Range Photogrammetry and 3D Imaging (3rd ed.). Berlin, Germany: De Gruyter.

Martin-Brualla, R., Radwan, N., Sajjadi, M. S., Barron, J. T., Dosovitskiy, A., & Duckworth, D. (2021). NeRF in the Wild: Neural Radiance Fields for Unconstrained Photo Collections. CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE.

Martindale, J. (n.d.). Mapping and Geographic Information Systems (GIS): What is GIS? Retrieved from https://researchguides.library.wisc.edu/GIS

Marvik. (2023, November 7). From zero to NeRF: what to expect data-wise on a NeRF project. Retrieved from https://blog.marvik.ai/2023/11/07/from-zero-to-nerf-wha-to-expect-data-wise-on-a-nerf-project/

Mason, A. (2015). Making 3D Models with Photogrammetry. Retrieved from The Haskins Society Journal web tutorial.

Matrone, F., Colucci, E., Iacono, E., & Ventura, G. M. (2023). The HBIM-GIS Maintenance Platform to Enhance the Maintenance and Conservation of Historical Built Heritage. Sensors.

Tancik, M., Weber, E., Ng, E., Li, R., Yi, B., Kerr, J., & Kanazawa, A. (2023). Nerfstudio: A Modular Framework for Neural Radiance Field Development. ACM SIGGRAPH 2023

Conference Proceedings (arXiv:2302.04264v4). ACM.

McGlone, J. C. (2013). Manual of Photogrammetry. Bethesda, MD: American Society for Photogrammetry and Remote Sensing (ASPRS).

McLine Studios. (2024, January). Fundamentals of Building Information Modeling (BIM). Retrieved from https://mclinestudios.com/fundamentals-of-bim/

Megat, N., QingDong, Y., Ten, D. W., LingLing, Y., Maarop, N., Rusli, H. M., & Lu, X. (2024). Design and Evaluation of Virtual Reality Application for Building Information Modelling. Open International Journal of Informatics.

Mehraban, M. H., Mirzabeigi, S., Wang, M., Liu, R., & Sepasgozar, S. M. (2025). Automated Image-to-BIM Using Neural Radiance Fields and Vision-Language Semantic Modeling. Buildings.

Mekawy, M. S. (2025). Building Information Modeling and Geographic Information Systems for Next-Generation Spatial Intelligence. Stockholm: Stockholm University.

Mikhail, E. M. (2001). Introduction to Modern Photogrammetry. New York: John Wiley & Sons.

Mildenhall, B. (2020). NeRF: Neural Radiance Fields. Retrieved from https://github.com/bmild/nerf

Mildenhall, B., Hedman, P., Martin-Brualla, R., Srinivasan, P., & Barron, J. T. (2021). NeRF in the Dark: High Dynamic Range View Synthesis from Noisy Raw Images. CVF CVPR. IEEE.

Misra, I., Girdhar, R., & Joulin, A. (2021, September 16). An End-to-End Transformer Model for 3D Object Detection. Retrieved from https://arxiv.org/pdf/2109.08141

Mittal, A. (2024). Neural Radiance Fields: Past, Present, and Future. arXiv preprint.

Mohammadrezaei, E., Ghasemi, S., Dongre, P., Gračanin, D., & Zhang, H. (2024, May). Systematic Review of Extended Reality for Smart Built Environments Lighting Design Simulations. Retrieved from https://arxiv.org/html/2405.06928v1

Montiel-Santiago, F. J., Hermoso-Orzáez, M. J., & Terrados-Cepeda, J. (2020, July). Sustainability and Energy Efficiency: BIM 6D. Study of the BIM Methodology Applied to Hospital Buildings. Sustainability. doi:10.3390/su12145731

Montoya-Sánchez, V., Camarretta, N., Ehbrecht, M., Schlund, M., Paterno, G. B., Seidel, D., & Zemp, D. C. (2024). Comparing airborne and terrestrial LiDAR with ground-based inventory metrics of vegetation structural complexity in oil palm agroforests. Ecological Indicators.

Mouha, R. A. (2021). Internet of Things (IoT). Journal of Data Analysis and Information Processing.

Moulon, P., Monasse, P., Perrot, R., & Marlet, R. (2016). OpenMVG: Open Multiple View Geometry. International Workshop on Reproducible Research in Pattern Recognition (pp. 60–74). Springer.

Muhammad Arbab Arshad, T. J. (2024). Evaluating Neural Radiance Fields for 3D Plant Geometry Reconstruction in Field Conditions. Plant Phenomics, 6, Article 0235.

Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant Neural Graphics Primitives with a Multiresolution Hash Encoding. ACM Transactions on Graphics (SIGGRAPH), 41(4). doi:10.1145/3528223.3530127

nerfstudio-project. (2023). Nerfstudio: A Modular Framework for Neural Radiance Field Development. Retrieved from https://github.com/nerfstudio-project/nerfstudio

Nex, F., & Remondino, F. (2014). UAV for 3D mapping applications: A review. Applied Geomatics, 6(1), 1–15.

Niedermayr, S., Stumpfegger, J., & Westermann, R. (2024). Compressed 3D Gaussian Splatting for Accelerated Novel View Synthesis. CVF CVPR. IEEE.

Niemeyer, M., Barron, J. T., Mildenhall, B., Sajjadi, M. S., Geiger, A., & Radwan, N. (2021). RegNeRF: Regularizing Neural Radiance Fields for View Synthesis from Sparse Inputs. CVF CVPR. IEEE.

Nixtla. (n.d.). MLP: Multi-layer perceptron for time series forecasting. Retrieved from https://nixtlaverse.nixtla.io/neuralforecast/models.mlp.html

Nyimbili, P. H. (2016). Structure from motion (SfM): Approaches and applications. In Proceedings of the International Scientific Conference on Applied Sciences (pp. 27–30). Antalya, Turkey.

Oechsle, M., Peng, S., & Geiger, A. (2021). UNISURF: Unifying Neural Implicit Surfaces and Radiance Fields for Multi-View Reconstruction. CVF ICCV. IEEE.

Oliveira, A. S. (2025). Photogrammetry in construction 2025: Industry survey results. Engineering, Construction and Architectural Management.

Osisanwo, F. Y., Izang, A., Kuyoro, S. O., & Chukwudi, J. H. (2016, December). Internet of Things (IoT): Bridging the Gap in Tertiary Educational Process. International Journal of Computer Trends and Technology, 42(2), 94–101.

Park, K., Sinha, U., Barron, J. T., Bouaziz, S., Goldman, D. B., Seitz, S. M., & Martin-Brualla, R. (2021). Nerfies: Deformable Neural Radiance Fields. CVF ICCV. IEEE.

Paul Schulz, T. H.-H. (2026). Automating 3D Dataset Generation with Neural Radiance Fields. ROBVIS2025 – Communications in Computer and Information Science, vol. 2629. Cham, Switzerland: Springer.

Pauwels, P., Bersselaar, E. v., & Verhelst, L. (2024). Validation of technical requirements for a

BIM model using semantic web technologies. Advanced Engineering Informatics. doi:10.1016/j.aei.2024.102426

Petrovska, I., & Jutzi, B. (2023). Geometric accuracy analysis between Neural Radiance Fields (NeRFs) and Terrestrial Laser Scanning (TLS). The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-1/W2-2023, pp. 1197–1204. ISPRS.

Petzold, F., Matthias, N., & Christoph, H. (2024). NERF2BIM: AI-Driven Detailing-on-Demand through Sustainable Point Cloud Surveys and Semantic 3D Understanding for Advanced Modeling of Existing Buildings. Retrieved from https://www.mdsi.tum.de/gni/gni-funded-projects/nerf2bim

Photogrammetry Planning Challenges. (2025, July). Retrieved from https://hub.construkted.com/guides/photogrammetry-challenges/

Piaseckienė, G. (2022). Dimensions of BIM in literature: review and analysis. Mokslas – Lietuvos Ateitis / Science – Future of Lithuania.

Pinho, L. B. (2021). Scan-to-BIM workflow: An overview and case study (Master's thesis). Department of Civil and Environmental Engineering. Milan: Politecnico di Milano.

PlanRadar. (2024, April). Virtual and Augmented Reality in Construction and Real Estate. Retrieved from https://www.planradar.com/ae/virtual-augmented-reality-construction-real-estate-industry/

Poelman, R., & Krevelen, R. v. (2010). A Survey of Augmented Reality Technologies, Applications and Limitations. The International Journal of Virtual Reality.

Pollution Sustainability Directory. (2025, November 28). Digital Construction Ecosystems. Retrieved from https://pollution.sustainability-directory.com/term/digital-construction-ecosystems/

Pumarola, A., Corona, E., Pons-Moll, G., & Moreno-Noguer, F. (2020). D-NeRF: Neural Radiance Fields for Dynamic Scenes. CVF CVPR. IEEE.

Qi, C. R., Su, H., Mo, K., & Guibas, L. J. (2017). PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. Computer Vision and Pattern Recognition (CVPR) (pp. 652–660). IEEE.

Qin, M., Li, W., Zhou, J., Wang, H., & Pfister, H. (2024). LangSplat: 3D Language Gaussian Splatting. CVF CVPR. IEEE.

Qin, R., Liu, H., & Li, J. (2024). Photogrammetry for digital twin creation: A 2024 review. ISPRS Journal of Photogrammetry and Remote Sensing, 208, 125–142.

Radl, L., Steiner, M., Parger, M., Weinrauch, A., Kerbl, B., & Steinberger, M. (2024, October). StopThePop: Sorted Gaussian Splatting for View-Consistent Real-time Rendering. ACM Transactions on Graphics, 43(4).

Rakib, F. H., Howlader, S., Rahman, M., & Bhuiyan, A. Y. (2019). Application of BIM Based Interoperability: A Case. Journal of Logistics, Informatics and Service Science, 6.

Rakotosaona, M.-J., Manhardt, F., Arroyo, D. M., Tombari, F., & Lepetit, V. (2023). NeRFMeshing: Distilling Neural Radiance Fields into Geometrically-Accurate 3D Meshes. Proceedings of the International Conference on 3D Vision (3DV). IEEE.

Ranftl, R., Bochkovskiy, A., & Koltun, V. (2021). Vision Transformers for Dense Prediction. CVF ICCV. IEEE.

Rauschnabel, P. A., Hinsch, C., Alt, F., Felix, R., & Shahab, H. (2022). What is XR? Towards a Framework for Augmented and Virtual Reality. Computers in Human Behavior.

Reed, N. (2012, December 31). The Radiance Field. Retrieved from https://www.reedbeta.com/blog/the-radiance-field/

Regueira-Gomez, L. E. (2025). SC-NeRF: NeRF-based point cloud reconstruction using a stationary camera for agricultural applications. arXiv.

Reinhardt, J. (2023, December 29). 2023 Level of Development (LOD) Specification. Retrieved from https://bimforum.org/2023-level-of-development-lod-specification/

Remondino, F. (2015). Heritage documentation and 3D recording: 10 years of practice. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-5/W7, pp. 355–360. ISPRS.

Remondino, F., & Campana, S. (2014). 3D Recording and Modelling in Archaeology and Cultural Heritage: Theory and Best Practices. Oxford: Archaeopress.

Remondino, F., & Stoppa, D. (2013). TOF Range-Imaging Cameras. Springer.

Remondino, F., Karami, A., Yan, Z., Mazzacca, G., Rigon, S., & Qin, R. (2023). A Critical Analysis of NeRF-Based 3D Reconstruction. Remote Sensing, 15(14).

Revizto. (2024, July 4). BIM standards: ISO 19650 and more. Different levels of BIM maturity model. Retrieved from https://revizto.com/resources/blog/bim-standards-level-of-information-bim

Revizto. (2025, October 28). What is a BIM File Format? BIM vs. CAD. Retrieved from https://revizto.com/resources/blog/bim-vs-cad-file-format

Rocha, G., & Mateus, L. (2024). Using Dynamo for automatic reconstruction of BIM elements from point clouds. Applied Sciences, p. 4078.

Roy Gosthipaty, R. A., & Ritwik. (2021, November 17). Computer Graphics and Deep Learning with NeRF using TensorFlow and Keras: Part 2. Retrieved from PyImageSearch.

Rubloff, M. (2024, April 1). InstantSplat: Sub Minute Gaussian Splatting. Retrieved from https://radiancefields.com/instantsplat-sub-minute-gaussian-splatting

Rubloff, M. (2023, June 12). What are the NeRF Metrics? Retrieved from https://radiancefields.com/what-are-the-nerf-metrics

Sacks, R., Eastman, C., Lee, G., & Teicholz, P. (2018). BIM Handbook: A Guide to Building Information Modeling for Owners, Designers, Engineers, Contractors, and Facility Managers. John Wiley & Sons. doi:10.1002/9781119287568

Salah, R., Géczy, N., & Károlyfi, K. A. (2026). Predictive hybrid scan-to-BIM method improves heritage building documentation completeness and accuracy. Scientific Reports.

Sampaio, A. Z., Sarmento, R. S., & Gomes, A. M. (2024). BIM Improved with VR and AR Technologies. Journal of Software Engineering and Applications.

Sarlin, P.-E., DeTone, D., Malisiewicz, T., & Rabinovich, A. (2020). SuperGlue: Learning Feature Matching with Graph Neural Networks. CVF CVPR. IEEE.

Scan to BIM Solutions. (2024, December 26). LiDAR Scanning vs Photogrammetry: Which Is More Effective for Aerial Mapping? Retrieved from https://scantobimsolutions.com/lidar-scanning-vs-photogrammetry-for-for-aerial-mapping/

Scan to BIM Solutions. (2025, July 1). What Are Lidar Scan to BIM Services and Why Your Project Needs Them? Retrieved from https://scantobimsolutions.com/what-are-lidar-scan-to-bim-services-and-why-your-project-needs-them/

Schenk, T. (2005). Introduction to Photogrammetry. Columbus: Ohio State University.

Schönberger, J. L., & Frahm, J.-M. (2016). Structure-from-Motion Revisited. Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. doi:10.1109/CVPR.2016.445

Schönberger, J., Zheng, E., Pollefeys, M., & Frahm, J.-M. (2016). Pixelwise View Selection for Unstructured Multi-View Stereo. European Conference on Computer Vision (ECCV), 9907.

Scopus. (2026). Bibliometric analysis of publications mentioning "Gaussian Splatting" (2023): 86 papers (Analysis conducted April 2026). Elsevier.

Scopus. (2026). Bibliometric analysis of publications mentioning "Gaussian Splatting" (2024): 472 papers (Analysis conducted April 2026). Elsevier.

Scopus. (2026). Bibliometric analysis of publications mentioning "Gaussian Splatting" (2026 Q1–Q2): 285 papers (Analysis conducted April 2026). Elsevier.

Scopus. (2026). Bibliometric analysis of publications with "photogrammetry" AND "BIM" (2026, partial year): 215 papers (Analysis conducted April 2026). Elsevier.

Seitz, S. M., Curless, B., Diebel, J., & Scharstein, D. (2006). A comparison and evaluation of multi-view stereo reconstruction algorithms. IEEE Conference on Computer Vision and Pattern Recognition, (pp. 519–528).

Sharma, A. (2024, November 26). The Annotated NeRF – Training on Custom Dataset from Scratch in PyTorch. Retrieved from https://learnopencv.com/annotated-nerf-pytorch/

Sharma, P. (2025, March 25). How to Transition from CAD to BIM: The Ultimate Guide. Retrieved from https://www.novatr.com/blog/transition-from-cad-to-bim-modeling

Shehadeh, A., Alshboul, O., Taamneh, M. M., Jaradat, A. Q., Alomari, A. H., & Arar, M. (2025). Advanced integration of BIM and VR in the built environment: Enhancing sustainability and resilience in urban development. Heliyon.

Shehadeh, V. S., Al-Hussein, O., & Mousavi, S. R. (2024). Integration of point cloud data with BIM for geometric quality assessment. Automation in Construction, 158, 105202.

Shi, J.-C., Wang, M., Duan, H.-B., & Guan, S.-H. (2024). Language Embedded 3D Gaussians for Open-Vocabulary Scene Understanding. CVPR 2024. IEEE.

Shishehgarkhaneh, M. B., Keivani, A., Moehler, R. C., Jelodari, N., & Laleh, S. R. (2022). Internet of Things (IoT), Building Information Modeling (BIM), and Digital Twin (DT) in Construction Industry: A Review, Bibliometric, and Network Analysis. Buildings.

Shobha. (2024, February). What is BIM (Building Information Modeling)? Retrieved from https://www.alphasand.in/blog/what-is-building-information-modeling

Siewczyńska, M., & Zioło, T. (2022). Analysis of the applicability of photogrammetry in building façade. Civil and Environmental Engineering Reports.

Snavely, N., Seitz, S. M., & Szeliski, R. (2006). Photo Tourism: Exploring photo collections in 3D. ACM Transactions on Graphics (TOG), pp. 835–846.

Sprintzeal. (2026). What Is NeRF (Neural Radiance Field)? Retrieved from https://www.sprintzeal.com/blog/neural-radiance-fields

Srivastava, M. C. (2025, November). Beyond Real: What Is Augmented Reality? Retrieved from https://www.mobileappdaily.com/knowledge-hub/augmented-reality

Succár, B. (2009). Building information modelling framework: A research and delivery foundation for industry stakeholders. Automation in Construction.

Sun, J., Shen, Z., Wang, Y., Bao, H., & Zhou, X. (2021). LoFTR: Detector-Free Local Feature Matching with Transformers. CVF CVPR. IEEE.

Survey, U.S. Geological. (2011). Using terrestrial light detection and ranging (LiDAR) technology for land-surface analysis in the Southwest (Fact Sheet 2011–3017). Reston, Virginia: U.S. Geological Survey.

Szeliski, R. (2010). Computer Vision: Algorithms and Applications. London: Springer.

Tang, I. (2025, April 22). A Comparison of Scan to BIM Workflows: Photogrammetry or LIDAR? Retrieved from https://www.bim.com.sg/blog/scan-to-bim/

Team Editorial. (2025, July 3). BIM files: the main BIM formats for design. Retrieved from https://biblus.accasoftware.com/en/bim-files-the-main-bim-formats-for-design/

Team, K. E. (2025, January 21). Neural Radiance Fields: A Quick Summary. Retrieved from https://www.kiriengine.app/blog/explained/what-are-neural-radiance-fields

Techno Struct Academy. (2024, December 4). Understanding LOD in BIM | Level of Development 100–500. Retrieved from https://www.technostructacademy.com/blog/understanding-lod-in-bim-level-of-development-100-500/

Toschi, I., Remondino, F., & Roncella, R. (2020). Photogrammetric point clouds for BIM applications. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B5-2020, pp. 35–42. ISPRS.

Triggs, B., McLauchlan, P., Hartley, R., & Fitzgibbon, A. (2000). Bundle adjustment — A modern synthesis. Berlin: Springer.

Trimble. (2021, November 16). From Level of Detail to Level of Information Need - adding the how to BIM processes. Retrieved from https://www.trimble.com/blog/construction/en-US/article/level-of-information-need

Trinh Van Hoa, Tri N.M. Nguyen, Le Van Phuc, & Ngo Chau Phuong. (2025). Integration of BIM and GIS for managing infrastructure project. Transportation Research Procedia, 85.

Tutorials Point. (2026, February). Data Preparation in Machine Learning. Retrieved from https://www.tutorialspoint.com/machine-learning/machine-learning-preparing-data.htm

University, G. (2022, November). What is Computer-Aided Design (CAD)? Retrieved from https://www.goodwin.edu/enews/computer-aided-design-definition/

Verbin, D., Hedman, P., Mildenhall, B., Zickler, T., Barron, J. T., & Srinivasan, P. P. (2021). Ref-NeRF: Structured View-Dependent Appearance for Neural Radiance Fields. CVF CVPR. IEEE.

Vi BIM. (2025, July 30). BIM Maturity Levels Explained: Level 0, 1, 2, 3. Retrieved from https://vibimglobal.com/blog/levels-of-bim-maturity/

Vi BIM. (2025, October). CAD to BIM Conversion: Benefits, Process, and Pricing. Retrieved from https://vibimglobal.com/blog/cad-to-bim-conversion/

Volk, R., Stengel, J., & Schultmann, F. (2014). Building Information Modeling (BIM) for existing buildings — Literature review and future needs. Automation in Construction, 38, 109–127. doi:10.1016/j.autcon.2013.10.023

Vosselman, G., & Maas, H.-G. (2014). Airborne and Terrestrial Laser Scanning. Whittles Publishing.

Wang, P., Liu, L., Chen, J., & Theobalt, S. (2022). NeRF-SR: High-Quality Neural Radiance Fields using Super-Sampling. Proceedings of the ACM Multimedia Conference. ACM.

Wang, P., Wu, P., Wang, J., Chi, H.-L., & Wang, X. (2018). A Critical Review of the Use of Virtual Reality in Construction Engineering Education and Training. International Journal of Environmental Research and Public Health.

Web of Science. (2026). Bibliometric analysis of publications with "Neural Radiance Field" (2024): 468 papers (Analysis conducted April 2026). Clarivate.

Web of Science. (2026). Bibliometric analysis of publications with "Neural Radiance Field" (2025): 412 papers (Analysis conducted April 2026). Clarivate.

Web of Science. (2026). Bibliometric analysis of publications with "Neural Radiance Field" (2026 Q1–Q2): 98 papers (Analysis conducted April 2026). Clarivate.

Westoby, M. J., Brasington, J., & Glasser, N. F. (2017). Applications of structure-from-motion photogrammetry in geomorphology. Progress in Physical Geography, 41(3), 280–307.

Westoby, M. J., Brasington, J., Glasser, N. F., Hambrey, M. J., & Reynolds, J. M. (2012). Structure-from-Motion photogrammetry: A low-cost, effective tool for geoscience applications. Geomorphology, 179, 300–314.

Whelan, T., Salas-Moreno, R., Glocker, B., Davison, A., & Leutenegger, S. (2016). ElasticFusion: Real-time dense SLAM and light source estimation. The International Journal of Robotics Research, 35(14). doi:10.1177/0278364916669237

White, R., Walsh, M., Ross, D. A., & Laughlin, R. (2026, March). From Reality to Virtual Reality: The Impact of 3DGS on Training, Education, and Beyond. Retrieved from https://sei.cmu.edu/blog/from-reality-to-virtual-reality-the-impact-of-3dgs-on-training-education-and-beyond/

Wolf, P., & Dewitt, B. A. (2014). Elements of Photogrammetry with Applications in GIS. New York: McGraw-Hill.

Wolfe, R. C. (2023). Understanding NeRFs. Retrieved from https://towardsdatascience.com/understanding-nerfs-2a082e13c6eb

Wong, M. O., Sun, Y., Ying, H., Yin, M., Zhou, H., Brilakis, I., & Lam, C. C. (2025). Image-based scan-to-BIM for interior building component reconstruction. Automation in Construction.

Wu, C. (2013). Towards linear-time incremental structure from motion. International Conference on 3D Vision (3DV), (pp. 127–134). Seattle, WA.

Wu, S., Hou, L., & Zhang, G. M. (2021). Proceedings of the 18th International Conference on

Computing in Civil and Building Engineering. Springer International.

Wu, T., Yuan, Y.-J., Zhang, L.-X., Yang, J., Cao, Y.-P., Yan, L.-Q., & Gao, L. (2024). Recent advances in 3D Gaussian splatting. Computational Visual Media, 10(4), 613–642. doi:10.1007/s41095-024-0436-y

Xiaofeng, L., Yiminxuan, L., Yongtong, H., Yu, L., Zhaoyin, Z., & Heap-Yih, C. (2025). Integrating CAD, BIM, immersive technology, and 3D Gaussian Splatting for construction model coordination with ISO 19650. Scientific Reports, 16. doi:10.1038/s41598-025-32244-y

Xiaohui Wang, Z. Y. (2026, January). LowLight-NeRF: A novel framework for high-fidelity neural rendering in challenging low-light. Neurocomputing.

Xinghua Gao, & Bozorgi, P. P. (2019). BIM-enabled facilities operation and maintenance: A review. Advanced Engineering Informatics.

Xu, Q. (2024). Few-shot NeRF by adaptive rendering loss regularization. Proceedings of the 18th European Conference on Computer Vision (ECCV 2024), (pp. 125–142).

Xu, Y., Zhang, T., & Chen, X. (2023). Advances in neural radiance fields for 3D reconstruction. Computers & Graphics, 112, 78–94.

Yan, C., Qu, D., Xu, D., Zhao, B., Wang, Z., Wang, D., & Li, X. (2024). GS-SLAM: Dense Visual SLAM with 3D Gaussian Splatting. CVF CVPR. IEEE.

Yang, L., Kang, B., Huang, Z., Zhao, Z., Xu, X., Feng, J., & Zhao, H. (2024). Depth Anything V2. Conference on Neural Information Processing Systems.

Yang, Z., Gao, X., Zhou, W., Jiao, S., Zhang, Y., & Jin, X. (2023). Deformable 3D Gaussians for High-Fidelity Monocular Dynamic Scene Reconstruction. CVF CVPR. IEEE.

Ye, M., Danelljan, M., Yu, F., & Ke, L. (2024). Gaussian Grouping: Segment and Edit Anything in 3D Scenes. ECCV 2024.

Ye, V., Li, R., Kerr, J., Turkulainen, M., Yi, B., Pan, Z., & Kanazawa, A. (2024, September 10). gsplat: An Open-Source Library for Gaussian Splatting. Journal of Machine Learning Research, 26(34), 1–17.

Yigitba, E., Nowosad, A., & Engels, G. (2023, July 13). Supporting Construction and Architectural Visualization through BIM and AR/VR: A Systematic Literature Review. Paderborn, Germany.

Yiheng, Z., Zhaofan, Q., Yingwei, P., Ting, Y., & Tao, M. (2023). Learning Neural Implicit Surfaces with Object-Aware Radiance Fields. International Conference on Computer Vision (ICCV 2023).

Yu, A., Fridovich-Keil, S., Tancik, M., Chen, Q., Recht, B., & Kanazawa, A. (2021). Plenoxels: Radiance Fields without Neural Networks. CVF CVPR. IEEE.

Yu, A., Li, R., Tancik, M., Li, H., Ng, R., & Kanazawa, A. (2021). PlenOctrees for Real-time Rendering of Neural Radiance Fields. ICCV. IEEE.

Yu, C., Verbree, E., Oosterom, P. v., & Pottgiesser, U. (2025). Geometric accuracy assessment of 3D Gaussian Splatting for built environment digitization. ISPRS Journal of Photogrammetry and Remote Sensing, 210, 245–260.

Yu, H., Liu, J., & Zhang, S. (2023). NeRF for large-scale scene reconstruction: Challenges and opportunities. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(11), 12890–12908.

Yu, Y., Verbree, E., Oosterom, P. V., & Pottgiesser, U. (2025). 3D Gaussian Splatting for Modern Architectural Heritage: Integrating UAV-Based Data Acquisition and Advanced Photorealistic 3D Techniques. AGILE GIScience Series.

Yu, Z., Sattler, T., & Geiger, A. (2024). Gaussian Opacity Fields: Efficient Adaptive Surface Reconstruction in Unbounded Scenes. ACM Transactions on Graphics, 43.

Zbirovsky, S. (2025). cloud2BIM: An open-source automatic pipeline for efficient conversion of large-scale point clouds into IFC format. arXiv.

Zhang, K., Riegler, G., Snavely, N., & Koltun, V. (2020). NeRF++: Analyzing and Improving Neural Radiance Fields. arXiv Preprint.

Zhang, R., Isola, P., Efros, A. A., Shechtman, E., & Wang, O. (2018). The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 586–595). IEEE.

Zhang, X., Meng, J., Xu, Z., Yang, S., Wu, Y., Wang, R., & Zhang, J. (2025). SecureGS: Boosting the Security and Fidelity of 3D Gaussian Splatting Steganography. ICLR 2025.

Zhang, Y., Ma, C., Gao, J., Lin, J., Jiang, X., & Wei, Z. (2025). HERO: Hardware-Efficient RL-based Optimization Framework for NeRF Quantization. arXiv:2510.09010.

Zhao, C., Sun, S., Wang, R., Guo, Y., Wan, J.-J., Huang, Z., & Ren, L. (2024). TCLC-GS: Tightly Coupled LiDAR-Camera Gaussian Splatting for Autonomous Driving. CVF International Conference on Computer Vision. IEEE.

Zhao, P., Li, C., Chen, L., Wang, X., & Jiang, J. (2025). A novel framework utilizing 3D Gaussian Splatting to construct building geometry for urban wind simulations. Sustainable Cities and Society.

Zhao, Z., Li, Y., Chen, X., Wang, X., & Jiang, H. (2026). Gaussian Building Mesh for urban-scale reconstruction. IEEE Transactions on Visualization and Computer Graphics. IEEE.

Zhenyu, L., Fu, C. C., Jiaying, Z., Zhaolun, L., Boyu, W., Mingzhu, W., & Jack, C. (2025). Multimodal LLM-driven language-embedded 3D Gaussian Splatting for semantic and realistic digitization of historical buildings. Automation in Construction, 181. doi:10.1016/j.autcon.2025.106628

Zhihong Chen, X. C. (2025). Neural radiance fields assisted by image features for UAV scene reconstruction. Scientific Reports, p. 30608.

Zhou, X., Lin, Z., Shan, X., Wang, Y., Sun, D., & Yang, M.-H. (2024). DrivingGaussian: Composite Gaussian Splatting for Surrounding Dynamic Autonomous Driving Scenes. CVF CVPR. IEEE.

Zhu, J., & Wu, P. (2022). BIM/GIS data integration from the perspective of information flow. Automation in Construction, 136, 104166.

Zhu, L., Hattula, E., & Raninen, J. (2024, November). Enhancing building extraction with Lidar and aerial image data. Retrieved from https://www.gim-international.com/content/article/enhancing-building-extraction-with-lidar-and-aerial-image-data

Zhu, Z., Fan, Z., Jiang, Y., & Wang, Z. (2024). FSGS: Real-Time Few-shot View Synthesis using Gaussian Splatting. ECCV.

Zimny, D., & Joanna Waczyńska, T. T. (2024, september). Points2NeRF: Generating Neural Radiance Fields from 3D Point Cloud. *Pattern Recognition Letters*. doi:DOI: 10.1016/j.patrec.2024.07.002

Zwicker, M., Pfister, H., Baar, J. v., & Gross, M. (2002). EWA Splatting. IEEE Transactions on Visualization and Computer Graphics, 8.