Applied Science and Innovative Research
ISSN 2474-4972 (Print) ISSN 2474-4980 (Online)
Vol. 10, No. 1, 2026
www.scholink.org/ojs/index.php/asir

Original Paper

3D Model Production Solutions from Satellite Imagery for Different Application Scenarios

Zheng Fang[1], Wenchao Xin[1] & Yan Gong[1*]

[1] China Siwei Surveying and Mapping Technology Co., Ltd., 100089, Beijing, China

[*] Corresponding author: Yan Gong

Received: December 6, 2025 Accepted: January 2, 2026 Online Published: January 26, 2026

doi:10.22158/asir.v10n1p55 URL: http://dx.doi.org/10.22158/asir.v10n1p55

Abstract

3D reconstruction from satellite imagery has become a core application in the field of geospatial information technology today. Its inherent advantage of global coverage makes it indispensable in numerous fields, including but not limited to smart city construction, geographic navigation systems, virtual tourism, disaster response and management, and military and national defense. This paper compares different 3D models converted from satellite imagery, analyzing the characteristics and distinctions of various 3D models. It provides a comprehensive and in-depth analysis of various technical solutions for creating 3D models from satellite imagery. Through different comparisons, it offers a technical selection guide for 3D models in different application scenarios, enabling users to understand form, structure, and environmental changes more intuitively and deeply, and to prioritize between "visual fidelity" and "data availability".

Keywords

Photogrammetry, Stereoscopy, SAR, InSAR, DSM, DEM, DTM, Ground Control Points (GCPs), Rational Polynomial Coefficients (RPCs)

1. Overview

1.1 Definitions and Differences Between 3D Models, Digital Elevation Model (DEM), Digital Surface Model (DSM), and Digital Terrain Model (DTM)

Before delving into 3D reconstruction techniques, it is essential to clarify several core concepts, as they represent the specific product types of 3D models in different application scenarios. The Digital Elevation Model (DEM) is typically used as a general term referring to any 3D computer graphics model representing elevation data (Digital elevation model, 2024). In practice, DEM is further divided into two more specific types: Digital Surface Model (DSM) and Digital Terrain Model (DTM) (Digital elevation model, 2024).

①**Digital Surface Model (DSM):** DSM represents the height of the Earth's surface, including all objects covering it, such as buildings, trees, and other artificial or natural features (Digital elevation model, 2024). Its data can accurately reflect the height of features like building roofs in urban areas or tree canopies in forests (Digital elevation model, 2024). Therefore, DSM is highly suitable for scenarios requiring consideration of surface cover, such as urban planning, flood inundation simulation, and radio network planning.

②**Digital Terrain Model (DTM):** Unlike DSM, DTM specifically refers to the bare-earth surface height model after removing all surface objects (such as buildings and vegetation) (Digital elevation model, 2024). Its data reflects the true ground elevation of the Earth. DTM is fundamental for applications such as hydrological analysis, geological studies, earthwork volume calculation, and forest floor management (Stereo satellite imagery, n.d.).

Therefore, distinguishing between DSM and DTM is the primary consideration in project selection, as project requirements are the fundamental driver for technical choice. If the project goal is urban digital twins or architectural planning, DSM or 3D models with building facades are needed (Stereo satellite imagery, n.d.). Conversely, if the project goal is hydrological, geological analysis, or forest management, an accurate DTM is required (Stereo satellite imagery, n.d.). In generating DTM, technologies capable of penetrating vegetation, such as Light Detection and Ranging (LiDAR) and specific radar techniques, have inherent advantages (JOUAV, 2025).

Thus, the selection of a technical solution must first match the needs of the final application scenario. For example, in densely vegetated areas, optical image photogrammetry struggles to obtain DTM, while LiDAR or InSAR provides better solutions (Wingtra, n.d.).

1.2 Types, Characteristics, and Roles of Optical and Radar Imagery in 3D Reconstruction

Satellite imagery used for 3D reconstruction mainly falls into two categories: optical imagery and radar imagery. These two types differ fundamentally in acquisition principles and data characteristics, determining their applicability in different technical solutions.

A. Optical Imagery: Optical imagery belongs to passive remote sensing technology, relying on sunlight as the source and capturing reflected visible, near-infrared, or thermal infrared light from the Earth's surface to generate images (TS2 Space, n.d.). Based on the number of spectral bands captured, optical imagery can be classified as panchromatic, multispectral, and hyperspectral imagery (Satpalda, n.d.). Panchromatic imagery typically has the highest spatial resolution but provides only black-and-white or grayscale information, making it excellent for feature identification and detailed mapping (Satpalda, n.d.). Multispectral and hyperspectral imagery provide rich color and spectral information, often used for vegetation analysis and land cover classification (Satpalda, n.d.). The greatest advantage of optical imagery lies in its rich visual detail and realistic texture information. However, its main limitation is susceptibility to weather conditions (e.g., clouds) and lighting, preventing all-weather data acquisition (JOUAV, 2025).

B. Radar Imagery: Radar imagery belongs to active remote sensing technology, acquiring surface information by emitting microwave signals and receiving backscatter from the satellite itself (TS2 Space, n.d.). Synthetic Aperture Radar (SAR) is the most common type (Satpalda, n.d.). Since microwaves can penetrate clouds and smoke and are unaffected by lighting conditions, SAR imagery enables all-weather, day-and-night data acquisition (Satpalda, n.d.). Radar imagery can even measure surface roughness and moisture. Its imaging results are visually distinct from optical photographs, typically presented as black-and-white images where water appears dark, and urban areas or mountains appear bright (TS2 Space, n.d.).

Therefore, data source selection constitutes a fundamental trade-off. The differences between optical and radar imagery form a basic technical trade-off. Users cannot simultaneously have the rich texture provided by optical imagery and the all-weather capability provided by radar imagery. If a project demands high visual fidelity for the final 3D model, it must accept the limitation of optical imagery being weather-dependent. Conversely, if a project requires acquiring surface elevation or deformation information at any time and under any weather condition, it must choose radar imagery and accept its lack of intuitive texture (TS2 Space, n.d.).

Thus, the choice of technical solution largely depends on whether the project prioritizes "visual fidelity" or "data availability."

2. In-Depth Analysis of Traditional Technical Solutions

2.1 Photogrammetry

Photogrammetry is the science of obtaining reliable measurements from photographs or digital images to generate 3D models or maps (Esri, n.d.). It is one of the most mature and widely used technologies in the field of 3D reconstruction. The core principle of photogrammetry is stereoscopy, which mimics human binocular depth perception by analyzing multiple images of the same area taken from different viewpoints to reconstruct 3D geometric information (Stereo satellite imagery, n.d.). Its workflow typically includes the following key steps:

A. Image Acquisition: This is the starting point of the entire process. It requires obtaining multiple stereoscopic or multi-viewpoint images of the same area with high overlap (typically recommended at 60%-80%) (Esri, n.d.). For satellite platforms, this is usually achieved by the same satellite capturing the same area multiple times at different times or from different viewpoints to obtain stereo image pairs (Stereo satellite imagery, n.d.).

B. Image Matching and Alignment: Specialized software (e.g., OpenDroneMap, AliceVision) automatically identifies unique feature points (e.g., corners, line segments, dense textures) in these images (Formlabs, n.d.). By comparing the same feature points appearing in multiple images, the software can precisely estimate the camera position and pose for each image, a process known as Structure from Motion (SfM) (Formlabs, n.d.).

C. Dense Reconstruction and Point Cloud Generation: After determining camera parameters, the software uses triangulation to calculate the 3D coordinates of these matched feature points, generating a sparse point cloud (Formlabs, n.d.). Subsequently, through more complex algorithms (such as Multi-View Stereo, MVS), the sparse point cloud is refined into a high-density point cloud, accurately reconstructing the 3D geometric shape of the scene (Formlabs, n.d.).

D. Model Generation and Texture Mapping: The final step involves generating a continuous mesh model from the dense point cloud data and mapping the original high-resolution optical image's texture information onto this mesh surface, resulting in a visually realistic 3D model with rich detail (Formlabs, n.d.).

The successful implementation of photogrammetry relies heavily on accurate modeling of image geometric relationships, where Ground Control Points (GCPs) and Rational Polynomial Coefficients (RPCs) are two core elements.

A. Ground Control Points (GCPs): GCPs are reference points on the ground with known precise geographic coordinates, typically collected in the field using high-precision surveying equipment (e.g., differential GPS) (JOUAV, n.d.). These points act as "anchors" for the 3D model. They are key for correcting satellite orbit errors and camera distortion, aligning the model with real-world coordinate systems (absolute accuracy) (JOUAV, n.d.). Studies show that using GCPs can significantly improve the accuracy of 3D models. For example, an evaluation study on stereo satellite imagery indicated that using 3 to 10 GCPs could improve horizontal absolute accuracy to about 0.6 meters and vertical absolute accuracy to about 0.8 meters, far exceeding errors of tens of meters without GCPs (JOUAV, n.d.).

B. Rational Polynomial Coefficients (RPCs): RPCs are a generic mathematical model describing the transformation relationship from 3D geographic coordinates (longitude, latitude, elevation) to 2D image coordinates (row, column) in the form of a ratio of two cubic polynomials (Esri, n.d.). This model is important because it replaces complex and often proprietary physical sensor models, becoming an industry standard for processing various satellite images (especially optical imagery) (Esri, n.d.). Satellite image providers usually supply RPCs along with the image data, but these original RPCs contain inherent errors (Hu, Shen, & Zhang, 2023). Therefore, RPCs need to be corrected using methods such as GCPs or Bundle Adjustment to eliminate systematic errors and improve positioning accuracy (Tao & Hu, 2001).

The greatest advantage of photogrammetry lies in its ability to generate 3D models with excellent visual fidelity and rich texture detail, giving it unique value in applications requiring high-quality visualization (JOUAV, 2025). Additionally, as it primarily relies on commercial-grade cameras and software, its overall cost is relatively low compared to active remote sensing technologies like LiDAR (JOUAV, 2025). However, high-precision photogrammetric reconstruction strongly depends on good lighting and weather conditions; clouds, shadows, and lighting variations can affect image matching quality (Formlabs, n.d.). Reconstructing areas with low texture or reflective/transparent surfaces (e.g., glass, water) remains technically challenging (Wingtra, n.d.). Furthermore, since optical imagery cannot penetrate dense vegetation, accurately generating true-ground DTM in these areas is difficult (JOUAV, 2025).

2.2 Radargrammetry

Radargrammetry is another important technical path for 3D reconstruction, with its core advantage being the ability to overcome the limitations of optical imagery regarding weather and lighting.

A. SAR (Synthetic Aperture Radar): SAR is an active remote sensing technology that continuously emits radio waves towards a target area from a radar antenna mounted on a moving platform (such as a satellite or aircraft) and receives the returned echoes (Synthetic-aperture radar, 2024). By synthesizing echoes received multiple times along the antenna's moving trajectory, SAR can simulate a large "synthetic aperture," thereby achieving two-dimensional surface images with higher resolution than the physical antenna size (Synthetic-aperture radar, 2024).

B. InSAR (Interferometric SAR): InSAR technology is a further development of SAR. It combines two or more SAR images of the same area taken from slightly different positions or at different times, using the phase difference of the returned signals to infer surface elevation information or surface deformation (National Aeronautics and Space Administration, n.d.). This phase difference, akin to human binocular disparity, can be used to generate Digital Elevation Models (DEM) (National Aeronautics and Space Administration, n.d.). If two images are taken from the same position but at different times, the phase difference can reveal minute surface displacements or deformation occurring between the two acquisitions, commonly used for monitoring earthquakes, landslides, and subsidence (National Aeronautics and Space Administration, n.d.).

The process of generating DEM using InSAR is relatively complex, typically involving a series of precise processing steps (OICRF, n.d.). First, high-precision co-registration of two SAR images is required to ensure pixel-level alignment. Subsequently, an interferogram is generated, encoding phase difference information in the form of colored fringes, which can be interpreted like contour lines on a map (National Aeronautics and Space Administration, n.d.). The next step is "phase unwrapping," a critical and complex procedure aimed at converting the phase difference information in the interferogram into continuous relative elevation values. Finally, by removing interfering factors such as flat-earth effect, topographic effect, and atmospheric effect, and performing geocoding, the final DEM product aligned with real geographic coordinates is generated (Lu & Kwoun, 2012).

The greatest advantage of InSAR is its all-weather, day-and-night operational capability (Wingtra, n.d.). Regardless of whether the target area is day or night, sunny or covered by thick clouds, InSAR can acquire reliable data (Synthetic-aperture radar, 2024). This makes it particularly suitable for mapping in areas with high cloud cover like tropical rainforests, where optical imagery acquisition is nearly impossible (LandInfo Worldwide Mapping, n.d.). However, InSAR generates DEM data; its output is essentially elevation information, lacking the color and texture information provided by optical imagery, and thus cannot be directly used for high-fidelity visualization (Satpalda, n.d.). Additionally, InSAR data processing is complex and requires specialized expertise. It is susceptible to interference from factors like atmospheric water vapor and ionospheric effects, leading to elevation errors (Lu & Kwoun, 2012). InSAR's ability to penetrate dense vegetation is also limited by its radar wavelength; for example, Cband radar faces challenges in penetrating vegetation (Arjasakusuma et al., 2021).

3. Exploration of Cutting-Edge Technologies

With the rapid development of computing power, artificial intelligence technologies represented by deep learning have begun to be widely applied in 3D reconstruction from satellite imagery, offering new ideas to overcome the limitations of traditional methods.

3.1 Deep Learning-Based Multi-View Stereo (MVS)

Traditional Multi-View Stereo (MVS) methods rely on manually designed feature extraction and matching algorithms (Stanford University, 2022). However, this approach often performs poorly when faced with complex lighting, shadows, and low-texture areas, and requires high preprocessing demands (A review of 3D reconstruction from high-resolution urban satellite images, 2023). The introduction of deep learning has fundamentally changed this landscape.

Principles and Development: Deep Learning (DL) MVS constructs end-to-end neural networks to automatically learn and extract more robust, discriminative deep features from input satellite images (Stanford University, 2022). Its core idea is to build a "cost volume" that evaluates the likelihood of pixel matching under different depth hypotheses by transforming and aggregating features from input images (Stanford University, 2022). By regularizing and regressing the cost volume, the network can directly output a depth map for the target image (Stanford University, 2022).

Representative Framework: Sat-MVSF (Satellite Multi-View Stereo Framework) is a deep learning MVS framework specifically designed for satellite imagery 3D reconstruction tasks (Wuhan University, n.d.). It consists of preprocessing, a dedicated Sat-MVSNet network, and a post-processing module (Zhang, Li, & Liu, 2023). A major innovation of this framework is the introduction of a differentiable RPC warping module, integrating the Rational Polynomial Coefficient (RPC) model directly into the deep learning pipeline (Zhang, Li, & Liu, 2023). This addresses the geometric issues of large field-of-view and non-central projection in satellite imagery, allowing the network to directly process raw satellite images without the tedious epipolar rectification required by traditional MVS methods (Zhang, Li, & Liu, 2023).

This framework demonstrates high generalization capability when processing multi-temporal, multi-viewpoint satellite imagery (Zhang, Li, & Liu, 2023). Even using models pre-trained on aerial imagery, combined with its self-refinement strategy, it can achieve reconstruction results on new satellite imagery scenes that outperform traditional methods (Zhang, Li, & Liu, 2023).

The core breakthrough of DL-MVS lies in "automation" and "generalization." Traditional MVS workflows rely on carefully tuned parameters and cumbersome preprocessing steps (Esri, n.d.). In contrast, deep learning MVS automates the feature extraction and matching process through end-to-end learning, greatly reducing manual intervention. Sat-MVSF, through modules like RPC warping, provides native support for the geometric characteristics of satellite imagery, giving its model the potential to train on one city and generalize to another (Gomez & Lopez, 2022). However, this method still faces challenges of high computational cost and GPU memory limitations, especially when processing large-sized images (Gomez & Lopez, 2022). This leaves room for subsequent more efficient technologies, such as neural rendering.

3.2 From Neural Rendering to Gaussian Splatting

Neural rendering technology is an emerging paradigm for scene representation and rendering, aiming to learn the geometric and appearance representation of a 3D scene from 2D images and generate realistic images from novel viewpoints.

Neural Radiance Field (NeRF) uses a multi-layer perceptron (MLP) network to represent a 3D scene as a continuous volumetric function (ResearchGate, n.d.). This network takes spatial coordinates and viewing direction as input and outputs the volume density and color at that point (ResearchGate, n.d.). Models like Sat-NeRF introduce NeRF into satellite imagery 3D reconstruction. By incorporating native satellite camera models (like RPCs) and shadow-aware lighting models, they successfully handle complex lighting and shadow variations in multi-temporal satellite imagery (de Franchis et al., n.d.).

Gaussian Splatting is a more efficient 3D reconstruction and rendering technology than NeRF. It uses a set of discrete, learnable 3D Gaussian primitives to represent the scene. These primitives can be understood as volumetric "building blocks" of the scene (Digital Sense, n.d.). Unlike NeRF, which requires expensive volumetric rendering and ray marching, Gaussian Splatting significantly accelerates training and rendering speed through efficient analytical projection and rasterization techniques (Digital Sense, n.d.).

The **EOGS (Earth Observation Gaussian Splatting)** framework is the application of Gaussian Splatting technology in the field of satellite imagery (Digital Sense, n.d.). It inherits the conceptual advantages of EO-NeRF in handling multi-temporal imagery but replaces the computationally expensive volumetric rendering pipeline with the more efficient Gaussian Splatting method (Digital Sense, n.d.).

Gaussian Splatting represents a critical breakthrough in the speed and accuracy of neural rendering techniques. Early NeRF-based satellite image reconstruction (like EO-NeRF) could generate high-fidelity 3D models, but its volumetric rendering computational cost was extremely high; training a 256x256 meter scene could take up to 15 hours (Digital Sense, n.d.). This inefficiency severely limited its practicality in large-scale, time-sensitive applications. In contrast, Gaussian Splatting (EOGS) avoids expensive ray marching, reducing training time from hours to mere minutes while maintaining accuracy comparable to NeRF (Hu, Shen, & Zhang, 2023). This significant efficiency improvement makes high-fidelity, large-scale satellite 3D reconstruction more feasible for commercialization and practical applications.

3.3 The Cost-Effective Future of Single-Image 3D Reconstruction

Single-image 3D reconstruction represents a technical path making an extreme trade-off between cost and efficiency.

This method uses deep learning models to infer 3D structure (such as building roof point clouds, height, and angles) from a single aerial or satellite image (Chen et al., 2025). Its core lies in training the model to learn the implicit correspondence between 2D image features and 3D geometric structure, thereby predicting depth and elevation without multi-view information (Chen et al., 2025).

The greatest advantage of single-image 3D reconstruction is its extremely low cost and high efficiency (A review of 3D reconstruction from high-resolution urban satellite images, 2023). It avoids the complexity and high cost of multi-view data acquisition, making it particularly suitable for applications requiring rapid, large-scale urban modeling, such as quickly updating city maps or conducting macro planning analysis (A review of 3D reconstruction from high-resolution urban satellite images, 2023). However, this method has inherent limitations. Since a single image lacks depth information, its reconstruction results are often incomplete; for example, it is difficult to reconstruct building facades or capture complex terrain details (Chen et al., 2025). Existing single-view methods also heavily depend on high-quality, large-scale training datasets and precise camera pose information (Chen et al., 2025). Although some methods can generate more detailed and precise roof point clouds by adding auxiliary conditions like edge maps, reconstructing complete building structures from a single image remains an unsolved challenge (Chen et al., 2025).

Single-image reconstruction is an extreme trade-off between low cost and high accuracy. Traditional multi-view methods pursue high accuracy and completeness, but this requires expensive data acquisition and processing. Single-image reconstruction takes the opposite approach, sacrificing some geometric completeness and accuracy in exchange for significant cost and time advantages (A review of 3D reconstruction from high-resolution urban satellite images, 2023). Although it cannot meet the needs of precise engineering surveying, it holds unique value for macro applications requiring rapid updates and coverage of large areas (such as urban planning trend analysis, traffic flow simulation). This provides users with a novel technical path choice: deciding between "accuracy first" and "cost first."

4. Comprehensive Evaluation and Comparison of Technical Solutions

4.1 Comparison of Key Metrics

We present a structured table to comprehensively compare the key metrics of the various technical solutions discussed above and provide a technical selection guide based on different application scenarios.

Technical Method	Data Source	Typical Accuracy (Vertical)	Cost/Efficiency	Output Type	Core Advantages	Main Limitations
Traditional Photogrammetry	Multi-viewpoint optical imagery	GCPs: ~0.8m (Kocaman, 2016)	Medium / Low (JOUAV, 2025)	Textured mesh, point cloud, DSM/DTM	High visual fidelity, rich texture, relatively lower cost than LiDAR (JOUAV, 2025)	Dependent on weather & lighting; difficult in complex/low-texture areas (Formlabs, n.d.)

InSAR	SAR stereo pair	GCPs: <1.5m (Kocaman, 2016)	Medium / Medium	DEM/DTM	All-weather, day-and-night capability; can monitor surface deformation (Wingtra, n.d.)	Lacks texture; complex data processing; susceptible to atmospheric effects (Lu & Kwoun, 2012)
Deep Learning MVS	Multi-viewpoint optical imagery	High (outperforms traditional methods) (Hu, Shen, & Zhang, 2023)	Medium / Medium	Point cloud, DSM	High automation; strong generalization ability; good robustness (Zhang, Li, & Liu, 2023)	High computational resource consumption; model constrained by GPU memory (Gomez & Lopez, 2022)
Neural Rendering (Gaussian Splatting)	Multi-viewpoint optical imagery	~1.37m (Digital Sense, n.d.)	High / High (Digital Sense, n.d.)	Textured mesh, DTM	Extremely fast training & rendering (minutes); high visual quality & accuracy (Digital Sense, n.d.)	Relies on multi-view data; still in rapid development (Digital Sense, n.d.)
Single-Image DL	Single optical image	Building height MAE: <5m (A review of 3D reconstruction from high-resolution urban satellite images, 2023)	Low / Very High (A review of 3D reconstruction from high-resolution urban satellite images, 2023)	Simplified point cloud, DSM	Extremely low data acquisition cost; fast processing; suitable for large-scale coverage (Chen et al., 2025)	Limited accuracy; incomplete results (lacks facades) (Chen et al., 2025)

Note. Typical accuracy values are based on corrected data or using GCPs. MAE stands for Mean Absolute Error.

4.2 Application Scenarios and Technical Selection Guide

Based on the technical characteristics and trade-offs revealed in the table above, the following selection recommendations can be provided for different types of projects:

A. High-Precision Engineering Surveying and Detailed Planning:

For projects with strict requirements on the absolute accuracy of 3D models, such as infrastructure construction and land surveying, traditional photogrammetry using Ground Control Points (GCPs) is strongly recommended. Additionally, in open areas, InSAR technology serves as a powerful supplement for generating high-precision DTM (Kocaman, 2016).

B. Large-Scale Urban Digital Twins and High-Fidelity Visualization:

If a project requires constructing a high-fidelity, visually stunning model covering an entire city and can accept limitations on weather conditions during data acquisition, neural rendering technology (e.g., Gaussian Splatting) is a highly promising choice (Digital Sense, n.d.). It achieves an excellent balance between efficiency and visual fidelity, making it ideal for scalable deployment.

C. Disaster Response and Large-Scale Deformation Monitoring:

When rapid impact assessment is needed after a disaster or continuous monitoring of ground subsidence, landslides, etc., is required, InSAR technology is the unique and optimal choice (Stereo satellite imagery, n.d.). Its all-weather capability ensures data can be acquired promptly in any emergency situation, a feature other technologies cannot replace.

D. Low-Cost, Rapid-Update Macro Applications:

For projects with limited budgets that require quick acquisition of broad urban overviews or macro trend analysis, more cost-effective solutions like single-image 3D reconstruction are worth considering (A review of 3D reconstruction from high-resolution urban satellite images, 2023). It trades acceptable accuracy loss for significant cost and time advantages.

5. Practical Tools and Platforms

To implement the technical solutions discussed above, a series of professional software tools and platforms are required. These can be divided into commercial solutions and open-source frameworks, each serving different user needs.

5.1 Commercial Solutions

The value of commercial solutions lies in their high degree of automation, integrated functionality, and stable performance, which significantly lowers the technical barrier for users.

A. ArcGIS Reality:

As Esri's reality capture software suite, ArcGIS Reality provides an end-to-end, full-scale workflow for processing data from drones, aerial photography, to satellite imagery (Esri, n.d.). It can automate the conversion of sensor data into high-precision orthophotos, DSMs, point clouds, and 3D meshes (Esri, n.d.). Its greatest advantage is deep integration with Esri's GIS platform, enabling seamless management, analysis, and sharing of geospatial data within a unified ecosystem (Esri, n.d.).

B. Maxar Precision 3D:

Maxar is a leading commercial satellite imagery provider. Its Precision3D product line offers a series of high-quality commercial 3D data products, including DSMs, DTMs, point clouds, and textured 3D meshes (Maxar Technologies, n.d.). Its patented Precision3D Registration (P3DR) technology is particularly notable (Maxar Technologies, n.d.). This technology can automatically register new images with a foundational 3D earth model without the need for manually setting Ground Control Points (GCPs), achieving 3-meter absolute accuracy (Maxar Technologies, n.d.).

The Value of Commercial Solutions: "Automation" & "Global Coverage"

In traditional photogrammetry workflows, finding and measuring GCPs is a time-consuming and labor-intensive task (Formlabs, n.d.). Technologies like Maxar's P3DR and automated workflows like ArcGIS Reality's directly address this pain point (Esri, n.d.). By leveraging powerful automation capabilities and pre-processed data, they greatly reduce user workload. This indicates a market trend towards providing easy-to-use, scalable, and globally available "plug-and-play" solutions that encapsulate complex technology in the background, allowing users to focus on the application itself rather than technical details (Esri, n.d.).

5.2 Open-Source Tools and Frameworks

Open-source tools offer great flexibility and customizability for researchers and users with limited budgets, serving as an important foundation for technological exploration and innovation.

A. OpenDroneMap (ODM):

ODM is an open-source command-line toolkit for processing aerial imagery (OpenDroneMap, n.d.). Utilizing open-source libraries like OpenSfM and OpenMVS, it can generate maps, point clouds, 3D models, and DEMs from images captured by drones, balloons, or kites (OpenDroneMap, n.d.). Its ecosystem also includes WebODM (providing a user-friendly web interface) and PyODM (Python SDK), offering choices ranging from basic processing to API integration (OpenDroneMap, n.d.).

B. AliceVision (Meshroom):

AliceVision is an open-source computer vision framework for 3D reconstruction and camera tracking (AliceVision, n.d.). Its user-friendly, node-based workflow software, Meshroom, allows users to configure and run complete photogrammetry processes through a visual interface. The strength of AliceVision lies in its customizability and extensibility. Users can modify underlying code or create custom workflows based on specific needs and integrate with other open-source software (like Blender) (AliceVision, n.d.).

The Open-Source Ecosystem Provides Flexible and Low-Cost Ground for Exploration

Commercial software offers convenience but is often expensive and operates as a "black box." Open-source projects like OpenDroneMap and AliceVision provide researchers and budget-conscious users with significant flexibility and customizability (JOUAV, 2025). This enables users to delve into underlying algorithms, adjust workflows for specific requirements, or even engage in secondary development, thereby driving technological innovation without relying on expensive commercial licenses.

5.3 Auxiliary Tools and Technologies

A high-precision 3D model does not rely solely on a single reconstruction algorithm; its final quality is the result of the synergistic effect of multiple auxiliary technologies.

A. GNSS Correction Technology:

The accuracy of position and attitude information during satellite image acquisition is crucial for 3D reconstruction (Esri, n.d.). GNSS correction technologies, such as Real-Time Kinematic (RTK) and Precise Point Positioning (PPP), significantly improve the absolute accuracy of geographic coordinates by using data from ground reference stations to eliminate satellite and atmospheric errors (Septentrio, n.d.). RTK is suitable for local areas, providing centimeter-level accuracy, while PPP offers global correction services with slightly lower accuracy and longer initialization times (Septentrio, n.d.). These high-precision position data, whether used to assist sensor positioning during acquisition or as Ground Control Points (GCPs) in post-processing, directly impact the final quality of the 3D model.

B. Electromagnetic Simulation Software:

From a broader perspective, the accuracy of 3D reconstruction is also closely related to the hardware design and calibration of the satellite sensors themselves (Ansys, n.d.). Professional electromagnetic (EM) simulation software, such as Ansys HFSS, is used by engineers to design and simulate high-frequency electronic products on satellites, such as antennas and radars, to ensure their performance and accuracy (Ansys, n.d.).

3D Reconstruction is a System Engineering of Multi-Technology Stack Collaboration

A high-precision 3D model is not the product of a single reconstruction algorithm; its final quality results from the collaboration of multiple technologies. From satellite hardware design (Ansys HFSS) to position correction during data acquisition (GNSS RTK/PPP), to final 3D model generation, the accuracy of each link has a chain effect on the final outcome (Septentrio, n.d.). This reveals a more macro perspective: successful projects require a systematic, end-to-end understanding of the entire geospatial data processing workflow.

6. Conclusion and Future Outlook

This paper has conducted a comprehensive and in-depth analysis of various technical solutions for creating 3D models from satellite imagery. The conclusion is that there is no "one-size-fits-all" best solution. Technology selection is a trade-off decision based on the project's specific requirements, budget, and time window.

A. Traditional technologies (such as photogrammetry and InSAR) remain solid and reliable choices. Photogrammetry excels in visual fidelity and rich texture, making it the preferred choice for high-quality visualization, albeit limited by weather and lighting conditions. InSAR, with its all-weather, day-and-night capability, holds irreplaceable advantages in applications like disaster response and surface deformation monitoring.

B. Cutting-edge technologies (such as deep learning and neural rendering) are fundamentally transforming the industry landscape by enhancing automation levels, generalization capabilities, and computational efficiency. Deep learning MVS frameworks (e.g., Sat-MVSF) enable end-to-end automated workflows, while neural rendering (e.g., Gaussian Splatting), through its extremely high processing efficiency, makes large-scale, high-fidelity 3D reconstruction more practical.

C. In terms of **practical tools**, commercial software (e.g., ArcGIS Reality, Maxar Precision3D) lowers user barriers by providing highly automated and integrated solutions, while open-source tools (e.g., OpenDroneMap, AliceVision) offer flexible, low-cost platforms for research and customization needs.

Future Developments:

Satellite imagery 3D reconstruction technology will continue to evolve in the following directions:

①**Multi-Source Data Fusion:** Combining the rich texture of optical imagery with the precise elevation data from radar/LiDAR will become an important direction for generating more complete and higher-quality 3D models (Stereo satellite imagery, n.d.). For example, using optical imagery for texture and LiDAR/InSAR data for DTM under vegetation can yield urban models that are both aesthetically pleasing and accurate.

②**AI-Driven Automation:** With the advancement of deep learning, 3D reconstruction workflows will become further automated, reducing manual intervention and increasing processing speed (Esri, n.d.). For instance, AI will be able to automatically identify and segment buildings, roads, and vegetation, and even predict complete building structures from single images, greatly improving data processing efficiency.

③**Towards Real-Time Rendering:** Efficient neural rendering technologies represented by Gaussian Splatting will push the generation of large-scale, high-fidelity 3D models from the "hour-level" to the "minute-level," enabling greater utility in real-time applications (Digital Sense, n.d.).

④**End-to-End Frameworks for Native Satellite Processing:** End-to-end frameworks like Sat-MVSF and Sat-NeRF, which can natively process satellite image geometry models, will continue to be the mainstream in future research and applications (Zhang, Li, & Liu, 2023). They will encapsulate complex geometric processing, allowing users to focus more on innovation at the application layer.

References

3D reconstruction. (2024, April 14). In *Wikipedia*. https://en.wikipedia.org/wiki/3D_reconstruction

7 best practices for data accuracy in satellite imagery analysis. (n.d.). *Map Library*. Retrieved from https://www.maplibrary.org/10521/7-best-practices-for-data-accuracy-in-satellite-imagery-analysis/

A review of 3D reconstruction from high-resolution urban satellite images. (2023). *arXiv preprint arXiv:2302.xxx*. https://bohrium.dp.tech/paper/arxiv/de35e8d26194b9f0386141d1a57e935f681525a1f0b85730347d2316b5a27443

AliceVision. (n.d.). *Photogrammetric computer vision framework*. Retrieved from https://alicevision.org/

Ansys. (n.d.). *Ansys HFSS | 3D high frequency simulation software*. Retrieved from https://www.ansys.com/products/electronics/ansys-hfss

Arjasakusuma, S., et al. (2021). Comparative analysis on digital surface model of urban area from Sentinel-1 SAR interferometry and aerial photogrammetry for disaster management. *E3S Web of Conferences*, *331*, 04017. https://www.e3s-conferences.org/articles/e3sconf/pdf/2021/107/e3sconf_icdmm2021_04017.pdf

Chen, Z., et al. (2025). AIM2PC: Aerial image to 3D building point cloud reconstruction. *arXiv preprint arXiv:2503.18527*. https://arxiv.org/html/2503.18527v1

de Franchis, C., et al. (n.d.). Sat-NeRF. Centre Borelli. Retrieved from https://centreborelli.github.io/satnerf/

Digital elevation model. (2024, May 28). In *Wikipedia*. https://en.wikipedia.org/wiki/Digital_elevation_model

Digital Sense. (n.d.). *3D terrain modeling with multi-date satellite images: Reconstruction, challenges, and future*. Retrieved from https://www.digitalsense.ai/blog/3d-terrain-modeling-with-multi-date-satellite-images-reconstruction

DJI Enterprise. (n.d.). Ground control points. *DJI Enterprise Insights*. Retrieved from https://enterprise-insights.dji.com/blog/ground-control-points

Eriksson, A. (2020). *3D reconstruction from satellite imagery using deep learning* [Master's thesis, KTH Royal Institute of Technology]. DiVA Portal. http://www.diva-portal.org/smash/get/diva2:1567722/FULLTEXT01.pdf

Esri. (n.d.). *Reality capture & photogrammetry software: ArcGIS Reality*. Retrieved from https://www.esri.com/en-us/arcgis/products/arcgis-reality/overview

Esri. (n.d.). *What is photogrammetry?—ArcGIS Pro documentation*. Retrieved from https://pro.arcgis.com/en/pro-app/latest/help/data/imagery/introduction-to-ortho-mapping.htm

Formlabs. (n.d.). *Photogrammetry: Step-by-step tutorial and software comparison*. Retrieved from https://formlabs.com/blog/photogrammetry-guide-and-software-comparison/

Gomez, A., & Lopez, J. (2022). An experimental comparison of multi-view stereo approaches on satellite images. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)* (pp. 1910-1919). https://openaccess.thecvf.com/content/WACV2022/papers/Gomez_An_Experimental_Comparison_of_Multi-View_Stereo_Approaches_on_Satellite_Images_WACV_2022_paper.pdf

Hu, X., Shen, X., & Zhang, J. (2023). A review of 3D reconstruction from high-resolution urban satellite images. *International Journal of Applied Earth Observation and Geoinformation*, *118*, 103228. https://www.researchgate.net/publication/368611875_A_review_of_3D_reconstruction_from_high-resolution_urban_satellite_images

JOUAV. (2025). *LiDAR vs. photogrammetry: The ultimate showdown for 3D mapping*. https://www.jouav.com/blog/lidar-vs-photogrammetry.html

JOUAV. (n.d.). A comprehensive guide to using ground control points for drone. Retrieved from https://www.jouav.com/blog/ground-control-points.html

Kocaman, S. (2016). Accuracy assessment of 3D model generation using high resolution stereo-optical satellite imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLI-B4*, 117-123. https://www.researchgate.net/publication/302499674_Accuracy_Assessment_of_3D_model_generation_using_High_Resolution_Stereo-Optical_Satellite_Imagery

LandInfo Worldwide Mapping. (n.d.). *Digital elevation models (DEMs) from satellite imagery*. Retrieved from https://landinfo.com/worldwide-mapping-products/digital-elevation-models-dems-from-satellite-imagery/

Large-Scale 3D Reconstruction from Multi-View Imagery: A Comprehensive Review. *Remote Sensing*, *16*(5), 773. https://www.mdpi.com/2072-4292/16/5/773

Li, R., & Zhang, L. (2024). DSM reconstruction from uncalibrated multi-view satellite stereo images by RPC estimation and integration. *Remote Sensing*, *16*(20), 3863. https://www.mdpi.com/2072-4292/16/20/3863

Liu, Y., et al. (2024). 3D building reconstruction from monocular remote sensing images with multi-level supervisions. *arXiv preprint arXiv:2404.04823*. https://arxiv.org/html/2404.04823v1

Lu, Z., & Kwoun, O. (2012). Digital elevation model generation from satellite interferometric synthetic aperture radar. In *Synthetic aperture radar for remote sensing* (pp. 321-342). Southern Methodist University. https://www.smu.edu/-/media/site/dedman/academics/departments/earthsciences/pdf/lu/098_lu_etal_insar_dem_2012.pdf

Maxar Technologies. (n.d.). *High-resolution satellite imagery*. Retrieved from https://www.maxar.com/maxar-intelligence/products/satellite-imagery

Maxar Technologies. (n.d.). *Precision3D data suite overview*. Retrieved from https://www.maxar.com/maxar-intelligence/products/precision3d-data-suite

National Aeronautics and Space Administration. (n.d.). Interferometry. *NASA Science*. Retrieved from https://science.nasa.gov/mission/nisar/interferometry/

Nearmap. (n.d.). *3D aerial imagery with AI data insights*. Retrieved from https://www.nearmap.com/blog/3d-aerial-imagery-with-ai-data-insights

OICRF. (n.d.). *InSAR operational and processing steps for DEM generation*. Retrieved from https://www.oicrf.org/-/insar-operational-and-processing-steps-for-dem-generation

OpenDroneMap. (n.d.). *Drone mapping software*. Retrieved from https://opendronemap.org/

Rational polynomial coefficient. (2023, October 1). In *Wikipedia*. https://en.wikipedia.org/wiki/Rational_polynomial_coefficient

ResearchGate. (n.d.). *Figure: ECEF vs. UTM-based representation of geographic 3D point coordinates…* Retrieved from https://www.researchgate.net/figure/ECEF-vs-UTM-based-representation-of-geographic-3D-point-coordinates-The-scene-has-an_fig1_372078399

Satpalda. (n.d.). *Photogrammetry services: 3D mapping & modeling*. Retrieved from https://satpalda.com/photogrammetry/

Satpalda. (n.d.). *Types of satellite imagery explained with use cases*. Retrieved from https://satpalda.com/types-of-satellite-imagery/

Septentrio. (n.d.). *GNSS corrections demystified*. Retrieved from https://www.septentrio.com/en/learn-more/insights/gnss-corrections-demystified

Stanford University. (2022). Towards deep multi-view stereo for 3D reconstruction from satellite imagery. *CS231A: Computer Vision, From 3D Reconstruction to Recognition*. https://web.stanford.edu/class/cs231a/prev_projects_2022/5CS231A_Final_Report.pdf

Stereo satellite imagery. (n.d.). *Satpalda*. Retrieved from https://satpalda.co/stereo-satellite-imagery/

Structure from motion (SfM). (n.d.). *Humboldt State University*. Retrieved from http://gsp.humboldt.edu/olm/Courses/GSP_216/lessons/SfM.html

Synthetic-aperture radar. (2024, June 10). In *Wikipedia*. https://en.wikipedia.org/wiki/Synthetic-aperture_radar

Tao, C. V., & Hu, Y. (2001). Understanding the rational function model: Methods and applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, *56*(1), 1-3. https://www.isprs.org/proceedings/xxxv/congress/comm4/papers/423.pdf

TS2 Space. (n.d.). *Satellite imagery: Principles, applications, and future trends*. Retrieved from https://ts2.tech/en/satellite-imagery-principles-applications-and-future-trends/

Wang, Y., et al. (2025). TS-SatMVSNet: Slope aware height estimation for large-scale earth terrain multi-view stereo. *arXiv preprint arXiv:2501.01049*. https://arxiv.org/html/2501.01049v1

Wingtra. (n.d.). *LiDAR vs. photogrammetry: What sensor to choose for a given application*? Retrieved from https://wingtra.com/lidar-drone/lidar-vs-photogrammetry-what-sensor-to-choose/

Wuhan University. (n.d.). Official implementation of Sat-MVSF. *GPCV|WHU-RS*. Retrieved from https://gpcv.whu.edu.cn/data/3d/mvs.html

Zhang, L., Li, R., & Liu, X. (2023). A general deep learning based framework for 3D reconstruction from multi-view stereo satellite images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *195*, 1-15. https://www.researchgate.net/publication/366682536_A_general_deep_learning_based_framework_for_3D_reconstruction_from_multi-view_stereo_satellite_images

Zhang, Y., & Zhang, J. (2020). Accuracy and efficiency: The comparison of different RPC parameters solving methods. *Journal of Geoscience and Environment Protection*, *8*(10), 84-94. https://www.scirp.org/pdf/gep_2020102015061745.pdf