Evaluating multi-view geometry for satellite-based 3D city modeling: towards 1+N constellation configurations

Xu Cheng[1], Xianfeng Huang[1], Yingdong Pi[1], Xinsheng Wang[1], Mi Wang[1]
[1] State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Keywords: Satellite photogrammetry, Multi-view stereo, Photorealistic 3D reconstruction, City modeling, Model quality assessment.

Abstract

The emergence of satellite constellations enables near-synchronous multi-view optical imaging, offering new opportunities for large-scale 3D city modeling. Yet a practically promising configuration, in which a primary near-nadir view is complemented by multiple oblique side-looking viewpoints, remains under-examined. This study develops a controlled semi-simulation framework to analyze how multi-view imaging geometry affects the recoverability of urban 3D structures. Under idealized conditions with imaging perturbations removed, e.g., radiometric, illumination, and sensor model errors, the experiments focus on three practical factors: the number of side-looking views, view obliqueness, and the constellation’s azimuthal orientation relative to the scene. With parameter sweep analysis, it reveals an asymmetric U-shaped trend between reconstruction performance and both the view count and the obliqueness: moderate angular diversity markedly strengthens urban scene recoverability. In contrast, large obliqueness reduces inter-view overlap and destabilizes matching, while excessive redundancy introduces consistency issues that ultimately degrade reconstruction performance. Furthermore, the results shows that geometric accuracy, completeness, and texture appearance each peak at different parameter combinations, revealing intrinsic trade-offs in multi-view urban reconstruction, as different evaluation criteria favor distinct optimal configurations. The study provides practical guidance for the geometric design and mission planning of multi-satellite constellations aimed at improving satellite-based 3D modeling in urban areas.

1. Introduction

Obtaining accurate 3D information on the land surface is a fundamental objective of Earth observation initiatives, underpinning applications in national security, geographic surveying, and smart city development (Bosch et al., 2016; Liu et al., 2021; Tuia et al., 2024). Recent advances in high-resolution optical satellite systems have enabled multi-angle image acquisition, making large-scale multi-view 3D reconstruction increasingly feasible (Beyer et al., 2018; De Franchis et al., 2014; Gómez et al., 2022; Kornus et al., 2006; Qin, 2016; Rupnik et al., 2018; Shean et al., 2016; Stucker and Schindler, 2022). However, despite the growing observational capacity, there remains a lack of systematic understanding of how imaging configurations affect reconstruction quality (Qin, 2019).

Satellite-based 3D reconstruction is highly sensitive to image acquisition configurations (Gómez et al., 2023; Tao et al., 2022). In practice, limitations in sensor performance (Peng et al., 2023; Yang et al., 2020), platform maneuverability (Yang et al., 2017), and orbital constraints lead to occlusion, constrained revisit conditions, insufficient view diversity, and reduced reconstruction accuracy. With a single satellite, obtaining a complete set of views is inherently difficult: even with pitch, yaw, and roll adjustments, many building facades cannot be observed. This blind-spot issue is particularly severe in urban environments with tall, closely spaced structures (Cao and Huang, 2021; Leotta et al., 2019).

The issue of insufficient observations has been partially resolved by the global increase in dense, low-orbit satellite constellations that offer more frequent data collection (Belward and Skøien, 2015). The reduced cost and enhanced performance of miniaturized satellite components have driven the creation of coordinated multi-satellite Earth-observation systems. Constellations such as PlanetScope, RapidEye, and SkySat (Bhushan et al., 2021) (U.S.), Pleiades Neo (France), and Jilin-1 (Li et al., 2021) (China) use satellite swarms or combinations of multiple satellites to capture imagery from various orbits and angles through multiple passes over the same area in a short span of days. However, multi-temporal imaging inevitably encounters radiometric and illumination variability, which complicates matching consistency and limits reconstruction performance in urban areas. In this context, the French National Centre for Space Studies (CNES) proposed the CO3D constellation design (Lebègue et al., 2020), which employs a diamond configuration of four satellites to enhance angular diversity and ensure temporal consistency in digital surface model (DSM) generation. However, the underlying principles of such a configuration are not explicitly formalized. Beyond empirically-driven fixed paradigms, a complementary, data-driven perspective on multi-view geometric configurations remains insufficiently explored, particularly in urban environments with significant occlusion.

A particularly promising imaging paradigm is the 1+N constellation, in which one near-nadir view is complemented by oblique side-looking views. This structure is both conceptually natural and operationally efficient: the near-nadir view provides a stable geometric anchor that preserves the global scene skeleton, while the side views supply diverse angular observations needed to resolve urban façades with strong relief and self-occlusion. Moreover, the near-nadir image serves as a geometric mediator between multiple oblique views, reducing their effective angular separation and ensuring sufficient overlap for stable multi-view stereo. Its simplest degenerate form, 1+2 with fixed viewing directions, corresponds to the classical tri-stereo imaging mode used in existing satellites. Generalizing this configuration to 1+N thus offers a flexible and scalable extension for future constellation design.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-235-2026

In this study, we focus on simulation-based evaluation of multi-view imaging geometry, particularly the 1+N constellation configuration, for satellite-based 3D reconstruction in urban area. The contributions are:

- We developed **an extensible satellite-image semi-simulation framework** that supports fully parameterized control over satellite viewing geometry. The framework integrates high-fidelity urban surface models to preserve building morphology and occlusion patterns, enabling reproducible generation of satellite imagery across a broad range of configurations. This general framework allows isolated analysis of geometric factors independent of radiometric or algorithmic noise.

- We revealed **a systematic characterization of how multi-view diversity affects reconstruction performance** under 1+N constellations. Through a wide parametric sweep over side-view counts and angular distributions, we reveal a robust asymmetric U-shaped dependency between view diversity and reconstruction accuracy: reconstruction improves with increased multi-view redundancy and angle, reaches an optimal configuration, and then degrades once the view quantity or angular obliqueness becomes excessively large.

- We provided **a first quantitative assessment of geometry-induced trade-offs across accuracy, completeness, and façade texture fidelity** in satellite-based urban 3D reconstruction. The complex building façades and severe occlusions in urban scenes makes completeness and texture fidelity essential alongside accuracy. We showed that optimal configurations differ across metrics, highlighting that excessive diversity simultaneously improves visibility but harms geometric stability. These results provide practical guidance for designing urban 3D imaging strategies and future 1+N satellite constellations.

2. Simulation framework

The framework generates satellite-like images from a high-fidelity oblique 3D urban model with full control over viewing parameters. As illustrated in Figure 1, the workflow consists of two components: image simulation and rational polynomial coefficient (RPC) estimation**.**

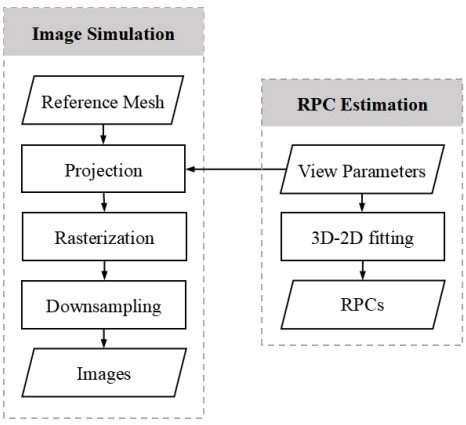

Figure 1. Workflow of the optical satellite image simulation framework.

Beginning with a textured mesh model, the scene is projected onto the image plane using a z-buffer algorithm (Catmull, 1974). Since the ground model typically has a finer resolution than the target satellite image, rasterization is performed at a higher intermediate resolution that matches the ground model. The rendered images are then downsampled to the target ground sampling distance (GSD). To derive the corresponding RPC parameters, we applied an iterative fitting method using nonlinear least squares optimization (Zhang et al., 2019). This produces radiometrically clean images that preserve façade visibility, object discontinuities and occlusion patterns, with the associated RPCs ensuring compatibility with typical satellite processing pipelines.

2.1 View Parameters

View parameters fully specify the transformation from the 3D urban surface model to the satellite image plane, and are divided into two categories as follows.

Parameters of position & orientation. Suppose each viewpoint lies on a hemisphere of radius equal to the orbital altitude H, with the optical axis directed toward the center of RoI. As illustrated in Figure 2, the viewing direction is defined by an azimuth angle θ in the x-y plane and an elevation angle φ measured from the horizontal plane, with the off-nadir-angle is given by $\alpha = \pi/2 - \varphi$. The hemispherical coordination system is centered at the target area, where the z-axis points upward, the x-axis points east, and the y-axis points north.

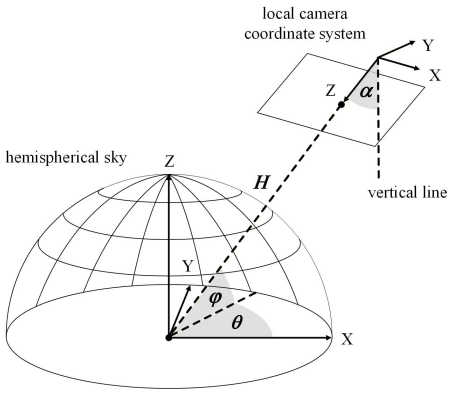

Figure 2. Parameters of position and orientation.

Parameters of sensor model. Though real optical sensors include both pushbroom and frame-based imaging systems, studies have shown that the imaging model of linear-array satellites can be approximated by a pinhole camera over limited scene extent (Zhang et al., 2019). Meanwhile, many emerging video satellites operate with frame-based CMOS detectors and inherently follow a central-projection geometry. For these reasons, we adopt a pinhole model parameterized by the specified focal length, pixel size (i.e., ground sampling distance), and image width and height.

2.2 Parameter space generation: 1+N configuration

This 1+N setting reflects a common scenario in satellite-based 3D reconstruction, where one near-nadir image serves as the geometric anchor and additional off-nadir views provide complementary angular diversity.

Given the vast combinations of possible N-view arrangements, two assumptions are introduced to make the solution feasible. First, we assume that the side-looking images are evenly distributed across the hemispherical sky, which is motivated by the aim of maximizing information acquisition. Though individual observation sites vary greatly in appearance and structure, when a sufficiently large number of locations are observed, the scenes can be considered statistically isotropic on average, meaning that they exhibit no preferred direction.

https://doi.org/10.5194/isprs-annals-XI-2-2026-235-2026

Second, we assume all the side views share the same off-nadir angle, which corresponds to the most elementary setting to ensure uniform azimuthal coverage and avoid directional bias. For clarity, we refer to this arrangement as the 1+N configuration throughout this paper.

To parameterize this configuration, we represent it as a triple (n, α, θ_0), where n denotes the number of side-looking views, α is the shared off-nadir angle, and θ_0 is the azimuth angle of the first side-looking image, as illustrated in Figure 3.

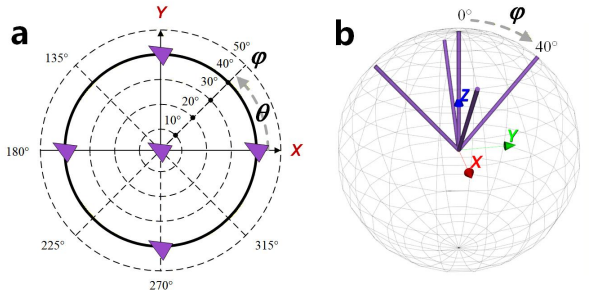

Figure 3. Parameter space of 1+N configuration. **a** Top view; **b** Frontal view.

3. Experimental setup

3.1 Image simulation parametric sweep

As illustrated in Table 1, we generated synthetic satellite images using a fixed set of viewing parameters (orbital altitude, GSD, focal length, and image size). The multi-view configurations were swept over three variables: the number of side views, the off-nadir angle, and the azimuth orientation, resulting in 180 candidate image sets.

It is worth noting that the orbital altitude is fixed under a simplified spherical model, where the sphere radius approximates the altitude, thereby eliminating image resolution variations induced by different off-nadir angles. Although orbital altitude affects GSD and image footprint, these effects can be interpreted as global scaling. The geometric relationships between viewpoints are preserved up to scale. Therefore, orbital altitude is treated as a fixed parameter to isolate multi-view geometric effects.

Table 1. Parameter settings for the simulation. (*) denotes parameters varied in the sweep.

Parameter	Symbol	Value / Range
orbital altitude	H	500 km
focal length	f	10^6 px
image size	$h \times w$	1200 × 1200 px
ground sampling distance	r	0.5 m
no. of side views*	n	{2, 3, 4, 5, 6}
off-nadir angle*	α	[0, 50°], sampled every 5°
azimuth orientation*	θ_0	[0, 360°], sampled every 30°

3.2 Reference 3D city model

We used a high-precision 3D urban model of Wuhan, China, as the reference ground truth for image simulation and reconstruction evaluation. As is shown in Figure 4, the region of interest (RoI) covered 0.5 km² urban area with rich façade structures and significant elevation variations. The model was generated from oblique photogrammetric techniques, which typically provide high-resolution, multi-angle imagery with minimal shadowing due to acquisition near local noon. As a result, both the geometric accuracy and texture quality of the reference model are high. The model was georeferenced using ground control points, achieving a checkpoint accuracy within 6 cm, and its average model resolution is approximately 3 cm, indicating high-fidelity appearance and detailed façade representation.

Figure 4. High-precision oblique photogrammetric urban model as reference.

3.3 Reconstruction pipeline

After simulating all images, the scene is reconstructed using the commercial software GET3D-SAT (Wuhan Daspatial Intelligence Technology Co., 2025), which implements a traditional pipeline consisting of bundle adjustment (BA) and multi-view stereo (MVS). These two components remain widely adopted in satellite-based 3D reconstruction due to their robustness and well-established accuracy. To this end, a unified reconstruction pipeline is adopted across all experiments to control algorithmic variability. Under this setting, the observed performance differences are primarily driven by geometric factors.

3.4 Evaluation metrics

Accuracy. We calculated accuracy only over the overlapping valid region between the reconstructed models and the ground truth, using the root mean squared error (RMSE) of the vertex-to-facet distance. The metric follows the evaluation protocols of widely used benchmarks, such as the IARPA MVS3DM (Bosch et al., 2016). However, since the outputs are mesh models rather than digital surface model (DSMs), the per-point error reflects spatial distance rather than elevation error**.**

Completeness. The completeness metric is the percentage of vertices from the ground truth data that are within a distance threshold t from the reconstructed models. The thresholds were set at 1 meter.

Texture Fidelity. Directly comparing the reconstructed model with original images is challenging due to the spatial discontinuity of texture coordinates. Inspired by the virtual rephotography approach (Waechter et al., 2017), we evaluated reconstruction quality by projecting the reconstructed model onto the virtual image plane of the original views. Specifically, we assessed the similarity between the reprojected and original images using structural similarity index measure (SSIM), where the computation was restricted to valid pixels defined by visibility masks.

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-235-2026 | © Author(s) 2026. CC BY 4.0 License.

4. Results and analysis

We first examine the main effects of the side-view count and the off-nadir angle on the reconstruction accuracy, completeness, and texture quality. Furthermore, a joint analysis was conducted since the influences of these two parameters are often intertwined. In addition, since the azimuth distribution only becomes meaningful once the number of views is fixed, we treat it as a conditional factor and do not analyze its influence separately.

4.1 Effect of side-view count on reconstruction quality

To isolate the main effect of the side-view count, the corresponding reconstruction performance were averaged over all combinations of off-nadir angles and azimuth rotation. Figure 5 shows the effect of side-view count on reconstruction quality.

As shown in Figure 5, the relationship between data redundancy and reconstruction quality is not linear. Reconstruction accuracy improves steadily with additional side views up to four, where the mean RMSE reaches the minimum. And the configuration with five side views achieves the lowest variance, indicating the most stable performance across different viewing realizations. Beyond this point, additional views introduce only marginal new information and may amplify matching inconsistencies, causing a slight increase in RMSE. This results in a shallow U-shaped trend, with four or five side views being the optimal configuration.

As shown in Figure 5, completeness exhibits a weaker dependency on the number of side views, presenting a counter-intuitive behavior. Though the minimum completeness value does increase with larger side-views count, the overall improvement is far less pronounced than in geometric accuracy. It should be noted that the absolute completeness values appear low because the reference ground truth covers a larger footprint than the reconstructed region, ensuring that side-view images, which naturally capture areas extending beyond the nominal region of interest, are fully supported by the reference model.

Though texture fidelity generally benefits from an increased number of views, the trend is not strictly monotonic. As shown in Figure 5, the mean SSIM fluctuates slightly within a narrow range. Given that the simulated dataset contains no illumination, shadow variations or imaging-model error, the instability arises mainly from view-dependent appearance changes and the accumulation of minor geometric inconsistencies during multi-view fusion. Overall, SSIM shows only limited sensitivity to the number of side views once radiometric inconsistencies and mapping deviations are excluded.

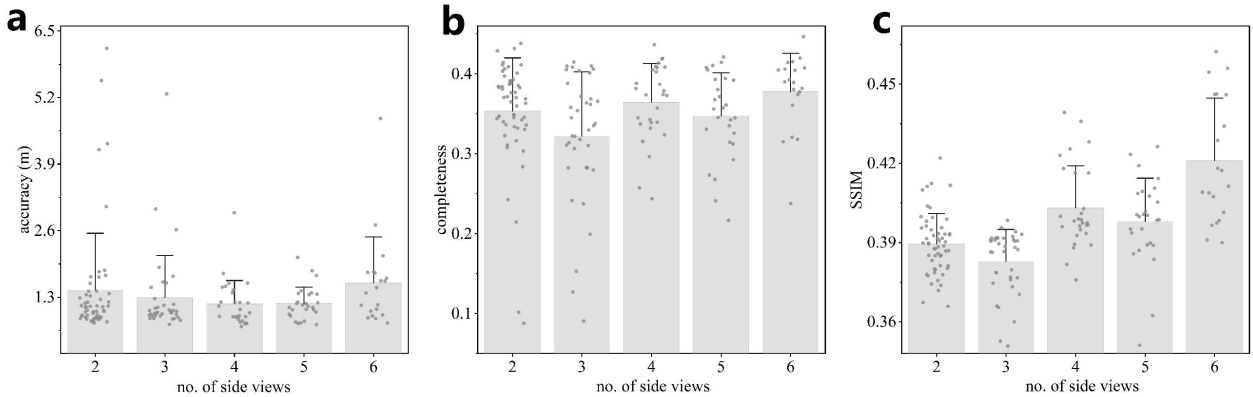

Figure 5. Effect of side-view count on **(a)** reconstruction accuracy, **(b)** completeness, and **(c)** texture quality (SSIM).

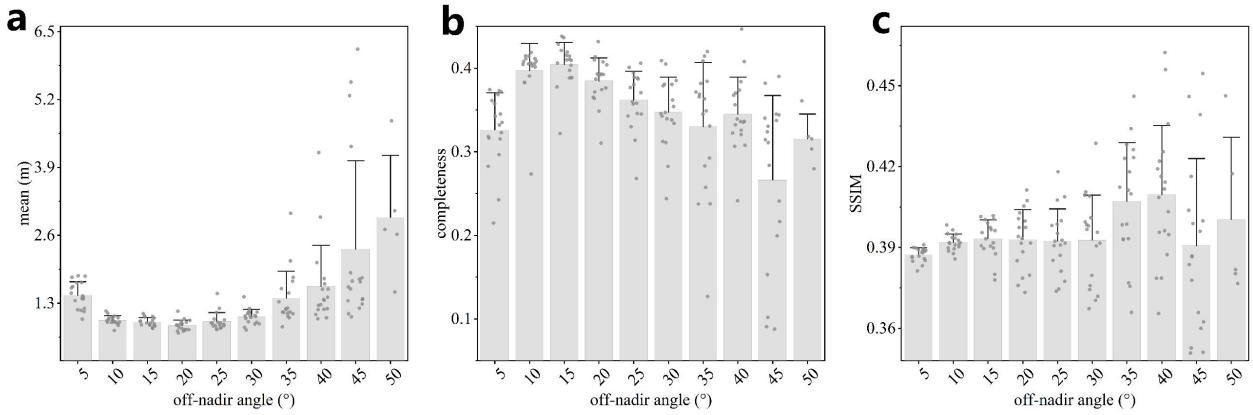

Figure 6. Effect of off-nadir angle on **(a)** reconstruction accuracy, **(b)** completeness, and **(c)** texture quality (SSIM).

4.2 Effect of off-nadir angle on reconstruction quality

To isolate the main effect of the off-nadir angle, the corresponding reconstruction performance were averaged over all combinations of side views count and azimuth rotation. Figure 6 shows the effect of off-nadir angle on reconstruction quality.

As shown in Figure 6a, reconstruction accuracy exhibits a clear dependence on the off-nadir angle. The mean error decreases from near-nadir imaging and reaches its minimum at approximately 20°, where the viewing geometry provides

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper.
https://doi.org/10.5194/isprs-annals-XI-2-2026-235-2026 | © Author(s) 2026. CC BY 4.0 License.

sufficient exposure to vertical surfaces while still preserving reliable inter-view appearance consistency for feature matching. Accuracy remains relatively stable between 10° and 30°, but degrades sharply at larger angles. Off-nadir angles of 40°-50° show both higher mean error and substantially larger variance. Overall, a moderate off-nadir angle near 20° yields the most reliable reconstruction quality.

As shown in Figure 6b, completeness exhibits a trend broadly consistent with that of accuracy. Reconstructions within the 10°-20° range achieve the highest completeness, as most surfaces remain visible and occlusion effects are minimal. When the off-nadir angle increases beyond 30°, side-facing surfaces dominate the footprint, causing more self-occlusion and reducing visibility. Consequently, completeness drops steadily and becomes more unstable, with a more pronounced decline at very oblique configurations.

As shown in Figure 6c, though the SSIM values exhibit only modest changes across different angles, a clear increase in variability can be observed. The increase can be attributed to the heterogeneous degradation of texture at large off-nadir angles. While well-observed surfaces preserve their structural appearance, oblique and partially occluded regions suffer from reduced resolution, foreshortening, and unstable dense matching. Consequently, SSIM exhibits noticeably larger standard deviations under oblique configurations.

4.3 Joint analysis and interaction effects

Figure 7 presents the joint effects of the number of side views and the off-nadir angle on reconstruction accuracy, completeness, and texture fidelity. We reverse the colomap of accuracy as we prefer smaller values, therefore the lighter color(yellow) means larger error.

As is shown in Figure 7a,b, reconstruction accuracy and completeness exhibit a well-defined optimal region in the central part of the parameter space. The poor performances along the right edges indicate that simply increasing the number of views leads to consistency issues.

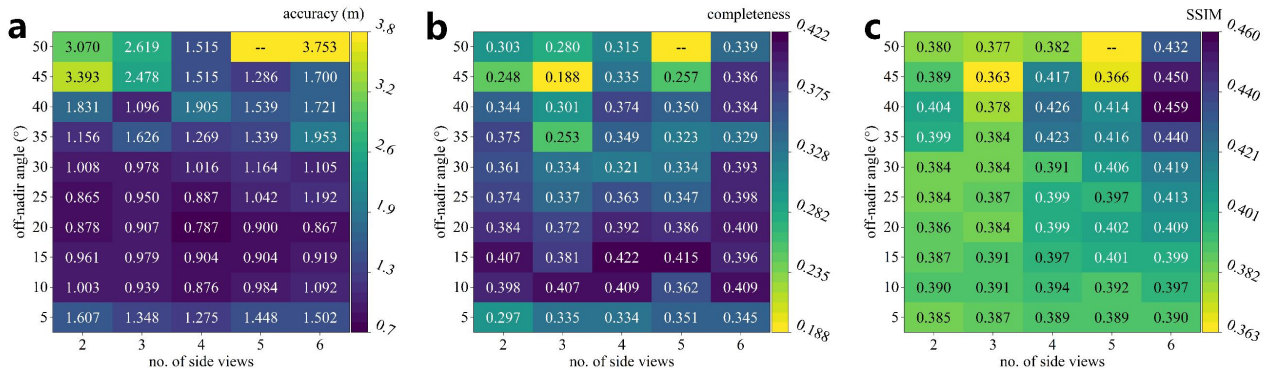

Figure 7. Joint effects of the number of side views and off-nadir angle on **(a)** reconstruction accuracy, **(b)** completeness, and **(c)** texture quality (SSIM). The combination of (n, α) = (5, 45°) failed the reconstruction.

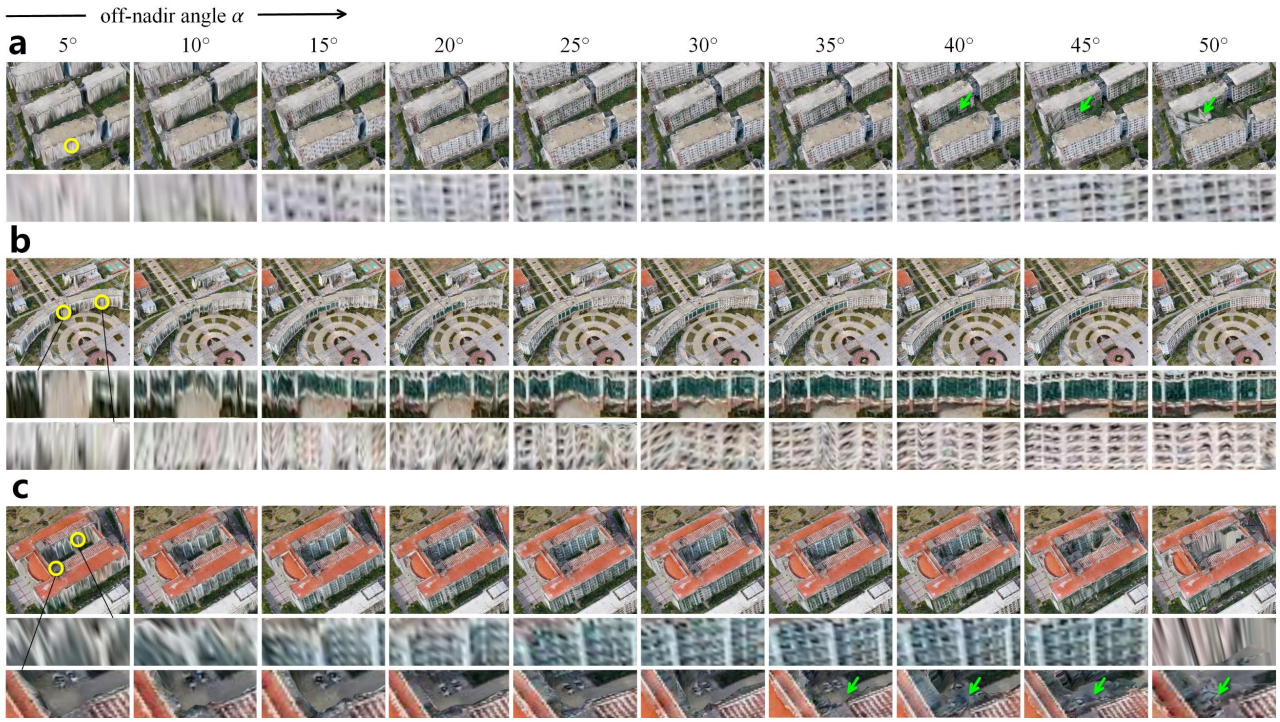

Figure 8. Detailed appearance of 3D building façades at different off-nadir angles, using four side-looking images**.** Magnified details are highlighted with yellow circles. Green arrows indicate areas with significant geometric structure errors. **a** Residential area with densely aligned houses; **b** Curved façade structure; **c** Courtyard (atrium) structure.

However, SSIM follows a different interaction pattern. As the off-nadir angle increases, the positive dependence of SSIM on the number of side views becomes increasingly evident. Nevertheless, SSIM becomes less meaningful in this situation as the underlying 3D geometry has already deteriorated.

To better illustrate the pattern of SSIM and texture quality in highly oblique angles, Figure 8 provides qualitative examples of reconstructed building façades under different off-nadir angles, with the number of side views fixed at four. As the off-nadir angle increases, the façade texture becomes clearer with reduced distortion. The significant geometric errors appear at $40°$ and larger angles (green arrows), yet the SSIM value at $40°$ is the highest within its column in the heatmap (Figure 6c), after which it begins to decline. This behavior reveals a delayed response to geometric degradation, indicating that SSIM can remain high even when the underlying geometry has already failed, and only drops once geometric distortions begin to visibly alter the texture.

5. Conclusions

This study presents a semi-simulation framework for analyzing multi-view imaging geometry in satellite-based urban 3D reconstruction. By isolating geometric factors, controlled experiments revealed an asymmetric U-shaped relationship between view diversity and reconstruction accuracy, with optimal performance achieved at moderate side-view angles around 15°-25° and four oblique observations, while excessive obliqueness degrades geometric stability. In dense urban environments, reconstruction completeness and texture fidelity exhibit different responses to geometric changes, highlighting a trade-off between visibility and geometric accuracy. Overall, off-nadir angle and view quantity dominate reconstruction performance, providing practical guidance for multi-satellite imaging design. Future work will consider radiometric coupling and extend the analysis to learning-based reconstruction methods.

Acknowledgements

This work was supported by the National Science Fund for Distinguished Young Scholars (Grant No. 62425102, China) and by the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS) Special Research Fund.

References

Belward, A.S., Skøien, J.O., 2015: Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. *ISPRS J. Photogramm. Remote Sens.*, 103, 115–128.

Beyer, R.A., Alexandrov, O., McMichael, S., 2018: The Ames Stereo Pipeline: NASA’s open source software for deriving and processing terrain data. *Earth Space Sci.*, 5, 537–548.

Bhushan, S., Shean, D., Alexandrov, O., Henderson, S., 2021. Automated digital elevation model (DEM) generation from very-high-resolution Planet SkySat triplet stereo and video imagery. *ISPRS J. Photogramm. Remote Sens.*, 173, 151–165.

Bosch, M., Kurtz, Z., Hagstrom, S., Brown, M., 2016: A multiple view stereo benchmark for satellite imagery. In: *2016 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pp. 1–9.

Cao, Y., Huang, X., 2021: A deep learning method for building height estimation using high-resolution multi-view imagery over urban areas: A case study of 42 Chinese cities. *Remote Sens. Environ.*, 264, 112590.

Catmull, E.E., 1974: A subdivision algorithm for computer display of curved surfaces. PhD thesis, The University of Utah.

De Franchis, C., Meinhardt-Llopis, E., Michel, J., Morel, J.M., Facciolo, G., 2014: An automatic and modular stereo pipeline for pushbroom images. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 2, 49–56.

Gómez, A., Randall, G., Facciolo, G., von Gioi, R.G., 2023: Improving the pair selection and the model fusion steps of satellite multi-view stereo pipelines. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pp. 6344–6353.

Gómez, A., Randall, G., Facciolo, G., von Gioi, R.G., 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. 844–853.

Kornus, W., Alamús, R., Ruiz, A., Talaya, J., 2006: DEM generation from SPOT-5 3-fold along track stereoscopic imagery using autocalibration. *ISPRS J. Photogramm. Remote Sens.*, 60, 147–159.

Lebègue, L., Cazala-Hourcade, E., Languille, F., Artigues, S., Melet, O., 2020: CO3D, a worldwide one one-meter accuracy DEM for 2025. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, 43, 299–304.

Leotta, M.J., Long, C., Jacquet, B., Zins, M., Lipsa, D., Shan, J., Xu, B., Li, Z., Zhang, X., Chang, S.-F., Purri, M., Xue, J., Dana, K., 2019: Urban semantic 3D reconstruction from multiview satellite imagery. In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1451–1460.

Li, W., Gao, F., Zhang, P., Li, Y., An, Y., Zhong, X., Lu, Q., 2021: Research on multiview stereo mapping based on satellite video images. *IEEE Access*, 9, 44069–44083.

Liu, Y., Zuo, X., Tian, J., Li, S., Cai, K., Zhang, W., 2021: Research on generic optical remote sensing products: a review of scientific exploration, technology research, and engineering application. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 14, 3937–3953.

Peng, T., Pi, Y., Chen, R., Pan, J., Cai, J., Wang, M., 2023: A combined side-slither relative radiometric calibration method

for non-collinear TDI-CCDS. *IEEE Geosci. Remote Sens. Lett.*, 20, 1–5.

Qin, R., 2019: A critical analysis of satellite stereo pairs for digital surface model generation and a matching quality prediction model. *ISPRS J. Photogramm. Remote Sens.*, 154, 139–150.

Qin, R., 2016: Rpc stereo processor (rsp)–a software package for digital surface model and orthophoto generation from satellite stereo imagery. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, 3, 77–82.

Rupnik, E., Pierrot-Deseilligny, M., Delorme, A., 2018: 3D reconstruction from multi-view VHR-satellite images in MicMac. *ISPRS J. Photogramm. Remote Sens.*, 139, 201–211.

Shean, D.E., Alexandrov, O., Moratto, Z.M., Smith, B.E., Joughin, I.R., Porter, C., Morin, P., 2016: An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery. *ISPRS J. Photogramm. Remote Sens.*, 116, 101–117.

Stucker, C., Schindler, K., 2022: ResDepth: A deep residual prior for 3D reconstruction from high-resolution satellite images. *ISPRS J. Photogramm. Remote Sens.*, 183, 560–580.

Tao, P., Xi, K., Niu, Z., Chen, Q., Liao, Y., Liu, Y., Liu, K., Zhang, Z., 2022: Optimal selection from extremely redundant satellite images for efficient large-scale mapping. *ISPRS J. Photogramm. Remote Sens.*, 194, 21–38.

Tuia, D., Schindler, K., Demir, B., Zhu, X.X., Kochupillai, M., Džeroski, S., Van Rijn, J.N., Hoos, H.H., Del Frate, F., Datcu, M., Markl, V., Le Saux, B., Schneider, R., Camps-Valls, G., 2024: Artificial Intelligence to Advance Earth Observation: A review of models, recent trends, and pathways forward. *IEEE Geosci. Remote Sens. Mag.*, 13, 119-141.

Waechter, M., Beljan, M., Fuhrmann, S., Moehrle, N., Kopf, J., Goesele, M., 2017: Virtual Rephotography: Novel View Prediction Error for 3D Reconstruction. *ACM Trans. Graph.*, 36(1), 8.

Wuhan Daspatial Intelligence Technology Co., 2025. GET3D-SAT software. https://daspatial.com/cn/cluster-satellite (December 2025).

Yang, B., Pi, Y., Li, X., Yang, Y., 2020: Integrated geometric self-calibration of stereo cameras onboard the ZiYuan-3 satellite. *ISPRS J. Photogramm. Remote Sens.*, 162, 173–183.

Yang, B., Wang, M., Xu, W., Li, D., Gong, J., Pi, Y., 2017: Large-scale block adjustment without use of ground control points based on the compensation of geometric calibration for ZY-3 images. *ISPRS J. Photogramm. Remote Sens.*, 134, 1–14.

Zhang, K., Snavely, N., Sun, J., 2019: Leveraging vision reconstruction pipelines for satellite imagery. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, pp. 0–0.