

Research on 3D Virtual Scene Reconstruction and Application Based on Multi-source Data Fusion

Hao Tang^{1,a}, Yuhang Wang^{1,b}, Xinwu Liu^{1,c}, Ziyi Wang^{1,d}, Chenghao Li^{1,e}, Zefan Zhang^{1,f}, Minfeng Xing^{1,g}, Shihua Li^{1,h*}

¹ School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China –
^atangh@std.uestc.edu.cn, ^b1357494918@qq.com, ^c805893709@qq.com, ^d1966422205@qq.com, ^e2146326331@qq.com,
^f1798339521@qq.com, ^gxingminfeng@uestc.edu.cn, ^hlishihua@uestc.edu.cn

Abstract

As digital cities evolve, the demand for 3D reconstruction increases, but challenges in accuracy and completeness remain. This study proposes a 3D virtual scene reconstruction and application framework, which overcomes the limitations of single-source data by incorporating multi-source data, improving accuracy, completeness, and interactivity. The method incorporates Unmanned Aerial Vehicle (UAV) photogrammetry, UAV oblique photogrammetry, and Backpack Laser Scanning (BLS) to generate high resolution mapping products, including Digital Orthophoto Map (DOM), Digital Surface Model (DSM), and dense point clouds. In this study, UAV images are utilized for 3D reconstructing large-scale scenes. For smaller scenes and complex individual buildings, two distinct data sources - oblique photogrammetry and BLS - are employed for modeling. The fusion of multi-source data addresses issues such as image blind spots and deformations, resulting in models with improved geometric accuracy and richer textural details. System development integrates a high resolution image viewer optimized with tiling technology and a 3D virtual reality interactive system constructed based on Unreal Engine 5, enabling immersive exploration and real - time interaction. This study offers a scalable solution for urban 3D reconstruction and provides tools for campus management and virtual tours, holding potential applications in the development of smart cities.

Keywords: 3D virtual campus, UAV data processing, 3D reconstruction, Mobile mapping technology, Point cloud processing

1. Introduction

With the development of digital and smart cities, urban simulation has evolved from 2D to 3D, and the demand for 3D reconstruction of urban scenes is increasing. As an essential information-based tool, 3D virtual systems are playing an increasingly important role in urban management and cultural dissemination (Tang and Wang, 2024). However, due to the limitations of data acquisition and processing technologies, the existing 3D virtual system development technologies still have shortcomings in terms of accuracy, completeness, and interactivity (Liu et al., 2023).

Reconstructing three-dimensional (3D) scenes is crucial for the development of 3D virtual systems. If the parameters of geographical objects are measured and 3D models are manually constructed, accuracy can be guaranteed. However, a considerable amount of time will be taken, and substantial labor costs will be incurred.

There are three main categories of methods for automatic 3D scene model construction, based on the data type: image-based, point-cloud-based, and multi-source data fusion-based reconstructions.

Using only image information is the most widely-used method for developing 3D urban systems at present. This method primarily employs Structure from Motion (SfM) or Multi-view stereo (MVS) algorithms to generate a large-scale 3D model in a short time from the aerial image sequences captured by Unmanned Aerial Vehicle (UAV) (Méndez-Barroso et al., 2018; Rublee et al., 2011; Seitz et al., 2006). The advantages of these methods are that the data acquisition is fast, it covers a large area, and the entire scene can be covered in a short time. Additionally, the sensor's characteristics allow it to capture texture information of geographical objects. However, when UAV capture images, visual blind spots often occur due to the mutual coverage and occlusion between geographical objects. Moreover, since this method relies solely on images to infer the properties of

geographical objects, errors are likely to occur in areas with poor image coverage. Therefore, although this method can quickly build large-scale models, its local accuracy is low, and there are many model deformations and distortions.

Regarding the reconstruction method that uses only point clouds, the current ways of collecting point cloud data mainly include: Terrestrial Laser Scanning (TLS), Backpack Laser Scanning (BLS), Mobile Laser Scanning (MLS), and UAV Laser Scanning (ULS), etc (Tang et al., 2024). Generally, point cloud data is used, followed by Delaunay triangulation to obtain the 3D model of geographical objects (Barber et al., 1996; Mao and Luo, 2025). Light Detection and Ranging (LiDAR), with its sensor characteristics, can provide high precision point clouds, which enables the generated models to also have high accuracy (Guo et al., 2021). Although LiDAR technology provides high-accuracy point clouds, its operation can be obstructed by obstacles such as pedestrians and trees, leading to missing data. Additionally, since LiDAR sensors are usually single-band, they cannot collect texture information of geographical objects. Therefore, while it can build high precision 3D models, these models may suffer from missing data and a lack of texture information due to data defects.

To overcome the limitations of single-source methods, this study proposes a multi-source data fusion algorithm that optimizes the registration process between UAV imagery and LiDAR point clouds, significantly enhancing model accuracy and detail. For instance, scholars Bodis-Szomoru et al. proposed the fusion of Airborne Laser Scanning (ALS) point clouds and MLS point clouds for modeling, which helps resolve the problem of over-smoothing in model surfaces caused by the use of ALS point clouds (Bodis-Szomoru et al., 2016). In addition, Liu et al. introduced a novel 3D model generation method based on the fusion of multi-source 3D data, including 3D point cloud data and 3D mesh data (Liu et al., 2023). In general, these methods require the fusion of multi-source data before model construction. Afterward, 3D models are built based on the point cloud data.

* Corresponding author

Although this approach offers high accuracy, it also introduces complexities in the construction process and large-scale model storage requirements.

This study proposes a multi-source data fusion framework, combining UAV imagery and BLS data to address the limitations of single-source methods, improving model accuracy and detail for large-scale scene reconstruction. For large-scale scenes, methods with lower detail precision but faster data acquisition are employed. In contrast, for irregular and complex buildings, methods with higher precision are adopted for construction. Furthermore, by integrating system development techniques, the development of a high resolution imagery viewer and a virtual reality interaction system has been realized. This research provides a reliable tool for scenarios such as campus management and virtual tours, and it can be further extended to the construction of urban 3D virtual scenes.

2. Study Data



Figure 1. Study area and data collection area. The red box is the oblique photogrammetry data acquisition range, and the blue box is the BLS data collection range.

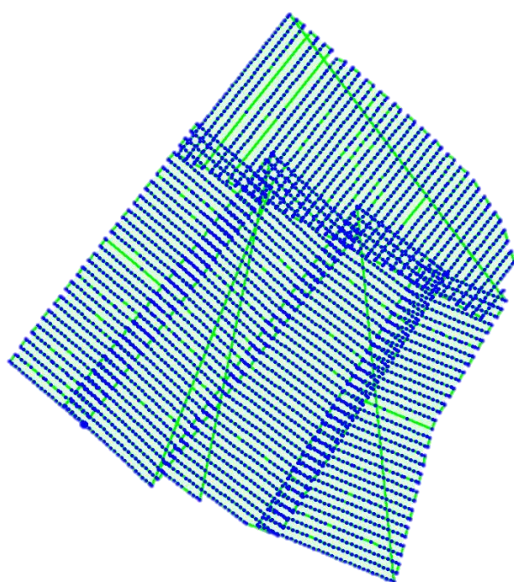


Figure 2. Top view of the initial image position. Green lines represent the route, and blue dot is the shooting location.

2.1 Study area.

The study area was selected in the Qingshuihe campus of University of Electronic Science and Technology of China, as shown in Figure 1. The campus spans approximately 309.06 hectares, with a total building area of about 1,490,000 square meters. The campus elevation ranges from 552.36 meters to 623.16 m. It is characterized by a predominantly modern architectural style, comprising structures of diverse heights and contemporary designs. It not only offers a wide range of data for the development of the 3D virtual campus system but also presents certain challenges in terms of data acquisition, processing, and modeling.

2.2 UAV photography data.

A DJI - Phantom 4 Pro UAV was used to obtain UAV images of the entire campus. The UAV was flown at an altitude of 100m, with a heading overlap of 80% and a side-to-side overlap of 75%. The UAV was flown at a speed of 10m/s, and the camera was a 1-inch CMOS, taking pictures at equal intervals along the direction of the course, with the lens pointing vertically downwards. A total of 4,732 photos were taken during the data acquisition process. The initial image position is shown in Figure 2.

2.3 UAV oblique photogrammetry data.

Oblique photogrammetry data were also collected using the DJI - Phantom 4 Pro. A total of three buildings in the study area were photographed. Among the three buildings, the irregular ones included the gymnasium and the library, as well as a regular building, as shown in red box of Figure 1. The UAV flight altitude was approximately 120 m, and photos were taken from five directions: front, back, left, right, and below. Each building had an average of about 500 images.

2.4 Backpack Laser Scanning Data.

To address the issue of insufficient accuracy in photogrammetry data, additional data were collected. The BLS data were acquired using the LiBackpack DGC50 system, which is equipped with two laser sensors and has a scanning distance of 120m. The system offers a relative accuracy of approximately 3cm, and the scanning frequency is 640,000 pts/s. Data from two buildings were collected to supplement the dataset, as shown in blue box of Figure 1. In the BLS data, the point cloud density is relatively high, providing more complete information about the building facades.

3. Method

This study focuses on 3D virtual scene reconstruction and application based on multi - source data fusion. The methodology consists of three stages: First, mapping data products, including the Digital Orthophoto Map (DOM), the Digital Surface Model (DSM), and the point cloud, are generated from processed UAV imagery. Second, 3D scene reconstruction is performed using three methods: UAV photography reconstruction, UAV oblique photogrammetry reconstruction, and multi - source data fusion reconstruction. These methods generate comprehensive and accurate 3D models. Third, system development integrates the generated data products and 3D models into user-friendly systems, including a high-resolution imagery viewer and a 3D virtual reality interactive system. The study aims to develop a comprehensive 3D virtual system that provides detailed, accurate, and interactive representations of urban scenes.

3.1 Mapping Data Product Generation

This section details the process of generating mapping data products from processed UAV imagery. The Digital Orthophoto Map (DOM) is generated to provide a high-resolution, orthorectified image of the study area. The Digital Surface Model (DSM) is produced to represent the elevation data of the surface features. Additionally, the point cloud of the entire study area is generated to offer a detailed 3D representation of the spatial distribution of objects and terrain. These products are crucial for subsequent 3D scene reconstruction and analysis.

3.1.1 DOM Generation. The DOM generation involves creating orthophoto imagery from the photos captured by UAV. DOM is imagery that has undergone geometric correction and can be used for precise measurement and analysis. First, the images captured by the UAV undergo preprocessing. Radiometric correction is applied to the images to eliminate the effects of illumination, atmosphere, and other factors. Subsequently, geometric correction is performed on the images using the camera's internal and external parameters (e.g., focal length, optical center) to remove geometric distortions caused by camera tilt and terrain undulations. The internal parameter matrix K is expressed as:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where f_x, f_y = focal length (in pixels) along the x-axis, y-axis
 c_x, c_y = the coordinate of the center of the image

Subsequently, image mosaicking is carried out. An optimization algorithm is employed to determine the seam positions between adjacent images, ensuring that the mosaicked image is seamless. Then, the blending algorithm is applied to the images at the seams to avoid discontinuities in grayscale values. Finally, the DOM can be generated.

3.1.2 DSM Generation. The generation of DSM involves creating a 3D representation of the objects' surface and its features, such as buildings and trees, using image data. The process begins with stereo matching, where feature points are extracted from the images using SIFT algorithm, and then matched across different views to establish correspondences (Lowe, 2004, 1999). The disparity, which is the difference in pixel coordinates of the same feature point in different images, is calculated.

Next, triangulation is performed using the disparity information and camera parameters to calculate the three-dimensional coordinates of the feature points. This results in a sparse point cloud. The depth Z can be calculated as follow:

$$Z = \frac{f \cdot B}{d} \quad (2)$$

where f = the focal length of the camera
 B = the baseline distance between the two cameras
 d = the disparity.

The point cloud is then filtered to remove noise and outliers, enhancing the accuracy and quality. This is followed by surface reconstruction, where the point cloud data is converted into a grid model. Then, the nearest neighbor interpolation is used to assign height values to each grid cell. After the above processing is completed, the DSM of the corresponding region is generated.

3.1.3 Point Cloud Generation. Point cloud generation involves creating a dense point cloud from the UAV image. the

dense point cloud generated accurately represents the geometric shape and spatial position of objects or environments. In this step, the dense point cloud is primarily generated based on the sparse point cloud from **Section 3.1.2**.

Specifically, to generate a dense point cloud, it is necessary to use the corresponding points in the sparse point cloud to infer the 3D positions of other neighboring points in the image. This involves depth estimation, which is the process of calculating the depth value for each pixel in the image. Depth estimation is carried out using the Semi-Global Matching (SGM) (Hirschmuller, 2008). The SGM algorithm approximates the 2D energy optimization problem by aggregating costs along multiple scanlines, thereby significantly reducing the processing time. First, for each pixel, the matching cost at different disparities is calculated. The matching cost is typically computed based on pixel intensity differences or mutual information. Then, the matching costs are aggregated along multiple directions to incorporate global contextual information. The aggregated cost $C(p, d)$ can be expressed as:

$$C(p, d) = \sum_{q \in N} \omega(q) \cdot cost(p, q, d) \quad (3)$$

where p = the current pixel
 q = the neighboring pixels
 N = the set of neighboring pixels
 $\omega(q)$ = the weight function
 $cost(p, q, d)$ = the matching cost between pixels p and q at disparity d

For each pixel, the disparity with the minimum aggregated cost is selected as the final disparity. Using the disparity and camera parameters, the depth Z for each pixel is calculated, resulting in a dense point cloud. After generating the dense point cloud, further post-processing is required to enhance the quality and accuracy, including filtering to remove noise and outliers.

3.2 3D Scene Reconstruction

This section outlines the method employed for 3D scene reconstruction. The process encompasses three primary approaches: UAV photography reconstruction, UAV oblique photogrammetry reconstruction, and multi-source data fusion reconstruction. These methods complement one another, providing different frameworks for reconstructing accurate 3D models.

3.2.1 UAV Photography Reconstruction. In this section, large-scale 3D models are generated based on the dense point cloud generated in **Section 3.1.3**. The process involves three main steps. First, the dense point cloud is transformed into a triangular mesh model using the Delaunay triangulation algorithm. This ensures that the resulting mesh has a well-structured topology and high geometric accuracy. Then, the generated mesh is optimized to reduce the number of polygons and enhance the smoothness and efficiency of the model. Laplacian smoothing algorithm is used for this optimization. Finally, the color information from the images is mapped onto the mesh model to create realistic textures.

3.2.2 UAV Oblique Photogrammetry Reconstruction. UAV oblique photogrammetry reconstruction is an important method for generating detailed and accurate 3D models of buildings. This technology uses images captured by UAVs from multiple angles to create a comprehensive 3D representation, which is particularly suitable for complex and irregularly shaped buildings (Verykokou and Ioannidis, 2018). Compared to

traditional orthophotography, oblique photogrammetry offers significant advantages in handling large scenes, providing richer details and greater geometric accuracy (Zhang et al., 2020). In this step, the same processing procedure as that for the photography data is applied to the oblique photogrammetry data. Since oblique photogrammetry captures images from multiple different angles, the dense point cloud it generates retains more building facade information, resulting in a more realistic model after triangulation.

3.2.3 Multi-source Data Fusion Reconstruction. Models reconstructed through imagery often suffer from poor accuracy in fine details and areas not covered by the camera, it is necessary to fuse other data for reconstruction. In this study, BLS was chosen to supplement the photogrammetry data.

Firstly, data registration is performed. The BLS data and the photogrammetry point cloud are dimensionally reduced to obtain their corresponding DSM. After dimensional reduction, normalization and binarization are carried out to highlight areas with significant elevation changes. Then, the KAZE operator is used for feature extraction on the binarized image (Alcantarilla et al., 2012). Subsequently, the transformation matrix is calculated based on the extracted features to match the images. Further, the transformation matrix obtained from image registration is used to guide point cloud registration, ultimately obtaining point clouds under the same coordinate system.

Next, the BLS point cloud is colored using the nearest-neighbor interpolation algorithm, based on the color of the photogrammetry point cloud. This results in two point clouds with RGB information.

Finally, the registered point cloud is down-sampled to reduce the data volume. After down-sampling, Delaunay triangulation algorithm is performed on the point cloud to obtain the final 3D model.

3.3 System Development

The development of the 3D virtual system involves the integration of high-resolution imagery and 3D models to create an interactive and immersive environment. This section outlines the key components of the system development, including the high-resolution imagery viewer and the 3D virtual reality interactive system.

3.3.1 High-resolution Imagery Viewer. Due to the high resolution of the DOM generated in *section 3.1.1*, it loads slowly in common image viewers. This section aims to develop a high-performance high-resolution imagery viewer and implement some interactive features.

In this study, high-resolution, large-scale imagery is compressed and divided into different levels using map tiling technology, thereby optimizing image access speed and memory usage (Clouston and Peterson, 2014; Lin et al., 2019). This enables rapid zooming in and out, panning, and high-detail viewing of imagery. Furthermore, based on viewing, building location and introduction functions are also provided.

3.3.2 3D virtual Reality Interactive System. This section, based on the models in *Section 3.2*, has developed a 3D virtual reality interactive system using Unreal Engine 5 (UE 5) (Li et al., 2023; Wang et al., 2024). The system allows character movements such as walking, jumping, and flying within the scene, offering an immersive user experience. Leveraging the capabilities of UE5, such as high-fidelity rendering and physics simulation, the system can respond to user interactions in real-time and view buildings and structures from different angles.

4. Results

4.1 Data Products

During the processing of UAV imagery, three types of data products were generated: the DOM, the DSM, and the point cloud for the entire area. These products were then used to generate the 3D model of the area. The DOM, DSM, and 3D model are shown in Figure 3 (a), (b), and (c), respectively. Among them, the resolution of both DOM and DSM is 0.02m. In the point cloud, there are a total of 545,197,552 points, with an average point cloud density is about 154.32 pts/m³.

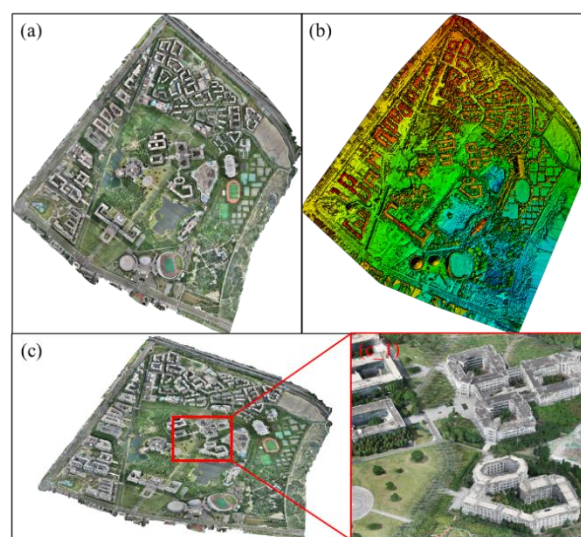


Figure 3. Data product result. (a), (b) and (c) is DOM, DSM and 3D model, respectively. (c_1) is local view of the model

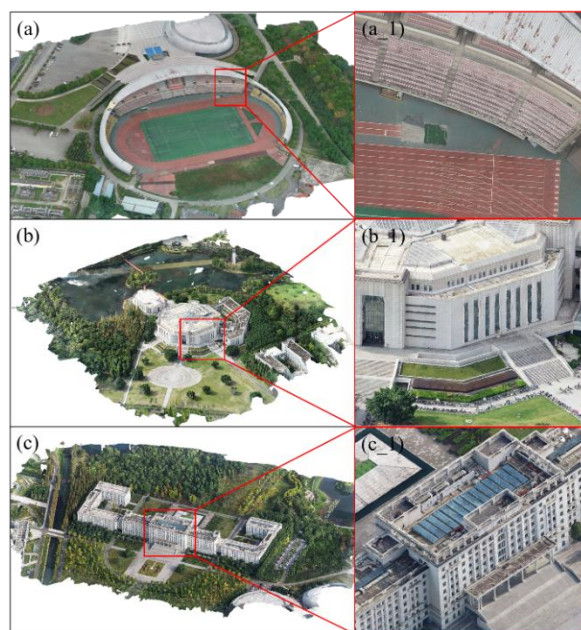


Figure 4. Oblique Photogrammetry reconstructed results. (a) and (b) represent irregular buildings, and (c) represents regular building. (a_1), (b_1), and (c_1) are the corresponding detail views.

4.2 3D Model

4.2.1 UAV Photography Model. The results of the generated 3D model are shown in Figure 3(c), with a detailed view in (c_1).

The model size is approximately $2193.79 \times 2464.76 \times 93.27$ m ($x \times y \times z$), consisting of about 1,000,000 faces. This model effectively represents the distribution of buildings, vegetation, and other land features across the entire study area. However, when it comes to individual buildings or land features, the level of detail is relatively poor, and there are noticeable deformations and other issues.

	Box Dimensions (m)			Faces
	X	Y	Z	
A	389.04	520.09	40.69	6,268,575
B	521.77	463.34	56.72	7,466,351
C	666.51	639.82	61.63	17,880,741
Average	525.77	541.08	53.01	10,538,556

Table 1. Oblique photogrammetry model parameter.



Figure 5. Point cloud registration result. RGB is shown for the photogrammetric point cloud and height information is shown for the BLS point cloud. (a_1) and (a_2) are overall views. (b_1) and (b_1) are top zoom views. (c_1) and (c_2) are side zoom views.

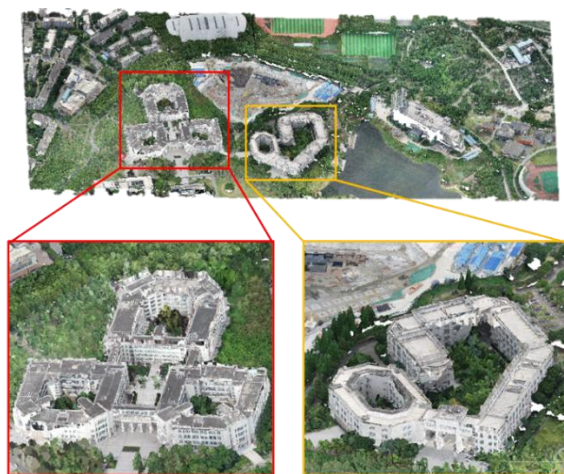


Figure 6. Multi-source Data Fusion reconstructed results.

4.2.2 UAV Oblique Photogrammetry Model. The results of oblique photogrammetry modeling are shown in Figure 4. This model provides clearer details of various buildings, as illustrated in the detail view on the right. Compared to the photogrammetry model, although this model has a smaller scale, it has more faces. The parameters of the three models are presented in Table 1, which correspond to the serial numbers in Figure 4. The average

size of the three models is $525.77 \times 541.08 \times 53.01$ m, and on average, they have 10,538,556 faces. In this model, it is possible to accurately measure parameters such as the length and size of land features. Moreover, this model can also be utilized for the development of a 3D virtual system.

4.2.3 Multi-source Data Fusion Model. 3D reconstruction through the fusion of multi - source data involves the registration of different point clouds. The registration results for point clouds of two different buildings are shown in Figure 5. As seen, both buildings have achieved good registration effects. The model constructed using the registered point clouds is shown in Figure 6. Compared to the oblique photogrammetry model, this model preserves more details. Additionally, the deformation issues present in the UAV photography blind spots have been largely addressed. The final model, constructed by fusing multi - source data, is approximately $607.35 \times 1133.71 \times 92.25$ m in size and consists of 56,926,748 faces.

4.3 System Development

The high-resolution image viewer results are shown in Figure 7, where (a), (b), and (c) represent different zoom levels. In this system, users can quickly zoom in and out of the map. They can also search and locate buildings by selecting building's names and view the corresponding building introductions.

The 3D virtual interaction system is shown in Figure 8. It allows users to perform actions such as walking, jumping and flying in the virtual scene. This system has strong interactivity.

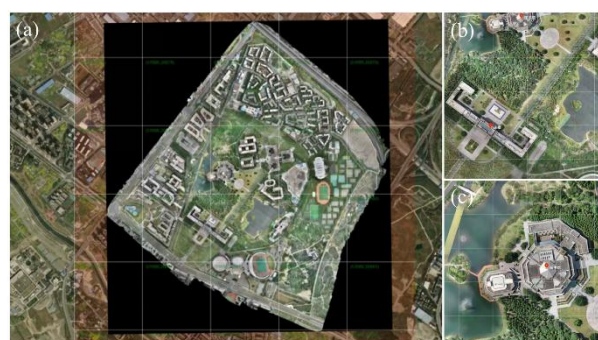


Figure 7. High resolution image viewer interface screenshot. (a), (b) and (c) represent different zoom levels.

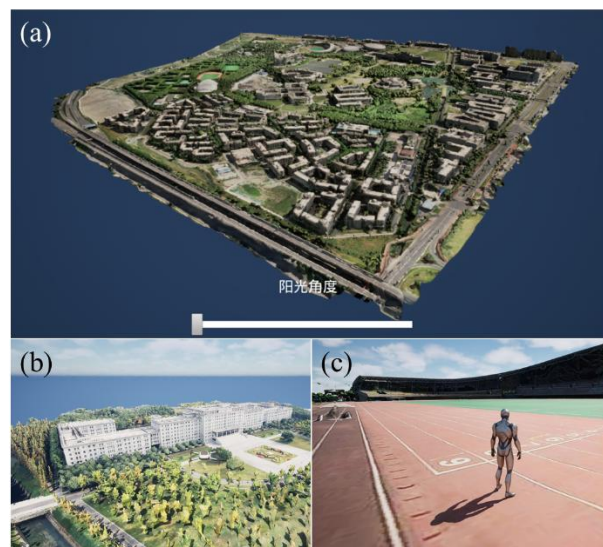


Figure 8. 3D virtual Reality Interactive System interface screenshot.

5. Discussion

5.1 Research Framework

The primary objective of this study is to reconstruct and apply 3D virtual scenes. To obtain comprehensive and precise data, this research employed a variety of data acquisition methods, including UAV photography, UAV oblique photogrammetry, and BLS. UAV photography provided extensive image data, while UAV oblique photogrammetry captured images of specific buildings from multiple angles. BLS was utilized to supplement the deficiencies in detail and accuracy of the photogrammetric data. This multi-source data acquisition approach not only ensured the comprehensiveness of the data but also enhanced its accuracy and reliability.

During the data processing stage, several data products were generated, including the DOM, DSM, and point cloud. The generation of DOM involved geometric correction and image mosaicking, ensuring precise measurement and analysis of the imagery. The DSM was produced through stereo matching and triangulation, utilizing camera parameters to generate a sparse point cloud, which was then filtered and subjected to surface reconstruction. The point cloud was generated using the SGM algorithm, producing a dense point cloud that accurately represented the geometric shape and spatial position of objects or environments.

In the 3D reconstruction phase, this study adopted three main methods: UAV photography reconstruction, UAV oblique photogrammetry reconstruction, and multi-source data fusion reconstruction. These methods complemented each other, providing 3D model reconstruction within different frameworks. UAV photography reconstruction involved converting the dense point cloud into a triangular mesh model, followed by optimization and texture mapping, to generate large-scale 3D models. UAV oblique photogrammetry reconstruction utilized images captured from multiple angles to produce detailed and accurate 3D models, particularly suitable for complex and irregularly shaped buildings. Multi-source data fusion reconstruction involved registering BLS data with photogrammetric point clouds to generate more precise 3D models, addressing the shortcomings of photogrammetric data in terms of detail and blind spots.

In the system development stage, this study integrated the generated data products and 3D models into user-friendly systems, including a high-resolution imagery viewer and a 3D virtual reality interactive system.

5.2 3D Model Reconstruction

While the system generally achieves satisfactory outcomes, certain errors and shortcomings still exist, which are primarily reflected in the model's inaccuracies.

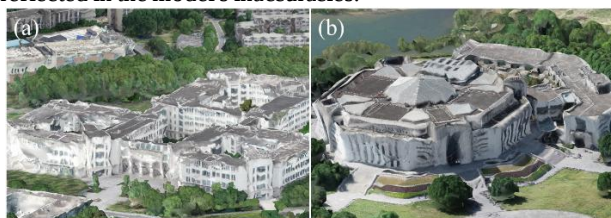


Figure 9. The area of the photogrammetry model that is obviously deformed

5.2.1 UAV photogrammetry model. Regarding the UAV photogrammetry model, despite its extensive coverage, it only provides nadir imagery. This type of data offers relatively good

representations of building rooftops, tree canopies, and prominent land features. However, it is less effective at capturing building facades and the understory structure of forests. These areas often exhibit poor or even absent data. As a result, noticeable deformations can be observed in the reconstructed models, as shown in Figure 9.

5.2.2 UAV oblique photogrammetry model. To address the limitations of the photogrammetry model, the oblique photogrammetry approach was employed to model certain land objects. The models constructed using this method have demonstrated relatively good performance in representing building facades, enabling accurate measurement of parameters such as object length and area. However, due to the limitations of UAV photography, significant errors remain in some blind spots. For instance, in areas like building passages, deformations or even voids can be observed, as shown in Figure 10(a), (b), and (c). Additionally, at the model's edges, insufficient photography coverage results in further deformations and voids, as shown in Figure 10(d).

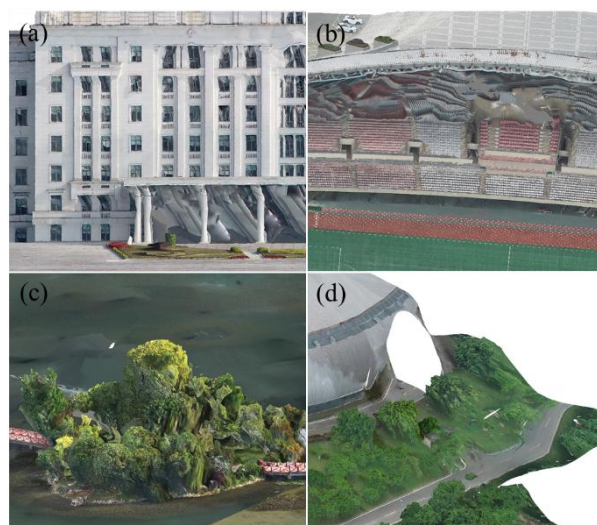


Figure 10. Examples of the main errors in the oblique photogrammetry model

5.2.3 Multi-source Data Fusion Model. For models generated through multi-source data fusion, the issues of model deformation and insufficient detail have been significantly mitigated. This method has compensated for the shortcomings of single-source modeling approaches. The main errors stem from the registrant process. However, this method has efficiency-related issues, such as high memory consumption and low generation efficiency. Future research should focus on developing more efficient algorithms to enhance the modeling process with multi-source data.

6. Conclusions

This study demonstrates the effectiveness of multi-source data fusion in 3D virtual campus reconstruction, addressing the limitations of single-source methods. By integrating UAV imagery, oblique photogrammetry, and BLS data, the framework achieves a balance between large-scale coverage and high-detail modeling. The developed systems—a high-performance imagery viewer and a UE5-based virtual reality platform—enhance usability and interactivity for urban management applications. Despite these advancements, challenges such as computational inefficiencies and registration inaccuracies underscore the need for optimized algorithms in future work. The research

underscores the importance of hybrid methodologies in geospatial modeling and provides a foundation for scalable smart city solutions.

Acknowledgements

The author would like to thank the Scientific research and education program of the School of Resource and Environment of the University of Electronic Science and Technology of China for providing experimental equipment and support.

References

- Alcantarilla, P.F., Bartoli, A., Davison, A.J., 2012. KAZE features, in: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (Eds.), *Computer Vision – ECCV 2012: Lecture Notes in Computer Science*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 214–227. doi.org/10.1007/978-3-642-33783-3_16
- Barber, C.B., Dobkin, D.P., Huhdanpaa, H., 1996: The quickhull algorithm for convex hulls. *ACM Trans. Math. Software* 22, 469–483. doi.org/10.1145/235815.235821
- Bodis-Szomoru, A., Riemenschneider, H., Van Gool, L., 2016: Efficient volumetric fusion of airborne and street-side data for urban reconstruction, in: 2016 23rd International Conference on Pattern Recognition (ICPR). Presented at the 2016 23rd International Conference on Pattern Recognition (ICPR), IEEE, Cancun, pp. 3204–3209. doi.org/10.1109/ICPR.2016.7900128
- Clouston, A., Peterson, M.P., 2014: Tile-based mapping with opacity, in: *Modern Cartography Series*. Elsevier, pp. 79–95. doi.org/10.1016/B978-0-444-62713-1.00006-4
- Guo, Q., Su, Y., Hu, T., Guan, H., Jin, S., Zhang, J., Zhao, X., Xu, K., Wei, D., Kelly, M., Coops, N.C., 2021: Lidar boosts 3D ecological observations and modelings: a review and perspective. *IEEE Geosci. Remote Sens. Mag.* 9, 232–257. doi.org/10.1109/MGRS.2020.3032713
- Hirschmuller, H., 2008: Stereo processing by semiglobal matching and mutual information. *IEEE Trans. Pattern Anal. Mach. Intell.* 30, 328–341. doi.org/10.1109/TPAMI.2007.1166
- Li, Y., Jiang, L., Xu, L., Xiangli, Y., Wang, Z., Lin, D., Dai, B., 2023: MatrixCity: a large-scale city dataset for city-scale neural rendering and beyond, in: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Presented at the 2023 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE, Paris, France, pp. 3182–3192. doi.org/10.1109/ICCV51070.2023.00297
- Lin, S.-S., Yang, J.-Y., Syu, H.-S., Lin, C.-H., Pai, T.-W., 2019: Automatic generation of puzzle tile maps for spatial-temporal data visualization. *Comput. Graphics* 82, 1–12. doi.org/10.1016/j.cag.2019.05.002
- Liu, W., Zang, Y., Xiong, Z., Bian, X., Wen, C., Lu, X., Wang, C., Marcato, J., Gonçalves, W.N., Li, J., 2023: 3D building model generation from MLS point cloud and 3D mesh using multi-source data fusion. *Int. J. Appl. Earth Obs. Geoinf.* 116, 103171. doi.org/10.1016/j.jag.2022.103171
- Lowe, D.G., 2004: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* 60, 91–110. doi.org/10.1023/B:VISI.0000029664.99615.94
- Lowe, D.G., 1999: Object recognition from local scale-invariant features, in: *Proceedings of the Seventh IEEE International Conference on Computer Vision*. Presented at the Proceedings of the Seventh IEEE International Conference on Computer Vision, IEEE, Kerkyra, Greece, pp. 1150–1157 vol.2. doi.org/10.1109/ICCV.1999.790410
- Mao, H., Luo, J., 2025: Approach to 3D SLAM for mobile robots based on point-line features and superpixel segmentation in dynamic environments. *Measurement* 247, 116742. doi.org/10.1016/j.measurement.2025.116742
- Méndez-Barroso, L.A., Zárate-Valdez, J.L., Robles-Morúa, A., 2018: Estimation of hydromorphological attributes of a small forested catchment by applying the structure from motion (SfM) approach. *Int. J. Appl. Earth Obs. Geoinf.* 69, 186–197. doi.org/10.1016/j.jag.2018.02.015
- Rublee, E., Rabaud, V., Konolige, K., Bradski, G., 2011: ORB: an efficient alternative to SIFT or SURF, in: 2011 International Conference on Computer Vision. Presented at the 2011 IEEE International Conference on Computer Vision (ICCV), IEEE, Barcelona, Spain, pp. 2564–2571. doi.org/10.1109/ICCV.2011.6126544
- Seitz, S.M., Curless, B., Diebel, J., Scharstein, D., Szeliski, R., 2006: A comparison and evaluation of multi-view stereo reconstruction algorithms, in: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1 (CVPR'06). Presented at the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1 (CVPR'06), IEEE, New York, NY, USA, pp. 519–528. doi.org/10.1109/CVPR.2006.19
- Tang, H., Li, S., Su, Z., He, Z., 2024: Cluster-based wood-leaf separation method for forest plots using terrestrial laser scanning data. *Remote Sens.* 16, 3355. doi.org/10.3390/rs16183355
- Tang, H., Wang, Z., 2024: Urban building coverage areas calculation and 3D modeling based on UAV LiDAR, in: IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Athens, Greece, pp. 8638–8641. doi.org/10.1109/IGARSS53475.2024.10642745
- Verykokou, S., Ioannidis, C., 2018: Oblique aerial images: a review focusing on georeferencing procedures. *Int. J. Remote Sens.* 39, 3452–3496. doi.org/10.1080/01431161.2018.1444294
- Wang, S., Zhang, J., Wang, P., Law, J., Calinescu, R., Mihaylova, L., 2024: A deep learning-enhanced digital twin framework for improving safety and reliability in human-robot collaborative manufacturing. *Rob. Comput. Integr. Manuf.* 85, 102608. doi.org/10.1016/j.rcim.2023.102608
- Zhang, X., Zhao, P., Hu, Q., Ai, M., Hu, D., Li, J., 2020: A UAV-based panoramic oblique photogrammetry (POP) approach using spherical projection. *ISPRS J. Photogramm. Remote Sens.* 159, 198–219. doi.org/10.1016/j.isprsjprs.2019.11.016