remote sensing

MDPI

Article

Improving Real-Scene 3D Model Quality of Unmanned Aerial Vehicle Oblique-Photogrammetry with a Ground Camera

Jinghai Xu [1,*], Suya Zhang [1], Haoran Jing [1], Craig Hancock [2], Peng Qiao [3], Nan Shen [1] and Karen B. Blay [2]

[1] School of Geomatics Science and Technology, Nanjing Tech University, 30 South Puzhu Road, Nanjing 211899, China; zhangsuya@njtech.edu.cn (S.Z.)
[2] School of Architecture, Building and Civil Engineering, Loughborough University, Sir Frank Gibb Building, Loughborough, Leics LE11 3TU, UK; c.m.hancock@lboro.ac.uk (C.H.); k.b.blay@lboro.ac.uk (K.B.B.)
[3] Science and Technology Development Center, Jiangsu Provincial Department of Housing and Urban-Rural Development, 287 North Jiangdong Road, Nanjing 210036, China
* Correspondence: xujinghai@njtech.edu.cn

Abstract: In UAV (unmanned aerial vehicle) oblique photogrammetry, the occlusion of ground objects, particularly at their base, often results in low-quality real-scene 3D models. To address this issue, we propose a method to enhance model quality by integrating ground-based camera images. This innovative image acquisition method allows the rephotographing of areas in the 3D model that exhibit poor quality. Three critical parameters for reshooting are the reshooting distance and the front- and side-overlap ratios of reshooting images. The proposed method for improving 3D model quality focuses on point accuracy, dimensional accuracy, texture details, and the triangular mesh structure. Validated by a case study involving a complex building, this method demonstrates that integrating ground camera photos significantly improves the overall quality of the 3D model. The findings show that optimal settings for reshooting include a distance (in meter units) of 1.5–1.6 times the camera's focal length (in millimeter units), a front overlap ratio of 30%, and a side overlap ratio of 20%. Furthermore, we conclude that an overlap rate of 20–30% in reshooting is a usable value.

Keywords: oblique photogrammetry; ground camera; reshooting parameters; real-scene 3D modeling; quality assessment

check for updates

Citation: Xu, J.; Zhang, S.; Jing, H.; Hancock, C.; Qiao, P.; Shen, N.; Blay, K.B. Improving Real-Scene 3D Model Quality of Unmanned Aerial Vehicle Oblique-Photogrammetry with a Ground Camera. *Remote Sens.* **2024**, *16*, 3933. https://doi.org/10.3390/rs16213933

Academic Editor: Filiberto Chiabrando

Received: 10 September 2024
Revised: 18 October 2024
Accepted: 18 October 2024
Published: 22 October 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Real-scene 3D models are precise digital models with a 3D structure, a natural appearance, and texture from real-scene geographic data [1]. They represent real, three-dimensional, and spatial–temporal information about human production, life, and ecological space. With the development of the "Realistic 3D China" project, they are considered a significant new kind of infrastructure in China [2]. The models reproduce real-world terrain, buildings, and environmental features with information technologies and provide important support for many application areas, such as disaster prevention and mitigation, urban planning, and cultural relic protection [3–6]. So, improving the accuracy and authenticity of these 3D models is an important issue [7,8].

Generally, real-scene 3D models are generated by oblique photogrammetry of unmanned aerial vehicles (UAVs). Fitted with multiple lens cameras on UAVs, the 3D modeling technique swiftly creates real-scene 3D models by gathering photos from multiple perspectives. The oblique photogrammetry approach can obtain large-area real-scene 3D model data and have the advantages of high realism, high accuracy, and high timeliness [9,10]. However, UAV flight is sensitive to the influence of occlusions, lighting conditions, and terrain restrictions of the survey areas, such as trees around a building. Sometimes, accurate and complete information on terrain or features even cannot be obtained by the UAV [11]. This often results in severe deformation and blurred textures in the 3D model [12].

Some other modeling techniques may complement the oblique photogrammetry method, such as ground-close-range photo modeling and laser scanning 3D modeling. Ground-close-range photo modeling is a method of three-dimensional modeling using ground photo data captured at close range [13]. The method has the advantages of relatively low cost, flexibility, and adaptability to complex terrain. It is suitable for high-precision small-scale modeling in ancient building restoration, landscape design, and small-scale engineering projects [14,15]. Three-dimensional modeling based on laser scanning is a method of obtaining surface point cloud data of a scene or object through a laser scanner and using these data to build a 3D model [16]. Arastounia et al. [17] elaborate on the working principle of 3D modeling based on ground laser point cloud data, the modeling characteristics, and different methods of point cloud data processing and 3D modeling. Lei et al. [18] and Henn et al. [19] illustrate through some modeling examples that the laser scanning modeling method has the advantages of high precision, the ability to capture details and complex structures, and not being affected by external factors such as light and weather. It is suitable for fields that require high precision and detail in heritage protection, engineering surveying, and geological exploration.

To solve the modeling limitation of only using UAV aerial photos, other data sources are considered supplementary data for joint modeling to leverage the full advantages of different types of data and improve data integrity [20,21]. Abayowa et al. [22] propose the fusion of UAV aerial photo and ground laser scanning point cloud data for modeling and demonstrate that the iterative nearest point algorithm can be used to register the two types of data accurately. Han et al. [23], Jo et al. [24], and Zhou et al. [25] realize the fusion modeling of UAV aerial photos and ground laser scanning point cloud data in traffic management, bridge construction, and cultural heritage protection, respectively. The accuracy analysis of the fused model confirms the feasibility of multi-source data fusion modeling. However, terrestrial laser scanners are more expensive, and their operation requires a professional background and technical training. In addition, the amount of point cloud data is huge, and the required data cleaning, registration, and splicing increase the workload and time cost [26,27]. In areas with rugged terrain or dense vegetation cover, laser scanners can be difficult to set up, making it challenging to obtain comprehensive and accurate data. So, we consider using a camera to reshoot areas corresponding to places of the real-scene 3D model with poor quality. By reshooting close-ups with the camera from the ground view, high-resolution photos are obtained that capture detailed information and surface features about survey objects. Compared with laser scanning equipment, the ground camera can be low-cost, simply operated, and easily deployed and adjusted in various complex terrains and environments [28]. Meanwhile, reshooting photos of specific areas results in a relatively small amount of photo data, making data processing and modeling easier. Studies show that it is possible to reshoot using cameras to solve the problem of "occlusion", restore the original structural characteristics of near-ground features, and more accurately reflect the colors and textures of a scene [29].

Many studies provide theoretical support for UAV aerial photo matching, fusion, and 3D modeling. Li et al. [30] present a real-time matching method for UAV aerial imagery to address the issue of insufficient image contrast in low-illumination environments. Shahid et al. [31] suggest a two-step matching framework, including CNN-based coarse-matching and fine-tuning standard off-the-shelf image-matching algorithms according to spectral, temporal, and flow features. Zhu et al. [32] describe an integration method for aerial and ground images that leverages photogrammetric mesh models for aerial–ground image matching. Zhang et al. [33] introduce a UAV aerial image stitching algorithm that utilizes semantic segmentation and the ORB (Oriented FAST and Rotated BRIEF) algorithm. Then, photo stitching and fusion are realized by a monotonic matrix and a weighted fusion algorithm. Sun et al. [34] propose a bundle adjustment approach specifically for oblique aerial photogrammetry modeling, designed to decrease the quantity of unknowns and simplify the dimensions involved in non-linear optimization. This method boasts a marked enhancement in efficiency and can be used under the limitation of computer memory.

Although these theoretical studies have made commendable progress, we still face challenges in how to solve the "occlusion" problem from application perspectives. These challenges include determining the kind of instrument that is suitable for ground photo acquisition, how to set key parameters for reshooting (i.e., reshooting route design, overlap ratios of reshooting photos), to what extent the 3D model's quality can be improved by fusing UAV aerial photos and ground camera photos, and how to realize fusion with various photogrammetric software programs, such as Pix4Dmapper, ContextCapture, and Agisoft Metashape. Unfortunately, these challenges may be difficult to solve through theoretical deduction, but experimental methods may accumulate application experience to deal with them. The application demand is tremendous, but exploration is limited. For example, Nesbit et al. [35] find that in complex terrain scenes, obtaining oblique photos at an angle of $\geq$5.0° between the camera axis and the lowest point can be used to supplement UAV aerial image modeling, thus reducing image data errors and improving spatial accuracy. Baramsyah et al. [36] recommend using the stripe method and convergence method for photo acquisition in order to obtain higher accuracy data when performing photogrammetry on slopes. In contrast to the current studies, we will focus on the application issue with an experimental method.

After the establishment of a real-scene 3D model, the quality of the model needs to be evaluated. Evaluating the quality of the model can help us better understand the differences and errors between the model and the real world. Cheng et al. [37] and Zhang et al. [38], respectively, combine geometric quality and structural similarity with the human visual system, fully considering the visual influence of the observer and the structural information retention to evaluate the quality of the model. Iwaszczuk et al. [39] evaluate the matching between photos and 3D building models by comparing the textures of different sequences with different temporal and directional parameters. Elberink et al. [40] comprehensively evaluate a real-scene 3D model's quality from the following aspects: topological structure relationship, point data accuracy, and point data average density. Tran et al. [41] indicate that the quality evaluation of real-scene 3D models can be carried out in the following aspects: qualitative evaluation and quantitative evaluation.

We used a complex building as a case study to explore how to improve the quality of real-scene 3D models through ground camera reshooting. This includes determining the reshooting route and parameters, using ContextCapture to achieve fusion, and analyzing the resulting improvement in model quality. The remainder of this paper is organized as follows: (1) the method for determining the parameters of ground camera reshooting is detailed in Section 2; (2) the proposal of the quality assessment method for real-scene 3D models is presented in Section 3; (3) the application of methods and optimization of ground camera reshooting parameters are shown in Section 4; and (4) the conclusions are finally reported in Section 5.

2. Route Design on Ground Camera Reshooting

2.1. Reshooting Instrument

We engineered a novel ground photography instrument that utilizes a conventional camera for capturing images. As shown in Figure 1, a Cannon EOS M6 camera was employed in this study, complemented by an array of accessories including a 5 m extendable selfie stick, a stable tripod, a smartphone, a 100 m measuring tape, painting fluid, and correction fluid. The instrument is operated using Camera Connect software Version 3.1.10.49, which facilitates the synchronization of the camera, selfie stick, and tripod in a vertical setup suitable for ground-level photography. This software also enables remote control of the camera to adjust photo capture and the angles of the shots.

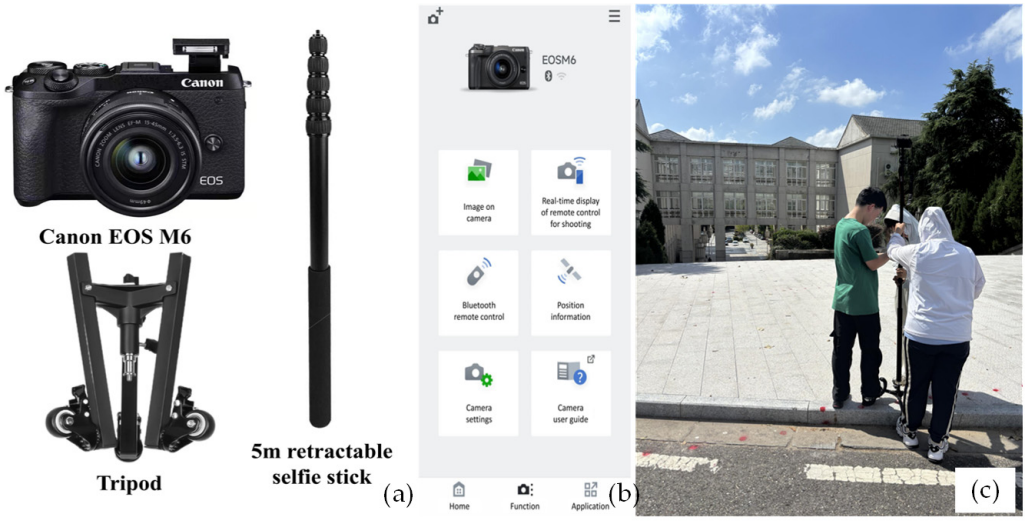


Figure 1. Components of the reshooting instrument: (**a**) EOS M6 camera, 5 m retractable selfie stick, and stable tripod; (**b**) the main interface of Camera Connect software; and (**c**) the reshooting instrument.

Detailed information about different parts of the ground photography instrument includes the following:

EOS M6 Camera: The EOS M6 is a regular single-lens reflex camera with a sensor size of 22.3 mm × 14.9 mm and a shooting capacity of 24.2 million pixels. This camera supports an ISO range of 100 to 25,600, which is suitable for a wide range of shooting scenarios. It is equipped with a Canon EF-M lens mount, and the commonly used standard zoom lens can cover a focal length range of 15 mm to 45 mm, making it suitable for lightweight photography needs that are convenient to carry.

Retractable selfie stick (5 m): The selfie stick consists of a selfie stick handle with a fixed length and a seven-section telescopic rod with a retractable range of 1 m to 5 m. The selfie stick is made of carbon fiber, which has excellent strength and rigidity. Compared with aluminum alloy selfie sticks, carbon fiber selfie sticks are more stable and have a higher load capacity. The selfie stick is lightweight, making it easy to carry and operate.

Camera Connect Software: Camera Connect is an application developed by Canon that aims to connect Canon cameras directly to smartphones or tablets through wireless connectivity, making photo shooting and sharing more convenient and real time. The app allows users to view the camera in real time and control the camera's shooting parameters remotely. The captured photos can be instantly shared with the smart device and post-processed.

2.2. Reshooting Parameters

2.2.1. Reshooting Distance

In photo acquisition, it is necessary not only to ensure that the details of the ground object can be reshot clearly but also to ensure that the ground object information can be obtained entirely. We call the distance from the camera to the object the "reshooting distance". According to the Chinese national standard "GB/T 7930-2008", the scale value of aerial photogrammetry is commonly 4–6 times the scale value of the corresponding topographic maps [42]. For example, suppose we want to make a topographic map with a scale of 1:500. In that case, the scale of aerial photogrammetry should be 1:2000–1:3000. The aerial photography scale is the ratio of the horizontal distance between the aerial photo line segment and the corresponding ground line segment. According to the principles of aerial photography and geometric optics, the formula is shown in Equation (1):

$$\frac{1}{m} = \frac{l}{L} = \frac{f}{H} \tag{1}$$

where m is the denominator of the scale value in aerial photogrammetry. l is the length of the line segment of the photo. L is the horizontal distance of the corresponding line segment on the ground. f is the focal length of the camera. H is the reshooting distance. Equation (1) provides a theoretical value for the reshooting distance, but the specific determination of the reshooting distance also needs to consider the terrain and topography of the study area.

2.2.2. The Overlap Ratio of the Reshooting Photos

To ensure accurate matching and stitching during reshooting and generate high-precision real-scene 3D models, it is necessary to ensure a certain ratio of overlap between the photos. Since the reshooting instrument faces the survey object, the stripe method can be used to reshoot the photo. The camera reshooting route is designed referencing the flight path of the UAV. Taking the facade of a building as a case study, the reshooting route is shown in Figure 2. Referring to UAV photogrammetry, we consider the up and down direction of the reshooting instrument as the reshooting front and the lateral direction of the reshooting instrument as the reshooting side in the process of reshooting. Reshooting points were established in both directions to capture overlapping images. Figure 3 shows the overlap ratio of the reshooting front and the reshooting side.

$$R_x = \frac{P_x}{L_x} \times 100\% \tag{2}$$

$$R_y = \frac{P_y}{L_y} \times 100\% \tag{3}$$

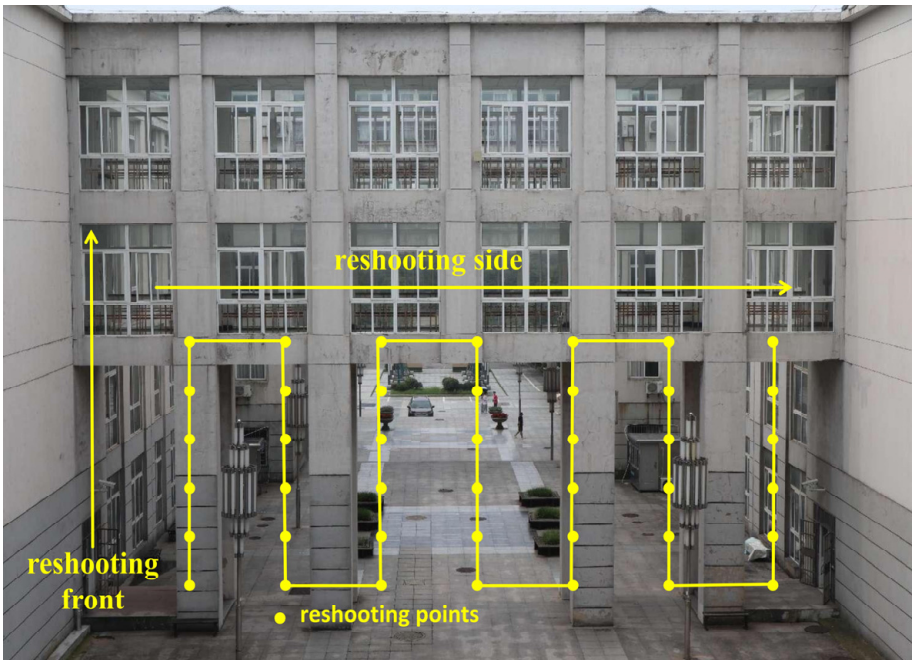


Figure 2. Reshooting route design for the ground camera.

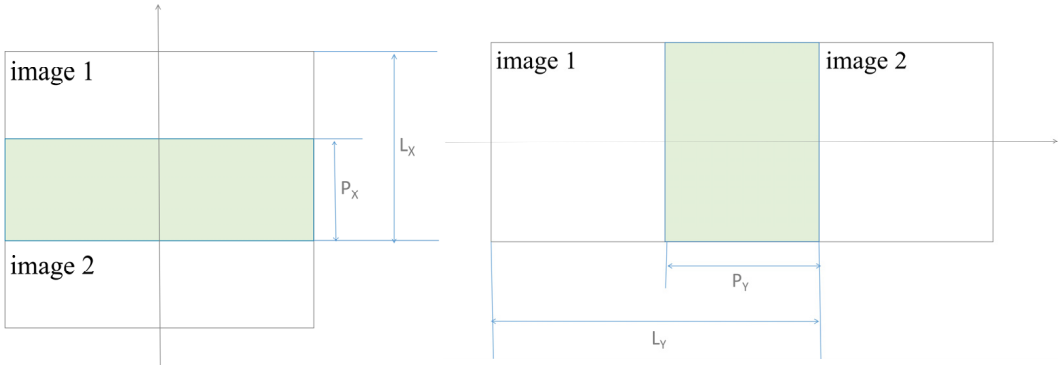


Figure 3. Illustration of the overlap ratio for reshooting photos: (**a**) the overlap ratio in the reshooting front direction and (**b**) the overlap ratio in the reshooting side direction.

Referring to the UAV photogrammetry method, the calculation formula is as shown in Equations (2) and (3). P_x is the overlapping length of the reshooting front. L_x is the height of the photo. R_x is the overlap ratio of the reshooting front. P_y is the overlapping length of the reshooting side. L_y is the width of the photo. R_y is the overlap ratio of the reshooting side. During photo acquisition, the overlap ratio of the reshooting front can be altered by controlling the longitudinal lifting distance of the reshooting instrument. Meanwhile, the overlap ratio of the reshooting side is changed by controlling the lateral translation distance of the reshooting instrument.

3. Evaluation Methods of a Real-Scene 3D Model

3.1. Quantitative Assessment Methods

3.1.1. Point Position Accuracy

Point position accuracy refers to the accuracy of each feature point in the model in geographic space, which can be evaluated from three directions as follows: E (east), N (north), and H (height). The selection of feature points needs to consider their representativeness and stability in the model. Feature points should have clear boundaries and distinct features to ensure that they can be accurately identified and measured across different photos.

The root mean square error (RMSE) was selected as the specific index for evaluation, comparing the degree of deviation between actual measured values and modeled values. The actual measured values were obtained using total stations, while the modeled values were extracted from the real-scene 3D model. The specific calculation formula is Equation (4). A smaller RMSE value indicates that the location and height of the points in the real-scene 3D model deviate less from the actual value, and the data accuracy is higher.

$$RMSE_T = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(T_i - \hat{T}_i\right)^2} \quad (4)$$

where T is the different orientations of the feature points in the geographic space. $RMSE_T$ is the root mean square error of the feature points in the T direction. n is the number of feature points. T_i is the actual measured value of the ith feature point in the T direction. $\hat{T}_i$ is the calculated value of the model of the ith feature point in the T direction.

3.1.2. Length Accuracy

Length accuracy refers to the accuracy of the linear dimensions of surveyed objects, such as the width and height of windows, evaluated by the deviation between actual ground survey values and 3D model measurements. When selecting experimental objects for evaluation, it is necessary to choose the critical objects in the 3D model and ensure that the object is easy to measure, such as windows, columns. By evaluating length accuracy, it is important to detect possible deviations, misalignments, or deformations of length

features in the model. The RMSE was also selected as the specific index of evaluation to calculate the root mean square error of the model length.

3.2. Authenticity Evaluation Method of the Real-Scene 3D Model

3.2.1. Texture Structure

In addition to high-precision location information, a high-quality real-scene 3D model should also have richer and clearer outline and texture information. The evaluation of the texture structure of the real-scene 3D model involves many important aspects. First, the clarity of the texture outline directly affects the presentation of surface details in the model. Clear textures accurately depict surface details and patterns, allowing for a more precise representation of the model's structure and features. Second, a high-quality 3D model should be naturally smooth at the joints of different parts to avoid significant fractures or adhesions, especially at the edges and corners of the model. In addition, during the texture mapping process, parts that are not properly mapped or covered will sometimes appear, resulting in hollow or unfilled areas on the surface of the model. Model distortion will directly affect the visual aesthetics of the model, affecting the realism and integrity of the model.

3.2.2. The Edge Length and Angle of the Triangular Mesh

In a real-scene 3D model, triangular mesh is a commonly used method of representing the surface of a 3D object. A triangular mesh is made up of a number of small triangles that describe the shape and surface of the entire model through their vertices, edges, and faces. The side lengths of the triangles in the triangular mesh affect the detail capture and visual realism of the model's surface. The smaller triangular side length captures the subtle changes and curvature of the model's surface more finely, helping to improve the visual quality of the model. The concentration of the side length of the triangle indicates that the triangle is uniform in size, no edges too long or too short appear, and the area and length of the feature are relatively balanced. Suppose the edge length of the triangle changes too much, especially in areas with significant curvature changes. This will affect the continuity of the model's surface and the natural visual flow. The angle of the triangle is also one of the critical factors. When the angular differences of adjacent triangles are small, it helps to maintain the continuity and smoothness of the surface. The more concentrated the triangle angle, the more balanced and regular the geometric shape, which also has an impact on the regularity of the shape of structures such as buildings. Angles that are too large or too small can cause the surface to stretch, twist, or deform.

4. Experiments and Analysis

4.1. Case Study

A building was selected as a case study, which covers an area of 7165 square meters, has a total of six floors, and a height of about 31 m. The building is complex for UAV photogrammetric survey and modeling in the following aspects: the building's structure and the terrain surrounding the building. The case study building belongs to the special-shaped category, and its shape is unique, as shown in Figure 4a. The main body of the building is divided into two relatively independent areas on the east and west sides, connected by two air corridors. The air corridor on the south side connects the fifth and sixth floors of the east–west area (see Figure 4c). The north air corridor connects the fourth and fifth floors of the east and west areas (see Figure 4e). The special shape of the building makes it prone to self-occlusion in the UAV survey, which can lead to poor quality in the corresponding 3D model.

In addition, the building is constructed on low-lying land on a hillside, which complicates the UAV survey. There are large slopes on the east and west sides, as shown in Figure 4b,d. The terrain elevation difference between the south (low) and the north (high) area of the building is about 10.8 m, and a long staircase connects them, as shown in Figure 4f.

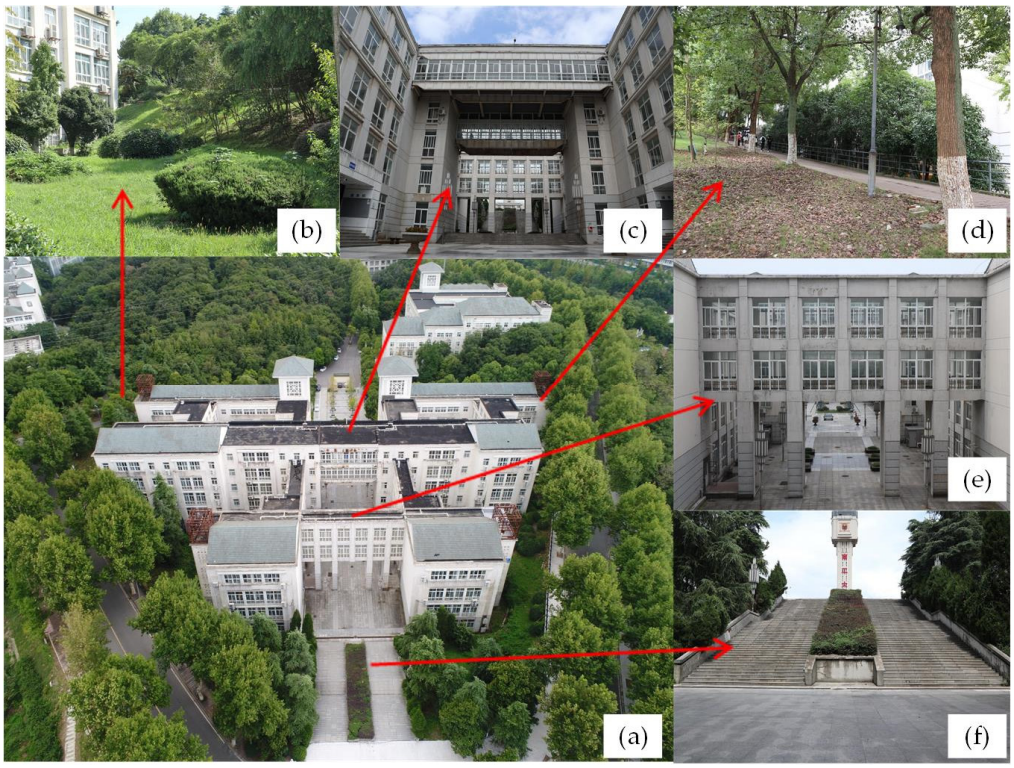


Figure 4. Overview of the case study building: (**a**) the appearance of the building; (**b**) the western slope near the building; (**c**) the south side aerial corridor in the building; (**d**) the eastern slope near the building; (**e**) the north side aerial corridor in the building; and (**f**) the long staircase on the north side of the building.

There are many trees around the building. When using a UAV to obtain aerial photos, the crowns and trunks of trees will block the bottom and side areas of the building. Dense trees and poor terrain limit the field of view of the UAV, which can easily lead to a lack of building information and blurred images. Furthermore, the texture of leaves and trunks is different from the surface texture of buildings. The modeling algorithm has difficulty distinguishing these textures, which easily leads to texture loss and affects the quality of the model. Additionally, the building walls are gray-white, and the window transparency features account for a large proportion, which is prone to mismatch during modeling.

4.2. Photo Acquisition and Modeling

4.2.1. UAV Photo Acquisition and Model

We use the DJI M300 UAV equipped with a five-lens oblique photogrammetric camera, the PSDK 102S V2, to obtain the oblique images of the ground objects, as shown in Figure 5. The camera, with 24.3 million effective pixels, can collect information from multiple angles including one vertical direction and four tilt directions. The angle between each tilt lens and the vertical lens is 45°. The focal length of the down-looking lens is 25 mm, and that of the side-looking lens is 35 mm. The image resolution of each lens is 6000 × 4000 pixels. The UAV, equipped with RTK-GNSS (Real-Time Kinematic, Global Navigation Satellite System) chips, can dynamically obtain the position of the aerial photos. The CORS (Continuous Operational Reference System) service for the RTK-GNSS is provided by a Chinese company, Qianxun SI (https://en.qxwz.com/). Their CORS service covers the whole of China and has an accuracy of 2 cm in plane and 5 cm in height.

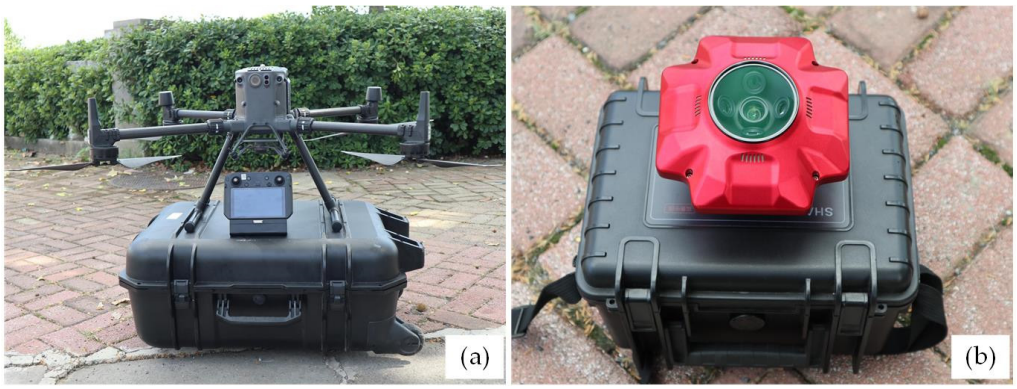


Figure 5. UAV photo acquisition equipment: (**a**) the DJI M300 UAV and (**b**) the oblique photographic camera.

We took the building as the center and expanded 30 m around as the survey area. Considering the accuracy of the real-scene 3D model and the size of the survey area, we designed the aerial flight altitude to be 96 m (height from the ground of the take-off point) and the ground resolution of the aerial photos to be 1.5 cm/pixel. In order to ensure the accuracy of point cloud preprocessing and data coverage, the side overlap rate was set to 70%, and the front overlap ratio was set to 70%. According to the route designed in Figure 6 for data collection, we obtained 1440 aerial photos (four groups of oblique photos and one group of vertical photos) and corresponding position (POS) data.



Figure 6. Aerial flight route. The building is the Renzhi Building, located at the Jiangpu Campus of Nanjing Tech University in Nanjing, Jiangsu Province, China.

The real-scene 3D model made by UAV aerial photos is shown in Figure 7. Because of the influence of trees on both sides, the characteristics of the building itself, and the topography of the long staircase, the quality of the back of the building is poor. The windows are severely deformed, and the connections between the windows and eaves are adhesive, as shown in Figure 7a. Additionally, the absence of photos at the bottom of the corridor leads to a significant void in the model, as illustrated in Figure 7b. The deformation of the column is significant, with voids appearing at the bottom, which affects the integrity of the column. In order to improve the quality of the model, we selected the back of the building as the study area and used the reshooting instrument to take more photos in the poor-quality area.



(**a**) (**b**)

Figure 7. Real-scene 3D model of the case study building generated by only UAV photos: (**a**) deformed windows in the model and (**b**) the corridor with a large void.

4.2.2. Acquisition of Reshooting Photos by the Ground Camera

The focal length of the ground camera (EOS M6) ranges from 15 mm to 45 mm. In this study, we found that if the focal length was too short, too many objects were captured in a photo. Conversely, if the focal length was too long, too few objects were captured. Neither of these situations was suitable for this study. Therefore, we selected a focal length between 28 mm and 40 mm. We first set the camera's focal length to 28 mm. According to Formula (1), the reshooting distance was calculated to be 56 m. The overlap ratio of the reshooting by the ground camera should be lower than aerial photos since they are just complements. However, we still used 70% as the overlap ratio of the reshooting front and reshooting side for the convenience of experiment analysis. So, we could resample the photos to simulate different overlap ratios for the ground camera. According to Formulas (2) and (3), the interval distance of the reshooting photos in the reshooting front direction and side direction were both set as 20 cm.

According to the reshooting route shown in Figure 2, reshooting points were set at the reshooting front and side, with an interval distance of 20 cm between the reshooting points. A total of 31 reshooting points was set at the reshooting front, and 10 reshooting points were set at the reshooting side. We used the reshooting instrument to reshoot photos at the reshooting points to construct the real-scene 3D model.

4.3. Real-Scene 3D Model Built by Fusing Photos of the Ground Camera and UAV

To improve the quality of the real-scene 3D model, we fused photos from the UAV and ground camera to build the 3D model. ContextCapture v4.4.12 software from the Bentley company was used to perform the process, including the aerial triangulation, corresponding point matching, and other operations, as shown in Figure 8.

The choice of tie points is critical and should be prominent and more accessible to identify, such as apparent signs, building corners, and building connections. The tie points should be evenly distributed to cover the entire modeling area. Since most of the study area has windows, gray-white walls, and columns, it was not possible to select the tie points on the transparent glass and the walls directly without any apparent signs. We chose the intersection of the window edges, the starting point, or end point of the concave and convex lines on the column that was easy to identify. A total of five tie points were selected,

as shown in Figure 9. We ensured that each tie point was contained in three–five photos so the study area could be observed and reconstructed from different angles.

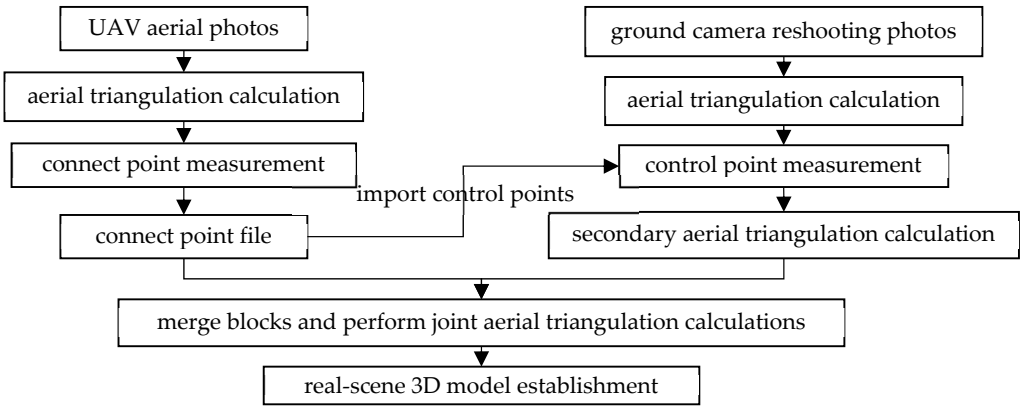


Figure 8. The process of constructing a fused model.



Figure 9. Distribution of the tie points.

4.4. Comparative Analysis of the Quality of the UAV Model and Fused Model

4.4.1. Accuracy of Point Coordination

We selected the intersection of window edges, the starting point or end point of the bump line on the column, and the corner point of the window as feature points to ensure the operability of the coordination measurement. A total of 26 feature points was selected and numbered from left to right and from top to bottom, as shown in Figure 10.

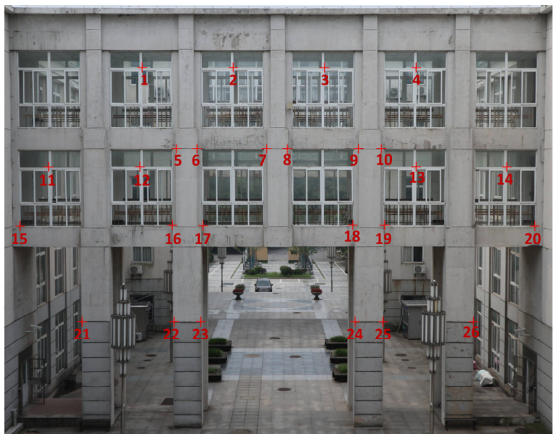


Figure 10. Distribution of feature points.

The actual coordinates of these feature points were obtained from ground surveys using an RTK-GNSS (Huace I90) and a total station (RTS 102R10). The CORS service provided by Qianxun SI determined the base station coordinates for the total station. Then, considering these coordinates as input parameters, a total station was used to measure the coordinates of the feature points. In these measurements, we used the CGCS2000 (China Geodetic Coordinate System 2000) coordinate system, a Gaussian–Kruger projection with a three-degree bandwidth, and reference the central meridian at 117° E. The UAV and the actual coordinate surveys have identical coordinate systems and projections. ContextCapture Viewer software allowed us to view the 3D model. We manually measured the coordinates of the feature points using the mouse-picking operation in the software and considered these measurements as the modeled values. The results are shown in Table 1.

Table 1. Coordination value of feature points (in meter units).

No.	The Actual Value of the Points			The Value of the Points in the UAV Model			The Value of the Points in the Fused Model		
n	N	E	H	N	E	H	N	E	H
1	3,552,879.164	654,179.799	60.177	3,552,879.17	654,179.86	60.52	3,552,879.19	654,179.87	60.58
2	3,552,879.134	654,175.814	60.179	3,552,879.14	654,175.88	60.53	3,552,879.16	654,175.90	60.58
3	3,552,879.046	654,171.845	60.168	3,552,879.09	654,171.92	60.55	3,552,879.05	654,171.93	60.57
4	3,552,879.013	654,167.815	60.179	3,552,879.04	654,167.89	60.55	3,552,879.01	654,167.88	60.57
5	3,552,879.504	654,178.308	56.642	3,552,879.47	654,178.38	57.01	3,552,879.50	654,178.40	57.06
6	3,552,879.486	654,177.432	56.638	3,552,879.45	654,177.50	57.01	3,552,879.49	654,177.49	57.05
7	3,552,879.448	654,174.308	56.640	3,552,879.40	654,174.37	56.99	3,552,879.45	654,174.39	57.05
8	3,552,879.441	654,173.455	56.624	3,552,879.42	654,173.53	57.01	3,552,879.44	654,173.53	57.03
9	3,552,879.397	654,170.297	56.620	3,552,879.35	654,170.35	56.97	3,552,879.38	654,170.34	57.02
10	3,552,879.387	654,169.429	56.622	3,552,879.36	654,169.49	57.01	3,552,879.37	654,169.51	57.03
11	3,552,879.214	654,183.932	55.862	3,552,879.24	654,184.00	56.28	3,552,879.29	654,183.99	56.27
12	3,552,879.177	654,179.905	55.889	3,552,879.21	654,179.99	56.26	3,552,879.22	654,179.99	56.30
13	3,552,879.009	654,167.919	55.893	3,552,879.03	654,168.02	56.29	3,552,879.00	654,167.91	56.29
14	3,552,878.954	654,163.924	55.880	3,552,878.97	654,164.00	56.24	3,552,878.92	654,163.91	56.28
15	3,552,879.299	654,185.200	53.300	3,552,879.22	654,185.27	53.72	3,552,879.35	654,185.25	53.73
16	3,552,879.215	654,178.531	53.329	3,552,879.15	654,178.61	53.68	3,552,879.24	654,178.61	53.75
17	3,552,879.212	654,177.190	53.327	3,552,879.16	654,177.25	53.71	3,552,879.26	654,177.27	53.75
18	3,552,879.110	654,170.590	53.339	3,552,879.08	654,170.67	53.73	3,552,879.13	654,170.64	53.75
19	3,552,879.088	654,169.201	53.328	3,552,879.03	654,169.26	53.69	3,552,879.09	654,169.24	53.74
20	3,552,878.995	654,162.539	53.318	3,552,878.95	654,162.62	53.72	3,552,878.97	654,162.62	53.72
21	3,552,879.553	654,182.481	49.111	3,552,879.43	654,182.54	49.39	3,552,879.63	654,182.56	49.57
22	3,552,879.501	654,178.455	49.093	3,552,879.38	654,178.51	49.39	3,552,879.56	654,178.48	49.54
23	3,552,879.482	654,177.276	49.090	3,552,879.44	654,177.34	49.46	3,552,879.53	654,177.33	49.53
24	3,552,879.397	654,170.478	49.097	3,552,879.33	654,170.55	49.45	3,552,879.44	654,170.51	49.52
25	3,552,879.376	654,169.276	49.097	3,552,879.33	654,169.32	49.46	3,552,879.41	654,169.36	49.52
26	3,552,879.318	654,165.299	49.079	3,552,879.29	654,165.36	49.47	3,552,879.32	654,165.35	49.49

According to Formula (4), the RMSE results in different directions between the UAV model and the fused model were calculated, as shown in Table 2. It can be seen from the table that, after photo fusion, the plane accuracy of the 3D model is well improved. We also notice that the elevation accuracy decreased in this experiment. There are probably several reasons for this decline, which include the following: (1) the height accuracy of the Qianxun CORS is only 5 cm; (2) the mouse-picking measurement method may introduce some errors; and (3) there are likely errors in camera self-calibration, which result in a vertical offset in the fused model.

Table 2. RMSE results of point position (in centimeter units).

Direction	UAV Model	Fused Model
N	5.2	3.5
E	7.0	6.6
H	36.9	41.5

4.4.2. Length Accuracy

Although the length can be calculated from the coordinates, we decided to measure the length of the experimental objects directly with a steel tape to avoid unnecessary correlation with the point accuracy evaluation. When conducting field measurements of experimental objects, it is necessary to ensure safety and operability. Meanwhile, in the established 3D

model, the window size and column width are easy to measure, but geometric objects such as walls are difficult to measure. So, we selected the width and height of a window and the width of a column as experimental objects that we considered to have fixed values throughout the wall. The column and window marked with a red star in Figure 11 were selected as places to determine the actual values. We measured the width and height values at those places with a steel tape three times and calculated an average value as their actual values, as shown in Table 3. And then, we measured the width and height in other places (a total of 24 places) in the UAV model and fused model using mouse-click operations. These places were evenly distributed in the whole wall and are marked with yellow and blue lines in Figure 11 (the meaning of the yellow and blue lines is the same in this experiment, but they are different in the experiments in Sections 4.5.2 and 4.5.3). The corresponding values are recorded in Table 3. The RMSE results of length, calculated by referring to Formula (4), are shown in Table 4. It can be seen that the lengths of the experimental objects in the fused model are closer to the actual values.



Figure 11. The spatial distribution of the experiment objects. The yellow and blue lines represent the experiment objects for the UAV and fused models. The lines marked with red stars indicate the locations for determining actual values. Additionally, the blue lines are further used for the experiments described in Sections 4.5.2 and 4.5.3.

Table 3. The actual values and measurement values of experimental objects (in meter units).

Object	The Actual Value	The Measurement Values in the UAV Model				The Measurement Values in the Fused Model			
Window width	0.60	0.57	0.60	0.64	0.62	0.60	0.60	0.60	0.60
		0.60	0.57	0.59	0.58	0.61	0.60	0.60	0.60
		0.61	0.61	0.56	0.60	0.60	0.59	0.59	0.60
		0.58	0.61	0.56	0.60	0.60	0.61	0.60	0.60
		0.59	0.61	0.58	0.60	0.60	0.60	0.59	0.60
		0.58	0.59	0.57	0.61	0.60	0.60	0.59	0.60
Window height	0.89	0.87	0.85	0.86	0.87	0.88	0.87	0.87	0.91
		0.84	0.87	0.94	0.85	0.92	0.88	0.91	0.91
		0.90	0.91	0.92	0.86	0.88	0.90	0.91	0.90
		0.91	0.85	0.87	0.87	0.90	0.90	0.90	0.91
		0.87	0.90	0.85	0.85	0.86	0.81	0.86	0.95
		0.88	0.91	0.85	0.84	0.90	0.86	0.91	0.90
Column width	1.18	1.23	1.20	1.24	1.21	1.21	1.21	1.21	1.20
		1.20	1.14	1.23	1.20	1.22	1.16	1.19	1.20
		1.16	1.19	1.20	1.14	1.22	1.19	1.16	1.15
		1.12	1.14	1.22	1.17	1.24	1.16	1.19	1.13
		1.10	1.15	1.20	1.16	1.22	1.14	1.20	1.19
		1.21	1.12	1.22	1.14	1.21	1.18	1.17	1.17

Table 4. RMSE results of length (in centimeter units). The UAV model represents the real-scene 3D model generated only using UAV aerial photos; the fused model represents the real-scene 3D model generated using UAV aerial photos and the photos from the ground camera.

Object	UAV Model	Fused Model
Window width	1.9	0.6
Window height	3.1	2.7
Column width	4.0	2.9

4.4.3. Texture Structure

In the UAV-only model, the columns are severely deformed, the windows appear hollow and lack clear texture, and the bottom of the corridor also appears hollow. By comparing the details, edges, and surface quality of the corresponding regions in the fused model, as shown in Figure 12, it is found that the column does not show severe deformation. There are no "holes" in the windows, and the geometry is more complete. There is no adhesion from the window to the eaves. The contour at the bottom of the corridor is relatively complete, and the part connected with the column is structured. Overall, the fused model has a sharper outline and texture. The quality of the poor parts of the initial model, such as windows, eaves, and the bottom of the corridor, is improved. A high ratio of restoration of the building itself is realized.



(**a**) (**b**)

Figure 12. Real-scene 3D models: (**a**) UAV model and (**b**) fused model.

4.4.4. Histogram of the Edge Length and Angle of the Triangular Mesh

After the real-scene 3D model was established, the triangular mesh of the building was created, as shown in Figure 13. As shown in the triangular mesh diagram, the number of triangular meshes in the fused model is more significant. The shape of the bottom of the corridor and the columns is more regular.

The histogram of the edge length and angle distribution is shown in Figure 14. The edge length of the triangular mesh of the UAV model is distributed in the range of 20–90 cm, with a high concentration in the range of 30–60 cm and a relatively concentrated ratio in the range of 60–80 cm. The edge length of the triangular mesh of the fused model is distributed in the range of 20–60 cm, and the concentration ratio is high in the range of 30–50 cm. It can be seen that the edge length of the fused model is more concentrated, indicating that the area and length of the building are more balanced. Fewer triangles are too slender or flat, which better reflects the balance and stability of the building structure. From the perspective of triangular mesh, the angle data of the UAV model and the fused model are concentrated in the range of 20–80°, and the highest concentration is in the range of 30–60°. These data show that the surface changes in both models are relatively gentle, with fewer abnormal spike objects, which is relatively close to the geometry of actual buildings. On the whole, the fused model has a richer level of detail, smoother and more consistent geometry, more vital expression ability, and better quality.

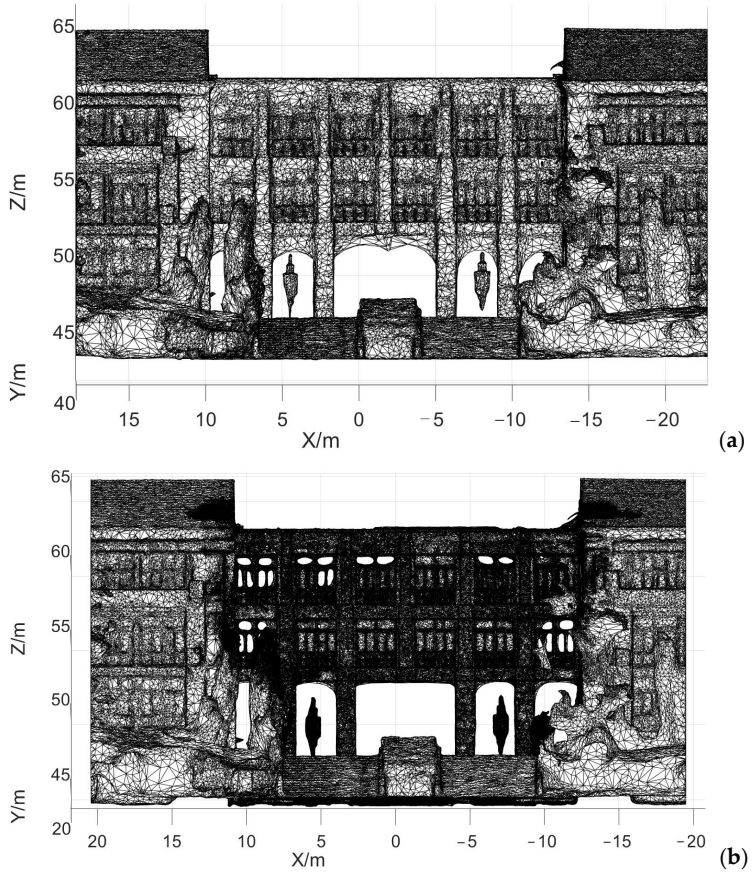


Figure 13. Triangular mesh diagram of the 3D model: (**a**) UAV model and (**b**) fused model.

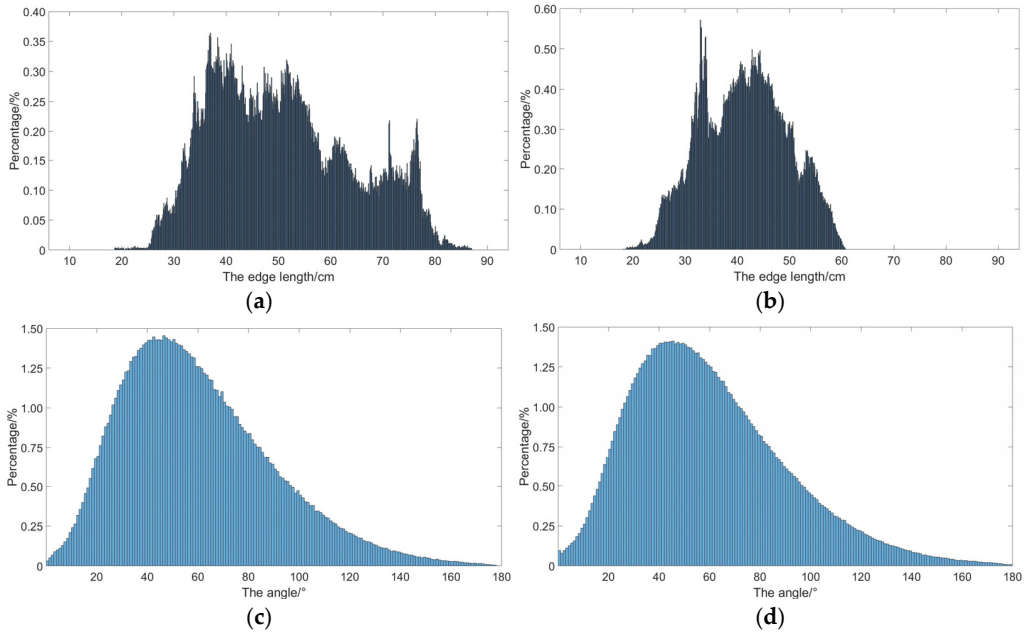


Figure 14. Histograms of edge length and angle distribution for UAV models (**a**,**c**) and fused models (**b**,**d**).

4.5. Optimization of the Ground Camera's Reshooting Parameters

4.5.1. Reshooting Distance

When taking reshooting photos, the closer the reshooting distance, the higher the photo clarity and the more comprehensive the texture details of the building facade. This

also means more on-site work needs to be performed. By changing the reshooting distance, we explored how to set the parameter to balance model quality and workload when the camera focal length is constant.

The corridor on the backside of the case study building served as an experimental area, as shown in Figure 4e. In the reshooting distance experiments, we considered the following parameters: the shortest reshooting distance, the longest reshooting distance, the interval distance for each reshooting experiment between the shortest and longest distances, and the overlap ratio of each experiment. Our on-site inspections revealed that the coverage area of each photo was limited when the reshooting distance was too close. To cover the target object fully, many photos must be taken, which would consume more manpower and resources. Meanwhile, a long staircase leading to higher ground is located 15 m away from the corridor, as shown in Figure 4f, which limited our ability to erect a reshooting instrument. Considering the influence of the workload and the long staircase, we decided that the shortest reshooting distance should be 40 m.

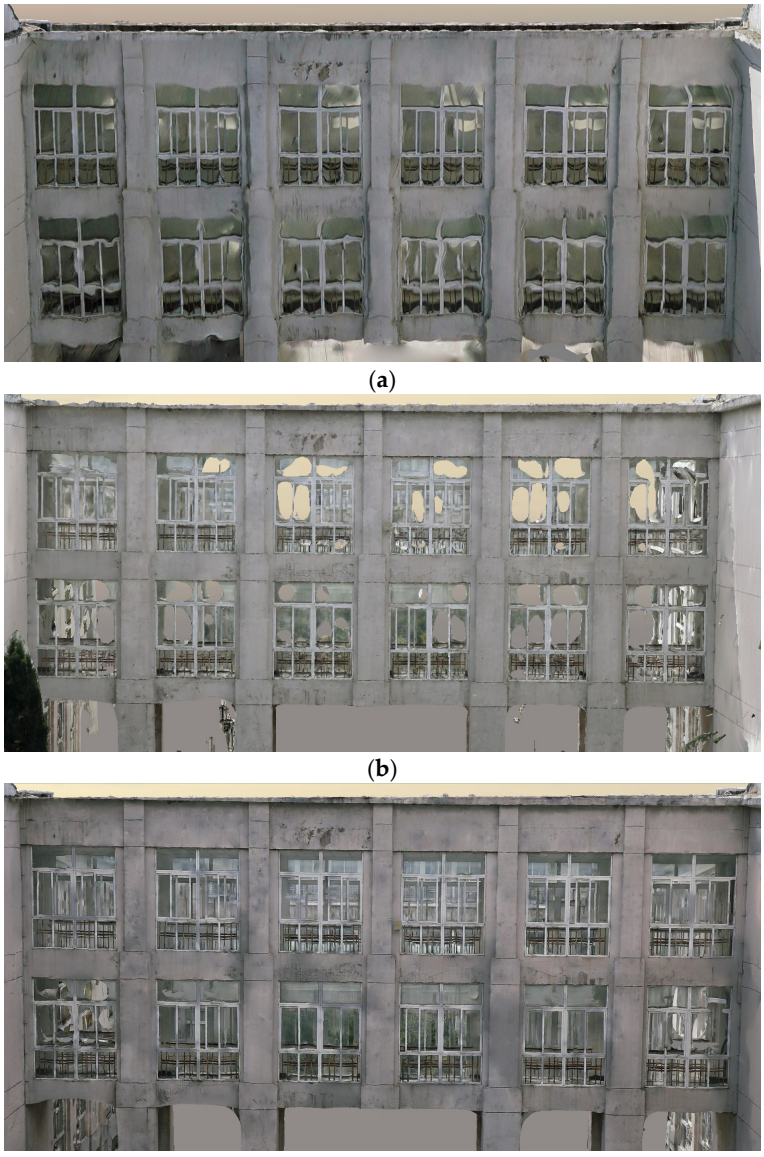
Similar to the shortest reshooting distance, the longest reshooting distance is also influenced by the following factors: the workload involved in reshooting, the focal length, and the terrain conditions during reshooting. We investigated the longest reshooting distances ranging from 50 m to 80 m and found that when the reshooting distance exceeded 70 m, the proportion of the target object in the reshooting photo became very small, regardless of the focal length. After considering these factors comprehensively, the longest distance was determined for different focal lengths, as shown in Table 5. Considering the on-site workload, we chose 5 m as the interval distance and reduced the overlap ratio to 50%. Despite this adjustment, the on-site workload remained heavy for each experiment, requiring almost 1–2 h even with such an interval distance and overlap.

Table 5. Parameters of different experiments.

No.	Focal Length/mm	Shortest Distance/m	Longest Distance/m	Interval Distance/m	Overlap Ratio
1	28	40	50	5	50%
2	32	45	55	5	50%
3	36	50	60	5	50%
4	40	60	70	5	50%

With those four groups of parameters, 12 experiments were conducted. After completing the on-site reshooting work, ContextCapture software was used to build the 3D models for the interior work. Consequently, 13 models were created, including one UAV model and 12 fused models. For example, the fused models with a 28 mm focal length, a 40 m reshooting distance, and a 50% overlap ratio, as well as those with a 28 mm focal length, a 45 m reshooting distance, and a 50% overlap ratio, are shown in Figure 15. The download link for the original figures of the 13 models is listed in the Supplementary Materials Section.

We qualitatively inspected these 3D models, including geometry shape, texture structure, and model details. Then, we determined the approximate range of the suitable reshooting distance for different focal lengths, as shown in Table 6. After that, a simple multiple relationship between the focal length and reshooting distance was deduced, as shown in Table 6.



(**a**)

(**b**)

(**c**)

Figure 15. Comparison of the UAV model and the fused models: (**a**) the UAV model; (**b**) the fused model with a 28 mm focal length, a 40 m reshooting distance, and a 50% overlap ratio; and (**c**) the fused model with a 28 mm focal length, a 45 m reshooting distance, and a 50% overlap ratio. The process of building the 3D model using ContextCapture software was time-consuming. Each model required approximately three hours. To expedite the experiment, we removed some UAV photos and retained only those directly related to the experimental area. Consequently, the UAV model in this figure is slightly different from the model in Figure 12a.

Table 6. Multiple relationships between the focal length and reshooting distance.

Focal Length/mm	Suitable Reshooting Distance/m	Multiple Relationship Between Focal Length and Reshooting Distance
28	40–45	1.4–1.6
32	45–50	1.4–1.6
36	55–60	1.5–1.7
40	60–65	1.5–1.7

To ensure that this quantitative relationship is applicable to each experiment, we intersected the four ranges in Table 6 and determined the final relationship as follows: the reshooting distance (in meter units) was approximately 1.5–1.6 times the camera's focal length (in millimeter units). Taking this relationship as the main guidance, we adjusted the suitable reshooting distance, as shown in Table 7. In this adjustment, we also considered the experiences from on-site work. The relationships for focal lengths of 30 mm, 34 mm, and 38 mm were deduced through interpolation and on-site work experiences. To express these empirical relationships in the formula, we fit Table 7 with a linear regression method, and the fitting result is shown in Figure 16.

Table 7. Optimal range for reshooting distance.

Focal Length/mm	Reshooting Distance Range/m
28	42–46
30	44–48
32	48–52
34	50–56
36	54–58
38	56–62
40	60–64

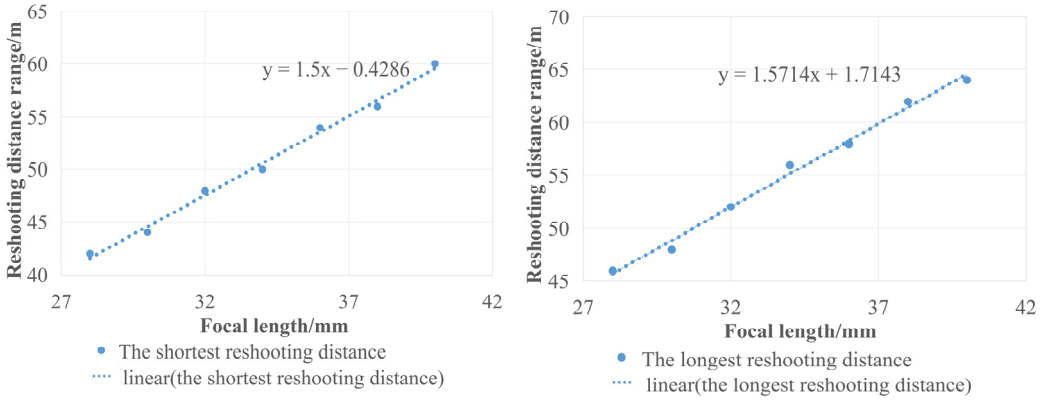


Figure 16. The relationship between the focal length (in millimeter units) of the ground camera and the optimal reshooting distance range (in meter units): (**a**) the relationship between the shortest reshooting distance and the focal length and (**b**) the relationship between the longest reshooting distance and the focal length.

4.5.2. The Overlap Ratio of the Reshooting Front

It is easy to understand that a high overlap ratio provides more feature points in the AT (aerial triangulation) process and improves the 3D model's quality. However, such a high overlap ratio also results in a tremendous workload not only on-site but also indoors. For example, in this study's fused modeling process, we use at least four tie points, and each point must have appeared on at least four photos. In this experiment, we decided to fix the focal length at 28 mm and adjust the overlap ratio and reshooting distance to find the optimal ratio value.

The physical size of the photos taken by EOS M6 is 56.4 cm (height) × 75.2 cm (width). So, when the overlap ratio of the reshooting front was 70%, the reshooting interval distance was approximately 20 cm according to Formula (2). The corresponding workload was very heavy in such situations. To analyze the optimal interval distance for the reshooting front, we set up three groups of comparative experiments, with different reshooting distances of 40 m, 45 m, and 50 m. Then, in each group, we set the overlap interval distances of the reshooting as 20 cm, 40 cm, and 60 cm. Their corresponding overlap ratios were approximately 70%, 30%, and 0% (since the height of the photo was only 56.4 cm).

To decrease the workload, we only selected 12 experimental objects to test the accuracy, as shown in Figure 11 (denoted as the blue lines). In contrast to RMSE, we just used the length deviation of different experimental objects between actual and model values to evaluate the optimal overlap ratio. The ratio corresponding to overall minimal deviations is the optimal one. We calculated the deviation in each experimental object at the different 3D models, as shown in Figure 17. There are 12 points on the horizontal axis, which represent 12 experimental objects. The values on the vertical axis represent different deviation values. The lines in different colors show the deviations in the 12 experimental objects of different 3D models. For example, the blue color represents the UAV model, and the orange color represents the fused model with a 20 cm overlap ratio. In the experiment, we found that the values of the experimental objects were small. So, the differences between actual and model values were also small, making it difficult to observe their deviations. We used the square of the difference between the actual and model values, as shown on the vertical axis in Figure 11. The squared values enlarged the differences, enabling us to identify the optimal ratio more easily.

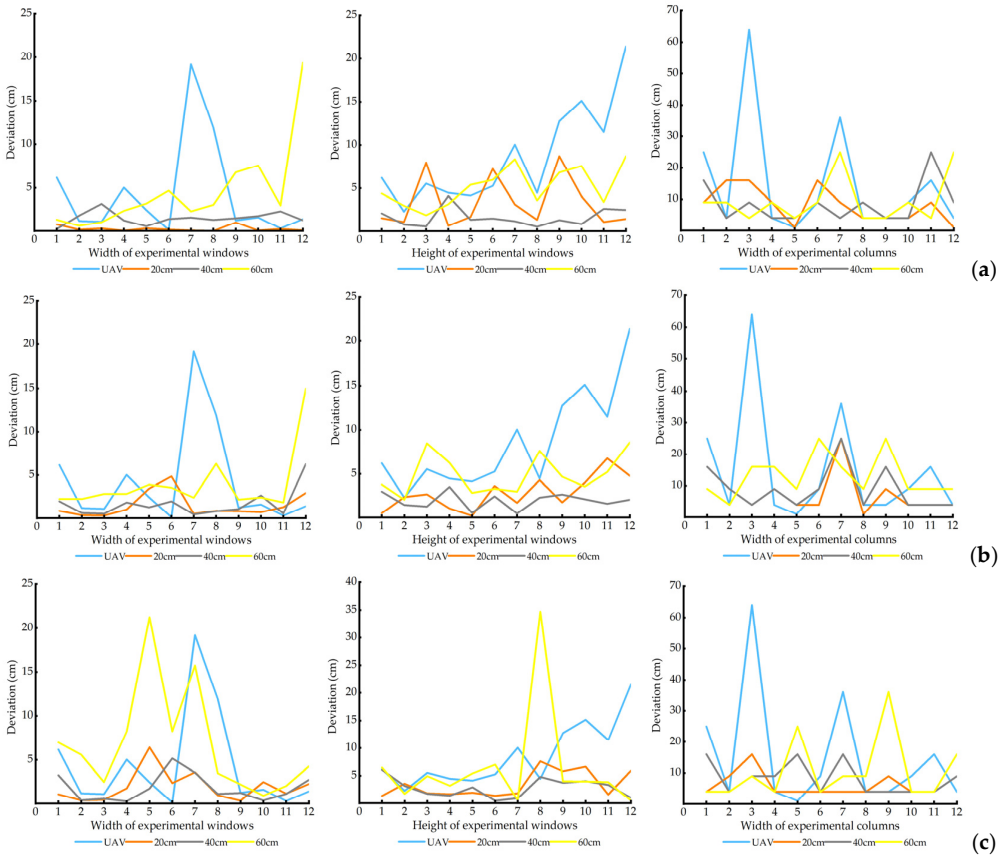


Figure 17. Deviation between the actual and modeled values of 12 experimental objects under different overlap interval distances and reshooting distances: (**a**) deviation in the experimental objects under a reshooting distance of 40 m; (**b**) deviation in the experimental objects under a reshooting distance of 45 m; and (**c**) deviation in the experimental objects under a reshooting distance of 50 m. Each point on the horizontal axis represents each experimental object; the value on the vertical axis represents the enlarged deviation value. Light blue represents the UAV model; orange represents the fused model with a 20 cm overlap interval distance (about 70% overlap ratio); gray represents the fused model with a 40 cm overlap interval distance (about 30% overlap ratio); and yellow represents the fused model with a 60 cm overlap interval distance (0% overlap ratio).

As can be seen in Figure 17, when the interval distance of the reshooting front is at 20 cm and 40 cm, the overall deviations fluctuate less, and no obvious singular values

appear across different reshooting distances. We considered the accuracy of the fused 3D model to be high. At the 60 cm interval distance for the reshooting front, although the accuracy is improved compared to the UAV model, the deviations fluctuate greatly, and singular values are sometimes even more noticeable. In general, we considered that the overlap ratio of the reshooting front is optimized at about 30% (40 cm in this study), which ensures the accuracy of the fused model and also improves the efficiency of reshooting.

4.5.3. The Overlap Ratio of the Reshooting Side

Similar experiments were conducted to explore the overlap ratio of the reshooting side. Compared with the experiment in Section 4.5.2, almost all experimental conditions were the same, except for the value range of the overlap ratio. Because the physical width of the photos differed from the height, we set the interval distances for the reshooting overlap as 20 cm (about 70% overlap ratio), 40 cm (about 47% overlap ratio), 60 cm (about 20% overlap ratio), 80 cm (0% overlap ratio, since the physical width of the photo was 75.2 cm), and 100 cm (0% overlap ratio, the reshooting terrain condition allows for more exploration compared with the reshooting front direction). After building real-scene 3D models, we conducted a similar qualitative examination and deviation calculation on the experimental objects. The results are shown in Figure 18.

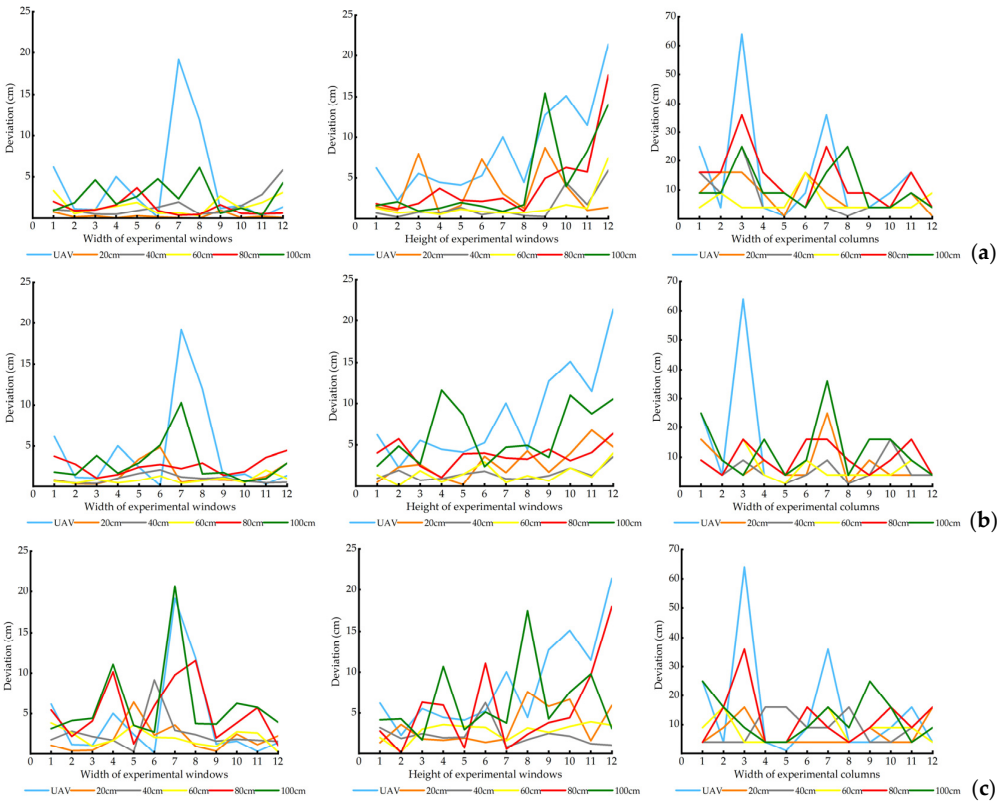


Figure 18. Deviation between the actual and modeled values of 12 experimental objects under different overlap interval distances and reshooting distances: (**a**) deviation in the experimental objects under a reshooting distance of 40 m; (**b**) deviation in the experimental objects under a reshooting distance of 45 m; and (**c**) deviation in the experimental objects under a reshooting distance of 50 m. Each point on the horizontal axis represents each experimental object; the value on the vertical axis represents the enlarged deviation value. Light blue represents the UAV model; orange represents the fused model with a 20 cm overlap interval distance (about 70% overlap ratio); gray represents the fused model with a 40 cm overlap interval distance (47% overlap ratio); yellow represents the fused model with a 60 cm overlap interval distance (about 20% overlap ratio); red represents the fused model with a 80 cm overlap interval distance (0% overlap ratio); and green represents the fused model with a 100 cm overlap interval distance (0% overlap ratio).

As can be seen in Figure 18, when the interval distance of the reshooting side is 20 cm, 40 cm, and 60 cm, the overall deviations fluctuate less, and no obvious singular values appear across different reshooting distances. At the interval distances of 80 cm and 100 cm, the accuracy is improved compared with the UAV model. However, the deviations fluctuate greatly, and the peak values are evident, which means the model quality is lower. So, we considered the optimal overlap ratio for the side direction in reshooting to be 20%, which ensures the accuracy of the fused real-scene 3D model and also reduces the reshooting workload.

4.6. Discussion

Integrating ground camera photos and UAV aerial photos for modeling may improve the quality of real-scene 3D models. We developed a ground camera reshooting instrument, which mainly consists of a Canon EOS M6 camera and a retractable selfie stick. Compared with other devices, such as laser scanners, this instrument has a low cost, is simple to operate, and can be easily deployed and adjusted in various complex terrains and environments. Because of the fact that the reshooting instrument is installed on the ground and occupies a small space, it is safer to acquire photos in densely populated areas than when using UAV.

Experiments in this study confirm that ground photos can greatly improve the quality of the real-scene 3D models. Figures 12 and 15 show that the improvement in the appearance of the 3D model is significant. For example, the deformations of windows and columns are greatly decreased, and the texture of the 3D model is clearer. The geometric shape becomes more accurate. Meanwhile, the point and length accuracies are also improved, as shown in Tables 2 and 4. However, compared with the improvement in appearance, the improvement in point and length accuracy does not seem as significant. The RMSE of height in Table 2 is decreased, which we believe is normal. The 26 test points in Figure 12a or Figure 15a should not be measurable since the UAV model is too blurry. However, we forcibly measured them using manual mouse-click operations in the ContextCapture viewer. We believe such manual operations greatly improve the accuracy of those points because of the identifying abilities of the human. An identical method was used to measure the coordinates of the test points in the fused model. Thus, the improvement (or even decline) in the accuracy of those test points mostly depends on the accuracy of the entire 3D model, which in turn depends on the accuracy of the Qianxun CORS service, which is 2 cm in the plane and 5 cm in height. The changes in RMSE in Table 2 basically follow these values. However, the absolute value of the RMSE of height is relatively high, even up to 4 dm. This may be caused by the following reasons: the height accuracy of GNSS is lower than that of the plane; the elevation difference of terrain in UAV survey; the impact of the atmosphere on UAV flight; and the errors in camera self-calibration, which may result in vertical offset. To decrease the possible systematic errors, we calculated the standard deviation estimation with Equation (5) and noticed that the standard deviation of the H direction improves from 3.1 cm to 1.6 cm. The results indicate that the improvement in the H direction of the fused model has consistency with other directions.

$$S_T = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\Delta T_i - \frac{1}{n}\sum_{i=1}^{n}\Delta T_i\right)^2} \quad (5)$$

where T is the different orientations of the feature points in the geographic space. S_T is the standard deviation of the feature points in the T direction, based on the average value of error. n is the number of feature points. ΔT_i is the deviation between the model value and the actual value of the ith feature point in the T direction.

In order to efficiently reshoot photos, we designed a reshooting route and optimized the core parameters, such as reshooting distance and the overlap ratios. The suitable parameters recommended in this study are actually a summary of experimental experience. We attempted to formulate the empirical results into a theoretical conclusion. For example, we used a linear regression model to fit the reshooting distance, as shown in Figure 16.

Although the regression result is preliminary, it inspires further exploration of the issue. We found that the optimal overlap ratios for reshooting in the front and side directions are different, specifically 20% for the front and 30% for the side. After further investigation, it was discovered that the difference probably stems from the physical size of the ground camera's photo. Hence, we propose a hypothesis that 20–30% is a usable overlap value for reshooting in both front and side directions since we did not investigate the photo sizes of most cameras in use.

We used a quality evaluation method for real-scene 3D models, which mainly includes the following aspects: quantitative evaluation and model authenticity evaluation. Quantitative evaluation methods included point position accuracy and length accuracy. The methods for evaluating model authenticity mainly included texture structure and the edge length and angle of the triangular mesh. This evaluation method provides a reference standard for the quality evaluation of real-scene 3D models.

Our study needs further exploration. At present, the adjustable height of the reshooting instrument is 3 to 5 m, which makes it difficult to capture the details of high-ground objects accurately. When the height of the reshooting instrument reaches a certain level, it will shake left and right because of the influence of the wind, affecting the quality of the reshooting photos and the overlap ratio between multiple photos. In future studies, it may be possible to consider replacing the selfie stick with other devices with a larger telescopic range and more stability. In addition, the current reshooting instrument is equipped with a single camera, and each photo covers a limited area. When the study area is large, a large number of reshooting points need to be set up, resulting in an excessive workload. In the future, the number of cameras carried on the same selfie stick can be increased to achieve the simultaneous reshooting of multiple lenses and angles. This can improve the coverage of features in each photo, reduce the workload, and improve the model's quality further. It is also possible to consider the impact of the camera tilt angle on reshooting.

Although we provide the empirical parameters for reshooting, it is crucial to locate the determined reshooting point's position accurately and quickly on the ground. A mobile app can be developed that combines mobile positioning functions in the future to achieve point layout and automatically collect reshooting points, saving time and effort. At present, the reshooting parameters were deduced based on a single experimental case study, and their applicability to other cases needs further exploration.

5. Conclusions

When a 3D model is established for surrounding features or complex terrain with many surrounding occlusions and large terrain changes, the use of UAV aerial photo modeling alone will lead to insufficient data integrity and poor model quality. With the more and more extensive application of real-scene 3D models, the requirements for model quality are becoming higher and higher. Many studies theoretically show that integrating ground camera photos and UAV aerial photos for modeling is a feasible method to improve the quality of real-scene 3D models. Aiming at exploring the issue from an application perspective, we use a regular camera to reshoot the poor-quality areas of a 3D model and optimize the camera reshooting parameters.

Based on the work in this study, we conclude that the optimal reshooting parameters are as follows: the reshooting distance (in meter units) is 1.5–1.6 times the camera's focal length (in millimeter units); the overlap ratio of the reshooting front is 30%, and the overlap ratio of the reshooting side is 20%. Furthermore, we conclude that an overlap rate of 20% to 30% in both the front and side directions for reshooting is a usable value for practical applications.

In order to evaluate the effect of reshooting, a quality evaluation method based on a real-scene 3D model is used. We evaluate the model from the aspects of point position accuracy, length accuracy, texture structure, and edge length and angle of the triangular mesh. It is found that the point position accuracy and length accuracy of the fused model are well improved, and the texture outline is clearer and more complete. Then, taking a

building as a complex case study, the reshooting parameters of the ground camera are verified, and the topological integrity and accuracy of the building model are enhanced, which provides a favorable reference for improving the quality of the real-scene 3D model. However, the results may be affected by the surrounding geography and the characteristics of the building. Meanwhile, the results from the case study may not be suitable for other scenarios.

In the future, it is necessary to carry out extended research on more scenarios to optimize the parameters of the reshooting further. In this study, the method of manually selecting the connection points is used to achieve the fusion of the two photos. After that, the modeling algorithm for UAV aerial photos and ground camera photo fusion can be further studied to achieve efficient photo fusion. Meanwhile, following the existing studies on building structure vibration [43,44], the method may be extended to monitor building deformations.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/rs16213933/s1. Figure S1: the UAV model; Figure S2: the fused model with a 28 mm focal length, a 40 m reshooting distance, and a 50% overlap ratio; Figure S3: the fused model with a 28 mm focal length, a 45 m reshooting distance, and a 50% overlap ratio; Figure S4: the fused model with a 28 mm focal length, a 50 m reshooting distance, and a 50% overlap ratio; Figure S5: the fused model with a 32 mm focal length, a 45 m reshooting distance, and a 50% overlap ratio; Figure S6: the fused model with a 32 mm focal length, a 50 m reshooting distance, and a 50% overlap ratio; Figure S7: the fused model with a 32 mm focal length, a 55 m reshooting distance, and a 50% overlap ratio; Figure S8: the fused model with a 36 mm focal length, a 50 m reshooting distance, and a 50% overlap ratio; Figure S9: the fused model with a 36 mm focal length, a 55 m reshooting distance, and a 50% overlap ratio; Figure S10: the fused model with a 36 mm focal length, a 60 m reshooting distance, and a 50% overlap ratio; Figure S11: the fused model with a 40 mm focal length, a 60 m reshooting distance, and a 50% overlap ratio; Figure S12: the fused model with a 40 mm focal length, a 65 m reshooting distance, and a 50% overlap ratio; Figure S13: The fused model with a 40 mm focal length, a 70 m reshooting distance, and a 50% overlap ratio.

Author Contributions: Conceptualization, J.X.; data curation, S.Z. and H.J.; formal analysis, S.Z. and P.Q.; methodology, J.X. and S.Z.; supervision, J.X.; validation, N.S. and H.J.; visualization, S.Z. and P.Q.; writing—original draft, J.X. and S.Z.; writing—review and editing, J.X., C.H. and K.B.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was jointly supported by The National Natural Science Foundation of China (Grant No. 42271420), the Natural Science Foundation of Jiangsu Province (Grant No. BK20220367), and the Science and Technology Plan Project of the Ministry of Housing and Urban-Rural Development of China (Grant No. 2022-K-041).

Data Availability Statement: The original contributions presented in this study are included in this article; further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to express their appreciation to the editor and the referees for their comments and suggestions that greatly improved this manuscript.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Deng, F.; Kang, J.H.; Li, P.L.; Wan, F. Automatic true orthophoto generation based on three-dimensional building model using multiview urban aerial photos. *J. Appl. Remote Sens.* **2015**, *9*, 095087. [CrossRef]
2. Chen, J.; Liu, J.J.; Tian, H.B. Basic Positioning and Technical Path of Realistic 3D China Construction. *J. Wuhan Univ. Inf. Sci. Ed.* **2022**, *47*, 1568–1575.
3. Cao, S.S.; Du, M.Y.; Zhao, W.J.; Hu, Y.G.; Mo, Y.; Chen, S.S.; Cai, Y.L.; Peng, Z.Q.; Zhang, C.Y. Multi-level monitoring of three-dimensional building changes for megacities: Trajectory, morphology, and landscape. *ISPRS J. Photogramm. Remote Sens.* **2020**, *167*, 54–70. [CrossRef]
4. Andersen, R.T.; Poulsen, E.S.; Pagola, A.M.; Medhus, B.A. Geophysical mapping and 3D geological modelling to support urban planning: A case study from Vejle, Denmark. *J. Appl. Geophys.* **2020**, *180*, 104130. [CrossRef]

5. Aubry, J.T.; Cerminara, M.; Jellinek, M.A. Impacts of Climate Change on Volcanic Stratospheric Injections: Comparison of 1-D and 3-D Plume Model Projections. *Geophys. Res. Lett.* **2019**, *46*, 10609–10618. [CrossRef]
6. Li, N.; Gong, X.Y.; Li, H.A.; Jia, P.T. Nonuniform multiview color texture mapping of photo sequence and three-dimensional model for faded cultural relics with sift feature points. *J. Electron. Imaging* **2018**, *27*, 011012. [CrossRef]
7. Shan, P.Y.; Sun, W. Research on 3D urban landscape design and evaluation based on geographic information system. *Environ. Earth Sci.* **2021**, *80*, 1–15.
8. Zhang, S.H.; Liu, C.; Haala, N. Guided by model quality: UAV path planning for complete and precise 3D reconstruction of complex buildings. *Int. J. Appl. Earth Obs. Geoinf.* **2024**, *127*, 103667. [CrossRef]
9. Cao, D.G.; Zhang, B.L.; Zhang, X.B.; Yin, L.; Man, X.C. Optimization methods on dynamic monitoring of mineral reserves for open pit mine based on UAV oblique photogrammetry. *Measurement* **2023**, *207*, 112364. [CrossRef]
10. Lai, H.Y.; Liu, L.P.; Liu, X.; Zhang, Y.M.; Xuan, X.H. Unmanned aerial vehicle oblique photography-based superposed fold analysis of outcrops in the Xuhuai region, North China. *Geol. J.* **2020**, *56*, 2212–2222. [CrossRef]
11. Gonzalez-Aguilera, D.; Crespo-Matellan, E.; Hernandez-Lopez, D.; Rodriguez-Gonzalvez, P. Automated Urban Analysis Based on LiDAR-Derived Building Models. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1844–1851. [CrossRef]
12. Wu, B.; Xie, L.F.; Hu, H.; Zhu, Q.; Yaud, E. Integration of aerial oblique imagery and terrestrial imagery for optimized 3D modeling in urban areas. *ISPRS J. Photogramm. Remote Sens.* **2018**, *139*, 119–132.
13. Kenarsari, E.A.; Vitton, J.S.; Beard, E.J. Creating 3D models of tractor tire footprints using close-range digital photogrammetry. *J. Terramech.* **2017**, *74*, 1–11. [CrossRef]
14. Santoši, Ž.; Šokac, M.; Korolija-Crkvenjakov, D.; Kosec, B.; Soković, M.; Budak, I. Reconstruction of 3D models of cast sculptures using close-range photogrammetry. *Metalurgija* **2015**, *54*, 695–698.
15. Pollefeys, M.; Gool, L.V.; Vergauwen, M.; Verbiest, F.; Cornelis, K.; Tops, J.; Kochet, R. Visual Modeling with a Hand-Held Camera. *Int. J. Comput. Vis.* **2004**, *59*, 207–232. [CrossRef]
16. Xu, J.H.; Jing, H.R.; Shen, N. Registration of oblique photography point clouds with terrestrial laser scanning point clouds based on geometric features of irregular building. *Surv. Rev.* **2024**, *56*, 509–524. [CrossRef]
17. Arastounia, M.; Lichti, D.D. Simultaneous identification, modeling and registration refinement of poles using laser scanning point clouds. *ISPRS J. Photogramm. Remote Sens.* **2021**, *181*, 327–344. [CrossRef]
18. Lei, W.; Zhu, S.J.; Jiang, C.; Li, J.Y.; Jiang, K.G.; Teng, C.Q.; Wei, T.; Guo, Q.B. Research on 3D Laser Scanning Monitoring Method for Mining Subsidence Based on the Auxiliary for Probability Integral Method. *KSCE J. Civ. Eng.* **2021**, *25*, 4403–4416.
19. Henn, A.; Gröger, G.; Stroh, V.; Plümer, L. Model driven reconstruction of roofs from sparse LIDAR point clouds. *ISPRS J. Photogramm. Remote Sens.* **2013**, *76*, 17–29. [CrossRef]
20. Jung, J.; Sohn, G. A line-based progressive refinement of 3D rooftop models using airborne LiDAR data with single view photory. *ISPRS J. Photogramm. Remote Sens.* **2019**, *149*, 157–175. [CrossRef]
21. Suveg, I.; Vosselman, G. Reconstruction of 3D building models from aerial photos and maps. *ISPRS J. Photogramm. Remote Sens.* **2003**, *58*, 202–224. [CrossRef]
22. Abayowa, O.B.; Yilmaz, A.; Hardie, C.R. Automatic registration of optical aerial photory to a LiDAR point cloud for generation of city models. *ISPRS J. Photogramm. Remote Sens.* **2015**, *106*, 68–81. [CrossRef]
23. Han, Y.T.; Feng, D.M.; Wu, W.W.; Yu, X.Y.; Wu, G.; Liu, J.X. Geometric shape measurement and its application in bridge construction based on UAV and terrestrial laser scanner. *Autom. Constr.* **2023**, *151*, 104880. [CrossRef]
24. Jo, Y.H.; Hong, S. Three-Dimensional Digital Documentation of Cultural Heritage Site Based on the Convergence of Terrestrial Laser Scanning and Unmanned Aerial Vehicle Photogrammetry. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 53. [CrossRef]
25. Zhou, Y.Z.; Huang, R.G.; Jiang, T.P.; Dong, Z.; Yang, B.S. Highway alignments extraction and 3D modeling from airborne laser scanning point clouds. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *102*, 102429. [CrossRef]
26. Huang, Y.C.; Ma, P.; Ji, Z.; He, L. Part-Based Modeling of Pole-Like Objects Using Divergence-Incorporatiod 3-D Clustering of Mobile Laser Scanning Point Clouds. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 2611–2626. [CrossRef]
27. Wang, F.Y.; Jiang, S.; Zhang, J. Automatic measurement of grid structures displacement through fusion of panoramic camera and laser scanning data. *Eng. Struct.* **2024**, *306*, 117701. [CrossRef]
28. Gašparović, M.; Malarić, I. Increase of readability and accuracy of 3D models using fusion of close range photogrammetry and laser scanning. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2012**, *39*, 93–98. [CrossRef]
29. Singh, S.P.; Jain, K.; Mandla, V.R. A new approach towards photo based virtual 3D city modeling by using close range photogrammetry. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *II-5*, 329–337. [CrossRef]
30. Li, W.; Liu, G.; Lu, K.; Chen, P.; Cui, J.; Yu, M.; Shen, P. Fast matching method of UAV aerial photography enhanced low illumination image. *Int. J. Aerosp. Eng.* **2022**, *1*, 9543893. [CrossRef]
31. Shahid, M.; Abhishek, B.; Channappayya, S.S. A Cross-Platform HD Dataset and a Two-Step Framework for Robust Aerial Image Matching. *IEEE Access* **2022**, *10*, 66153–66174. [CrossRef]
32. Zhu, Q.; Wang, Z.D.; Hu, H.; Xie, L.F.; Ge, X.M.; Zhang, Y.T. Leveraging photogrammetric mesh models for aerial-ground feature point matching toward integratiod 3D reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2020**, *166*, 26–40. [CrossRef]
33. Zhang, G.X.; Qin, D.Y.; Yang, J.Q.; Yan, M.Y.; Tang, H.P.; Bie, H.Z.; Ma, L. UAV Low-Altitude Aerial photo Stitching Based on Semantic Segmentation and ORB Algorithm for Urban Traffic. *Remote Sens.* **2022**, *14*, 6013. [CrossRef]

34. Sun, Y.B.; Sun, H.B.; Yan, L.; Fan, S.Y.; Chen, R. RBA: Reduced Bundle Adjustment for oblique aerial photogrammetry. *ISPRS J. Photogramm. Remote Sens.* **2016**, *121*, 128–142. [CrossRef]
35. Nesbit, R.P.; Hugenholtz, H.C. Enhancing UAV–SfM 3D Model Accuracy in High-Relief Landscapes by Incorporating Oblique photos. *Remote Sens.* **2019**, *11*, 239. [CrossRef]
36. Baramsyah, H.; Rich, L. Applicability Assessments of Close-Range Photogrammetry for Slope Face 3D Modelling. *Aceh Int. J. Sci. Technol.* **2019**, *8*, 144–151. [CrossRef]
37. Cheng, L.; Zhang, W.; Zhong, L.; Du, P.; Li, M. Framework for Evaluating Visual and Geometric Quality of Three-Dimensional Models. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 1281–1294. [CrossRef]
38. Zhang, F.; Li, X.Y.; Liu, H.; Li, W.; Hu, W.; Li, H.C. Three-dimensional terrain model multiview quality assessment considering human visual system. *J. Appl. Remote Sens.* **2015**, *9*, 90. [CrossRef]
39. Iwaszczuk, D.; Stilla, U. Camera pose refinement by matching uncertain 3D building models with thermal infrared photo sequences for high quality texture extraction. *ISPRS J. Photogramm. Remote Sens.* **2017**, *132*, 33–47. [CrossRef]
40. Elberink, O.S.; Vosselman, G. Quality analysis on 3D building models reconstructed from airborne laser scanning data. *ISPRS J. Photogramm. Remote Sens.* **2010**, *66*, 157–165. [CrossRef]
41. Tran, H.; Khoshelham, K.; Kealy, A. Geometric comparison and quality evaluation of 3D models of indoor environments. *ISPRS J. Photogramm. Remote Sens.* **2019**, *149*, 29–39. [CrossRef]
42. *GB/T 9730-2008*; General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China; National Standardization Administration of the People's Republic of China. Specification for Aerophotogrammetric Office Operation of 1:500 1:1000 1:2000 Topographic Maps. Standards Press of China: Beijing, China, 2008.
43. Han, Y.T.; Wu, G.; Feng, D.M. Vision-based displacement measurement using an unmanned aerial vehicle. *Struct. Control Health Monit.* **2022**, *29*, e3025. [CrossRef]
44. Shan, J.Z.; Huang, P.C.; Loong, N.C.; Liu, M. Rapid full-field deformation measurements of tall buildings using UAV videos and deep learning. *Eng. Struct.* **2024**, *305*, 117741. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.