

AN OBJECT-ORIENTED UAV 3D PATH PLANNING METHOD APPLIED IN CULTURAL HERITAGE DOCUMENTATION

Xinyue Liu¹, Zheng Ji^{1,*}, Hao Zhou¹, Zuxun Zhang¹, Pengjie Tao¹, Ke Xi¹, Lingfeng Chen¹, Jose Marcato Junior²

¹ School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China,
(lxy2015, jz07, haozhu47)@whu.edu.cn, zhangzx@cae.cn, (pjtao, xike19, whuchenlingfeng)@whu.edu.cn

² Faculty of Engineering, Architecture and Urbanism and Geography, Federal University of Mato Grosso do Sul,
Campo Grande 79070-900, Brazil, jose.marcato@ufms.br

KEY WORDS: UAV, Path planning, Cultural heritage, Object-oriented, Structure line.

ABSTRACT:

Since image-based 3D model reconstruction can faithfully recover the real texture of cultural heritage with high accuracy, it is widely used in cultural heritage documentation. Given the complexity of manual image acquisition at present, we propose an object-oriented unmanned aerial vehicle (UAV) path planning method to obtain close-up and high-resolution images for the 3D reconstruction of cultural heritage. Four basic geometric classes are defined and can be automatically divided or interactively defined on the surface of an initial coarse model. We propose the concept of aerial strip unit in conventional photogrammetry to generate multiple regular strip units for photography. The optimal flight path connecting each unit is generated considering the obstacle avoidance and the shortest distance. Based on the self-developed 3D engine, we take the Ancient City of Ping Yao and Yellow Crane Tower in China as two cases to design the UAV 3D path planning. Experimental results show that, compared with general planning methods, our method can improve the flight efficiency of UAV and the visual fineness of the reconstruction results.

1. INTRODUCTION

In recent years, the documentation of cultural heritage has received increasing attention. The refined 3D reconstruction is an essential task, which plays a great role in analyzing and preserving cultural heritage. Although terrestrial laser scanners (TLS) can directly obtain high-precision coordinates of the target surface, it is not applicable to large targets and usually has a high acquisition cost. With the rapid development of UAV technology, multi-rotor UAVs are widely used to acquire close-range images with high resolution in the current 3D reconstruction of cultural heritage due to their low cost, small size, and high flexibility (Murtiyoso and Grussenmeyer, 2017). Image-based 3D reconstruction methods combining structure from motion (SfM) and multi-view stereo (MVS) approaches can generate detailed 3D point clouds and surface models with only monocular cameras (Aicardi et al., 2018). However, since the key of these algorithms is to extract geometric information in multi-view images, the quality of the reconstruction results depends heavily on the acquired image quality (Goesele et al., 2007). Therefore, the photographic path planning of UAV is a prerequisite for improving the quality of 3D reconstruction.

The path planning methods of conventional aerial photogrammetry and oblique photogrammetry are mostly horizontal or terrain-following. These methods usually perform regular path like a grid or circle beyond a certain distance above the target for flight safety. However, for the refined 3D reconstruction of some unconventional ground or intricate artificial objects, the images obtained by these path planning methods cannot solve the self-occlusion problems or loss of detail caused by remote photography. To solve such problems, the concept of nap-of-the-object photogrammetry (NoOP) was proposed, and the corresponding path planning and post-processing methods have also been researched and implemented

(He, 2019). NoOP requires UAVs to fly close along the surface of the target and photograph toward the target.

In some previous works on cultural heritage reconstruction, the UAV paths were designed based on target characteristics and reconstruction requirements. For the reconstruction of building facades and other 2D objects (Bolognesi et al., 2015; Bakirman et al., 2020), the path planning approach is straightforward because only a few factors such as image overlap and flight height need to be considered. For some 3D targets, researchers usually design the camera network according to the reconstruction requirements (Pan et al., 2019; Jo and Hong, 2019) and then perform the flight mission in manual mode (Themistocleous et al., 2015; Sun and Zhang, 2018; Federman et al., 2018), which requires experienced operators, but it will still bring discrepancies to the actual work due to the site conditions and manual operations.

In the current urban 3D reconstruction work, to ensure high accuracy and full coverage, the method based on a coarse proxy model is the most used solution (Roberts et al., 2017; Smith et al., 2018). This coarse-to-fine planning strategy first reconstructs the target to generate an initial coarse model, then designs an optimal path according to the coarse proxy model, and finally performs dense image acquisition. This work analyses the surface of the initial model to determine the viewpoint position and orientation that can cover the target. Many crucial parameters that can affect the reconstruction quality are considered, such as image overlap, parallax angle, and observation angle (Hoppe et al., 2012; Hepp et al., 2018; Koch et al., 2019). In addition, it is necessary to generate a collision-free optimal path with guaranteed quality of the reconstruction results, which is generally defined as the shortest flight length or minimum energy consumption (Zheng et al., 2018). Several studies consider the shortest path, obstacle avoidance, and turning angle (Hepp et al., 2018; Yan et al., 2021; Zhang et al., 2021).

* Corresponding author

In this paper, we propose an object-oriented UAV 3D path planning method, which divides the targets into points, lines, planes, and body objects according to the coarse proxy model and generates the flight path composed of elementary strip units to meet the accuracy requirements of refined 3D reconstruction. We considered two architectural heritage sites in China to show the results and application of 3D reconstruction using the proposed path planning method.

2. UAV PATH PLANNING METHODS

In order to obtain high-quality images for refined 3D reconstruction, UAV path planning is required. The path planning methods need to consider the limitations of the flight platform, camera, and other hardware equipment. In addition, the image requirements for 3D reconstruction, such as the image overlap and the ground sampling distance (GSD), need to be considered. The initial model of the target is acquired prior to the flight to plan the UAV path.

The conventional aerial photogrammetry uses a 2D flight mode, in which the UAV flies at a fixed altitude and photographs the target vertically downward or at a fixed tilted angle. This method is suitable for relatively flat areas, and the flight height and strip spacing are planned according to a 2D map. This method only needs to solve the position (x_i, y_i) and flight height H of the UAV. When there are large terrain undulations in the target area, the images will have large geometric deformation and resolution differences, reducing reconstruction accuracy.

The 2.5D flight mode is to change the fixed flight height H to variable height h_i according to the elevation of the target area based on the 2D mode. This path planning method takes the digital elevation model (DEM) as the initial model to determine h_i to keep the photographic distance constant. The problem of inconsistent image resolution is solved to some extent, but the image distortion due to the fixed camera angle still exists.

Different from the above flight modes, the 3D path planning in NoOP needs to solve the UAV position (x_i, y_i, z_i) in 3D space and the camera angle $(pitch_i, yaw_i)$ towards the target surface according to the initial 3D model of the target. The path file transmitted to the flight control equipment includes the continuous track positions and the camera orientation parameters. Multi-rotor UAVs with the real-time kinematic (RTK) module can ensure the accuracy of the UAV position and achieve close-up photography of up to 5 m to obtain ultra-high-resolution images at the millimeter level. At the same time, the gimbal with a continuously controllable angle has high flexibility, and the camera angle can be adjusted according to the planning result to achieve object-oriented photography. The characteristics of different flight modes are listed in Table 1.

Flight mode	Initial model	Parameters to be solved
2D	2D map	(x_i, y_i, H)
2.5D	DEM	(x_i, y_i, h_i)
3D	3D model	$(x_i, y_i, z_i), (pitch_i, yaw_i)$

Table 1. The characteristics of different flight modes.

3. THE OBJECT-ORIENTED 3D PATH PLANNING METHOD

This paper proposes and implements an object-oriented 3D UAV path planning method, which can effectively acquire images to obtain the surface information of the target with millimeter-level accuracy and provide a data basis for subsequent high-accuracy

3D reconstruction and texture and geometric information preservation. Compared with the existing research and commercial software, our method is universal. It supports multiple types of geographic data and implements the planning according to the objects. The generated flight path is more in line with flight control and can be modified for obstacle avoidance. Fig.1 shows the overall pipeline of our proposed method.

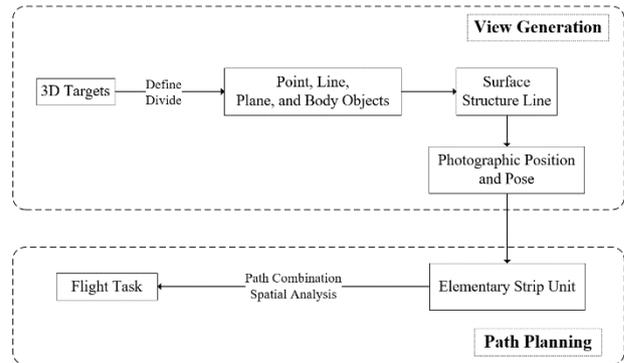


Figure 1. The overall pipeline of the proposed method.

The flight path is planned based on a geometric scene proxy (Smith et al., 2018), which can be generated from multiple kinds of data. For different types of targets, we define four basic geometric classes and generate the corresponding surface of the instantiated object. Then, we generate the discrete surface structure lines according to the reconstruction requirements, and the position and orientation of object-oriented viewpoints are determined. A group of viewpoints constitutes the elementary strip unit in photogrammetry. The final flight path is formed by the spatial analysis and path combination of the strip units.

3.1 The Coarse Proxy Model

The whole process is implemented in our own engine, independent of other open-source 3D frameworks that most platforms rely on. Our engine supports various geographic data formats, including DEM, laser scanning data, .osgb data, and building information modeling (BIM) data such as .ifc and .fbx data. In essence, the above data can be transformed into a 3D mesh model. The lack of surface information and normal vector information in point cloud data can be solved by normal vector estimation and Poisson surface reconstruction. For the BIM models in the local coordinate system, we implement a unified orientation and coordinate conversion module based on the PROJ library (Proj Contributors, 2022) to conduct the conversion from the local coordinate system to the projection coordinate system and then to the world coordinate system.

In addition, China is advocating a new type of fundamental surveying and mapping and constructing a 3D real scene database. The relevant geographic databases are constantly being improved, and many cities have built their fundamental 3D models that can be used as the coarse proxy model. We can quickly obtain a coarse model for areas without a proxy model by oblique photogrammetry with regular baseline or manual mode path planning methods.

3.2 Interactive Definition of Photographic Objects

Many flight path planning software does not choose the photographic object in 3D space. It defines a plane area and then realizes the 3D image acquisition by specifying the flight height. This method is not intuitive, and it is difficult to deal with

irregular objects, resulting in a complex operation and extra learning costs.

We define four basic geometric classes for object definition, namely point, line, plane, and body classes. The definition of basic geometric classes facilitates the subsequent surface division and grouping of different objects and determines the surface points to be photographed of the target through discretization. In NoOP, it is obvious that the surface normal vector of the object can be directly converted to the photographic direction. Each target can be flexibly defined as a combination of the basic geometric objects to achieve complete coverage, including the local details.

The point class is mainly used in areas that require special supplementary photography. By the ray intersecting method, the position of the point object on the mesh model surface can be determined, and the normal vector of the surface is usually taken as the photographic direction. For points with multiple normal vectors, such as edge points or corner points, we provide a convenient interactive way to determine the photographic direction by modifying the pitch and yaw angle of the camera, as shown in Fig.2 (a).

The line class is mainly applied at discontinuities of the target's boundaries. The line object is determined by selecting two points on the mesh model surface, and the photographic direction is normalized by the mean of the normal vector on the surface of the two points. The photographic direction can also be modified with two degrees of freedom through the interaction, as shown in Fig.2 (b).

The plane class is mainly applied to a large range of flat surfaces. After selecting the plane object, the photographic direction is directly determined by the normal vector of the plane, as shown in Fig.2 (c).

The body class is composed of a prism, which is defined in two ways. Input the number of faces of the prism, define the center point, and then stretch outward from the center point to form the top surface. Finally, stretch it in the height direction to form the prism, as shown in Fig.2 (d). Alternatively, an irregular plane is firstly defined by the similar method of plane object definition, and then the prism is formed in the height direction, as shown in Fig.2 (e). The two methods are for regular and irregular prisms uniformly called the sketch body model. The surface of the sketch body model can represent the target surface to be photographed, which is the combination of multiple continuous plane objects mathematically. The photographic direction of each plane object is parallel to its normal vector. Usually, we do not generate a bottom surface of the body object because most UAVs currently do not support upward photography.

This section describes more about how to interactively define photographic objects, which is time-consuming in the experiment. Although it is a manual operation, the convenient interaction enables us to quickly determine the photographic objects, especially the combined use of the body object with the point and line objects. The body object is for a wide range of all-around coverage, and the point and line objects are for local details. In terms of ease of use and time consuming, our integrated approach is similar to the currently available automatic algorithms, and even for flawed initial models, human intelligence can be extremely useful due to the interaction. The applicability of our approach is also higher than that of automatic methods.

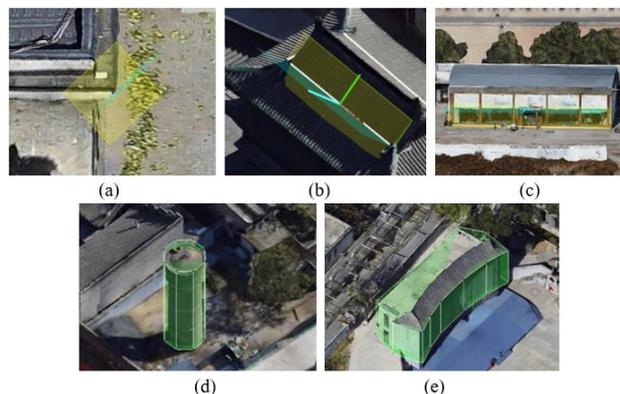


Figure 2. Illustrations of different geometric classes for object definition. (a) Point; (b) Line; (c) Plane; (d) Regular prism; (e) Irregular prism.

3.3 Automatic Extraction of Photographic Objects

We also propose the automatic method to realize the extraction of objects. We hope to automatically extract the object based on the four geometric classes from the surface of the photographic target.

The existing methods mainly sample the surface of the 3D mesh model directly to generate sampling points, which correspond to our point objects (Smith et al., 2018; Zhou et al., 2020). Considering the need to subsequently generate paths that are more compatible with UAV flight, which is explained in the next section, more effort was put into the automatic extraction of line, plane, and body objects.

For line objects, we use contour lines to divide each target in height, divide it at the top along the short side of the enveloping rectangle of the cross-section, and cooperate with the discrete point clustering algorithm. The Douglas-Peucker algorithm is used to realize the extraction of line objects, as shown in Fig. 3.

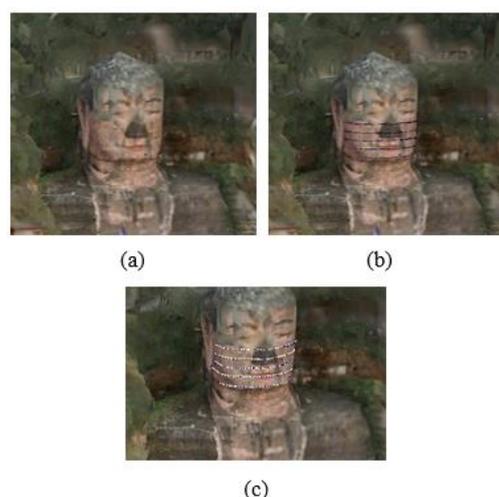


Figure 3. The target (a), intersection points of contour lines with the surface (b), and generated line objects (c).

The extraction methods of plane and body objects, which correspond to the existing algorithms such as polygon reduction, facade extraction, convex hull, and model monomer, are being further integrated. The usability and practicality have been greatly improved with the manual interaction mode provided.

3.4 Elementary Strip Unit Planning

The flight path planned by the methods that rely on discrete viewpoints has some defects. Constrained by aerodynamics, UAV is relatively stable and energy-saving when flying at a fixed altitude. Constant changes in flight height will bring additional acceleration and deceleration processes, resulting in increased energy consumption and reduced UAV working time, which should be avoided for small UAVs with limited battery endurance. Zhou et al. (2020) used Poisson disk sampling to obtain the sampling point objects on the target surface at different heights. The subsequent generation and optimization of the viewpoints are based on these points. After these operations, the viewpoints and photographic directions tend to be fragmented. Zheng et al. (2018) used the contour lines segmentation method to segment the buildings, but it does not consider the increase in energy consumption caused by the UAV height changes when generating the path.

Based on the conventional aerial photogrammetry, we introduce the concepts of surface structure line and elementary strip unit. The objects of the four geometric classes defined above can essentially be transformed into the structure lines or a combination thereof. For example, the point object is a structure line with a length of 0, and the polyline, plane, and body object are a combination of multiple structure lines. Considering the accuracy requirements of reconstruction, the parameters such as image overlap, parallax angle, and resolution are determined, and the structural lines (l_1, \dots, l_n) of the target are sampled. The normal vector corresponding to each structure line is (IN_1, \dots, IN_n) , and the discrete points $p_i = (p_{i1}, \dots, p_{in})$ of each structure line l_i are generated. The sampling calculation is defined as

$$\begin{aligned}
 s. t. \quad h_{overlap} &= 1 - \frac{dis(l_i, l_{i+1})}{H} = 50\% ; \\
 w_{overlap} &= 1 - \frac{dis(p_{i,j}, p_{i,j+1})}{W} \in (50\%, 75\%) ,
 \end{aligned} \tag{1}$$

where $W = GSD * pixels_w$, and $H = GSD * pixels_H$. The viewpoints generated by a structure line constitute the elementary strip unit.

For the structural lines l_i , assume that the Y direction is the true north direction, and the viewpoint v_{ij} is represented by its position (x_{ij}, y_{ij}, z_{ij}) and photographic direction $(pitch_{ij}, yaw_{ij})$, which are calculated by

$$\begin{aligned}
 pitch_{ij} &= \varphi; \quad yaw_{ij} = \kappa; \\
 (x_{ij}, y_{ij}, z_{ij}) &= p_{ij} + IN_i * d, \\
 s. t. \quad IN_i &= \text{Normalize}(\sum_{k=0}^n normal_k),
 \end{aligned} \tag{2}$$

where $normal_k$ is the normal vector of the triangular surface s_k crossed by l_i , and $d = \frac{focalLength}{pixelSize} * GSD$. The generation of structure lines, aerial strips, and viewpoints are shown in Fig.4.

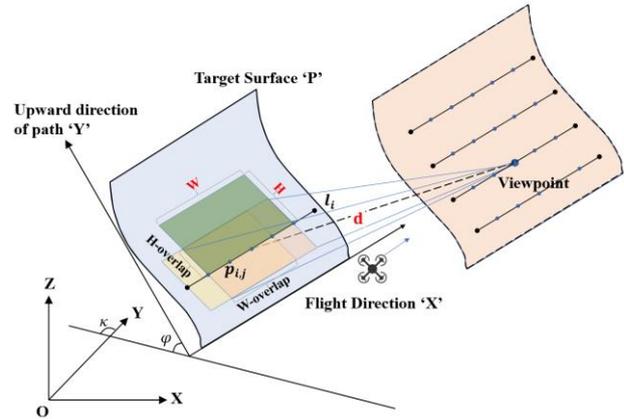


Figure 4. The generation of structure lines, aerial strips, and viewpoints.

We apply a convergent photographic to improve coverage for some special areas such as corners, especially for the facade that cannot be observed by vertical photography. When generating the viewpoints in the convergent photography, the pitch and yaw angle can be added or subtracted an angle to change the photographic position, which is equivalent to manual interaction to modify IN_i .

The elementary strip unit is generated in consideration of the reconstruction requirements. For example, the side overlap is generally set at 50%, and the heading overlap is set at 50%–75%. The resolution can be set to millimeter-level accuracy as required. It eliminates the need for additional optimization of the viewpoints. The elementary strip unit generated from the most frequently used plane and body objects can ensure that each strip unit is horizontal, which greatly improves the efficiency of UAV flight. Fig. 5 shows the comparison of our path with the paths of other methods.

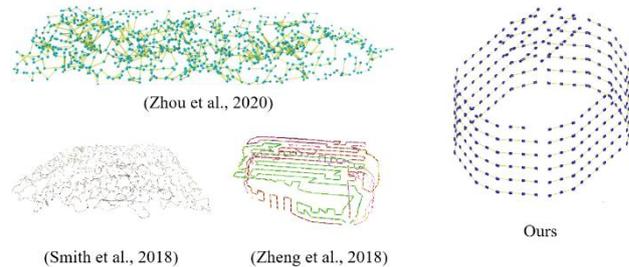


Figure 5. Comparison of paths generated by different methods.

3.5 Path Planning and Obstacle Avoidance

The path generated for geometric objects is still discrete in essence, consisting of elementary strip units. The conventional and oblique photogrammetry do not consider path planning and obstacle avoidance because the strip units are relatively regular, and the flight altitude is much higher than the photographic target. However, since the close-up strips are more discrete, and the flight altitude is almost the same as the target height, the obstacle avoidance and optimal path problem need to be considered to form the final path.

We need to determine whether it crosses an obstacle such as a wall for each elementary strip unit. Here we only need to determine whether there is an intersection between the strip and the triangle within the range of each corresponding model block. For such strips, we provide corresponding interactive editing

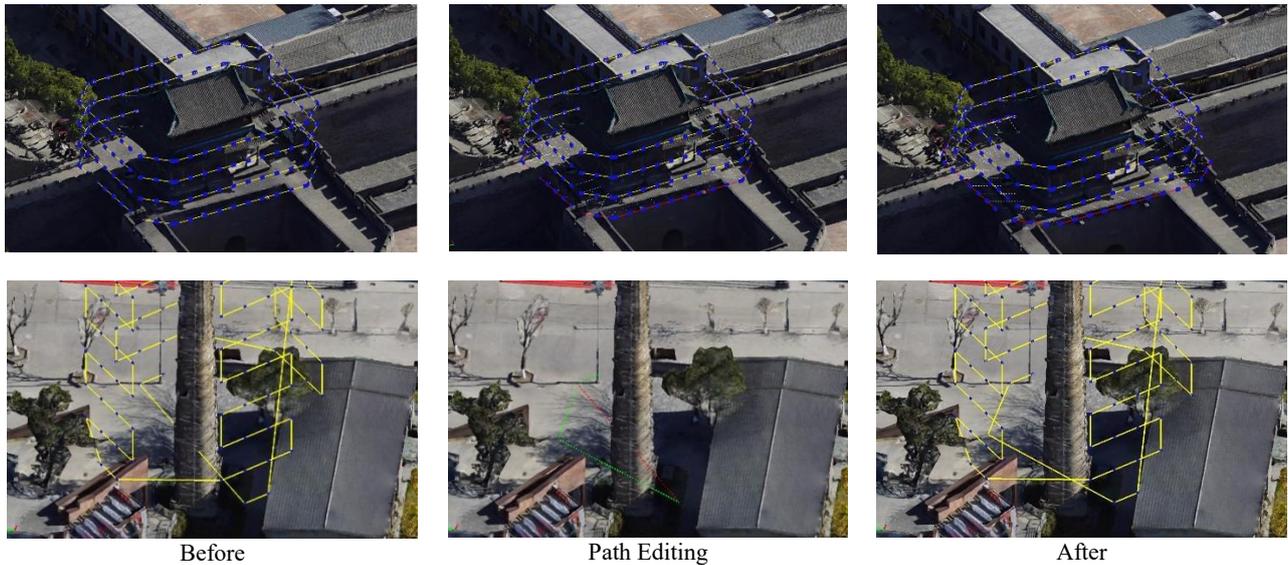


Figure 6. Path editing methods. The first row shows the method of lifting the overall strip geometry whereas the second row shows the method of modifying the position of local points.

methods to lift the overall strip or modify the position of local points, as shown in Fig. 6.

For the strip completely within the range of obstacles, that is, the strip does not intersect with the triangle but is invalid; we judge whether the strip crosses the obstacle through the connection between the elementary strip units. For obstacles that do not exist in the initial model, manual field observation is required.

For path planning, we construct a heterogeneous traveling salesman problem (TSP) with the graph-based data structure. The elementary strip unit in 3D space is taken as a node in the graph to solve the shortest path connection problem. Considering the energy consumption of UAV, for the cost between every two nodes, we consider the horizontal distance, vertical distance, and angle change. Set $L = \{L_1, \dots, L_n\}$ to be the set of all strips. For each strip L_i , mark its two endpoints and denote the i -th strip by $L_i[(X_1, Y_1, Z_1), (X_2, Y_2, Z_2)]$. Then the distance between two strips and the angle change can be calculated. Set $E = \{(r, s) : r, s \in L\}$ to be the set of connections between strips. The cost function is defined as

$$\Delta\tau_{rs} = \alpha * \Delta XY_{rs} + \beta * \Delta Z_{rs} + \gamma * \varphi_{rs}, \quad (3)$$

where ΔX_{rs} is the horizontal distance between two connecting points of strip L_r , L_s , and ΔZ_{rs} represents the vertical distance. φ_{rs} is the angle of UAV flying from L_r to L_s . α , β , and γ are the weights of the three factors, respectively

For distance calculation, we reset the weight to 3 in the height direction and 1 in the horizontal direction according to experience. It should be noted each of our nodes is an elementary strip unit with a starting and an ending point. The starting and ending points are not determined before planning, so the cost $\Delta\tau_{rs}$ between every two nodes is calculated with four different values. In addition, we need to consider whether the path crosses the obstacle when calculating its cost. The cost is set to be infinite for the elementary strip unit that is completely within the obstacle range. We take the minimum cost and update the cost between

nodes after each node is connected to generate the best path connection scheme iteratively.

4. EXPERIMENTS AND DISCUSSION

4.1 Two Experimental Cases

Among the existing automatic algorithms, the process of UAV path generation is extremely complicated. It is difficult to complete all processes in a program, including scene management, flight area division, path generation, and obstacle avoidance optimization, making it difficult to be conveniently and effectively applied in the protection of cultural heritage. Based on the engine framework we constructed, we took the Ancient City of Ping Yao and Yellow Crane Tower in China as two cases to perform the integration process of the whole 3D planning. The UAV images were acquired by DJI Phantom 4 RTK, which has a centimeter-level positioning system and a 3-axis gimbal system (Dji, 2018).

The Ancient City of Ping Yao, located in central Shanxi Province, China, is a traditional city built in the 14th century. It is a collection of ancient walls, streets, shops, dwellings, and temples. It was included in the UNESCO World Heritage List in 1997. In October 2021, part of the city walls collapsed due to heavy rainfall, which further made people realize the importance of the preventive protection of cultural heritage relying on 3D reconstruction documentation. However, its complicated urban composition and architectural structure also put forward higher requirements for UAV path planning tasks. Firstly, we quickly obtained the initial model of the Ancient City of Ping Yao by vertical photography, as shown in Fig. 7 (a). Much of the facade information is missing due to the vertical photography. There are some isolated towers in the scene, but most are low buildings close together, which is not suitable for close-up parallel photography. Therefore, for the isolated towers, we used the body class to define the outer surface of the tower and quickly generate the close-up path, as shown in Fig. 7 (b). For the area with dense buildings, in addition to the vertical photography, close-up convergent photography was adopted; that is, the main optical axis is not perpendicular to the surface, and finally, a collision-free path can be generated, as shown in Fig. 7 (c) (d).

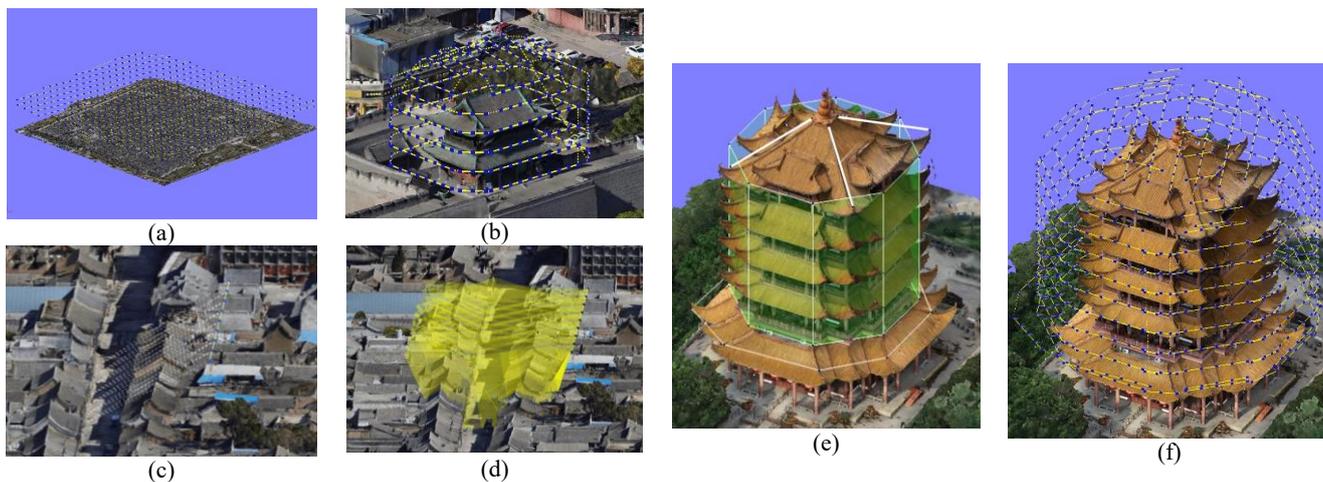


Figure 7. Different path planning methods. (a) The vertical photography; (b) Path planning for the body object; (c) (d) The convergent photography; (e) Path planning for the combination of body and line objects; (f) The generated path for the Yellow Crane Tower.

The Yellow Crane Tower in Hubei Province, which was first built in 223 AD, is also of high cultural value as a magnificent building. First, we took photos around the Yellow Crane Tower by specifying the radius and center to generate the initial model. The initial model has facade information due to the wraparound photography, but it is not detailed enough. Besides, the building is not cylindrical, and the circumferential flight mode leads to inconsistent resolution of the building, which is a common problem in reconstructing irregular targets. On this basis, we used the body class to plan a more detailed path for the facade and the top surface. For the top details, the line class was used for further planning, as shown in Fig. 7 (e), to improve and unify the image resolution of the building surface. The complete path generated is shown in Fig. 7 (f), where solid lines represent elementary strip units and dotted lines represent transition paths between strip units.

4.2 Comparison of Coarse-to-fine

We carried out path planning for the Ancient City of Ping Yao and Yellow Crane Tower, respectively, and transmitted the generated continuous track positions and camera orientation parameters as path files to the flight controller for flight and photography tasks. We present the reconstruction results of images obtained using our planning method and compare them with the reconstruction results of regular baselines, as shown in Fig. 8.

The Ancient City of Ping Yao is a large scene with a complicated composition of buildings. For its 3D reconstruction, the current planning method is usually oblique photography following the regular grid above the scene. This planning method will lead to the loss of building facade information.

Through comparison, it can be found that our method can also perform refined reconstruction for the area under the eaves,

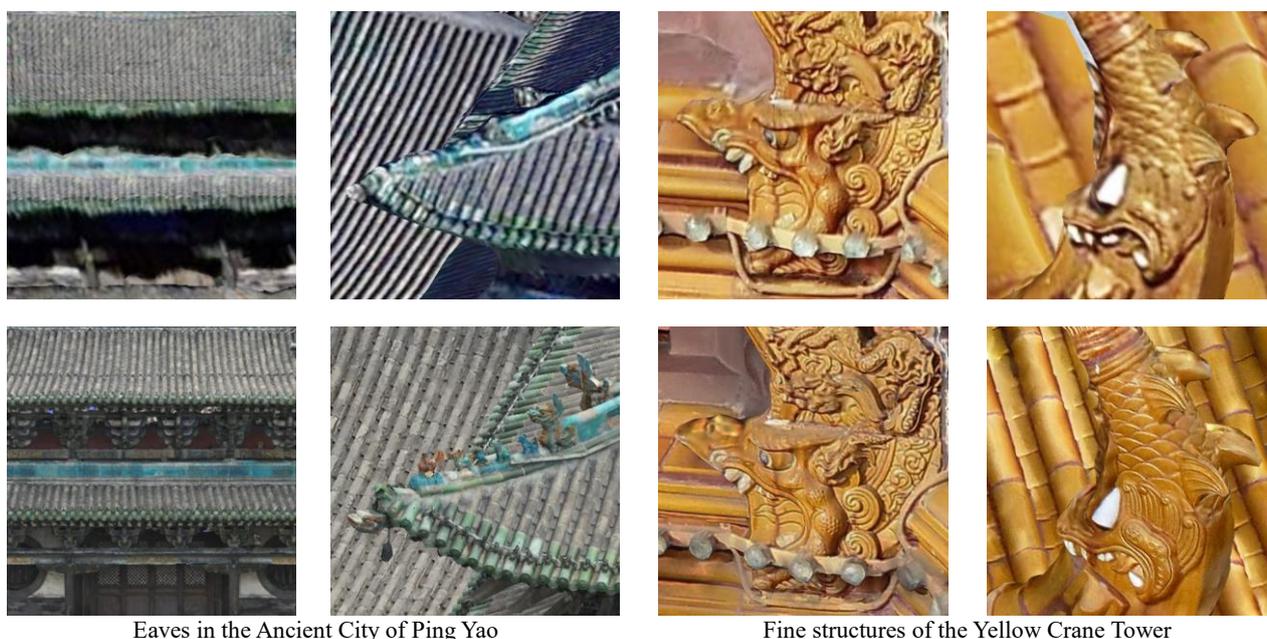


Figure 8. Comparison of reconstruction results of different path planning methods. The first row shows the results of regular baselines whereas the second row shows the results of our path planning method.

which better solves the problem of texture loss caused by the building self-occlusion. In addition, the fineness of the reconstruction results is significantly improved due to the close-up photography.

A circular wraparound baseline is usually adopted for the path planning of specific buildings such as the Yellow Crane Tower. Although this type of method can obtain the facade information of the target, the reconstruction of the details is not sufficient because it is not planned for the characteristics of the target. We can see that our method can better recover fine structures through comparison.

4.3 Facade Images Generation

The archaeological line drawing is an important form of archiving cultural relics. The orthographic projection principle is adopted to accurately recover the appearance features of cultural relics in the form of lines. Compared with images and 3D models, line drawings can more accurately represent the characteristic information and structural relations of each part of the artifact. The previous manual mapping method is inefficient and may cause new damage to cultural relics. The refined 3D model completely recovers the high-resolution surface information of cultural relics and can generate facade images with detailed features for feature extraction and line drawing.

In the engine we build, we can achieve an orthophoto image of the target surface from any direction. The orthographic image obtained by vertical projection is the digital orthophoto map (DOM) in the conventional photogrammetry, while the facade image can be obtained by horizontal projection. The method is as follows: Orthographic projection is applied to the surface of the model to obtain the surface within the range of the current view, which can generate an initial image. The required facade image can be obtained after the image dodging processing, as shown in Fig. 9.

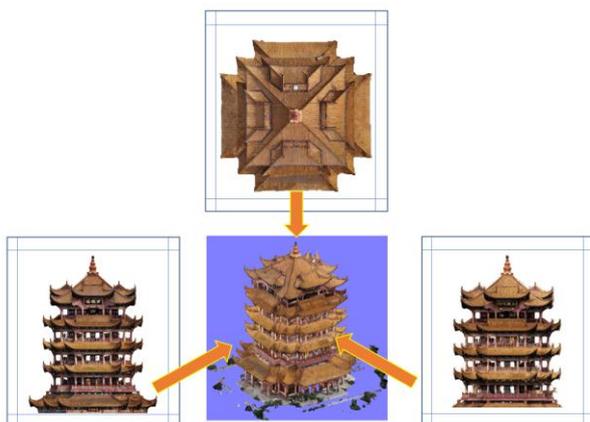


Figure 9. Facade images of the Yellow Crane Tower.

5. CONCLUSIONS

In this paper, we propose an object-oriented path planning method and provide an integrated operation tool for the close-up path planning of refined 3D reconstruction of cultural heritage. We define four basic geometric classes, which support the interactive definition and automatic division of photographic objects on the initial models of various data types. According to the requirements of refined 3D reconstruction, we calculate the resolution, overlap, and exterior orientation parameters of the camera and then generate the elementary strip units suitable for

UAV flight. We adopt a path optimization to generate the shortest path of UAV flight considering obstacle avoidance. Taking the Ancient City of Ping Yao and Yellow Crane Tower as examples, experiments show that compared with general planning methods, our method is more suitable for UAV flight, and the obtained 3D model has excellent visual fidelity. However, there are some limitations to our approach. The robustness of automatic division suitable for all kinds of scenes and data is limited, especially in the extraction of photographic objects. At present, more emphasis is placed on manual interaction, so we are working on improving the universality of our method. In the future, our main work will focus on integrating more intelligent algorithms to improve planning efficiency further.

ACKNOWLEDGEMENTS

The authors would like to acknowledge DJI Enterprise of Shenzhen DJI Technology Co., Ltd. for providing experimental equipment and data.

REFERENCES

- Aicardi, I., Chiabrando, F., Maria Lingua, A., and Noardo, F., 2018. Recent trends in cultural heritage 3D survey: The photogrammetric computer vision approach, *Journal of Cultural Heritage*, 32, 257-266. doi.org/10.1016/j.culher.2017.11.006.
- Bakirman, T., Bayram, B., Akpınar, B., Karabulut, M. F., Bayrak, O. C., Yigitoglu, A., and Seker, D. Z., 2020. Implementation of ultra-light UAV systems for cultural heritage documentation, *Journal of Cultural Heritage*, 44, 174-184. doi.org/10.1016/j.culher.2020.01.006.
- Bolognesi, M., Furini, A., Russo, V., Pellegrinelli, A., and Russo, P., 2015. Testing the low-cost RPAS potential in 3D cultural heritage reconstruction, *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*. doi.org/10.5194/isprsarchives-XL-5-W4-229-2015.
- DJI, 2018. Phantom 4 RTK - DJI. dji.com/phantom-4-rtk (10 Jan 2022).
- Federman, A., Shrestha, S., Quintero, M. S., Mezzino, D., Gregg, J., Kretz, S., and Ouimet, C., 2018. Unmanned Aerial Vehicles (UAV) Photogrammetry in the Conservation of Historic Places: Carleton Immersive Media Studio Case Studies, *Drones*, 2, 18. doi.org/10.3390/drones2020018.
- Goesele, M., Snavely, N., Curless, B., Hoppe, H., and Seitz, S. M., 2007. Multi-View Stereo for Community Photo Collections, *2007 IEEE 11th International Conference on Computer Vision*, 14-21 Oct. 2007, 1-8. doi.org/10.1109/ICCV.2007.4408933.
- He, J., 2019. Nap-of-the-Object Photogrammetry and Its Key Techniques, Wuhan University. 10.27379/d.cnki.gwhdu.2019.000509.
- Hepp, B., Nissner, M., and Hilliges, O., 2018. Plan3D: Viewpoint and Trajectory Optimization for Aerial Multi-View Stereo Reconstruction, *ACM Transactions on Graphics*, 38, Article 4. doi.org/10.1145/3233794.
- Hoppe, C., Wendel, A., Zollmann, S., Pirker, K., Irschara, A., Bischof, H., and Kluckner, S., 2012. Photogrammetric camera network design for micro aerial vehicles, *Computer vision winter workshop (CVWW)*, 1-3.

- Jo, Y. H. and Hong, S., 2019. Three-Dimensional Digital Documentation of Cultural Heritage Site Based on the Convergence of Terrestrial Laser Scanning and Unmanned Aerial Vehicle Photogrammetry, *ISPRS International Journal of Geo-Information*, 8, 53. doi.org/10.3390/ijgi8020053.
- Koch, T., Körner, M., and Fraundorfer, F., 2019. Automatic and Semantically-Aware 3D UAV Flight Planning for Image-Based 3D Reconstruction, *Remote Sensing*, 11, 1550. doi.org/10.3390/rs11131550.
- Murtiyoso, A. and Grussenmeyer, P., 2017. Documentation of heritage buildings using close-range UAV images: dense matching issues, comparison and case studies, *The Photogrammetric Record*, 32, 206-229. doi.org/10.1111/phor.12197.
- Pan, Y., Dong, Y., Wang, D., Chen, A., and Ye, Z., 2019. Three-Dimensional Reconstruction of Structural Surface Model of Heritage Bridges Using UAV-Based Photogrammetric Point Clouds, *Remote Sensing*, 11, 1204. doi.org/10.3390/rs11101204.
- PROJ contributors, 2022. PROJ coordinate transformation software library. Open Source Geospatial Foundation. proj.org (10 Jan 2022).
- Roberts, M., Shah, S., Dey, D., Truong, A., Sinha, S., Kapoor, A., Hanrahan, P., and Joshi, N., 2017. Submodular Trajectory Optimization for Aerial 3D Scanning, *2017 IEEE International Conference on Computer Vision (ICCV)*, 22-29 Oct. 2017, 5334-5343. doi.org/10.1109/ICCV.2017.569.
- Smith, N., Moehrle, N., Goesele, M., and Heidrich, W., 2018. Aerial path planning for urban scene reconstruction: a continuous optimization method and benchmark, *ACM Transactions on Graphics*, 37, Article 183. doi.org/10.1145/3272127.3275010.
- Sun, Z. and Zhang, Y., 2018. Using Drones and 3D Modeling to Survey Tibetan Architectural Heritage: A Case Study with the Multi-Door Stupa, *Sustainability*, 10, 2259. doi.org/10.3390/su10072259.
- Themistocleous, K., Ioannides, M., Agapiou, A., and Hadjimitsis, D. G., 2015. The methodology of documenting cultural heritage sites using photogrammetry, UAV, and 3D printing techniques: the case study of Asinou Church in Cyprus, *Third International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2015)*, 953510. doi.org/10.1117/12.2195626.
- Yan, F., Xia, E., Li, Z., and Zhou, Z., 2021. Sampling-Based Path Planning for High-Quality Aerial 3D Reconstruction of Urban Scenes, *Remote Sensing*, 13, 989. doi.org/10.3390/rs13050989.
- Zhang, H., Yao, Y., Xie, K., Fu, C.-W., Zhang, H., and Huang, H., 2021. Continuous aerial path planning for 3D urban scene reconstruction, *ACM Transactions on Graphics*, 40, Article 225. doi.org/10.1145/3478513.3480483.
- Zheng, X., Wang, F., and Li, Z., 2018. A multi-UAV cooperative route planning methodology for 3D fine-resolution building model reconstruction, *ISPRS Journal of Photogrammetry and Remote Sensing*, 146, 483-494. doi.org/10.1016/j.isprsjprs.2018.11.004.
- Zhou, X., Xie, K., Huang, K., Liu, Y., Zhou, Y., Gong, M., and Huang, H., 2020. Offsite aerial path planning for efficient urban scene reconstruction, *ACM Transactions on Graphics*, 39, Article 192. doi.org/10.1145/3414685.3417791.