MODEL-BASED MULTI-UA V PATH PLANNING FOR HIGH-QUALITY 3D RECONSTRUCTION OF BUILDINGS

S. Zhang1 *, W. Zhang1, C. Liu2

1 School of Geospatial Engineering and Science, Sun-YatSen University, Zhuhai, China - (zhangsh52.zhangwm25)@mail.sysu.edu.cn
2 College of Surveying and Geo-informatics, Tongji University, Shanghai, China - liuchun@tongji.edu.cn

KEY WORDS: Multi-UAV, Coarse model, Path planning, 3D reconstruction, Photogrammetric constraints.

ABSTRACT:

Unmanned aerial vehicle (UAV) photogrammetry is widely used for acquiring high-quality 3D models of urban areas. However, the completeness and quality of the reconstructed model can be affected by complex or concave structures when using common flight paths. Furthermore, flight paths with multi-altitude and multi-attitude waypoints can be time-consuming and may not be efficiently implemented by a single drone. To address these challenges, we propose a model-based multi-UAV path planning method for capturing images that enable complete and precise 3D reconstruction of buildings. Our method analyzes the geometry of the input coarse model and plans viewpoints based on the reconstructability related to completeness and precision. By solving a multiple travelling salesmen problem (mTSP), individual flight paths for each UAV are generated. We conducted a real-world experiment comparing the performance and efficiency of the proposed method with two existing solutions. Results of the proposed method produces a more complete 3D reconstruction with fewer images compared to the existing methods. It also shows that using two UAVs with the proposed method can significantly reduce the overall time required for 3D reconstruction.

1. INTRODUCTION

Unmanned aerial vehicle (UAV) photogrammetry is a primary method for acquiring high-quality 3D models of urban areas, which serve as spatial references for various urban applications and management tasks, such as construction site monitoring and building information modeling (BIM) (Anwar et al., 2020; Alshawabkeh et al., 2021). Drones equipped with gimbaled cameras offer high maneuverability, enabling them to capture images at close range and from various angles (Nex et al., 2022), resulting in the acquisition of ultra-high resolution 3D models (Cramer et al., 2018). However, when it comes to urban buildings, which often have complex or concave structures, commonly used UAV path planning methods (such as nadir, oblique, or orbit modes) may lead to incomplete or imprecise 3D reconstructions due to self-occlusion (Jiang et al., 2021).

In urban environments, the use of larger or fixed-wing UAVs, which can carry more powerful mapping sensors, is limited due to the complexity of the environment and regulations (Kandeel et al., 2021). The limited flying range of small drones restricts their mapping capabilities. Furthermore, these small mapping drones lack smart features such as real-time obstacle avoidance and can only follow pre-planned waypoints. Therefore, allocating the mapping task to multiple UAVs and planning individual waypoints for each UAV is a practical solution to improve the efficiency of data capture for high-quality 3D reconstruction of urban areas (Otto et al., 2018).

Effectively acquiring high-quality 3D models for complex buildings involves two main aspects: fine-grained viewpoint planning and multi-UAV path planning. In terms of viewpoint planning, the challenge of occlusion caused by complex build-

* Corresponding author

ings needs to be addressed first (Cabreira et al., 2019). Model-based UAV path planning methods analyze the geometry of the input coarse model (or proxy model), which can be obtained from previous exploration flights (Hepp et al., 2019) or existing models such as satellite reconstructions or BIM (Zhou et al., 2020; Tan et al., 2021). Subsequently, an optimal subset of viewpoints is selected from a candidate set that is uniformly and densely initialized in the navigable space (Roberts et al., 2017). The selection of viewpoints is often guided by a customized reconstructability index, which takes into account factors such as parallax angle, baseline, and ground sampling distance (GSD) in Structure from Motion-Multi-View Stereo (SfM-MVS) theory (Zhang et al., 2020; Peng and Isler, 2019). To address the trade-off between intersection angle and image matching quality, Smith et al. (2019) proposed a parallax weighing function (Wenzel et al., 2013). The problem of viewpoint selection involves non-linear constraints, making it NP-hard, while the objective function is often submodular (Koch et al., 2019). However, the computational complexity can be reduced while maintaining a certain level of global optimality (Krause and Gholvin, 2014). Li et al. (2022) formulated a max-min optimization problem to efficiently plan viewpoints in 3D space. Nevertheless, widely used image overlap-based reconstructability may lead to redundant viewpoint planning and still lacks the guarantee of complete reconstruction. In summary, the reconstructability index should be carefully designed to balance reconstruction quality and computational cost.

To increase the efficiency of image capture for high-quality 3D reconstruction of buildings, a practical solution is to utilize multi-UAV path planning. In conventional nadir image capture, a boustrophedon coverage flight path is commonly employed. The survey area is divided into sub-regions or sub-strips, with each UAV assigned to cover a specific sub-region (Avellar et al., 2015; Cabreira et al., 2019). However, it is important to note that viewpoints for high-quality 3D reconstruction are irregularly distributed. Therefore, the task allocation of multiple UAVs needs to be carried out at the waypoint planning stage.
level rather than the geometry level. This problem can be formulated as a multiple traveling salesman problem (mTSP) or a vehicle routing problem (VRP). The essence of the problem is that multiple drones start at the same point and need to reach multiple waypoints, with each waypoint being accessed by only one drone. The objective of multi-UAV path planning for high-quality 3D reconstruction is to allocate their paths in order to achieve the lowest total cost (Coutinho et al., 2018).

For multirotor drones, energy consumption is largely dependent on flight time and has little to do with the attitude of the drone (Liu et al., 2020). Therefore, the main objective of multi-UAV path planning for high-quality 3D reconstruction should be to minimize the total time taken. However, it is important to note that this problem is NP-hard, and commonly used methods such as genetic algorithms and simulated annealing algorithms can only achieve approximate optimization (Aggarwal et al., 2000). Recent advancements in linearization and mixed integer programming have further improved the optimality of the solutions (Stanek et al., 2019). However, for mapping missions, conflicts between UAVs and the buildings should also be considered and resolved.

In this paper, we propose a model-based multi-UAV path planning method for capturing images that enable complete and precise 3D reconstruction. The proposed method analyzes the geometry of the input coarse model and plans viewpoints based on the reconstructability related to completeness and precision. Individual flight paths for each UAV are generated by solving the multiple traveling salesman problem (Bektas, 2006). The workflow concept is illustrated in Figure 1.

The viewpoint \( C \) in this study has five degrees of freedom, consisting of three degrees of freedom for the position \( C^P = \{x, y, z\} \) and two degrees of freedom for the attitude \( C^A = \{\phi, \kappa\} \). The roll angle \( \omega \) of the camera is assumed to be zero due to onboard stabilization. The objective of the viewpoint selection is to choose the minimum number of viewpoints \( C^*_c \) from the candidate set \( C_0 \) to fully cover the target surface by planning two viewpoints to observe each sampling point. The viewpoint set \( C^*_c \) should satisfy the objective function and constraints:

\[
C^*_c = \text{arg min} \ |C_e| \quad \text{s.t.} \quad f(C^*_c) \geq q
\]  

(1)

In Equation 1, \( |C_e| \) is the number of viewpoints. \( f(C^*_c) = \sum_{i=1}^{n} R(C_i) \) is the overall reconstructability, and \( q \) is the target reconstruction percentage. Considering that there may be highly occluded parts that cannot be observed by the UAV, \( q \) can be reduced by a factor of 0.9. \( R(C_i) \) is the reconstructability of a viewpoint \( C_i \):

\[
R(C_i) = \arg \max_{s} \frac{s \omega}{s} + \sum_{j} \sin \psi_j
\]  

(2)

2. METHOD

The workflow of the proposed method is illustrated in Figure 2. This method begins by taking a coarse mesh model of the target as input, which can be obtained from BIM (building information modeling) or satellite reconstruction. Candidate viewpoints are then densely initialized in the normal direction of each facet of the target. The distance is calculated using the sensor size, focal length, and the required GSD. Subsequently, the reconstructability of each viewpoint is evaluated by considering factors such as the GSD, baseline, intersection angle, and the estimated error ellipsoid (Zhang et al., 2020). Viewpoints are selected from the candidate viewpoint set to ensure complete coverage of the target, resulting in a strengthened camera network. Next, the error ellipsoid of all target points is estimated through photogrammetric forward intersection. Targets with an error ellipsoid larger than the given threshold are identified as low-precision. Additional viewpoints are planned to optimize the viewing geometry of these targets, thereby increasing the reconstruction precision. The viewpoints planned in the previous two steps are combined to form the optimized camera network. Finally, the viewpoints are divided and connected to form the flight path for each individual UAV. The details of the workflow are explained as follows.

![Figure 1. Concept of the proposed method.](image)

![Figure 2. The proposed workflow.](image)
The reconstructability function (Equation 2) considers the ratio of the observed sampling points \( s \) from viewpoint \( C \) to all the points \( s \), and \( \psi \) is the angle between the light ray direction and the principal direction of the error ellipsoid of point \( s \). Other reconstructability factors are transformed into constraints for the optimization problem. The geometry for this viewpoint selection step is shown in Figure 3. The viewpoint candidate (grey) is planned alongside the target point normal \( n \), and its distance to the target \( d \) depends on the required GSD. \( m \) is the light ray direction. \( \theta \) is the angle between \( m \) and \( n \), which is set to 40° according to (Hoppe et al., 2012). \( N \) is the principal direction of the error ellipsoid, and \( \psi \) is the angle between \( m \) and \( N \).

This step aims to select additional viewpoints based on the pre-estimated camera network \( C \). The new viewpoints planned in this step, \( C_p \), are finally combined with the viewpoints for completeness, \( C_r \), to form the final optimized camera network for UAV path generation \( C = C_r \cup C_p \).

### 2.3 Path Generation for Multiple UAVs

This step involves planning flight paths for \( n \) UAVs using the camera network \( C_r \) obtained from the previous steps. We assume that the UAVs are multirotor drones equipped with gimbal cameras, allowing them to fly in 3D space and capture images from various attitudes by yawing and rotating the gimbal.

To begin, we construct a graph \( G(V, E) \) where each viewpoint \( C_r \) in the set \( C_r \) corresponds to a vertex \( V \). The edges \( E \) represent the cost between vertices. Assuming there are \( n \) UAVs required to traverse all viewpoints in \( C_r \), the takeoff positions for each UAV are denoted as \( S = \{S_1, S_2, ..., S_n\} \), where \( \forall S_i \not\in C_r \). The path for the \( i \)th UAV is denoted as \( P_i = \{S_1, E_1, ..., E_k, S_i\} \), indicating that the UAV takes off from \( S_i \), follows the edges \( E_k \), and returns to the takeoff position. We define the set of break-points \( B = \{B_1, B_2, ..., B_{n-1}\} \) to divide the path \( P \) into \( P_i \) segments.

The objective of using multiple UAVs for the flight path is to save time. For multirotor UAVs, the flight time is primarily determined by the length of the path (Liu et al., 2020). Therefore, we set the cost \( E \) as the Euclidean distance between vertices, and the objective function is to minimize the overall length:

\[
\min \sum_{i=1}^{n} L(P_i) \tag{5}
\]

where \( L(P_i) \) represents the length of the \( i \)th path. Assuming that the maximum flying range of each drone is \( d_m \), we introduce a penalty coefficient \( k_d \) to adjust the distance correlation function beyond the range path. This adjustment effectively modifies the value of the objective function to increase the constraint:

\[
L(P_i) = \begin{cases} 
L(P_i) & L(P_i) \leq d_m \\
 k_d(L(P_i) - d_m) + d_m & L(P_i) > d_m 
\end{cases}
\]

In Equation 6, if the path for a certain UAV exceeds the range, the cost of its excess part increases by a factor of \( k_d \). If an

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLVIII-1-W2-2023-1923-2023 | © Author(s) 2023. CC BY 4.0 License.
obstacle is present between two vertices, the distance is adjusted using an obstacle-avoidance trajectory, which can be planned using an algorithm such as A*. Finally, a multiple traveling salesmen problem is formulated and solved on the graph G to obtain the flight path for each UAV.

3. EXPERIMENT AND RESULTS

A real-world experiment was conducted to evaluate the performance and efficiency of the proposed method. The experiment compared the proposed method with the Nap-of-the-object photogrammetry planning (NFP) method (Wang et al., 2022) and Agisoft Metashape’s method (MS) (Agisoft LLC, 2022). The target building was a multipart Chinese-style building with dimensions of approximately 155 × 155 m. A lower-resolution model reconstructed from nadir UAV images was used to simulate the input satellite model. Two DJI Phantom 4 RTK drones were used for image capture, with an average capture distance of 30 m and a GSD of approximately 1.5 cm. The overlap settings for NFP and Metashape were set to 65%. The sidelap for NFP was set to 50%, while Metashape does not provide such a setting. Figure 5 shows the viewpoints planned by each comparing method. Agisoft Metashape was used for model reconstruction.

To evaluate the absolute accuracy, 20 reference points were signalized with the real-time kinematic (RTK) system, and 5 of them were ground control points (GCPs) (Figure 6). The ground truth model was reconstructed from oblique imaging and georeferenced with all reference points, ensuring high completeness and accuracy.

Table 1 presents the properties of the flight paths. The proposed method planned only 62 images, while NFP planned 169 images and MS had 192 images, resulting in the longest flight path. The flight time was significantly lower for the proposed method due to the fewer viewpoints. The “Proposed-DUAL” paths were implemented by individual drones (No. 1 and 2), which saved approximately 160 seconds compared to a single drone.

Table 1. Flight path comparison.

<table>
<thead>
<tr>
<th>Name</th>
<th># of images</th>
<th>Length(m)</th>
<th>Flight time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFP</td>
<td>169</td>
<td>2300</td>
<td>1667</td>
</tr>
<tr>
<td>MS</td>
<td>192</td>
<td>2814</td>
<td>1874</td>
</tr>
<tr>
<td>Proposed</td>
<td>62</td>
<td>1746</td>
<td>549</td>
</tr>
<tr>
<td>Proposed-DUAL-1</td>
<td>28</td>
<td>1674</td>
<td>354</td>
</tr>
<tr>
<td>Proposed-DUAL-2</td>
<td>34</td>
<td>1575</td>
<td>384</td>
</tr>
</tbody>
</table>

The visual comparison of the dense image matching (DIM) point cloud is shown in Figure 7. The reconstruction from NFP’s method has many defects on the green roof and red façade, while the proposed method produces a more complete result with only some defects on the roof. The MS path with 192 images appears to be the most complete one. The strip-like viewpoint distribution of NFP’s method may not be as good as the even distribution achieved by Metashape’s method. The Proposed-DUAL method had exactly the same viewpoints as the Proposed method, but it was implemented with two drones working collaboratively. There was one more set of interiors to solve, and the reconstruction looks similar to that from a single drone.

The DIM point cloud from each flight path was compared with the reference model, and the quantitative results are demonstrated using three indicators: precision, recall, and F1-score. Precision P quantifies the number of reconstructed points that are close to the ground truth points, while recall R quantifies the number of ground truth points that are close to a reconstructed point, which can also be interpreted as the completeness and an indicator of the overall absolute accuracy. The F1-score, $F = \frac{2PR}{P+R}$, can be used as a performance measure for 3D reconstruction (Hepp et al., 2019). Table 2 presents the results. The quantitative results align with the visual comparison. Among all three flight paths, the proposed method achieves the best results for all three indicators. The NFP method has a higher precision but lower recall compared to the MS method, despite using fewer images. The strip-like viewpoints of the NFP method may assist in accurate image matching and exterior estimation, resulting in better precision. Compared to the single drone method, the results from the two drones method (Proposed-DUAL) are slightly lower, possibly due to the introduction of another set of camera interior.

Table 2. Quantitative comparison of DIM point cloud.

<table>
<thead>
<tr>
<th>Name</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFP</td>
<td>36.26</td>
<td>26.91</td>
<td>0.309</td>
</tr>
<tr>
<td>MS</td>
<td>34.94</td>
<td>28.42</td>
<td>0.313</td>
</tr>
<tr>
<td>Proposed</td>
<td><strong>41.32</strong></td>
<td><strong>30.30</strong></td>
<td><strong>0.350</strong></td>
</tr>
<tr>
<td>Proposed-DUAL</td>
<td>40.98</td>
<td>29.90</td>
<td>0.346</td>
</tr>
</tbody>
</table>

Regarding the absolute accuracy, we compared the errors in the XY, Z, and total dimensions for the checkpoints. The results are presented in Table 3. It is worth noting that the XY errors obtained from all flight paths are relatively large, considering the centimeter-level GSD. This might be attributed to systematic errors between the reference points and the initial exteriors obtained from the drone. All the compared methods yielded similar results, with only a few centimeters of difference. However, the proposed flight path implemented with two drones exhibited slightly better performance in terms of XY and total error.

Table 3. Absolute error comparison via checkpoints (cm)

<table>
<thead>
<tr>
<th>Name</th>
<th>XY error</th>
<th>Z error</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>NFP</td>
<td>48.35</td>
<td>8.98</td>
<td>49.18</td>
</tr>
<tr>
<td>MS</td>
<td>40.95</td>
<td>5.40</td>
<td>41.30</td>
</tr>
<tr>
<td>3D</td>
<td>42.99</td>
<td>6.23</td>
<td>43.44</td>
</tr>
<tr>
<td>3D-DUAL</td>
<td><strong>40.38</strong></td>
<td>5.69</td>
<td><strong>40.78</strong></td>
</tr>
</tbody>
</table>

4. CONCLUSIONS

In this paper, we proposed a model-based multi-UAV path planning method for capturing high-quality 3D reconstructions of buildings. The proposed method utilizes a coarse model as input and incorporates photogrammetric constraints and error ellipsoids to guide viewpoint selection. The flight paths for multiple UAVs are then planned by solving a multiple traveling salesman problem. The results demonstrate that the proposed method achieves a more comprehensive 3D reconstruction using fewer images compared to existing methods such as NFP and Metashape. Additionally, the efficiency of image capture is increased through the planning of paths for multiple UAVs.

However, there are still some limitations that need to be addressed. Firstly, the proposed method assumes that the input coarse model is error-free. In practice, the coarse model may contain errors up to several meters, which can impact the quality and safety of the reconstruction. Therefore, future work will consider the uncertainty associated with the input model.
Secondly, the current multi-UAV path generation method only considers obstacle avoidance. However, there may be potential conflicts between UAVs during operation, necessitating mission planning that avoids collisions between UAVs. To address this, future work will introduce additional dynamic constraints to ensure safety.

5. ACKNOWLEDGEMENTS

The research in this paper was supported by the Fundamental Research Funds for the Central Universities, Sun Yat-sen University (Grant No. 23ptpy98).

References


Spruyt, V., 2015. How to draw an error ellipse representing the covariance matrix.


