GCP Deployment and Recognition System based on Light-Marker UAV

Ruikai Li [1], Yi Wan [1], Dong Wei [1], Yongjun Zhang [1], Ziqian Huang [1], Xiaotong Li [1], Changjun Zhu[1]

[1] WHU, School of Remote Sensing and Information Engineering, 430079, Wuhan, Hubei, China – (liruikai, yi.wan, weidong, zhangyj, huangziqian, lixiaotong, whursgis-zcj)@whu.edu.cn

Keywords: UAV Photogrammetry, Ground Control Points, Light-Marker UAV, Target Encoding, Target Decoding, Graph Matching.

Abstract

This paper addresses the heavy reliance on manual operations in control point acquisition for UAV photogrammetry and proposes an encoded control point deployment and recognition method based on a Light-Marker UAV (LMUAV). Conventional approaches rely on manual placement of control points and manual identification and measurement in images for aerial triangulation, resulting in low efficiency. To address this limitation, an LMUAV equipped with an LED array actively broadcasts its positional information as quaternary optical signals. The observing UAV performs coarse localization of the target region by integrating communication priors with the imaging model, followed by light spot segmentation and graph construction within the region of interest (ROI). Node correspondences are then recovered by constructing a template graph and an observation graph and applying Reweighted Random Walks (RRWM) graph matching. The matching robustness is further enhanced by incorporating directional point constraints and RANSAC-based geometric filtering. Based on the recovered correspondences, the encoded information is decoded through color recognition and validation, enabling automatic control point recovery. Experimental results in a cross-flight-line scenario with a single target UAV demonstrate that the proposed method achieves stable node matching and encoding–decoding, with a sequence-level accuracy of 76.32%, and a final effective decoding rate of 71.05%, while maintaining centimeter-level positioning accuracy, thereby validating its effectiveness for automatic control point acquisition in UAV mapping.

1. Introduction

With the rapid development of UAV technology, UAV photogrammetry has become an important means for acquiring high-resolution spatial information and has been widely applied in terrain mapping, urban reconstruction, disaster monitoring, and geospatial data updating (Colomina and Molina, 2014; Berra et al., 2020; Kovanič et al., 2023). In typical photogrammetric workflows, Ground Control Points (GCPs) are used to establish geometric constraints between images and real-world coordinates, thereby improving the accuracy of aerial triangulation and 3D reconstruction (Wolf et al., 2014; Luhmann et al., 2019; Sanz-Ablanedo et al., 2018).However, conventional GCPs require manual deployment and field measurement, which are time-consuming and costly in large-scale or complex environments. In scenarios such as mountainous regions, disaster sites, or offshore operations, manual deployment is not only inefficient but may also pose safety risks. Therefore, reducing reliance on manually deployed control points and improving automation in UAV mapping has become an important research direction.

To address these challenges, existing studies have explored alternatives to reduce or replace GCPs. One approach employs high-precision Global Navigation Satellite System (GNSS) techniques for direct georeferencing (Zeybek and Şanlıoğlu, 2023; Han et al., 2024; Atik and Arkalı, 2025). Another focuses on optimizing the number and spatial distribution of control points (Harwin and Lucieer, 2012; Atik and Arkalı, 2025). Additionally, virtual control points or multi-sensor data (e.g., stereo vision and LiDAR) have been introduced to further reduce field workload (Nex and Remondino, 2014; Kregar et al., 2025). However, in complex environments or under high-precision requirements, these approaches still cannot fully eliminate the need for manually deployed control points. On this basis, visual-based methods have been further explored for automatic detection and identification of control points. A common approach deploys artificial targets with distinctive visual features, such as checkerboard patterns, circular targets, or coded markers, enabling automatic localization in images (Luhmann et al., 2019). In recent years, encoded visual markers such as AprilTag and ArUco markers (Olson, 2011; Garrido-Jurado et al., 2014) have been introduced to improve automation. By embedding identification codes into visual patterns, these markers enable both localization and identity recognition of control points, thereby facilitating automatic correspondence establishment. Although these methods improve recognition efficiency, they still rely on manual ground deployment.

To improve target observability under long-range conditions, active visual targets such as luminous devices, blinking markers, or optical beacons have been introduced (Li et al., 2025). Compared with passive markers, active light sources exhibit higher contrast and visibility in complex backgrounds or low-light environments, and have been applied in UAV localization, robotic navigation, and cooperative systems. However, existing studies mainly focus on improving detection and recognition, while the issue of manual control point deployment remains largely unresolved. Therefore, achieving an integrated solution for both deployment and recognition is still a key challenge, especially under complex observation conditions.

To address this problem, this paper extends control points from the ground to the air by employing a flying platform with known positions as a dynamic reference. In this framework, a LMUAV serves as an active optical node, broadcasting its identity and telemetry through an encoded visual target that is autonomously captured by an observing UAV. Compared with traditional GCPs, such aerial control points eliminate manual deployment and can be autonomously positioned, significantly improving the flexibility and efficiency of the entire workflow. Based on this idea, this paper proposes an aerial encoded control point deployment and recognition method based on a LMUAV. A fixed-geometry LED array is mounted on the LMUAV, where positional information is encoded into color sequences using quaternary optical signals. A communication link between the LMUAV and the observing UAV provides spatial priors to estimate the LMUAV location in the image and extract the

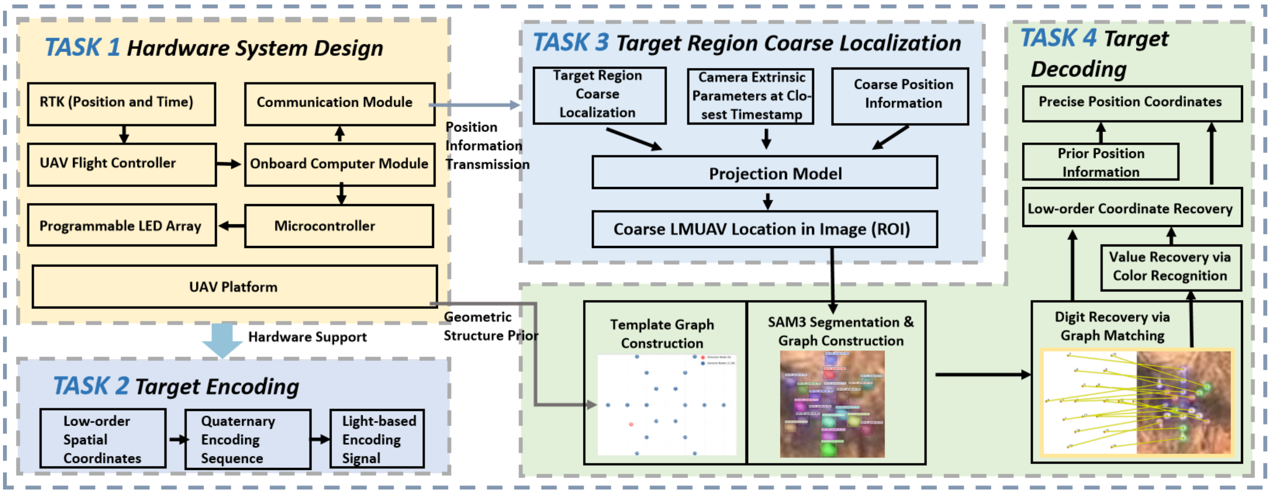

Figure 1. Overall workflow of the LMUAV-based aerial control point system. The proposed system comprises four main stages: hardware system design, target encoding, target region coarse localization, and target decoding. It forms a closed-loop pipeline from position information encoding to image-based decoding.

corresponding ROI. For control point decoding, a graph-theoretic representation models the geometric topology of light points, with node mappings established through graph matching. The encoded sequence is then decoded through color recognition, and the precise position of the LMUAV is determined by integrating communication information. The main contributions of this paper are summarized as follows:

1. An aerial encoded control point method for UAV cooperative mapping is proposed.
2. A quaternary optical encoding scheme is designed to represent positional information.
3. A graph matching-based decoding method with geometric priors is proposed to recover node correspondences from structural relationships, followed by color-based decoding.

Experiments were conducted in urban, wilderness, and hilly terrains, each covering approximately 1 km². A LMUAV was used to deploy 23 control points in a cross-flight-line manner. The LMUAV sequentially maneuvered between control points, while the observing UAV remained stationary and performed observations after each deployment, enabling asynchronous coordination (Zhang et al., 2022). Experimental results demonstrate that the proposed method achieves reliable decoding performance and centimeter-level positioning accuracy, while significantly improving efficiency compared to traditional manual approaches.

2. Method

As shown in Fig 1,.the framework consists of four main components: hardware system design, Target Encoding Based on Quaternary Light Indicators, coarse localization of the target region, and Geometric prior-guided graph matching for decoding. These components jointly accomplish the complete process from positional information generation, spatial constraint establishment, to visual decoding.

2.1 Hardware System Design

To enable in-air position broadcasting and visual recognition of the target UAV, this paper designs the LMUAV hardware system and establishes a wireless communication link based on a digital video transmission system. The proposed system converts high-precision positioning information of the UAV into observable optical encoded targets, providing effective prior information for subsequent visual detection and decoding.

In terms of hardware architecture, the LMUAV integrates several key components, including a flight control and positioning module, an onboard computing module, a light encoding module, and a communication module. The flight control system ensures stable flight and attitude control, while the positioning module provides real-time three-dimensional positional data. As shown in Fig. 2(b), a high-precision GNSS module with Real-Time Kinematic (RTK) capabilities is employed to provide continuous and robust positioning information, which serves as the input for the encoding process. As shown in Fig. 2(a), the onboard computing unit integrates the onboard computing module, facing vision sensor, and communication module. The onboard computing module is responsible for acquiring and processing positioning data, and generating the corresponding LED encoding sequence according to predefined encoding rules. The encoded information is transmitted to the microcontroller via serial communication, where it is used to control the LED array. A downward-facing vision sensor is employed to evaluate ground surface conditions and assess landing feasibility. As shown in Fig. 2(c), based on the Arduino Nano platform, the microcontroller performs timing control and driving of the LED array, converting the encoded sequence into specific RGB light patterns. The light encoding module consists of multiple LED array units (each comprising WS2812 RGB LEDs) arranged in a regular geometric configuration. Each LED pixel can be independently controlled in color, enabling information encoding through combinations of different color states (Worldsemi, 2020). With the millisecond-level refresh of WS2812 LED units, the LMUAV can broadcast its spatial position in real time during the flight. To enhance visibility under long-range observation conditions, the spacing between LED groups is carefully designed to ensure sufficient distinguishability at distances on the order of 100 meters. For communication, a digital video transmission system is adopted to establish a wireless link between the LMUAV and the observing UAV. This link is used to transmit positional and status information of the LMUAV, allowing the observing UAV to obtain a coarse estimate of the target location, thereby reducing the visual search space and improving detection efficiency. As shown in Fig 3, the target UAV is captured in side view and top view.

The overall system operates as follows: the onboard computing module continuously receives positioning data from the flight controller and converts it into encoded light sequences. The encoded data are then transmitted to the microcontroller via a serial interface, which drives the LED array to display the corresponding color patterns. Meanwhile, the positioning data are simultaneously transmitted to the observing UAV through the communication link to assist in coarse localization of the target region. Through this process, the LMUAV continuously broadcasts visible light encoded signals in the air, enabling intuitive representation and transmission of its positional information.

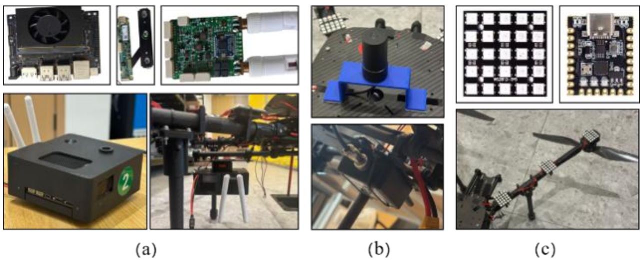

Figure 2. (a) Onboard Computing Unit; (b) GNSS module; (c) Composition and Deployment of the LED Encoding Module

Figure 3. Physical photo of the LMUAV (Side view and Top view)

2.2 Target Encoding Based on Quaternary Light Indicators

To enable the target UAV to actively broadcast its positional information in the air, a quaternary light encoding method based on an RGB LED array is proposed. This method quantizes the spatial coordinates of the target UAV into a discrete data sequence and maps it to corresponding light color patterns. Considering that the observing UAV and the target UAV operate within a relatively limited spatial range, it is unnecessary to encode the full absolute coordinates. Instead, only the low-order components of the coordinates (the residues after modulo operation) are encoded, which effectively reduces the required encoding length.

At a given time, the position of the LMUAV in the WGS84 coordinate system can be expressed as (λ, ϕ, h). To facilitate digital encoding, the longitude, latitude, and altitude are first discretized according to predefined resolutions, resulting in integer representations L_λ, L_ϕ, L_h, which denote the quantized integer values of longitude, latitude, and altitude, respectively. The number of encoding digits for longitude, latitude, and altitude are set to $n_\lambda = 5$, $n_\phi = 6$, and $n_h = 7$. Accordingly, the maximum representable range of each component is given by: $N_\lambda = 4^{n_\lambda}, N_\phi = 4^{n_\phi}, N_h = 4^{n_h}$ .Each integer coordinate component can then be decomposed into an integer quotient and a low-order remainder as follows:

$$\begin{cases} L_\lambda = q_\lambda N_\lambda + L'_\lambda \\ L_\phi = q_\phi N_\phi + L'_\phi \\ L_h = q_h N_h + L'_h \end{cases} \tag{1}$$
$$0 \le L'_\lambda < N_\lambda, 0 \le L'_\phi < N_\phi, 0 \le L'_h < N_h$$

where L'_λ , L'_ϕ , L'_h = low-order encoded components (remainders after modulo operation) of longitude, latitude and height, respectively.
q_λ, q_ϕ, q_h = the corresponding quotient terms.

Since only the low-order components are retained during encoding while the high-order information is not directly transmitted, the encoding scheme exhibits a periodic property. When the actual geographic coordinates exceed the representable range of a single encoding cycle, different spatial positions may correspond to identical encoded sequences. To resolve the ambiguity caused by this periodicity, coarse position information obtained through communication is introduced during decoding, enabling the unique recovery of the true coordinates.

The coordinate encoding process further requires converting the integer coordinate values along the three spatial dimensions into fixed-length quaternary strings. Let the quaternary conversion function be defined as $Q(v, \mathrm{n})$, where v is a non-negative decimal integer and n represents the output string length. If the converted string contains fewer than n digits, zero-padding is applied to the left to satisfy the length requirement. The encoded sequences for the three dimensions are then determined and concatenated into a complete sequence , as follows:

$$\begin{cases} S_\lambda = Q(L'_\lambda, 5) \\ S_\varphi = Q(L'_\varphi, 6) \\ S_h = Q(L'_h, 7) \end{cases} \tag{2}$$
$$D = S_\lambda \parallel S_\varphi \parallel S_h, \quad d_i \in \{0,1,2,3\}, i = 1,2,\dots,18 \tag{3}$$

where S_λ , S_φ , S_h = quaternary sequences of longitude, latitude, and height
D = concatenated encoding sequence
d_i = i-th element of the sequence D

In the light mapping stage, each quaternary symbol is assigned to a specific color state of an LED. The mapping is defined as follows: *0→white,1→red,2→green,3→blue.* Each element of the 18-digit data sequence is mapped to the color state of a corresponding encoding LED, forming the resulting light-based encoded pattern.

Figure 4. The directional indicator Light

Directional LED and Validation Mechanism, in addition to the 18 encoding LEDs, an extra LED is introduced as a directional indicator, which is positioned between two adjacent arms of the UAV, as shown in Fig. 4. This LED serves a dual purpose. On one hand, it provides a reference for determining the orientation of the LED array, enabling recovery of the starting position and ordering of the encoded sequence in the image. On the other hand, its color simultaneously functions as a validation bit for the entire data sequence, allowing verification of the decoding result. Let the validation value corresponding to the directional LED be denoted as p. In this work, a modulo-4 checksum is adopted, defined as:

$$p = \left(\sum_{i=1}^{18} d_i \right) mod 4 \quad (4)$$

After applying the same color mapping rule as used in the encoded light array for visualization, the observing UAV can recover the observed check value p_{obs} and the 18-bit decoded sequence d_i^{dec}. The check value p_{dec} is recomputed from the decoded sequence d_i^{dec}. A decoding result is regarded as valid when p_{dec} equals p_{obs} ; otherwise, it indicates potential recognition errors or sequence misordering, and the result should be rejected or re-evaluated. Through this design, the directional node simultaneously provides orientation reference and check information, enabling consistency verification of the decoded sequence and enhancing the robustness of the decoding process.

2.3 Coarse Localization of the Target Region

To reduce the search space for subsequent visual recognition, the proposed method utilizes the positional information transmitted by the LMUAV through the communication link to perform coarse localization of the target in the image. The LMUAV transmits its positioning data along with timestamps through the communication link, while the observing UAV records the image acquisition time during capture. Since both timestamps are derived from the RTK module, a unified time reference is established. The positioning data corresponding to the current image is obtained by selecting the record with the minimum temporal difference. Given that the positioning information of the target UAV is provided in the WGS84 coordinate system(λ, ϕ, h), it is first converted into a three-dimensional Cartesian representation in a local world coordinate system, denoted as $P = (X_t, Y_t, Z_t)$ .Using the camera intrinsic and extrinsic parameters, the 3D position of the LMUAV can be projected onto the image plane through the standard perspective projection model. However, projecting a single point is insufficient to represent the spatial extent of the target in the image. Therefore, considering the physical size of the LED array, a local region centered at P is constructed and projected onto the image plane to obtain the corresponding ROI.

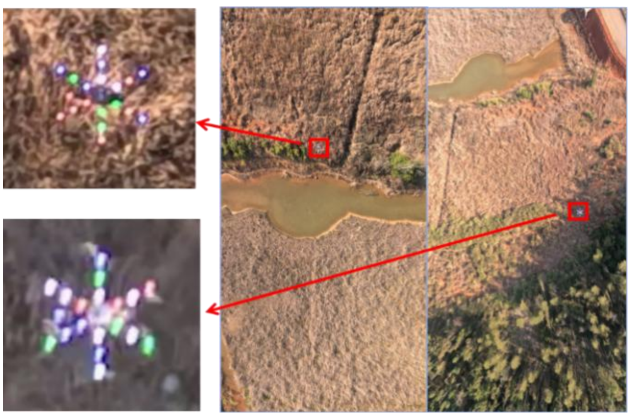

Figure 5. Extract the ROI from the full image.

2.4 Geometric prior-guided graph matching for decoding

To decode the LMUAV information from observation images, we propose a graph matching-based method incorporating geometric priors. The multi-colored LED array presents two primary challenges: (1) topological correspondence, mapping detected points to the template array to determine their positional order (bit significance); and (2) color identification, extracting encoded values from each LED. Specifically, a template graph and an observation graph are constructed to establish point correspondences. By integrating these matching results with color recognition, the complete encoded sequence is reconstructed.

2.4.1 Graph Construction: The graph construction aims to represent the LED array as a structured graph, such that the rigid geometric relationships among LED points can be recovered through graph matching. In this work, two types of graphs are constructed: the template graph and the observation graph.

The template graph is derived from the physical design of the LMUAV. Since the LED positions on the UAV are fixed, the template graph is constructed based on the geometric layout of the LED array, with coordinates normalized according to the UAV dimensions. Each node corresponds to an LED position, and the distance between adjacent nodes on the same arm is set to 1. A visualization of the template graph is shown in Fig. 6.

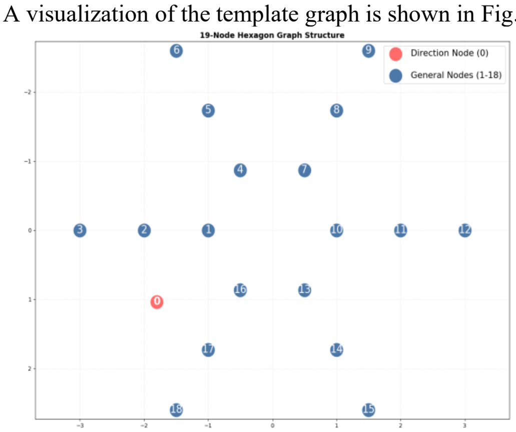

Figure 6. The visualization of the template graph

The observation graph is constructed based on detected LED centroids. Specifically, the Segment Anything Model3 (SAM3, Carion et al., 2025) is employed for instance segmentation within the ROI to extract luminous LED regions. As a zero-shot segmentation foundation model, SAM3 identifies LED instance masks without task-specific fine-tuning, demonstrating high robustness in complex lighting environments. Given that LEDs manifest as high-intensity regions with distinct boundaries, SAM precisely delineates their morphology. The pixel centroid of each segmented region is then computed as a graph node to form the observation graph, as illustrated in Fig. 7.

https://doi.org/10.5194/isprs-archives-XLIX-B2-2026-1167-2026 | © Author(s) 2026. CC BY 4.0 License.

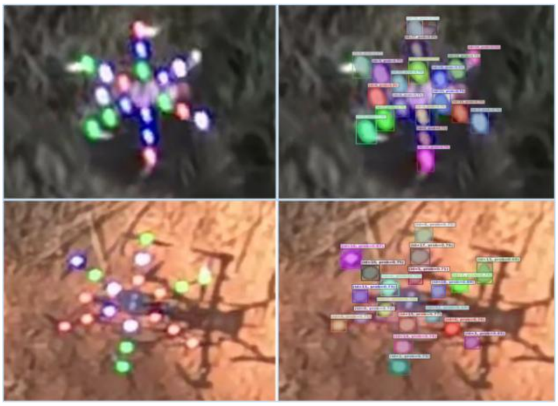

Figure 7. LEDs segmentation results based on SAM3

The node sets of the template and observation graphs are defined as:

$$\begin{cases} V^T = \{ v_i^T \mid i = 0, \ldots, 18 \} \\ V^O = \{ v_j^O \mid j = 1, \ldots, m \} \end{cases} \quad (5)$$

where $m =$ the number of nodes in the observation graph.

2.4.2 Graph Matching with Geometric Priors : After constructing the template and observation graphs, the goal is to establish node correspondences to recover the spatial positions of the LED array. Due to noise, missing detections, and geometric variations, direct matching is challenging. To address this, a graph matching framework with structural constraints is adopted. A fully connected set of candidate correspondences is first constructed by pairing each node in the template graph with all nodes in the observation graph. To eliminate scale differences, node coordinates are normalized using the median of pairwise distances between centroids. Conflict constraints are introduced to enforce one-to-one matching and prevent multiple assignments to the same node. Based on these constraints, an affinity matrix is constructed to measure structural consistency between candidate correspondences. The graph matching problem is then formulated as an optimization over the affinity matrix and solved using the RRWM algorithm (Cho et al., 2010). To further improve robustness, a geometric consistency filtering step is applied using RANSAC. A homography transformation is estimated to model the relationship between the template and observation graphs, and outliers are removed based on reprojection error.

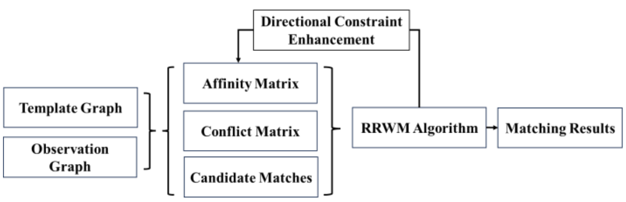

Figure 8. The two-stage graph matching strategy

Due to the axial symmetry of the LED array, a single graph matching process may produce multiple equivalent solutions. To resolve this ambiguity, a two-stage graph matching strategy is proposed. In the first stage, an initial graph matching is performed to identify the direction point. Due to its asymmetric spatial configuration, the direction point exhibits distinctive geometric relationships that enable it to be reliably identified during graph matching. In the second stage, an angular constraint is introduced based on the direction point. Specifically, the line connecting the array center and the direction point is taken as the reference direction, and the polar angles of all nodes are computed relative to this direction in both graphs. The relative angular differences between corresponding nodes are then evaluated for each candidate match, and only those satisfying a predefined angular consistency threshold are retained. Candidate matches that violate this constraint are discarded or penalized during affinity construction, thereby reducing ambiguous correspondences. Based on the filtered candidate set, the graph matching process is performed again to obtain the final node correspondences. The angular relationships are defined as follows:

$$\theta_k^G = \arctan 2\left(y_k^G - y_c^G,\, x_k^G - x_c^G\right) \quad (6)$$

$$\phi_k^G = \theta_k^G - \theta_d^G \quad (7)$$

where $G \in \{T, O\}$ = template graph and observation graph
k = index of a point, where k = d denotes the direction point and k = i, j denote node indices
(x_c^G, y_c^G)= center of graph G
$\left(x_k^G, y_k^G\right)$ = the coordinates of node k in graph G
${\theta_d}^G$= polar angle of the direction point with respect to the graph center
${\theta_k}^G$ = polar angle of a node with respect to the graph center
ϕ_k^G = relative angle with respect to the direction reference

For a candidate matching pair $\left(v_i^T, v_j^O\right)$, the pair is retained if the difference between their relative angles with respect to the reference directions satisfies:

$$\mid \phi_i^T - \phi_j^O \mid < \tau \quad (8)$$

Candidate correspondences satisfying the angular consistency constraint are retained, while inconsistent ones are penalized in the affinity matrix construction. This suppresses ambiguous matches during the subsequent RRWM optimization. After filtering, RRWM is performed again to obtain the final correspondences. The proposed two-stage strategy effectively resolves symmetry-induced ambiguity while preserving global structural consistency, resulting in more stable and reliable matching results. The recovered correspondences further determine the topological ordering of LED nodes, enabling subsequent decoding of the encoded sequence.

2.4.3 Color-based Decoding: To improve the robustness of LED color recognition under complex imaging conditions, a clustering-based color classification method is adopted. Considering that the central regions of LED blobs are prone to overexposure, pixel regions corresponding to each segmented LED are first extracted. A ring mask is then constructed via successive erosion operations, and color statistics are preferentially computed from edge pixels. When insufficient pixels are available, the full mask region is used as a fallback. This strategy effectively reduces the influence of saturated central pixels. For feature representation, multiple color descriptors are extracted for each LED region, including mean and median values in the RGB space, normalized chromaticity components, and HSV-based features such as average saturation and value. This multi-dimensional representation captures absolute color intensity, relative channel proportions, and illumination characteristics. Based on these features, all detected LED points within a single frame are clustered into four categories using the K-Means algorithm. Since only four colors (white, red, green, and blue) are used, clustering leverages relative color differences within the same frame, eliminating the need for predefined thresholds. The cluster indices are then

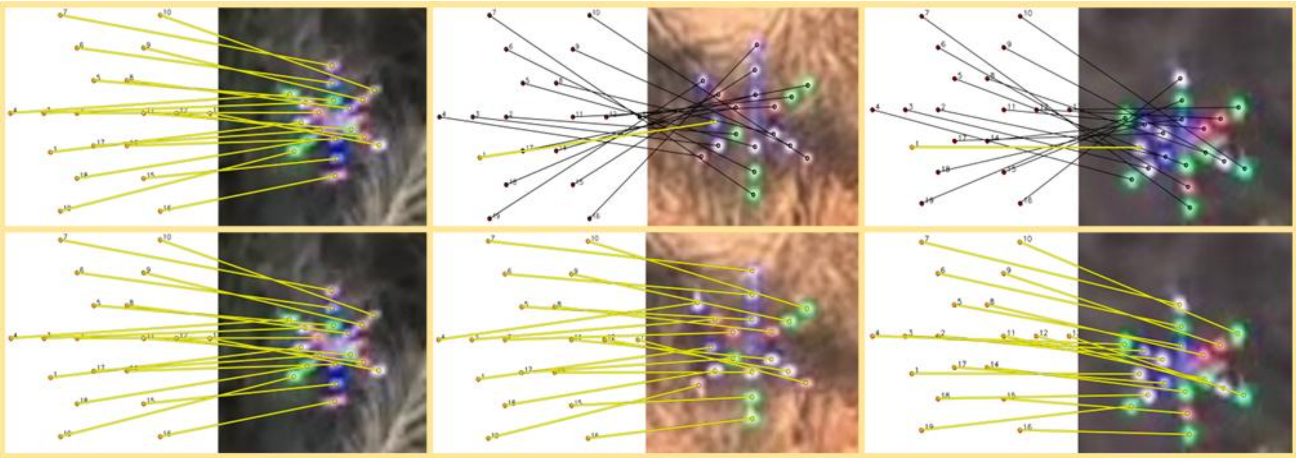

Figure 9. Top: RRWM results. Bottom: Two-stage RRWM results. The proposed method reduces ambiguity caused by structural symmetry via direction-point-based angular constraints. The recovered geographic coordinates (longitude, latitude, and height) are $29.793299°/113.901801°/39.14\,\mathrm{m}$, $29.793383°/113.903488°/47.69\,\mathrm{m}$, and $29.793238°/113.900360°/106.74\,\mathrm{m}$, which are consistent with the ground truth.

mapped to physical color labels based on the characteristics of cluster centers. Specifically, the cluster with high brightness and low saturation is identified as white, while the remaining clusters are assigned to red, green, and blue according to their dominant RGB channels.

By exploiting the global consistency of LED colors within a single frame, the proposed method avoids the sensitivity of threshold-based approaches to illumination variation and noise, providing stable and reliable color recognition for subsequent decoding.

3. Experiment

3.1 Experimental Test

In the experiments, a LMUAV is used for sequential control point deployment by hovering at predefined locations within the target area. Control points are arranged in the overlapping regions of adjacent flight strips to enable multi-view observation and provide a unified encoding reference. The observation UAV performs coordinated image acquisition once each control point reaches a stable state. Experiments are conducted in urban, wilderness, and hilly terrains (~1 km² each), where 23 control points and 45 images containing encoded LED structures are collected, with ground truth encoding sequences recorded for evaluation. As shown in Fig. 10, the observation UAV follows the flight trajectory, while the target UAV moves sequentially according to the control points.

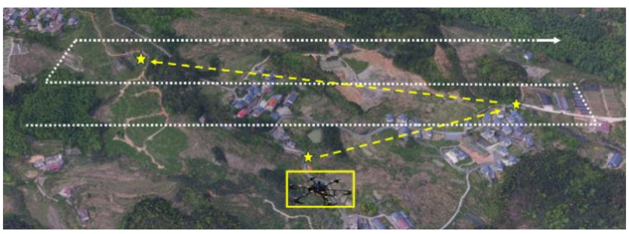

Figure 10. The observation UAV follows the flight trajectory, while the target UAV moves sequentially according to the control points.

3.2 Ablation Study of Graph Matching

In the processing pipeline, LED points are first detected using a segmentation model. Images with no fewer than 19 detected LEDs are retained, resulting in 38 valid samples. Observation graphs are constructed and matched with the template graph. The matching performance is evaluated using accuracy based on ground-truth correspondences:

$$\text{Accuracy} = \frac{X^T GT}{\sum GT} \tag{9}$$

where GT = a binary indicator vector of ground-truth correspondences.

To evaluate the contributions of the second-stage matching strategy and RANSAC-based geometric filtering, an ablation study is conducted with four configurations, as summarized in Table 1: (1) baseline RRWM; (2) RRWM with RANSAC-based geometric filtering; (3) RRWM with the proposed second-stage matching strategy; (4) the complete method. The angular threshold is set to $\tau = 20°$ in (3) and (4). The performance of each configuration is evaluated in both matching and decoding tasks.

Method	Accuracy(%)
RRWM	51.47
RRWM+RANSAC	52.88
Two-stage RRWM	95.59
Two-stage RRWM+RANSAC	98.03

Table 1. Ablation study results

The baseline RRWM suffers from ambiguity due to structural symmetry, resulting in low accuracy. RANSAC provides limited improvement. In contrast, the proposed two-stage strategy significantly improves performance by reducing ambiguity through direction-point-based constraints. The complete method achieves the best results, demonstrating improved robustness. As illustrated in Fig. 9, the node correspondence results under different strategies are compared. The first row shows the results of the baseline RRWM, where noticeable ambiguities occur due to the inherent symmetry of the geometric structure. The second row shows the results of the Two-stage RRWM with the direction-point constraint. By exploiting the geometric prior provided by the direction point, angular constraints are imposed on the candidate matches, effectively reducing ambiguities caused by rotational symmetry.

3.3 Comprehensive Evaluation of the Decoding Pipeline

To comprehensively evaluate the decoding capability of the proposed method in real-world scenarios, a quantitative analysis of the overall decoding pipeline is conducted in terms of pipeline availability, decoding accuracy, verification effectiveness, and overall system performance. The system consists of 19 light nodes, among which 18 nodes are used for data encoding, while the remaining direction node serves both orientation indication and verification purposes. Therefore, the encoding accuracy is evaluated based on the 18 data bits, while the verification mechanism is assessed separately.

Pipeline Availability. The decoding pipeline requires sufficient LED detections (≥19 nodes). The pipeline availability is defined as:

$$R_{\mathrm{valid}} = \frac{N_{\mathrm{valid}}}{N_{\mathrm{all}}} \tag{10}$$

where N_{valid} = samples with at least 19 detected LEDs.
N_{all} = total number of samples.

This metric reflects the proportion of samples that can enter the complete decoding pipeline in real scenarios.

Decoding Accuracy. After graph matching and color recognition, the recovered 18-bit encoding sequence is evaluated. Then:

$$Acc_{\mathrm{seq}} = \frac{N_{\mathrm{success}}}{N_{\mathrm{valid}}} \tag{11}$$

where N_{success} = correctly decoded samples (18-bit sequence)

This metric evaluates whether a complete decoding is achieved for each image and is a key indicator of practical usability.

Verification Effectiveness. To enhance the reliability of decoding results, a modulo-4 verification mechanism is introduced based on the check bit. The observed check value $p_{\mathrm{obs}}$is compared with the decoded check value p_{dec}. A decoding result is considered valid if $p_{\mathrm{dec}} = p_{\mathrm{obs}}$; otherwise, it is rejected. Based on this mechanism, Accuracy After Verification is defined:

$$Acc_{\mathrm{pass}} = \frac{N_{\mathrm{correct}}^{\mathrm{pass}}}{N_{\mathrm{pass}}} \tag{12}$$

where $N_{\mathrm{correct}}^{\mathrm{pass}}$ = correct samples after verification
N_{pass} = samples passing verification

This metric indicates the reliability of the final output results.

Overall System Performance. the effective decoding rate is defined to measure the proportion of samples that produce correct outputs among all valid samples:

$$Acc_{\mathrm{final}} = \frac{N_{\mathrm{correct}}^{\mathrm{pass}}}{N_{\mathrm{valid}}} \tag{13}$$

This metric provides a comprehensive evaluation of the overall usability of the system.

The statistical counts and evaluation metrics are summarized in Table 2 and Table 3. Out of 45 samples, 38 samples ($R_{valid} = 84.44\%$) satisfy the decoding condition, indicating that the proposed pipeline is applicable to most real scenarios. Among them, 29 samples are correctly decoded, achieving a sequence-level accuracy of $Acc_{seq} = 76.32\%$. After applying the verification mechanism, the accuracy of accepted results increases to $Acc_{pass} = 87.10\%$, demonstrating improved reliability. As a result, the final effective decoding rate is $Acc_{final} = 71.05\%$, achieving a balance between reliability and completeness of the decoding results. This demonstrates that the verification mechanism improves result reliability at the cost of reduced coverage (rejecting a small number of correctly decoded samples).

Symbol	Description	Value
N_{all}	Total number of samples	45
N_{valid}	Samples entering decoding pipeline	38
N_{success}	Correctly decoded samples	29
N_{pass}	Samples passing verification	31
$N_{\mathrm{correct}}^{\mathrm{pass}}$	Correct samples after verification	27

Table 2. Summary of statistical counts

Metric	Description	Value
R_{valid}	$N_{\mathrm{valid}}/N_{\mathrm{all}}$	84.44%
Acc_{seq}	$N_{\mathrm{success}}/N_{\mathrm{valid}}$	76.32%
Acc_{pass}	$N_{\mathrm{correct}}^{\mathrm{pass}}/N_{\mathrm{pass}}$	87.10%
Acc_{final}	$N_{\mathrm{correct}}^{\mathrm{pass}}/N_{\mathrm{valid}}$	71.05%

Table 3. Evaluation metrics of the decoding pipeline

3.4 Qualitative Results

To qualitatively demonstrate the decoding process, representative examples are shown in Fig. 11 and Table 4. For each sample, the detected LED nodes and the recovered node correspondences are visualized. After LED detection, graph matching, and color recognition, a complete quaternary encoding sequence is obtained. The sequence consists of 19 symbols, including 18 data symbols and 1 verification symbol. The verification symbol is used to check the consistency of the decoded sequence. After removing the verification symbol, the remaining sequence is divided into three subsequences, S_λ, S_ϕ, and S_h , corresponding to longitude, latitude, and height, respectively. These subsequences are then converted from quaternary representations into integer values, denoted as L'_λ, L'_ϕ, and L'_h , representing the low-order components of the coordinates. Based on the prior coarse position information (λ_0, ϕ_0, h_0) , the quotient terms corresponding to the actual encoding cycle are first determined. These quotient terms are then combined with the decoded low-order values to recover the complete integer representations $\left(L_\lambda, L_\phi, L_h\right)$. The recovered integer coordinates are then transformed into continuous geographic coordinates (λ, ϕ, h) . The results show that the recovered positions are consistent with the actual UAV locations. This demonstrates that the proposed method can reliably reconstruct the target position from the encoded light sequence, completing the full process from observation to coordinate recovery.

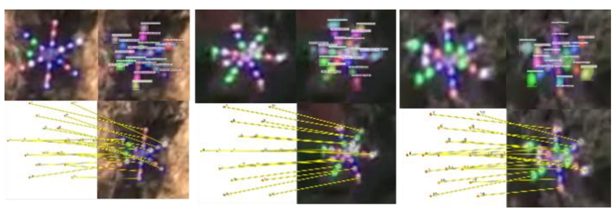

Figure 11. Representative examples of the decoding pipeline

Complete encoding sequence	0 10223 111021 0331322	3 11333 312200 1022201	0 10223 123021 0331322
$p_{\text{dec}} = p_{obs}$	True	True	False
Prior position information	113.9065568° 29.7940299° 158.13m	113.906561° 29.7940286° 40.33m	
S_{λ} S_{φ} S_{h}	10223 111021 0331322	11333 312200 1022201	
L'_{λ} L'_{λ} L'_{λ}	957 1229 15857	453 1177 3962	
q_{λ} q_{ϕ} q_{h}	1112368 72739 0	1112368 72739 0	
Recovered coordinates	113.9065789° 29.7940173° 158.57m	113.9065285° 29.7940121° 39.62m	

Table 4. Representative examples of coordinate recovery

4. Conclusion

This paper presents a control point deployment and recognition method based on a LMUAV. By leveraging an LED-encoded target and a geometry-aware graph matching decoding strategy, the proposed approach achieves accurate control point recognition and matching. The method eliminates the need for manual control point deployment and provides a high level of automation, demonstrating its practical potential. Nevertheless, the method is sensitive to LED detection quality and occlusion, and its performance may degrade under high noise or severe missing observations. Future work will focus on improving the robustness of the perception module to enhance system reliability under challenging conditions.

References

Atik, M.E., Arkalı, M., 2025. Comparative assessment of the effect of positioning techniques and ground control point distribution models on the accuracy of UAV-based photogrammetric production. *Drones*, 9(1), 15. doi.org/10.3390/drones9010015

Berra, E.F., Peppa, M.V., Mills, J.P., Moore, P., 2020. Advances and challenges of UAV SfM MVS photogrammetry and remote sensing: Short review. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*XLII-3/W12, 267–272. doi.org/10.5194/isprs-archives-XLII-3-W12-267-2020

Carion, N., et al., 2025. Segment Anything with Concepts (SAM 3). *Meta Superintelligence Labs*. https://ai.meta.com/research/publications/sam-3-segment-anything-with-concepts/

Cho, M., Lee, J., Lee, K.M., 2010. Reweighted random walks for graph matching. In: *Proc. ECCV*, pp. 492–505. doi.org/10.1007/978-3-642-15555-0_36

Colomina, I., Molina, P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* 92, 79–97. doi.org/10.1016/j.isprsjprs.2014.02.013

Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F.J., Marín-Jiménez, M.J., 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognit.* 47(6), 2280–2292. doi.org/10.1016/j.patcog.2014.01.005

Han, S., et al., 2024. Enhancing direct georeferencing using real-time kinematic UAV imagery in large-scale infrastructure applications. *Drones*, 8(12), 736. doi.org/10.3390/drones8120736

Harwin, S., Lucieer, A., 2012. Assessing the accuracy of georeferenced point clouds produced via multi-view stereopsis from UAV imagery. *Remote Sens.* 4(6), 1573–1599. doi.org/10.3390/rs4061573

Kovanič, Ľ., Blišťan, P., Urban, R., et al., 2023. Review of photogrammetric and lidar applications of UAV. *Appl. Sci.*13(11), 6732. doi.org/10.3390/app13116732

Kregar, K., et al., 2025. Combining UAV photogrammetry and TLS for change detection in challenging coastal environments. *Drones*, 9(4), 228. doi.org/10.3390/drones9040228

Luhmann, T., Robson, S., Kyle, S., Boehm, J., 2019. *Close-range photogrammetry and 3D imaging*. 2nd ed. De Gruyter, Berlin/Boston. doi.org/10.1515/9783110607253

Nex, F., Remondino, F., 2014. UAV for 3D mapping applications: A review. *Appl. Geomat.* 6, 1–15. doi.org/10.1007/s12518-013-0120-x

Olson, E., 2011. AprilTag: A robust and flexible visual fiducial system. In: *Proc. IEEE Int. Conf. Robot. Autom.* pp. 3400–3407. doi.org/10.1109/ICRA.2011.5979561

Sanz-Ablanedo, E., Chandler, J.H., Rodríguez-Pérez, J.R., Ordóñez, C., 2018. Accuracy of UAV photogrammetry: The role of ground control points. *Remote Sens.* 10(10), 1606. doi.org/10.3390/rs10101606

Wolf, P.R., Dewitt, B.A., Wilkinson, B.E., 2014. *Elements of photogrammetry with applications in GIS*. 4th ed. McGraw-Hill, New York.

Worldsemi, 2020. WS2812B Intelligent Control LED Datasheet. World Semiconductor Co., Ltd. https://cdn-shop.adafruit.com/datasheets/WS2812B.pdf

Zeybek, M., Şanlıoğlu, İ., 2023. Improving the spatial accuracy of UAV platforms using direct georeferencing methods: An application for steep slopes. *Remote Sens.* 15(10), 2700. doi.org/10.3390/rs15102700

Zhang, K., et al., 2022. Optimization of ground control point distribution for UAV surveying. *Sustainability* 14(15), 9505. doi.org/10.3390/su14159505