Research Article

Xu Chen*

UAV patrol path planning based on machine vision and multi-sensor fusion

https://doi.org/10.1515/comp-2022-0276
received December 25, 2022; accepted March 29, 2023

Abstract: With the rapid development of unmanned aerial vehicle (UAV) technology, there are more and more fields of UAV application. This research mainly discusses the UAV patrol path planning based on machine vision and multi-sensor fusion. This article studies how to apply ultrasonic, a classic ranging sensor, to obstacle avoidance of UAVs. The designed ultrasonic obstacle avoidance system is a complete set of hardware and software systems. The hardware part consists of a forward ultrasonic module and a central signal processing module. Among them, a single-axis stabilization gimbal is designed for the forward ultrasonic module, which decouples the attitude angle of the UAV and the pitch detection angle of the ultrasonic sensor. In the central signal processing module, Kalman filtering is performed on the ultrasonic data in the four directions of front, rear, left, right, and left, and the obstacle avoidance control signal is sent to the flight controller according to the filtered sensor data. At the same time, a human–computer interaction interface is also designed to set various parameters of the obstacle avoidance system. For the route planning method of the tower, the routine steps are used to inspect the tower with a single-circuit line, and the specific targets are the insulator string, the ground wire, and the conductor. In this study, the average statistical result of the straight-line distance of the UAV patrolling 100 m is 99.80 m, and the error is only 0.2%. The fusion obstacle avoidance control method based on machine vision is suitable for the engineering application of UAV perception obstacle avoidance. The obstacle avoidance method adopted in this article can be extended to most flight control platforms, and it is a control method with broad application prospects.

Keywords: machine vision, multiple sensors, drone inspection, path planning

1 Introduction

UAV is a kind of unmanned aerial vehicle that can fly autonomously, semiautomatically, or remotely. It has the characteristics of being light weight, small size, high mobility, good concealment, and strong adaptability. With the rapid development of the economy, almost all industries, including electric power inspection, air detection, and drone logistics, have various needs for drones. However, manual operation of obstacle avoidance has the problems of low efficiency, poor effect, instability, and high cost, which seriously affect operation efficiency. Regarding the problem of UAV obstacle avoidance control, a lot of theoretical research and practical tests have been carried out at home and abroad, and various obstacle avoidance control methods have been tried, including classical ultrasonic feedback obstacle avoidance control, structured light binocular obstacle avoidance control, proportional guidance obstacle avoidance control, and fuzzy control. Some of them have also had certain achievements, but there are also some shortcomings, such as the large interference of ultrasonic echoes, the large interference of strong light outside the structured light, and the unstable actual effect of intelligent control.

The path planning technology of the UAV is to plan the optimal global path or the safest local path for the UAV in a three-dimensional (3D) environment, considering the environmental threats, the physical conditions of the UAV, the limitations of the flight area, and many other factors. This enables the UAV to avoid all kinds of dangers and complete the predetermined mission objectives. In the task, due to the requirements of the collection and fusion of sensor information and the control of the UAV group consistency, it puts forward relatively high requirements for the information exchange ability of the system. Among them, wireless communication, as the main method of communication between UAV and base station or UAV, greatly affects the overall system performance. At the same time, due to the limited size of the UAV and the limited size of the energy module it can carry, this has also become one of the bottlenecks in
the performance of the UAV. Finally, in the traditional communication field, the scarcity of wireless spectrum also has a great performance impact in the field of UAV communication [1].

The path planning technology, namely the shortest path search algorithm, is a hot issue in many technical fields at present, with broad application prospects and scientific research value. This article proposes a method to jointly invoke the electromagnetic domain and geographic domain degrees of freedom to optimize the UAV path planning strategy. When UAVs perform reconnaissance missions, the environment they often face is a multi-domain complex environment. Among them, the electromagnetic domain has two characteristics in most cases. First, the electromagnetic environment changes dynamically. The time interval of change in the electromagnetic domain is often less than milliseconds, and this environment puts forward extremely high requirements for the real-time performance of the strategy planning algorithm. Second, the environment may contain complex disturbances. In practical problems, UAVs often work in adversarial environments. In the electromagnetic domain, UAVs will face unintentional interference from natural and civil environments, as well as intentional interference from adversaries. Therefore, while taking into account the reconnaissance task, the UAV needs to choose an appropriate communication strategy to deal with interference, such as frequency domain maneuvering. At the same time, through communication strategies such as frequency domain maneuvering, the feasible strategy space of UAV path planning is increased, which can improve the performance of reconnaissance tasks. After 400 frames, when the scale changes greatly, the error rate can be controlled below 0.6, and the coverage rate can be above 0.3.

2 Related work

As a means of information exchange between UAVs and ground systems, communication technology is an important part of UAVs. With the development of wireless communication, network communication, and satellite communication technology, UAV communication technology has also been improved. Yi et al. proposed a new method for distributed multi-target tracking of multistatic radar. The method is based on fused posteriors using generalized covariance intersection of multi-objective densities in a multi-objective Bayesian filtering scheme. The solution they proposed is particularly suitable for sensor fusion with posterior density. The posterior density is parameterized as a generalized multi-Bernoulli (GMB) distribution, which is an unlabeled version of the variational object, co-occurrence and variational objects density by discarding labels. To obtain a closed-form solution for fused GMB densities, they used an efficient approximation of the densities. The approximate density is another GMB density that preserves the first moment (intensity or probability hypothesis density) and cardinality distribution of the original density. Therefore, it is called a second-order approximation of the GMB density [2]. Weon and Lee believe that driverless vehicles have higher requirements on the reliability and recognition performance of the road environment and driving conditions. Since a single sensor cannot accurately identify various driving conditions, a recognition system using only a single sensor is not suitable for autonomous driving due to the uncertainty of recognition. They developed an autonomous vehicle that uses sensor fusion of radar, lidar, and vision data, which is coordinate-corrected via global positioning system (GPS) and inertial measurement unit (IMU). Deep learning and sensor fusion improve the recognition rate of stationary objects in the driving environment (such as lanes, signs, and crosswalks), and accurately identify dynamic objects (such as vehicles and pedestrians). Through actual road tests, they verified that the unmanned autonomous driving technology developed in this research meets the reliability and stability requirements of the NHTSA Level 3 autonomous driving standard [3]. Elgharbawy et al. present a new method for validating multi-sensor data fusion algorithms in complex automotive sensor networks. Multi-sensor fusion plays a central role in enhancing interpretation of traffic conditions, facilitating reasoning and decision-making. As such, it plays an important role in the continued innovation of advanced driver assistance systems (ADASs) that are paving the way for autonomous driving. They introduced a real-time framework that can benchmark the performance of fusion algorithms at the electronic system level using hardware-in-the-loop (HiL) co-simulation. Their research provides a quantitative approach to the trade-off between physical realism and computation for real-time synthetic simulations. Their proposed framework illustrates a general architecture for ADAS sensor error injection for robustness testing of systems under test. They constructed a lemniscate model with error to find multivariate outliers with Mahalanobis distance [4]. Ding et al. proposed a longitudinal vehicle speed estimator based on multi-sensor fusion for four-wheel independent drive electric vehicles by utilizing GPS, BeiDou navigation and positioning (GPS-BD) module, and low-cost IMU. To accurately estimate vehicle speed, they first proposed a method combining wheel speed and GPS-BD information to compensate for the effect of road
slope on the GPS-BD module output horizontal speed and IMU longitudinal acceleration. Then, they synthesized a multi-sensor fusion-based longitudinal vehicle speed estimator using three virtual sensors that generated three longitudinal vehicle speed trajectories from multiple sensor signals. Finally, through HiL tests, they verified the accuracy and reliability of the proposed longitudinal vehicle speed estimator under different driving conditions [5]. The goal of Dawood et al. is to develop an ensemble model based on image processing techniques and machine learning to automatically achieve consistent peeling detection and numerical representation of damage in subway networks. The ensemble model includes a hybrid algorithm, interactive 3D representation, and is supported by regression analysis to predict spalling depth. First, the red, green, and blue images are preprocessed by a hybrid algorithm to remove image noise and enhance key cues related to peeling. Second, a spalling processor is designed to detect defect properties, thereby providing a 3D visual model of the defect. Third, they use a new regression analysis model combined with image processing techniques in intensity curve projection to measure the depth and severity of spalling damage [6]. The UAV provides a stable experimental platform for the whole system, receives the attitude adjustment command sent by the remote controller, and calculates the duty ratio of the four motors, thereby controlling the four motors to provide the corresponding lift force to the body.

3 UAV patrol path planning and design

3.1 Machine vision

UAV path planning is to study the problem of searching the optimal path from the starting point to the destination point of the UAV. A path with low cost not only saves the cost of UAV operation, but also increases the success rate of the UAV to complete the task. A path with high security can also improve the survival rate of the UAV.

According to the properties of the expectation operator [7]:

\[ \text{MSE} = \varepsilon \{ s^2(t) \} - 2 \varepsilon \{ s(t) y(t) \} + \varepsilon \{ y^2(t) \} = T_1 + T_2 + T_3. \]  

Class 1 pixel count is as follows:

\[ W_1^{(k)} = \sum_{i=1}^{L} \text{Hist}(i). \]  

Class 1 mean gray value is as follows:

\[ M_1^{(k)} = \sum_{i=1}^{L} \frac{i \cdot \text{Hist}(i)}{W_1^{(k)}.} \]  

Class 1 variance is as follows:

\[ \sigma_1^{(k)} = \sum_{i=1}^{L} (i - W_1^{(k)})^2 \cdot \frac{\text{Hist}(i)}{W_1^{(k)}.} \]  

The gradient of the image function \( f(x,y) \) at point \((x,y)\) is as follows:

Size is as follows:

\[ g(x, y) = \sqrt{G_x^2 + G_y^2} = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2}. \]  

Direction is as follows:

\[ \varphi(x, y) = \arctan(G_y/G_x). \]  

UAV path planning mainly includes flight reconnaissance environment information acquisition, flight constraint simulation, and route planner design. This article uses cascade proportion integral differential control for the control of the UAV, that is, the control method of the inner loop and the outer loop is adopted, and the attitude system is controlled by the inner loop. The control objects are the attitude, yaw rate, and vertical speed of the UAV; the position system is controlled in the outer loop, and the control objects are the horizontal speed, heading angle, and height of the UAV. The outputs of the inner loop control are control torque and vertical thrust, which are converted to motor voltages by using static mixing matrices or by feeding them into the transpose model of the propulsion system [8,9].

3.2 Multi-sensor fusion

The flight environment information acquisition mainly depends on the support of other ground and high-altitude equipment and systems. The UAV environment information is obtained through radar microwave and other technical means, and the model that can be calculated by computer is constructed through mathematical modeling. The infrared thermal imaging system converts the infrared radiation naturally emitted by the object into an infrared thermal image visible to the human eye through an internal optical conversion mechanism. Therefore, the determination of the position of the object in the infrared image is the same as the determination of the position of the visible light image, which is based on the simplified pinhole imaging principle. It is assumed that the infrared camera and the visible light camera are placed in such a way that the optical axes of the two cameras are parallel,
and the distance between them is fixed and as small as possible [10]. The working principle of the infrared imaging system is shown in Figure 1.

When a square-sized area with a side length of |PQ| is selected, the actual area size is \( d_1^2 \), and the total number of pixels in the image is \( m_1^2 \), which is represented as follows:

\[
m_1^2 = \left( \frac{f_1}{z_1} \ast \frac{1}{d_{lf}} \right)^2 d_{lf}^2.
\]

Taking a square area whose side length is \( |PQ| \), set the physical distance between two adjacent pixels (assuming the horizontal and vertical directions are the same) as \( d_{vl} \), and the actual area of the square at this time is \( d_1^2 \). The total number of pixels in the image is \( m_2^2 \), and the formula is as follows [11]:

\[
m_2^2 = \left( \frac{f_2}{z_2} \ast \frac{1}{d_{vl}} \right)^2 d_{vl}^2.
\]

When there are two images of the same size of the actual area, the ratio of the total number of imaged pixels is as follows:

\[
\frac{m_1^2}{m_2^2} = \left( \frac{f_1}{f_2} \ast \frac{z_1}{z_2} \ast \frac{d_{vl}}{d_{lf}} \right)^2.
\]

The modular parameters are as follows [12]:

\[
F_i(n) = I_i(n), \quad L_i(n) = \sum W_{ik} Y_i(n - 1), \quad U_i(n) = F_i(n)(1 + \beta L_i(n)),
\]

\[
Y_i(n) = \begin{cases} 
1, & \text{if } U_i(n) > E_i(n - 1) \\
0, & \text{else,}
\end{cases}
\]

\[
E_i(n) = \exp(-a_E)E_i(n - 1) + V_i Y_i(n),
\]

\[
T_{ij}^{l,k} = T_{ij}^{l,k}(n - 1) + Y_{ij}^{l,k}(n).
\]

The measurement frequency of IMU is very high, but the measurement process will be disturbed by Gaussian error and its own random walk. To this end, the IMU measurement model can be described by using a mathematical formula [13]:

\[
\tilde{\alpha}_W(t) = \alpha_W(t) + \eta_W(t),
\]

\[
\tilde{\beta}_W(t) = \beta_W(t) + \eta_W(t).
\]

Assuming that the time interval between two IMU measurements is \( \Delta t \), then:

\[
R_{WB}(t + \Delta t) = R_{WB}(t) \exp \left( \int_t^{t + \Delta t} \omega_W(\tau) \, d\tau \right),
\]

\[
\nu_W(t + \Delta t) = \nu_W(t) + \int_t^{t + \Delta t} a_W(\tau) \, d\tau,
\]

\[
p_W(t + \Delta t) = p_W(t) + \int_t^{t + \Delta t} v_W(\tau) \, d\tau
\]

\[+ \int_t^{t + \Delta t} a_W(\tau) \, d\tau^2.
\]

3.3 Simulation platform based on MFC and OpenGL

In order to effectively realize the algorithm simulation, a 3D visual dynamic algorithm simulation platform based on Microsoft Foundation Classes (MFC) and OpenGL is constructed in this article. The simulation platform integrates five modules: the main control program, the data

![Figure 1: The working principle of the infrared imaging system.](image-url)
receiving module, the output display module, the initialization module, and the path planning module [14]. These five modules constitute the overall operation system of the software, and each module has a strong correlation. The function of the whole platform is the calculation and display of the path planning of the UAV [15]. Most of the available path planning algorithms can be used for single UAV path planning. The software as a whole achieves good UAV path planning human–computer interaction standards. The overall structure of the platform is shown in Figure 2. The simulation platform uses the view/document structure method in MFC to create the human–computer interaction interface of the software, and the interface is simple and easy to understand. This allows users to perform various map parameters, algorithm preferences, and viewing angle adjustments in the main control window without needing to know the specific content of the algorithm. This saves the need to fully understand the program to adjust the parameters in the algorithm [16].

3.4 Path optimization and formulation

The complexity of the environment, the constant interference of various threat sources on the UAV, and the diversity and multi-level of tasks, the cruise, communication, task execution, and other requirements of the UAV are very strict. The GPS and navigation system on the drone are all composed of electronic equipment. These devices and data communication systems may experience inaccurate data transmission and equipment failures after strong electromagnetic interference. The inspection of the transmission line cannot be completed safely, and it may even cause damage to the line and tower due to misoperation [17]. In order to obtain the speed, attitude, and position information of the drone, a sensor plug-in needs to be built in Gazebo. This includes an IMU for angular velocity and angular acceleration, a barometer for altitude measurement, ultrasonic sensors for assisting takeoff and landing, magnetic sensors for heading, and GPS for position and velocity. At the same time, two plug-ins were added to Gazebo to calculate the propulsion and drag acting on the aircraft for the four motor voltages and wind vector, respectively [18,19].

A single UAV has a single performance and limited load. When the task is complex, a single UAV cannot complete. At this time, it is necessary to consider how to solve the problem through the cooperation of multiple UAVs. The distribution track of the magnetic induction intensity around the transmission line is an ellipse, the maximum value of the magnetic induction intensity is near the wire, and the spatial distribution of the magnetic induction intensity is shown in Figure 3.

The purpose of UAV path planning is to propose a strictly arranged movement, which is to arrange a feasible flight route for a UAV to fly from one place to another. By using differential GPS positioning technology, UAVs can achieve outdoor centimeter-level positioning accuracy. This can completely ensure that the drone maintains a safe distance from the line and tower during the inspection of the power line, and will not endanger the safety of the line. Therefore, the use of differential GPS positioning technology can solve the problem caused by the positioning

**Figure 2:** Overall structure of the platform.
error of UAV in power line inspection [20]. The differential GPS system is mainly composed of a base station, a digital radio station, and a rover. The differential GPS system is shown in Figure 4.

### 3.5 UAV patrol path planning

The body attitude is directly calculated from the accelerometer and gyroscope output data according to the kinematic and dynamic equations. Therefore, the accuracy and stability of the sensor output data have a decisive effect on the working state of the aircraft. In practice, the measurement accuracy of the sensor chip itself has limitations. In addition, there is mechanical vibration during the flight, which causes the noise of different frequencies to be mixed in the original data of the sensor [21]. The flight control unit needs to select an appropriate data calibration method to optimize the data [22]. The flight control unit uses the low-pass filtering and moving average filtering methods to process the accelerometer data, uses the temperature drift removal method to process the gyroscope data, and finally uses the accelerometer data to correct the gyroscope data to obtain relatively credible original data [23,24].

The traditional path usually consists of a series of segmented straight lines connecting each waypoint, which link the starting point and the ending point.

Mapping the input of a linear problem to a nonlinear feature space, we obtain the following:

\[
    w = \alpha \sum \phi(x_i).
\]

(21)

Doing the dot product of all samples, and the result being stored in the kernel function \(K\), we obtain the following:

\[
    K = \kappa(x_i, x_j).
\]

(22)

The regression function \(f(z)\) of the target block image is as follows:

\[
    f(z) = w^Tz = \sum \alpha \kappa(z, x_i),
\]

(23)

\[
    \alpha = (k + \lambda I)^{-1} y,
\]

(24)

where \(K\) is the kernel correlation matrix of all training samples.

Diagonalizing \(f(z)\), we have the following:

\[
    f'(z) = k \odot \alpha.
\]

(25)

### 3.6 Collision avoidance

Obstacle avoidance requires the UAV to perceive the surrounding environment. It is assumed that there are suitable sensors on the UAV to sense the distance between

---

**Figure 3:** The distribution of the magnitude of the magnetic induction in space.

**Figure 4:** Differential GPS system.
the obstacle and the UAV in a certain range. When building static obstacle models of various irregular shapes in many literature, the usual way to deal with them is to expand them according to the size of the UAV. This method uses the circumscribed circle as the obstacle area and judges whether collision avoidance is necessary according to the relationship between the projected length of the distance between the UAV and the center of the circumscribed circle in the direction of the UAV's forward speed and the safety distance \[25\].

The obstacle is modeled as a circle, and under some obstacle avoidance algorithms, the evasive motion of the UAV can be made as an arc, and then it can benefit from the navigation control law suitable for tracking the circular route \[26\]. During the flight of the UAV, if it is detected that there is a no-fly zone ahead, the original path is abandoned by a detour. It selects an appropriate point in the vicinity of the no-fly zone, and it re-plans the path online according to the target point position, the newly selected node, and the current position. When a new obstacle is detected, this process is repeated, or a new route is generated according to a certain avoidance route template. Another way is to change the heading temporarily and return to the original path to continue the movement after bypassing the no-fly zone \[27,28\]. Generally speaking, the track below a certain threshold is expensive for UAVs, and this distance is usually considered as the minimum range.

**Figure 5:** Changes in the three directions of X, Y, Z.
4 Results of UAV patrol path planning

As a complex system composed of multiple UAVs, the research of multiple UAV formation involves the intersection of control, communication, aerodynamics, artificial intelligence, and other disciplines. If the base station coordinates set in the rover program are too far from the true value, the distribution of GPS positioning is relatively scattered. After further analysis, the distribution of GPS coordinates is decomposed into three directions of $X$, $Y$, and $Z$, and the three directions of $X$, $Y$, and $Z$ are obtained. The changes in each direction are shown in Figure 5.

It can be seen from Figure 5 that there is a maximum offset of $\pm 0.5$ m in the $X$ direction and a maximum offset of $\pm 0.5$ m in the $Y$ direction. The change in the $Z$ direction is larger than that in the $X$ and $Y$ directions, and the maximum offset value reaches $\pm 1$ m. It can be seen that when the set value of the base station is quite different from the real value, the maximum positioning error of the differential GPS reaches $\pm 1$ m, and the error is large. One datum is sampled every second, and the mean and standard deviation of the distribution in the three directions of $X$, $Y$, and $Z$ are obtained statistically, and the results are shown in Table 1.

Then analyzing the data in the $Y$ direction, it can be seen from Figure 6 that it is similar to the situation in the $X$ direction. The shape of the trigonometric function cannot be seen in the distribution diagram in the $Y$ direction. The curve burrs are very large, indicating that the positioning accuracy in the $Y$ direction is very poor. The $Z$-direction distribution of GPS single-point positioning is shown in Figure 6.

At the same time, the maximum offset and minimum offset data in the three directions of $X$, $Y$, and $Z$ during the circular motion are recorded. These data correspond to the radius of the offset circle, as shown in Table 2.

The content of UAV flight inspection includes tower foundation inspection (13 photos), ground wire inspection between stalls (video shooting), and defect inspection. The specific content is shown in Table 3.

<table>
<thead>
<tr>
<th>Direction</th>
<th>Mean (m)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>-0.03808</td>
<td>0.14830</td>
</tr>
<tr>
<td>$Y$</td>
<td>0.17610</td>
<td>0.33463</td>
</tr>
<tr>
<td>$Z$</td>
<td>0.14470</td>
<td>0.33131</td>
</tr>
</tbody>
</table>

Table 1: Mean and standard deviation of the distribution in the three directions of $X$, $Y$, and $Z$

![Figure 5: UAV patrol path planning diagram](image1)

![Figure 6: GPS single-point positioning Z direction distribution](image2)
The camera used in the experiment was calibrated using the checkerboard calibration board. There are many tools for camera calibration; Matlab is mainly used here for calibration, and the calibration results of ZED Stereo Camera are shown in Table 4.

The global navigation satellite system (GNSS)-IMU positioning results have certain errors, but the relative distance measurement results are relatively accurate: The average measurement value of the 100 m straight-line distance reaches 99.80 m, and the error is only 0.2%. In addition, it can be seen from data set 2 that the GNSS measurement results will gradually converge and become more accurate after the pause, and the error of 0.17 m is relatively small. The distance detection under different motion conditions is shown in Table 5.

For a frame of data with about 25,000 inserted point clouds, the method of de-redundant map stitching only needs to add 6784.5 data points on average. It can be seen that this method can effectively reduce the amount of map data from the source while ensuring the accuracy of spatial modeling. Combined with the use of filtering methods, a concise spatial 3D model can be finally obtained, and if necessary, the octree structure can be further transformed into a spatial grid map model. The change in the number of point clouds is shown in Figure 7.

Four videos shot by themselves are used as the test set to simulate the videos shot by the camera mounted on the UAV. The tempered glass insulators and composite insulators were photographed when the scale became larger or smaller, which was convenient to analyze the specific test effect. The detailed introduction of the four video sequences mentioned in this article is shown in Table 6.

For video sequence 1, the test object is a tempered glass insulator. The resolution of the video sequence is 544 pixels × 960 pixels, and the purpose is to test the adaptability of the algorithm when the scale of the tempered glass insulator becomes larger. The test results of video sequence 1 are shown in Figure 8.

It can be seen that at the 0th frame, the scale error rate is relatively large, reaching 0.27. The pixel error of the center point also reached 28 pixels. The target tracking frame is larger than the actual target area, especially the upper part of the frame identifies a part of the background in the tracking frame. But as it gets closer and closer to the target, the scale of the target gets bigger and bigger, and the position of the recognition gets more and more accurate. The scale error can be controlled below 0.05, which means that there is a small error

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Tower number</th>
<th>Shooting standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>001#</td>
<td>Take a clear photo of the sign at the tower sign</td>
</tr>
<tr>
<td>2</td>
<td>002#</td>
<td>Take four photos each at 30° above</td>
</tr>
<tr>
<td>3</td>
<td>003#</td>
<td>the four corners of the tower (it can reflect the surrounding environment of the tower, the whole picture of the tower, whether the tower material is lost, whether the insulator string has self-explosion, etc.)</td>
</tr>
<tr>
<td>4</td>
<td>004#</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>005#</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>006#</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Maximum and minimum offsets in X, Y, and Z directions

<table>
<thead>
<tr>
<th>Direction</th>
<th>Max (m)</th>
<th>Min (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>3.256</td>
<td>−1.349</td>
</tr>
<tr>
<td>Y</td>
<td>2.566</td>
<td>−2.106</td>
</tr>
<tr>
<td>Z</td>
<td>3.157</td>
<td>−5.227</td>
</tr>
</tbody>
</table>

Table 3: Inspection content

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Tower number</th>
<th>Shooting standard</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>001#</td>
<td>Take a clear photo of the sign at the tower sign</td>
</tr>
<tr>
<td>2</td>
<td>002#</td>
<td>Take four photos each at 30° above</td>
</tr>
<tr>
<td>3</td>
<td>003#</td>
<td>the four corners of the tower (it can reflect the surrounding environment of the tower, the whole picture of the tower, whether the tower material is lost, whether the insulator string has self-explosion, etc.)</td>
</tr>
<tr>
<td>4</td>
<td>004#</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>005#</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>006#</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Calibration results of ZED cameras

<table>
<thead>
<tr>
<th>Camera internal parameters</th>
<th>$F_x$</th>
<th>$F_y$</th>
<th>$C_x$</th>
<th>$C_y$</th>
<th>$K1$</th>
<th>$K2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numerical value</td>
<td>707.699</td>
<td>707.692</td>
<td>707.671</td>
<td>337.655</td>
<td>−0.177</td>
<td>0.0347</td>
</tr>
<tr>
<td>Binocular camera extrinsic parameter $R$ (rotation vector)</td>
<td>−0.00777</td>
<td>0.012069</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binocular camera external parameter $T$</td>
<td>119.697</td>
<td>0.01269</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Distance detection in different motion situations

<table>
<thead>
<tr>
<th>Data set</th>
<th>Sports situation</th>
<th>Distance between the beginning and end of the track (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Walking slowly</td>
<td>98.08</td>
</tr>
<tr>
<td>1</td>
<td>Walking fast</td>
<td>98.92</td>
</tr>
<tr>
<td>2</td>
<td>End stop</td>
<td>100.17</td>
</tr>
<tr>
<td>3</td>
<td>Join bump 1</td>
<td>101.4</td>
</tr>
<tr>
<td>4</td>
<td>Join bump 2</td>
<td>99.77</td>
</tr>
<tr>
<td>5</td>
<td>Join the speed</td>
<td>99.21</td>
</tr>
<tr>
<td>6</td>
<td>1 min pause in between</td>
<td>100.21</td>
</tr>
<tr>
<td>7</td>
<td>Turn around</td>
<td>99.94</td>
</tr>
<tr>
<td>8</td>
<td>Eight characters</td>
<td>100.51</td>
</tr>
</tbody>
</table>
between the area of the tracking frame and the actual area of the target. The pixel error of the center point is also about 5 pixels, all of which are below 30 pixels. For a video sequence of 544 pixels \( \times \) 960 pixels, the center point pixel error is relatively small. The coverage rate of the tracking frame also reaches more than 0.95, which proves that the tracking frame of the project implemented in this article can cover the entire tracking target more comprehensively. As a complex system composed of multiple UAVs, the research of multiple UAV formation involves the intersection of control, communication, aerodynamics, artificial intelligence, and other disciplines.

For video sequence 3, the video sequence used in this article adds a certain background interference, and two composite insulators are used for testing, which increase the difficulty of testing. It can be seen that the test results are relatively good at the beginning. After 400 frames, when the scale changes greatly, the error rate can be controlled below 0.6, and the coverage rate can be above 0.3. The results of the test are not ideal, but the target that needs to be identified can still be initially tracked. The test results of video 4 are shown in Figure 9.

The obstacle avoidance process of the UAV with a flight speed of 7 m/s is shown in Figure 10. The entire obstacle avoidance process takes less time to avoid two obstacles due to the faster flight speed. At present, after decades of continuous research and development and actual combat accumulation, the single UAV system has achieved certain technical maturity, while multi-UAV formation flight is a new research field involving multi-disciplines.

5 Discussion

IMU is a necessary sensor for UAV to achieve stable flight. Usually, motion unit is composed of accelerometer and gyroscope. The accelerometer is used to measure the acceleration of the object, and the gyroscope is used to measure

<table>
<thead>
<tr>
<th>Serial number</th>
<th>Testing object</th>
<th>Testing purposes</th>
<th>Resolution (pixels ( \times ) pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Tempered glass insulator</td>
<td>Algorithm adaptability when the insulator scale becomes larger</td>
<td>544 ( \times ) 960</td>
</tr>
<tr>
<td>2</td>
<td>Tempered glass insulator</td>
<td>Algorithm adaptability when insulator scale becomes smaller</td>
<td>960 ( \times ) 544</td>
</tr>
<tr>
<td>3</td>
<td>Composite insulator</td>
<td>Algorithm adaptability when the insulator scale becomes larger</td>
<td>960 ( \times ) 544</td>
</tr>
<tr>
<td>4</td>
<td>Composite insulator</td>
<td>Algorithm adaptability when insulator scale becomes smaller</td>
<td>960 ( \times ) 544</td>
</tr>
</tbody>
</table>
the angular velocity of the object. Accelerometers can be divided into capacitive, piezoelectric, and piezoresistive types according to the measurement method. The measurement principle is generally to convert the acceleration into an electrical signal by using the principle that the acceleration will change some electrical characteristics of the measuring element. Common micro-electro mechanical system accelerometers mostly use capacitive measurement methods, which use the capacitance changes generated by electrode movement to detect acceleration values. With the change of acceleration, the movable electrode changes its position in the damping mechanism composed of air and reed, that is, changes the distance between the upper and lower fixed electrodes, which causes the change of the capacitance value between the electrodes.

It uses depth extraction and obstacle distance measurement in the global field of view. Although the obstacle detection range is wider, in the process of obstacle threat perception on low-altitude routes, the ground or ground objects may be misjudged due to the pitch of the aircraft. How to achieve an effective obstacle recognition effect under the uncertain interference information of low-altitude routes is a problem that needs to be further discussed. UAV line patrol has the characteristics of convenience, efficiency, and practicality. Moving the harsh environment in the field to the control room, or even the house, UAV equipped with some infrared, ultraviolet, camera, and other equipment can locate and find hidden faults in the budding state in time, and eliminate the faults to protect the safe operation of transmission lines. This is a power
line inspection method with potential for development. In order to improve the efficiency of line inspection, it is divided into long-distance inspection and short-distance inspection according to the length of the distance. Then, the equipment is used to first query the abnormal points of transmission line faults, and then hovering over the fault points to take pictures in all directions, so as to obtain clearer information in an all-round way, and prepare for subsequent image processing and decision-making.

With the further development of image processing technology, UAV will be a better and more efficient way of power line inspection, and will gradually become the development direction of power line inspection, opening a new era of power line inspection. In many applications of drones, the camera sensor is inseparable, and the camera sensor can obtain more abundant information than other sensors. Based on machine vision, the information captured by the camera can be extracted and processed to realize the application function of the UAV. Compared with the target tracking task in the common scene, the target tracking task in the UAV scene is more complicated and difficult. In addition to the need to overcome challenges such as illumination changes, scale changes, object occlusion, and deformation during object tracking, it is also necessary to overcome the disturbance of the image blur caused by the jitter when the UAV is flying, and the image angle change caused by the yaw movement, which makes the target tracking task more difficult. In addition, considering the computing power of the UAV processor, there are higher requirements for the complexity and real-time performance of the tracking algorithm.

Figure 9: Video 4 test results.
Image filtering also causes the edge information of some target objects in the image to be weakened while denoising. In addition, the quality of the ambient lighting will cause insufficient image contrast, which will affect the further processing of the image. Image enhancement processing can improve people's sensory experience, and at the same time can highlight features well, which brings a lot of convenience to image analysis and processing. The main realization of image enhancement is to highlight the useful feature information to make the blurred image clear, and at the same time, it can be operated in a targeted manner according to specific requirements. Finally, the effect of reducing the feature degree of useless objects, increasing the difference between different features, improving the image quality, and increasing the amount of information can be achieved, and the optimization results of the recognition accuracy can be improved.

Generally speaking, the larger the memory occupied by the image content, the more comprehensive information it can carry, and it also means that it will increase the burden on computer operations. So this article crops the image before image processing and divides the original image into different regions according to different characteristics. The feature points in the original image are preserved and the relevant structural features are left at the same time, and the image information is reduced to reduce the workload and improve the efficiency. Therefore, image segmentation is a very important step in machine vision, and it is also the premise of automatic image analysis.

6 Conclusion

This article mainly introduces the flight principle and target tracking principle of UAV. The flight principle of the UAV mainly introduces how the UAV realizes the flight attitudes such as hovering, elevating motion, pitching motion, rolling motion, and yaw motion. The target tracking principle of UAV mainly introduces how the target position is converted from the 2D image coordinate system to the 3D UAV coordinate system. The installation angles of the binocular camera, optical flow sensor, and lidar are calibrated. At the same time, the 3D point cloud of feature points recovered in oriented FAST and rotated BRIEF - simultaneous localization and mapping can also improve the UAV's ability to perceive 3D obstacles. Aiming at the optimization of route planning and image recognition algorithm, the images collected today are of low definition, which are related to flight speed, altitude, and image processing.

Funding information: The author(s) received no financial support for the research, authorship, and/or publication of this article.
Conflict of interest: The authors declare that there are no conflict of interest regarding the publication of this article.

Ethical approval: This article does not contain any studies with animals performed by any of the authors.

Informed consent: No conflict of interest exists in the submission of this manuscript, and manuscript is approved by all authors for publication. I would like to declare on behalf of my co-authors that the work described is original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part. All the authors listed have approved this manuscript.

Data availability statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

References


