Eagle eye Aerial robot team Technical Paper

Sun Bo
Ningxia university/Institute of physics and electrical engineering
Zhou Changle
Ningxia university/Institute of physics and electrical engineering
Ge Congle
Ningxia university/Institute of physics and electrical engineering

0 ABSTRACT

The International Aerospace Robot Contest was founded in 1991, the first time in 2012, the establishment of the Asia-Pacific Division in China, with the US Division simultaneously. The purpose of the contest is to promote the development of aerial robot technology, its tasks are the current unmanned aerial vehicle technology cannot be achieved, In accordance with the rules of the game, the air robots need to complete the very challenging task, The 7th generation task is on the UAV Indoor autonomous navigation presents new challenges. According to the task of the competition, the intelligent four-rotor platform is built. The platform uses the onboard processor to realize the upper function of positioning and navigation, and equipped with a variety of sensors, the optical flow sensor and IMU data fusion to achieve the positioning hover, Identification and navigation control through the monocular identification of the target, the use of PID controller for tracking and driving a set of programs, experiments show the feasibility of the program.

1 INTRODUCTION

a) Statement of the problem

During the past time, Our team did some works that acquired information as well as discussion. Lastly, We summarized the following question:

1.How to identify the ground robots;

2.How to control the aircraft to hover,land,turn around etc. ;

3.How to avoid obstacles effectively;

b) Conceptual solution to solve the problem

Question 1: Using OpenMV module to identify the target, because the OpenMV module built-in processing chip can be processed to the good image back to the development board, do not take up the onboard processor resources can be followed by the implementation of the follow-up operation, and OpenMV built A function of the call procedures, the use of simple and convenient. But in the development process found that OpenMV can only use python language for development, python implementation is relatively slow, so this design to give up the use of OpenMV module and OpenCV for C + + programming.

Question 2: In the indoor experiment, the autonomy takeoff and hover process, there will be aircraft cannot hover in the designated location, the use of optical flow sensor positioning.

Question 3: For the obstacle avoidance we use a lidar, the laser radar is based on the time of flight (TOF, time of flight), by measuring the laser time to carry out thedistance d = ct / 2, the laser radar, including the transmitter and the receiver, the transmitter irradiates the target with a laser, and the receiver receives the backlit light. The mechanical lidar includes a mechanical mechanism with a mirror in which the beam rotates so that the beam can cover a plane so that we can measure the distance information on a plane.

c) Figure of overall system architecture

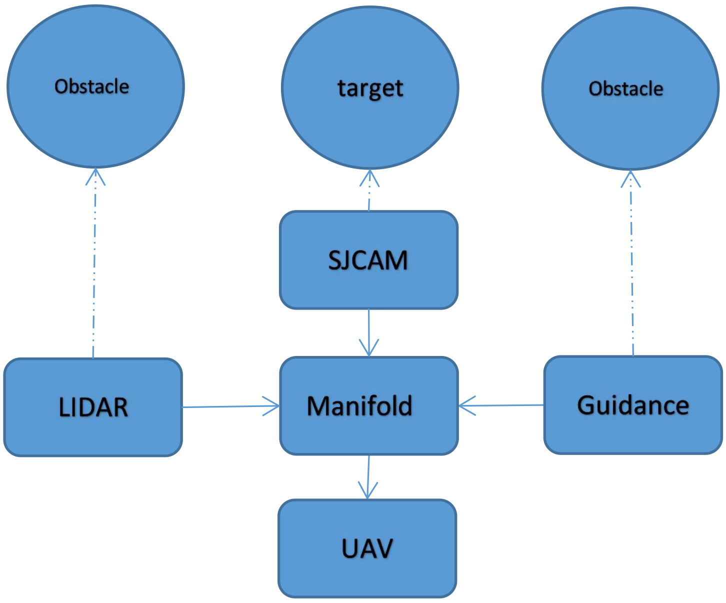
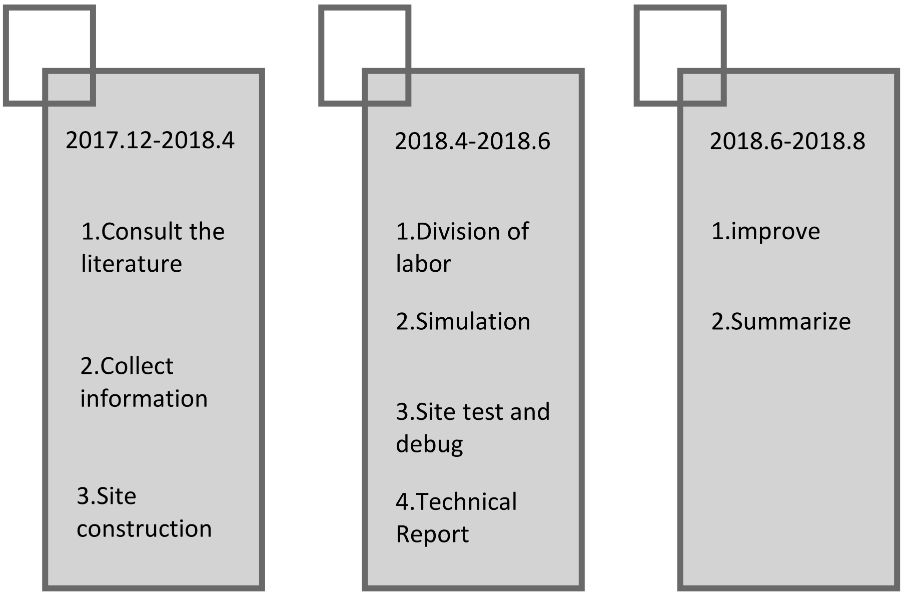


Fig.1 overall system architecture

d）Yearly Milestones



2017.12-2018.4	2018.4-2018.6	2018.6-2018.8
1.Consult the literature	1.Division of labor	1.improve
2.Collect information	2.Simulation	2.Summarize
3.Site construction	3.Site test and debug	
	4.Technical Report	

Fig.2 yearly milestones

2 equipments

We choose DJI M100 as the vehicle platform.The main structural features of the four rotor craft are the basic layout of the axisymmetric, the four rotor crosses are evenly distributed.

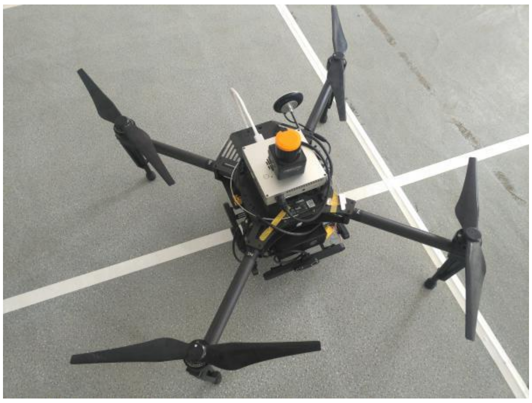


Fig.3 Aircraft system

Our air vehicle mainly contains the rotor, flight stability control equipment(IMU), power supply, landing gear and other components. Landing gear is composed of four groups of support plate, landing gear and the body is screw, nut and spring connection, When the landing gear pressued, the spring was pulled to achieve the role of shock absorption. The landing gear is the bearing part of the weight of the whole body. The landing gear is made of composite material, which has good support stability and vibration absorption performance.

a) Guidance

Fig4. Guidance vision&ultrasonic system

Guidance is a revolutionary visual sensing system. With a powerful processing core, integrated visual cameras, ultrasonic sensors, and the most advanced computer vision algorithms in the world.

Guidance continuously scans the nearby environment and detects obstacles in real time. When used with a DJI flight controller, it can tell your flight system to automatically avoid collision, even at high speeds.

b)Lidar

Fig5. HOKUYO 10lx 2D LIDAR

hokuyo company ust-10lx 2d laser scanning ranging product has 10m , 270 °

measurement range , dc12v / 24v input , 25ms scan time , ip65 protection level , non - contact measurement . high precision , high resolution and wide field of view provide good environment recognition capability for autonomous navigation robot ; Compact design saves installation space , low weight , low power consumption .

c)momocular

Fig6. SJCAM monocular

SJCAM M10 is a momocular camera with a micro-USB,900 mAh,support for 12MP ,10MP,8MP,5MP,3MP,1.3MP video format.

d)Manifold

Manifold is a new generation of heterogeneous processor development boards produced by DJI, From the parameter point of view, The overall performance of Manifold is basically the same as the current mid end smartphone, It is equipped with a frequency 2GHz A15 four core processor and 2GB ram motherboard with Ethernet interface, 2 USB 3 ports, 2 USB 2 ports, 1 HDMI and 1 GPIO interface video interface, You can even install a CPU cooling fan, and the memory chip needs to be purchased and added. Because of the arm architecture, Manifold can run operating systems designed based on the arm architecture, including Debian, Ubuntu, and Android.

Fig7. Manifold embedded computer

3 vision algorithm

Firstly,to maintain UAV stability in air,optical flow (a)is used.To find the irobot,we choose HSV color model(b).to avoid the incoming obstacles,automatic obstacle avoidance system is fatal to make sure the UAV works in the zone we want ,we choose SVM method(d) to recognize the boundary.

a)optical flow

Optical flow refers to the movement of objects, faces or sides in the field, which is produced by the movement of the camera or object, and now is mainly used for the detection and estimation of moving objects, the division of objects and so on. The optical flow sensor collects the surface image of the object at a certain rate, and then analyzes the image digital matrix. Since the two adjacent images always have the same characteristics, by comparing the position change information of these feature points, we can judge the average motion of the surface features of the object. The result of this analysis will eventually be converted into two-dimensional coordinate offset, And in the form of pixel values stored in a specific register, to achieve the detection of moving objects.

In this design, for the adjacent two frames acquired at time t and time, the brightness of the points x, y on the image plane produces an increment, assuming that the brightness is represented, and the velocity of the light flow in the horizontal and vertical component：

$$u = \frac{dx}{dt}, v = \frac{dy}{dt}$$

And then find the optical flow field speed, but only one point is not enough, so the need for multiple points to detect, in order to determine the exact optical flow information. There are two large directions in the calculation of the optical flow: one is the dense light flow, which considers all the brightness information of a region: the second is the thinning flow, which is the use of the characteristics of the information to characterize the whole object, And the optical flow information of the feature point is used as the optical flow information.

The feature-based optical flow algorithm is a better algorithm than the algorithm based on the brightness-based template matching algorithm. The feature-based optical flow algorithm is used to extract the feature points in the image, and the motion information of the feature points is used to characterize the light of the whole image Flow information. Common feature extraction is extracted by Harris corner point, FAST feature point extraction, SIFT feature point and SUFT feature point extraction. After extracting the feature points of the image, it is necessary to calculate the optical flow of the feature points to determine the optical flow field of the image. At present, the more effective algorithm is Kucas-Kanade algorithm. The algorithm is based on the assumption that adjacent small motion, adjacent frame brightness is constant and adjacent frame space consistency. The optical flow equation is established in the image feature point and its neighborhood, With the velocity component representing the light flow in the horizontal and vertical directions, then：

$$\begin{bmatrix} I_{x1} & I_{y1} \\ I_{x2} & I_{y2} \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_{t1} \\ I_{t2} \\ \vdots \end{bmatrix};$$

Fig8. optical flow matrix

Lucas-Kanade algorithm using the principle of least squares method find the minimum:

$$A\bar{u} = \mathrm{b} \Rightarrow A^T A\bar{u} = A^T b \Rightarrow u = (A^T A)^{-1} A^T b$$

Then then to solve the point of the optical flow vector, for each feature point, you can find such an optical flow vector. According to different locations to extract a number of optical flow vector points, and then obtain a relatively accurate optical flow field.

The visual range of the aircraft platform is different from that of the ground robot, and there are six degrees of freedom for the aircraft, namely the translation of the x, y and z axes and the rotation of the x, y and z axes, respectively. For a four-rotor, the translation of the x and y axes is achieved by the pitch and roll angles generated by the rotation of the y and x axes. For this task, because the experimental site is 20m * 20m square site, which are the same length of the side length is 1m grid, through the camera to do visual odometer to estimate the location of the estimated aircraft, because the height of the aircraft has known , So the number of lattices that can be calibrated is constant, and every time you move a lattice coordinate plus 1 method as a visual odometer. The change of the attitude of the aircraft will inevitably affect the image flow field, it is possible that the aircraft does not move, but the optical flow field has produced a lot of changes, will cause a lot of false estimates, so here with the IMU data on the attitude change Caused by changes in the optical flow field made compensation.

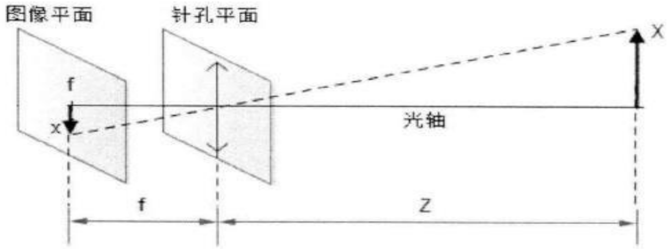


Fig9. imaging diagram

The optical flow field is generated by the motion projection in the three-dimensional world on the image plane. If P is a point in the camera coordinate system, the coordinates are, f is the focal length, Z is the distance from P to the camera, and P is at the plane The projection point can be expressed as：

$$P = \mathrm{f}\,\frac{P}{Z}$$

F is the distance from the pinhole plane to the image plane, and the coordinates of the z-direction of the point P are constant. For the query of the rigid bod kinematics, the relative motion between the camera and the point P can be expressed as:

$$V = -T - w \times P$$

Where W is the angular velocity of the aircraft and T is the translation vector of the motion of the aircraft. In this case, the relationship between the optical flow of the plane P and the velocity of the P in the camera coordinate system is obtained, where K is the velocity of the Z axis.

$$\frac{flow}{\Delta time} = \dot{p} = v - f\frac{ZV - V_z P}{Z^2}$$

The above two formulas are written as follows:

$$\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} = -\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} - \begin{bmatrix} w_y Z - w_z Y \\ w_z X - w_x Z \\ w_x Y - w_y X \end{bmatrix}$$

$$v_x = f\frac{ZV_x - V_z X}{Z^2}, \quad v_y = f\frac{ZV_y - V_z Y}{Z^2}$$

Put these two formulas and $X = x\frac{z}{f}, Y = y\frac{z}{f}$ into it

$$\begin{bmatrix} V_x \\ V_y \\ V_z \end{bmatrix} = -\begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} - \begin{bmatrix} w_y Z - w_z Y \\ w_z X - w_x Z \\ w_x Y - w_y X \end{bmatrix}$$

$$v_x = f\frac{ZV_x - V_z X}{Z^2}, \quad v_y = f\frac{ZV_y - V_z Y}{Z^2}$$

As the plane of the optical flow velocity can be obtained by the target matching or feature point method can be seen as the camera's translation vector is the required amount for the aircraft body rotation angle can be measured by the gyroscope in the IMU, x , Y is the coordinate of the optical flow point in the image plane, which can be obtained from the image. So that W can find the camera's translation vector.

In summary, the realization of the monocular visual odometer based on the optical flow method can be divided into the following steps:

1. Image plane optical flow field calculation, the specific method can be used chess board matching method and feature point based method.

2. Image optical flow vector purification, the specific method can be used in the text, that is, the first to find the average of all optical flow vector, in the exclusion of light away from the average value of the external point, and then find the remaining optical flow vector as the whole image Of the optical flow vector.

3. Attitude compensation and scale reduction, the specific method is the optical flow vector contains the camera rotation caused by the optical component, according to the gyroscope and camera focal length to compensate, while the height of information through the optical flow velocity scale reduction.

4. The velocity of the optical flow is integrated to obtain the motion path of the camera.

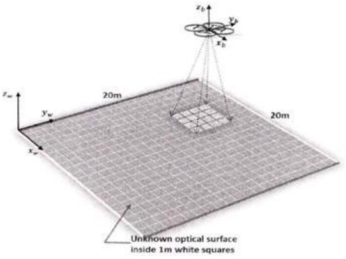


Fig9. visual odometer

When The Quadrotor without GPS in indoor, it will go with the position drift, The first step is to solve hovering problems when the Quadrotor is to accomplish a particular task, that is, in the Uncontrolled quantity, The Quadrotor remains at the current point. For Underactuated systems such as The Quadrotor, the output of the controller is always matched with each barrel roll, pitch, orientations, and the oil port from four control channels, named; the input of the Controller is through kinds of sensors to testing to get the position and the estimated value of the speed.

For the Quadrotor which the component of the height is above the average, so it put to use a simple on-off controller to change the height. also put ordinary P-controller to control the directions.

In our trial, for control the position X,Y, try PD-control first.In the tiral the quadrotor is over speed in Error zero result in the control procedure come to vibration even unstability. To solve this problem, adjust the control strategy ,use the cascade control , which the P-controller in inner race and the PD-controller in outer race; .the sensor of the quadrotor cannot provide the observation in x,y,v_x,v_y direct, so evaluate the parameters is indeed. The evaluate of the speed v_x, v_y is from series Kalman filtering.The evaluate of the position is from integration. As is full designed ,the formula As shown in equation

$$\mathrm{v}_{\mathrm{x}}^{*} = k_p \cdot \left(x^{*} - x\right) - k_d \cdot v_x$$

$$u_\theta = k_p \cdot \left(v_x^{*} - v_x\right)$$

b) Target Identification

To achieve the Ground robot’s Tracking and touch, the first thing is to discriminate the terrain mobile robot ,due to the rules, the Ground robot i-creater is behind the mark plate which in red or green , it’s a convenience for the unmanned aerial vehicle’s discriminate. This system design is using opencv to recognize the image. Opencv, Many common algorithms in image processing and computer vision have been realized, is an open source computer vision library, it comes from series c++ and c-function. The best choice for this target identification is opencv.

The image recognition model is used to RGB (red, green, blue) model and HSV (hue, saturation, brightness) , RGB is widely used in color monitor and color video camera, Our usual pictures are usually RGB models. The HSV model is more consistent with the way people describe and interpret colors, The color description of HSV is natural and very intuitive to people.

HSV color model evolved from CIE color space, It is used in the user intuitive color description method, HVC ball type color solid color with Munsell system is closer. (as shown in Figure HSV, color six pyramids) Only the HSV color model is an inverted cone Liuling, equivalent to only half of the Munsell color solid ball (southern hemisphere), so don't contain black pure color in a color of the top surface of the Liuling cone plane. In the HSV color model six Ling cone, hue (H) at the top surface of the cone plane parallel to the six Ling color, they are around the center axis V rotation and change, Red, yellow, green, cyan, blue, magenta, six standard color are separated by 60 degrees. The color brightness (b) changes along the central axis V Liuling cone from top to bottom, the center shaft top White (v = 1), the bottom is black (v = 0), they said the gray color color department. Color saturation (s) changes along the horizontal direction, closer to the center axis of the Liuling cone color, the saturation is low, the center hexagon color saturation is zero (s = 0), with the highest value of V = 1 coincides, highest saturation of color in the hexagonal frame edge line (s = 1).

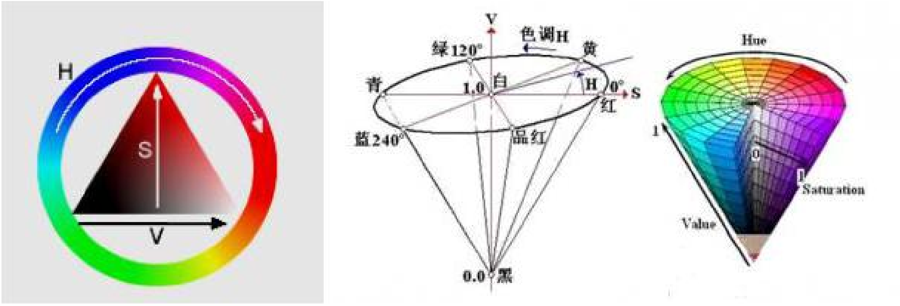
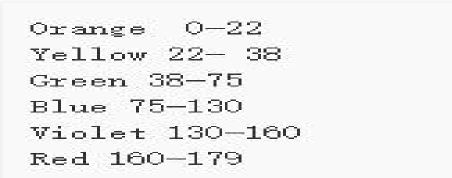


Fig10. HSV colour gamut

Hue (h:hue): the use of angle measurement, ranging from 0 degrees to 360 degrees, to counter clockwise calculated from the red, red and green is 0 degrees, 120 degrees, 240 degrees. They are blue colored yellow: 60 degrees, 180 degrees for cyan, magenta saturation (300 degrees; s:saturation): ranges from 0 to 1, the greater the value, the more color saturation. The brightness (v:value): range 0 (black) to 255 (white).

This will be a HSV model of the RGB steering adjustment range of the various HSV values can be very good to the color you separated for complete identification of the specific color, we know that the H component can represent an object's color, but the value of S and V are in a certain range, because as the representative of s is represented the H by the color white and mixed degree,it’s also said the s is small, the color is more pale, which is more shallow; V represents h represented by the degree of mixing the color and black, also said that the V is smaller, the color is more and more black. After the trail, the value recognition of blue h is from 100 to 140, s and V are 90 to 255. The value of some of the basic colors of H can be set as follows:



Orange	0—22
Yellow	22— 38
Green	38—75
Blue	75—130
Violet	130—160
Red	160—179

Fig11. HSV colour range list

Firstly, histogram equalization of color images is made, Converting RGB to HSV, Gets the two color image of the target color, Then through the combined channel, open and close to operation it,Remove the image noise, some connected connected domain, to identify the target.

The identification and tracking of the ground robot is very important in this mission. It requires high precision of tracking, and the ability to recognize and predict the position of the irobot when obstructed by obstacles. So we use the camshift algorithm based on the Kalman filter algorithm to track the target, and achieve a good tracking effect.

Camshift algorithm is an improved algorithm on meanshift algorithm, also called mean shift algorithm. It is an iterative step, that is, to calculate the offset mean of the current point, to move the point to the offset mean, and then to use it as a new starting point to move on. Until the final condition is met.

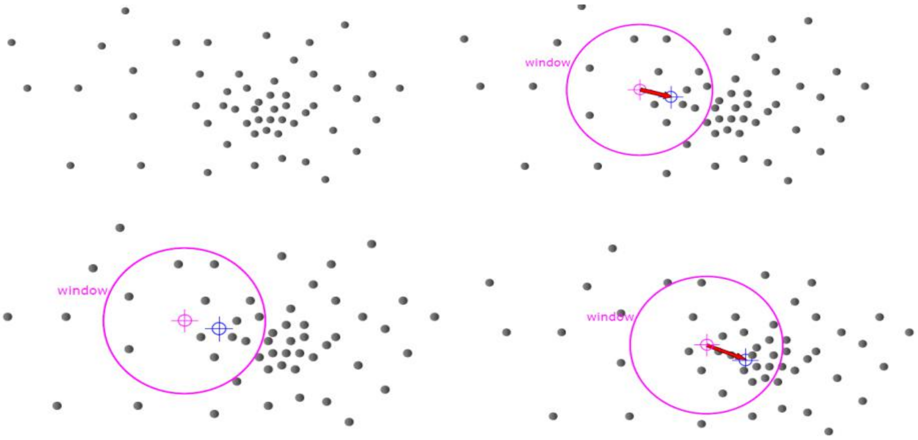


Fig16. Mean shift process of camshift algorithm

The kernel function for the probability density of the characteristics of the required region is estimated as:

$$\hat{f}(x)=\frac{\sum_{i=1}^{n}K(\frac{x_i-x}{h})\omega(x_i)}{h^d\sum_{i=1}^{n}(x_i)}$$

For $f(x)$ gradient function $\nabla f(x)$ is estimated as:

$$\nabla\hat{f}(x)=\frac{2\sum_{i=1}^{n}(x-x_i)k'(\left\|\frac{x_i-x}{h}\right\|^2)\omega(x_i)}{h^{d+2}\sum_{i=1}^{n}\omega(x_i)}$$

The center of camshift recognition is always moving towards the direction of gradient estimation function $\nabla\hat{f}(x)$, that is, moving towards the direction of the maximum probability density of the target feature, so the tracking of the target is realized. The following is the flow chart of the implementation of the camshift program:

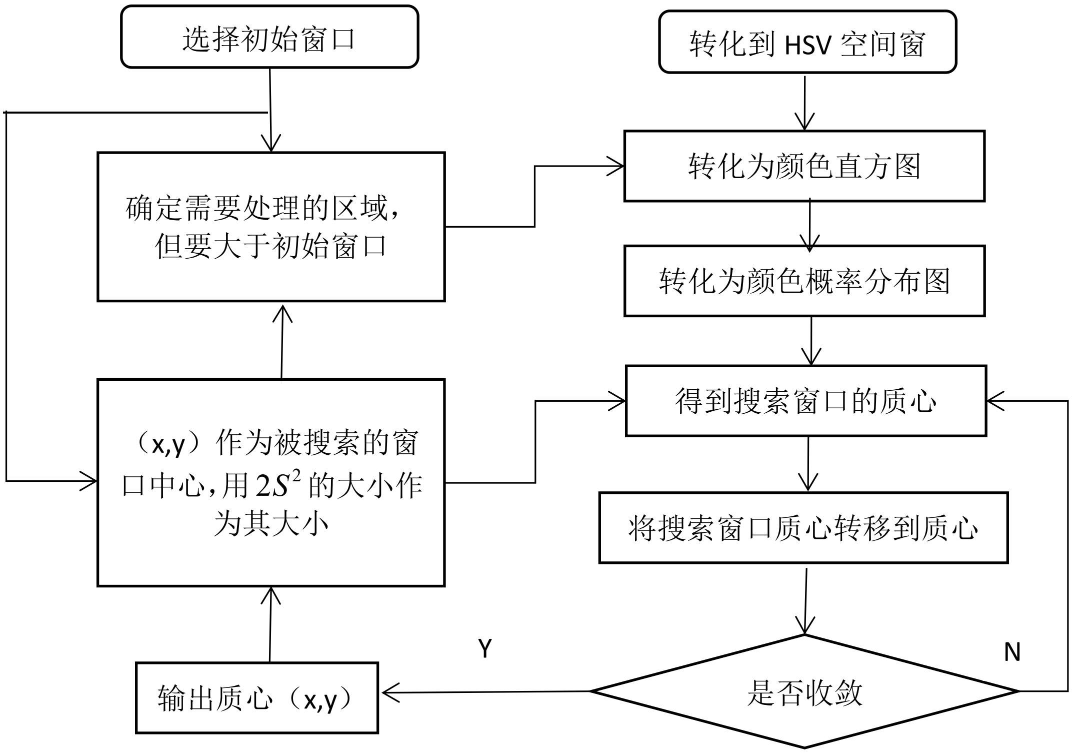

选择初始窗口
确定需要处理的区域，但要大于初始窗口
（x,y）作为被搜索的窗口中心，用$2S^2$的大小作为其大小
输出质心（x,y）
转化到 HSV 空间窗
转化为颜色直方图
转化为颜色概率分布图
得到搜索窗口的质心
将搜索窗口质心转移到质心
是否收敛
Y
N

Fig17. Program implementation flow chart of camshift algorithm

However, there are obstacles in the actual competition, which may block the sight of the aerial robot, so in order to ensure the effect, we add the Kalman filtering algorithm based on camshift.

The Kalman filter proposed by Professor Kalman can predict the state of moving target at the next moment. Kalman filter uses state equation and measurement data to estimate the state of the system. The target tracking problem is abstracted into a discrete control process system, whose state equation and measurement equation are:

$$X(k) = AX(k-1) + BU(k) + W(k)$$

$$Z(\mathrm{k}) = HX(k) + V(k)$$

The Kalman filter combines covariance matrix to estimate the optimal position of the system. Kalman filter consists of five key steps, which can be used to predict sheltered targets, and the results are shown in the figure:

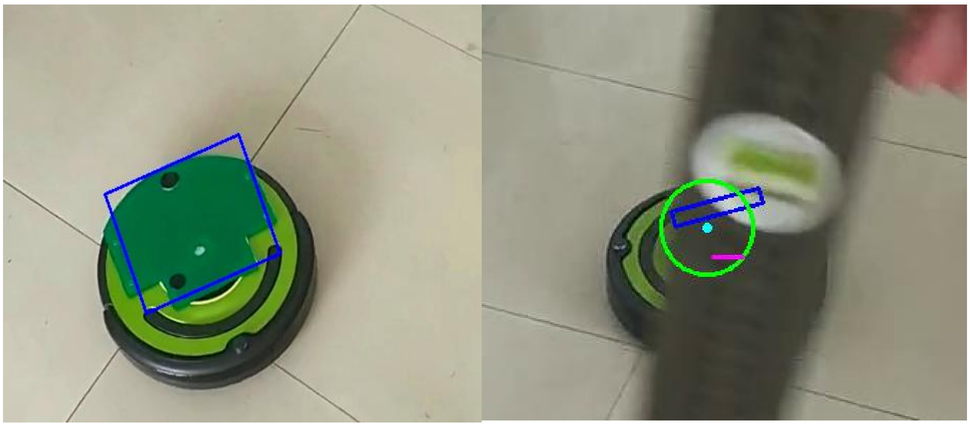


Fig18. KF effect comparison

c)Threat Avoidances

Automatic obstacle avoidance system is an important safety guarantee for the successful completion of the flight mission, which reflects the intelligence and safety of the aircraft to a great extent. Obstacle detection technology in the automatic obstacle avoidance system has a vital position. We use infrared radar to monitor environmental obstacles to real-time obstacle avoidance. Infrared environmental radar is based on TOF (Time of Flight) principle, with a unique optical, electrical, design, in order to achieve stability, precision, high sensitivity and high-speed distance measurement. TOF is the abbreviation of Time of Flight technology, that is, the sensor sends modulated near-infrared light, reflecting the object after reflection, the sensor by calculating the light transmission and reflection time difference or phase difference, to convert the distance of the scene to be shot to produce Depth information.

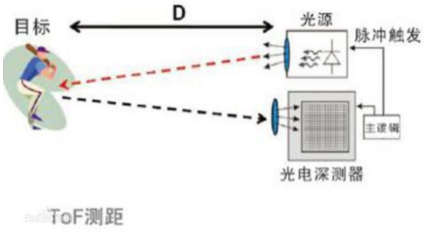


Fig12.TOF algorithm

d)boundary detection

A four-rotor drone is required not to cross the boundary for more than 10 seconds, otherwise it is considered a foul, which requires the UAV to be able to accurately identify the boundary. In boundary detection, support vector machine (SVM) method is a very classical machine learning method. It has strong advantages in small sample size and nonlinear pattern recognition, such as global optimization and strong applicability. Simple structure and short learning and prediction time. In this paper, the svm method is used to detect the boundary. The key of the detection is the training of the svm classifier, and the premise of the training is the selection of the sample feature, and the hog feature is selected to describe the sample. In order to extract the hog features of the local region of the image containing ground targets, the image is first divided into small connected areas, which are called "cell units", and then the histogram statistics of the gradient direction of pixels in cell are obtained. Finally, the information is connected together to form a high dimensional feature. By its extraction process as shown in the figure

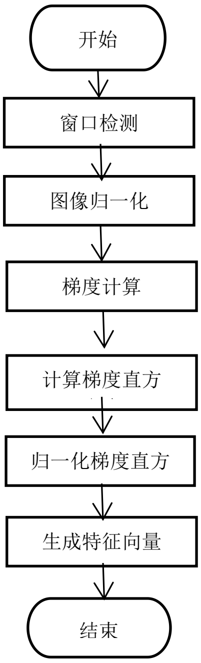


Fig13.flow path of boundary detection

The collection of hog features was completed, and the classification and training of svm samples began. The principle of classifying the planes of two groups of samples is essentially to find a hyperplane which can maximize a certain value, which is the minimum distance between the hyperplane and all the training samples.

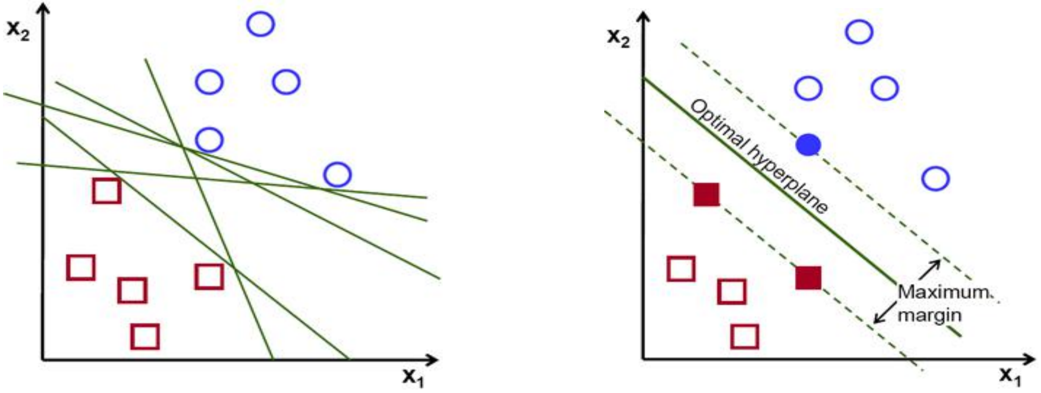


Fig14. diagrammatic sketch of SVM

We can use $f(x) = \beta_0+\beta\mathrm{T}$ to deliver the hyperplane ,β represent for weight vector, β_0 is the bias.our main target is to calculate β and β_0。

Generally $|\beta_0+\beta\mathrm{T}x|=1$ is used to express optimal hyperplane,x represent the points that are close to the hyperplane,called support vector，the smallest distance from x to（β， β_0）is：

$$\text{distance}=\frac{|\beta_0+\beta\mathrm{T}x|}{\|\beta\|}\approx\frac{1}{\|\beta\|}$$

The interval is M, which is two times as close as the closest distance:

$$\text{M}=\frac{2}{\|\beta\|}$$

The final maximization m is transformed with additional constraints into a minimization function $L(\beta)$, The condition that a hyperplane classifies all training samples correctly:

$$\min_{\beta,\beta_0} L(\beta)=\frac{1}{2}\|\beta\|^2 \ \textit{subject to}\ y_i(\beta^{\mathrm{T}}x_i+\beta_0)\geq 1\ \ \forall i$$

y_i represents the class marking of sample，This is a Lagrange optimization problem. By Lagrange multiplier method,the optimal hyperplane can obtain the weight vector β and β_0。

These are the principles of SVM, and the SVM program flow chart is as follows:

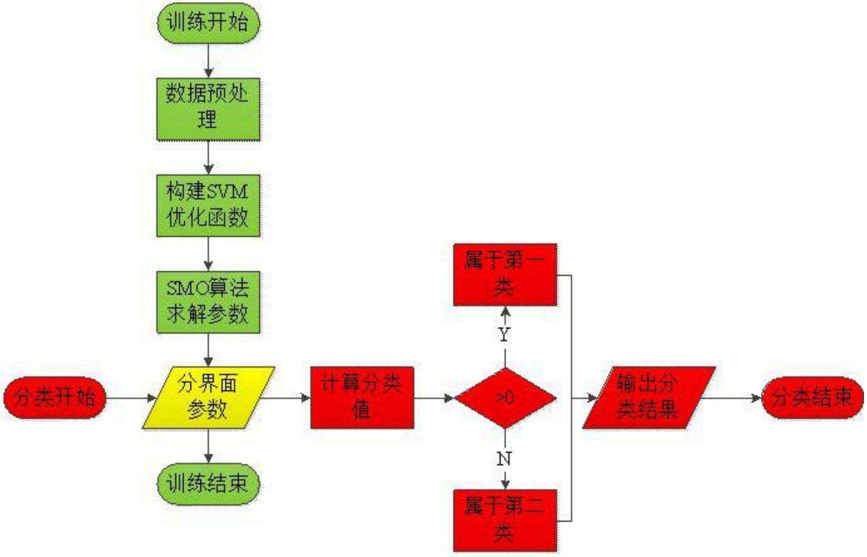


Fig15. flow path of SVM

4 UAV control

a)ROS

The ROS is controlled by the nodes between nodes are connected, the program is written in the ROS node, the node as a talker data port and flight connection as a listener, to realize the control program through the ROS master node. Each communication port requires a launch file to open the launch of listener, to be able to communicate.

The establishment of communication is the reference flight and ground station mavlink communication protocol, mavlink protocol was first proposed by the Zurich Federal Institute of computer vision and geometry of the experimental group Lorenz Meier was released on 2009, and follow the LGPL protocol. Mavlink protocol is a higher level of open source protocol on the basis of serial communication, mainly used in MAV (micro aerial vehicle) communication on.Mavlink is for small aircraft and ground station (or other vehicle) who often used to make data communication for transmitting and receiving rules and add a check (checksum). The mavlink based and ROS operating system, and install the relative dependencies, can form new communication protocols, namely mavros.

The preparation of ROS node in the first set of mavros each of the "message", and then send complete instructions for flight control, to achieve the purpose of control.

b)DJI SDK

b1)SDK structure

The DJI Onboard SDK is an open source software library that enables computers to communicate directly with a selection of DJI aircraft and flight controllers over a serial interface. The SDK gives access to aircraft telemetry, flight control and other aircraft functions, meaning a developer can use the SDK to attach their own computing device onboard an aircraft and use it to control flight.

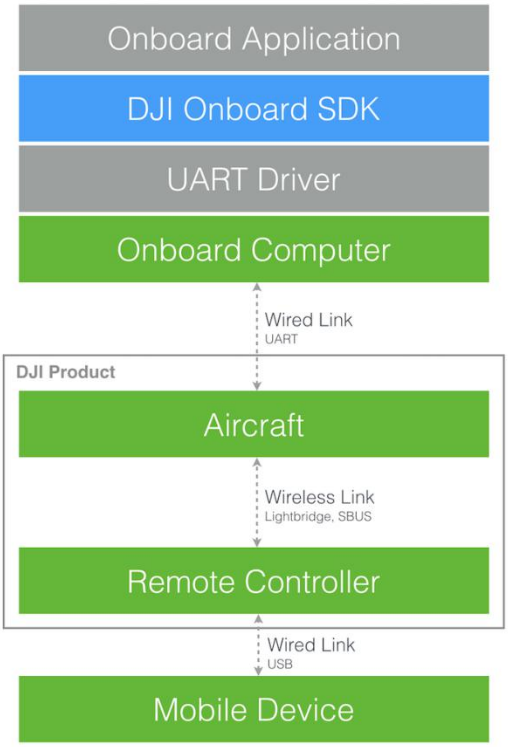


Fig19. The connection between DJI SDK and aircraft

b2)API Hierarchy

An application accesses the DJI product's capabilities through the main Vehicle class in the DJI Onboard SDK.

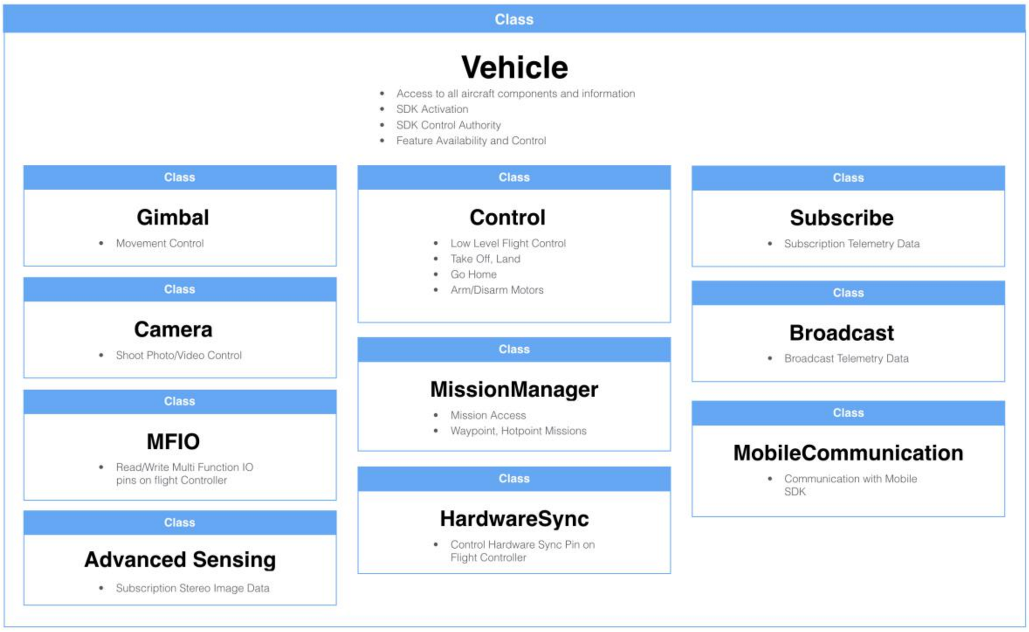

Class
Vehicle
Access to all aircraft components and information
SDK Activation
SDK Control Authority
Feature Availability and Control
Class
Gimbal
Movement Control
Class
Camera
Shoot Photo/Video Control
Class
MFIO
Read/Write Multi Function IO pins on flight Controller
Class
Advanced Sensing
Subscription Stereo Image Data
Class
Control
Low Level Flight Control
Take Off, Land
Go Home
Arm/Disarm Motors
Class
MissionManager
Mission Access
Waypoint, Hotpoint Missions
Class
HardwareSync
Control Hardware Sync Pin on Flight Controller
Class
Subscribe
Subscription Telemetry Data
Class
Broadcast
Broadcast Telemetry Data
Class
MobileCommunication
Communication with Mobile SDK

Fig20. DJI OSDK API structure

The Vehicle class is used to administer the SDK (e.g. SDK activation), and also provides the objects to control and receive state information from the product.

5 OPERATIONS

a) Flight Preparations

IARC7 mission requirements have each side for 20 m square area in an indoor, area is one of the side length is 1 m small square, as shown in the figure below. Due to the venue is similar to the camera calibration of Zhang's, so the camera calibration adopts the method.

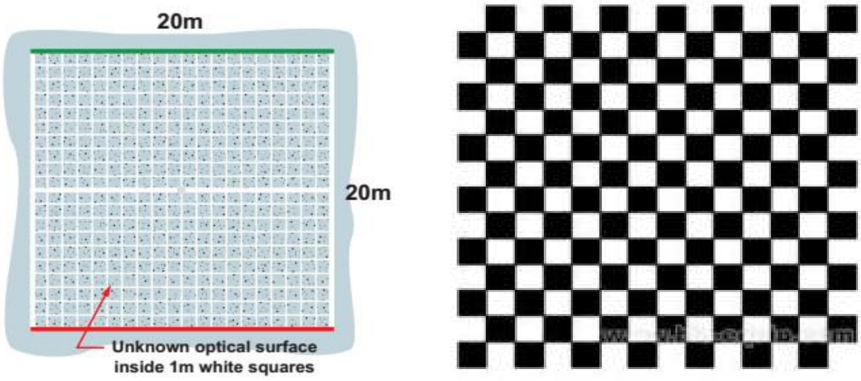

20m
20m
Unknown optical surface inside 1m white squares

Fig21. the Zhang's calibration and the checkerboard

Here "calibration " also called "the Zhang's calibration", in 1998,The method is came up with the professor,Zhang's,single face checkerboard camera calibration method, the Zhang's calibration method has been widely used as a kit or encapsulated function.

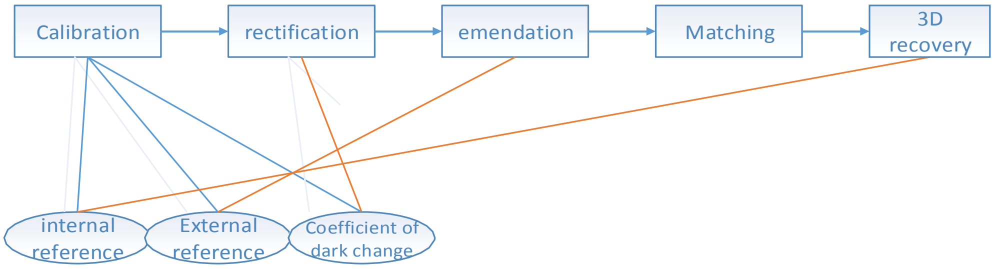

Calibration
rectification
emendation
Matching
3D recovery
internal reference
External reference
Coefficient of dark change

Fig22. the flow path of Zhang's calibration

It is clear from the figure that the calibration of the internal parameters, external parameters and distortion coefficient, is the need for picture correction, camera correction and 3D recovery basis, there is no good calibration, will not be able to complete 3D reconstruction.

Homography: A projection map defined as a plane to another plane in computer vision.In fact, it is the spatial plane of the three-dimensional point and the camera plane two-dimensional point mapping, Professor Zhang Zhengyou in the paper <A-flexible-new-technique-for-camera-calibration> talked about the camera model, So easy to get:

$$s\tilde{m} = A\begin{bmatrix} R & t \end{bmatrix}\tilde{M}$$

Where the homogeneous coordinates of the three-dimensional point M of the spatial plane represent the coordinate points of the world coordinate system（X,Y,Z,1）, The homogeneous coordinates of m represent the pixel coordinates (u, v, 1) of the image plane, R denotes the rotation matrix, t denotes the translation matrix, S denotes the scale factor, and A denotes the internal reference of the camera, as follows:

$$A = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

For the matrix A, where $(u_0.v_0)$ is the principal point position, α and β are the dimension factors of the image in the u axis and the v axis. The γ parameter describes the twist factor of the two coordinate systems (that is, the pixel in the XY direction on the scale deviation). In order to facilitate the operation, the scale factor does not change the coordinate value for homogeneous coordinates.

Since the calibration is planar, we can construct the world coordinate system on the plane of Z = 0 and then perform the homography calculation. Let Z = 0 convert the above formula into the following form:

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A\begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix}\begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = A\begin{bmatrix} r_1 & r_2 & t \end{bmatrix}\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

So, the relationship between M on a template and the point m on its image, then there are now:

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = H\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

H is the homography matrix, which is described here is the space between the three-dimensional point and the camera plane of the relationship between the three-dimensional point. And the camera plane can be obtained through the image processing, the space of the three-dimensional point can also be obtained by doing a good board, according to the location of different locations taken by multiple photos, you can get a different H matrix, by the above formula Analysis of H is a 3x3 matrix, and there is an element as a homogeneous coordinate, so that H has 8 unknowns to be solved. Write H as a matrix of 3x3:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

If the third row of the third column of H elements into 1, then the other 8 can be expressed by, so by eight degrees of freedom.

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

From the above matrix we can find two equations with eight unknowns, so we need at least four points to find H.

②Use the constraints to solve the internal reference matrix

$$\begin{bmatrix} h_1 & h_2 & h \end{bmatrix} = \lambda A\begin{bmatrix} r & r & t \end{bmatrix}$$

H matrix is the union of internal and external parameters, we finally get the internal reference and external parameters, through the above formula can be two constraints. 1, r_1, r_2 orthogonal: $r_1r_2 = 0$.2, the rotation vector is modulo 1, that is, $| r_1 | = | r_2 | = 1$. So you can r_1, r_2 replaced by h_1, h_2 and A were combined to express. $r_1 = h_1A^{-1}$, $r_2 = h_2A^{-1}$. Can get:

$$\begin{aligned} h_1^T A^{-T} A^{-1} h &= 0 \\ h_1^T A^{-T} A^{-1} h_1 &= h_2^T A^{-T} A^{-1} h_2 \end{aligned}$$

In the formula, h_1, h_2 is solved by the soliton, the unknown quantity only the internal matrix A. The inner array A contains five parameters: α, β, u_0, v_0, γ. If you want to completely solve the five unknowns, you need three single should of the matrix. The three singularity matrices can produce six equations under two constraints. So that you can solve all the five internal reference. The three singularity matrices are obtained by passing three checkers from different directions.

③*Based on the internal reference matrix*

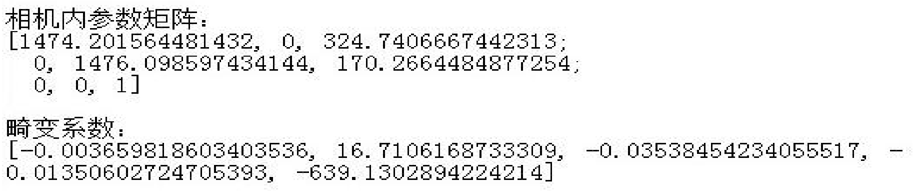
Through the above calculation, you can get the camera's internal array. For external array, we can achieve the following formula:

$$\begin{bmatrix} \mathrm{h}_1 & h_2 & h_3 \end{bmatrix} = \lambda A \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}$$

To simplify:

$$\begin{aligned} r_1 &= \lambda A^{-1} h_1 \\ r_2 &= \lambda A^{-1} h_2 \\ r_3 &= r_1 \times r_2 \\ t &= \lambda A^{-1} h_3 \end{aligned} \quad 其中\lambda = 1/\|A^{-1} h_1\| = 1/\|A^{-1} h_2\|$$

The results of the experiment calibration are shown below：



```
相机内参数矩阵：
[1474.201564481432, 0, 324.7406667442313;
  0, 1476.098597434144, 170.2664484877254;
  0, 0, 1]

畸变系数：
[-0.003659818603403536, 16.7106168733309, -0.03538454234055517, -
0.01350602724705393, -639.1302894224214]
```

\

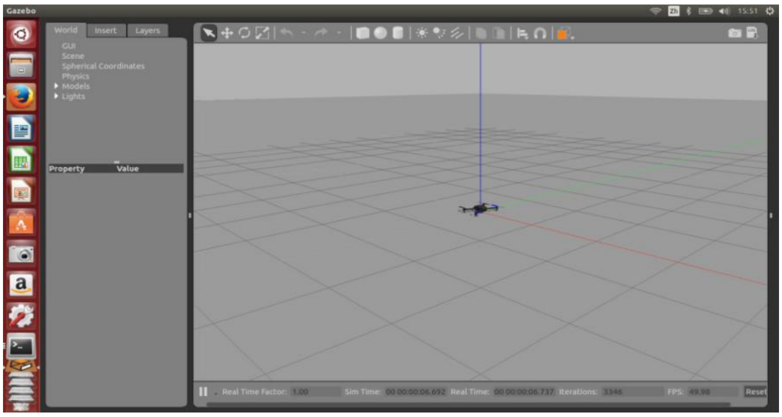
After calibration of the site, through the visual control below, after each passing a grid, the coordinates plus 1, and through the visual capture system to estimate the movement of the ground robot trajectory. According to the calibration results, the UAV flew to the designated area to drive the ground robot.

b)Safety

For remote control switch and the procedure of problem, the design set up the hand of the remote control automatic switch block, sets the five channels to offboard and manual mode, after the experiment and testing, free flight mode switch, and once the program is written into the program out of control will automatically be hovering in the current point and switch to manual mode, waiting for the remote control instruction.

c)Modeling and Simulation

Considering the security of experimental equipment, as well as to reduce the risk.Therefore, we first adopt is a computer simulation, and then transplanted into the true experiment operation on board, we use the simulation software of gazebo, installed first open gazebo will download the built-in model (see chart).Then install protobuf library, as a gazebo interface tools, install DJI M100 native firmware, and ros and gazebo interface tool, and establish the basic urdf model, this is iris four rotor as the basic model.



Real Time Factor: 1.00 Sim Time: 00 00:00:06.692 Real Time: 00 00:00:06.737 Iterations: 3346 FPS: 49.98

Fig23. Gazebo UI

c) Testing

d) ①Laboratory equipment components

Flight control	DJI M100
Airborne plate	DJI Manifold
laser radar	HOKUYO 10lx
Ultrasonic module	Guidance
camera	SJCAM
Luminous flux	Guidance

②Experimental site

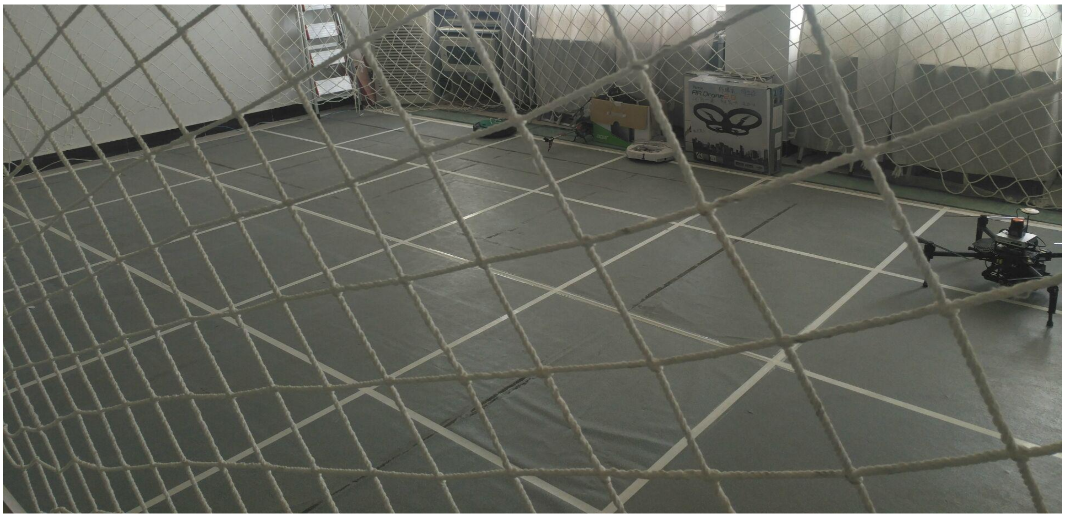


Fig23. experimental site

③Experiment

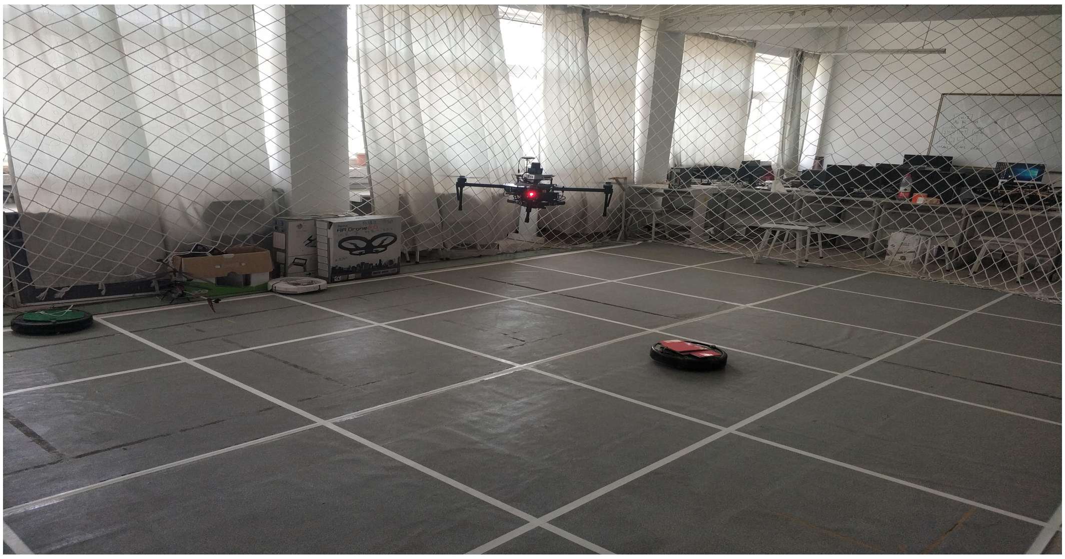


Fig24. experiment procedure

④Experimental results

After the early preparation of the experiment, as well as computer simulation. By field flight test at the experimental site, we have reached the effect that the air robot can drive the ground robot. But in the drive in the path of the optimization needs to be improved.

6 CONCLUSION

The system design for robot contest seventh generation task, the international air aircraft in an open environment for the game without auxiliary scene in the positioning and navigation, aircraft and ground moving target the interaction of these two technical requirements.Mainly discussed the construction of the intelligent vehicle and, based on multi-sensor vehicle positioning and hover, monocular target recognition with the PID control of the aircraft motor to drive strategy, combined with ilf camera visual positioning method of the experimental area for calibration, the experiment preliminary certified the rationality of the system, basic meet the experimental requirements.