Technical Report of Civil Drone for the IARC 2019

Zheng Tan
Civil Aviation University of China

ABSTRACT

Compared with the previous 7 missions, the Mission 8 of IARC has three new challenges: multi-UAV collaborative ability, non-electronic human-computer interaction ability and obstacle avoidance ability to mobile targets. In this paper, assignment problem model is used to solve the interaction ability of different aerial robots when they perform different tasks in the same airspace; voice recognition function is realized by hardware unit with core chip LD3320; image acquisition and mosaic is realized by OpenCV vision library and ROS system. The experimental data analysis proves the rapidity and superiority of each module's function.The report describes the UAV system designed by our team for the eighth generation mission and the UAV's approach from perception to cognition.

1 INTRODUCTION

1.1 Problem Description

Mission 8 requires a team member to work with a team of UAVs within 8 minutes to retrieve a specific item from a 28-by-15-meter area patted by 4 enemy drones. Enemy drones could attack the contestant with a laser beam (once every 5 seconds). If the contestant is illuminated 10 times, the game will be stopped. Our UAVs must have adequate safety measures (such as putting propellers in a protective cover) and can only be controlled by non-electronic means such as gesture or voice commands. Participants retrieve specific items from four boxes with combination locks located in the venue. Part of the QR code is placed next to each box. The four parts of the QR code should be correctly assembled and scanned to obtain the final digital password. This mission requires at least three new behaviors of UAVs, namely, non-electronic man-machine interaction ability, precise positioning and navigation in indoor unmarked scenes and autonomous coordination ability of UAVs. So our drones have to be smart enough to make the right decisions at the right time.

1.2 Solutions to Problems

1.2.1 Speech Recognition Technology

The purpose of speech recognition is to enable robots to understand human language and respond correctly to human intentions [1]. Speech recognition can be roughly divided into two stages: 1) Training stage. The acoustic model is obtained by modeling. 2) Identification stage. Search algorithm is used to obtain the optimal solution. The overall process is shown in *Figure 1* below:

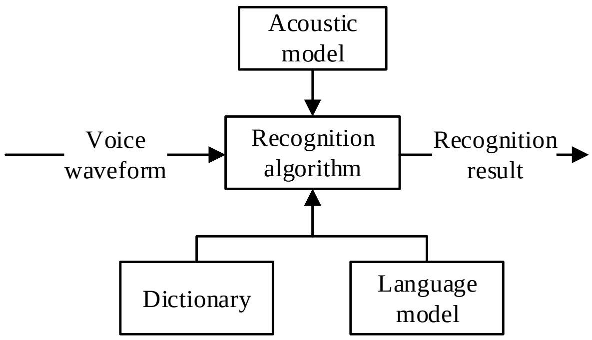

Figure 1. Overall flow of speech recognition.

The mathematical model of weight vector in speech recognition is shown as follows:

$$W^* = \arg\max_{W} P(\mathrm{W} \mid \mathrm{X}) = \arg\max_{W} \frac{P(\mathrm{X} \mid \mathrm{W})\mathrm{P}(\mathrm{W})}{P(\mathrm{X})} = \arg\max_{W} P(\mathrm{X} \mid \mathrm{W})\mathrm{P}(\mathrm{W}) \tag{1}$$

Where, X represents speech signal and W represents text sequence. $P(\mathrm{X} \mid \mathrm{W})$ represents the language model, the probability that the sequence of words is considered A sentence; $\mathrm{P(W)}$ represents the acoustic model, the probability that the sequence of words will be translated into speech signals.

1.2.2 Wireless Communication Technology

We use NRF24L01 module to complete man-machine communication and multi-machine communication, and frequency hopping communication mode is proposed, whose principle is as follows: To complete the communication between 1 primary node and 10 secondary nodes, 10 secondary nodes need to be set as different channels, while the primary node can complete the communication by setting corresponding target channel. Only one primary node can communicate with a specific node at a time.

1.2.3 Multi-drone task assignment

The process of *Multi-drone* task assignment is as shown in *Figure 2*:

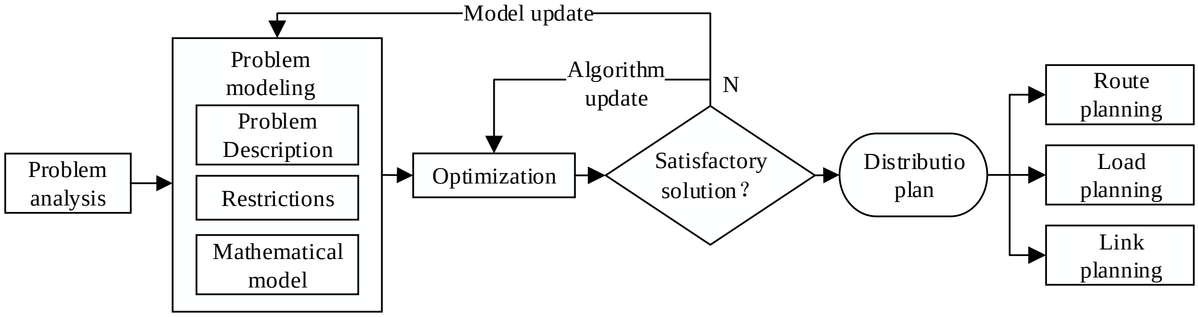

Figure 2. Multi-drone task assignment process.

Adopt the assignment model and algorithm as the solution:

$$
\begin{cases}
x_{ij} = \begin{cases} 1, \text{ When No.} i \text{ UAV is assigned task } j \\ 0, \text{ otherwise} \end{cases} \\
\sum_{i=1}^{n} x_{ij} = 1, j = 1, 2, \cdots, \mathrm{n} \\
\sum_{j=1}^{n} x_{ij} = 1, i = 1, 2, \cdots, \mathrm{n} \\
\min C = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij}
\end{cases} \tag{2}
$$

For *n* UAVs and *n* tasks, c_{ij} represents the cost of doing task *j* for UAV *i*, x_{ij} represents the decision state of whether task *j* is assigned to UAV *i* with the value of 1 or 0 [2][3], C represents the total cost of the current task.

Where constraint condition indicates that the a task can only be completed by one UAV, and a UAV can only complete one task. The feasible solution satisfying the constraint conditions can also be described as a matrix:

$$
\mathbf{X} = (x_{ij})_{n \times n} = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{nn} \end{bmatrix} \tag{3}
$$

The transformation steps of variable matrix are as follows :

(1) Transform the coefficient matrix of the assignment problem so that at least one element of each row and column is "0".

①Let the coefficient matrix of each row of elements to subtract the smallest element of the row;

②Let each column of the coefficient matrix subtract the smallest element of the column.

(2) From the first row, if there is only one zero element in the row, parenthesis the zero element, and draw a line to cover the column of the zero element in the parenthesis. If there is no zero element in the row, or if there are more than two zero elements in the row (those crossed out are not included), proceed to the next row and proceed to the last row.

(3) Process the column elements using the method described in (2) and obtain the optimal solution.

2 UAV SYSTEM DESIGN SCHEME

2.1 Configuration Instructions

2.1.1 The UAV Configuration

We chose two independent four-rotor UAV platforms, as shown in *Figure 3*, which are made of carbon fiber tubes, carbon fiber sheets and aluminum alloy connectors. The materials we use are as strong and light as possible, which can save power and increase the flight time of the UAV.

Figure 3. UAV platforms.

The accessories and parameters of the four-rotor UAV are shown in *TABLE 1*:

TABLE 1. ACCESSORY PARAMETERS OF UAV

Accessories	Parameters
The motor	V350 kv:580
The propeller	1238 carbon fiber propeller
electronic speed controller	30A electronic speed controller
The battery	8000mah 25C
Flight control	Pixhawk

2.1.2 Other Hardware Configuration

Other hardware equipments are empty end computer, camera, laser radar, optical flow sensor, etc. Hardware structure of the system is shown in *Figure 4*. The empty end computer adopts Intel's NUC microcomputer, which is mainly used for data processing and programming. Lidar is a two - dimensional Lidar with obstacle avoidance as its main function. The camera adopts wide-angle USB camera, which is mainly used to identify two-dimensional code; The optical flow sensor adopts px4flow sensor, which is mainly used for navigation and positioning of UAV. NRF24L01 wireless module is used for wireless transmission of voice commands; The ground stations monitor the UAV's flight status in real time. The remote control is mainly used to switch the flight mode of the UAV and prevent the UAV from landing control in case of emergency.

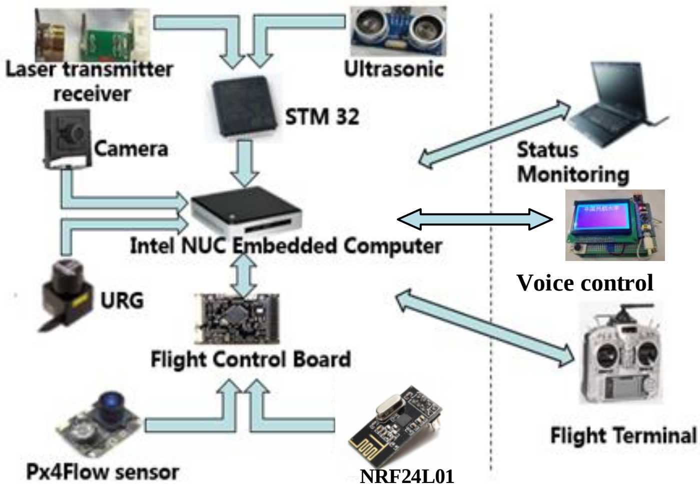

Figure 4. Hardware structure of the system.

2.2 Flight Control System

2.2.1 Location of UAV Based on Improved Optical Flow Method

Optical flow sensor and inertial navigation are used for UAV positioning. Therefore, we propose an improved optical flow method for UAV positioning. Firstly, we use median filter to filter common noise. At the same time, the effect of light intensity on optical flow is reduced by progressive method. Then we introduce the Hessian matrix to eliminate the outliers, and use the three-step search block matching method to calculate the optical flow. Finally, the discrete Kalman filter is used to fuse optical information and rate gyroscope information to estimate the UAV's altitude, attitude and speed and to perform UAV positioning. The UAV positioning flow chart using the improved optical flow method is shown in *Figure 5*.

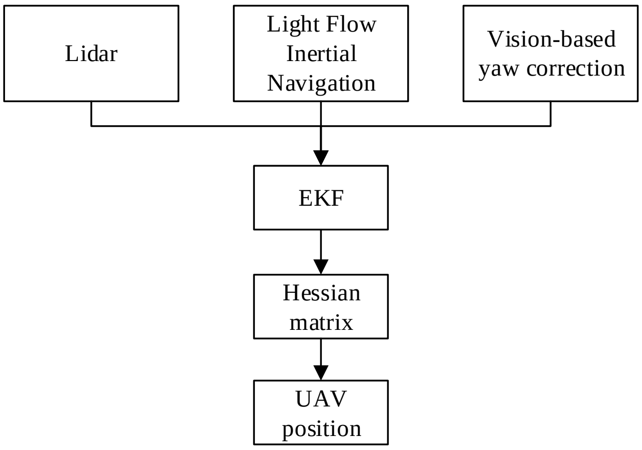

Figure 5. Location flow chart based on improved optical flow method.

2.2.2 THE OVERALL ARCHITECTURE OF THE SYSTEM

According to the characteristics of the eighth generation task, the unmanned aerial vehicle system is divided into three levels: perception level, decision level and behavior level. The overall framework of the system is shown in *Figure 6*.

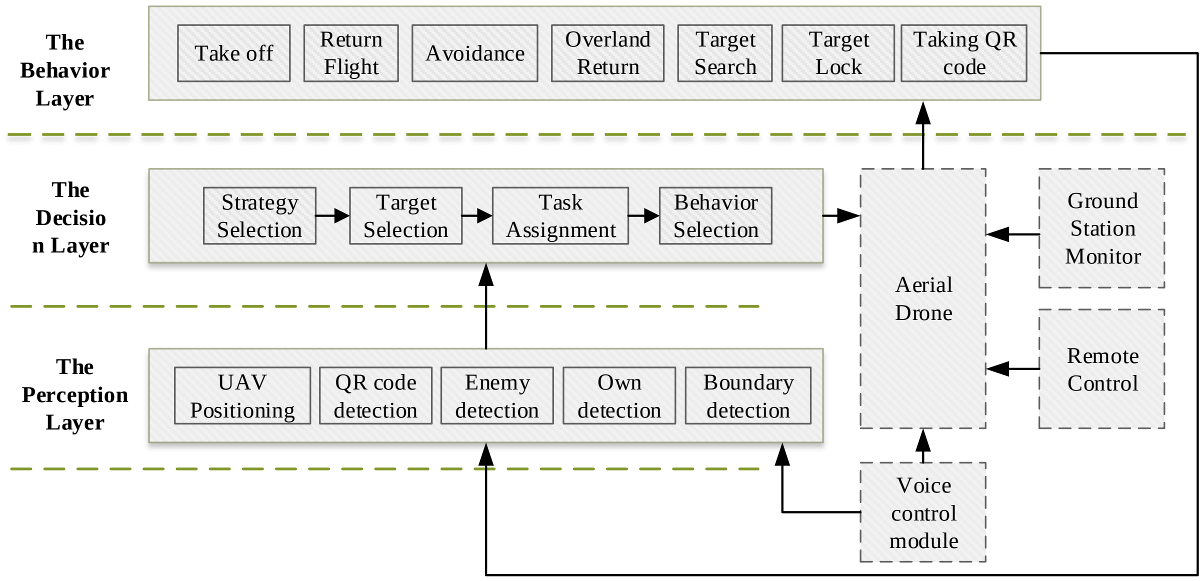

Figure 6. Architecture diagram of the system.

Perception layer main function is the positioning of the UAV itself, QR code detection, sentry drones, their detection and edge detection. The data obtained by the perceptual layer are preprocessed and transmitted to the decision-making layer. In the sequential decision-making of the decision-making layer, the strategies, control objectives and control behaviors adopted are finally determined according to the situation. The output of the decision-making layer is passed to the behavior layer. The behavior layer controls the UAV to execute corresponding behaviors according to the decision results.

3 MISSION SOLUTIONS

3.1 Ground QR code perception

The wide-angle camera installed on the UAV is used to realize the perception of two-dimensional code and environmental information on the ground, as shown in *Figure 7*. When the UAV flies over the QR code, the QR code image is converted into grayscale image and binarization processing by using the robot operating system (ROS) and OpenCV vision library. Then, search the maximum and sub-large squares in the graph. When the centers of the two squares coincide and the number of interior angles is 4, the image will be saved to the database. After that, Angle adjustment and region segmentation are carried out on the collected images. After obtaining the four-part incomplete two-dimensional code, the code is pieced together according to the arrangement rules. Finally, the assembled two-dimensional code is recognized and the password number is obtained.

Figure 7. Wide-angle camera.

3.2 Voice control module

3.2.1 Overall scheme and architecture of modules

The core processing unit of the module is STM32F103C8T6, a 32-bit processor based on the ARM cortex-m3 kernel. This module takes the dialog management unit as the center and realizes the speech recognition function through the hardware unit with the LD3320 chip as the core.

The basic process of speech recognition is shown in *Figure 8*. Speech recognition consists of two stages: training and recognition. Whether training or recognition, it is necessary to preprocess input speech and extract features. The specific work in the training phase is to input the training speech several times through the user, obtain the feature vector parameters through preprocessing and feature extraction, and finally achieve the purpose of establishing the reference model library of the training speech through feature modeling. In the recognition stage, the main work is to compare the similarity measurement between the input speech feature vector parameters and the reference model in the reference model library, and then output the input feature vector with the highest similarity as the recognition result.

According to the object of recognition, The existing speech recognition technology can be divided into two categories: specific person recognition and non-specific person recognition.

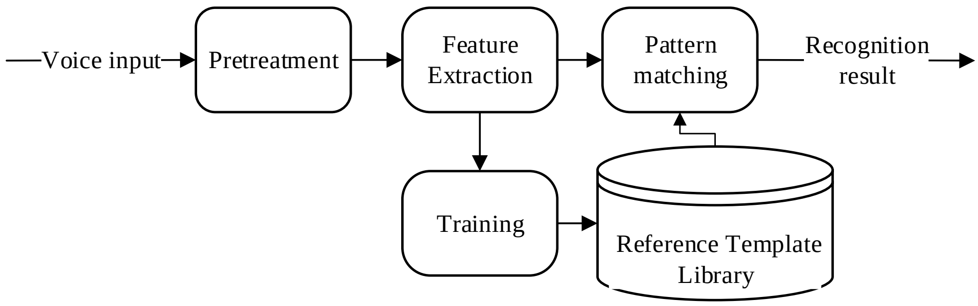

Figure 8. Basic process of speech recognition.

3.2.2 Hardware circuit design

As shown in *Figure 9*, the hardware circuit mainly includes the main control core and the speech recognition part transmitted to the main controller in parallel. After processing, the main controller will send command data to USART, which connects NRF24L01 module to realize wireless communication with the UAV.

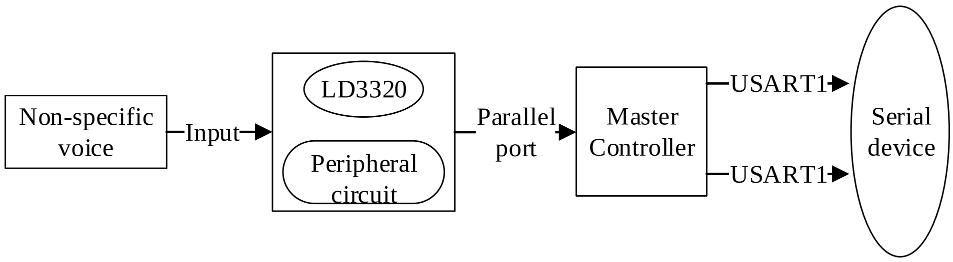

Figure 9. Overall architecture of hardware circuit.

3.2.3 Speech recognition program design

The design of speech recognition program refers to the LD332X development manual [4]. In this paper, interrupt mode is adopted. Its workflow is divided into general initialization -- speech recognition with initialization -- writing recognition list -- start recognition -- response to interrupt.

(1) General initialization and speech recognition initialization. In the initialization program, the main task is soft reset, mode setting, clock frequency setting, FIFO setting.

(2) Write the recognition list. The rule for lists is that each identifying entry corresponds to a specific number (1 byte), which can be the same or discontinuous, but value must be less than 256(00H ~ FFH). This chip supports up to 20 identification items, each identification item is standard mandarin Chinese. A series of different Numbers is used to identify entries in the contest, and *TABLE 2* is a simple example.

TABLE 2 IDENTIFY EXAMPLES

Serial Number	Details
0	All of them
1	All take off
2	advance
3	back
4	rising

(3) Start identifying. The recognition of speech can be started by setting several relevant registers. The relevant process is shown in *Figure 10*. ADC channel is microphone input channel, and ADC gain is microphone volume. "[fThresh=X]" represents the threshold corresponding to this command word (value range: "-32768" to "32767", default value is "0". The lower the score, the lower the reliability of the said recognition).After model identify any right or wrong command words , a score will be feedback. Actually, only score is higher than the threshold, the identification result will be feedback by system. According to the actual situation (typically threshold is set in a smaller value first, then statistical distribution can be gained based on large quantities of test data, and then set reasonable values).Adjusting the actual value of the value is helpful for adjusting each command word recognition and greater sensitivity.

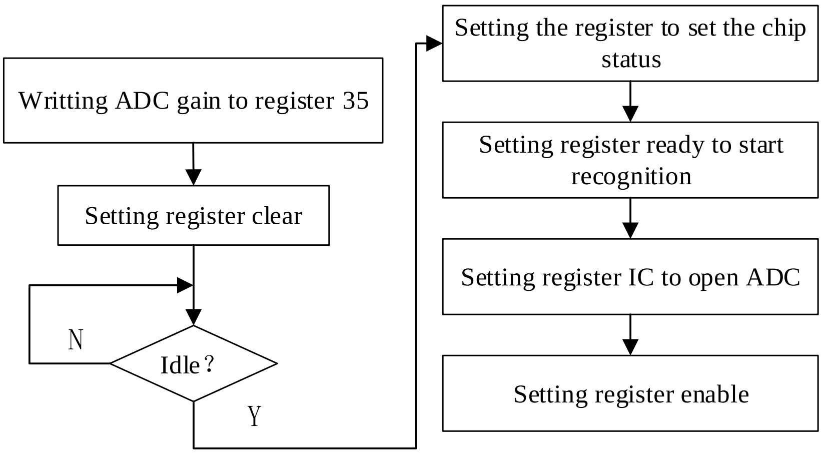

Figure 10. Identification process.

(4) Response interrupted. If the microphone picks up the sound, an interruption signal will be generated regardless of whether the normal result is recognized. The interrupt program analyzes the result according to the value of the register. Reading the value of the BA register, you can know that there are several candidate answers, while the answer in the C5 register is the one with the highest score and the most likely correct answer.

3.2.4 Integral module design

In order to facilitate voice command control, operators can intuitively see the sent voice content. We designed a set of voice control module integrating display, voice processing and wireless communication, as shown in *Figure 11* below.

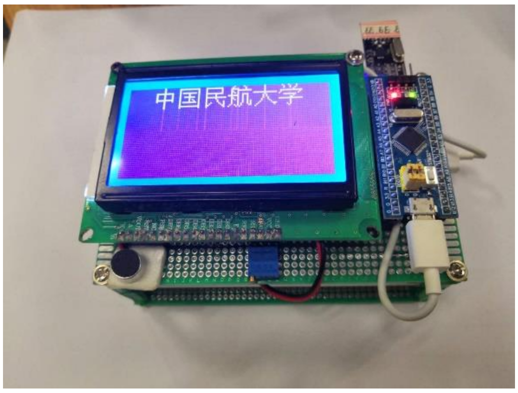

Figure 11. Voice control module.

When operating the voice input of team members, the screen will display the recognition result in real time according to the voice input, as shown in *Figure 12* below. According to the feedback information, judge the correctness of the input instruction, and send the correct instruction.

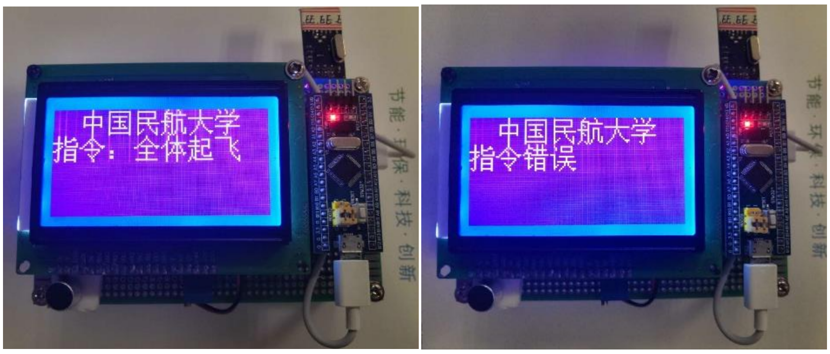

Figure 12. Speech recognition feedback results. a) Correct instructions. b) Wrong instruction.

3.3 Decision-making of UAV

The decision-making problem in Mission 8 can be divided into three major aspects: strategic decision-making, goal decision-making and control behavior decision-making. These three decisions are made sequentially and belong to the Sequential Decision (SD) problem. SD is a decision-making method for the optimization of stochastic or uncertain dynamic systems[5]. Its characteristics are that the system is dynamic, sequential and the turntable that may appear next is random or uncertain.

In Mission 8, the decision-making process of strategy selection is carried out. Firstly, make a decision on the UAV searching target strategy. Secondly, the lock on the QR code is selected as the target under the selected strategy. Finally, the target is positioned according to its position until the image is taken. The overall decision-making process is shown in *Figure 13*.

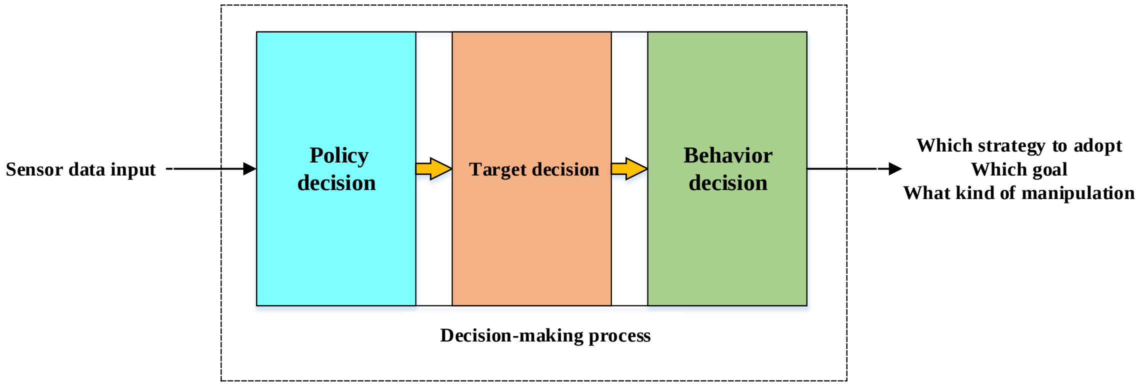

Figure 13. Overall process diagram for decision making.

3.4 Behavioral control of drones

According to the competition tasks, the behaviors of UAVs are divided into seven basic behaviors, namely, taking off and entering, returning, avoiding obstacles, returning over territory, target search, target locking and image acquisition. Some typical behaviors are analyzed below.

3.4.1 Target search behavior
Since RRT algorithm is fast and suitable for planning including kinematics and dynamics constraints, the idea of this algorithm is applied to the UAV target search process. After estimating the approximate position of the target robot on the ground, if no target is found when the UAV flies to this position, then RRT algorithm strategy is adopted to search the area until the target is found or the next command is received [8]. RRT search schematic diagram is shown in *Figure 14*.

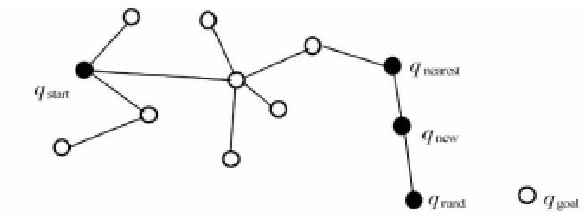

Figure 14. Schematic diagram of RRT search.

RRT search steps
Step1. Select the starting point qstart as the root node;
Step2. Randomly select the sampling point qrand in the space;
Step3. Search the nearest xnearest to the sampling point in the generated random tree;
Step4. Select a new node qnew by a certain step size on the line of qnear and qrand;
Step5. Repeat the above steps until the target point is reached or less than the given distance threshold from the target point.

3.4.2 Obstacle avoidance
As shown in *Figure 15*, when the drone encounters one or two obstacles, it chooses to bypass. That is according to the set threat area, the drone bypasses the threat area to continue other behaviors to achieve obstacle avoidance behavior.

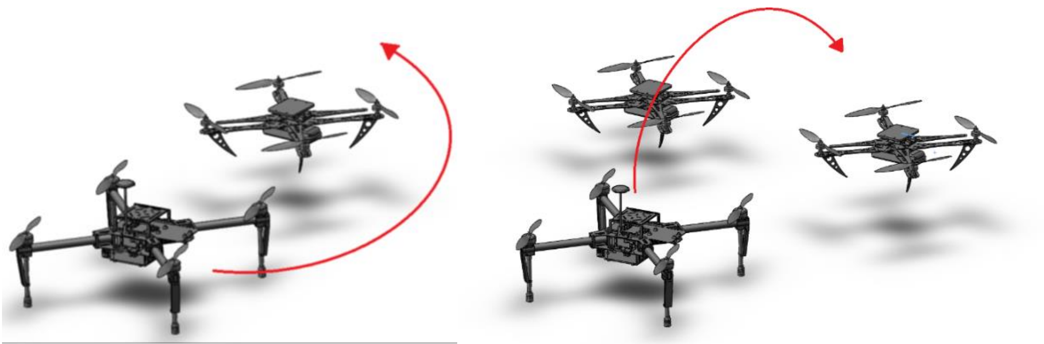

Figure 15. Schematic diagram of obstacle avoidance mode. a) UAV bypass obstacle avoidance. b) UAV flying high obstacle avoidance.

3.5 Security measures

3.5.1 Structure design of UAV
In order to increase the safety of UAV and avoid the paddle wounding personnel who rotate at high speed when the UAV is out of control, an anti-collision ring surrounded by blades is designed by using CAD drawing software. As shown in the figure below, the anti-collision ring

surrounds the entire blade, which improves the safety of the system.

As shown in *Figure 16*, the UAV will avoid obstacles by ascending or descending the flying height of itself when the drone is surrounded by multiple obstacles and the distance between the obstacles is less than the set threat range.

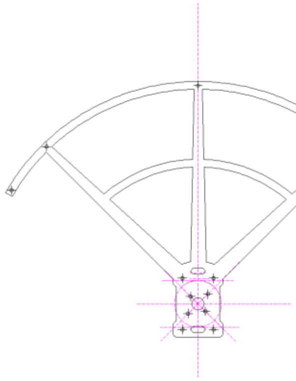

Figure 16. Anti-collision ring of UAV.

3.5.2 Scheme of electromagnetic interference / radio frequency interference

We adopt interference shielding circuit and instruction intervention to solve the interference problem. The contents of the instruction intervention are as follows:

When the drone is out of bounds or out of control, Issue the command "stop" from the voice control module and flight control system, and UAV would land in place at that time. The overall effect is shown as *Figure 17*.

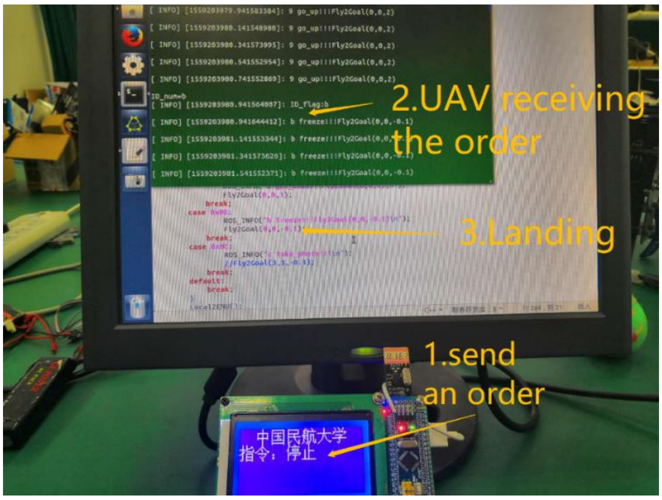

Figure 17. End drone action through voice control.

4 EXPERIMENTS AND RESULTS

4.1 QR Code Identification Experiment

With the gradual construction of the UAV system, we also did some experiments. In the visual recognition of the ground QR code, we can completely segment the four-part defective QR code and identify it after splicing. *Figure 18* shows the picture of the disabled QR code detected by the UAV camera. The green part is the largest square found. Corner detection is performed on this, and Figure 4-2 shows the result of the inner square corner detection. According to this, the scale of the QR code unit area can be obtained, and then the partial missing QR code is segmented, as shown in Figure 4-3. In the same way, the other three parts of the two-dimensional code are obtained, and the splicing and recognition are completed, and the password is obtained.

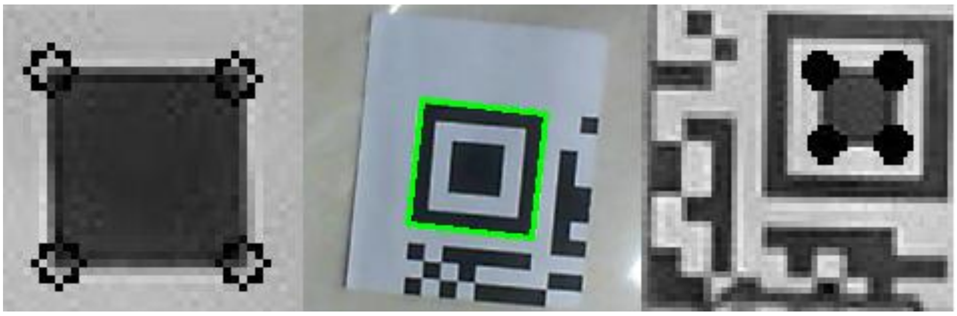

Figure 18. Experimental results. a) Detected incomplete QR code.
b) Internal square corner detection result. c) Split the partially broken QR code.

4.2 Speech Recognition Experiment

In order to ensure the speech recognition rate, stability and response time of the designed speech recognition module, this paper tests the speech recognition module described accordingly. The experiment environment is quiet home environment and noisy hospital environment. A total of 8 voice commands, 10 times tests for each voice command. The total number of experiments for each specific person in each environment is 80, and the number of successful recognition is recorded. The test results are listed in *TABLE 3*.

TABLE 3. STATISTICS OF SUCCESS NUMBER OF EXPERIMENTS

Command	Quiet environment			Noisy environment		
	Non-specific voice 1	Non-specific voice 2	Non-specific voice 3	Non-specific voice 1	Non-specific voice 2	Non-specific voice 3
Take off	9	9	8	9	9	8
Go ahead	9	8	9	8	8	8
Back	9	10	10	9	9	9
Fly left	9	9	9	9	9	9
Fly right	10	10	10	8	9	8
Rise	9	9	8	7	8	6
One meter down	9	8	9	8	8	8
Stop	10	9	10	10	9	10
Recognition rate	92.5	90	91.25	85	86.25	82.5

Among the three non-specific voice in the test, only non-specific voice 1 is female, the others are male. It can be seen from the data in the table that the speech recognition rate for non-specific people can reach more than 90% in a quiet environment, and in noisy environments it can reach 82.5% or more. In terms of recognition rate, the speech recognition rate is lower in a noisy environment than in a quiet environment. In terms of stability, the stability of the system is good in a quiet environment, the voice is said once, and the module can make a correct response at most 2 times. In the noisy environment, the stability of the system has decreased. Individual voice commands need to be said to be 3 or even more than 3 times to be accurately identified by the module. In terms of timeliness, the voice in a quiet environment can guarantee the real-time response of the system, the response time generally does not exceed 1 s, and the response time in a noisy environment is relatively long.

4.3 UAV Flight Experiment

According to the mission requirements, we did some basic experiments, including automatic takeoff, forward and backward, and the experimental effect is shown in Figure X.

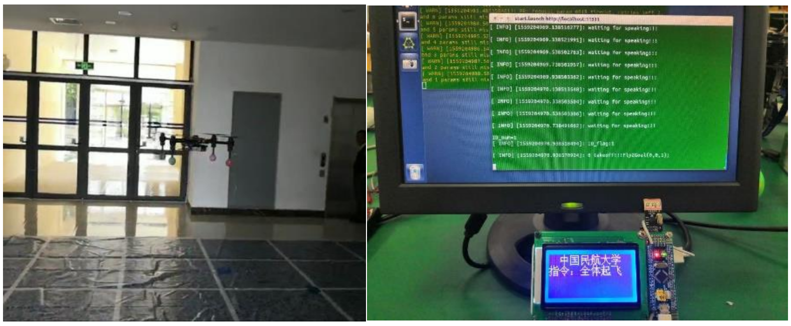

Figure 19. Vertical landing effect.

5 Summary

According to the rules of the competition, we designed our own UAV systems and mission solutions, and the feasibility of the solution was verified by the initial experiments. UAV can realize autonomous flight in unknown environment and make autonomous behavioral decisions according to environmental information, such as obstacle avoidance and target search. The next direction of our efforts is to improve the performance index of each subsystem to successfully complete the requirements of Mission 8.

6 References

[1] Lee K F, Hon H W. Speaker-independent phone recognition using hidden Markov models[J]. Acoustics, Speech and Signal Processing, IEEE Transactions on, 1989, 37(11): 1641-1648.

[2] LEE D J. Nonlinear estimation and multiple sensor fusion using unscented information filtering[J]. IEEE Signal Processing Letters,2008,15: 861-864.DOI: 10.1109/ LSP. 2008. 2005447.

[3] 王林,王楠,朱华勇,等.一种面向多无人机协同感知的分布式融合估计方法[J].控制与决策, 2010,25(6):814-820.WANG Lin,WANG Nan,ZHU Huayong,et al. Distributed fusion estimation algorithm for multi-UAVs cooperative sensing[J]. Control and Decision,2010,25(6): 814-820.

[4] ICRoute .LD332X 开 发 手册 [EB/OL] .[2010-03-10] .http ://www .icroute .com/web

cn/DownLoad .html .
[5] The navigation and control technology inside the ar.drone micro UAV . J.-P.Bristeau. 18th IFAC World Congress. 2011.
[6] A Comparative Study of Optical Flow and Traditional Sensors in UAV navigation. Chao Haiyang,Yu Gu,Jason Gross ,etc. American Journal of Infection Control . 2016.
[7] An Open Source and Open Hardware Embedded Metric Optical Flow CMOS Camera for Indoor and Outdoor Applications. Honegger D,Meier L,Tanskanen P,Pollefeys M. IEEE International Conference on Robotics and Automation ICRA . 2013.
[8] Online learning of robust object detectors during unstable tracking. Kalal Z, Matas J, Mikolajczyk K. Proceedings of the IEEE On-line Learning for Computer Vision Workshop. 2009.