Technical Paper of NWPU-VWPP team for the 8th International Aerial Robotics Competition

Kehan Xue, Quantao Wang, Xiaopeng Yan

Northwestern Polytechnical University

ABSTRACT

The International Aerial Robotics Competition is the world's top air robot competition. This paper shows the solutions and progress made by the NWPU-VWPP team in solving the International Aerial Robotics Competition 8th mission. We will present our scheme from the following points: Overall System Design, Navigation and State Estimation, Mission Package related algorithm implementation, and the Strategy System. We achieved a vision-based autonomous navigation and obstacle avoidance function for UAVs in an indoor open environment without GPS. And we implemented a non-electronic human-computer interaction interface. We also proposed a powerful dynamic strategy scheduling system and added people to the decision loop.

INTRODUCTION

Problem Statement

The main content of the International Aerial Robotics Competition 8th mission is that a contestant works in cooperation with a team of drones, trying to retrieve a particular item in a site where the "enemy" has a drone patrol, while ensuring that it is not subdued by the "enemy". In addition to the technical characteristics required by the previous series of tasks (such as full autonomy, obstacle avoidance, navigation guidance, etc.), mission 8 also requires non-electronic human-computer interaction, fleet coordination, air target identification and other capabilities. The increasing autonomy of UAVs and the instantaneous, relevant and accurate information exchange between UAVs can bring great advantages to the competition. Getting a larger amount of information can provide an overwhelming advantage for the team. And it is necessary to consider the complexity of the command and control. The unique and most important challenge facing the competition in this space is the development of appropriate dynamic strategy planning capabilities between people and drones.

Conceptual Solution

The core problem of this generation of missions is to build a cluster of intelligent aircraft that can realize non-electronic interaction with people, and to add human links in mission planning to achieve dynamic mission planning capabilities in a dynamic environment.

SYSTEM OVERVIEW

First, a stable hardware system is the foundation of everything. Next, through the Driver Layer, which can communicate with the Hardware Layer, we build the Function Layer required for the mission 8. Finally, the data of all the services was streamed into the Strategy Layer. The Strategy Layer is responsible for all strategy scheduling and outputting all the final control signals.

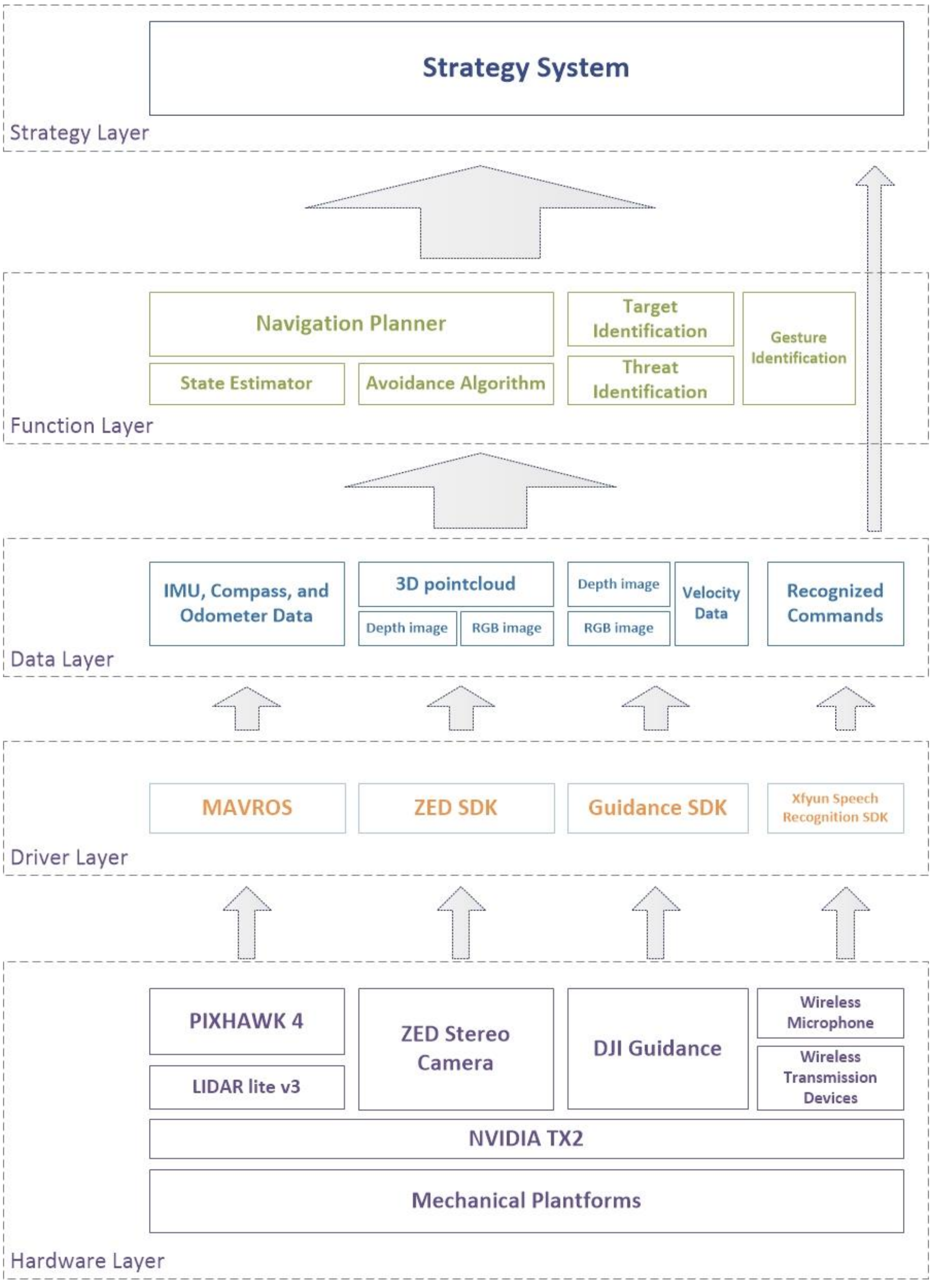

Figure 1. Overall system architecture

We chose our hardware devices according to our requirements. After building the Driver Layer, we implement our mission-related algorithms based on the data we get from them. We made the UAV achieve a better state estimation by streaming the output data of the DJI Guidance into the PX4 Autopilot system. Our navigation and obstacle avoidance algorithms are based on the 3DVFH*[1][2] algorithm with Octomap[2][3]. Through the iFLYTEK's SDK to achieve speech recognition. For the target identification and threat detection, we use a Mobile-SSD[4] and a yolov2-tiny network structure[5], both are Deep Learning based methods. The method mixed with Convolutional Pose Machines[6] and traditional geometric method is used to identify gestures. Our Strategy Layer based on finite automatic state machine. Bring the human into the dynamic decision system. Through these efforts, our UAVs are highly autonomous and as much as possible to complete the eighth generation of missions.

Yearly Milestones

Since our team is participating in the competition for the first time, a lot of work has been carried out during the less than one year of preparation. In order to improve the efficiency of development, we modularize the development tasks. It is mainly divided into the bottom layer design, localization and navigation system development, target recognition module development, human-machine interface development, strategy system development and overall architecture, six parts. In the end, we successfully achieved the localization and navigation functions of the UAVs indoor, the detection and recognition functions of targets and threats, the human-machine interface functions and the strategy system. And completed their integration.

AIR VEHICLE

Description of Configuration/Type

Mechanical Design

The mechanical design of the air robot was modified based on the frame of the DJI Matrice 100 aircraft. Following its overall design approach, we used its stands, motors and ESC (Electronic Speed Control), but made a new center board structure to place our own hardware. The DJI M100's fuselage frame has a very good aerodynamic performance and shock absorption design, which can reduce a lot of trouble during the subsequent development process.

Figure 2. Our UAV's design

Hardware System

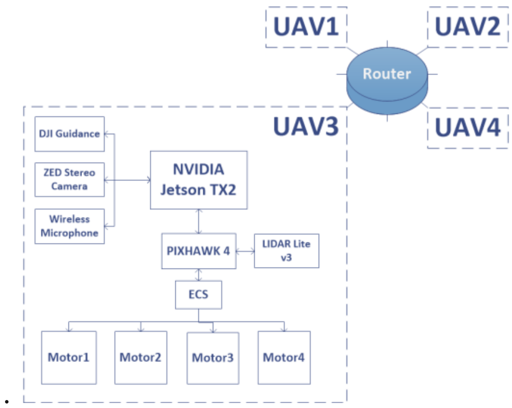

Figure 3. Electronic system structure

We use the DJI 3510 motors and DJI E SERIES 620D ESC on the UAVs, used to provide sufficient power, which can allow the aircraft to maintain its balance between maneuverability and stability as much as possible under heavy load.

We use PIXHAWK 4 as our flight control board. Pixhawk 4 is based on the Pixhawk-project FMUv5 open hardware design and runs PX4 Autopilot on the NuttX OS. PIXHAWK 4 has a redundant design of dual IMUs. The IMUs in it is ICM-20689 and BMI055. The type of Magnetometer is IST8310. The open source features and good stability of PIXHAWK 4 made us finally choose it.

In order to make our aircraft better capable of achieving high altitude indoors, we have installed the Lidar lite v3 module for our aircraft. Lidar lite v3 is a High-performance Optical Distant Measurement Sensor. Lidar lite v3 is extremely small in size and weight. It has a range of 40 meters. It’s accuracy is plus or minus 2.5 cm. It is an ideal module for indoor height setting on an aircraft.

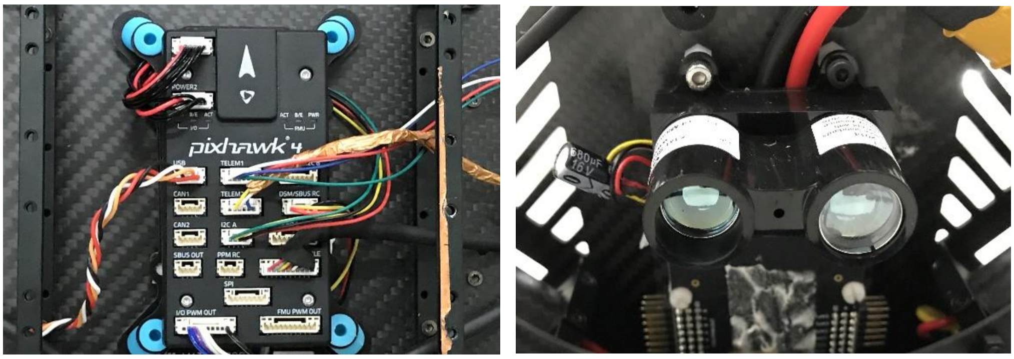

Figure 4. PIXHAWK 4 and Lidar lite v3 on our UAV

The onboard computer uses NVIDIA TX2. NVIDIA Jetson TX2 is the fastest, most power-efficient embedded AI computing device. This 7.5-watt supercomputer on a module brings true AI computing at the edge. NVIDIA Jetson TX2's high computing performance and low power consumption, and system compatibility is very good, we chose it. NVIDIA Jetson TX2 runs all the main programs, processes data from ZED, DJI Guidance, and communicates with PIXHAWK via the serial port. And by connecting the TX2 on each UAV to the same LAN, low-power communication between each UAV is achieved.

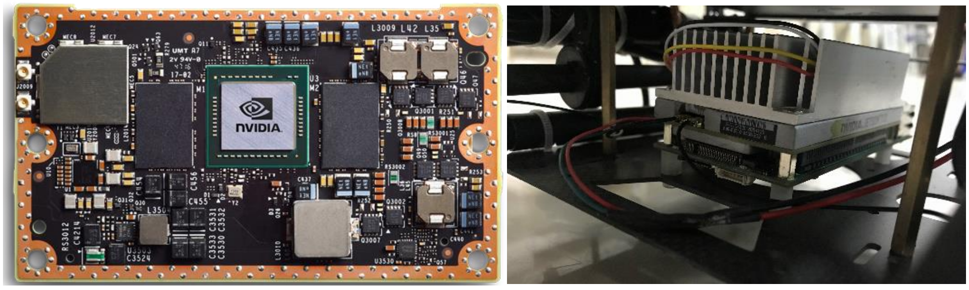

Figure 5. NVIDIA Jetson TX2 and it on our UAV

We use ZED Stereo Camera to help our aircraft sense the environment. ZED is the world's fastest depth camera. With its 16:9 native sensors and ultra sharp 6 element all glass lenses, we can capture 110° wide-angle video and depth. ZED cameras can output high quality depth maps and dense point cloud data. We can use this data to implement the navigation and obstacle avoidance functions of the aircraft.

We equip our UAV with the DJI Guidance system. The Guidance uses a high-precision stereo vision algorithm with a near-ground positioning accuracy of up to centimeter. Localization information is available in complex terrain and high-speed flight conditions. Its output information includes 8 RGB images and depth images, as well as estimated speed information of the aircraft. We use it to estimate the state of the aircraft in an open indoor environment.

Figure 6. ZED Stereo Camera and DJI Guidance system

Flight Control System

Navigation/State Estimation System

The navigation system is based on the 3DVFH* algorithm to realize real-time obstacle avoidance and local path planning. The 3DVFH* obstacle avoidance algorithm suitable for real-time application on UAVs combines the ideas behind the previously presented 3DVFH+ and the VFH* algorithm with a novel memory strategy. The 3DVFH* algorithm computes obstacle avoidance maneuvers in a purely reactive manner without the need to build a global map of the environment. The green line in the *Figure 7* below is the planned path.

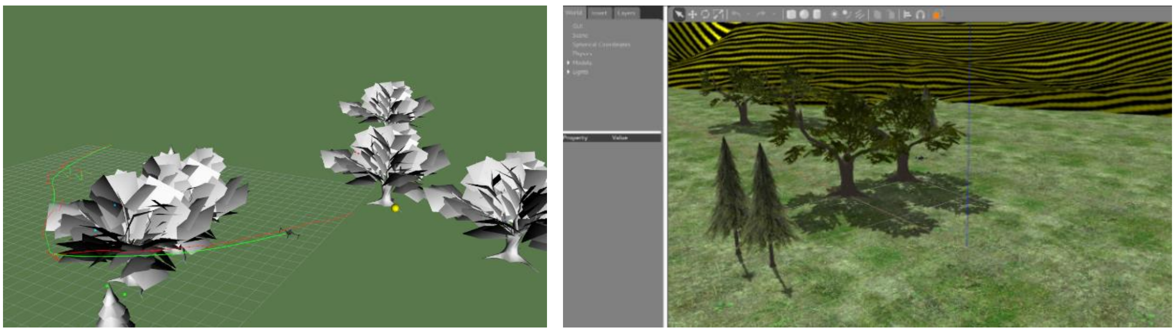

Figure 7. Simulation of 3DVFH algorithm in gazebo with ROS*

Attitude/Position Control System

The estimates come from EKF. This is a standard cascaded position-velocity loop. Depending on the mode, the outer (position) loop is bypassed (shown as a multiplexer after the outer loop). The position loop is only used when holding position or when the requested velocity in an axis is null. The integrator in the inner loop (velocity) controller includes an anti-reset windup (ARW) using a clamping method.

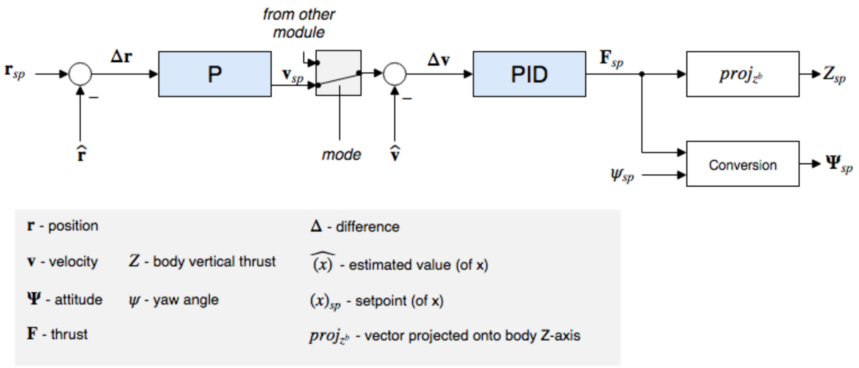

Figure 8. Attitude and Position Control Structure

MISSION PACKAGE

Perception System

Target Identification and Behavior

The main objectives of this task are our team member and QR code, as well as storage boxes for QR code. Among them, for the recognition of the participant and storage boxes with QR code, our team adopts the MobileNet_SSD[4] object recognition model. The MobileNet_SSD is a lightweight deep network model for mobile terminals, it can reach a considerable speed on our UAV platform.

For the recognition of QR code, our UAV captured QR code directly above the storage box, then processed the image, partitioned the part containing QR code information from the image, transmitted four QR code images to the same UAV for QR code combination to obtain all the codes. The figure below is the specific algorithm processing flow.

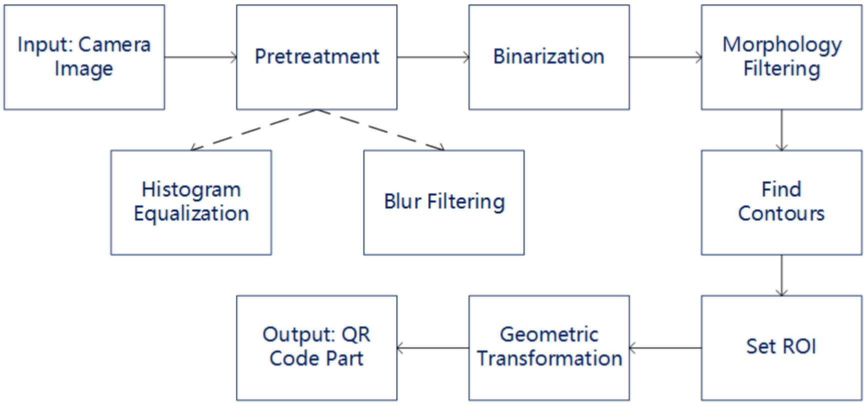

Figure 9. Image segmentation QR code

Threat Identification and Behavior

The threat in this mission is four enemy UAVs. Threat identification, which means the identification of " Enemy" UAVs. Our team uses the Yolov2-tiny target detection model built under the Darknet framework to identify the enemy UAV. Yolov2-tiny has the characteristics of fast processing speed and real-time performance in embedded system, making it suitable for the application of UAV platform. In the actual test process, the processing speed of 70 frames per second can be achieved on GPU GTX 1050, and on the TX2 airborne processor of UAV platform, the real-time processing effect can also be achieved, which can adapt to the scene requirements of UAV fast moving.

By training in the UAV data set produced by our team, the model can accurately identify the enemy UAV in the video of mission 8 of IARC official website, and the recognition probability is over 90%. The position of the enemy UAV in the image captured by our UAV camera can be obtained by machine learning. And the coordinate information of multiple UAVs can be fused. The position of the enemy UAV can be obtained by the transformation of the camera coordinate system and the world coordinate system, thus providing the threat location information for our team's strategy system.

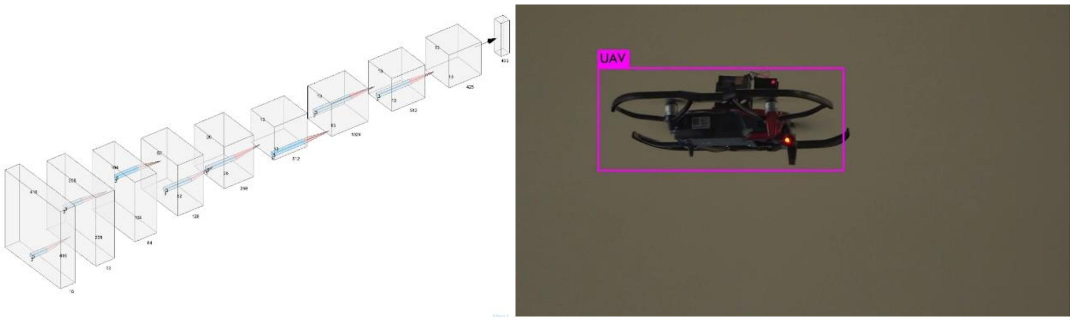

Figure 10. Yolov2-tiny model structure and the object recognition result

Gesture Identification and Behavior

Recognition of our team's gestures is based on image processing, using OpenCV computer vision library for the process, and machine learning human posture analyses, using the Tensorflow framework of the Convolutional Pose Machines[6]. These two methods have their own advantages and disadvantages, the former is fast, while the latter has strong environmental robustness and many kinds of recognition, so our team adopts a combination of the two methods to carry out gesture recognition of our team members. The image processing step of gesture recognition following the figure below:

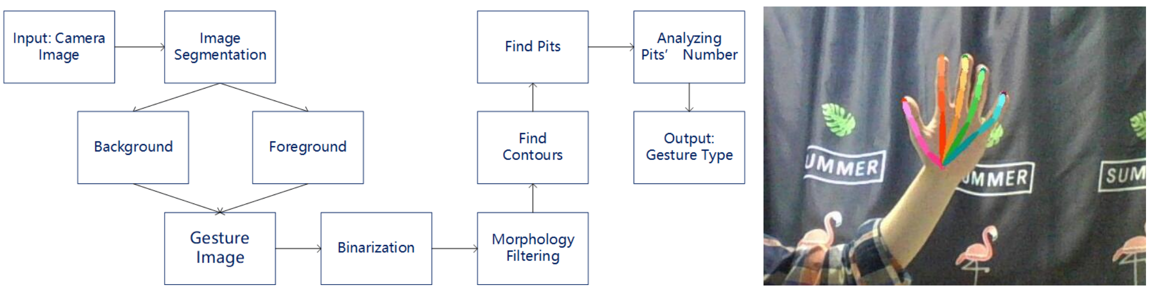

Figure 11. gesture recognition by only using OpenCV and the result

Communications System(s)

The *Figure 12* below shows the structure of our communication system.

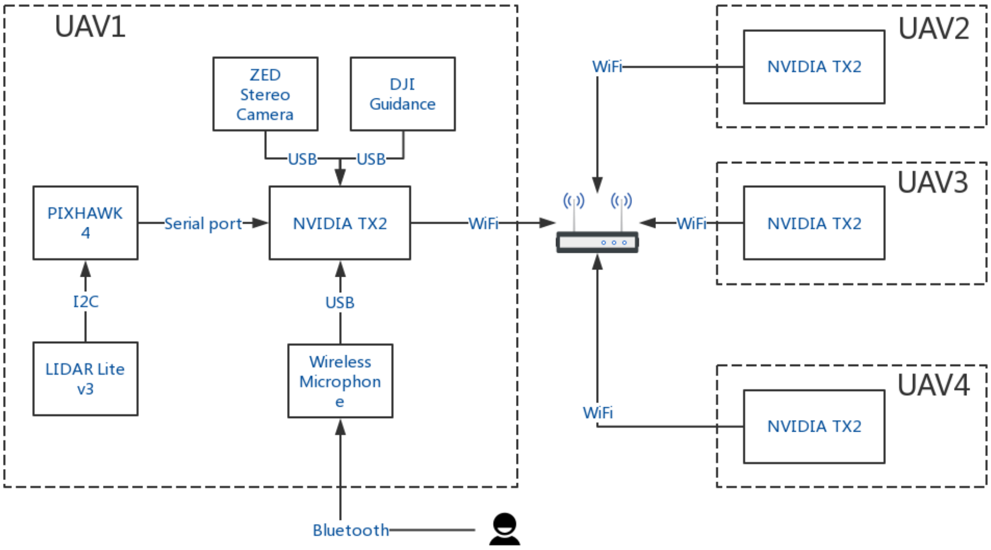

Figure 12. Communications system

Strategy System

In order to realize multi - machine and human collaborative decision-making, this paper subdivides the upper level strategy. The pure strategy layer will obtain processed arena information and human-computer interaction information, and control strategy through perfect control action[9]. Its basic structure is shown in the figure below:

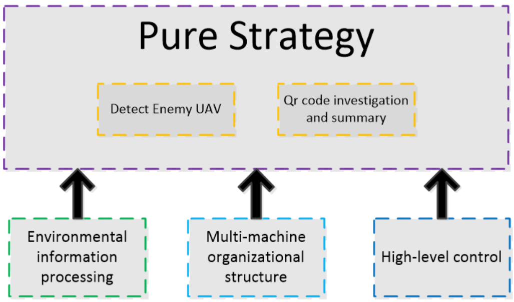

Figure 13. High-level policy infrastructure

Centralized and distributed multi-machine organizational structure

The centralized organization structure and distributed fault-tolerant mechanism are adopted in this paper to ensure the simplicity and reliability of strategy development[7]. While the dimension explosion caused by the unlimited expansion of the number of individuals is not considered because of the fixation of the aircraft number, the centralized organization has obvious advantages. At the same time, due to the uncertainty of the arena environment, abnormal communication is common, which poses a great threat to the coordination of aircraft, individual strategies are implemented on individual aircraft as a basic strategy[10][11].

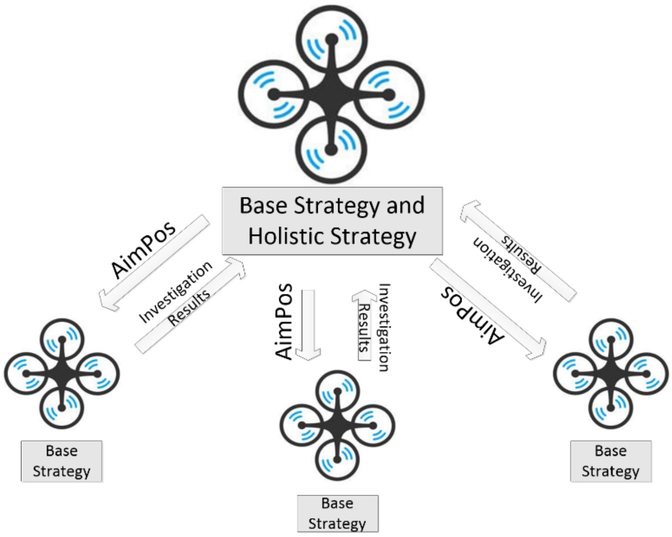

Figure 14. Multi-machine organization structure

Environmental information processing

According to the structure of aircraft perception, the field information that can be obtained directly at present mainly includes the position and posture of our aircraft, the depth point cloud map directly in front of the aircraft (for obstacle avoidance and path planning), the image and depth map of the aircraft in four directions (at most 20 meters with limited perspective), the image and depth map of the lower view and other information. After image processing and information extraction, we can obtain the following effective information:

There are four possibilities for information related to the location of human and enemy robots:

$$\begin{cases} do\ not\ see \\ finding\ the\ tar\ and\ angle\ but\ no\ depth\ by\ guidence \\ finding\ the\ tar, angle\ and\ depth\ by\ guidence \\ finding\ the\ tar, angle\ and\ depth\ by\ zed \end{cases}$$

After that, through the above information, the positions of the four enemy aircraft can be comprehensively calculated. On the basis of the comprehensive calculation, two kinds of predictions are required:

$$\begin{cases} enemy\ track\ \ prediction \\ human\ track\ \ prediction \end{cases}.$$

High-level control

According to the processed environment information, basic control interface and strategy, three high-level control actions are designed to complete the overall strategy.

(1) cruise

The first is to determine the formation, in order to ensure the normal flight of our aircraft, we need to consider an important factor: aircraft overlooking distance should be greater than the diameter of the aircraft. On this basis, the formation should be determined to increase the field of vision, increase the flexibility of maneuvers, and to provide certain cover surface. For this reason, we designed a 2-1-1 cruise formation, which takes people's activities as tracks and follows people's movements for accompanying cruise.

(2) track and block

In the process of tracking the blocking, the blocking point needs to be determined. There are three conditions to consider:

$$\begin{cases} our\ aircraft\ i\text{s in}\ the\ line\ of\ obstruction \\ our\ aircraft\ is\ within\ \text{max blocking radius} \\ the\ closer\ you\ get\ to\ the\ enemy\ without\ a\ collision, the\ better\ the\ blocking\ effect \end{cases}$$

The maximum blocking radius is as follows:

$$R_{\max} = \frac{V_{\mathrm{home}}}{V_{opp}} \cdot Dis_{opp}$$

(3) Confirmation and detection of QR code

In order to realize the task of piecing the incomplete two-dimensional code information placed in different areas into complete information, this process was divided into four steps. The host distributes the detection number, the slave confirms the detection number, the master and slave conduct the detection, and the slave reports the detection information in four steps.

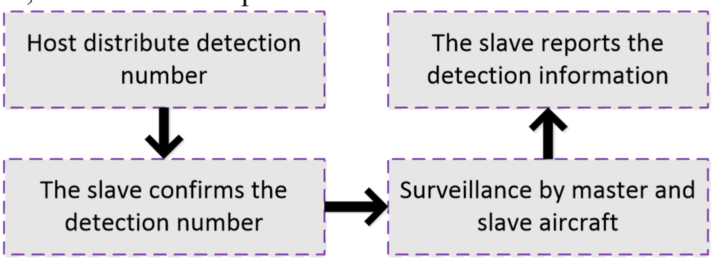

Figure 15 Two-dimensional code acquisition process

(4) Strategy organization transition based on finite automatic state machine

Complex strategies are difficult to achieve in one step, requiring a lot of manual involvement rather than machine automation. In fact, humans currently have too many advantages over machines in overall control of strategy. Only in the optimization of strategy details, the setting of strategy conditions and other tedious tasks requiring trial and error, can machines better replace human beings.

In order to combine human work with machine work, a decision network based on hierarchical finite automatic state machine is proposed. In essence, the decision network proposed here adjusts the state classification mode originally belonging to the whole into the boundary classification mode of multiple strategies, increasing the number of classification, but reducing the difficulty of classification, so as to realize the scientific switching in the process of policy switching. This applies the transition principle of a finite automatic state machine[8].

However, due to the complexity of the state space, the state space needs to be further abstracted to adapt to the discretization of the priority automatic state machine. With the use of a finite automatic state machine, the organizational form of host policy is shown in *Figure 16* on a macro level. The policy transformation conditions represented by 1 to 12 will become the focus of the study and guide the strategy to proceed according to the plan. This approach ensures that the details of the policy are fully considered and effectively isolates the multi-tier policy structure.

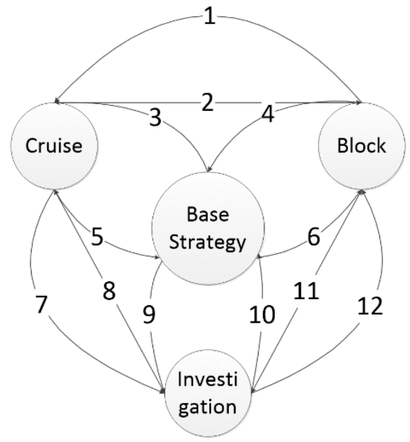

Figure 16. Strategy transformation automaton

RISK REDUCTION

Safety

We set the UAV's obstacle avoidance action to the highest priority. Ensure that the aircraft avoids all unnecessary problems under normal circumstances. Our aircraft adopts a passive risk protection approach. As the *Figure 2* shown, we equip our aircraft with reliable paddles to minimize damage to the outside world in the event of an accident.

Modeling and Simulation

As the *Figure 7* shown, we tested the aircraft's obstacle avoidance algorithm on the gazebo simulation platform. And in the usual development process, when the physical aircraft is running, try to test the flow of the algorithm in gazebo as much as possible to ensure that no low-level errors will occur.

CONCLUSION

This paper describes the design of the aircraft and the algorithm for IARC mission 8. So far, we have implemented most of them. Platform construction and development of each module and basic completion, ongoing debugging of the strategy and integration of the entire system. In the coming period, we will continue to work hard on our development work.

REFERENCES

[1]. Tanja Baumann: Obstacle Avoidance for Drones Using a 3DVFH* Algorithm, ETHZ

[2]. Simon Vanneste, Ben Bellekens, and Maarten Weyn, "3DVFH+: Real-Time Three-Dimensional Obstacle Avoidance Using an Octomap"

[3]. Armin Hornung, Kai M. Wurm, Maren Bennewitz, Cyrill Stachniss, Wolfram Burgard: "OctoMap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees", Autonomous Robots (2013)

[4]. Jonathan Huang, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer, Zbigniew Wojna, Yang Song, Sergio Guadarrama, Kevin Murphy: "Speed/accuracy trade-offs for modern convolutional object detectors", arXiv preprint arXiv: 1611.10012, 2017.4.25

[5]. Joseph Redmon, Ali Farhadi: "YOLO9000: Better, Faster, Stronger", arXiv preprint arXiv: 1612.08242, 2016.12.25

[6]. Shih-En Wei, Varun Ramakrishna, Takeo Kanade, Yaser Sheikh: "Convolutional Pose Machines.", arXiv preprint arXiv: 1602.00134, 2016.4.12

[7]. Perolat J , Leibo J Z , Zambaldi V , et al. A multi-agent reinforcement learning model of common-pool resource appropriation[J]. 2017.

[8]. Capuano N, Chiclana F, Fujita H, et al. Fuzzy Group Decision Making With Incomplete Information Guided by Social Influence[J]. IEEE Transactions on Fuzzy Systems, 2018, 26(3): 1704-1718.

[9]. Canellas M C , Feigh K M , Chua Z K . Accuracy and effort of decision making strategies with incomplete information[C]// 2014 IEEE International Inter-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA). IEEE, 2014.

[10]. Cong-Cong L , Rodriguez R M , Luis M , et al. Consensus building with individual consistency control in group decision making[J]. IEEE Transactions on Fuzzy Systems, 2018:1-1.

[11]. Xu X H , Du Z J , Chen X H . Consensus model for multi-criteria large-group emergency decision making considering non-cooperative behaviors and minority opinions[J]. Decision Support Systems, 2015, 79:150-160.