SYSU *Swift* TECHNICAL REPORT MISSION 8

INTERNATIONAL AERIAL ROBOTICS COMPETITION 2019

Sun Yat-sen University*
Guangzhou, Guangdong Province, 510006, P.R.China

ABSTRACT

The Sun Yat-sen University Swift Aerial Robotics Team has constructed a cluster of quadcopter aircrafts for the fulfillment towards Mission 8 of the International Aerial Robotics Competition. To achieve the fluent cooperations among multi-aircrafts and human under the complex circumstances that exists opposing enemy aircrafts, not only the ego perceptions and control, but also the real-time intercommunications and analysis of the shared information contribute to the top level decisions for the aircrafts' formation, herein assuring the success. This paper presents both the hardware and software frameworks of the project named *Chicken* for Mission 8. With the integration of the multi-sensors, Omni-Terrain Height Estimator and Localization Module promote the control stability of the vehicles. A flexible task distribution for aircrafts' formation network and non-electronic interaction between the vehicles and human are designed to agilely plan with strong robustness.

1 Introduction

Mission 8 of the IARC challenges participators to develop high intelligent aerial robotics which can choose and drive robots to a specific side in an indoor arena. In the environment with non-electronic command and control, such as speech control or gesture control, the aerial robotics are required to locate the necessary component and conunter the attacks of enemies, together with identify the password to help you get the component. Your aerial robotics also can heal some wounds if they are in proximity to you and are commanded to do so. A pratical system architecture and new strategy should be developed.

To fulfill the task, we design a system which consists of several modules, including positioning, flight control, swarm interaction, QR code detection, obstacle avoidance, trajectory planning and safety system. At first we should control the aircrafts and formation, which we intend to use speech control. And in order to help it arrive the correct position and identify the seperated QR code, positioning system, obstacle avoidance and QR code detection should be developed. Meanwhile, people need to avoid killed by sentry robots, the aircrafts should be equipped with heal laser.

To achieve the above functions, optical flow, IMU data, lidar and UWB device are combined to construct a positioning system. Then, we implement flight control based on PID rather than original fight controller. And we set up the whole software system framework based on ROS, which can be distributedly managed.

2 Hardware Structure

The hardware system is the fundamental electronic base of an aerial robot. Choosing appropriate devices as well as constructing a hardware system, with low cost, low power consumption and excellent performance for the task is of great significance.

2.1 Platform

2.1.1 Air Vehicle

In allusion to Mission 8, XAircraft-X650 is chosen as the platform of the aerial robots. This aircraft is a quadcopter which is simpler and more stable and robust than many other structures. The frame has a maximum size of

* **Team Leaders:** Yinyi Liu, Xiangwei Zeng

55cm, which is within the competition limitation. The empty frame weight is only 880g, with maximum take-off weight 720g, which enables all needed onboard devices to be equipped.

Figure 1: Platform.

2.1.2 Airborne Computer

The aircraft utilizes an onboard computer as the main intelligent processor for information integration. The modal of airbrone computer is known as NVIDIA Jetson TX2, which is an embedded system-on-module with Dual-core NVIDIA Denver 2 + Quad-core ARM Cortex-A57, 8GB 128-bit LPDDR4 and integrated 256-core Pascal GPU. Besides, Jetson TX2 has a small size of 4 inches, which suits for the aircraft board.

All the data collected by the sensors are gathered together and processed in Jetson TX2. Image processing, trajectory generating and upper layer strategy are implemented in this powerful core.

2.2 Equipments

2.2.1 Control Unit

We use DJI-N3 AutoPilot as our flight controller, and it is responsible for executing commands from Jetson TX2. Besides, remote control can take over control of the flight since there is a selector of control method in the control system. This method can ensure safety when something unexpected happens.

DJI-N3 has 2 IMU inside and can extend high-performance external sensors. The IMU combines accelerometers, gyroscopes and magnetometers, which offer data of attitude and acceleration to show the current state of aircraft. The frequency of data transmission from IMU to Jetson TX2 is 100Hz in single floating-point. The data from IMU is utilized and integrate with other information, such as UWB location, optical flow, to realize localization.

2.2.2 Positioning System

Positioning is the basis of flight control, obstacle avoidance, strategy implementation and thus the basis of aerial vehicle to accomplish the Mission 8. Positioning system should be stable and reliable enough. We come up with a solution based on UWB(ultra-wideband) localization technology, optical flow integration positioning and height mesurement with lidar, which combines the advantages of three technologies.

Sensors:

1.UWB device: RoboMaster UWB Positioning System Set

2.Lidar : DE-LiDAR-TF02

3.Camera : CM3-U3-13Y3C-CS

4.IMU : DJI N3 autopilot

First we deploy a UWB base station at the midpoint of the four sides of the stadium, and carry a UWB tag module on each aircraft and pass the UART interface. The location information is sent to Jetson TX2 to determine the exact location of the aircraft. At the meanwhile, color camera heads towards the ground to collect images of the ground. With the data of attitude from IMU, airborne computer can execute optical flow calculation. Optical flow and the location information from UWB are discussed further in the chapter of software structure.

Figure 2: UWB device.

DE-LiDAR-TF02 is selected for measuring height, as the altitude imformation is vitally important. TF02 is a lidar sensor using laser source to detect the altitude. It provides high sensitivity and the ability of measuring area is up to 22m with high precision reaching centimeter level. Moreover,the maximum sampling frequency is 100Hz and it only weigh 50g.

2.2.3 Obstacle Detection System

Considering obstacle aircrafts should be encountered in the competition, obstacle detection system is constructed to detect obstacle in all directions and deal with the crash threat. Intel RealSense camera D435 is utilised to obtain depth imformation which could be used to detect obstacles and calculate distance of obstacles by algorithm.

2.2.4 Vision System

The vision system is the pivotal part of the aerial robot for Mission 8. Cameras with broad horizon and less smear are therefore selected for collecting vision imformation.

In order to capture images with high resolution of the ground and improve the accuracy of QR code detection, we choose CM3-U3-13Y3C-CS colour camera with high resolution of 1280*1024. It has a maximum frame rate of 149fps and global shutter.

2.2.5 Communication System

The aircrafts communicate with ground computer with WiFi. Mission 8 requires four aerial robots controlled in a good flying formation, in order to detect four parts of QR code simutaneously. Each aircraft is equipped with one dual band wireless network card in airborne computer with WiFi. A ground router is connected to networks cards and ground computer by antenna, so that the airborne and ground computer are settled in the same small area LAN.

One aircraft stands out to be the leader, which receives the command from ground computer directly, and passes on to the rest three followers. Followers should only execute the command transmitted by leader aircraft. All data are transmitted by 802.11X protocol.

Within the LAN, the flight status, position and altitude for example, and obstacles information are shared among four aircrafts, together with providing videos stream to mobile device.

2.2.6 Healing Laser

In order to heal person after being hit by sentry kill beam, we set a module called healing laser controlled by STM32F103. It shoots healing beam in the frequency of 13kHz. The person will die after being hit by kill beam 10 times, so we can ultilize healing laser to increase the health point and obtain short immune.

3 Software Structure

The Vehicle utilizes NVIDIA Jetson TX 2 as OnBoard Processor, running Robotics Operate System *(ROS)*, which is the fundamental environment of our software project named *Chicken*. The framework is designed as the levels of perception, control, planning and decision. The functions of our codes entail four main parts: positioning system, obstacle avoidance, QR code detection and swarm interaction.

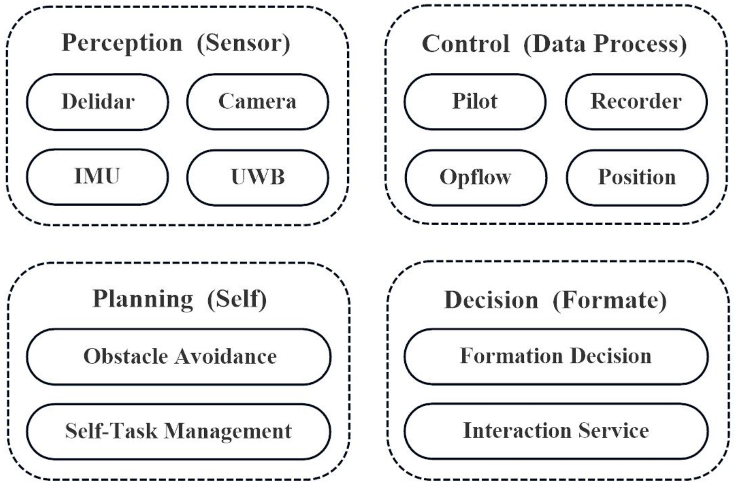

Figure 3: Software Framework Based on ROS.

3.1 Control System

We build a comprehensive control system as Figure 4 shows. The control system contains Omni-Terrain Height Estimator which estimates the altitude, Positioning Module which determines the flight states, Auto-Pilot Module which can control the aircraft using PID controller, Formate and Obstacle Avoidance which affect the aircraft to execute the control.

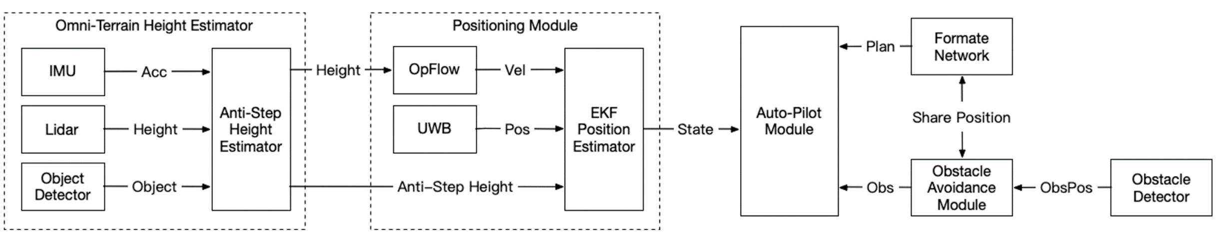

Figure 4: Comprehensive Control System.

3.1.1 PID controller

The control system is shown below. We chose a traditional but effective control theory, PID algorithm, to realize the flight control after trying kinds of methods, leading to a quite stable control effect.

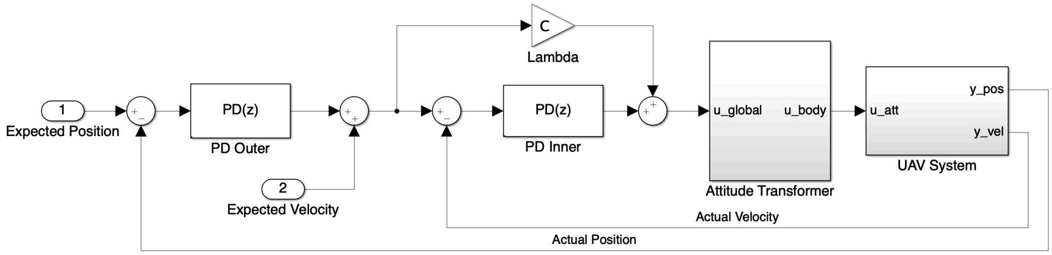

Figure 5: PID Control Flow.

3.1.2 Omni-Terrain Height Estimator

In GPS-denied Environment, height information is based on line-radar or ultrasonic sensor. While vehicle hovers through the uneven terrain with object bumps, step wave of height data occurs, which leads to a vehicle climb-out due to a rapid change in the vertical dimension. In this section, we fuse the data from IMU (with barometer inside), Lidar (line-radar for height measurement) and Object Detector (image analysis) to eliminate the mutation for omni-terrain in order to smooth the height control loop stably.

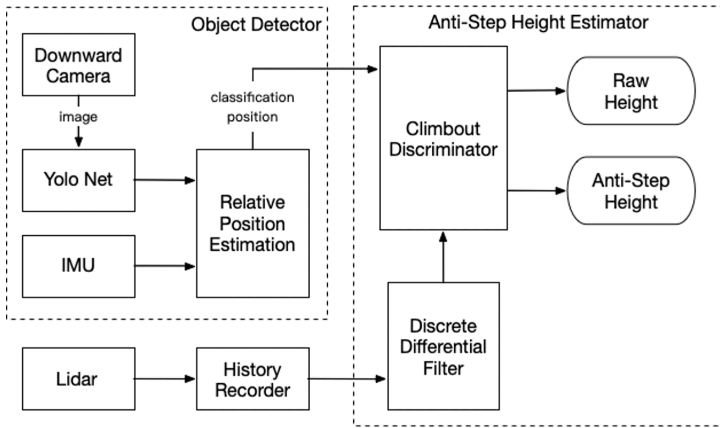

Figure 6: Complement Details for Omni-Terrain Height Estimator.

See Figure 6. To achieve this estimator, three functional parts are built: Object Detector, Lidar Measurement and Anti-Step Height Estimator. YOLO v3 Net is a fast CNN structure applied for real-time object detection. Combined with inertia measurement, relative spatial distribution towards camera view could be determined. With the classification data and historic sequence of vertical height data, Climbout Discriminator predicts the approximate anti-step height relative to the previous plain level. The output of raw height and anti-step height could be used for optical flow estimation and auto pilot control respectively.

3.2 Positioning System

3.2.1 Control

Our flight controller is DJI-N3. This flight controller has complete Software Development Kit, which can meet our team's needs and other extend functions.

We using Onboard-SDK to realize stable flight of aircraft and mission plan.

Introduction of SDK

DJI-SDK provides many interfaces for flying. There are three parts, including Motor Control, Beginning and Ending Flights and Flight Control. Motor control provides APIs which can control motors. With motor control we can make the aircraft move flexible. Beginning and Ending Flights can make the aircraft take off and land easily and safely.

Flight control provides APIs to control the aircraft in a high layer control commands. So we can send flight control commands directly using DJI Onboard SDK APIs.

Stable flight

To solve the problem of weak GPS signals indoor, our team use UWB device and optical flow technology to get the exact position. So we have to use additional PID control loop of position and velocity to make the aircraft take off and land stably.

Mission plan

In order to complete Mission 8, we set finite state machine for aircraft. The state of aircraft is in the loop of Disarm, Armed, Takeoff, Hover, Landing, Mission as shown in Figure 7. By doing this we can achieve any task we planned.

Control flow

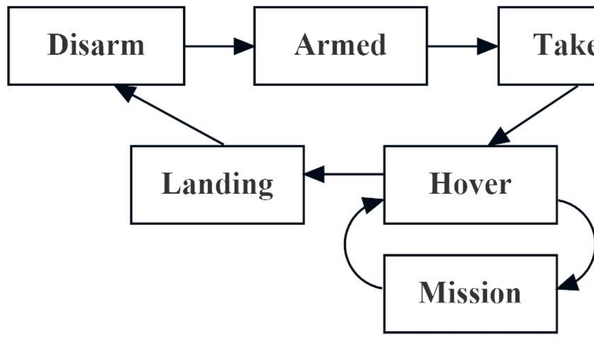

Figure 7: The state of aircraft.

We control the aircraft according to the following procedures. At the beginning we get the paraments of PID control loop and initialize the loop. Then we subscribe topics that provide necessary data, including the status, position, velocity, height, decesion of aircrafts. Meanwhile, we initialize the flight status and obtain the control permission of aircrafts. Finally we come into the state machine.

3.2.2 Optical flow

In Mission 8, in order to achieve accurate formation and target tracking, obstacle avoidance, we need to realize the indoor positioning technology of the drone. So we use UWB technology and optical flow technology here. The common use of technology improves system redundancy and is more suitable for competition requirements.

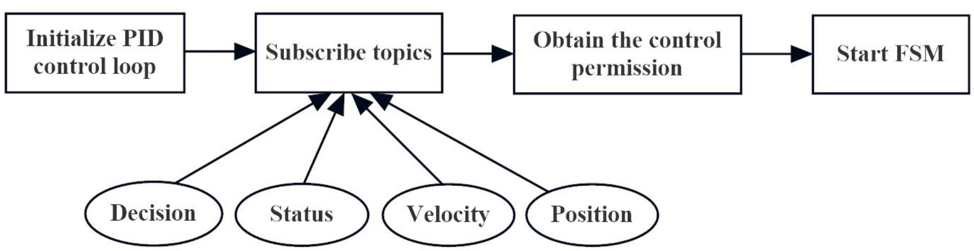

Figure 8: Flow chart.

Capture image

First step is to use the camera to read the image, then the image using the correction matrix to correct image. The radial distortion and tangential distortion could be fixed by calculating the distortion coefficient $(k_1, k_2, p_1, p_2, k_3)$ and lens distortion will be calibrated by the correction matrix $\begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$.

Calculate the feature points

Using the KLT (Kanade-Lucas-Tomasi Tracking Method) method for feature points detection. KLT method means calculate the eigenvalues of the gradient matrix. The point that the eigenvalues considered large and close is the correct feature point we are looking for.

Calculate the optical flow

After the feature point is obtained, the motion speed (pixel speed) is calculated using the Lucas-Kanade optical flow method.

Convert pixel speed to the actual speed

In order to calculate the actual horizontal offset, here we use the principle of pinhole imaging. First you need to know the basic parameters of the camera. The focal length, the actual distance between each pixel, and the distance between the focus and the object, the actual speed can be solved by the principle of similar triangles.

Table 1: Definitions.

f	Focal length
$l1$	Camera pixel distance
d	Distance from object to focus
$l2$	Actual distance
v	Actual speed

We can caculate the distance from object to focus d and actula speed v according to formula (1) and (2)

$$d = \frac{l2}{l1} f \tag{1}$$

$$v = \frac{d}{\Delta t} \tag{2}$$

3.3 Obstacle Avoidance

The main idea of obstacle avoidance is to navigate the aircraft through a planned path from the start point to goal position. Artificial Potential Fields (APF) is a reactive approach to solve this problem since the trajectories are not planned explicitly but obtained while executing actions by differentiating the potential function. A potential function is a differentiable real-valued function whose output can be taken as the energy and its negative gradient as net-force acting on the autonomous vehicle. Two different additive potential functions are presented in this report: attractive and repulsive functions. Therefore, the obstacle avoidance problem can be formulated as finding a function that represents the energy of the system and generating a force on the vehicle so that the energy of the system is minimized only at the goal position.

The intuition behind the additive potential functions is attracting the robot to the goal position and repelling the robot from the obstacles. The potential functions can be constructed as $U(q) = U_{att}(q) + U_{rep}(q)$, which is the sum of attractive and repulsive potentials.

The Attractive Potential: The essential requirement of attractive potential function is that$U_{att}(q)$should be monotonically increasing with the current distance $d(q, q_{goal})$. A combination of conical and quadratic forms of attractive potential function is presented here.

$$U_{att}(q) = \begin{cases} \frac{1}{2}\epsilon d^2 & d \leq d^*_{goal} \\ \epsilon d^*_{goal} - \frac{1}{2}(d^*_{goal})^2 & d > d^*_{goal} \end{cases} \tag{3}$$

Where $d = d(q, q_{goal})$, ϵ is the attraction gain, and d^*_{goal} is the threshold after which quadratic function changes to the conical form. The gradient of attractive potential is obtained by:

$$\nabla U_{att}(q) = \begin{cases} \epsilon(q - q_{goal}) & d(q, q_{goal}) \leq d^*_{goal} \\ \frac{\epsilon d^*_{goal}(q - q_{goal})}{d(q, q_{goal})} & d(q, q_{goal}) > d^*_{goal} \end{cases} \tag{4}$$

The Repulsive Potential: The behavior of a proper repulsive potential function should look like the behavior of a tightened spring, or the magnets getting closer to each other so that the obstacle avoidance can be satisfied. In this report, the following formula is utilized:

$$U_{req}(q) = \begin{cases} \frac{1}{2}\eta(\frac{1}{D(q)} - \frac{1}{Q^*})^2 & D(q) \leq Q^* \\ 0 & D(q) > Q^* \end{cases} \tag{5}$$

Gradient of the function is the repulsive force:

$$\nabla U_{req}(q) = \begin{cases} \eta(\frac{1}{Q^*} - \frac{1}{D(q)})\frac{1}{D^2(q)}\nabla D(q) & D(q) \leq Q^* \\ 0 & D(q) > Q^* \end{cases} \tag{6}$$

where Q^* is the distance threshold for an obstacle to create a repulsion effect on the robot and is the repulsion gain.

The total repulsive potential field can be obtained by summing up the potentials caused by all of the obstacles,

$$U_{rep} = \sum_{k=1}^{N} U_{rep_i}(q) \tag{7}$$

and the movement is realized by following the negative gradient of the sum of attractive/repulsive potentials or simply the following sum:

$$-\nabla U(q) = -\nabla U_{att}(q) - \nabla U_{rep}(q) \tag{8}$$

3.4 QR Code detection

It is not a simple matter to extract four seperated QR code parts.

Firstly, we need to locate to the code display, and get position of display to facilitate the further extraction of features of the code. Secondly, the approximate location area of QR code display is further searched to determine the rotated rectangular containing "the quadrant". Then the convex hull with the largest area is obtained and the approximate boundary point is obtained. Then the boundary points are used to calculate the exact four corners of the boundary. Finally, the corrected figure is obtained through perspective transformation, and the data of QR code is obtained through zbar library identification.

The relative size of the quadrant within the image taken by the plane is not large, and it is relatively easy to extract. Select the appropriate threshold. Since the particial occlusion is pure black, it is easy to extract the general range. For the result contours, we compute and select the one with the largest area as the target.

After obtaining the rough region of the particial occlusion, we use the gaussian blur filter function to operate the binarization image, and use the appropriate convolution kernel to expand the image. The region with the largest area is selected as the concentrated region of the QR code. Finally, the minimum rectangle of the region is used as the rectangular region of the QR code.

The rectangular region of the QR code has included all the information of the QR code, and the convex hull contour with the largest area was drawn. Five corner points were found, including three boundary points and two relay points.

From the Angle of five Angle points more than 110 degrees point concentration, accurate calculation of another corner point. With the other three corner points , we sort the corner points according to their relative positions, and perspective transformation is carried out on the sorted point set. At this point, we have accurately extracted the QR-code. The four QR codes are spliced together , and finally the QR-code data is identified by the zbar library.

3.5 Swarm interaction

3.5.1 Speech recognition

In order to allow our aerial robots to be controlled with non-electronically method, we add speech recognition function for our project. The player on the arena will transmit speech signals to control aerial robots when necessary, thereby guiding the robots to accomplish the mission successfully. The speech recognition module mainly includes two processes: speech recognition and instruction transmission.

We use speech recognition SDK provided by Baidu to complete speech recognition program. When the speech recognition program works, the speech signal will be uploaded to the Baidu server, and the text corresponding to the speech will be received after recognition. With fidelity of sound captured by the microphone and high recognition rate of Baidu API itself, the accuracy of speech recognition can be ensured. In addition, we have designed a set of rules for parsing text into instructions. Based on this set of rules, the program converts the text into a form of instructions to control four drones in the arena.

With the purpose of allowing the aerial robot to concentrate on the task, we use Surface Go as the mobile side to capture speech. Under the ROS structure, the mobile Surface Go can communicate with four aircrafts in the air through topic. The mobile device will publish the message to topic when captured speech signals, then aerial robots can subscribe the topic and perform a corresponding action according to the instruction.

3.5.2 Strategy of aircraft cluster

Aircraft's communication is based on ROS. It has a component called Multimaster-fkie, that is, systems build from two or more ROS networks, each with its own rescore node. It can periodically send multicast messages to the

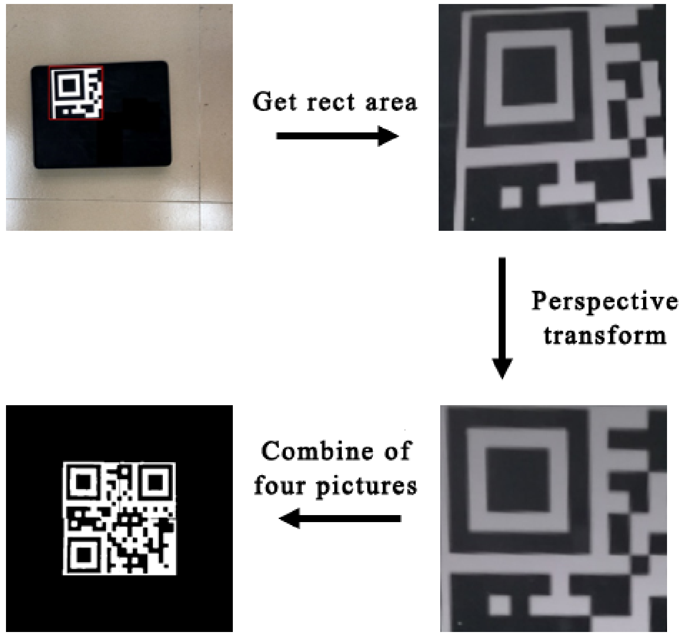

Figure 9: QR code processing flow.

common network to make the other possible ROS masters aware of its presence and detect any other ROS master. The topic namespace "/network/" is how we distinguish the topics inside aircraft and the topics published in the common network, that is, all of the topics in the network have naming format like "/network/". Therefore, the topics inside the aircrafts and the topics in the network will never have conflict and they perform their duties through cooperation.

In order to achieve swarm interaction, we need a cluster module to plan for the routines of aircrafts between speech module and controlling system. That's why we need clustering strategy, called formate module. Our formate module is executed and conducted by aircraft 1, which is the leader aircraft. Figure 10 is the framework of this strategy.

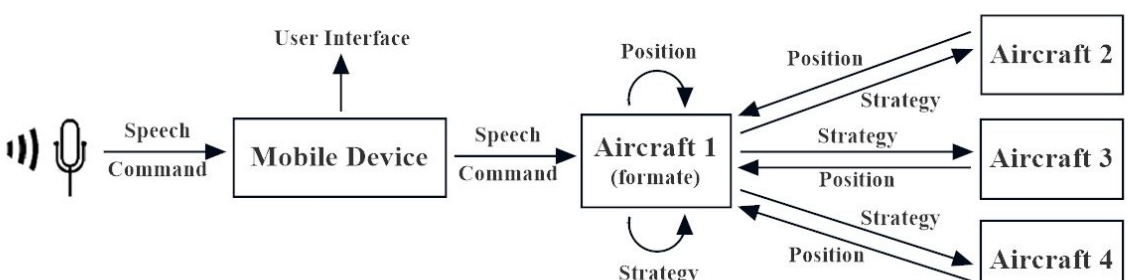

Figure 10: Strategy.

Formate module obtain positions of four aircrafts while running. After receiving speech commands, formate module will translate it into the changes of Aircrafts target positions and send them to four aircrafts respectively. When four aircrafts receive them, they will change them into controlling strategy to reach the targets.

Our basic commands consist of taking off, landing, hovering, going up, going down, shifting, executing preset formations and so on. Moreover, we add some auxiliary speech commands to help aircrafts to accomplish the mission more conveniently, such as catching QR codes, controlling lazers.

4 Experiments

4.1 Simulation

We build a simulation environment based on V-REP. Using ROS Interface in V-REP, we can simulate the scene with some of the critical modules of the project. To simulate real environment, we place ground, aircrafts, screens with QR codes, black boxes and so on in the 3D simulation space. Meanwhile, we exchange information such as current positions, aircrafts' targets, cameras' information by ROS Interface, so we can test network module, speech module, strategy module and QR code detection module. After several tests, we find that our aircrafts can reach the goals above the screen following the fixed route by person in about 4 minutes.

However, the differences between simulated environment and actual environment mainly embodies in positioning error, aircrafts' hardware performance and controlling parameters, etc. So other modules shouldn't be tested in simulation and should be debugged in actual tests.

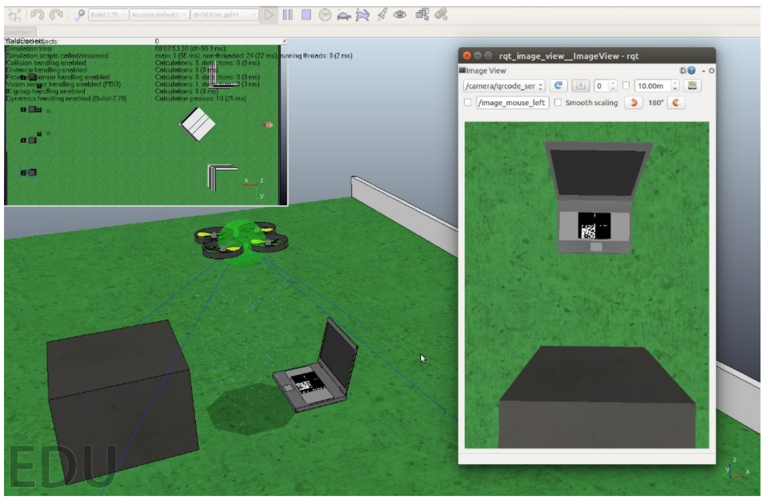

Figure 11: Simulator.

4.2 Sensors and modules test

4.2.1 Onboard sensors

Camera

The camera mounted on the side of the aircraft is looking directly downward, providing video stream containing ground information with 30-60 fps.

Due to the size limit of the camera, certain distortion to images would exists and we have to calibrate the camera output using the data collected from shooting the standard checkerboard pattern when simulating the actual flight circumstance. The calibrated images would be sent to image transport for QR code recognition and optical flow calculation.

Lidar

The DE-LiDAR TF02 is firing downward as well to figure out the exact altitude that the aircraft is located at. This easy-to-use lidar can be tested by comparing the serial output against the manual measurement that read from a ruler. And the precision of this kind of lidar is 1cm. We ultilize it to limit the height of aircraft within 4m.

Precision location

The DJI UWB system is particularly handy in short-range indoor positioning applications. During our test flights, four base stations were set at the corners of the arena and the onboard receiver was trying to find the relative location of the aircraft while communicating with the base stations. The test result could be verified by measure the exact location of the aircraft and compare it with the systems output.

Movement tracking

The inertial measurement unit inside the flight controller combined with the global navigation satellite system module that stick out should be able to report the motion stats about the aircraft, such as angular rate, magnetic field and acceleration.

4.2.2 Self-pilot System

Flight control

The OSDK-ROS interface provided by DJI became the bridge between our applications and the hardwares. We have made necessary modifications to the SDK to meet the real demand, including the PID thrust management. In order to verified the respecting behaviors of the high-level control signals, the simulator was developed to monitor all the actions. The simulator, which is designed in V-REP, reproduced the complete competition arena for the most part. We can preview partial process in the simulator.

Data fusion

To accurately determine the status of the aircraft when it's executing the tasks, multiple raw data sources are considered together to minimize the error. The on-field flight tests were repeatedly performed to collect data for analysis. We deliberately constructed the recorder module for data collection. The tests proved that it's reliable in the competition.

4.2.3 Automatic task distribution

Network

To achieve the communication between each system via WLAN, we utilized the multimaster-fkie package to publish and subscribe topics remotely. Considering the range limit and stability of wireless connections, we tested the connections in real exercises to optimize the location of the base stations and its antenna type.

Mobile device

The mobile module helps provide the real-time information to running team member. The functions of camera view port, voice recognition status and QR code assemble result are verified during our flight test.

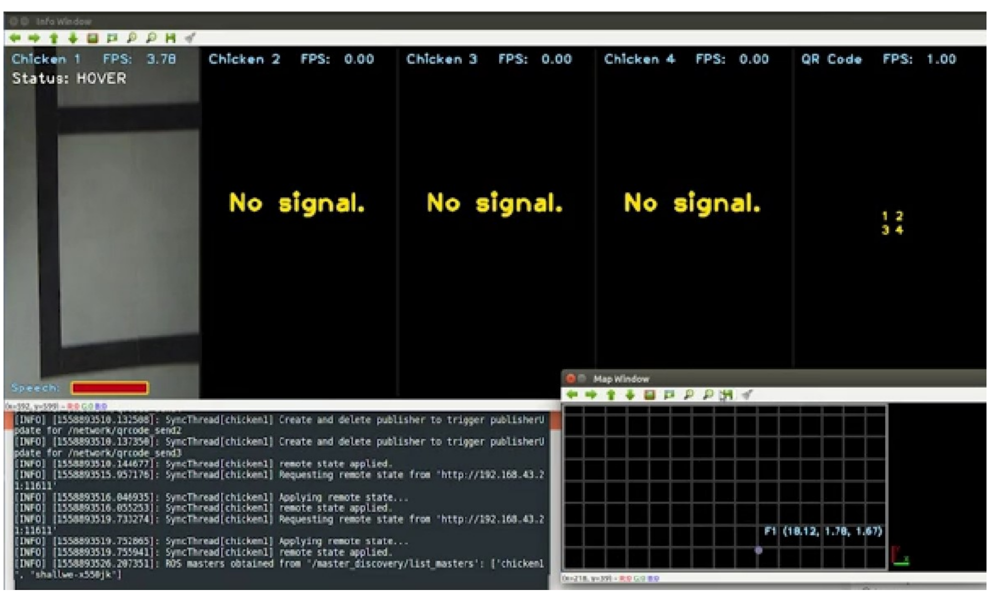

Figure 12: User Interface on Mobile Device.

4.3 Flight test

Eight flight tests have been conducted every week. The flight tests focused on checking UWB Positioning System, QR code detecting and reading system, Optical Flow data collecting and other data collecting system. The tests were conducted both indoors and outdoors, in order to check the system stability in complicated environments. All problems encountered during the flight test have been solved gradually.

Figure 13: Real flight test.

5 Conclusion

In this paper, the technical details of an aircraft system are presented. The Mission is divided in some subproblems at different levels to overcome. We solve each subproblem separately and combine the modules at the end to complete the mission. Up to now, most of the modules are almost finished and need to be tested and combined. The system are expected to operate normally in the competition.