Technical Paper for the International Aerial Robotics Competition-Mission 8

Ziwen LIU, Zongyuan YAN, Zefan LI, Luyao YANG
Nanjing University of Aeronautics & Astronautics, Nanjing, China

ABSTRACT

In the 8th mission in International Aerial Robotics Competition (IARC) a fleet of quad-rotors will demonstrate behaviors, including man-machine interaction via non-electronic command and control, fused sensory enhancement of a human operator by a fleet of aerial robots, swarm interaction, aerial target designation and head-to-head interaction with opposing aerial robots. The Zero-One Team form Nanjing University of Aeronautics and Astronautics has prepared four quad-rotors with added improvements in object recognition, obstacle avoidance, communication and human-robot interaction. The flight platform and technical implementation will be introduced in this paper.

I. INTRODUCTION

1.1 Problem Statement

As the development of the hardware performance and software framework, hands-free interaction has become a forefront in our daily lives. Against this background, the 8th mission of IARC has proposed some enhanced behaviors like non-electronic man-machine interaction and swarm interaction. Our Zero-One Team from NUAA will focus on the voice command reception, target recognition and flight task assignment to complete the task in Mission 8.

1.2 Conceptual Approach

DJI Matrice 100 is used as a platform to develop the aerial robots. We use DJI Guidance as a navigation and obstacle-avoidance module, and Manifold as a onboard computer. With low-latency, high-frequency sensor telemetry, aircraft control, video feeds and powerful SDK, we can easily extend the software and hardware capabilities, and focus on the obstacle-avoidance algorithm, voice command process, object recognition and other technologies that needed to be added or developed.

II. AERIAL VEHICLE

2.1 Platform

DJI Matrice 100 is a stable and reliable flight platform, which has all of DJI's easy-to-fly technology built in. It includes the flight controller, propulsion system, GPS, DJI Lightbridge, a dedicated remote controller, and a rechargeable battery. This system automatically manages the most complex tasks required for flight, so that developers can focus on their own work.

Figure 1 DJI Matrice 100

2.2 Guidance

DJI Guidance is a revolutionary visual sensing system consists of five sensor modules and one central processor. Even without GPS, DJI Guidance can achieve hovering that is accurate to within centimeters. Even when flying at high speeds, high-precision stereo algorithms provide positioning information over nearly any terrain. Its vision positioning system is effective at altitudes of up to 20 meters. Besides, 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 void collision, even at high speeds.

Figure 2 DJI Guidance

2.3 Onboard Computer

The Manifold is a powerful onboard computer that transforms drones into autonomous robots that turn your vision into reality. It is a high-performance embedded computer specially designed for the DJI Onboard SDK. The Manifold is expandable and easily integrated with DJI's enterprise platforms, making it the go-to computer for any user who wants to build customized, autonomous drone solutions.

Figure 3 DJI Manifold

2.4 Audio Receiver and Language Process Model

ReSpeaker 4-Mic Array for Raspberry Pi is a quad-microphone expansion board for Raspberry Pi designed for AI and voice applications. This means that we can build a more powerful and flexible voice product that integrates Amazon Alexa Voice Service, Google Assistant, and so on. It has 4 microphones, 3meters radius voice capture, 2 grove interfaces and software algorithm: VAD (Voice Activity Detection), DOA (Direction of Arrival) and KWS (Keyword Search).

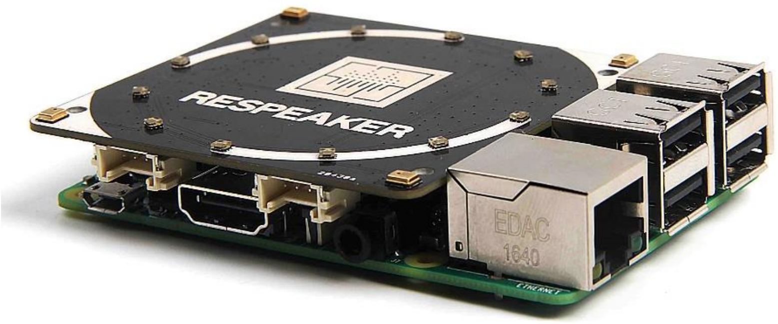

Figure 4 ReSpeaker 4-Mic Array

III. SOFTWARE SYSTEM OVERVIEW

3.1 Localization

The aerial robots use the DJI Guidance positioning module. The Guidance uses a high-precision stereo vision algorithm with a near-ground positioning accuracy of to centimeters. Positioning information can be provided in complex terrain and high-speed flight conditions, The effective height of the visual positioning system is up to 20 meters. In addition, DJI Guidance can output speed, acceleration and attitude information, fusing with onboard IMU and compass we can get accurate position and navigation information from Guidance.

3.2 Natural Language Process and Control

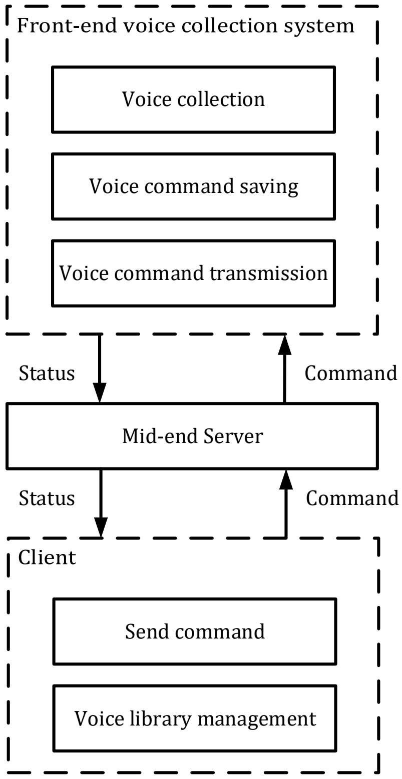

Figure 5 Voice Transmission Structure

The structure of the voice transmission system based on the mobile communication network is as shown in figure 5.

The voice collection system collection system mainly implements three functions of voice data collection, voice data storage, and voice data transmission. The voice collection system completes voice collection, compression coding, saving, and transmission according to user requirements. Besides, it can realize communication with the server, voice file uploading management and other functions.

The server mainly completes the data transfer between the front-end voice collection system and the client. Since the network address of the client and the front end is not fixed in a large network environment, a transit design is required to complete the information interaction. And the server has its own storage space, so that it can also be used as the mid-end storage of voice data.

The client mainly completes the real-time interaction between the system and the user. The client can realize real-time control of the front-end acquisition module and real-time monitoring, downloading, and saving of voice data. When the client receives the user's control command, the client establishes a connection with the server to send the user instruction to the server. The client can also display status information of the front-end device, and send commands to manage voice files in the server according to the needs of the user.

3.3 Obstacle Avoidance

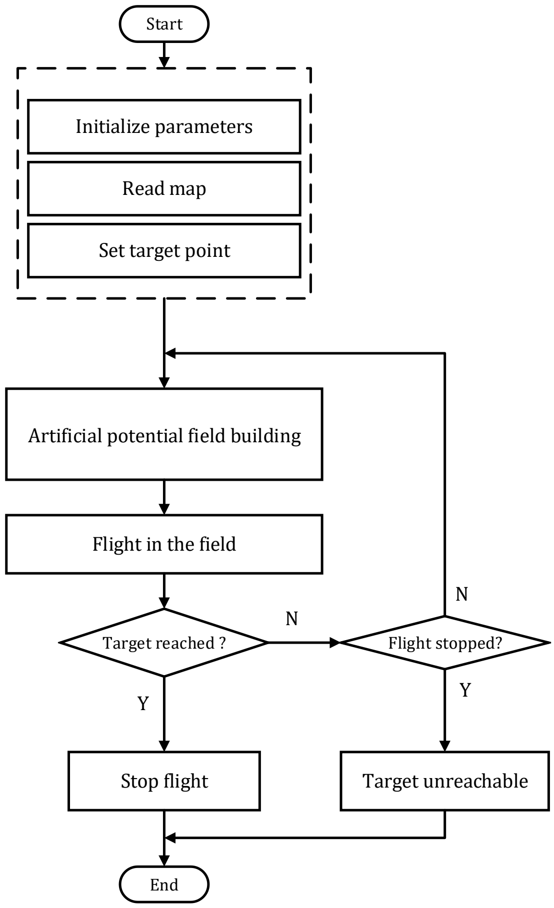

Figure 6 Obstacle Avoidance Algorithm Structure

In the case that the aerial robots can perceive the surrounding obstacles, the path of the drone to the target point is automatically planned by the artificial potential field method. This method offers navigation information by setting the force field function in the planning space and finding the gradient (combination force) on the resultant field. Setting the gravitational field at the target position, the gravitational field will affect the whole space. The distance between each position and the target point will have different gravitational field values, thus generating gravitational force toward the target. Setting the repulsive field at the obstacle position, the repulsive field is only a certain range of action, which produces a repulsive force that keeps the object away from obstacles, thereby preventing the object from colliding with the obstacle. The object will move along the gradient of the resultant force field under the combined force of gravity and repulsion. If the object doesn't fall into the abnormal situation of the local minimum point, the object will eventually reach the target. This method is intuitive and computationally fast, but the artificial potential field method also has some limitations. The most representative limitations are the local minimum and the unreachable target.

In order to enhance the ability of artificial potential field method to plan path and avoid obstacles, the traditional force field function is improved, and a new gravitational field function and repulsive field function are introduced. Figure 1 shows an algorithm for artificial potential field path planning.

3.4 Communication and Collaboration

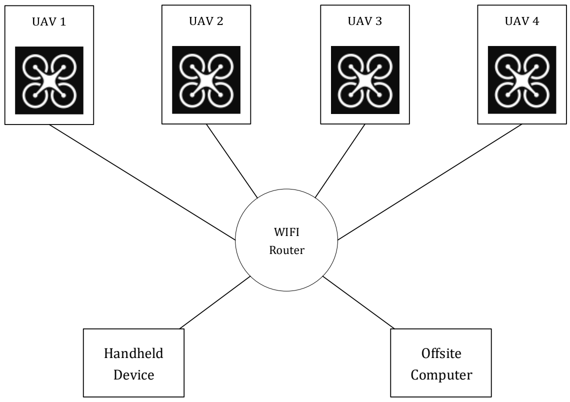

Figure 7 Communication Framework

As shown in the figure 7, the communication structure consists of four parts: WIFI router, handheld device, a fleet of aerial vehicles and an offsite computer.

The WIFI router adopts 802.11n standard, the theoretical maximum transmission speed reaches up to 600Mbit/s, and can support the mutual communication of any two terminals.

The handheld segment can receive and display the video and other information needed for the mission transmitted by four aerial robots, and the handheld terminal can also receive the task information (such as QR code decoding information) processed by the offsite computer. Due to the limitation of the non-electronic command and control, the handheld won't send any electronic control command but the holder's non-electronic information (such as voice or gesture) can be collected and transmitted to the offsite computer.

The offsite computer needs to receive the voice signal from the task executor, and the intelligent processing and analysis is converted into control commands and send to four aerial robots, such as receiving a voice command like "Robot one take a picture". Offsite computer can recognize this command and assists the robot to make decisions. Besides, offsite computer also needs to receive data from aerial robots, such as a part of QR code image, and can complete the splicing and processing of the missing QR code.

Beside transmitting real-time data to the task executor or offsite computer, four aerial robots need to transfer their working status to each other, such as "the enemy is approaching robot 1" or "robot 2 finishes recognizing QR code". In addition, four aerial robots also need to accept real-time instructions from the task executor or computer, such as need to rescue, landing control and other control instructions.

3.5 Object and QR Code Recognition

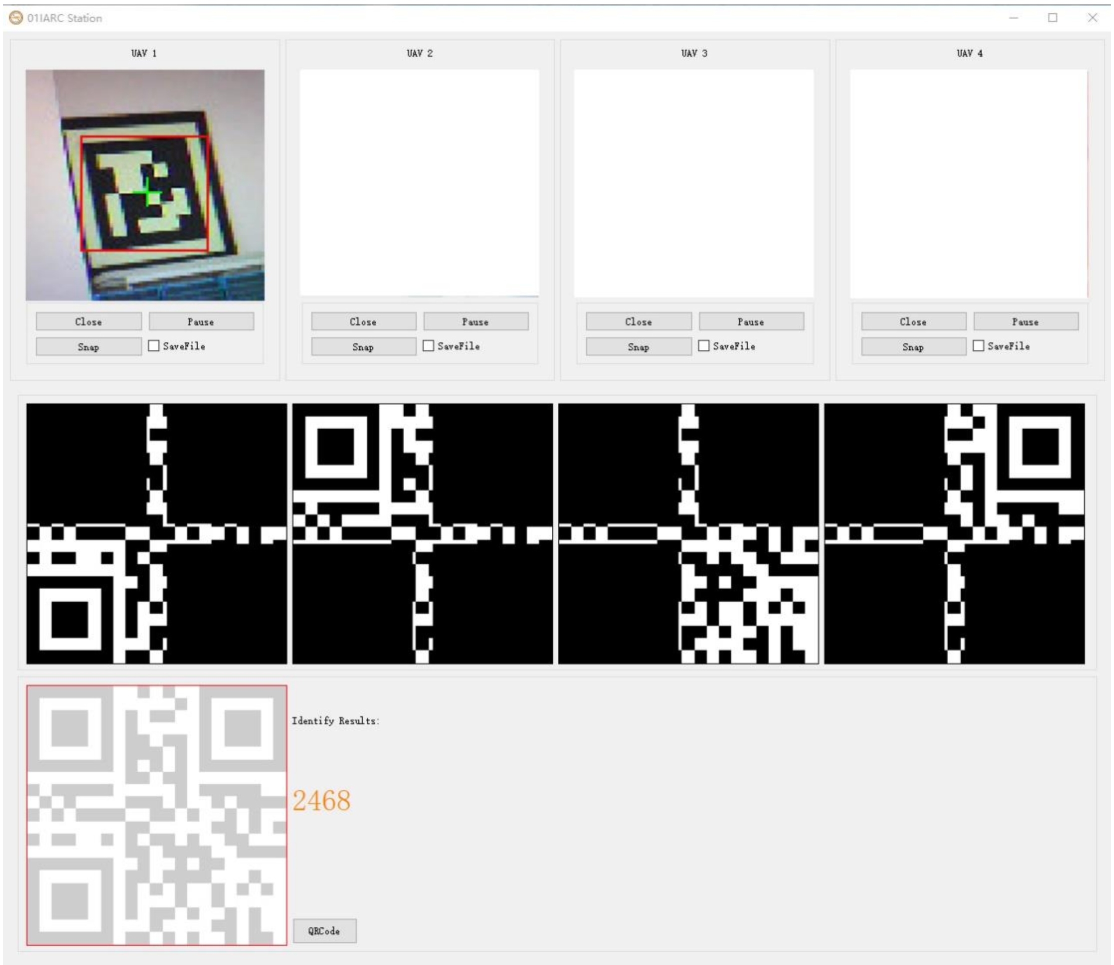

Figure 8 Handheld Device Software

As shown in the Figure 8, we develop a handheld device software to receive each aerial robot's first person view image to help combine four different parts of the QR code. The FPV image can be displayed in the square under "UAV1" ~ "UAV4", when an aerial robot recognition part of the QR code it will hold the position and draw a red rectangle for the QR code (or part of QR code). After that image in the red rectangle will be snapped and displayed in the lower part of the software. Finally, four parts of the QR code will be combined and the information will be extracted to assist competitor getting object in the target case.

IV. CONCLUSION

In this paper, we presented the technical details of four quadrotor system capable of receiving voice command and control. The quad-rotor aircraft's platform is DJI Matrice 100, the indoor navigation and obstacle-avoidance module is DJI Guidance, Manifold is implemented as onboard computer to process man-machine and machine-machine interaction. Besides, we add voice process module to send non-electronic command. The experiment results show that our quad-rotor aircraft can autonomously navigate in unknown indoor environments without GPS signals, and can response the non-electronic command to finish the appointed tasks.

REFERENCE

1. Harris C, Stephens M. A combined corner and edge detector[C]. Proceedings of the 4th Alvey Vision Conference. 1988:147-151.

2. Horn B K, Schunck B G. Determining Optical Flow. Artifical Intelligence, 1981, 17(1-3): 185-204.

3. Lucas B, Kanade T. An Iterative Image Registration Technique with An Application to Stereo Vision. Proc. of DARPA IU Workshop, 1981: 121-130.

4. Nister D. An efficient solution to the five-point relative pose problem. IEEE Transactions on pattern analysis and machine intelligence, 2004, 26(6):756-770.

5. Guo Li, Ang Haisong, Zheng Xiangming. Monocular Vision based Indoor Micro Air Vehicles' Motion Estimation and Structure Recovery. Journal of Nanjing University of Aeronautics & Astronautics. Vol.44, No.2, 2012.

6. Levenberg K. A method for the solution of certain non-linear problems in least squares. Quarterly Journal of Applied Mathematics II (2) (1944) 164–168.