Warehouse Small Cargo-carrying UAV Design and Environmental T265 Camera Placement Angle Study

Ang He[1], Xiangda Wang[2], Xinyu Song[1], Hongwei Huang[1, *], Peng Liu[3]
[1]*College of Information Engineering, Shanghai Maritime University, Shanghai, China*
[2]*School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China*
[3]*College of Transport & Communications, Shanghai Maritime University, Shanghai, China*
**Corresponding Author.*

Abstract: The Intel RealSense Tracking Camera T265 is a tracking camera that uses its own Vi-Slam algorithm to output horizontal and vertical coordinates, which has a wide range of applications in drones, unmanned boats, and unmanned vehicles. In this paper, we utilize the T265 and the Pixhawk4 (PX4) flight controller to build a Robotics Operating System (ROS)-based, fixed-point cruise-capable, self-taking-off and landing unmanned aerial vehicle (UAV) for cargo loading. The T265 is a fisheye black and white camera with powerful visual simultaneous localization and mapping (Slam) localization and a large visual range. The T265 features the Movidius Myriad 2 visual processing unit. Slam construction based on the comparison of feature points to output coordinates. Nevertheless, when the environment has a single color and insufficient feature contrast, it is easy to cause a SLAM error, leading to drifting of the drone's coordinates, which is very dangerous. It is easy to injure drone operators and other pedestrians accidentally, and at the same time, it is often accompanied by drone crashes that cause economic losses. We experimented with placing the T265 camera at multiple angles to test the best placement angle for more accurate drone positioning. The article also describes the mechanical mechanism design, hardware and software design of the cargo UAV, the optimal placement angle was verified in the environment described in the paper.

Keywords: T265; Unmanned Aerial Vehicle (UAV); Visual SLAM; Robot Operating System (ROS)

1. Introduction

Autonomous drones are being used more and more widely, both for military and civilian purposes, and when cruising indoors at a fixed point, precise coordinates and positioning become the key to whether or not the drone works well [1]. Unmanned Aerial Vehicle (UAV) technology is becoming increasingly important in the aviation and all kind of industry [2], Intel RealSense T265 uses proprietary VI-SLAM algorithms to estimate linear and angular position and velocity [3] which Can be used as a perceptron for indoor autonomous cruising UAV. In the field of small cargo drones, the structural design of the cargo drone directly determines the weight of the cargo that can be carried. On a micro drone with a motor diameter of 410mm, we have an on-board computer, fixed altitude radar, T265 camera, pixhawk4 flight control, servos, and more to achieve what we want to achieve. The T265 camera is prone to Slam error when the environment has a single color and the environment has fewer feature points, such as in a warehouse, resulting in the UAV's coordinates shifting, and in the automation of drone programs. The T265 data was collected using custom Python software that utilized the pyrealsense2 library from Intel (version 2.36.0). This software recorded streams from the accelerometer and gyroscope, as well as the VI-SLAM position and velocity estimates, and saved them to disk in a binary format. The information captured by the OTS cameras was initially sent to Vicon Blade software. In Blade, a rigid body was fitted to the infrared optical marker data. The positional data of this rigid body was then transmitted through the Robot Operating System (ROS) middleware using a custom wrapper. Data from the IMU inside the perambulator were read out and processed using a custom software suite written in Julia, developed by Trium Analysis Online GmbH [4]. During the automation of the drone programs, the consequences of drone coordinate shift can be fatal. When the UAV has only one sensor T265 or the data weight given by T265 in the T265 fusion algorithm is large, it is easy for the UAV to lose control of the phenomenon, causing danger to the UAV pilot and pedestrians, and also causing economic losses, therefore, when using a single sensor T265 and the UAV is in the environment with a single color and fewer environmental characteristics, to find out the best angle for the UAV to be placed is necessary, which may be helpful for the underfunded groups that have difficulty in doing fusion algorithms. At the same time, we also give a micro UAV mechanical structure design scheme, which may be helpful for the power system, low gravity UAV models want to complete the autonomous movement, autonomous cruise delivery of goods.

2. System Structure

In this paper, the hardware of our UAV mainly consists of Pixhawk4 flight control, flight power system, on-board computer, fixed-height radar, and T265 black-and-white fisheye camera. Our onboard computer is the NUC11PAHi7 equipped with 32GB operating memory, providing powerful computing power for algorithmic calculations. The system is UBUNTU 18.04 with Robot operating system (ROS).Using Tiger Motor MT-2216 to make power. The finalization radar takes the TFmini-S LIDAR ranging sensor, we will use the T265 to get the horizontal and vertical coordinates of the UAV state, and use the finalization radar TFmini-S to get the altitude coordinates of the UAV, so that the UAV doesn't fly out diagonally when the T265 has a Slam error. The hardware structure of the airplane is shown in Figure 1.

3. UAV Mechanical Structure Design

For small cargo-carrying uav, the mechanical structure design is very important to be able to carry the weight of the cargo, including the loading space at the bottom, while the use of each hardware and components should be calculated for the weight to match the lifting force of the motors, so as to avoid the lack of lifting force that prevents the drone from taking off. At the same time, we have to consider where to put the on-board computer (NUC), NUC compared to the ordinary on-board drone microcontroller such as raspberry pi or jetson nano, the volume is larger, but provides a higher arithmetic support, so we also need to solve the NUC larger volume, weight and the problems arising from the larger. So we shelled the on-board computer (NUC) and placed it above the Pixhawk flight controllers and placed the altimetry radar (TFMINI) between the battery and the upper connector plate, and mounted the model LiPo battery and the cargo carrier underneath.

Figure 1. Hardware Structure Diagram of UAV

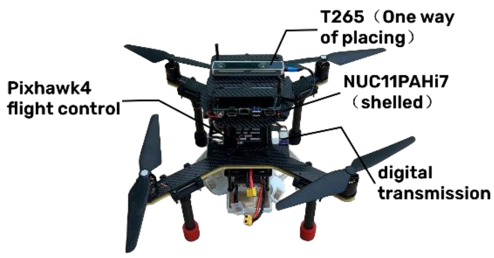

Figure 2. Frontal Structure of the UAV

Figure 2 shows one way of placing the T265, and we will introduce other placement angles and experimental environments later. The drone's diagonal two brushless motors are 410mm apart and the paddles are 10-inch paddles.

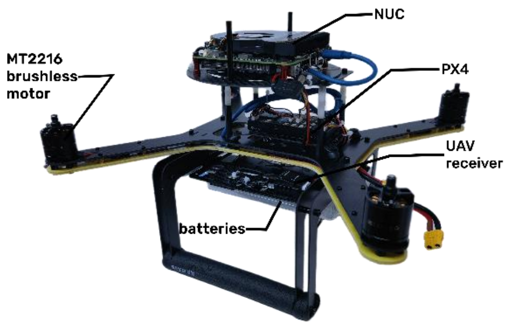

Figure 3. Side Structure of the UAV

As shown in Figure 3, this is the side structure diagram of our designed UAV, which shows the structure of the UAV in a more three-dimensional way, and through Figure 3, we can see the placement of our UAV's receiver, battery, flight control, and on-board computer.

4. QGroundControl

QGroundControl, an open source simulation environment for flight control of multiple aircraft [5]. It allows operators to build complex tasks, including tasks and other program items, by adding a task designer.

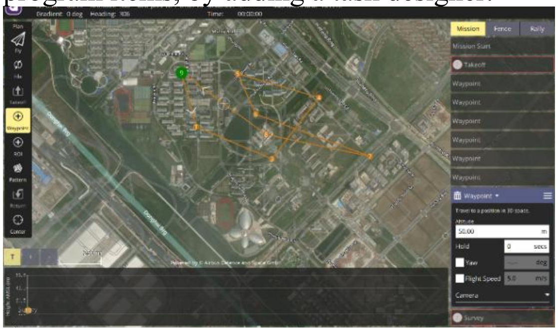

Figure 4.QGroundControl Interface Schematic

As shown in Figure 4, taking Shanghai Maritime University Lingang Campus as an example, QGroundControl can be used to plan the trajectory of the drone according to the GPS to set the cruise point, and we also need to use QGroundControl to update the firmware and Vehicle Setup, which also includes the channel adjustments and command definitions.

5. Improving Kannala-Brandt Model

Fisheye camera lenses do not strictly adhere to the assumptions of specific projection models, such as isometric, isostereoscopic, orthogonal, stereoscopic, or linear projections. Instead, fisheye lenses exhibit their own unique projection characteristics. These lenses are designed to create wide-angle images by capturing a very wide field of view. Therefore Kannala-Brandt [6] proposed a general form of estimation for different types of fisheye cameras, and Kannala-Brandt model includes four distortion parameters $[k_1\, k_2\, k_3\, k_4\,]$ and four other parameters $\left[c_x\, c_y\, f_x\, f_y\right]$ [7]. The parameters of the model are schematically shown in equation (1)(2)(3)(4).

$$\Pi(x,i) = \begin{bmatrix} f_x d(\alpha) & \frac{x}{r} \\ f_y d(\alpha) & \frac{y}{r} \end{bmatrix} + \begin{bmatrix} c_x \\ c_y \end{bmatrix} \quad (1)$$

$$r = \sqrt{x^2 + y^2} \quad (2)$$

$$\alpha = atan2(r,z) \quad (3)$$

$$r(\alpha) = k_0(\alpha) + k_1\alpha^3 + k_2\alpha^5 + k_3\alpha^7 + k_4\alpha^9 \quad (4)$$

In Equation 4, α is the angle of incidence of the light ray and $r(\alpha)$ is the projection of the incident light ray on the fisheye image plane.

The camera internal reference matrix represent-tation is shown in Equation 5:

$$\begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \quad (5)$$

The f_x in pixels in the matrix is the ratio of the camera focal length f to the camera cmos parameter d_x. The physical meaning of d_x is the actual length of each pixel, which can be in mm/pixel. c_x, c_y denote the camera principal points, i.e., the coordinates of the intersection of the optical center with the image plane, in pixels. However, in warehouses or similar areas with a single color and few features, a slam error may result.

6. Experimental Environment

In this paper, the experiments and the environments discussed in the paper are mainly warehouses or similar scenarios with a single color (Black and white from T265 view, taking grayscale) and few feature points. In the process of competition or scientific research, we usually put cross labels with different colors into the ground to help the VI-Slam algorithm of T265 in order to perfectly reflect the excellent features of T265, and try to make T265 not lose coordinates.

Figure 5 shows the environment in which our UAV is conducting the experiment, (a) an image of the UAV while it is in the air and (b) an image of the UAV while it is flying close to the ground. Turning the UAV camera downwards may result in a small image input view, which will be discussed in more detail later.T265 sets the position of the drone at the beginning of the program as the origin, and then calculates the trajectory, coordinates of the drone by its own algorithm and imu. So when there are no extra sensors for fusion, choosing the right camera placement is likely to be effective in reducing the number of slam error occurrences.

7. Positioning of the T265 Camera and Analyze

We tried a total of two T265 camera placements as **Figure 6** showed, directly below the nose of the drone, and directly above it, and we tried to analyze through these why it produces a slam error causing the drone's coordinates to shift severely.

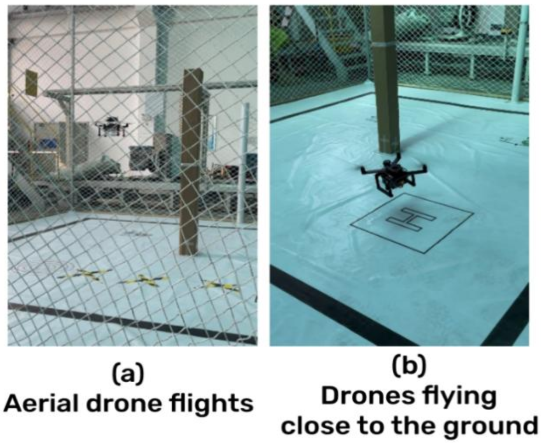

Figure 5. UAV Experimentation Environment (Warehouse)

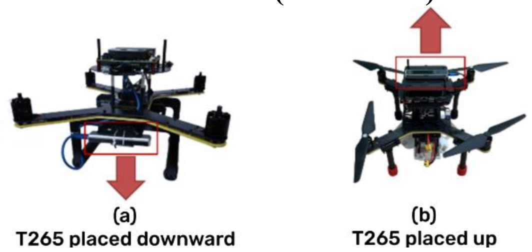

Figure 6. T265 Placement Down and Placement Up

We put the drone in the experimental field with everything ready, and we will experiment five times with both T265 placement modes. We set up five coordinate points, including the drone's left-right movement and up-down movement to test it, and the image received by the T265 fisheye camera will be changed in the process of moving. We will write the written speed and path into a program, and control the attitude of the aircraft through the PX4 flight controller, fixed-point cruise, at the same time, if a Slam error occurs, there will be obvious prompts, and at the same time, the probability of coordinate drift, which means that the UAV does not follow the prescribed route.

Table 1. Comparison of T265 Placed Up and Down on a Drone

	1	2	3	4	5	Fulfillment(s)
down	F	F	F	S	F	1
up	S	S	S	S	S	5

The S in Table 1 stands for success, i.e., all five navigation points we designed from start to finish were successfully completed and landed, and the F stands for failure, i.e., there was a Slam error. We can find that the success rate of the T265 facing up is much higher than the success rate of the T265 mounted with the nose of the drone facing down. It is worth mentioning that in our experiments, we found that when the T265 was mounted with the nose facing down, it tended to have a slam error when taking off to the first preset point, and we felt that a possible reason for this was that the height of the drone was shorter, and the distance between the T265 camera and the ground was closer, so that there were fewer feature points to see, fewer items to characterize, and therefore the target was often lost.

Figure 7. Expansion of Analog T265 Image Input

Figure 7 shows an image of what we can see when we simulate the T265 facing downward, with the visual input angle of the T265 getting larger and more feature points and colors as the drone is raised. Since the Realsense viewer is not good to open during the drone flight, so we use a normal image to replace the black and white fisheye image of the T265 input, of course the display may be more complicated because the angle of the fisheye is bigger, and it is converted into a planar image by mathematical formulas, which can be clearly seen that the image input is extended, and it can be seen that the color of the input image has become richer (equivalent to grayscale in T265), and there are more feature points, but according to the experiments, often the drone will lose coordinates at the beginning of the takeoff, i.e., there will be a Slam error from the first image to the second image, which will lead to the drone blowing up, so there should be enough feature points present at the beginning of T265, and there will be a Slam error between the first and second image, which will lead to the drone blowing up, and there will be a Slam error between the first and second image, and there will be a Slam error between the first and second image. Therefore, there must be enough feature points at the beginning of T265 takeoff. When T265 is placed on top of the airplane and facing

upwards, there will be more feature points because there will be lights 4on the roof of a warehouse and nets in a flight lab, so it is more appropriate to place T265 upwards in this kind of environment.
Additionally, the light will also have an impact on the T265's slam build, and strong light exposure may make the input image true, so it is also necessary to try to avoid strong light exposure.

8. Path Plainning

Unmanned aerial vehicles (UAVs) are one of the most challenging and promising technologies in the field of aeronautics, and path-planning tasks are fundamental to UAV mission execution [8-9]. In this project, in order to ensure the delivery drop point of the delivered goods, our UAV path planning method is relatively simple, using a fixed-point cruise, when arriving at the coordinate point where the items need to be delivered, the items are vertically dropped down through the servo.

9. Conclusion

In this paper, we introduce a hardware building scheme and related hardware facilities for a small warehouse cargo-carrying UAV, and also for the T265 black-and-white fisheye camera carried by the UAV to be placed at two separate angles and compare their success rates, which proves that the unique VI-SLAM algorithm of the T265 requires visual image inputs with many feature points. And in the environment described in the article, mounting the T265 upwards towards the top might be a better option. The downside is that in this setting, we have only examined two angles of T265 placement, and there may be other, more appropriate placement angles. Through this paper, related research workers can use the conclusions drawn in this paper as a reference to be used as a basis when building similar UAVs again. The effect of a single sensor as a sensing unit of a UAV is limited, so multi-sensor fusion may be one of the future research directions, and fusing multiple sensors can help to help the UAV to better localize, reduce the number of failures, and run the program more stably.

References

[1] Lea Matlekovic, Filip Juric and Peter Schneider-Kamp, Microservices for autonomous UAV inspection with UAV simulation as a service, Simulation Modelling Practice and Theory, Volume 119, 2022, 102548, 1569-190X.
[2] Kong Foo Keong, Chai Chang Sa'ar, Suaathi Kaliannan, and Nurul 'Ulyani Mohd Najib, The Adoption of Unmanned Aerial Vehicles (UAV) Technology in the Construction Industry: Construction Stakeholders' Perception, IOP Conference Series: Earth and Environmental Science, Volume 1238, 012024, page 1-3.
[3] Peter Hausamann, Christian B. Sinnott, Martin Daumer, and Paul R. MacNeilage, Evaluation of the Intel RealSense T265 for tracking natural human head motion. Scientific Reports volume 11, Article number: 12486 (2021).
[4] Peter Hausamann, Christian B Sinnott, Daumer, Martin Daumer and Paul R MacNeilage. Evaluation of the Intel RealSense T265 for tracking natural human head motion. Sci Rep 11, 12486 (2021).
[5] Cristian Ramirez-Atencia, David Camacho, Extending QGroundControl for Automated Mission Planning of UAVs, Sensors, 2018; 18(7): 2339.
[6] J. Kannala and S. S. Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. IEEE transactions on pattern analysis and machine intelligence, 28(8):1335–1340, 2006.
[7] P. Furgale, J. Rehder, and R. Siegwart. Unified temporal and spatial calibration for multi-sensor systems. In 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 1280–1286, Nov 2013.
[8] Yijing Zhao, Zheng Zheng, Yang Liu, Survey on computational-intelligence-based UAV path planning, Knowledge-Based Systems, Volume 158, 2018, Pages 54-64, ISSN 0950-7051.
[9] Zhangjie Fu, Jingnan Yu, Guowu Xie, Yiming Chen, and Yuanhang Mao. A Heuristic Evolutionary Algorithm of UAV Path Planning, Wireless Communications and Mobile Computing. Volume 2018, Article ID 2851964.

