Autonomous quad copter control for person tracking

Sieuwe Elferink

1 Introduction

Autonomous quad copters are currently a topic of interest in the research community. This is because they can be useful in situations like delivering packages or finding casualties after a natural disaster.

Fontys is also interested in researching Autonomous quad copters. For this the research project *Autonomous drone's* in cooperation with Delta and the professorship of Teade Punter was created. To demonstrate autonomous control capability in this project a Proof Of Concept is created. In this Proof Of Concept a quad copter is tasked in following a moving person in a open field fully Autonomous.

This document will describe the system created to demonstrate this Proof Of Concept. The document will first describe the sensors and hardware used of processing the data. Then the AI based vision pipeline and control structure are described. As last performance results and recommendations for improving the system are described.

2 Sensors

When looking at a quad copter there are 4 axis of movement. These are Roll, Yaw, Pitch and Height. While all are controllable using software, they are not all needed. To achieve any location in 3D space only Yaw, Pitch and Height are needed. This is because the quad copter can fly in a polar like ordinate system by having its angle changed using Yaw and its distance changed using Pitch. The height can be adjusted by throttling all motors. This already removes the Roll axis greatly reducing the complexity of the control system.

All 3 axis need to be controlled to achieve fully autonomous flight. This needs to be done based on sensor data from its surrounding. The system will consist of 3 main sensors. These are: A RGB based camera for the Yaw axis, A lidar based distance sensor for the Roll axis and a barometer for the height.

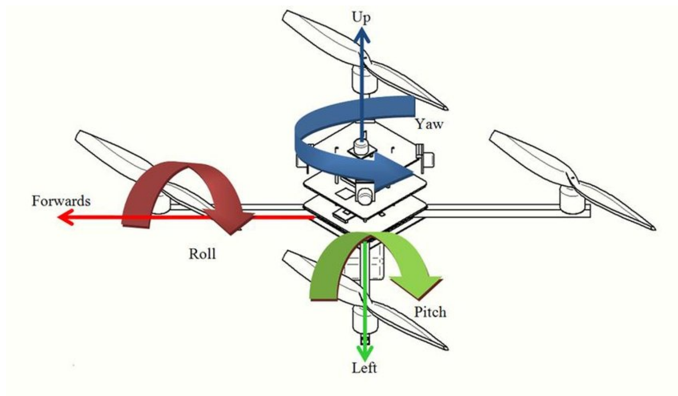

Figure 1: 3 axis of movement for a quad copter

The RGB camera uses a IMX219 Sony sensor capable of 1080p at 60fps. This makes it together with its CSI interface possible to quickly sample the surrounding environment to have smoother control. The camera module used is also fitted with a wide angle 160 fov lens making it capable to monitor a larger area. This makes it better to find targets which are not in the center of the quad copter.

The Lidar used is a TF Luna solid state Lidar sensor. It has a range of up to 8 meter and a accuracy up to 1m. Also the sample area of 2.5cm * 2.5cm makes is suitable for larger targets. A non solid Lidar which can sample 360 degree is not used since distance is only needed for one angle which is in front of the quad copter. Having a 360 Lidar will only complicate the system.

3 Processing Hardware

Having a as low as possible Latency for the entire control pipeline is very important. The higher the Latency is the slower the system will respond. In order to decrease latency all sensor data processing and control is done on the quad copter. To achieve this a small footprint and light yet powerful computer is needed. For this task the Jetson Nano 4GB is used. It has great performance in for example AI because of its integrated GPU. It is also small, light and can be powered using 5V. It also has some embedded features like exposed uart pins making it capable of communicating with the TF Luna Lidar.

4 System Architecture

The control system takes the sensor data from the sensors described above and will process these into positional movement commands that the quad copter can perform. The system works in multiple states as shown in Figure 2.

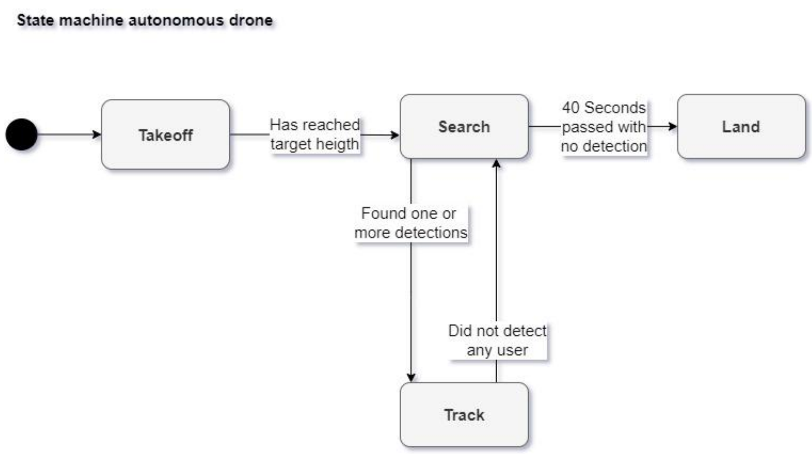

Figure 2: Control system state diagram

In the *TakeOff* state the quad copter will perform flight checks. After this the quad copter will arm its motors and takeoff. The state is excited when the quad copter reaches it target altitude.

In the *Search* state the quad copter will attempt to find a person to track. First the quad copter hovers for 10 seconds in the same location while running an AI based object detection model to find a person in the camera view. If after 10 seconds no valid detection's are found, then the quad copter will perform a slow rotation in the yaw axis to find a person. If after 30 seconds no valid detection's are found, then the quad copter will go to the *Land* state. If a valid detection is found within the 40 seconds, then the quad copter will go into the *Track* state.

In the *Track* state the quad copter will run the main vision and control algorithms described in section Vision And Control. The quad copter will use the sensor data from the lidar and RGB camera to keep the person in the center of the quad copter. If the person is lost during tracking, then the quad copter will return to the *Search* state.

In the *Land* state the quad copter will perform a landing in the current location. After landing the quad copter will be disarmed and the program will exit.

5 Vision And Control

The Vision And Control system is responsible for processing the sensor data and converting this data into usable movement commands for the quad copter's flight computer. The system proposed in this document is explained below.

First the raw camera feed is read using a csi interface and down scaled to 300 * 300 pixels. After this a TensorRT optimized CNN model from Google called Mobile-net is used to find objects in the camera feed. Because the Mobile-net network is trained on more classes then only humans, the detection's are filtered to only include humans.

From the list with detection's the first detection is used. From this detection the 4 corner coordinates of the surrounding bounding box are used to calculate the center point of the detection. A second static point that represents the center of the camera feed has been created at the startup of the system. This static point represents the quad copters heading since the camera is mounted in line with the front of the quad copter. The distance between the two center points in the X axis is calculated. This distance becomes larger when the detection is further away from the center of the drone and becomes closer when the detection is closer to the center of the drone. This makes the value suitable for the yaw control.

To enable control in 3 dimensions, depth vision is needed. This is achieved using the solid state Lidar mounted in the same direction as the RGB camera. This Lidar is positioned and calibrated such that the center point of the RGB camera feed overlaps the 2.5cm*2.5cm IR square of the Lidar at a distance of 2.5 meter. This makes that the object where the center of the RGB camera feed is pointed has the distance given by the Lidar at that moment. To make sure the Lidar only gives valid depth information when the lidar is pointed towards a person, the Lidar distance will only be used when the center of the RGB camera is within the bounding box of the detection being tracked. Since the lidar distance represents the distance between the quad copter and the person being tracked, it can be used for controlling the pitch axis.

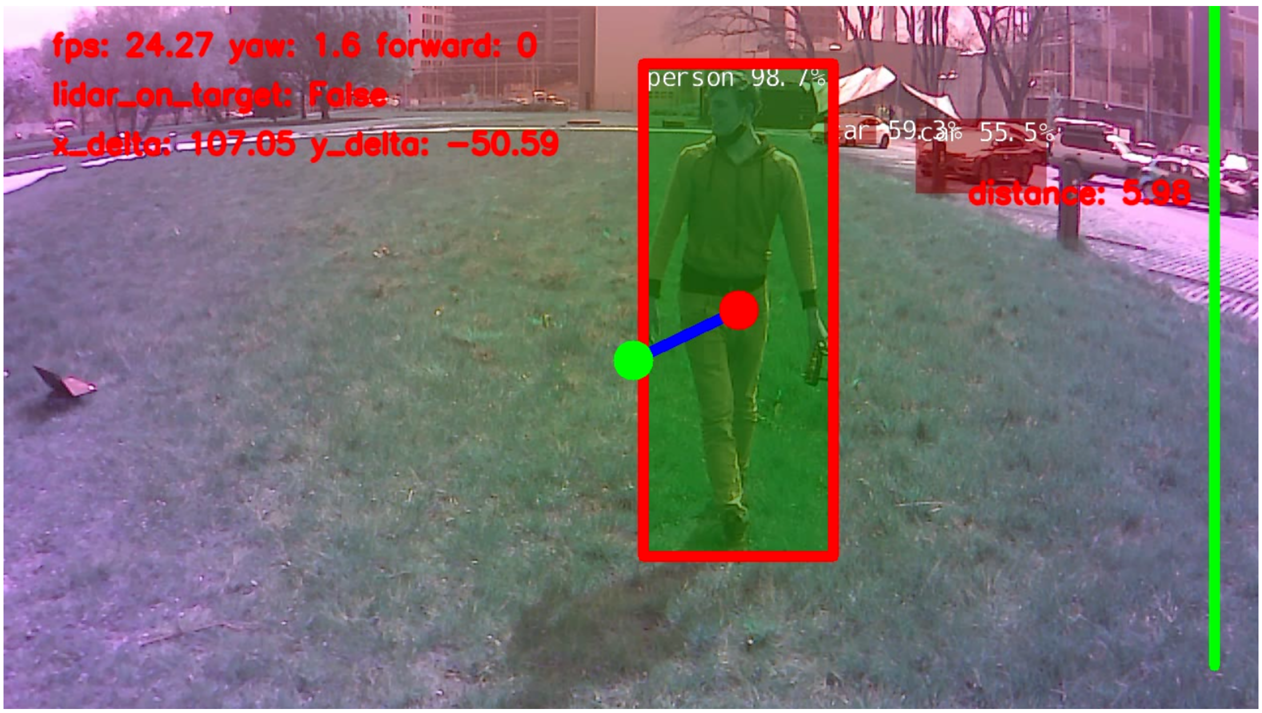

Figure 3: Visualization of the vision system

The center point distance and the Lidar distance are both filtered using two separate moving average filters of 5 steps. This removes spikes in the raw data because of for example faulty model detection's or faulty Lidar data such that they don't disturb the control of the quad copter.

The distances for yaw and pitch are then fed into their respective separate PID controller. Both PID controllers where manually tuned by visual observation and by recording the yaw and pitch responses to later plot these for finer adjustment. The results from the PID controllers are the speed in m/s for the pitch axis and the difference in heading between the current heading and desired heading in degree. These values are both send to the quad copter to finalize the control loop.

Because of optimizations in the entire loop, the system is capable of running at an average of 24hz. Trying to increase this frequency will give negligible better control since the physical response time of the quad copter is not sufficient to keep up with faster incoming commands.

The height or altitude of the quad copter during tracking is set to a static height of 3 meter. The altitude is kept static using a PID controller located in the flight computer. The main control loop described above is separate from the control loop for altitude control.

6 Performance Validation

Performance Validation has been done by testing the system performance in multiple real world situations. During testing the movement of the quad copter is observed while the quad copter is following a person which is moving in a random pattern. Below are the results.

The system can detect and follow a person through a wide range of movements. When a person moves straight forwards from the front of the quad copter, the system can keep the correct distance when a person is walking at maximum 8km/h. Speeds below 2km/h are possible but the system will show less smooth control. When the person is walking towards the quad copter, then the system will move backwards to ensure a safe following distance.

When a person is walking to the left or right while keeping the same distance from the quad copter, then the system can keep the person in the camera view if the person is walking at a maximum of 6 km/h. Also when a person walks to the right and quickly changes directions to the left, then the system can responds quickly and adjust accordingly.

When a person is combining the two types of movement, then the system can also keep track of a person. Especially when a person moves forward at a fast walking pace and makes small changes, around 40 degree, compared to the walking direction, then the system can smoothly track the person over long distances. The system starts to struggle when the change in walking direction is over 75 degree while walking forward. When this happens the system first needs to adjust its heading to make sure the Lidar is pointed on the target. During this time the quad copter cant move forward because the Lidar data is not valid. Because the person is walking forward in the same time the distance between the person and quad copter becomes much larger than the 2.5 meter. After the rotation is complete the system aggressively moves forward to get back to the 2.5 meter follow distance. This non smooth control results sometimes in the loss of tracking.

7 Conclusion

This research describes a novel and practical applied vision and control pipeline for controlling autonomous quad copters in the task of following a person in a large empty field. While the system is usable in some applications right now it can be a great base for further work to especially improve the system response for fast movements. In order to do this improvements in the depth measuring hardware needs to be made. The single solid state Lidar is often not pointed towards the person during flight returning unsuitable invalid data. These gaps in depth data makes it difficult for the control system to respond correctly.