

PAPER • OPEN ACCESS

Face Tracking for Flying Robot Quadcopter based on Haar Cascade Classifier and PID Controller

To cite this article: A S Priambodo et al 2021 J. Phys.: Conf. Ser. 2111 012046

View the article online for updates and enhancements.

You may also like

- Correlation analysis of Land Surface Temperature (LST) measurement using DJI Mavic Enterprise Dual Thermal and Landsat 8 Satellite Imagery (case study: Surabaya City) R R Darettamarlan, H Hidayat and M R Darminto
- Rapid site assessment in a small island of the Philippines contaminated with mine tailings using ground and areal technique: The environmental quality after twenty years

D B Senoro, K L M de Jesus, C A Yanuaria et al.

 - <u>Plane Accuracy of Aerophotogrammetry</u> without Image Control Points of DJI <u>PHANTOM 4 RTK UAV</u> Chun Xiu, Zhengxu Li, Haiqin Duan et al.





This content was downloaded from IP address 149.113.110.81 on 28/01/2025 at 02:15

Face Tracking for Flying Robot Quadcopter based on Haar Cascade Classifier and PID Controller

A S Priambodo¹, F Arifin², A Nasuha³ and A Winursito⁴

^{1,2,3,4} Electronics and Informatics Engineering of Education Department, Universitas Negeri Yogyakarta, Indonesia

E-mail: ardyseto@uny.ac.id

Abstract. The fundamental aim of this research is to develop a face detection system for a quadcopter in order to follow the face object. This research has two main stages, namely the face detection stage and the position control system. The face detection algorithm used in this research is the haar cascade method which is run using the python and OpenCV programming languages. The algorithm worked well, getting around 16fps on a low spec computer without a GPU unit. The results of the face detection algorithm are proven to be able to recognize faces from the camera installed on the DJI Tello mini drone. The mini drone was chosen because it is small and light, so it is harmless, and testing can be carried out indoors. Besides, the DJI Tello can be programmed easily using the python programming language. The drone's position is then compared with the set point in the middle of the image to obtain errors so that control signals can be calculated for up/down, forward/backward, and right/left movements. From the testing results, the response speed that occurs in the right/left and up/down movements is less than 2 seconds, while for the forward/backward movement, it is less than 3 seconds.

1. Introduction

The development of science and technology that is increasingly advanced in recent years can benefit all aspects of human life. Especially the rapid developments in computers, electronics, mechanics, and robotics over the last few years convince us that more robot-human interactions are not a dream but a possibility that could happen in the future. A robot is a set of mechanical devices that can perform physical tasks with human supervision and control or a predetermined program.

Robots help humans a lot, including doing dangerous work to reduce the risk that can injure humans [1]. Robotics technology has been widely applied to production machines to produce products in large quantities. Robots have several advantages, including being fast, thorough, working full time and automatically, and producing uniform products regardless of fatigue and time [2]. However, robots sometimes require periodic maintenance at certain times so that the results are carried out optimally. Many types of robots have been interesting lately; one type of robot that is widely researched and studied is a flying robot. This type of flying robot is widely researched because it brings many application opportunities.

An unmanned Aerial Vehicle (UAV) is a vehicle that can fly without a human crew when carrying out flight missions. UAVs are grouped into two types, namely the fixed-wing aircraft type and the rotarywing multirotor type. Multicopter is a Vertical Take-Off and Landing (VTOL) vehicle because the UAV can take off and land vertically without the need for a large runway [3]. One type of VTOL UAV is a

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Published under licence by IOP Publishing Ltd 1

Quadcopter with four motors that have the characteristics of good maneuverability and high stability. Quadcopter has many advantages such as low energy consumption, efficient system, and simple operation. The quadcopter can be controlled remotely, so it can be used to replace the human role. Multicopter UAVs in mapping, aerial monitoring, and handling natural disasters have better stability than aircraft-type UAVs [4].

The computing capabilities of computers have increased, and the price-to-performance ratio of computer systems has decreased, thereby contributing to the development of broader computer applications. Humans have the sense of sight, one of the natural sensors imitated in developing computer capabilities. The camera is a sensor that can replicate how the eye works, which can be implemented in computer applications. The data captured by the camera is raw and meaningless information. Then the next step is to develop a computer application to understand the information in the image captured by the camera. The utilization of images like this is included in the field of computer vision.

One of the research topics in computer vision that is often studied in the last few decades is face detection. Face detection is used to determine the presence or absence of a face in an image, and this section is the first step in the identification process, so its existence is vital [5]. Detecting faces is very easy for humans; however, this task is very complicated for the computer because there is some complexity associated with location, viewing angle, light, and occlusion. Of the many face detection methods, the most frequently used method is the Viola-Jones method [6]. The Viola-Jones method was created by Paul Viola and Michael Jones in 2001. The face detection process is performed by classifying an image based on simple feature values through a classifier formed from training data. This method combines the concept of Haar, Integral Image, and AdaBoost features, which are then processed into a Cascade Classifier [7]. The main part of the Viola-Jones algorithm is computation and feature selection. Using Integral Image to extract Haar features has accelerated computation time compared to a per-pixel calculation. The main advantage of the Viola-Jones algorithm is that it is robust, which has a high detection for face tracking in images with a low error rate.

Several studies applying image processing to quadcopters are the development of quadcopters as traffic controllers, air pollution monitoring, and area mapping [8]. In addition, other studies have also used quadcopters to detect, localize, and recognize trees. In this study, we propose a quadcopter system that can follow faces that have proven to be effective in real-time experiments on a computer with limited specifications. Quadcopter motion control uses a proven PID control algorithm that can work well on a comprehensive system with good results.

2. Material and Methods

2.1. Materials

This research uses a commercial mini quadcopter made by DJI in collaboration with Ryze Robotics, the DJI Tello, shown in figure 1. This mini quadcopter was chosen because it offers the best solution when flying around people without giving danger to the people around them. The small size of the DJI Tello with dimensions of 98×92.5×41 mm and a weight of 80 grams makes this quadcopter also practical and easy to fly indoors. The DJI Tello is equipped with a propeller guard to protect the propeller when it hits an object and protect object. The DJI Tello is equipped with a 3-axis gyroscope and accelerometer sensor that functions to obtain the angular position and speed of the vehicle, a barometer sensor that is used for altitude from the vehicle, and a downward-facing infrared sensor to maintain a more precise flight altitude compared to a barometer.



Figure 1. Mini Quadcopter DJI Tello

The main advantage of this small drone is the extra features such as the Vision Positioning System (VPS) and an onboard camera. Combining these features with an advanced control system makes this mini quadcopter hover in a quiet position. The onboard camera contained in this mini drone has specifications of 5-megapixel photos and 720p live video streams. The maximum flight time is 13 minutes, with a maximum distance of about 100m. This drone has been equipped with a failsafe protection feature, which means that when the connection is lost, the drone will land safely. In addition, another safety feature is that when the drone hits an object hard enough, the drone motor will stop rotating. This drones have several fascinating features which make them the perfect candidate for the drone, such as affordability, relatively smaller size, programmable with python and swift, embedded camera, Intel processor for stable flight and turbulence reduction.

This drone control uses a wifi network with an application installed on an Android or iPhone. DJI released a Software Development Kit (SDK) with various read and write commands over UDP communication. Thanks to Damià Fuentes Escoté for porting the SDK into the python programming language [9]. By utilizing the djitellopy library, we can access sensor data on the DJI Tello and give orders to move according to the control signal we provide. There are two basic network streams: a dual-ended full-duplex connection for sending and receiving commands between the laptop and drone and a one-way half-duplex connection for streaming video from the drone to a laptop connection. After establishing a fairly stable connection with the live streaming function, python and OpenCV were used for facial recognition on video streaming from the drone.



Figure 2. System Proposed Overview

The complete proposed system is shown in Figure 2. A DJI Tello drone that can be communicated via a Wi-Fi network is connected to a laptop. The camera from DJI Tello captures an image transmitted to a laptop for further processing using a computer vision algorithm whose details are discussed in the method section. The computer used in this project has detailed specifications, as shown in Table 1.

Tabel 1. Computer Specification

2111(2021)012070 $401.10.1000/1/72-0370/2111/1/012070$	2111 (2021) 012046	doi:10.1088/1742-6596/2111/1/012046
--	---------------------------	-------------------------------------

CPU	Intel i5-7200U (2C 4T, Max 3.10 Ghz)
RAM	DDR4 12GB (8GB+4GB)
Storage	SSD SATA 500GB
Connectivity	WiFi 2.4Ghz
Webcam	720p
Operating System	PopOS! 20.04 LTS

2.2. Methods

Face detection is a way of localizing and extracting facial areas for facial recognition purposes. Facial recognition is a technology in image processing (computer vision) that can identify a person's identity or information from the face. This technology is still very broad in its use in other fields, including security, robotics, or the health sector.

Face detection in this research uses the Haar Cascade classifier algorithm. The Haar Cascade algorithm applies a cascade function to train images through 4 main stages: (1) determining Haar features, (2) creating integral images, (3) Adaboost training, and (4) classifying using a cascading classifier. This process is shown in figure 3. The important thing that must be considered in using this algorithm is that it requires many datasets with positive facial images and non-face negative images for the classification training process.



Figure 3. Stages of Haar Cascade Classifier

The first step in the haar cascade algorithm is to collect the haar features. Haar feature is a calculation performed on the adjacent square area at a certain location in a detection window. This calculation involves adding up the pixel intensities in each section and calculating the difference between the summed results. Some types of haar features are Edge features, line features, four-rectangle features. This feature will experience problems when executed for large images, so an integral images process is needed because the number of operations is reduced.

Integral image aims to speed up the process of calculating haar features. In general, an integral image is an image whose value for each pixel is an accumulation of the values of the top and left pixels. It should be noted that almost all haar features used have low performance when detecting objects. So we need haar features that match the object to be detected. The use of Adaboost here is the answer to selecting the optimal haar feature that can be used. The AdaBoost algorithm is an algorithm that builds strong classifiers by combining several simple (weak) classifiers. Weak classification results are created by moving the window over the input image and calculating the Haar feature for each subsection. This difference is compared with the studied threshold that separates non-objects from objects. Since this is a "weak classifier", many Haar features are required for accuracy to form a strong classifier.

The combination of Cascade of Classifier is the last stage in the haar cascade method. By combining classifications in a cascade structure or Cascade of Classifier, the speed of the detection process can be increased, namely by focusing on areas in the image that are likely to be. This classifier is done to determine where the object is looking for is in an image. The classification in this algorithm consists of three levels where each level produces a sub-image that is believed to be not an object. This process is done because it is easier to judge that the sub-image is not the object wanted to detect than whether the sub-image is the object wanted to detect.

2111 (2021) 012046 doi:10.1088/1742-6596/2111/1/012046



Figure 4. Error Calculation Methods (a) Error Calculation in X and Y Axis; (b) Error Calculation in Distance Between Drone and Face

In order to be able to do tracking, a parameter is needed that becomes a reference for the quadcopter in making movements. The method proposed in this study is to create a set point in the form of a bounding box with a certain size located in the center of the image, as shown in figure 4. The object detection bounding box results are compared with the middle bounding box so that the error rate of the quadcopter's position against the object's face is known. There are three parameters, namely x, y, and d, representing the distance on the x and y axes, and d or distance, the distance between the object and the quadcopter. This illustration is shown in Figure 4. This error is then used as a reference in calculating the PID control system to generate a control signal.



Figure 5. Flowchart of the face tracking algorithm for quadcopter

In a PID control system, a variable e represents the error obtained from the difference between the actual value of the position and the desired set point r(t). The output of the control system is denoted by

y(t). each component of each Proportional, Integral, and Derivative control is added to produce a control signal. The general form of the equation of the PID controller [10] is written as follows:

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt}$$

The algorithm used in this study is shown in Figure 5. In general, it is an iterative process of face recognition, and calculation of control signals used to control forward/backward, right/left, and up/down movements. The API has covered the process of requesting images from the camera and sending motion commands on djitellopy via Wi-Fi communication.

3. Result and Discussion

The face detection system is built using the python programming language with Open-CV and djitellopy libraries. The test is carried out in several poses and conditions. The results of this test are shown in Figure 7. It can be seen in the test results of 4 poses that the Haar Cascade algorithm can detect faces well.



Figure 6. Face Detection Testing for Some Poses

The next test is an error test that applies the method described in the method section. There are three error variables in the X, Y axes and the distance between the drone and the face. The error results are displayed on the screen to make it easier to observe, denoted by eX, eY, and eD. It can be seen in Figure 7 that there is an error on the X-axis of -24, which means to the right 24 pixels; the error on the Y-axis is 8, which means upwards of 8 pixels. Error at a distance is calculated using To calculate the distance error using a comparison of the area of the reference box with the results of the face reading. The test results show that the eD value is -86, which means the drone is too close to the face.



Figure 7. Face Detection and Error Calculation (eX, eY, and eD) Test Using Webcam and Image

After getting the error value from the actual condition of the drone, the control signal can be calculated so that the drone can follow the face. PID control system is used to deal with this problem. The test results show the optimal parameter values of Kp, Ki, and Kd in table 3. It should be noted that

each movement on the quadcopter has its control system so that there are nine parameters. The movement results in each action can be seen in Figure 8, where when a change occurs, the error enlarges, and as soon as possible, the drone will correct the error so that it goes to zero. From Figure 8, it can be seen that the corrections made by the designed control system can work well.

Tabel 2. PID Parameter		
Kp – forward / backward	1.05	
Ki – forward / backward	0.02	
Kd – forward / backward	0.1	
Kp-right / left	0.65	
Ki – right / left	0.004	
Kd – right / left	0.08	
Kp - up / down	1	
Ki – up / down	0	
Kd – up / down	0.2	



Figure 8. Error in X, Y and D Position

The last test is to test the system with various displacements from the face and then observe the drone's movement whether it can follow the face or not. For observations to be made, additional features are needed, namely the storage of video results. The process of storing video results needs a little adjustment related to the defined fps parameter. When the fps parameter is too large compared to the actual fps, the video recording will experience acceleration. The defined fps parameter must match the actual fps.

The results of several tests are shown in Figures 9 - 11, wherein Figure 9 observations are made on right-left movements, and the results are that the drone can follow the face. In Figure 10, observations are made on the up-down movement by changing the position of the face down, and it is proven that the drone can follow these changes. The last test is related to the distance between the face and the drone, as shown in Figure 11. The drone can move back and forth by adjusting the set point that has been determined.

2111 (2021) 012046 doi:10.1088/1742-6596/2111/1/012046



Figure 9. Face Tracking Movement of Quadcopter in Right/Left (a) Initial Position; (b) Face Change Position to The Right; (c) Quadcopter Follow the Face



Figure 10. Face Tracking Movement of Quadcopter in Up/Down (a) Initial Position; (b) Face Change Position to The Bottom; (c) Quadcopter Follow the Face



Figure 11. Face Tracking Movement of Quadcopter in Forward/Backward (A) Initial Position; (B) Face Change Position Closed to the Drone; (C) Quadcopter Follow the Face

4. Conclusion

Haar cascade classifier method can be used for face detection in images from webcams with low computer specifications because this method is simple and lightweight classification method. The haar cascade algorithm is designed using a dataset of facial images so that only facial images can be detected and other objects cannot be detected. Sometimes, the face is not detected due to poor lighting, or the algorithm does not get the feature is looking for at that position. The drone position is calculated based on the actual distance to the set point located in the center of the image and then becomes the basis for calculating control signals for up/down, forward/backward, and right/left movements. The PID control system that is designed to work well in real-time is proven from drone testing results that it can follow objects when faces change positions. Many object detection algorithms have better performance but require large computations, so the selection of the haar cascade method is due to the limitations of the computing system owned. Improvements that can be made are to add a recognition algorithm so that the drone can follow a predetermined face object.

5. References

[1] Galin, R.R. and Meshcheryakov, R.V., 2020. Human-robot interaction efficiency and human-

robot collaboration. In Robotics: Industry 4.0 Issues & New Intelligent Control Paradigms (pp. 55-63). Springer, Cham.

- [2] Stano, G. and Percoco, G., 2020. Additive manufacturing aimed to soft robots fabrication: A review. *Extreme Mechanics Letters*, p.101079.
- [3] Rashad, R., Goerres, J., Aarts, R., Engelen, J.B. and Stramigioli, S., 2020. Fully actuated multirotor UAVs: A literature review. *IEEE Robotics & Automation Magazine*, 27(3), pp.97-107.
- [4] Dündar, Ö., Bilici, M. and Ünler, T., 2020. Design and performance analyses of a fixed wing battery VTOL UAV. *Engineering Science and Technology, an International Journal*, 23(5), pp.1182-1193.
- [5] Liu, Z., Qi, X. and Torr, P.H., 2020. Global texture enhancement for fake face detection in the wild. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8060-8069).
- [6] Viola, P. and Jones, M.J., 2004. Robust real-time face detection. *International journal of computer vision*, 57(2), pp.137-154.
- [7] Soo, S., 2014. *Object detection using Haar-cascade Classifier*. Institute of Computer Science, University of Tartu, 2(3), pp.1-12.
- [8] Ostojić, G., Stankovski, S., Tejić, B., Đukić, N. and Tegeltija, S., 2015. Design control and application of quadcopter. *International Journal of Industrial Engineering and Management*, 6(1), pp.43-48.
- [9] D. F. Escoté, djitellopy, (2021), GitHub repository, https://github.com/damiafuentes/DJITelloPy
- [10] Priambodo, A.S., Cahyadi, A.I. and Herdjunanto, S., 2017. Perancangan Sistem Kendali PD untuk Kestabilan Terbang Melayang UAV Quadcopter. Departemen Jurnal Teknik Elektro dan Teknologi Informasi. Fakultas Teknik, Universitas Gadjah Mada.