DRONE BASED PEOPLE COUNTER AND TRACKING USING UNIQUE ID USING OPENCV PYTHON

By

HARIYARAN BHASKAR

Thesis submitted in fulfilment of the requirements for the Bachelor degree of Engineering (Honours) (Aerospace Engineering)

June 2020

ENDORSEMENT

I, Hariyaran S/O Bhaskar, hereby declare that all corrections and comments made by the supervisor and examiner have been taken consideration and rectified accordingly.

(Signature of Student)

Date: 15/07/2021

PROFESOR MADYA DR. ELMI ABU BAKAR
Timbalan Dekan
(Penyelidikan, Inovasi Dan Libatsama Industri-Komuniti)
Pusat Pengajian Kejuruteraan Aeroangkasa
Universiti Sains Malaysia
Kampus Kejuruteraan
14300 Nibong Tebal

(Signature of Supervisor)

Name: Assoc. Prof Dr. Elmi Abu Bakar
Date: 15/07/2021

(Signature of Examiner)

Name: Dr Ho Hann Woei
Date:

DECLARATION

This thesis is the result of my own investigation, except where otherwise stated and has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any other degree.

ACKNOWLEDGEMENT

I am using this opportunity to express my deepest gratitude to my thesis advisor and final year project main supervisor, Assoc. Prof Dr. Elmi Abu Bakar and assistant supervisor Dr. Nishat Akhtar. With their patience guidance throughout the year has led me to the completion of this final year project. Their vast knowledge in the field of image processing and thought-provoking suggestions has guided me towards the successful completion of this project. Both their advice and comment are valuable and useful which help in my thesis writing. They were patient in monitoring me in order to ensure that I am on the right track along the way of completion of this project.

Thank you to the examiner, Dr. Ho Hann Woei, for the involvement in my final year project and thesis assessment.

Moreover, I would like to show my gratitude to Mr. Mohamad Zihad bin Mahmud and other lab technicians for their help regarding the final year project. Without him, I would not be able to complete my final year project as he helps me regarding the type of drone suitable for my project and some methodology related informations. I would like to thank the final year project lecturers and other lecturers who have given the talk and guideline for the final year project and thesis writing.

Finally, yet importantly, I would like to thank both my beloved parents who have always supported me throughout the completion of the project. Their love, steady encouragement, and support have helped me keep a positive attitude in completing this project. I would like to dedicate this thesis to my parents and my family.

.

DRONE BASED PEOPLE COUNTER AND TRACKING USING UNIQUE ID USING OPENCV PYTHON

ABSTRACT

The purpose of this thesis is to implement an algorithm for drone-based square boundaries using Python and OpenCV. This thesis provided a detailed method in implementing the algorithm for face and person recognition by using haar cascade classifier feature detection and contour approximation. The algorithm is developed so that it is capable to detect various objects by applying a bounding box on the frame image. This project uses Python as its programming language and OpenCV as an open-source library for programming. The image is taken from the DJI Tello drone. The acquired image is then converted to grayscale. Gaussian filter is used for image smoothing and noise removal. The canny edge detector is used for the recognition of an object's edges. Upon implementation, the contours are performed for further analysis and recognition of the person shape. The crosshairs are drawn on the frame for aiming purposes. The testing is done on different images' characteristics to verify the required features and the problems come out. Also, the experiments are done by using the various value of epsilon to estimate the accuracy of the detection. The angle of projection and the distance between the drone and the object are included in observation. The outcome of this project reflects the object detection technique which will potentially improve the machine vision and will subsequently contribute to the development of the image processing in artificial intelligence.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

UAV	:	Unmanned Aerial Vehicle
HDR	:	High Dynamic Range
RGB	:	Red-Green-Blue
DJI	:	Da Jiang Innovations
USM	:	University Sains Malaysia
2D	:	2-Dimensional
GUI	:	Graphical User Interface
OpenCV	:	Open Source Computer Vision
FPS	:	Frame Rate Per Second
JPG/JPEG	:	Joint Photographic Experts Group
Ms	:	Milliseconds
Mbps	:	Megabits per second
m	:	Metre
CPU	:	Central Processing Unit

LIST OF SYMBOLS

f_i	:	Number of pixels having grey level '*i*'
N	:	Number of pixels in the images
δ	:	Difference between pixels
$l(x)$	:	Pixel value of point x
i	:	Index of the cluster
t_p	:	Threshold point
G_σ	:	Gaussain factor
a	:	Amplitude
Q	:	Quality factor
t	:	Image without noise
s	:	Filtered image
$\sigma_{s,t}$	:	Covariance between 2 images
σ_s^2	:	Variance of filtered image
σ_t^2	:	Variance of source image without noise
$G_\sigma(x, y)$	:	Gaussian Kernel function

CHAPTER 1

INTRODUCTION

1.1 Research background

An unmanned aerial vehicle (UAV), called a drone has been developed for many research areas. It can be controlled by a user, or a program designed for an auto-flying mission such as the military, disaster rescue facility, agriculture, and transportation. Drones entered the commercial space as exciting toys, slowly transforming into a multi-billion-dollar industry with many commercial applications ranging from photography to military surveillance. Recently, the tracking of people and objects detection are seen as the vital factors for surveillance in the smart-city and can be used to increase safety for the digital society. This can be implemented to support the closed-circuit television (CCTV) systems. Blind spots in CCTV coverage can compromise coverage and lead to failure. Today with crimes rising there is a need to find other techniques providing better surveillance. Artificial Intelligence, with its recent advancements and disruptive technology, has been a game changer for the drone industry. AI has opened doors in this domain to avenues that were unimaginable just a few years back. AI can replace humans at various levels of commercial drone use—they can autonomously control the drone flight, analyse sensor data in real time or even examine the data post-flight to generate insights. At any of these levels, it is often required to identify and track a person-of-interest around the drone through the data captured by its sensors, making people detection and tracking fundamentally important to impart artificial intelligence to a drone. The general idea of this project is to include this image processing algorithms and multiple proposed methodology to detect, count, and track a number of people in a designated area using a drone.

1.2 Problem Statement

The machine has a high capability of detection and recognition of objects as well as human but could not reach the capability to understand the natural environment. Regardless of the complexity of the world, irregularity in the properties of surrounding objects, it's very difficult for a computer to understand nature like humans. However, the reliable learning process for computer vision using an object's properties such as form, size, texture, and other related properties that improve with the correct results of detection and classification. Mostly, when several objects belong to a single frame, existing technology systems may not be able to detect the objects properly (Tiwari & Singhai, 2017). So, we suggested implementing drone-based image processing relevant to computer technology on this project.

1.3 Research objective

A research study designed to implement and assess the performance of the proposed technique involves the following research objectives:

I. To develop and implement an algorithm for detecting and counting people using Background Subtraction and Haar Based classifier.

II. To develop target tracking algorithm and determine the motion of people by incorporating unique ID.

III. To implement the proposed target detection algorithm for drone-based real-time people tracking and estimate the accuracy.

1.4 Research scope

This work is to recognize and detect a person within a square boundary by using the contour properties. We propose an algorithm that can detect approximately several people in a frame while outcoming the people counter.

Here, the footage of pedestrians walking or people in a crowd is acquired from the drone will be applied with the Haar-Cascade algorithm to recognize various the frame acquired by the drone. Next, pre-processing the images to resize and remove the noise from corrupted. Furthermore, convert the image to grayscale, blur it with Gaussian filter, and detect the human features including full body detection by applying the Haar-Cascade detection algorithm. Then, apply the contours to compute the bounding rectangle and assign a unique number of ID to each person detected.

Various image processing algorithms for detecting person were proposed and implemented in a parallel manner over multiple compute cores. This work more focuses on a person detection and tracking based on a square boundary using contours properties in which the live feed will be acquired from the drone simultaneously.

1.5 Thesis outline

This thesis contains five major chapters that include an introduction, literature review, methodology, results and discussions, and lastly conclusions and recommendations.

Chapter 1 introduces the main idea of the project which is briefly explained and some information regarding the research background, problem statement, research objective, research scope and limitations, and thesis outline.

Chapter 2 describes the literature review. The proposed processing of images based on person detection is briefly clarified and where it typically relates to the application. Also, what are the person recognition features. The various techniques and methods for detecting a person will be explained in this chapter. Recent studies were also discussed on object detection. It provides a decent understanding of how to develop and implements a new approach from the current method.

In Chapter 3, the proposed method and technique used for person detection and tracking based unique ID are presented. Briefly explained about installing the software, image acquisition, image pre-processing, image filtering, Haar-cascade algorithm, and assigning unique ID for tracking for this project.

Chapter 4 discussed the results of the experiment obtained. The images are composed of different elements. The frames and multiple elements in the algorithm are manipulated for evaluation of accuracy. The detection of accuracy is related to the projection angles and the range between the drone and the target. The detection of accuracy is also calculated in different brightness and light conditions. The result of the analysis is shown in the figure, and the collected data is represented in the table.

Chapter 5 concludes all the findings and recommendations for future works.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

Nowadays machines have high capability of detection and recognition of objects as well as human. However due to complexity of the world, irregularity in the properties of surrounding objects, it is very difficult for a computer to interpret and process nature like humans. However, the reliable learning process for computer vision using an object's properties such as form, size, texture, and other related properties that improve with the correct results of detection and classification. Mostly, when several objects belong to a single frame, existing technology systems may not be able to detect the objects properly. Review of literature is done to know the state of the art for methods of object detection so that improvements can be made to current methods.

2.2 Face Detection

For face detection, the Viola-Jones face detector (Viola & Jones, 2004) is well-known method that first proposed for object and then for pedestrian detection (Viola et al., 2003) and nowadays, this technique, called Haar cascade classifier, has become a standard technique for face detection. The Viola-Jones face detector computes feature vector based on the Haar feature. It calculates from the rectangle detector or sub-window. The detector scans through the image with the size of 24→ 24 pixel resolution. The size of the detector will increase by 1.2 times. Then, the set of the feature vector is given to the AdaBoost classifier which is the weak classifier. This approach can process in real-time and get high precision.

Dalal and Triggs (Dalal & Triggs, 2005) proposed a feature extraction method called histograms of oriented gradients (HOG) and first designed for human detection. This method can use the shape for calculation of the intensity of gradients. Each sub-sampling of images will change the edge images as a cell, in which the orientation of the edge image will be calculated for each cell. Then the orientation can lead to create the histogram and send to learn with the support vector machine (SVM) using the linear kernel. The SVM is used as a binary classifier to be trained on positive and negative images, accordingly it is used to classify each of the sub-window images. Moreover, many research uses the HOG-SVM method (Bristow & Lucey, 2014; Dadi & Mohan Pillutla, 2016) for face detection as well.

In every discussed face detection method, filtering the image as pre-processing process is necessary. It can be done for noise removal, blur removal, edge detection etc. Linear and non-linear filters are the algorithms which are used for filtering. Right filter should be selected for any specific purpose. If the image or input given has less amount of noise but the magnitude is high, then non-linear filters are used whereas linear low-pass filter is sufficient when the input given contains noise in large amount but the magnitude of noise is low. Linear filters are the most frequently used filters as it is simplest and fastest. Unlike non-linear filters, the linear filtering is done through applying the algorithm on the neighbour pixels of the input pixels in the image. The neighbourhood pixels are identified through their locations which are relative to the input pixel. Works related to multiple face detection method and multiple filtering techniques are discussed in upcoming section.

2.3 Related work

2.3.1 Image Filtering Technique

In (King, 2015), max-margin object detection (MMOD) is proposed for finding the object. In this approach, the non-overlapping sub-window image slide through image and uses the window scoring function to find the fewest possible detection mistakes. As for the feature extraction method, the spatial pyramid bag-of visual-words model combined with the HOG descriptor method are computed from each pixel location. The MMOD method is performed well on the face detection dataset and benchmark (FDDB).

In addition, in (Binangkit & Widyantoro, 2017), a color image is converted into the HSV color space and an extracted area of the traffic light color such as red, yellow and green colors as the region of interest (ROI). The pixels of traffic light colors are represented as the set of feature vectors. Then the feature vector is given to the machine learning technique to create a model that uses it to classify if the ROI is that of the traffic light or not

It is different to Dalal and Triggs (Dalal & Triggs, 2005) because all area of the image will be used to analyze with HOG and then applied to the linear function which can develop to finding the objects. Their cases study has three samples, consist of TU Darmstadt cows, INRIA pedestrains and face detection dataset and benchmark (FDDB).

Binagkit and Widyantoro (Binangkit & Widyantoro, 2017) propose how to detect to the color of a traffic light, which the image is changed to being HSV color space. After that it is set the period of color such as red, yellow and green for extraction region of interests (ROI). The ROI is selected to learning with SVM by using RBF kernel.

Background subtraction (Saravanakumar et al., 2010) is popularly used technique for segmenting out objects of interest in a frame. For this technique involves subtracting an