IMPROVED PEDESTRIAN DETECTION AND DISTANCE DETECTION WITH STEREO CAMERA USING COMPUTER VISION

THE WEI XIN

A project report submitted in partial fulfilment of the
requirements for the award of the
Master of Engineering (Computer and Microelectronics System)

School of Electrical Engineering
Faculty of Engineering
Universiti Teknologi Malaysia

FEBRUARY 2022

DEDICATION

This thesis is dedicated to my father, who taught me that the best kind of knowledge to have is that which is learned for its own sake. It is also dedicated to my mother, who taught me that even the largest task can be accomplished if it is done one step at a time.

ACKNOWLEDGEMENT

Foremost, I would like to express my sincere appreciation to my supervisor, Dr. Usman Ullah Sheikh for his motivation, immense knowledge and providing me a very good basic knowledge about stereo vision and guiding me to complete the experiment. His guidance had helped me to complete my project in time and writing this thesis. Other than my supervisor, I would like to thank Dr. Ab Al-Hadi Ab Rahman for his informative advice given during the viva session and questions that let me understand the part I had neglected inadvertently.

Besides my supervisor, I would like to thank my classmate: Khor Chao Liang on giving companion and discussion along the whole period during the project. I would also like to thank my friends that gave me useful advice on the project and spent time being my model for detection. Lastly, I want to thank my parents for their support and encouragement.

ABSTRACT

Advanced Driver Assistant System (ADAS) is an advanced technology that provides assistance in driving and parking functions. The main purpose of ADAS is to increase road safety and enhance comfort of the driver. ADAS system generally is implemented along with sensors and cameras, to detect obstacles or pedestrian and respond accordingly. ADAS had been proven to reduce road fatalities by minimizing human error. One of the advanced systems used for ADAS nowadays is stereo vision cameras, which is set up using two identical cameras to produce a 3D image. The main highlight of a stereo vision is that a 3D image has depth which gives higher accuracy compared to a 2D. Nowadays, cases of road accidents have been increasing every year. According to World Health Organization (WHO), approximately 1.35 million people die each year due to road accidents. Among these number of cases, more than half of the deaths were among pedestrians, cyclists, and motorcyclists. Meanwhile currently most of the detection was done on sunny day, as it is much more challenging to detect pedestrian during complicated environment such as rainy day. The flow of the project is firstly to set up a pair of stereo cameras on the dashboard of a stationary car to detect pedestrians. Using OpenCV with Python language, stereo matching is completed to obtain a disparity map and depth map. Once the depth map is obtained, distance can be calculated. To increase the accuracy on detecting pedestrians, Haar Cascade classification was implemented. After that, a major image training, testing and validation is performed. In this thesis, a detection of pedestrian in rainy day environment is implemented using stereo vision cameras. This work had proved that depth extraction using the custom stereo camera can work in a rainy day with an error of 0.2m accuracy off and Haar Cascade classification was successfully implemented to detect pedestrians.

ABSTRAK

Sistem Pembantu Pemandu Lanjutan (ADAS) ialah teknologi canggih yang menyediakan bantuan dalam fungsi pemanduan dan tempat letak kereta. Tujuan utama ADAS adalah untuk meningkatkan keselamatan jalan raya dan meningkatkan keselesaan pemandu. Sistem ADAS secara amnya dilaksanakan bersama dengan penderia dan kamera, untuk mengesan halangan atau pejalan kaki dan bertindak balas dengan sewajarnya. ADAS telah terbukti dapat mengurangkan kematian jalan raya dengan meminimumkan kesilapan manusia. Salah satu sistem canggih yang digunakan untuk ADAS pada masa kini ialah kamera penglihatan stereo, yang disediakan menggunakan dua kamera yang sama untuk menghasilkan imej 3D. Sorotan utama penglihatan stereo ialah imej 3D mempunyai kedalaman yang memberikan ketepatan yang lebih tinggi berbanding dengan 2D. Kini, kes kemalangan jalan raya semakin meningkat setiap tahun. Menurut Pertubuhan Kesihatan Sedunia (WHO), kira-kira 1.35 juta orang mati setiap tahun akibat kemalangan jalan raya. Di antara bilangan kes ini, lebih separuh daripada kematian adalah dalam kalangan pejalan kaki, penunggang basikal dan penunggang motosikal. Pada masa ini kebanyakan pengesanan dilakukan pada hari yang cerah, kerana ia adalah lebih mencabar untuk mengesan pejalan kaki semasa persekitaran yang rumit seperti hari hujan. Aliran projek itu mula-mula menyediakan sepasang kamera stereo pada papan pemuka kereta pegun untuk mengesan pejalan kaki. Menggunakan OpenCV dengan bahasa Python, pemadanan stereo dilengkapkan untuk mendapatkan peta perbezaan dan peta kedalaman. Setelah peta kedalaman diperoleh, jarak boleh dikira. Untuk meningkatkan ketepatan mengesan pejalan kaki, pengelasan Haar Cascade telah dilaksanakan. Selepas itu, latihan imej utama, ujian dan pengesahan dilakukan. Dalam tesis ini, pengesanan pejalan kaki dalam persekitaran hari hujan telah dilaksanakan menggunakan kamera penglihatan stereo. Kajian ini telah membuktikan bahawa pengekstrakan ralat menggunakan kamera stereo tersuai boleh berfungsi pada hari hujan dengan ketepatan 0.2m sahaja dan pengelasan Haar Cascade berjaya dilaksanakan untuk mengesan pejalan kaki.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADAS	-	Advanced Driver Assistance System
CCF	-	Convolution Chanel Features
CNN	-	Convolution Neural Network
HOG	-	Histogram of Oriented Gradient
LBP	-	Local Binary Pattern
OCSVM	-	One-Class Stereo Vision Vector Matrix
R-CNN	-	Region Based Convolutional Neural Networks
ROIs	-	Region of Interests
SSD	-	Single Shot Detection
stereoBM	-	Stereo Block Matching
stereoSGBM	-	Stereo Semi-Global Block Matching
SVDD	-	Support vector data description

LIST OF SYMBOLS

B	-	Base line
f	-	Focal length
fx	-	Focal point x
fy	-	Focal point y
cx	-	Offset of principle point of x-axis

CHAPTER 1

INTRODUCTION

1.1 Introduction

Advanced driver-assistance system (ADAS) is an advanced system that provides convenience to autonomous users in term of safety and ease. The essential safety-critical applications of ADAS are blind spot detection, lane departure warning or correction, traffic sign recognition, automatic emergency brake and pedestrian detection or avoidance. The main purpose of the application mentioned is to ensure the traffic safety [1]. In the past, the automotive safety uses a passive safety which is able to minimize injury when accidents happen. ADAS had proved that it has successfully reduce the number of accidents and fatalities. Basically, there are six autonomous vehicle levels for driving, which are from level 0 to level 5. Level 0 to Level 2 requires human to monitor the driving environment. While Level 3 to Level 5 of the automated system will monitor the driving environment. Meanwhile shifting towards fully autonomous cars, the vehicle is capable of sensing the environment and operate without human involvement [2].

ADAS is implemented with various types of sensors to achieve a 360-degree vision for self-driving cars. Figure 1.1 shows an overview of features and functions of ADAS, and respective sensors used for each feature. From Figure 1.1, it can be seen that there are three main sensors used in ADAS, which are LIDAR, camera, and ultrasound sensor. However, there are more sensors being introduced recently to improve the system in terms of accuracy and reducing cost, such as thermal sensor, night vision camera and birds eye camera. Meanwhile, object detection feature is an important feature in ADAS for pedestrian detection. Over the years, the training of pedestrian detection for ADAS was done using Convolutional Neural Network (CNN) engine. The accuracy and speed of detection are the main factors that being emphasized in those implementations [3].

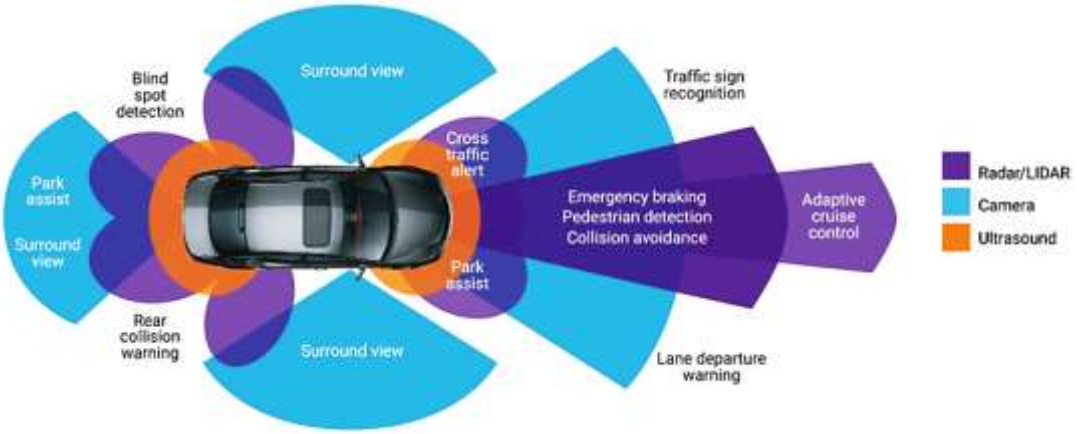

Figure 1.1 Sensors of ADAS [3]

In this thesis, pedestrian detection is the main focus. Stereo camera is used due to lower cost compared to other types of sensors. Using stereo camera can be calculate depth. The stereo camera used in this project is mainly setup by using two cameras on a car dashboard, and this setup is used to obtain the dataset for pedestrian detection during rainy day. By undergoing Haar feature-based cascade classifiers training and classification using machine learning (OpenCV with Python), the system is able to detect pedestrians and calculate the distance between vehicle and pedestrian. More accidents happen during rainy weather. Hence this thesis will focus on pedestrians detection on rainy day

1.2 Problem Statement

According to Ministry of Transport Malaysia, the number of road accidents had increased within 10 years. In Figure 1.2, can see that the statistics recorded for Malaysia road accidents had increased from 414,421 cases to 567,516 cases from year 2010 to year 2019. Which means that in year 2019, there is an average 1555 cases of accidents per day in Malaysia. Although the road fatalities rate had reduced until the lowest rate of 6,167 in year 2019 but there is still a huge number of accidents. The graph of road fatalities is shows in Figure 1.3 [4]. Meanwhile, rain increases the

possibilities of accident. According to a data collected from National Highway Traffic Safety Administration, there are more accident happen on a rainy day. Even so, there are many research papers that had been done the experiment in a good and sunny weather, where an adverse weather had not been considered.

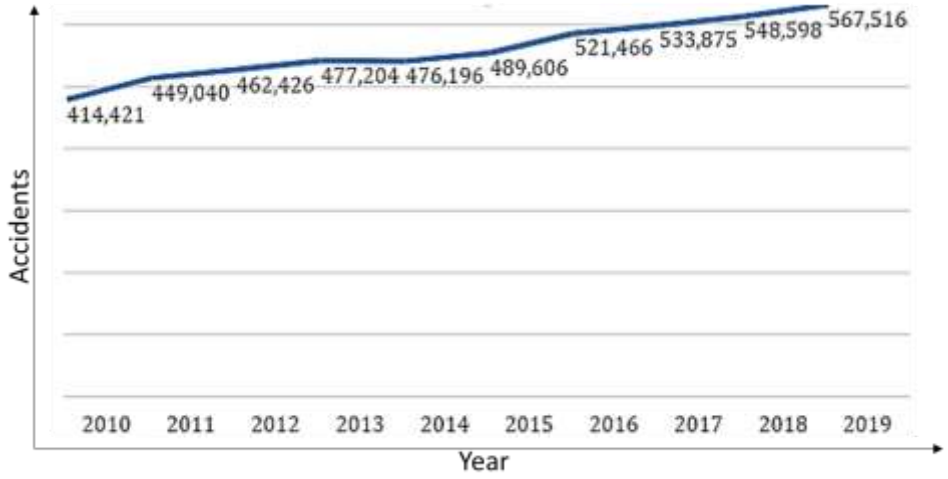

Figure 1.2 Malaysia road accident rate from year 2010 to year 2019 [4]

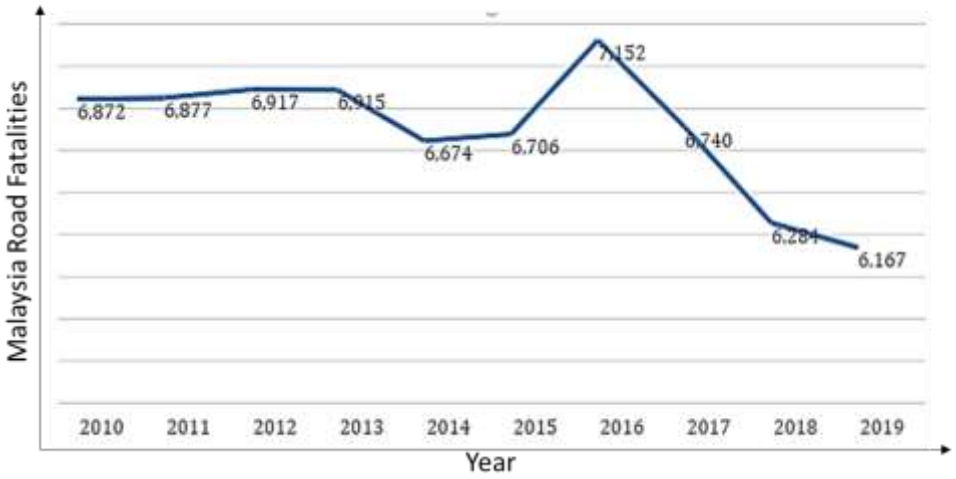

Figure 1.3 Malaysia road fatalities rate from year 2010 to year 2019 [4]

In previous works the detection was only done on detecting one nearest pedestrian at one time, which is not sufficient for safety precaution. In addition, the detection was done on the disparity map instead of a major target image. The accuracy can be increased if there are both detections happening in the disparity map and on

alpha major image which maybe the left image. Below is the summary list of problem statements:

(a) Most research using stereo camera is used outside the vehicle.

(b) Detection of pedestrian is only done in sunny/good weather.

(c) No label for detection.

1.3 Research Objectives

The objectives of the research are:

(a) To propose an improved pedestrian detection method using stereo vision camera in rainy weather.

(b) To implement the pedestrian and distance calculation together in one framework.

1.4 Research Scope

The scope of the research is:

(a) Detect and measure distance for pedestrians.

(b) Detection will be done by using a stereo camera on the vehicle's dashboard.

(c) Camera calibration will be performed using OpenCV with Python language.

1.5 Report Organization

This report consists of five chapters. Chapter 1 which is titled as “Introduction” describes the overview of the project. In this chapter, the subchapters include problem statement, research objectives, and research scope. These should help on providing basic understanding about this research field, target of study, and the methods and techniques that had been applied to this project.

In Chapter 2, previous studies that are related to ADAS, pedestrian detection, stereo vision sensor, machine learning, and depth detection are presented. Firstly, several studies about pedestrian detection are presented and followed by some research related to stereo vision sensor.

In Chapter 3, the method of calibration and implementation process of handling a stereo camera with OpenCV and Haar cascaded classifier to perform pedestrian detection are presented. This chapter includes the steps on the setup and calibration of the stereo camera. Followed by the steps to train the classifier.

Chapter 4 which is entitled as “Result and Discussion” will clearly describe and provide the result obtain from the project. Along with the screenshot of disparity maps generated and screenshots of pedestrians detected on an image. Discussion about the results and justification are given.

Lastly, Chapter 5 will be the conclusion of the whole work and future work will also be explained and described in this chapter.

REFERENCES

[1] Intellias. (2021). How to Build ADAS Technology for Autonomous Driving OEMs and Tier 1 vendors should build ADAS technology to beat the competition and meet safety requirements. Retrieved May 5, 2021, from https://www.intellias.com/how-to-build-adas-technology-for-autonomous-driving/

[2] Perforce Software. (2020). What are the 6 Autonomous Vehicle Levels for Driving? https://www.perforce.com/blog/qac/6-levels-of-autonomous-driving

[3] Synopsys. (2021). What is ADAS ? https://www.synopsys.com/automotive/what-is-adas.html

[4] Ministry Of Transport Malaysia. (2020). Road Accidents and Fatalities in Malaysia. MINISTRY OF TRANSPORT MALAYSIA. https://www.gulf-times.com/story/656066/Road-accidents-and-fatalities-drop-in-Qatar#:~:text=A total of 44 cases,down by 6.7%25 in 2019.

[5] ZMP INC. (2021). ADAS (Advanced Driver-Assistance Systems , Basic information on development of advanced driving support system). ZMP INC. https://www.zmp.co.jp/en/knowledge/adas_dev

[6] e-con Systems. (2018). What is a stereo vision camera? Retrieved May 5, 2021, from https://www.e-consystems.com/blog/camera/what-is-a-stereo-vision-camera/

[7] Perforce Software. (2020). What are the 6 Autonomous Vehicle Levels for Driving? https://www.perforce.com/blog/qac/6-levels-of-autonomous-driving

[8] Stackable. (2014). Building and calibrating a stereo camera with OpenCV (<50€). https://erget.wordpress.com/2014/02/01/calibrating-a-stereo-camera-with-opencv/

[9] Eser, A. Y. (2020). The Depth I : Stereo Calibration and Rectification. https://python.plainenglish.io/the-depth-i-stereo-calibration-and-rectification-24da7b0fb1e0

[10] Eser, A. Y. (2021). OpenCV Camera Calibration. https://aliyasineser.medium.com/opencv-camera-calibration-e9a48bdd1844

[11] Opencv, L. (2020). Camera Calibration using OpenCV. Learn OpenCV. https://learnopencv.com/camera-calibration-using-opencv/

[12] A. shakeri, B. Moshiri and H. G. Garakani, "Pedestrian Detection Using Image Fusion and Stereo Vision in Autonomous Vehicles," 2018 9th International Symposium on Telecommunications (IST), 2018, pp. 592-596, doi: 10.1109/ISTEL.2018.8661069.

[13] Z. Chen and X. Huang, "Pedestrian Detection for Autonomous Vehicle Using Multi-Spectral Cameras," in IEEE Transactions on Intelligent Vehicles, vol. 4, no. 2, pp. 211-219, June 2019, doi: 10.1109/TIV.2019.2904389.

[14] J. Li, H. Liu, Z. Sun and R. Huang, "Real Time Obstacle Estimation Based on Dense Stereo Vision for Robotic Lawn Mowers*," 2019 IEEE International Conference on Robotics and Biomimetics (ROBIO), 2019, pp. 1186-1191, doi: 10.1109/ROBIO49542.2019.8961557.

[15] Nguyen, U., Rottensteiner, F., & Heipke, C. (2019). Confidence-Aware Pedestrian Tracking Using a Stereo Camera. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 4(2/W5), 53–60. https://doi.org/10.5194/isprs-annals-IV-2-W5-53-2019

[16] H. Xue, D. Q. Huynh and M. Reynolds, "Pedestrian Tracking and Stereo Matching of Tracklets for Autonomous Vehicles," 2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring), 2019, pp. 1-5, doi: 10.1109/VTCSpring.2019.8746329.

[17] Dairi, A., Harrou, F., Senouci, M., & Sun, Y. (2018). Unsupervised obstacle detection in driving environments using deep-learning-based stereovision. Robotics and Autonomous Systems, 100, 287–301. https://doi.org/10.1016/j.robot.2017.11.014

[18] Emani, S., Soman, K. P., Sajith Variyar, V. V., & Adarsh, S. (2019). Obstacle detection and distance estimation for autonomous electric vehicle using stereo vision and DNN. Advances in Intelligent Systems and Computing, 898(January), 639–648. https://doi.org/10.1007/978-981-13-3393-4_65

[19] C. Damian, D. Grigorescu, I. Ghindă and M. Robu, "Using Mono and Stereo Camera System for Static and Moving Objects Detection," 2019 International Conference on Electromechanical and Energy Systems (SIELMEN), 2019, pp. 1-5, doi: 10.1109/SIELMEN.2019.8905820.

[20] Paolo Galeone. (2018). *Camera calibration guidelines Mar*. https://pgaleone.eu/computer-vision/2018/03/04/camera-calibration-guidelines/#:~:text=A value %3C 0.5 is good,given input images on average.

[21] cv :: StereoBM Class Reference. (2022). Open Source Computer Vision. https://docs.opencv.org/3.4/d9/dba/classcv_1_1StereoBM.html

[22] Bogdan, O., Eckstein, V., Rameau, F., & Bazin, J. C. (2018). A deep learning approach for automatic intrinsic calibration of wide field-of-view cameras. *Proceedings - CVMP 2018: 15th ACM SIGGRAPH European Conference on Visual Media Production, December*. https://doi.org/10.1145/3278471.3278479