UNIVERSITI TEKNIKAL MALAYSIA MELAKA

DEPTH MAP RECONSTRUCTION FROM STEREO IMAGE USING SUM OF ABSOLUTE DIFFERENCES

This report is submitted in accordance with the requirement of the Universiti Teknikal Malaysia Melaka (UTeM) for the Bachelor of Electronic Engineering Technology (Telecomunication) with Honours.

by

EFFY SHAZRIN BIN SABTUHE

B071710271

950930155041

FACULTY OF ELECTRICAL AND ELECTRONIC ENGINEERING TECHNOLOGY

2021

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS LAPORAN PROJEK SARJANA MUDA

Tajuk: A STUDY OF DEPTH MAP RECONSTRUCTION FROM STEREO IMAGES USING SUM OF ABSOLUTE DIFFERENCES

Sesi Pengajian: 2020/2021

Saya **EFFY SHAZRIN BIN SABTUHE** mengaku membenarkan Laporan PSM ini disimpan di Perpustakaan Universiti Teknikal Malaysia Melaka (UTeM) dengan syarat-syarat kegunaan seperti berikut:

1. Laporan PSM adalah hak milik Universiti Teknikal Malaysia Melaka dan penulis.
2. Perpustakaan Universiti Teknikal Malaysia Melaka dibenarkan membuat salinan untuk tujuan pengajian sahaja dengan izin penulis.
3. Perpustakaan dibenarkan membuat salinan laporan PSM ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. **Sila tandakan (X)

☒	SULIT*	Mengandungi maklumat yang berdarjah keselamatan kepentingan Malaysia sebagaimana yang termaktubalam A RAHSIA RASMI 1972.
☐	TERHAD*	Mengandungi maklumat TERHAD yang telah ditentukan organisasi/badan di mana penyelidikan dijalankan.
☒	TIDAK TERHAD	

Yang benar,

.....

EFFY SHAZRIN BIN SABTUHE

Alamat Tetap:

BKM 0769 KG BUKIT KALAM

87008, W. PERSEKUTUAN

LABUAN

Tarikh: 29/03/2021

Disahkan oleh penyelia:

DR. EDSTAM ASSENDI BIN HAMZAH
Pensyarah Kanan
Jabatan Teknologi Kejuruteraan Elektronik & Komputer
Fakulti Teknologi Kejuruteraan Elektrik & Elektronik
Universiti Teknikal Malaysia Melaka

Tarikh: 29/03/2021

*Jika Laporan PSM ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali sebab dan tempoh laporan PSM ini perlu dikelaskan sebagai SULIT atau TERHAD.

DECLARATION

I hereby, declared this report entitled Depth Map Reconstruction From Stereo Images Using Sum Of Absolute Differences is the results of my own research except as cited in references.

Signature: ……

Author : EFFY SHAZRIN BIN SABTUHE

Date: 29/03/2021

APPROVAL

This report is submitted to the Faculty of Electrical and Electronic Engineering Technology of Universiti Teknikal Malaysia Melaka (UTeM) as a partial fulfilment of the requirements for the degree of Bachelor of Electronics Engineering Technology (Telecommunications) with Honours. The member of the supervisory is as follow:

Signature: ....

Supervisor:

DR. ROSTAM AFFENDI BIN HAMZAH
Pensyarah Kanan
Jabatan Teknologi Kejuruteraan Elektronik & Komputer
Fakulti Teknologi Kejuruteraan Elektrik & Elektronik
Universiti Teknikal Malaysia Melaka

ABSTRACT

Projek ini menyediakan kaedah untuk menyelesaikan masalah korespondensi ketika memadankan gambar stereo menggunakan algoritma Sum of Absolute Differences (SAD). Perisian MATLAB menyediakan alat tersebut. Pemetaan perbezaan dihasilkan melalui algoritma pemadanan blok Sum of Absolute Differences (SAD). Konvensional SAD biasanya diimbas di seluruh gambar stereo untuk mengetahui perbezaan garis antara gambar yang ditangkap di kiri dan kanan, dan kemudian memperoleh peta perbezaan yang setara, yang boleh mengakibatkan masa berlalu yang tinggi. Terdapat empat langkah asas dalam kaedah penglihatan stereo untuk mencitrakan pembinaan semula. Biasanya terdapat empat fasa ini. Penyelewengan gambar yang diambil dari lensa kamera diekstrak terlebih dahulu pada langkah penyimpangan. Langkah seterusnya adalah menyesuaikan jarak dan sudut ketinggian antara kedua gambar kamera untuk menentukan panjang fokus dan sumbu epipolar, seperti yang diduga di luar. Perbandingan antara gambar kiri dan kanan akan dihitung dalam tahap korespondensi dan digunakan untuk mengukur peta perbezaan. Kaedah ini sering dikenali sebagai saling beroperasi. Projek ini dapat mengembangkan algoritma pemadanan stereo menggunakan jumlah perbezaan mutlak untuk melakukan penyusunan peta mendalam dari algoritma yang dicadangkan dan algoritma stereo sederhana mengira hasil yang setanding dengan canggih terkini di penanda aras Middlebury. Kaedah projek ini menggunakan antara muka pengguna grafik yang dikembangkan dari persekitaran MATLAB yang menghitung pemetaan perbezaan dengan menggunakan jumlah perbezaan mutlak. Projek ini telah memenuhi semua objektif saya dan hasil yang saya dapat untuk mencapai penanda aras yang ditetapkan. Dalam pelbagai domain, seperti penjejakan, navigasi robot, dll., Algoritma ini dapat digunakan dengan lebih baik. Tidak ada petunjuk kedalaman yang lebih baik atau sangat diperlukan daripada petunjuk kedalaman yang lain. Petunjuk mempunyai kelebihan dan kekurangan tersendiri.

ABSTRACT

This project provides a method for solving the issue of correspondence when matching the stereo image using the algorithm Sum of Absolute Differences (SAD). MATLAB software provides the tool. The mapping of disparities is produced through the block matching algorithm Sum of Absolute Differences (SAD). The SAD conventional is usually scanned in entire stereo images to figure out the discrepancy in lines between the captured images on the left and right, and then obtains the equivalent map of disparity, which can result in a high elapse time. There are four basic steps in a stereo vision method for imaging the reconstruction. There are typically four phases of this The distortion of the captured images from the camera lens is extracted first in the un distortion step. The next step is to adjust the distance and the elevation angle between the two camera images to determine the focal length and the epipolar axis, as presumed beyond. The comparisons inter the left and the right image will be calculated in the correspondence stage and used to measure the map of disparities. This method is often known as being interoperable. This project was able to develop stereo matching algorithm using sum of absolute differences to do the depth map reconstruction from the proposed algorithm and a simple stereo algorithm compute results comparable to current state-of-the-art on Middlebury benchmark. The method of this project is using a graphical user interface that developed from MATLAB environment that compute the disparity mapping by using sum of absolute differences. This project has answered all my objectives and the results I got to achieve the set benchmark. In various domains, such as tracking, robot navigation, etc., these algorithms can be better used. There is no better or indispensable depth cue than another depth cue. A cue has its own perks and disadvantages.

DEDICATION

I would like to dedicated and special thanks to

My beloved father and mother,

To my beloved family, my respected lecturer and fellow friends

And for the rest, might Allah have blessed you

Thanks for all the guided and support

ACKNOWLEDGEMENTS

First and foremost, I am very grateful to the almighty ALLAH S.W.T for letting me to finish my Final Year Project 1 and give me strength that I need in order to fulfill my duty as an Electronic student.

Special thanks to my supervisor Ts. Dr. Rostam Affendi Bin Hamzah for the guidelines and teach me during completing my final year project for duration 1 year. May Allah always blessed you and I never forget your advice and guidance during I make a mistake.

Thanks also to all of the kind lecturers in Bachelor of Electronic Engineering Technology (Telecommunication) with Honours section for their accommodation, suggestion and opinion during the project progress in university. In particular, I would like to thank all the staff and technicians, for their cooperation, indirect or directly contribution upon completing my project.

Most importantly, special thanks also to my family for their external support when I told them I wanted to continue my education and especially my friends for encouraging me to finish this project. Without their support the ideas could not have been realized.

TABLE OF CONTENTS

LIST OF FIGURE

CHAPTER 1

INTRODUCTION

1.0 Introduction

This chapter provides a summary of using sum of absolute differences (SAD) in determining the disparity mapping. The problem background and problem statement are described next. This is followed by the objective and the focuses of this research which involves the study of depth map reconstruction from stereo images utilise sum of absolute differences (SAD).

1.1 Background

Matching stereo is an issue in finding correspondence between two input pictures. This is one of the fundamental issues of computer vision for a wide variety of applications and has thus been thoroughly studied in the area of computer vision for many years. For each point in the left image, stereo matching consists of finding its corresponding in the right one. The dissimilarity between those points' horizontal distances is the disparity. A map of disparity consists of all attribute outcomes of disparity within an image. Such a map is essentially a representation of the perceived scene profundity. Hence, the maps of inequalities were used to resolve issues effectively.

There are three specific groups of techniques to match the stereo: area-based, phase-based, and feature-based techniques. The most beneficial area-based techniques in real-time stereo vision are SAD-based implementations because they can be executed directly in hardware. The computations needed in tenure of design units are readily, as there are only summaries and absolutism values. Simultaneous design units may be used to manage different ranges of disparities and to reduce the required computational time.

In general, it is essential to have at least two images for reconstruction of an object, please consider the relevant points on these images and to construct a map of depth. The quality of the acquired images can be determined by several factors: lighting, object shape, camera angle and camera focal length. Many stereo matching algorithms are based on the constraints of similarities, epipolarity, consistency, continuity and ordering and taking the basic process: matching cost calculation; matching cost aggregation; variance calculation; refinement of differences. For the images on the left and right, the matching expense is determined by the difference in the gray values of the corresponding pixels.

Traditional stereo matching methods, such as Sum of Square Differences (SSD) and Sum of Absolute Differences (SAD), They are designed for simplistic scenes and cannot handle less textured images, and are prone to light and noise variations. On the other hand, normalized cross-correlation (NCC) can avoid the noise better, and light variants would not be impacted, but it is a computation-intensive approach. Matching stereo consists of finding the corresponding in the right image for every juncture in the left image. The difference between all points in horizontal intervals is the variance. A map of disparity consists of all the probable values of disparity within an image. A rather map is essentially a depiction of the perceived scene size. Although this technique is successful in images without texture, it varies depending on centralized sample pixel, and therefore the corresponding results can be reduced if the noise affects the central pixel.

1.2 Problem Statement

Computer vision is actually a vital field of study. This requires techniques for the acquisition, processing, interpretation and comprehension of images. Computer vision approaches seek to use different mathematical methods to model a complex visual environment. Some of the objectives of computer vision are to describe the world we see on the basis with one or even more images and to restructure its properties, including the distribution of light, colour and shape. Stereo vision is a computer vision field that tackles a major research issue: Tri-dimensional point reconstruction maps the depth estimate.

A stereo vision system is made up of a stereo camera, including two cameras located horizontally (one on the left and one on the right). Then, the two captured images simultaneously by such cameras are analyzed to retrieve visual depth information. The challenge is to identify the correct method for estimating the discrepancies between the views displayed in the two images to map (e.g. plot) the environment's correspondence (e.g. disparity). Intuitively, a map of disparities represents the equivalent pixels, which are horizontally moved between the left and right images. Every year, new methods and techniques are developed to solve this problem, and show a trend towards improving accuracy and time consumption.

1.3 Objective

The objective of this project comes from some short of research and problem statement. This objective showed the reason and the outcomes of this project. The objective had state below:

- To develop stereo matching algorithm using Sum of Absolute Differences (SAD).
- To reconstruct depth map from the proposed algorithm.
- To analyze the performance of the proposed algorithm using Standard benchmarking evaluation system.

1.4 Project Scope

The aim of stereo matching is to draw distinctions from two very similar perspectives. Current stereo matching methods can induce incorrect matching when there are loads of disruptions in image noise and disparity. This project to achieve disparity (difference) maps is a set of extraordinary issues which remain to be addressed in stereo vision systems. Various techniques are used to provide maps of disparities.

Commonly preferred is the method that is based on stage geometry that is given by matching pixels between two images. The SAD algorithm attempts to modify the manner in which the disparity is measured by using the edge information derived from the captured stereo images. While less computational loads are needed by the edge operator used in the proposed process, the disparity map may be depicted with a low elapsing period while retaining the equal efficacy of the reconfiguration map. This tool is provided by MATLAB software.

1.5 Project Outlines

In chapter 1, it will be explaining briefly about the possibility of the project. Tasks background will be discussed in this part. This part will concentrate on the outlines of the outline of the undertaking, specifying the objective, the issue articulation, and the scope of the project.

In chapter 2, this section is about the idea, hypothesis, and some characteristic of equipment that utilized as a part of this task. This part also contains a meaning of term used as a part of this undertaking and furthermore discusses about the idea of the research and how it identified with the theory.

In chapter 3, this section will explain about the methodology. Methodology chapter is a step that need to be follow and detailed reports of studies that need to be complete to achieve the objective. This chapter describes the methodology required to complete the project and discusses the creation of the project.

In chapter 4, this section is about the result and discussion. It will explain the data that has been collected after the project has been completed. The comparison of the data will be discussed. The testing, implementing and troubleshooting throughout this project will be explained in this chapter.

In chapter 5, this section is about conclusion. It will explain the completed project either the objective is achieved or not. Recommendation and future improvement also can be suggested and discussed in this chapter.

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

The chapter will address the literature review where information on the relevant project is based. It will cover both the concept based on the previous project and the theory to achieve this project. This study focuses on the algorithm matching the Stereo and the method used for this related project.

2.1 Related Work

2.1.1 Depth Estimation in Disparity Mapping

In recent decades' numerous algorithms and applications have been presented to estimate disparity maps. According to (A.Qayyum, 2015) in method for monitoring trees and vegetation is proposed, based on satellite images. The stereo matching algorithms are calculated based on stereo satellite images to gage disparity maps. The estimate to depth map of the level of the trees and plants at the base points is inversely rate able to the level of disparity. Dynamic programming (DP) and block matching with energy minimization was proposed for the calculation of the depth map.

The pedestrian detection method is proposed in (Dongyoung, 2017) focused on a complex map of inequalities for Smarter Vehicles. The dense map of disparities is used to make pedestrian identification more effective. The method consists of several phases, the identification of hazard areas using information on road and column identification characteristics, the classification of pedestrian zones using segmentation based on complex maps with disparities and classification with pedestrians using the ideal criteria.

The work of (A. K. Wassim, 2016) is an object tracking investigation where a stereoscopic camera detects objects, make it a low cost solution for target detection. Its goal is to verify and track objects in a video frame throughout the video without previous knowledge of the objects. Calculate the map of disparities using a couple of stereo images. Instead, to detect object blobs, the map of disparities undergoes a profound segmentation and the corresponding area in rectified stereo image is the object of interest. Furthermore, in (Okae J. D., 2017) An effective optimization is given to increase the efficiency of the stereoscopic vision system and to precisely compare the measured map of disparities with the actual depth data.

With modern technology and artificial knowledge in a methodology (Lobaton Q. G., 2017) in a value multiples stereo disparity approach is proposed for the identification of sturdy obstacles for autonomous driving applications in outdoor scenes. The calculation of disparity is affected greatly by reflections, loss of texture and repeated patterns of objects. This can lead to misleading estimates, that can lead to some bias in approaches to obstacle detection that make use of the map of disparities. In order to address this problem, a new research is suggested instead of calculating the disparity of a single attribute, which uses the variety of candidates for each point of the picture. Based on a mathematical analysis distinguished by the success of various metrics, these are selected: the number and distance between the candidates according to the actual value of the disparity. It continues to create a location map that provides an estimate of the obstacles.

However, in (Z. Liang Y. F., 2018) explain the different phases in which the depth estimation is carried out: first, the extraction of functions, initial measurement and final optimization of the estimated depth. In (G. Song H. Z., 2017) suggest a new method of optimizations using tight smoothing constraints obtained in a neural network. The aim is to provide rigorous softening of the output disparity map. In this study, the first move was to question the architecture of CNN, called DD-CNN, to determine whether the disparities are discontinuous. Practice of this method was performed using real data from stereo data from Middlebury.

In the next stage, they identify an objective functions that consists of a given term using the (Zbontar J. a., 2015) method and a term designed to penalize disparities. These studies have found that the subnetwork framework produces a depth map through an architects that encode and decodes information based on DispNetCorr1D (N. Mayer E. l., 2016). Finally, a dual-structured network is built in (W. Luo, A. G. Schwing and R. Urtasun, 2016). The system requires an input image that moves through a finite range of layers followed by standardization and a linear rectified construct. For each layer, various filters were tested in their experiments, and the parameters were exchanged between both the two structures.

Figure 2.1: (a) combined images; (b) depth map; (c) 3D model of the depth map.

2.1.2 Stereo matching based on SAD

The stereo vision is the most important field of computer vision and it provides various algorithms for computing the map of disparity. Through using two stereo images, the scene Depth can be obtained from two separate points with some displaced values. The correlation values of the left image compared with the right image are the result of stereo matching. The photos depth map is determined using the disparity map or stereo matching. The stereo matching function for computing the exact depth map is very difficult. According to (P, 2017) proposed a cost aggregation approach based on the segment-tree for non-local stereo matching, leading to improvements in both the precision of disparities and the processing speed. In placing more emphasis, (K. Zhang Y. F., 2017) a cross-scale design was suggested to enhance the cost averaging for effective stereo matching. (Cigla, 2015) Recursive edge-aware (REAF) filters provided for precise and effective stereo matching.

In the case of the global stereo matching, the disparities are calculated by decreasing the global energy feature. An algorithm to identifying depth discontinuities from pairs of stereo images. Their approach can handle dynamic programming acceleration. A new stereo matching based on segmentation using graph cuts, which is used by assigning disparity planes to each segment to achieve the optimal solution. (Weijer, 2015) Suggested integration of cost allocation-filtering approaches and global energy minimization approaches to encourage increased stereo matching using a two-stage energy minimization algorithm using MRF modelling. Its solution can be used to successfully address the problem of stereo matching in occlusion areas.

In several practical applications (C. Kwan, 2018), It is essential that the images are synchronized with different views. According to various perspectives, the parallax poses severe problems in the synchronization of images, as most capture algorithms only refer to pictures with flattened material. Problem of registration for images having different content depths. This can be seen that the communication of features can either be done in the foreground or in the background, not both.

Its technique could be used to effectively address the issue of stereo matching in occluded regions. (LeCun, 2015) Used the Neural Network Convolution approach to predict matching image patches and calculated stereo matching costs, further optimized by cross-based expense agglomeration and sub-global matching. However, (W. Luo, 2016)proposed a deep learning network to produce accurate results efficiently on GPU. Use of the matching approach semi-global, the author (Pollefeys, 2017) presented an estimation method for learning based penalties to predict detailed estimates of dense disparity map.

While previous approaches can effectively yield precise disparities in stereo matching, It isn't easy to enforce them, and complex scenes may fail to implement them. In addition, the learning-based methods are not reliable, but are dependent on the training data. In this project, they suggest a stable and efficient SAD-based algorithm. My method is simple to apply and my results are somewhat close to those produced by the use of state-of-the-art models in the public data collection.