Faculty of Electrical and Electronic Engineering Technology

DEVELOPMENT OF DEPTH MAP RECONSTRUCTION ALGORITHM FROM STEREO IMAGES USING LOCAL-BASED TECHNIQUE

IZZAT BIN MOHAMAD SAUPI

Bachelor of Computer Engineering Technology (Computer Systems) with Honours

2022

DEVELOPMENT OF DEPTH MAP RECONSTRUCTION ALGORITHM FROM STEREO IMAGES USING LOCAL-BASED TECHNIQUE

IZZAT BIN MOHAMAD SAUPI

A project report submitted
in partial fulfillment of the requirements for the degree of
Bachelor of Computer Engineering Technology (Computer Systems) with Honours

Faculty of Electrical and Electronic Engineering Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2022

UNIVERSITI TEKNIKAL MALAYSIA MELAKA
FAKULTI TEKNOLOGI KEJUTERAAN ELEKTRIK DAN ELEKTRONIK

BORANG PENGESAHAN STATUS LAPORAN
PROJEK SARJANA MUDA II

Tajuk Projek : DEVELOPMENT OF DEPTH MAP RECONSTRUCTION ALGORITHM FROM STEREO IMAGE USING LOCAL-BASED TECHNIQUE

Sesi Pengajian : 2021/2022

Saya IZZAT BIN MOHAMAD SAUPI mengaku membenarkan laporan Projek Sarjana Muda ini disimpan di Perpustakaan dengan syarat-syarat kegunaan seperti berikut:

1. Laporan adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan laporan ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (✓):

☐	**SULIT***	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
☐	**TERHAD***	(Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
☑	**TIDAK TERHAD**	

Disahkan oleh:

(TANDATANGAN PENULIS)

Alamat Tetap: LOT 2826,
KAMPUNG BARU
LUBOK SETOL,
17200 RANTAU PANJANG,
KELANTAN

Tarikh: 04/01/2022

(COP DAN TANDATANGAN PENYELIA)

Ts. DR. ROSTAM AFFENDI BIN HAMZAH
Dean
Faculty of Electrical & Electronic Engineering Technology
Universiti Teknikal Malaysia Melaka

Tarikh: 11/1/2022

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

DECLARATION

I declare that this project report entitled "DEVELOPMENT OF DEPTH MAP RECONSTRUCTION ALGORITHM FROM STEREO IMAGES USING LOCAL-BASED TECHNIQUE " is the result of my own research except as cited in the references. The project report has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : 

Student Name : IZZAT BIN MOHAMAD SAUPI

Date : 04/01/2022

APPROVAL

I hereby declare that I have checked this project report and in my opinion, this project report is adequate in terms of scope and quality for the award of the degree of Bachelor of Computer Engineering Technology (Computer Systems) with Honours.

Signature : 

Supervisor Name : .....

Date : 11/1/2022

Ts. DR. ROSTAM AFFENDI BIN HAMZAH
Dean
Faculty of Electrical & Electronic Engineering Technology
Universiti Teknikal Malaysia Melaka

Signature : .....

Co-Supervisor : .....

Name (if any)

Date : .....

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

ABSTRACT

This project provides a method for solving the issue of correspondence when matching the stereo image using local-based techniques. This technique is called "Sum of Absolute Differences" (SAD). MATLAB software provides the tool. A map of disparities is produced through the block matching algorithm, Sum of Absolute Differences (SAD). For example, there are four basic steps in the stereo vision method for imaging the reconstruction. There are typically four phases to this. The distortion of the captured images from the camera lens is extracted first in the undistortion 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 comparison between the left and 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 a stereo matching algorithm using the sum of absolute differences to do the depth map reconstruction from the proposed algorithm and a simple stereo algorithm that computes results comparable to the current state-of-the-art on the Middlebury benchmark.

ABSTRAK

Projek ini menyediakan kaedah untuk menyelesaikan isu surat-menyurat apabila memadankan imej stereo menggunakan teknik berasaskan tempatan. Teknik ini dipanggil "Sum of Absolute Differences" (SAD). Perisian MATLAB menyediakan alat tersebut. Peta ketaksamaan dihasilkan melalui algoritma padanan blok, Jumlah Perbezaan Mutlak (SAD). Sebagai contoh, terdapat empat langkah asas dalam kaedah penglihatan stereo untuk pengimejan pembinaan semula. Biasanya terdapat empat fasa untuk ini. Herotan imej yang ditangkap daripada kanta kamera diekstrak terlebih dahulu dalam langkah tidak herot. Langkah seterusnya ialah melaraskan jarak dan sudut ketinggian antara dua imej kamera untuk menentukan jarak fokus dan paksi epipolar, seperti yang diandaikan di luar. Perbandingan antara imej kiri dan kanan akan dikira dalam peringkat surat-menyurat dan digunakan untuk mengukur peta jurang. Kaedah ini sering dikenali sebagai saling boleh kendali. Projek ini dapat membangunkan algoritma pemadanan stereo menggunakan jumlah perbezaan mutlak untuk melakukan pembinaan semula peta kedalaman daripada algoritma yang dicadangkan dan algoritma stereo ringkas yang mengira hasil yang setanding dengan keadaan terkini pada penanda aras Middlebury.

ACKNOWLEDGEMENTS

First and foremost, I am very grateful to the almighty Allah S.W.T for letting me finish my Bachelor Degree Project and giving me the strength that I need to fulfil my duty as an electronic student.

A special thanks to my supervisor, Ts. Dr. Rostam Affendi Bin Hamzah, for his guidance and instruction while I was completing my Bachelor Degree Project over the course of two semesters. May Allah always bless you and I will never forget your advice and guidance when I make a mistake.

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

Finally, thanks also to all the kind lecturers in the Bachelor of Electronic Computer Engineering Technology (System Computer) with Honours section for their accommodation, suggestions, and opinions during the project progress in university. I would like to thank all the staff and technicians for their cooperation, whether indirect or direct, in completing my project.

TABLE OF CONTENTS

PAGE

DECLARATION

APPROVAL

DEDICATIONS

ABSTRACT **i**

ABSTRAK **ii**

ACKNOWLEDGEMENTS **iii**

TABLE OF CONTENTS **i**

LIST OF TABLES **iii**

LIST OF FIGURES **iv**

LIST OF SYMBOLS **vi**

LIST OF ABBREVIATIONS **vii**

LIST OF APPENDICES **viii**

CHAPTER 1 INTRODUCTION **9**

1.0 Introduction 9
1.1 Background 9
1.2 Problem Statement 10
1.3 Project Objective 11
1.4 Scope of Project 12

CHAPTER 2 LITERATURE REVIEW **13**

2.0 Introduction 13
2.1 Related Work 13
2.1.1 Depth Estimating in Disparity Mapping 13
2.1.2 Stereo matching based on SAD 15
2.1.3 3D Reconstruction 18
2.2 Previous Work 19
2.2.1 Stereo Image and Depth Map Generation for Image with Different Views and Resolutions 19
2.2.2 Improve Accuracy of Disparity Map for Stereo Image using SIFT and Weighted Color Model 21
2.2.3 Ceation of Depth Map from Stereo Images of Faces for 3D Model Reconstruction 22
2.2.4 Computer Vision Based Distance Measurement System using Stereo Camera View 23

2.2.5 Disparity map estimation with deep learning in stereo vision. 24
2.2.6 Efficient Binocular Stereo Matching Based On Sad and Improved Census Transformation 25
2.2.7 3D Distance Measurement Accuracy On Low-Cost Stereo Camera 26
2.3 Summary 27

CHAPTER 3 METHODOLOGY 28
3.0 Introduction 28
3.1 Project Overview 28
3.2 Flowchart represent process of project 28
3.3 Block Diagram of Project 29
3.4 Program Development 30
3.5 Software Implementation 30
3.6 Stereo Vison Principle Depth Map Generation 31
3.7 Sum of Absolute Differences (SAD) 33
3.8 The Examples of SAD in MATLAB 34
3.9 Summary 35

CHAPTER 4 RESULTS AND DISCUSSIONS 36
4.0 Introduction 36
4.1 Dataset 36
4.2 GUI demonstrates an efficient stereo matching algorithm. 38
4.3 Depth Mapping Propagation 41
4.4 Summary 49

CHAPTER 5 50
5.1 Conclusion 50
5.2 Future Works 51

REFERENCES 52

APPENDICES 58

LIST OF TABLES

TABLE	TITLE	PAGE
Table 4.1	Dataset from Middlebury.	37
Table 4.2	Results of algorithm.	42
Table 4.3	Comparison Disparity with difference window size.	43
Table 4.4	3D Reconstruction.	47

LIST OF FIGURES

FIGURE	TITLE	PAGE
Figure 2.1	a) Combined images ; b) depth map ; c) 3D model of the depth map	15
Figure 2.2	Comparisons of stereo matching.	17
Figure 2.3	Standard benchmarking images provided in the Middlebury Stereo Vision datasets.	18
Figure 2.4	Signal flow of a new stereo image formation and depth map generation system.	20
Figure 2.5	Procedures in stereo image formation function.[17]	21
Figure 2.6	Circle indicates bounded area suggested by SIFT. Patch window (M). Square indicates “disparity max” area.[32]	22
Figure 2.7	Camera calibration using a chessboard.	23
Figure 2.8	Calculating the depth map. (T-Stereo base, Z – distance).	23
Figure 2.9	Block diagram of stereo vision system establish.	24
Figure 2.10	Computing inequality between two points on disparity maps.	24
Figure 2.11	Components of a typical convolutional network.	25
Figure 2.12	Architecture of the convolutional network.	25
Figure 2.13	Flowchart stereo matching.	26
Figure 2.14	Calibration pattern : (a) Chessboard (b) Feature Descrptor	26
Figure 3.1	Project development flowchart.	29
Figure 3.2	Block diagram of project.	29
Figure 3.3	MATLAB.	31
Figure 3.4	Two cameras’ location and their projection planes image.	32
Figure 3.5	Relationship between disparity and depth.	33
Figure 3.6	Process Disparity Map	34
Figure 4.1	Graphical user interface for background stereo vision.	38

Figure 4.2 View images from the GUI. 39

Figure 4.3 Grayscale stereo matching 40

Figure 4.4 Color stereo matching 41

Figure 4.5 Depth-mapping reconstruction 41

LIST OF SYMBOLS

LIST OF ABBREVIATIONS

SAD	-	Sum Absolute Difference
2D	-	Two Dimensional
3D	-	Three Dimensional

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
Appendix 1	Gantt Chart Project Development PSM 1	58
Appendix 2	Gantt Chart Project Development PSM 2	59
Appendix 3	Code of GUI	60
Appendix 4	Code for SAD algorithm	61

CHAPTER 1

INTRODUCTION

1.0 Introduction

This chapter summarizes the procedure for finding the disparity mapping using the local-based approach known as the Sum of Absolute Differences (SAD). Besides, the problem's background and problem statements are explained next. Following the research's objective and main purpose, which is to investigate depth map reconstruction from stereo images using the Local-Based Technique.

1.1 Background

In finding correspondence between two input pictures, there is an issue with matching stereo. This is one of the fundamental issues of computer vision for a wide variety of applications and has thus been thoroughly studied in computer vision for many years.[1] Stereo matching is the process of determining which point in the right picture corresponds to which position in the left image. The disparity, which is the similarity in horizontal distances between those sites, A disparity map is a collection of all the disparity-related outcomes associated with a picture. This style of map is essentially a depiction of the scene's perceived profundity. As a result, inequality maps were employed to address challenges successfully.

There are three types of stereo matching techniques: area-based, phase-based, and feature-based. Due to the fact that SAD-based applications may be run directly on hardware, they are the most beneficial area-based techniques in real-time stereo vision. Since there are

only summaries and absolutism principles, the computations required in the tenure of design units are easily accessible. Simultaneous design units may be used to accommodate diverse sets of inequalities while still reducing computing time.

In general, at least two images are required for the reconstruction of an object; please consider the related points on these images and create a depth map. Several factors influence image quality, including illumination, object type, camera angle, and camera focal length. Many stereo matching algorithms are based on the constraints of similarities, epipolarity, consistency, continuity, and ordering and take the basic steps of matching cost calculation, matching cost aggregation, variance calculation, and refinement of differences.[2] The matching cost for the images on the left and right is measured by the disparity in grey values of the respective pixels.

Traditional stereo matching techniques, such as Sum of Square Differences (SSD) and Sum of Absolute Differences (SAD),[3] are built for simple scenes and are susceptible to light and noise variations. Normalized cross-correlation (NCC), on the other hand, can avoid noise better and does not affect light variants, but it is a computationally expensive technique. Stereo matching is all about locating the equivalent picture in the right image for each junction in the left image. The variance is the difference between all points in horizontal intervals. A disparity map is made up of all the possible disparity values within a picture. A road map is nothing more than a depiction of the scene's apparent size. While this method works well in images with no texture, it varies depending on the centralised sample pixel, so the resulting effects can be minimised if noise influences the central pixel.

1.2 Problem Statement

Computer vision is an important area of research.[4] This necessitates image collection, delivery, presentation, and comprehension strategies. Computer vision

approaches use a variety of mathematical methods to recreate a dynamic visual environment. One of the aims of computer vision is to describe the environment we perceive using one or more pictures and to reconstruct its features, such as light, color, and shape distribution. Stereo vision is a branch of computer vision that handles a significant research challenge: depth estimation is mapped using tridimensional point reconstruction.

A stereo vision system is comprised of a stereo display and two sensors positioned horizontally (one on the left and one on the right). Following that, the two pictures acquired concurrently by such cameras are analyzed to yield visual depth information. The objective is to identify the most efficient way of computing the differences between the two photos in order to chart (e.g., plot) the environment's correlation (e.g., disparity). A chart of inequality intuitively depicts the horizontally displaced pixels between the left and right pictures. Each year, new ideas and strategies are developed to handle this issue, with a focus on accuracy and efficiency.

1.3 Project Objective

The aim of this project is derived from some analysis and a problem statement. This goal explained the aim of the project as well as its results. The following was mentioned as the goal:

a) To develop stereo matching algorithm using Sum of Absolute Differences (SAD).

b) To reconstruct depth map from the proposed algorithm.

c) To analyze the performance of the proposed algorithm using Standard benchmarking evaluation system.

1.4 Scope of Project

Stereo matching's objective is to discriminate between two very close perspectives. When there are many picture noise and divergence disturbances, conventional stereo matching algorithms might result in erroneous matching. This endeavor to build divergence (difference) maps demonstrates a number of outstanding issues that remain unanswered in stereo vision systems. Numerous strategies are used to generate disparity maps.

The stage geometry produced by comparing pixels between two photographs is a frequently used method. By using edge information extracted from recorded stereo images, the SAD algorithm tries to alter the way disparity is computed. While the suggested technique uses a less computationally intensive edge operator, the disparity map may be rendered with a short elapsed time while retaining the same potency as the reconfiguration map. This capability is provided by the MATLAB program.

.

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

This section will analyze the literature review on which the necessary project information is based. It will cover both the previous project's concept and the theory required to complete this project. This research focuses on the algorithm that combines the Stereo and the technique that was used for this related project.

2.1 Related Work

2.1.1 Depth Estimating in Disparity Mapping

Several techniques and applications for estimating disparity maps have been proposed in recent decades. A system for monitoring trees and plants based on satellite photos is presented[5]. To generate disparity maps, stereo matching algorithms are computed from stereo satellite images. The estimated depth map of the level of trees and plants at the base points is inversely proportional to the level of disparity. The computation of the depth map was suggested using dynamic programming (DP) and block matching with energy reduction.[6]

[7] provides a pedestrian recognition system based on a complex map of inequalities for Smarter Vehicles. The dense map of disparities is used to improve pedestrian identification. The technique is divided into several phases: hazard area identification using information on road and column identification characteristics, pedestrian zone classification

using segmentation based on complicated maps with disparities, and pedestrian zone classification using ideal criteria.

[8] conducted an object tracking experiment in which a stereoscopic camera identifies objects, resulting in a low-cost method for target identification. Its purpose is to check and track things in a video frame over the duration of the film without prior knowledge of the items. Using a handful of stereo images, compute the disparity map. Instead, to detect object blobs, the disparity map is deeply segmented, and the appropriate region in the corrected stereo picture represents the item of interest.

A methodology [9] in a value multiples stereo disparity method is presented for the detection of sturdy obstacles for autonomous driving applications in outdoor scenarios using modern technologies and artificial knowledge. Reflections, texture degradation, and repetitive object patterns all have a significant impact on disparity calculation. This can cause wrong estimations, which can contribute to bias in techniques to obstacle detection that use the disparity map. To overcome this issue, instead of calculating the disparity of a single attribute, a new study is proposed that employs a wide range of possibilities for each location in the picture. These are chosen based on a mathematical analysis distinguished by the success of several metrics: the number and distance between the candidates based on the actual value of the discrepancy. It continues to generate a location map that estimates the obstacles.

However, [10] describe the many phases of depth estimation: first, function extraction, initial measurement, and final optimization of the predicted depth. [11] propose a new optimization method based on tight smoothing constraints achieved in a neural network. The goal is to offer rigorous softening of the output disparity map. The initial step in this investigation was to investigate the CNN design, known as DD-CNN, to see if the

inequalities were discontinuous. This approach was tested using real data from Middlebury's stereo data.

Then, using the [12] technique, they define an objective function that consists of a given term and a term designed to penalize disparities. Finally, a dual-structured network is constructed [13]. The system requires an input picture that goes through a finite number of layers before being standardized and linear corrected. In their studies, several filters were evaluated for each layer, and the parameters were exchanged between 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

Stereo vision is the most significant field of computer vision, and it provides many techniques for constructing the disparity map. The scene Depth may be derived from two different locations with certain displaced values combining various stereo images. The correlation values of the left image compared to the right image are the outcome of stereo matching. The disparity map or stereo matching is used to determine the depth map of a photograph.[6] It is quite difficult to compute the accurate depth map using the stereo matching function. In order to place more focus, [14] proposed a cross-scale design to improve cost averaging for effective stereo matching. [15] Recursive Edge-Aware Filters (REAF) allowed for accurate and accurate stereo matching.[2]

The disparities in global stereo matching are estimated by reducing the global energy feature. A method for detecting depth discontinuities in stereo images. Their method is capable of handling dynamic programming acceleration. A new stereo matching method