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

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

Date : 11/1/2022

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

LIST OF TABLES

TABLE **TITLE** **PAGE**

LIST OF FIGURES

LIST OF SYMBOLS

LIST OF ABBREVIATIONS

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

LIST OF APPENDICES

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 based on graph cuts segmentation that achieves the best result by assigning disparity planes to each segment. [16] Using a two-stage energy reduction algorithm based on MRF modelling, it was suggested that cost allocation-filtering techniques and global energy reduction techniques be used to encourage improved stereo matching. Its solution is effective in solving the problem of stereo matching in occlusion areas.

The images must be synced with different views in many practical applications [17]. According to several viewpoints, parallax causes serious issues in image synchronization because most capture algorithms only refer to flattened images. Registration issue for images with different content depths. As can be seen, feature communication can be done in either the foreground or the background, but not both.[17]

Its method might be applied to solve the problem of stereo matching in obstructed areas. [18] Predicted matching picture patches using Neural Network Convolution and computed stereo matching costs, which were further improved using cross-based expense agglomeration and sub-global matching. [13], on the other hand, suggested a deep learning network that could deliver correct results quickly on a GPU. The author [19] provided an estimation method for learning based penalties to anticipate detailed estimations of dense disparity map using a semi-global matching technique.

While previous methods can effectively provide exact stereo matching disparities, they are difficult to enforce, and complicated scenarios may fail to apply them. Furthermore, learning-based approaches are unreliable since they rely on training data. They propose a SAD-based method that is both stable and efficient in this project. My approach is straightforward to use, and the results are comparable to those obtained using state-of-the-art models in public data collecting.[2]

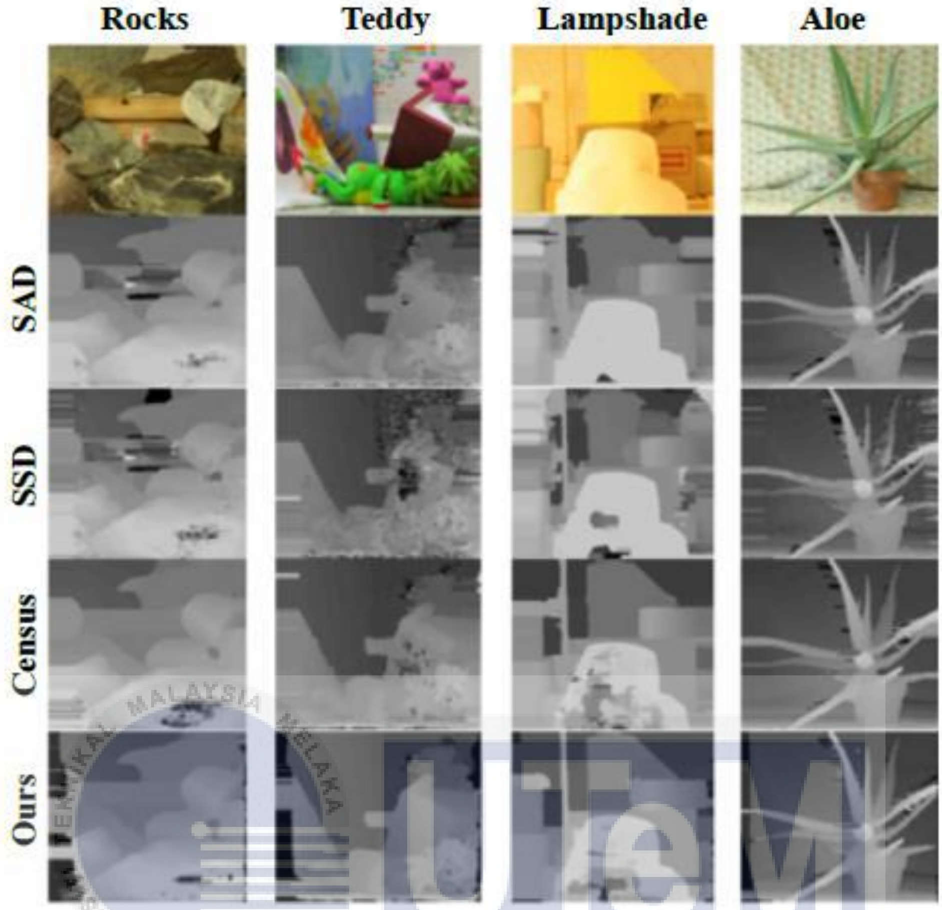


Figure 2.2 Comparisons of stereo matching.

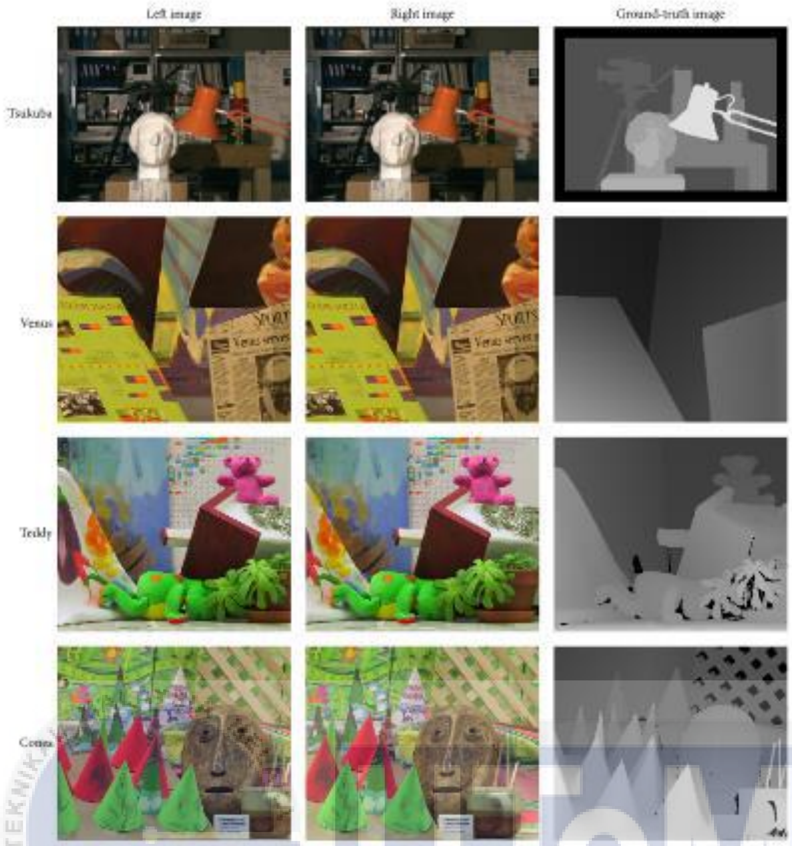


Figure 2.3 Standard benchmarking images provided in the Middlebury Stereo Vision datasets.

2.1.3 3D Reconstruction

3D Reconstruction in a Single-View Theoretically, recovering 3D structure from single-view photos is an ill-posed issue.[20] Many attempts to solve this problem have been made, such as ShapeFromX [21] [22], where X can represent silhouettes [23], shading [24], and texture. These methods, however, are rarely relevant in real-world situations since they all need strong presumptions and extensive natural imagery skills [25].

By using a single-view image as an input, 3D-Vale-GAN [26] uses GAN and VAE to generate 3D structures with the performance of generative adversarial networks (GANs) [27] and variational autoencoders (VAEs). 3D-VAE-GAN, on the other hand, contains class marks for restoration. MarrNet [26] reconstructs 3D objects by estimating depth, surface normal, and silhouettes from 2D photos, which is challenging and often results in significant distortion. OGN and O-CNN use Octree to represent higher resolution 3D volumetric objects with a constrained memory budget [28]. The points in the point cloud representation, however, have a large degree of independence due to the minimal relations between them. As a result, these methods will be unable to reliably restore 3D quantities.

SfM and SLAM multi-view 3D reconstruction methods are effective in a variety of situations. These methods compare image characteristics and estimate the camera's pose for each image. When many viewpoints are separated by a wide margin, however, matching becomes difficult. Furthermore, scanning all an object's surfaces before reconstruction is often impossible, leading in partial 3D shapes with occluded or hollowed-out areas [29].

Large-scale datasets of 3D CAD models have been used to power deep-learning-based methods for 3D reconstruction. Both 3D-R2N2 and LSM use RNNs to derive 3D shape from single or multiple input images, with outstanding results. Even yet, RNNs are time-consuming and permutation-variant, resulting in inconsistency in reconstruction outcomes. 3DensiNet [30] uses max pooling to aggregate features from several images. Max pooling, on the other hand, extracts maximum values from features while ignoring other useful characteristics for 3D reconstruction.

2.2 Previous Work

2.2.1 Stereo Image and Depth Map Generation for Image with Different Views and Resolutions

Some early findings in developing stereo images and depth maps of various resolutions and view angles using uncalibrated images, according to [17][31]. Their goal is to see whether images of various resolutions and angles of view can be merged to generate high-quality stereo images and depth maps. To start, using a high-resolution image to improve a low-resolution image (assuming the low-resolution image is inaccurate and lacks generality) (assuming it is the right one). To obtain stereo and depth maps of the same quality, combine the changed left image with the previous, high-resolution right image.[17]

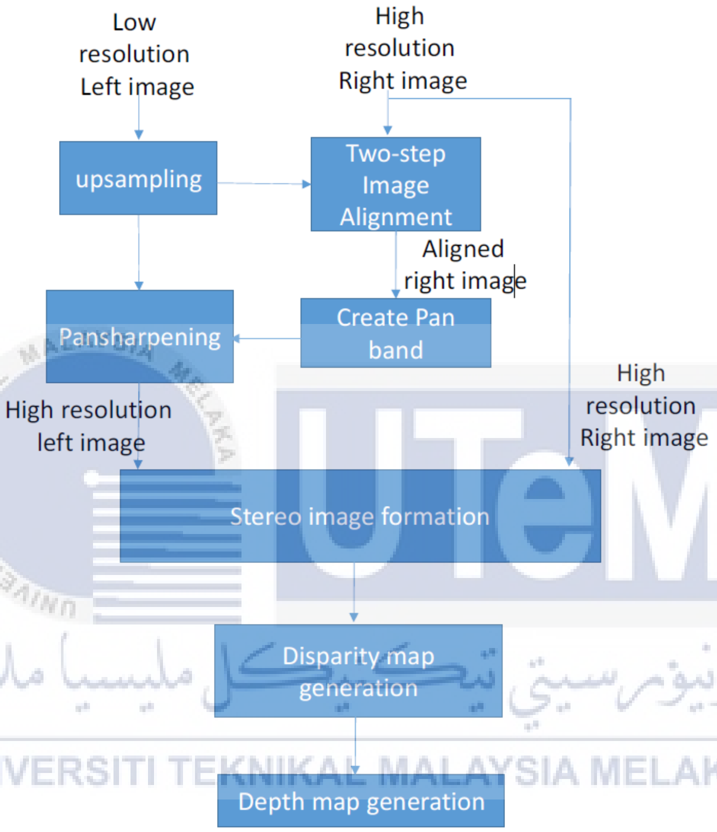


Figure 2.4 Signal flow of a new stereo image formation and depth map generation system.

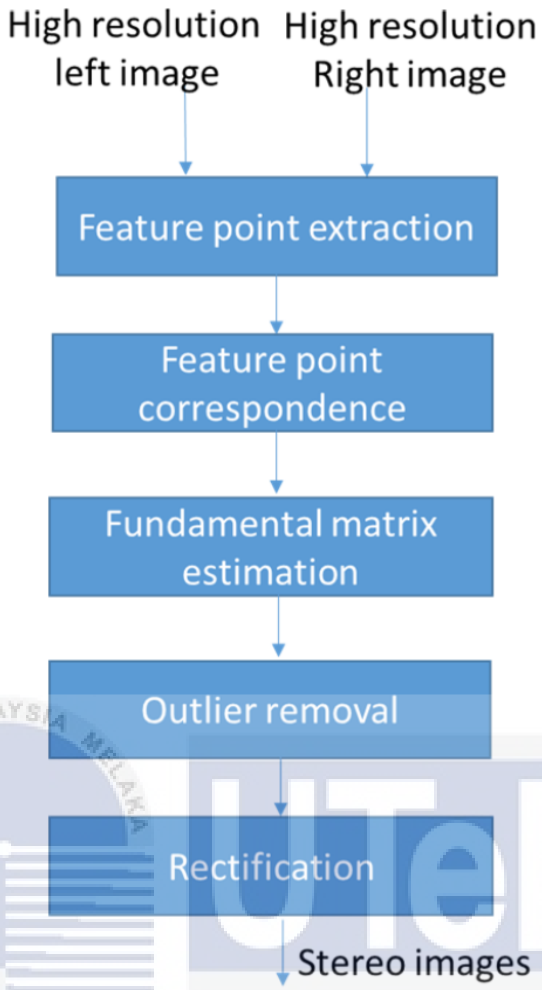


Figure 2.5 Procedures in stereo image formation function.[17]

2.2.2 Improve Accuracy of Disparity Map for Stereo Image using SIFT and Weighted Color Model

According to the author [32], SIFT algorithms and weighted YCbCr are used to minimize computation time and improve accuracy while generating a disparity map from stereo images.[32] To match the objects, SIFT is used to link the search area in the left and right images. Weighted YCbCr is used to improve matching accuracy.[32] As a result, the study of disparity maps determines the speed and accuracy of depth estimate.

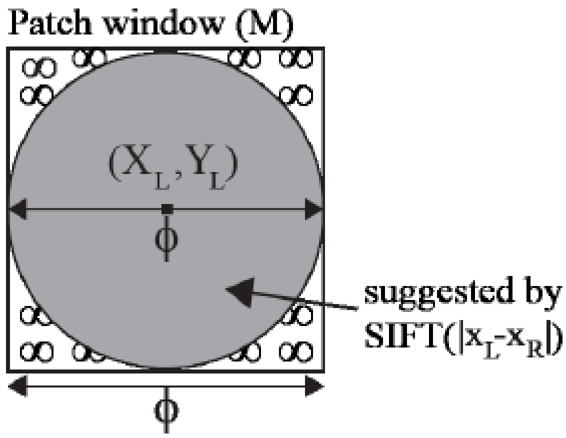


Figure 2.6 Circle indicates bounded area suggested by SIFT. Patch window (M). Square indicates "disparity max" area.[32]

2.2.3 Ceation of Depth Map from Stereo Images of Faces for 3D Model Reconstruction

Stereo cameras are one of the ways for reconstructing 3D face models that are now accessible. In most cases, the reconstruction procedure entails several steps [33]. A virtual scene with stereo cameras and a human head was created using the "Autodesk 3Ds Max" 3D editor. The proposed method has previously been tried with two "VISAR" cameras and a "Arduino Micro" microcontroller. Using the "Arduino Micro" microcontroller, the "Arduino" programmed allows cameras to be integrated. An application named "FlyCap" is used to collect the images from the cameras. Because the first images have distortions, the algorithm's first step is to calibrate the camera. For calibration, similar dots are placed on all stereo images. Following that, those spots are used to determine the degree of disruption, and the images are corrected as needed. For computing the depth map, the corrected images are used. The depth map of the faces is constructed using the front half-tone images.

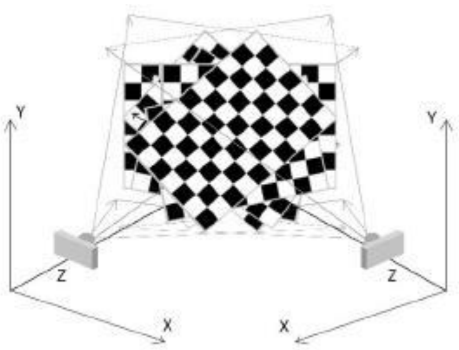


Figure 2.7 Camera calibration using a chessboard.

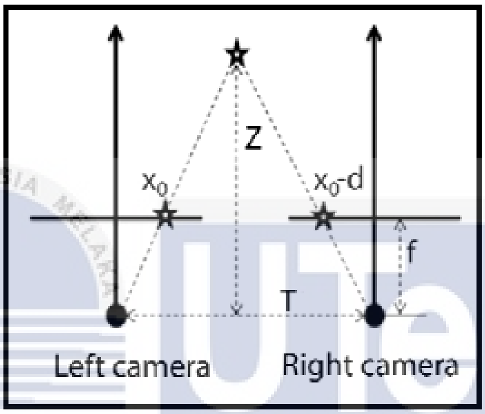


Figure 2.8 Calculating the depth map. (T-Stereo base, Z – distance).

2.2.4 Computer Vision Based Distance Measurement System using Stereo Camera View

Only stereo cameras can produce computer vision that is like human vision. A computer vision system is designed to measure object distances using the stereo camera system. This is confirmed by [34], which states, "Measure the distance from the stereo camera to the screen of the produced face images". The range measurements are then carried out on the photographs taken with the stereo camera system, with the differences between the frames being measured. The tool's function was measured in the final step by comparing a calculated Euclidean distance to the real distance values for the suggested computer vision system.[34]

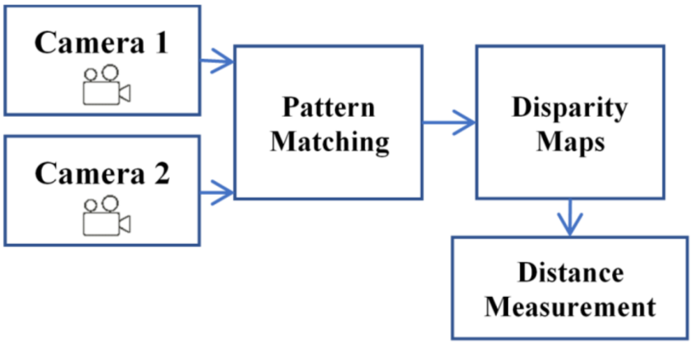


Figure 2.9 Block diagram of stereo vision system establish.

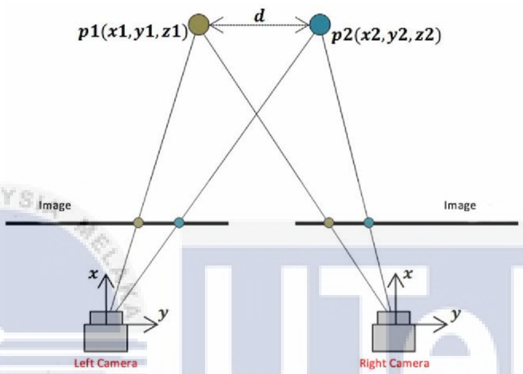


Figure 2.10 Computing inequality between two points on disparity maps.

2.2.5 Disparity map estimation with deep learning in stereo vision.

It has been demonstrated how a new completely networked convolutional architecture can estimate the disparity map between stereo pictures. With the information obtained, it is easy to observe how post-processing of the source images allows for the discernment of edges in the images, which appears to be the main difficulty to be addressed as a future stage in the research to provide more detailed findings in [35]. The disparity map is produced from the corrected stereo image pair using this method.[36] To anticipate stereo depth images, they proposed a new convolution layer-based neural network architecture. To train the network, the Middlebury databases were used to equate predicted map error with a known map of disparities.

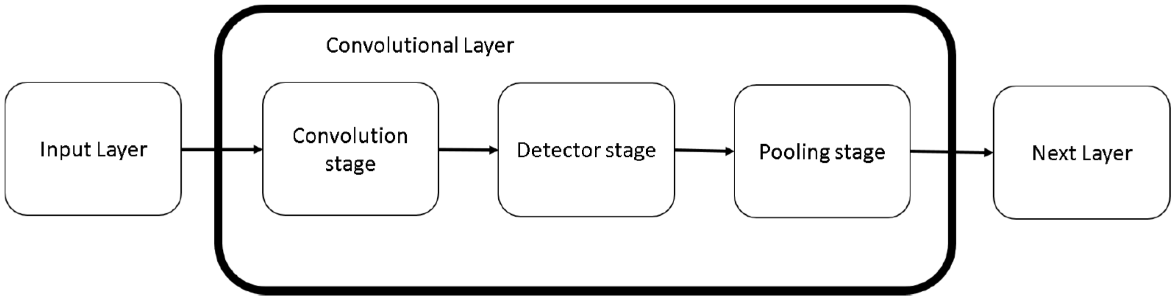


Figure 2.11 Components of a typical convolutional network.

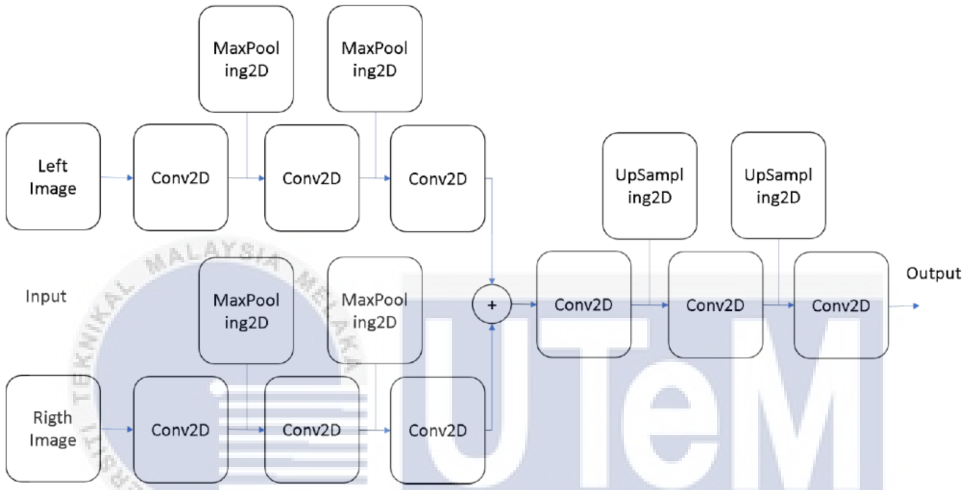


Figure 2.12 Architecture of the convolutional network.

2.2.6 Efficient Binocular Stereo Matching Based On Sad and Improved Census Transformation

A new SAD-based stereo matching algorithm and improved census transformation were suggested in [2]. They've also improved stereo matching performance by implementing innovative bilateral and selective filters to reduce noise and differences. First, do improved census transformation by combining SAD and improved census transformation,[2] and then they estimate equivalent costs. Finally, they stack up all the costs and compute the differences. They also propose that enhanced selective filters and bilateral be used to enhance the quality of disparities to generate more disparities.

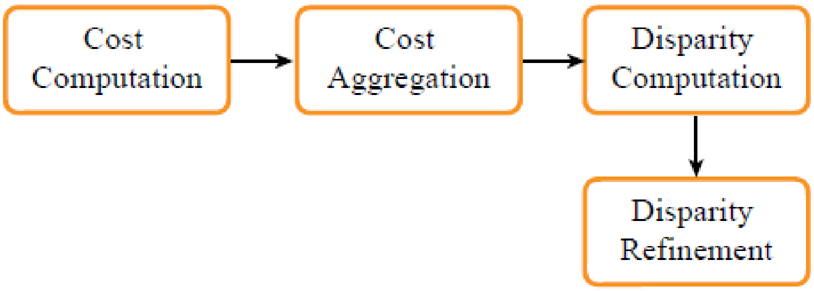


Figure 2.13 Flowchart stereo matching.

2.2.7 3D Distance Measurement Accuracy On Low-Cost Stereo Camera

In the Logitech HD Webcam C270.0 offers a low-cost 3D sensor using a pair of mid-range webcams.[37] Two calibration techniques were evaluated based on their performance when applied to the custom 3D sensor. The first approach considered, which used the checker board pattern, outperformed the second way, which utilized the particular pattern. The former has the shortest MRE of 0.1272 pixels and produces high-quality 3D photos that accurately portray the genuine physical 3D scene.

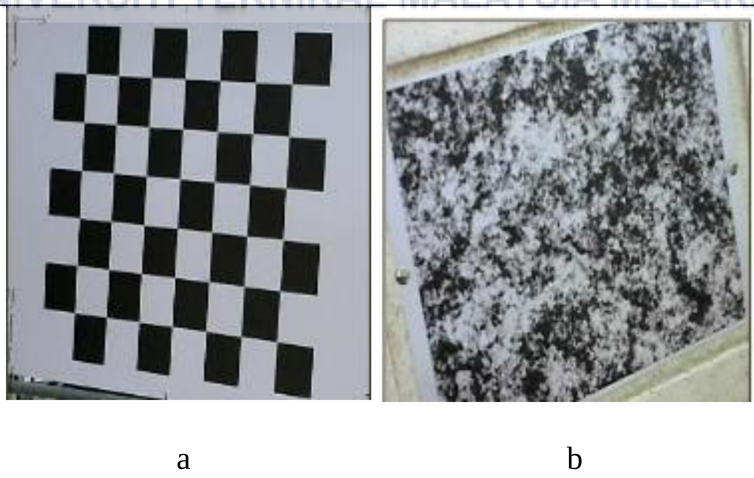


Figure 2.14 Calibration pattern : (a) Chessboard (b) Feature Descrptor

2.3 Summary

The predicted depth map of tree and plant level at the basis points is inversely proportional to the amount of discrepancy. The disparity map is used to enhance pedestrian recognition. Compute the disparity map using a small number of stereo images chosen based on mathematical analysis and the success of numerous criteria. The disparity map, often known as stereo matching, is used to determine the depth map of an image. Using the stereo matching function to generate an accurate depth map is a tough task. 3D reconstruction needs strong assumptions as well as significant natural imagery capabilities. It works well in several circumstances. For each image, these techniques compare image attributes and estimate the camera's point. Using uncalibrated images, create stereo images and depth maps of varying resolutions and view angles. SIFT methods and weighted YCbCr are used to minimize computation time and improve accuracy when constructing a disparity map from stereo pictures. Stereo cameras are one of the new methods for reconstructing 3D face reconstructions. Only stereo cameras can produce computer vision that is comparable to human eyesight. The disparity map is constructed using this method from the rectified stereo image pair. It improves stereo matching performance by incorporating novel bilateral and selective filters to reduce noise and differences. After being tested on the custom 3D sensor, two calibration methods were explored.

CHAPTER 3

METHODOLOGY

3.0 Introduction

This section will outline the project's flow and the route that must be taken to accomplish the project's objectives. Additionally, it covers the design of the experiment's concept, which will be discussed in this chapter. It will walk you through the process of developing the software that will be utilized to accomplish the project's goals. Each item in this procedure is critical since it will determine if the project's outcome is effective. This section illustrates the technique using a flowchart and a block diagram.

3.1 Project Overview

The technique and procedure of the project will be enhanced as a consequence of this chapter. Researchers use this practice to ensure that their aims are satisfied and that the finest possible findings are produced. There is an introduction, a protocol, and a flowchart for the project. This chapter is crucial before going on to the following chapter, which focuses on locating and evaluating information.

3.2 Flowchart represent process of project

Every project that is developed must have a flowchart to help the reader understand the project's function through a visual diagram. It is possible to have a better understanding of how a process works by using a flowchart, and it may also be used to improve the process. The disparity map is generated by first converting the left and right stereo pictures to grayscale images and then using the SAD technique to those images.

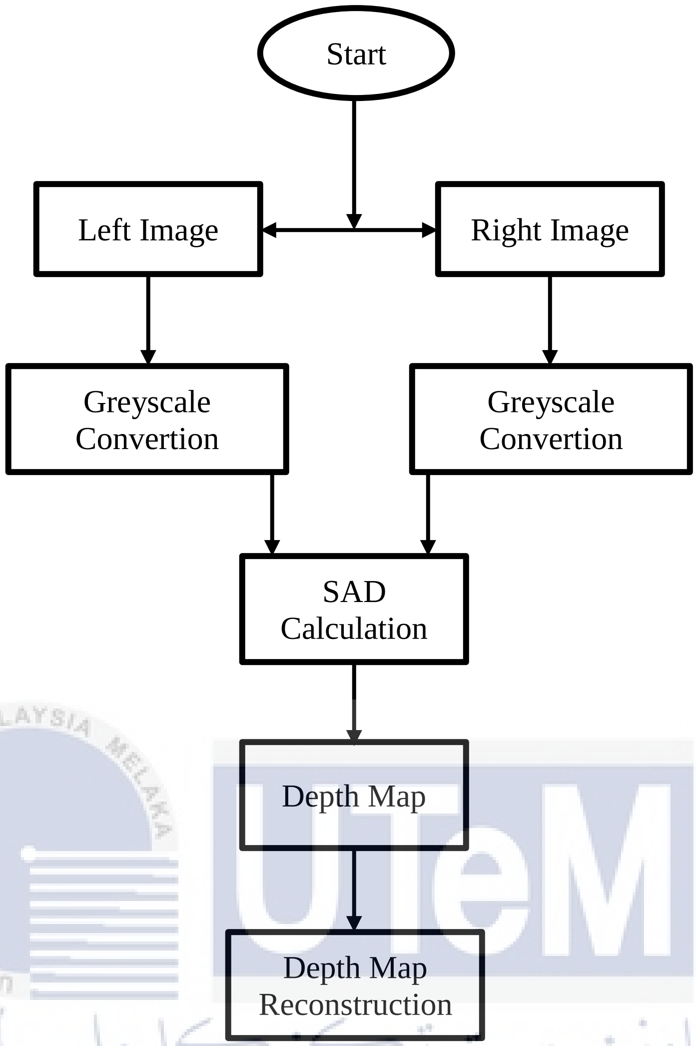


Figure 3.1 Project development flowchart.

3.3 Block Diagram of Project

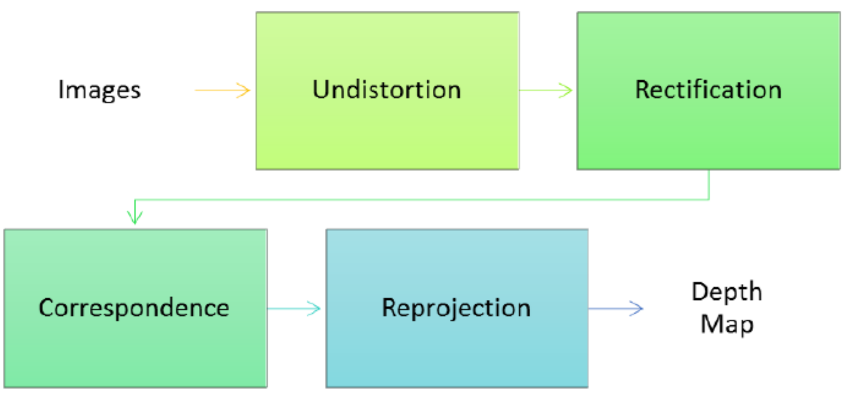


Figure 3.2 Block diagram of project.

The reconstruction of a stereo vision system is commonly mapped in four basic steps, as shown in Figure 3.2. In most cases, there are four steps. The camera lens distortion of the collected images is first eliminated from the undistortion stage. It is expected that the elevation angle and distance between two camera pictures must be changed in a future phase in order to achieve the focal length and epipolar axis.

The next step is to calculate the similarity between the left and right pictures, which will be utilized to generate the disparity map. This strategy is said to as interoperable. Finally, using the triangle property, the disparity map may be recreated, a process known as reprojection.

3.4 Program Development

The coding for program development will be based on the data from the standard measurement. The evolution of the code will be influenced by references from the internet, articles, and journals. In addition, MATLAB software will be used to create the calculation model. After that, the program will be updated to include the automated measurement coding.

3.5 Software Implementation

MathWorks created MATLAB, also known as Matrix Laboratory, which is a multi-numerical computing system and a patented programming language. This software can do a variety of mathematical activities, including matrix operations, function plotting, data implementations of algorithms, and the combination of programs written in different languages.

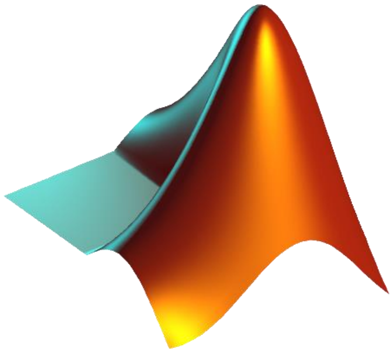


Figure 3.3 MATLAB.

This project will make use of the Image Processing Toolbox. Many features are available in this toolbox, including visualization, analysis, image processing, and algorithm building. Additionally, this toolset is capable of doing noise reduction, three-dimensional image processing, image registration, and sculptural alteration.

In addition, Image Acquisition Toolbox is included in this project. Users will be able to connect to and configure external hardware properties using this toolbox, which will provide functionalities. The toolbox may also generate code to automate the user's acquisitions.

3.6 Stereo Vison Principle Depth Map Generation

Principal directions displacement cameras are used in stereo vision to record the same view from two distinct angles on the same scene. There are several similarities and few variations between the two images captured. In human perception, the brain combines the two images to produce a 3D model of the things being observed. The 3D representation for the observed objects is generated in computer vision by discovering correlations between stereo images and analyzing them with projective geometry. Stereo reconstruction challenges are like stereo pair problems.

While moving the two images over each other, the computer compares the images to discover the corresponding bits. The magnitude of the difference between the two corresponding pixel resolutions is known to as "disparity," and it is related to the object's distance. A "disparity map" is a representation of the displacement set between the matched pixels. When picture pixels have a high disparity, the item seems closer to and sharper in the disparity map; when the disparity is low, the object appears darker.

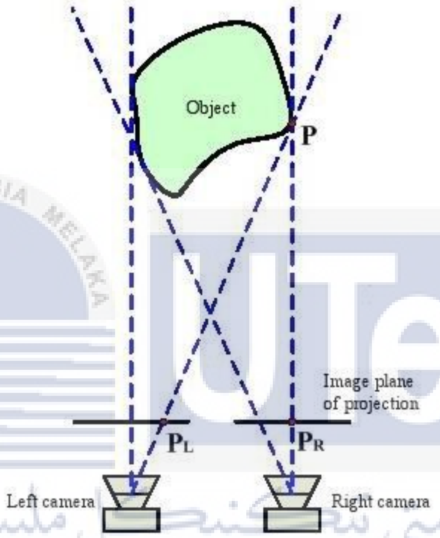


Figure 3.4 Two cameras' location and their projection planes image.

As seen in Figure 3.4, stereo pictures may be formed by combining two identical cameras. These cameras all revolve in the same direction on the same plane. To make projection modeling easier, cameras are equipped with picture planes in front of them.

Disparity is the difference between two pixels in a stereo picture pair that correspond to the same physical location. After the stereo pictures are formed, a feature correspondence method is necessary to identify the pixels that relate to the same physical location. A map of disparities is constructed for each pixel in the picture based on the correspondence feature's findings.

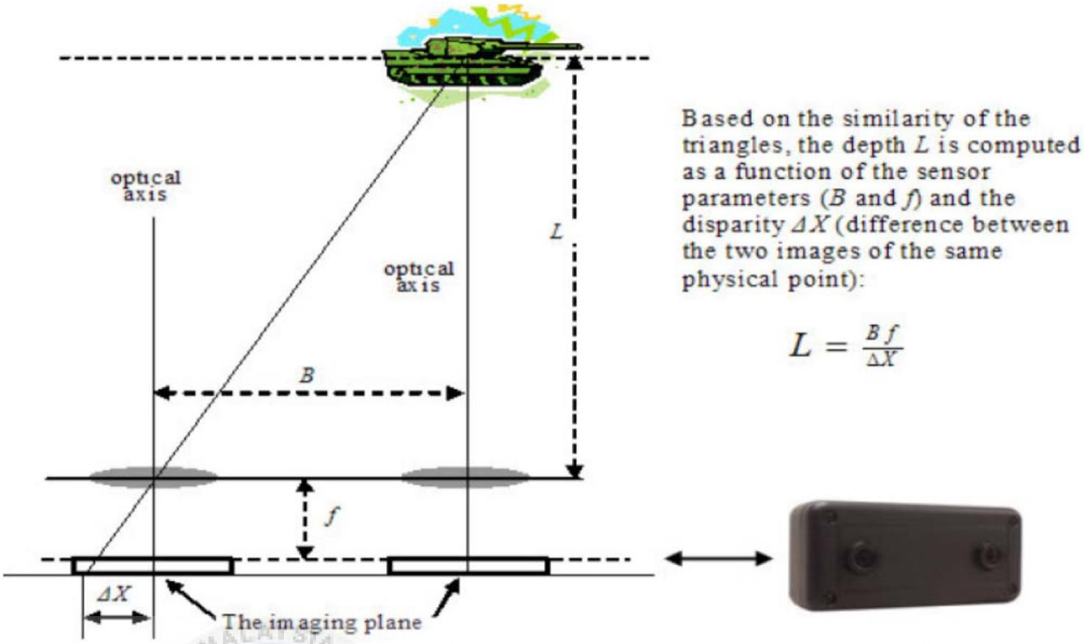


Figure 3.5 Relationship between disparity and depth.

3.7 Sum of Absolute Differences (SAD)

The map of disparities is formed by moving the source window in the left image pixel-by-pixel over the aim windows in the right image within a different search scope called disparity range and calculating (SAD) for each moving phase, with the overall SAD computed between two windows representing one's similarity. The disparity (d) of the reference image's central pixel is represented by the amount of difference between matched windows. This procedure is done for each pixel in the reference picture in order to determine their difference values. Sort the data to get a visual representation of the differences. The SAD is obtained from the following equation:

$$SAD = \sum(i,j) \in W(x,y)|Il(i,j) - Ir(i,j+d)| \quad (1.0)$$

$$SAD: \sum\big(Il(x,y), Ir(x+d,y)\big) = \sum|Il(x,y) - Ir(x+d,y)| \quad (2.0)$$

Where: W (x, y) is a window encompassing the coordinates (x, y), d is the disparity value, and Il and Ir are the left and right picture's respective intensity values. This approach was created using the MATLAB application and tested with various window widths and stereo picture pairings to demonstrate its results. (Tsukaba) is one of the Stereo Pairs with a scale of 384x288.

3.8 The Examples of SAD in MATLAB

Left Image

Right Image

True disparity map

Figure 3.6 Process Disparity Map

As shown in Figure 3.6, the matching algorithm produces a differential diagram that may be shown as an image, showing the 2D picture in 3D (left and right). The disparity is proportional to the concentration, and the target can be seen approaching the stereo cameras and emerging in a light grey hue.

3.9 Summary

This chapter presents the proposed methodology in order to develop depth map reconstruction. A flowchart of project development shows how the flow process from stereo images to depth map reconstruction.. Besides, a block diagram shows the steps or level of project development which are undistortion, rectification, correspondence and reprojection. Then, MATLAB as the platform which is can run the project development. The principles of stereo visson and Sum Absolute Difference explained in this chapter. Example of Sum Absolute Difference also given in the end of this chapter.

CHAPTER 4

RESULTS AND DISCUSSIONS

4.0 Introduction

In this chapter, it will cover about the finding initial result and the analysis from the collected data. The summary and the analysis for the project will be discussed at the last part of this chapter. The experiment are run on a Intel Core i5-5200 with system on a Nvidia Geforce 920M with 8.00 GB memory(RAM). Experiment is implemented on MATLAB R2020a.

4.1 Dataset

Stereo images carefully selected from the Middlebury dataset are test data. Such pictures, which are shown in Table 4.1, are called Teddy, Tsubaka, Cone and Venus. Ground facts for the depth of such photos are included in their metadata, and each image determines the greatest discrepancy to be searched for accuracy.

Table 4.1 Dataset from Middlebury.

	Left	Right
Teddy		
Tsubaka		
Cone		
Venus		

4.2 GUI demonstrates an efficient stereo matching algorithm.

Middlebury dataset was used to evaluate the stereo vision technology. MATLAB® R2020a was used to implement the kinematics methods. As seen in Figure 4.1, a graphical user interface is constructed in the MATLAB environment. The graphical user interface (GUI) retrieves photos from the database. The SAD algorithm generates the scene's depth map.

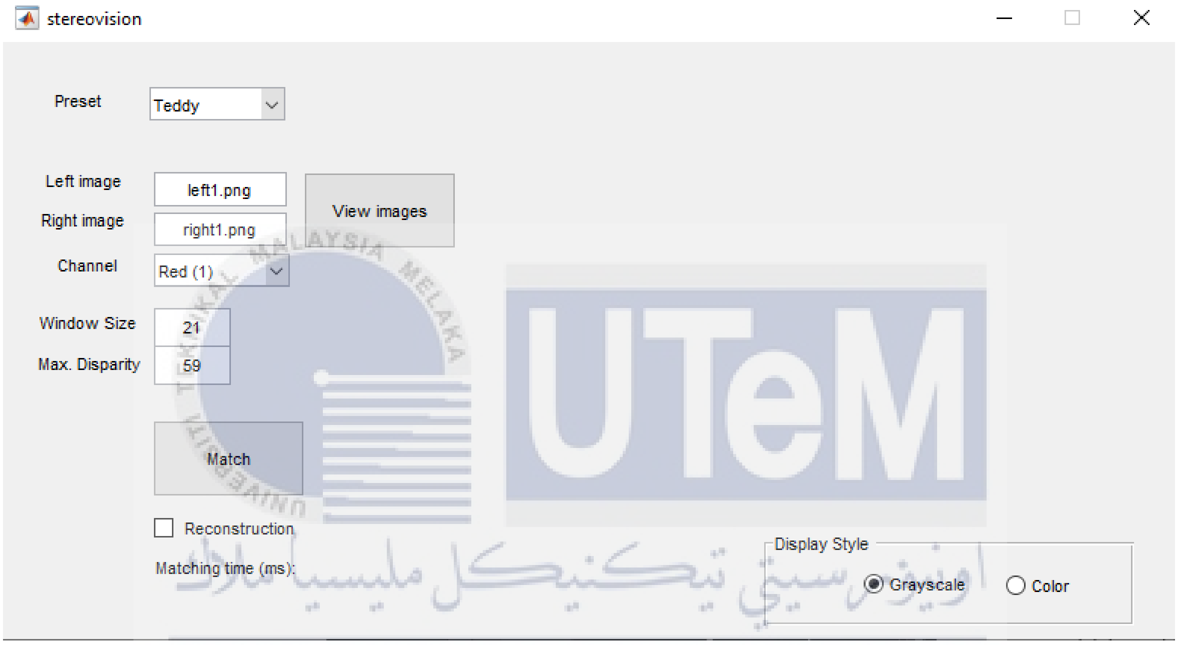


Figure 4.1 Graphical user interface for background stereo vision.

Background The outline and graphical user interface (GUI) for the depth mapping reconstruction procedure are shown in Figure 4.1. The disparity approach employs stereo imaging to identify comparable picture information around matching points in the left (reference) and right (target) stereo images. As long as sufficient and evenly dispersed points are recognized in both photos, this approach produces comparable results to target calibration.

The correspondence issue is simplified to a one-dimensional search for correspondences in the resultant corrected pictures. The relative locations of related points, whose point coordinates are specified in relation to the primary points of the left and right pictures, provide the depth information (disparity D) of the associated pixels. I designed a block matching approach that takes use of the surrounding pixels' attributes and makes use of a few statistical measures, such as the correlation C of grey values, where the corresponding pixels are determined by the greatest correlation of blocks.

At the upper left of the GUI, the preset feature allows me to pick photographs that I already stored in the database during the MATLAB coding process. The view images block to view the images we choose for example, I choose the Teddy pictures then it will show the I left and right images that I choose as show in Figure 4.3. For window size and Max disparity, the data already set at the coding in MATLAB. I already set up the data of the value of Max disparity and the window size. When we select the preset (picture), the value of the max disparity and window size will produce automatically at the GUI.

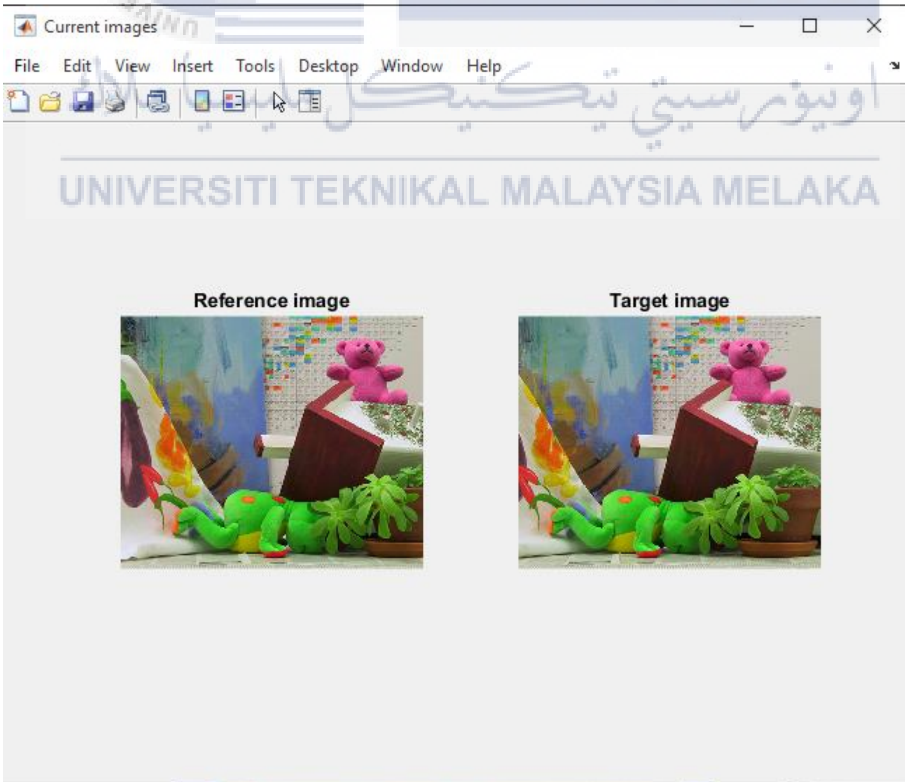


Figure 4.2 View images from the GUI.

The button *Match* at the GUI background for do the stereo matching and tick the *Reconstruction* to see the depth mapping reconstruction for the pictures was select at the stereovision GUI. For the display result for stereo matching can be choose either Greyscale in Figure 4.3 or Color in Figure 4.4 disparity display. For the disparity map reconstruction result as shown in Figure 4.5.

Figure 4.3 Grayscale stereo matching

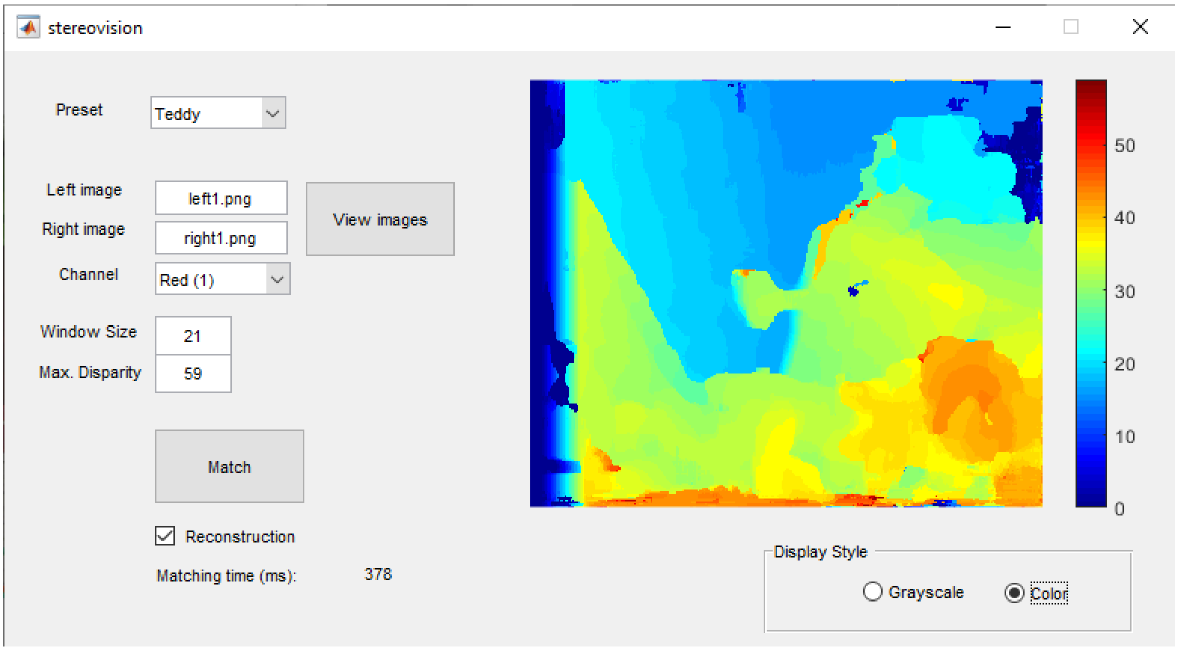


Figure 4.4 Color stereo matching

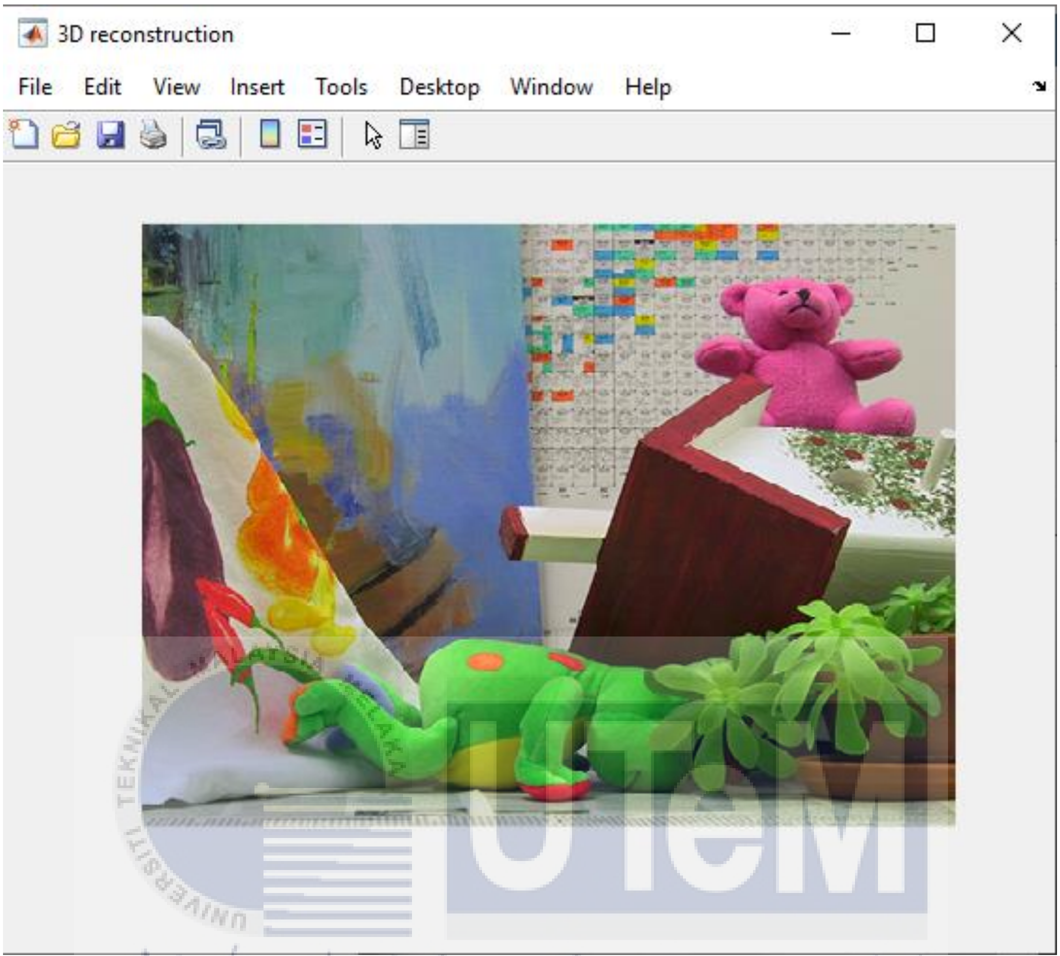


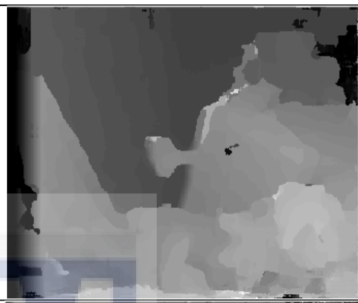
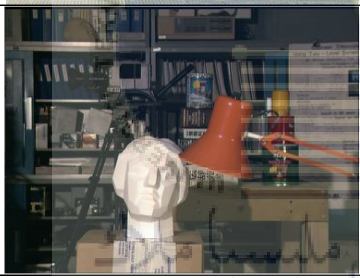
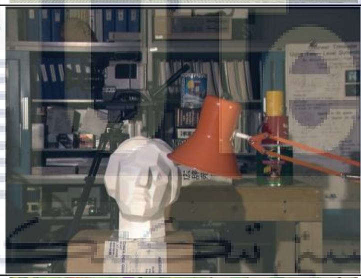
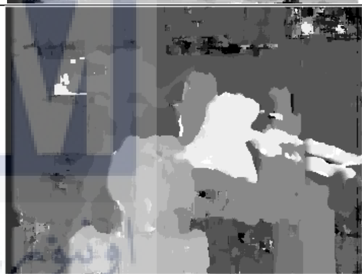
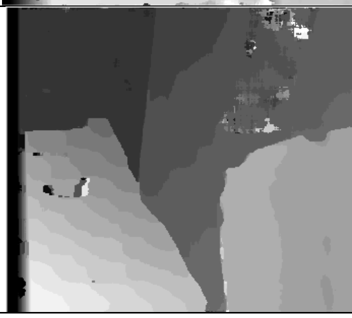
Figure 4.5 Depth-mapping reconstruction

4.3 Depth Mapping Propagation

As described in the preceding subsection, the correlation routine can be performed by calculating disparity values for each pixel on the range of rectified stereo images. As seen in Table 4.2, the disparity map is formed of the depth values calculated for each pixel in the two-dimensional picture. The darker intensities reflect items that are farther away from the camera or sensor, while the lighter intensities reflect objects that are closer to the camera or sensor.

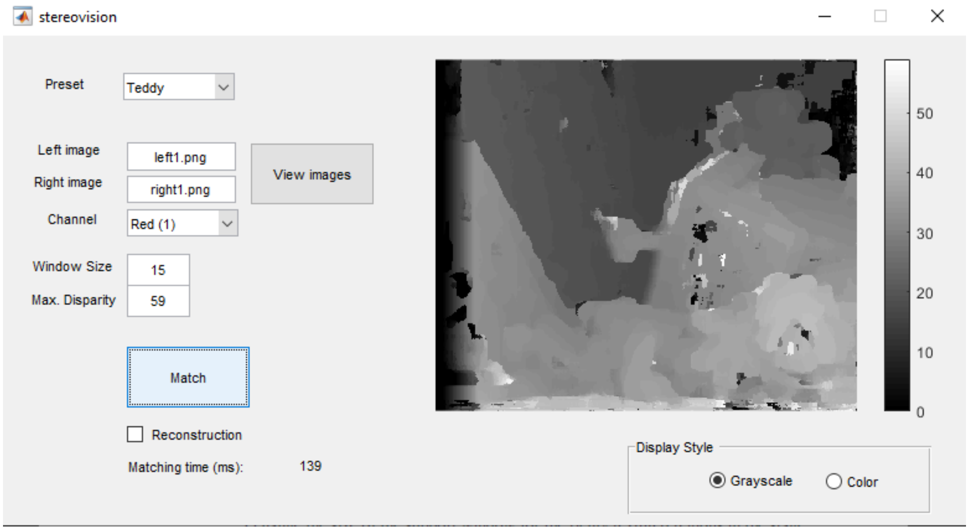
In the Tsukuba image, for example, the table lamp object is whiter or brighter than the object further away from the lamp. The Venus picture disparity map then shows that the paper is brighter than the order object and has a darker area to the left of the picture where the disparity could not be calculated. The Cones image is like the Venus image in that there are some areas with black or darker areas that cannot be calculated by the disparity. Finally, the teddy bear picture shows that the frog doll is lighter than the teddy bear. That is, the teddy bear is further away from the camera or sensor than the frog doll.

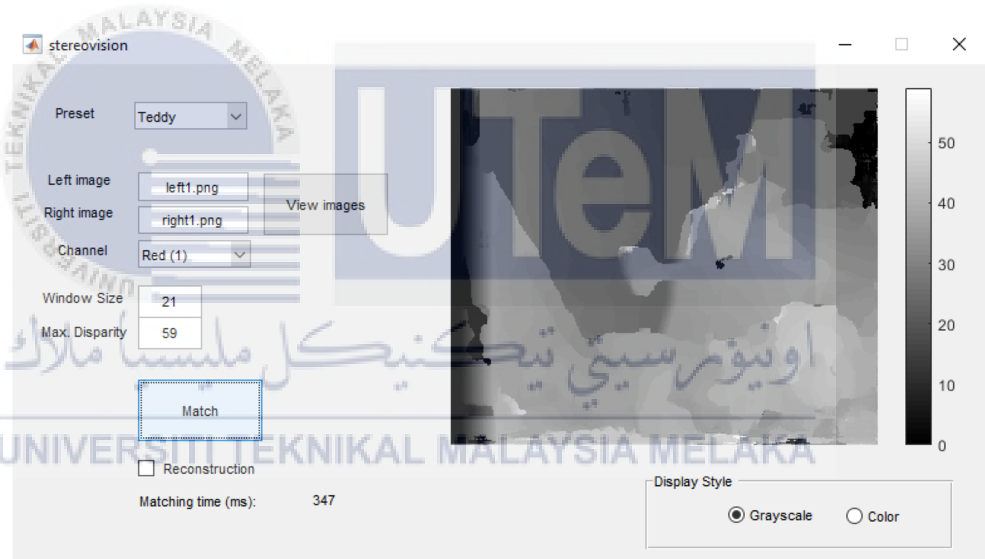
Table 4.2 Results of algorithm.

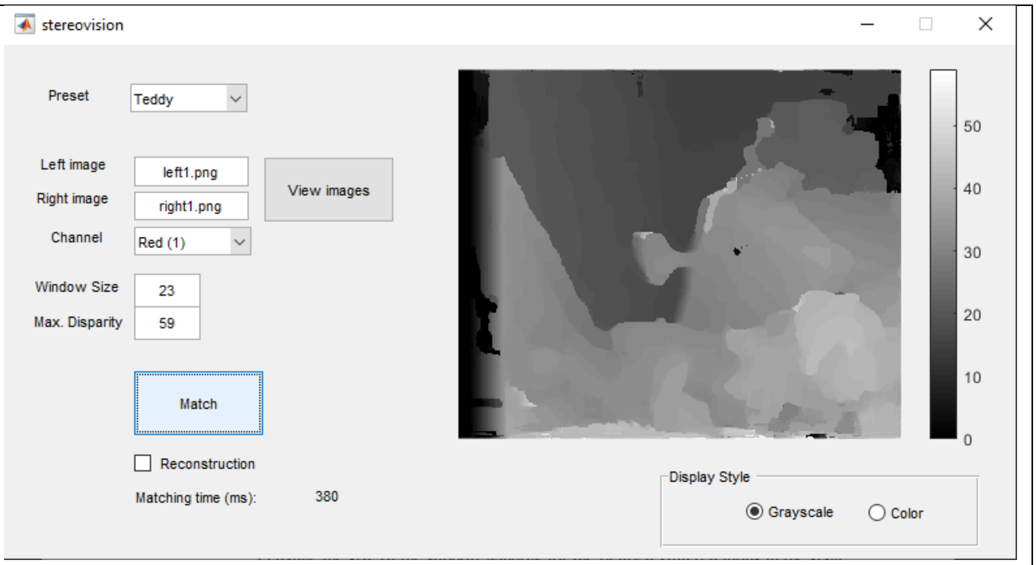
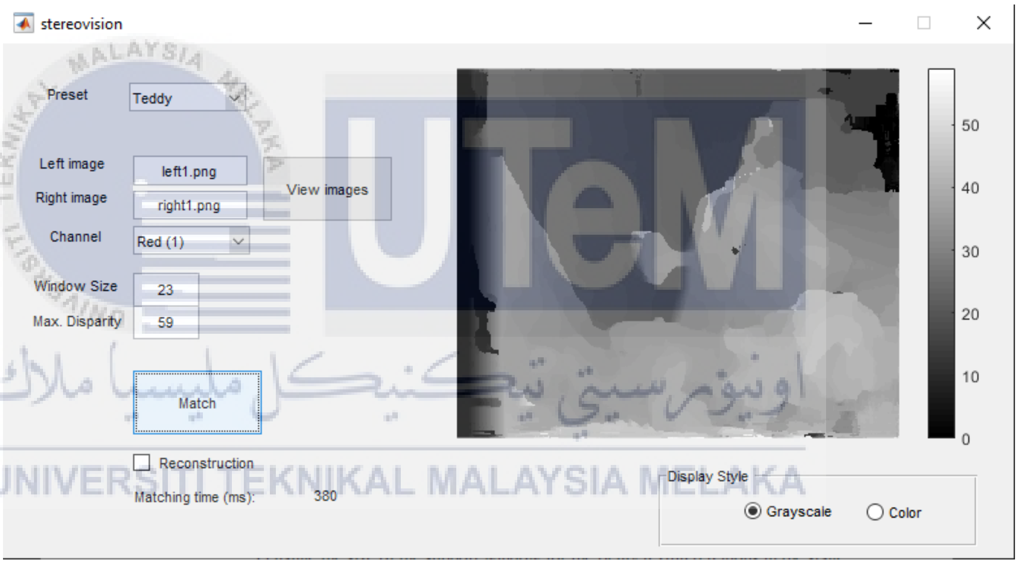
	Left	Right	Results
Teddy			
Tsubaka			
Cone			
Venus			

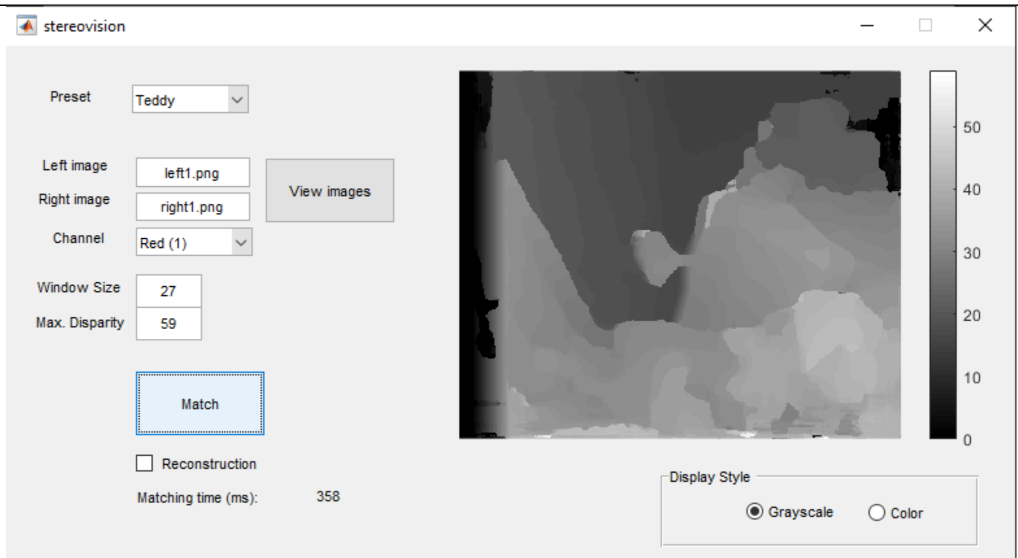
During the step of cross-based support window building, I adjust the size of the support window for the densely textured areas. I raised the window size from 15, 17, 19, 21, 23, 25, and 27 to 25, 23, 25, and 27. As seen in Table 4.3, these factors influence the tiny difference in the accuracy disparity between teddy photos.

Table 4.3 Comparison Disparity with difference window size.

Size Window	Results
15	
17	

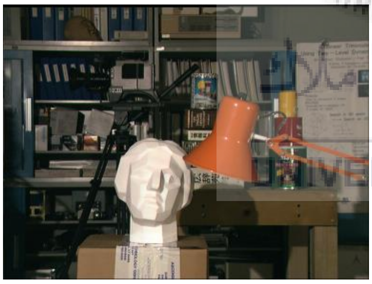
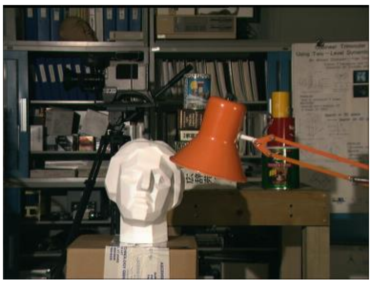
<table>
<tr>
<td>19</td>
<td>
</td>
</tr>
<tr>
<td>21</td>
<td>
</td>
</tr>
</table>

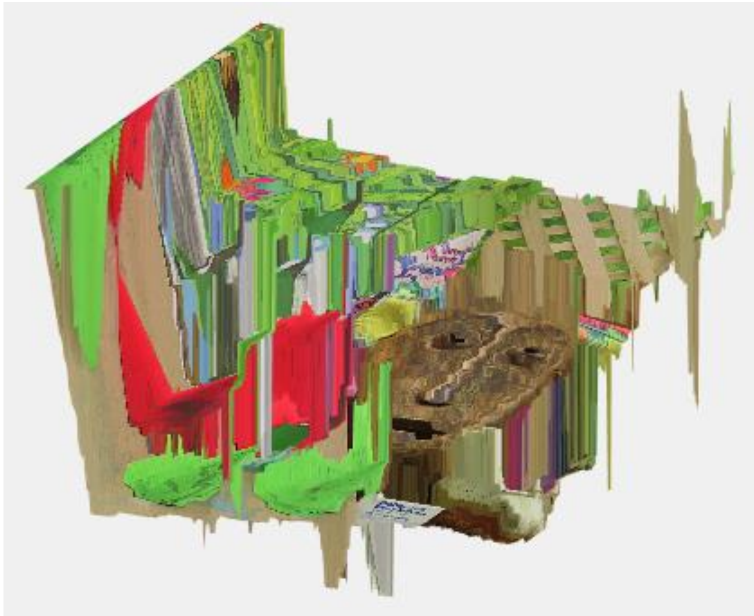
23	stereovision Preset Teddy Left image left1.png Right image right1.png View images Channel Red (1) Window Size 23 Max. Disparity 59 Match Reconstruction Matching time (ms): 380 50 40 30 20 10 0 Display Style Grayscale Color
25	stereovision Preset Teddy Left image left1.png Right image right1.png View images Channel Red (1) Window Size 23 Max. Disparity 59 Match Reconstruction Matching time (ms): 380 50 40 30 20 10 0 Display Style Grayscale Color

27	

Each pixel may be re-projected from the disparity map onto a collection of data points in the 3D point cloud coordinate system using the (x, y, z) co-ordinates denoting its position in 3D space. The 3D scenes captured from the stereo images are represented in Table 4.3. The results obtained are realistically precise, describing the scenario depicted very well and interpreting even intense viewpoints correctly. In Table 4.3 represent a couple of high objects positioned against the wall. For example, the cones image we can clearly see the distance and position of each cone, so we can know which object is closer to the camera or sensor. This same level of precision is achieved in all the image from Middlebury dataset and the results are quite satisfying.

Table 4.4 3D Reconstruction.

Stereo Image	3D Reconstruction
 	
 	




to be such a
SPORT
Venus serves


to be such a
SPORTS
Venus serves


SPORT

4.4 Summary

This chapter details the steps necessary to accomplish the project. The experiment was run using MATLAB R2020a, and this project made use of a MATLAB-created graphical user interface. The test data was carefully chosen from the Middleburry dataset, which includes Tsukuba, Cones, Venus, and Teddy. After selecting the image to be simulated, a graphical user interface (GUI) created in MATLAB will display the depth map reconstruction. The result of the disparity mapping algorithm will be displayed after selecting a picture from the GUI background. The disparities for each pixel of the corresponding intensity are revealed by a grayscale disparity map as a result of the stereo matching process. Sections that are lighter are closer to the camera, whereas areas that are darker are farther away. In the absence of dark regions, the discrepancy could not be measured. To summarise, in each of the images, MATLAB is used to locate matched points, which is greatly aided and accelerated by the epipolar constraint concepts. It keeps calculating the difference between the respective lines, giving me the depth of the object in question. This method can be repeated for each pixel in the stereo images that provide my final solution to 3D reconstruction.

CHAPTER 5

5.1 Conclusion

We provide a stereo matching technique in this work that is based on disparity propagation and the sum of absolute differences. Disparity mismatches in the original disparity map are repaired by disparity propagation from neighboring trustworthy points. This is accomplished by extracting disparity subsets for reliable pixels and constructing a new cost volume appropriately. "Depth Mapping Reconstruction from Stereo Images" is the project's purpose. "was a success when the Sum of Absolute Differences was used. The stereo matching technique was developed in MATLAB using the Sum of Absolute Differences (SAD). Then, using the Middlebury benchmark, a basic stereo algorithm achieves performance similar to the state-of-the-art. The correctness of the propagation may be ensured by using pixelwise line segments, which are both fast and accurate. This project accomplished all of my goals, and the outcomes allowed me to fulfill the benchmark. (D. Scharstein and R. Szeliski, 2015) substantiate the findings shown in Table 4.2. (XingzhengWang, 2016). Each of these algorithms is subject to a unique set of constraints. These algorithms have the potential to improve performance in several disciplines, including tracking, robot navigation, and so on. There is no greater or more significant depth cue than another. A cue offers several benefits and drawbacks.

5.2 Future Works

Some suggestions are provided as a resource for researchers who will be working on this project in the future. In the future, I hope to apply my algorithm to real-world problems, and I intend to port it to graphics hardware. After optimization and integration into a hardware framework for an indoor robot navigation system, the suggested SAD algorithm will be implemented.

REFERENCES

[1] D. K. Hoa, L. Dung, and N. T. Dzung, “Efficient determination of disparity map from stereo images with modified sum of absolute differences (SAD) algorithm,” *Int. Conf. Adv. Technol. Commun.*, pp. 657–660, 2013, doi: 10.1109/ATC.2013.6698196.

[2] Y. Zhang, W. Chen, H. Liu, J. Liu, and H. Du, “Efficient Binocular Stereo Matching Based on Sad and Improved Census Transformation,” *Proc. - Int. Conf. Mach. Learn. Cybern.*, vol. 2019-July, pp. 1–6, 2019, doi: 10.1109/ICMLC48188.2019.8949324.

[3] S. Thakkar and T. Huff, “Internet streaming SIMD extensions,” *Computer (Long. Beach. Calif).*, vol. 32, no. 12, pp. 26–34, 1999, doi: 10.1109/2.809248.

[4] R. A. Hamzah and H. Ibrahim, “Literature survey on stereo vision disparity map algorithms,” *J. Sensors*, vol. 2016, 2016, doi: 10.1155/2016/8742920.

[5] A. Qayyum, A. S. Malik, M. N. B. M. Saad, F. Abdullah, and M. Iqbal, “Disparity map estimation based on optimization algorithms using satellite stereo imagery,” *IEEE 2015 Int. Conf. Signal Image Process. Appl. ICSIPA 2015 - Proc.*, pp. 127–132, 2016, doi: 10.1109/ICSIPA.2015.7412176.

[6] A. Qayyum, I. Razzak, A. S. Malik, and S. Anwar, “Fusion of CNN and sparse representation for threat estimation near power lines and poles infrastructure using aerial stereo imagery,” *Technol. Forecast. Soc. Change*, vol. 168, no. September 2020, p. 120762, 2021, doi: 10.1016/j.techfore.2021.120762.

[7] C. H. Lee and D. Kim, “Dense disparity map-based pedestrian detection for intelligent vehicle,” *2016 IEEE Int. Conf. Intell. Transp. Eng. ICITE 2016*, pp. 108–111, 2016, doi: 10.1109/ICITE.2016.7581317.

[8] W. A. Khan, D. R. Pant, B. Adhikari, and R. Manandhar, “3D object tracking using disparity map,” *Proceeding - IEEE Int. Conf. Comput. Commun. Autom. ICCCA 2017*, vol. 2017-Janua, pp. 1015–1018, 2017, doi: 10.1109/CCAA.2017.8229982.

[9] Q. Ge and E. Lobaton, “Obstacle detection in outdoor scenes based on multi-valued stereo disparity maps,” *2017 IEEE Symp. Ser. Comput. Intell. SSCI 2017 - Proc.*, vol. 2018-Janua, pp. 1–8, 2018, doi: 10.1109/SSCI.2017.8280990.

[10] Z. Liang *et al.*, “Learning for Disparity Estimation Through Feature Constancy,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 2811–2820, 2018, doi: 10.1109/CVPR.2018.00297.

[11] G. Song, H. Zheng, Q. Wang, and Z. Su, “Training a convolutional neural network for disparity optimization in stereo matching,” *ACM Int. Conf. Proceeding Ser.*, pp. 48–52, 2017, doi: 10.1145/3155077.3155083.

[12] J. Žbontar and Y. Le Cun, “Computing the stereo matching cost with a convolutional neural network,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, no. 1, pp. 1592–1599, 2015, doi: 10.1109/CVPR.2015.7298767.

[13] W. Luo and A. G. Schwing, “(Michael recom)Efficient Deep Learning for Stereo Matching,” [Online]. Available: http://www.cs.toronto.edu/.

[14] K. Zhang, Y. Fang, D. Min, L. Sun, S. Yang, and S. Yan, “Cross-Scale Cost Aggregation for Stereo Matching,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 5, pp. 965–976, 2017, doi: 10.1109/TCSVT.2015.2513663.

[15] C. Çiǐla, “Recursive edge-aware filters for stereo matching,” *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, vol. 2015-Octob, pp. 27–34, 2015, doi: 10.1109/CVPRW.2015.7301363.

[16] M. G. Mozerov and J. Van De Weijer, “Accurate stereo matching by two-step

energy minimization," *IEEE Trans. Image Process.*, vol. 24, no. 3, pp. 1153–1163, 2015, doi: 10.1109/TIP.2015.2395820.

[17] C. Kwan, B. Chou, and B. Ayhan, "Stereo Image and Depth Map Generation for Images with Different Views and Resolutions," *2018 9th IEEE Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2018*, pp. 573–579, 2018, doi: 10.1109/UEMCON.2018.8796832.

[18] J. Žbontar and Y. Lecun, "Stereo matching by training a convolutional neural network to compare image patches," *J. Mach. Learn. Res.*, vol. 17, pp. 1–32, 2016.

[19] A. Seki, M. Pollefeys, T. Corporation, E. T. H. Zürich, and Microsoft, "SGM-Nets: Semi-global matching with neural networks," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, no. 1, pp. 6640–6649, 2017, doi: 10.1109/CVPR.2017.703.

[20] H. Xie, H. Yao, X. Sun, S. Zhou, and S. Zhang, "Pix2Vox: Context-aware 3D reconstruction from single and multi-view images," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2019-Octob, pp. 2690–2698, 2019, doi: 10.1109/ICCV.2019.00278.

[21] J. T. Barron and J. Malik, "Shape, illumination, and reflectance from shading," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 8, pp. 1670–1687, 2015, doi: 10.1109/TPAMI.2014.2377712.

[22] S. Savarese, M. Andreetto, H. Rushmeier, F. Bernardini, and P. Perona, "3D reconstruction by shadow carving: Theory and practical evaluation," *Int. J. Comput. Vis.*, vol. 71, no. 3, pp. 305–336, 2007, doi: 10.1007/s11263-006-8323-9.

[23] E. Dibra, H. Jain, C. Öztireli, R. Ziegler, and M. Gross, "Human shape from silhouettes using generative HKS descriptors and cross-Modal neural networks," *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*, vol. 2017-Janua, pp. 5504–5514, 2017, doi: 10.1109/CVPR.2017.584.

[24] S. R. Richter and S. Roth, "Discriminative shape from shading in uncalibrated illumination," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 07-12-June, pp. 1128–1136, 2015, doi: 10.1109/CVPR.2015.7298716.

[25] Y. Zhang, Z. Liu, T. Liu, B. Peng, and X. Li, "RealPoint3D: An Efficient Generation Network for 3D Object Reconstruction from a Single Image," *IEEE Access*, vol. 7, pp. 57539–57549, 2019, doi: 10.1109/ACCESS.2019.2914150.

[26] J. Wu, Y. Wang, T. Xue, X. Sun, W. T. Freeman, and J. B. Tenenbaum, "MarrNet: 3D shape reconstruction via 2.5D sketches," *Adv. Neural Inf. Process. Syst.*, vol. 2017-Decem, no. Nips, pp. 541–551, 2017.

[27] I. Goodfellow *et al.*, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020, doi: 10.1145/3422622.

[28] M. Tatarchenko, A. Dosovitskiy, and T. Brox, "Octree Generating Networks: Efficient Convolutional Architectures for High-resolution 3D Outputs," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2017-Octob, pp. 2107–2115, 2017, doi: 10.1109/ICCV.2017.230.

[29] B. Yang, S. Rosa, A. Markham, N. Trigoni, and H. Wen, "Dense 3D Object Reconstruction from a Single Depth View," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 12, pp. 2820–2834, 2019, doi: 10.1109/TPAMI.2018.2868195.

[30] M. Wang, L. Wang, and Y. Fang, "3DensiNet," no. 2, pp. 961–969, 2017, doi: 10.1145/3123266.3123340.

[31] C. Kwan, B. Chou, and B. Ayhan, "Enhancing Stereo Image Formation and Depth Map Estimation for Mastcam Images," *2018 9th IEEE Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2018*, pp. 566–572, 2018, doi: 10.1109/UEMCON.2018.8796542.

[32] O. U. Pramote and P. Piamsa-Nga, "Improve accuracy of disparity map for stereo

images using SIFT and weighted color model," *Proc. 2015-7th Int. Conf. Knowl. Smart Technol. KST 2015*, pp. 109–114, 2015, doi: 10.1109/KST.2015.7051470.

[33] O. Krutikova, A. Sisojevs, and M. Kovalovs, "Creation of a Depth Map from Stereo Images of Faces for 3D Model Reconstruction," *Procedia Comput. Sci.*, vol. 104, no. December 2016, pp. 452–459, 2016, doi: 10.1016/j.procs.2017.01.159.

[34] E. Dandil and K. K. Cevik, "Computer Vision Based Distance Measurement System using Stereo Camera View," *3rd Int. Symp. Multidiscip. Stud. Innov. Technol. ISMSIT 2019 - Proc.*, 2019, doi: 10.1109/ISMSIT.2019.8932817.

[35] V. M. Mendoza Guzmán, J. M. Mejía Muñoz, N. E. Moreno Márquez, P. I. Rodríguez Azar, and E. Santiago Ramírez, "Disparity map estimation with deep learning in stereo vision," *CEUR Workshop Proc.*, vol. 2304, pp. 27–40, 2018.

[36] B. Ayhan and C. Kwan, "Mastcam image resolution enhancement with application to disparity map generation for stereo images with different resolutions," *Sensors (Switzerland)*, vol. 19, no. 16, 2019, doi: 10.3390/s19163526.

[37] M. F. Abu Hassan, A. Hussain, M. H. Md Saad, and K. Win, "3D distance measurement accuracy on low-cost stereo camera," *Sci. Int.*, vol. 29, no. 3, pp. 599–605, 2017.

[38] J. Wu, C. Zhang, T. Xue, W. T. Freeman, and J. B. Tenenbaum, "Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling," *Adv. Neural Inf. Process. Syst.*, no. Nips, pp. 82–90, 2016.

[39] J. Wu, C. Zhang, T. Xue, W. T. Freeman, and J. B. Tenenbaum, "Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling," *Adv. Neural Inf. Process. Syst.*, no. Nips, pp. 82–90, 2016.

[40] J. Wu, C. Zhang, T. Xue, W. T. Freeman, and J. B. Tenenbaum, "Learning a probabilistic latent space of object shapes via 3D generative-adversarial modeling,"

Adv. Neural Inf. Process. Syst., no. Nips, pp. 82–90, 2016.

[41] X. Wang, Y. Tian, H. Wang, and Y. Zhang, “Stereo matching by filtering-based disparity propagation,” *PLoS One*, vol. 11, no. 9, pp. 1–8, 2016, doi: 10.1371/journal.pone.0162939.

[42] Scharstein, D., Szeliski, R. A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms. International Journal of Computer Vision 47, 7–42 (2002). https://doi.org/10.1023/A:1014573219977

[43] D. Scharstein and R. Szeliski, "High-accuracy stereo depth maps using structured light," 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings., 2003, pp. I-I, doi: 10.1109/CVPR.2003.1211354.

APPENDICES

Appendix 1 Gantt Chart Project Development PSM 1

Activity	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14
Confirm the title														
Study Journal														
Drafting literature review														
Update literature review														
Methodology														
Introduction														
Report Chapter 1-4														
Preparation slide														
Presentation BDP 1														

Appendix 2 Gantt Chart Project Development PSM 2

Activity	W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14
Study Journal														
Update Literature Review														
Study GUI														
Run Code														
Report Chapter 4-5														
Preparation slide & Poster														
Presentation BDP 1														

Appendix 3 Code of GUI

```
str = get(hObject, 'String');
val = get(hObject,'Value');
% Set current data to the selected data set.
switch str{val};
case 'Teddy'
    mystring = 'left1.png'; set(handles.edit3,'UserData',mystring); set(handles.edit3,'String',mystring);
    mystring = 'right1.png'; set(handles.edit4,'UserData',mystring); set(handles.edit4,'String',mystring);
    myval = 21; set(handles.edit1,'UserData', myval); set(handles.edit1,'String', int2str(myval));
    myval = 59; set(handles.edit2,'UserData', myval); set(handles.edit2,'String', int2str(myval));
case 'Tsukuba'
    mystring = 'left2.png'; set(handles.edit3,'UserData',mystring); set(handles.edit3,'String',mystring);
    mystring = 'right2.png'; set(handles.edit4,'UserData',mystring); set(handles.edit4,'String',mystring);
    myval = 11; set(handles.edit1,'UserData', myval); set(handles.edit1,'String', int2str(myval));
    myval = 15; set(handles.edit2,'UserData', myval); set(handles.edit2,'String', int2str(myval));
case 'Cones'
    mystring = 'left3.png'; set(handles.edit3,'UserData',mystring); set(handles.edit3,'String',mystring);
    mystring = 'right3.png'; set(handles.edit4,'UserData',mystring); set(handles.edit4,'String',mystring);
    myval = 21; set(handles.edit1,'UserData', myval); set(handles.edit1,'String', int2str(myval));
    myval = 59; set(handles.edit2,'UserData', myval); set(handles.edit2,'String', int2str(myval));
case 'Venus'
    mystring = 'left4.png'; set(handles.edit3,'UserData',mystring); set(handles.edit3,'String',mystring);
    mystring = 'right4.png'; set(handles.edit4,'UserData',mystring); set(handles.edit4,'String',mystring);
    myval = 25; set(handles.edit1,'UserData', myval); set(handles.edit1,'String', int2str(myval));
    myval = 19; set(handles.edit2,'UserData', myval); set(handles.edit2,'String', int2str(myval));
end
guidata(hObject,handles)
```



```
handles.left = get(handles.edit3,'UserData');
handles.right = get(handles.edit4,'UserData');
left=imread(handles.left); right=imread(handles.right);
figure(4), set(figure(4),'Name', 'Current images'), set(figure(4), 'NumberTitle', 'off');
subplot(121); imshow(left); title('Reference image');subplot(122); imshow(right); title('Target image')
```

Appendix 4 Code for SAD algorithm

```
function [spdmap, dcost, pcost, wcost] = stereomatch(imgleft, imgright, windowsize,
disparity, spacc)

% Set Parameters
WS  = uint16(windowsize);                % Set window size, must be uneven
WS2 = uint16( ( WS - 1 ) / 2 );          % Half window
D   = uint16(disparity)+1;               % number of disparities

% Read image sizes
heightL = uint16( size( imgleft, 1 ) );    heightR = uint16( size( imgright, 1 ) );
widthL  = uint16( size( imgleft, 2 ) );    widthR  = uint16( size( imgright, 2 ) );
if ( heightL ~= heightR  ||  widthL ~= widthR )
    error('Height and width of left and right image must be equal');
end

% Initialization
pcost = zeros( heightL, widthL, D, 'uint8' );
wcost = zeros( heightL, widthL, D, 'single' );
dmap  = zeros( heightL, widthL, 'uint8' );
dcost = zeros( heightL, widthL, 'single' );
h = zeros(WS,WS,'double'); h(1,1) = 1; h(1,WS) = -1; h(WS,1) = -1; h(WS,WS) = 1;

% Calculate pixel cost
for Dc = 1 : D
   maxL = widthL + 1 - Dc;
   pcost(:, Dc : widthL, Dc ) = imabsdiff( imgright( :, 1 : maxL), imgleft( :, Dc :
widthL) );
end

% Calculate integral cost
icost = single(pcost);
icost = cumsum( cumsum( icost ), 2 );

% Calculate window cost
wcost = imfilter(icost,h,'same','symmetric');

% Search disparity value
[ dcost(:,D+WS2:widthL), dmap(:,D+WS2:widthL)] = min( wcost(:,D+WS2:widthL,:),[], 3
);
for j=WS2+1:D+WS2
    [ dcost(:,j), dmap(:,j)] = min( wcost(:, j, 1 : (j - WS2) ),[], 3 );
end

% Adjust disparity map
warning off;
spdmap = single(dmap-1);

% Subpixel interpolation

if spacc==1
for j=D+1:widthL
for i=1:heightL
if dmap(i,j)>1 && dmap(i,j)<D
p = polyfit2((single(dmap(i,j)-2:dmap(i,j))),shiftdim(single(wcost(i,j,dmap(i,j)-
1:dmap(i,j)+1)),1),2);
```

```
temp=roots(p);
spdmap(i,j)=real(temp(1));
end
end
end
end
warning on;
```