

**3D Reconstruction for Simultaneous Localisation and Mapping  
System (SLAM) in Robotics using Webcam Images**

by

Benny Sia Wan Fung

Dissertation submitted in partial fulfilment of

the requirements for the

Bachelor of Engineering (Hons)

(Electrical and Electronics Engineering)

JUNE 2010

Universiti Teknologi PETRONAS

Bandar Seri Iskandar

31750 Tronoh

Perak Darul Ridzuan

# **CERTIFICATION OF APPROVAL**

3D Reconstruction for Simultaneous Localisation and Mapping System in  
Robotics Using Webcam Images

by

Benny Sia Wan Fung

A project interim submitted to the  
Electrical and Electronics Programme  
Universiti Teknologi PETRONAS  
in partial fulfilment of the requirement for the  
BACHELOR OF ENGINEERING (Hons)  
(ELECTRICAL AND ELECTRONICS ENGINEERING)

Approved by,

---

(Dr. Mohamad Naufal bin Mohamad Saad)

UNIVERSITI TEKNOLOGI PETRONAS  
TRONOH, PERAK

June 2010

## **CERTIFICATION OF ORIGINALITY**

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

---

BENNY SIA WAN FUNG

## **ACKNOWLEDGEMENT**

First and foremost, I would like to give praise and gratitude to God for His blessings and strength throughout this final year project. Also, I would also like to sincerely thank Dr.Naufal as the most dedicated supervisor, constantly providing me guidance and assistance, teaching and monitoring me throughout my entire final year project. In addition to that, my appreciation and gratitude is extended to Mr. Fabrice Meriaudeau and LE2I as the collaborator of this project. Not forgetting all my colleagues for always supporting me in various ways. My appreciation would be incomplete without giving credit to UTP in preparing me for this challenge by equipping me with the foundations and essential skills for self-learning. Also, sincerely thanking Electrical and Electronics Engineering Department's lecturers and technicians constantly being there to provide support and assistance whenever needed. Last but not least, a sincere gratitude to those who was involved directly or indirectly throughout the course of this program. Thank you.

## **ABSTRACT**

The purpose of this research is to study on 3D reconstruction using webcam images in indoor environment. The research is focused on computer vision and image processing for robotics. It aims to design a simple but efficient 3D reconstruction technique which could be used as part of the Simultaneous Localization and Mapping (SLAM) system in robotics and autonomous vehicles. Among the various sensors that can be used to conduct 3D reconstruction, this project focuses solely on using webcam as an imaging sensor. This is due to the fact that it is widely used and easily operable. However, it has its own limitations including low resolution and restricted scope of view. Simultaneously, at the end of this study, it is planned to design the software architecture for a fully automated 3D reconstruction system which can be plug and play for designated webcams. In this study, extensive research will be done followed by simulation and testing for each algorithm. At the end of the study, a comparison of results will be done to obtain the best fitted 3D reconstruction model to suit our purpose.

## TABLE OF CONTENTS

<b>CERTIFICATION OF APPROVAL</b> .....	ii
<b>CERTIFICATION OF ORIGINALITY</b> .....	iii
<b>ACKNOWLEDGEMENT</b> .....	iv
<b>ABSTRACT</b> .....	v
<b>LIST OF FIGURES</b> .....	viii
<b>CHAPTER 1 INTRODUCTION</b> .....	<b>1</b>
1.1 Background of Study .....	1
1.2 Problem Statement .....	2
1.3 Objectives.....	2
1.4 Scope of Study .....	3
<b>CHAPTER 2 LITERATURE REVIEW</b> .....	<b>4</b>
2.1 Epipolar Geometry .....	4
2.2 Feature Point Matching .....	6
2.2.1 <i>Feature Point Detection</i> .....	6
2.2.2 <i>Information Extraction</i> .....	7
2.2.3 <i>Points Matching</i> .....	8
2.3 Correlation Matching .....	9
<b>CHAPTER 3 METHODOLOGY</b> .....	<b>10</b>
3.1 Procedure Identification .....	10
3.2 Tools and Equipments.....	11
3.3 Project Timeline .....	11

<b>CHAPTER 4 RESULT AND DISCUSSION .....</b>	<b>12</b>
4.1 System Algorithm .....	12
4.2 Susan Detector Result .....	13
4.3 Feature Point Matching Result.....	15
4.4 Reconstruction of 3D Points .....	16
4.4.1 Camera Calibration .....	16
4.4.2 Fundamental Matrix Estimation .....	17
4.4.3 Essential Matrix Estimation.....	17
4.4.4 Estimating the Rotation and Translation Matrix .....	18
4.4.5 Triangulation Method .....	20
4.5 User Interface .....	24
 <b>CHAPTER 5 CONCLUSION AND RECOMMENDATION.....</b>	 <b>25</b>
5.1 Conclusion .....	25
5.2 Recommendation.....	25
 <b>REFERENCES.....</b>	 <b>26</b>
<b>APPENDICES .....</b>	<b>28</b>
APPENDIX I: GANTT CHART .....	29

## LIST OF FIGURES

Figure 1 : Epipolar Geometry [5].....	4
Figure 2 : Two circular mask on a simple image.....	6
Figure 3 : Project flow diagram .....	10
Figure 4 : Illustration of System Algorithm.....	12
Figure 5: Corners marked in red detected using Susan Detector.....	13
Figure 6: Before smoothing .....	14
Figure 7: After smoothing.....	14
Figure 8: Feature Points Matching Result.....	15
Figure 9: Points matching in images with maximal number of details.....	15
Figure 10: 21 Calibration Images.....	16
Figure 11: Four different possible perspectives [10] .....	18
Figure 12: Two images with 18 matched points marked in red.....	22
Figure 13: 3D Reconstructed Points and Camera Locations .....	23
Figure 14: User Interface .....	24



# **CHAPTER 1**

## **INTRODUCTION**

### **1.1 Background of Study**

Throughout the past decades, the need for automated systems has increased drastically. One of the fields that are constantly developing includes robotics and automated vehicles. In robotics and automated vehicles, 3D reconstruction is the core for robotics localization technology which provides the environment modeling for navigation and obstacle avoidance.

Hence, 3D reconstruction has been a common research area in computer vision following the advancements in robotics. There are a lot of papers written on this research subject, including several great introductions that can be studied and easily understood [1-2]. However, the problems of 3D reconstruction with images are still not completely solved until today.

Also, since many of the smaller robotic units use webcams to navigate around in indoor environment since it's fairly affordable and easy to install, this study attempts to address the problems with 3D reconstruction using a fairly low end image sensor which is the webcam.

## **1.2 Problem Statement**

Since most research on 3D reconstruction is very specific and specially tailored to a custom requirement or specific robots, many of the techniques used are not applicable under other circumstances [3-4]. Hence, this research attempts to design a technique that fits a very user friendly imaging sensor which is the webcam.

However, due to noise and error of the sensor readings, it is impossible to obtain a perfect estimation of a robots pose and position. Also, if the image is captured while the robot is moving, it is almost certain that the estimation is error prone. Moreover, this image processing has to take place in real time, time consumption is a very important factor.

## **1.3 Objectives**

Upon completing the project, several objectives need to be achieved. The objectives of study are as follows:

- To technically study the basic concept of 3D reconstruction technique used in robotics and automated vehicles.
- To design an effective 3D reconstruction technique for unknown indoor environment using webcam images.
- To create a user interface to simulate the selected 3D reconstruction technique.

## **1.4 Scope of Study**

In order to narrow down the area of research, the scope of study is carefully tailored to focus on 3D reconstruction of unknown indoor environments. In indoor environments, it is assumed that the surroundings are sufficiently bright with a constant lighting. Also, the assumption that no other moving objects are in the environment was made.

Moreover, the sensor that is being used is webcam which can be calibrated. This is to ensure that the intrinsic parameters of the camera is known or can be obtained by simple calibration method. Once the camera is calibrated, we assume that all intrinsic parameters remain constant throughout the experiments and studies.

Finally, this research on 3D reconstruction is tailored for robotics navigation and object avoidance only, allowing a wider margin of error for the precision of the reconstructed scenes.

# CHAPTER 2

## LITERATURE REVIEW

### 2.1 Epipolar Geometry

To understand the basics of stereovision and 3D reconstruction, it is important to first study on the epipolar geometry [5]. Hence, a brief review of the epipolar geometry is covered in this section.

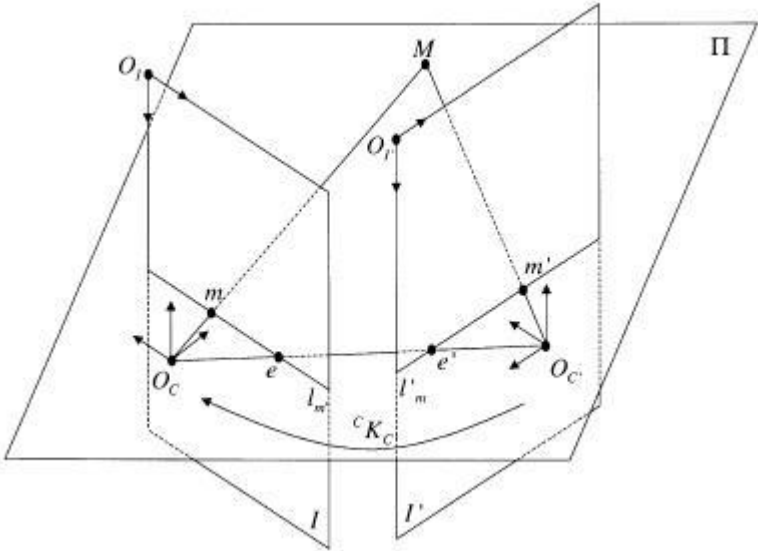


Figure 1 : Epipolar Geometry [5]

As described in Figure 1, any 3D object point  $M = [X, Y, Z, 1]$  expressed with respect to a world coordinate system  $\{W\}$ , it can be related to the 2D projection on an image plane in pixels  $m = [x, y, 1]$  through the following equation,

$$sm = PM \tag{1}$$

where  $s$  is a scale factor and  $P$  is a  $3 \times 4$  matrix that is composed of the intrinsic parameter,  $K$ , and the extrinsic parameters  $D$  as follows,

$$P = KD \tag{2}$$

Note that  $K$  relates the optical and internal geometry of the camera while  $D$  relates both the camera coordinate system and the world coordinate system.

The parameter  $D$  can be further decomposed into a 3x3 rotation matrix,  $R$ ; and a 3x1 translation matrix,  $t$ .

$$D = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \quad (3)$$

With this, the epipolar geometry describes the geometry between two different positions of a mobile camera. The object point  $M$  and its 2D projections  $m$  and  $m'$  on both image planes forms a new plane. This plane will intersect both image planes, resulting in two epipolar lines  $l_m$  and  $l_{m'}$  respectively.

All of these epipolar geometry can be expressed in a single matrix, called the Fundamental matrix,  $F$  relating any two image points corresponding to an object point in space.

$$m^T F m' = 0 \quad (4)$$

Also, when the intrinsic parameters of the camera(s) are known, we can simplify the above equation to

$$Q^T E q' = 0 \quad (5)$$

where  $q = K_c^{-T} m$ ,

$q' = K_c^{-1} m'$ , and

$E = [t R]$ .

This gives very crucial information, when the Fundamental matrix of a calibrated camera is known; the Essential Matrix,  $E$  which describes the displacement of the camera can be obtained.

## 2.2 Feature Point Matching

Feature point matching technique is deemed suitable in this scenario since it requires minimal computational power and hence increases the speed of computation as well. This method is done by taking several images of the same scene from different angle.

Then, an attempt to identify certain feature point is made. This feature points are selected as points of interest that could be easily seen from different angles by the moving robot. Finally these feature points are being matched with its corresponding feature points from several other images. Next, the epipolar geometry is used to calculate the camera's relative position and reconstruct its surrounding.

### 2.2.1 Feature Point Detection

To begin with, feature points are defined as points/pixels which contain unique traits in an image. In this project, feature points are chosen to be "corner points". To identify these corner points, the Susan Detector [6] was tested. The algorithm is explained as follows.

Figure 2 shows that for a pixel in question (nucleus), a circular mask is used. For every pixel within the mask, the pixel's brightness (its gray level) is compared with the nucleus's brightness. If the difference in brightness falls within a defined threshold,  $t_1$  sets at 90, it is counted, and otherwise it is ignored.

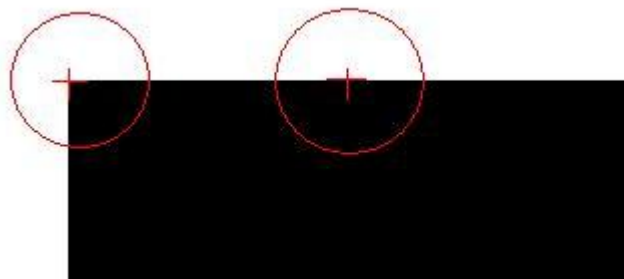


Figure 2 : Two circular mask on a simple image

The total count will give an area for which the pixels are having about the same brightness as the nucleus, called the USAN area [6]. If we divide the USAN area with the mask size in use, the ratio of salient degree ( $sd$ ) from 0 to 1 which defines the USAN area will be obtained. This information is then assigned to the nucleus.

After scanning through the whole image, the resulting image will show the salient degree of each point. Theoretically, an edge gives a salient degree of 0.5 while a 90 degree corner gives a salient degree of 0.25. With this in knowledge, the corner points can be extracted from others by using this geometrical threshold,  $t_2$  sets at 0.4.

Now with the feature points obtained, focus will be to determine its characteristics that are of interest and store them for matching purposes in later stages.

### *2.2.2 Information Extraction*

From each feature point detected in an image, a set of data is being extracted and stored in a set of tables. These data, includes Gray level, Salient Degree and Direction of Corners. The gray level and salient degree of a pixel is as explained before, while this section focuses solely on the Direction of Corners [6].

Considering a USAN area for a particular pixel (nucleus), the direction of corner is computed from the vector from nucleus to the center of gravity of the USAN area. To compute the center of gravity, there are two distinct cases which require two different computation methods:

a) Case 1:

*The USAN area is not symmetrical in the mask, hence giving a centre of gravity which is located away from the nucleus. This is generally the case for corners and edges in most cases.*

*Dealing with this, we obtain a sum of vector from nucleus to each pixel in the USAN area and divide it with the total sum of pixels accounted for, which gives us the direction of corner instantly.*

b) Case 2:

*The USAN area is symmetrical in the mask, giving a centre of gravity which is located relatively near to the nucleus itself. This is normally generated from noise pixels or a thin white line cutting across the mask through the nucleus.*

*To deal with this, compute the moment from the USAN area, and then compute its direction. This involves a rather complicated algorithm.*

### *2.2.3 Points Matching*

Now with these information extracted from points in two images, comparison of these information can be done to match the points as long as all these information do not differ more than allowed by the threshold set. The threshold is set respectively:

- a) Gray level :  $|G_1 - G_2| \leq t_1,$
- b) Salient Degree :  $|SD_1 - SD_2| \leq t_2,$
- c) Direction of Corner :  $|DC_1 - DC_2| \leq t_3.$

With this, points from two images of the same object from different viewpoints can be matched. A test is done on this method since it is the best fitted method, details of the result will be discussed in the following chapters.



### **2.3 Correlation Matching**

The next method studied is the conventional method of conducting a correlation matching [7] from several images of the same scene taken from multiple angles to obtain the matched sections from images. Using correlation, it is possible to obtain corresponding points between these images by doing a pixel to pixel check. Knowing the epipolar geometry, an estimation of the relative position of the camera and its surrounding can be obtained.

Since correlation method picks up any points that match regardless whether the point is actually of interest it gives a denser mapping compared to feature point detection. While it could be an advantage, it is also a drawback from computational power's point of view. The more points being needed to be analysed, the more complicated the algorithm would be and this would prolong the process time which is critical for robotics application.

## CHAPTER 3 METHODOLOGY

### 3.1 Procedure Identification

The flow of the project is being explained in the flow chart in Figure 3.

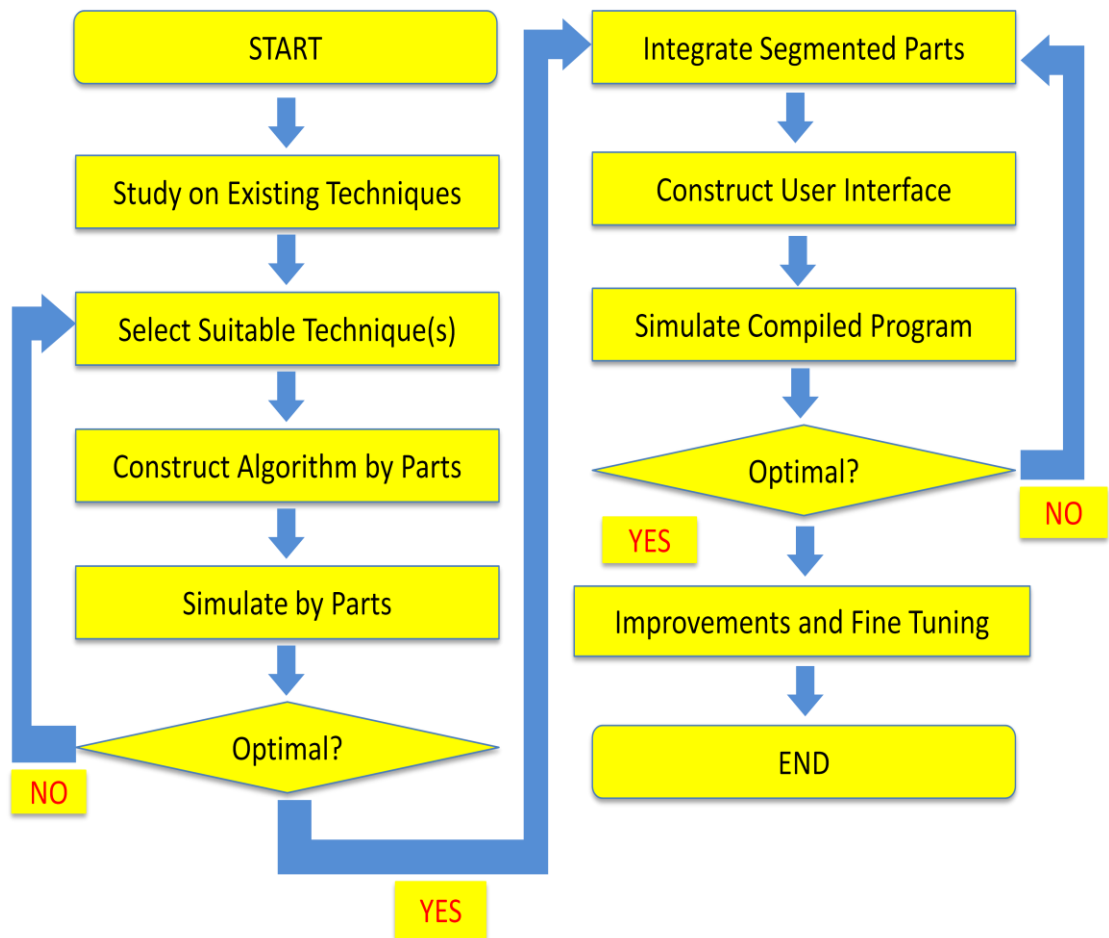


Figure 3 : Project flow diagram

### **3.2 Tools and Equipments**

There are several tools and software required in order to implement this project which are as follow:

- Webcam - used as image sensor (input) for the project.
- Visual Studio 2005 Software – Used to create application and integrates the webcam to the laptop.
- Matlab 2003 Software – Used for certain image processing functions and to plot 3D points for simulations.

### **3.3 Project Timeline**

The project is being planned and a timeline is plotted using Gantt Chart (Appendix I). Following the Gantt chart, the project will be ensured to progress smoothly according to the assigned timeline.

## CHAPTER 4

### RESULT AND DISCUSSION

#### 4.1 System Algorithm

The system can be broken down into several steps where each step has to be fulfilled before the next step can be taken in order to reach the final goal of having 3D points reconstructed from several 2D images. These steps are being illustrated in Figure 4.

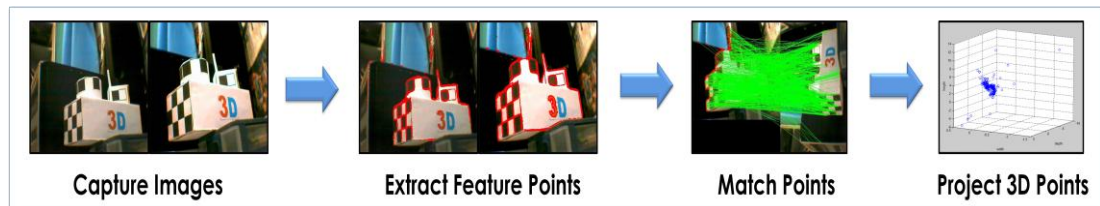


Figure 4 : Illustration of System Algorithm

First of all, images are captured from 2 different perspectives, and then the points of interest will be extracted from both of these images. Next, image point pairs are being matched in order to re-project the 3D points to obtain its coordinates respectively.

The results of these steps will be further discussed in the following sections of this report.

## 4.2 Susan Detector Result

The following is a result of the feature point extraction method. Using the Susan detector, the corner points of an image captured by the webcam is being processed and marked in red as shown in Figure 5.



Figure 5: Corners marked in red detected using Susan Detector

However, the main variable affecting this algorithm is its mask size. If a small mask (e.g. 7x7pixels) is used, it detects almost all edges rather than just the corner points. If a bigger mask (e.g. 41x41pixels) is used, it works perfectly identifying most of the corner points with minimal faulty points, but with the drawback of processing time which is extremely slow. This causes a dilemma of speed versus accuracy as always. To solve this, a looping which scans through the image 1<sup>st</sup> time using a small mask size of 7x7 is done.

Then, at all the points in interest as marked by the 1<sup>st</sup> scan, a second scan is done with a bigger mask size of 41x41px. This reduces the scanning for 52000px (280x200px) with a mask size of 41x41px to an average of less than 10000px only. This gives a 5 times increase in speed.

Also, for better accuracy, several improvements were made. First of all, for calculating the USAN area, a smoothing effect rather than a step effect can be obtained by a simple threshold method. Hence, the equation for USAN area is modified to be an exponential function [6].

This will give a smoother response due to its exponential nature. Now, the USAN area will not be given in discrete numbers but rather it will be changed to a floating point with accuracy up to 6 decimal points.

With this accuracy, the geometric thresholding window can be narrowed down from 0.4 down to 0.255 to obtain relatively accurate corner points and reducing multiple repetitions over the same corner. The effect of smoothing is shown in Figure 6 and Figure 7.



Figure 6: Before smoothing    Figure 7: After smoothing

The elimination of multiple repetitions over the same point is highly desirable for later process of image matching. Only a single point for each corner is needed to be matched to increase its accuracy. With this feature points and its information that characterize each corner points, the information stored is being compared and points are matched from two images.

### 4.3 Feature Point Matching Result

Figure 8 shows the result of the feature point matching algorithm. The points that are matched are linked with a green indication line for easier cross reference.

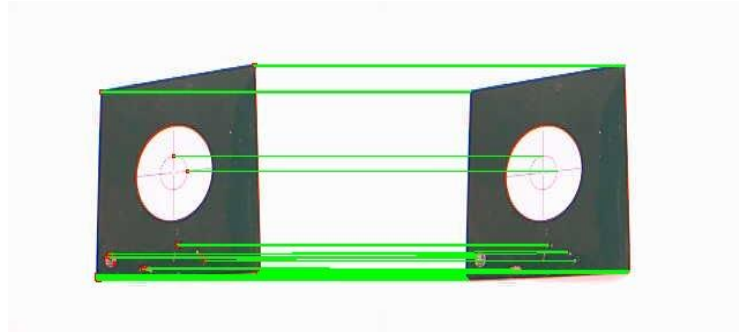


Figure 8: Feature Points Matching Result

Figure 9 shows the results of a further test which is conducted to obtain a match in images that consist of more details.

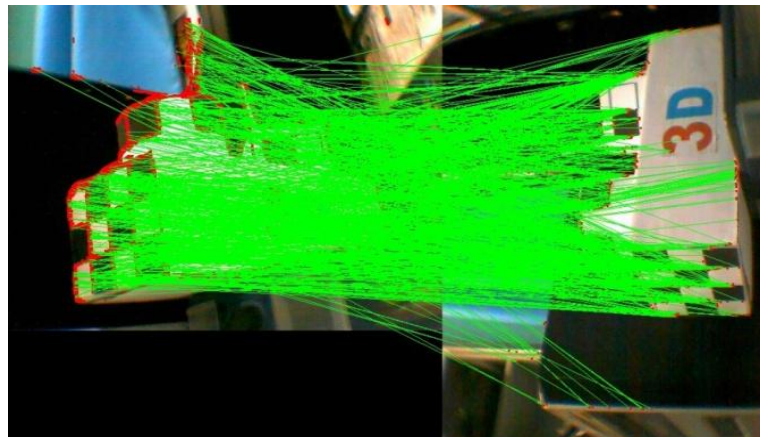


Figure 9: Points matching in images with maximal number of details

In a more complicated scene as shown in Figure 9, the results show that the matching is very accurate. One of the weak points is the number of points being detected is quite scarce, this might be a major problem in later stages when reconstruction of these points is done. This will be tested and rectified when the whole system is being integrated.

## 4.4 Reconstruction of 3D Points

### 4.4.1 Camera Calibration

Firstly, calibration is conducted to obtain the Intrinsic Parameters,  $K$ . This is done by using Joaquim Salvi's Matlab Fundamental Matrix Estimation ToolBox [8].

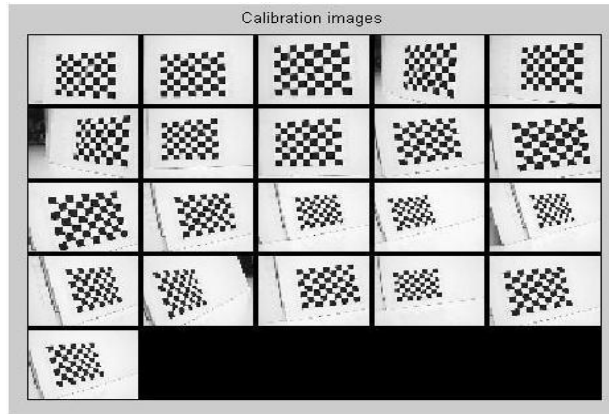


Figure 10: 21 Calibration Images

After fixing the manual zoom range of the webcam, 21 images of the calibration pattern were captured as shown in Figure 10. These images were fed into the system giving us the calibration data as follows.

$$\textbf{Focal Length: } \quad \textbf{fc} = [ 513.46484 \quad 515.04309 ] \pm [ 5.19743 \quad 4.72052 ]$$

$$\textbf{Principal point: } \quad \textbf{cc} = [ 180.04619 \quad 127.87672 ] \pm [ 5.93725 \quad 5.04804 ]$$

These data is then fit into the intrinsic matrix,  $K$  as shown.

$$K = \begin{bmatrix} 513.4648 & 0 & 180.04619 \\ 0 & 515.043 & 127.87672 \\ 0 & 0 & 1 \end{bmatrix}$$



#### 4.4.2 Fundamental Matrix Estimation

The fundamental matrix,  $F$  is estimated using the pre-installed library, *OpenCV tool*. Feeding in all the point coordinates for every pair of matched points obtained using the feature point matching algorithm as explained, `cvFindFundamentalMatrix` will provide an output of a 3x3 matrix.

However, it is important to ensure that it only output one matrix since it is possible to produce from zero output up to 3 fundamental matrices. This is due to the amount of inaccuracy matching which is fed in as input.

#### 4.4.3 Essential Matrix Estimation

With  $K$  and  $F$  obtained, we are able to obtain  $E$  using the following equation,

$$E = K^T F K \quad (6)$$

This step is rather straightforward. The essential matrix contains information on the intrinsic parameters of the camera as well as the epipolar geometry of the system in question.

#### 4.4.4 Estimating the Rotation and Translation Matrix

Having the essential matrix, computation of the rotation matrix,  $R$  and translation matrix,  $t$  can be done [9]. First, determine the Singular Value Decomposition of  $E$ , where

$$E = USV^T \quad (7)$$

From there, obtain the Rotation Matrix,  $R$  as follows

$$R = UWV^T \text{ or } UW^T V^T \quad (8)$$

And the translation matrix is  $t = u_3$  or  $-u_3$

where  $u_3$  is the last column of  $U$ , and  $W = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ .

With this method, instead of one solution, four different sets of solutions for  $R$  and  $t$  is possible ( $P_A, P_B, P_C$  and  $P_D$ ) resulting from the possibility of four different views as shown in Figure 11.

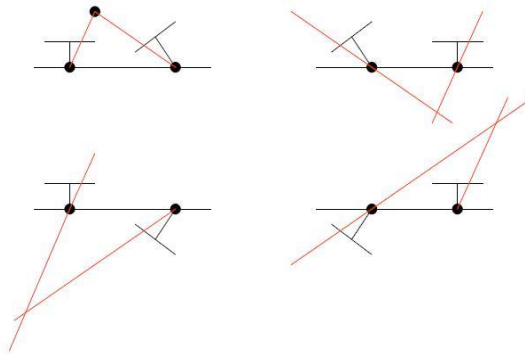


Figure 11: Four different possible perspectives [10]

However, this can be solved by using back projection technique. There are a set of rules as detailed by David Nister in his research paper [11]. Assuming  $P_A$  gives one configuration of the  $R$  and  $t$ , solve for the true configuration by just using one point.

This point is triangulated using ( $[I | 0], P_A$ ) to obtain the space point  $X$ . The method of triangulation will be detailed in the next subsection. With this point, a cheirality test is conducted following the conditions below.

Let  $c_1 = X_3X_4$  and  $c_{21} = (P_AX)_3X_4$ .

- a) If  $c_1 < 0$ , the point is behind the first camera. If  $c_2 < 0$ , the point is behind the second camera.
- b) If  $c_1$  and  $c_2$  is greater than 0, then  $P_A$  and  $X$  is the true configuration.
- c) If  $c_1 < 0$  and  $c_2 < 0$ , reflection  $Hr = \text{diag}(1, 1, 1, -1)$  is applied and we obtain  $P_B$ .
- d) If  $c_1c_2 < 0$ , the twist  $H_t$  is applied to obtain  $P_C$  and the point  $H_tX$ .

If  $X_3(H_tX)_4 > 0$ ,  $P_c$  and  $H_tX$  is the true configuration. Else, reflection  $H_r$  is applied and we get  $P_D$ .

#### 4.4.5 Triangulation Method

First of all, there are two parts where triangulation is needed; the first part is to triangulate to obtain the real  $R$  and  $t$  and the second part is to triangulate to obtain the real 3D point coordinates for all the points that have been identified and matched before.

To begin with the first part, it is known that

$$w \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} X_i, \quad (9)$$

this gives three equations which can be solved simultaneously. Solving this and factoring  $X_i$ , the final equation as follows,

$$\begin{bmatrix} P_3 x_i - P_1 \\ P_3 y_i - P_2 \end{bmatrix} X_i = 0. \quad (10)$$

Similarly, solving for the corresponding point viewed from the second camera gives,

$$\begin{bmatrix} P'_3 x'_i - P'_1 \\ P'_3 y'_i - P'_2 \end{bmatrix} X_i = 0. \quad (11)$$

Combining these results, the following equation in the form  $A x = 0$  is obtained, where

$$A = \begin{bmatrix} P_3 x_i - P_1 \\ P_3 y_i - P_2 \\ P'_3 x'_i - P'_1 \\ P'_3 y'_i - P'_2 \end{bmatrix}. \quad (12)$$

Normalising  $A$ ,  $A_{norm}$  is obtained in the form,

$$A_{norm} = \begin{bmatrix} \frac{1}{\|A_1\|} A_1 \\ \frac{1}{\|A_2\|} A_2 \\ \frac{1}{\|A_3\|} A_3 \\ \frac{1}{\|A_4\|} A_4 \end{bmatrix}. \quad (13)$$

Lastly, the singular value decomposition of  $A_{norm}=USV^T$  is obtained, and the 3D point coordinate  $X$ , is taken from the last column of  $V$ .

Now, after the cheirality test is done using the obtained results, the second part of triangulation to obtain all the 3D point coordinates can be done. For this purpose, the intrinsic parameters  $K$  obtained before is multiplied with  $P$  using simple matrix multiplication method before the triangulation process. This will transform the points from camera perspective to the world view perspective which gives the 3D point coordinates.

An experimental result is shown in Figure 12, where two images of the same object are taken from two slightly different views using the calibrated webcam.



Figure 12: Two images with 18 matched points marked in red

These images are scanned and 18 point pairs are found and matched. From these matched points, the system then computes the Fundamental Matrix, and the corresponding 3D point coordinates.

This 3D point coordinates is then fed into Matlab solely for plotting purposes. Figure 13 shows the 3D plotted points of the object in blue as well as the camera positions which is in red.

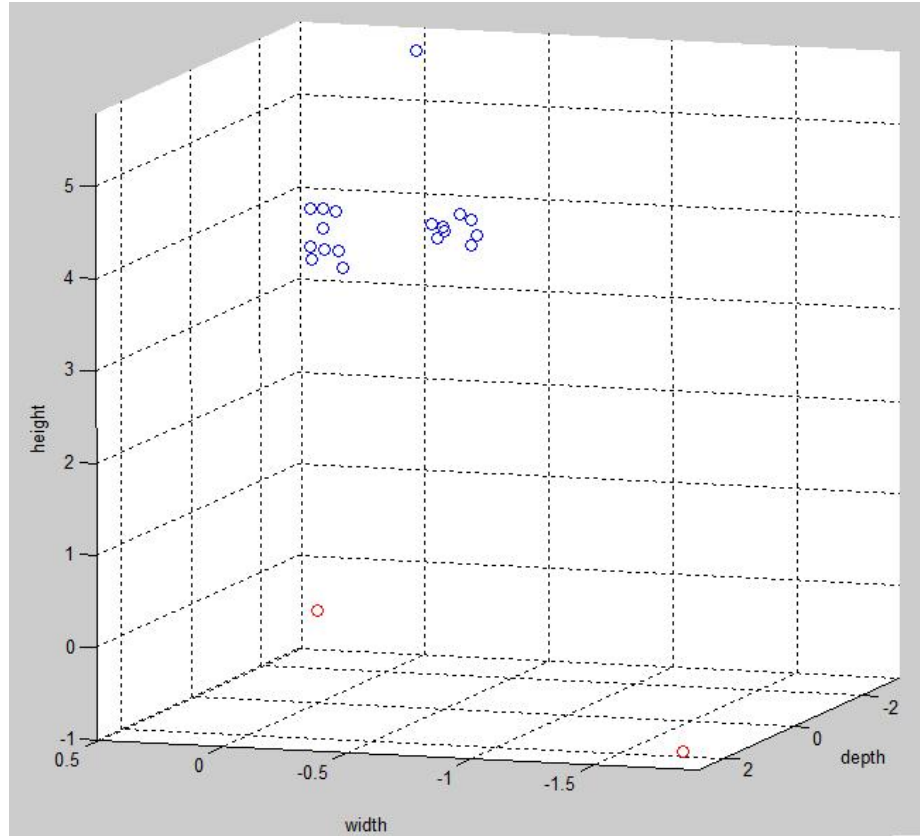


Figure 13: 3D Reconstructed Points and Camera Locations

## 4.5 User Interface

To integrate the webcam and the application to allow an easier and faster application usage, a user interface has been designed and created. This is done using MFC application in Visual Studio 2005.

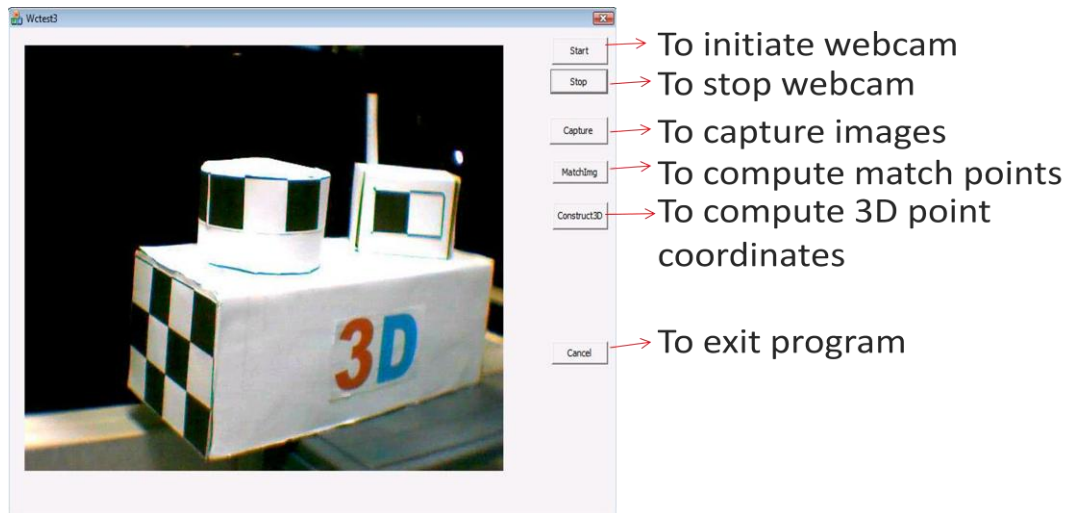


Figure 14: User Interface

As shown in Figure 14, the user interface allows video stream to be shown on screen. The function of each button is also shown in Figure 14, where all parts of the system can be run by a simple click of buttons provided within the user interface.



## **CHAPTER 5**

### **CONCLUSION AND RECOMMENDATION**

#### **5.1 Conclusion**

Nowadays, most robotics or automated vehicles utilise image sensors to navigate itself in an unknown environment. However, the costs of these image sensors are very high and it requires a very high computational power. In order to overcome this problem, an alternative is to use a relatively cheap image sensor which is the webcam. From this research study, it is proven that the navigation system can be designed using the webcam. Also, this study shows that the 3D reconstruction techniques can be improved to minimize the computational power and time as well.

From the research and testing, simulations of the project from points matching to 3D reconstruction are completed in parts. The feature point matching algorithm is being implemented with accurate results. Moreover, the 3D reconstruction technique provides a very good estimation of the scene as well. In conclusion, the main objectives of this project have been achieved. Also, it would be relevant and up to date with the current technology in robotics and computer vision. I believe that this study will benefit the university as well as giving me more insights on robotics and computer vision.

#### **5.2 Recommendation**

At the end of this project, I am certain that a whole robotics navigation system can be implemented using the 3D reconstruction techniques that I have completed. Further fine tunings of the parameters can be made by calibration to suit different environment needs. Also, this system can be expanded to include robotics mapping functions in the future.

## REFERENCES

- [1] Zhengyou Zhang, Rachid Deriche, Olivier Faugeras et al, “A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry”, INRIA Sophia-Antipolis, Sophia-Antipolis, France, May 1994.
- [2] A. Fusiello, E. Trucco, A. Verri, “A Compact Algorithm for Rectification of Stereo Pairs”, Machine Vision and Applications, March 2000.
- [3] David M. Cole, Paul M. Newman. “Using Laser Range Data for 3D SLAM in Outdoor Environments”. Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference. May 15-19, 2006
- [4] Remo Ziglier et al, “3D Reconstruction Using Labelled Image Regions”, Eurographic Symposium on Geometry Processing, USA, pp.1-12, 2003.
- [5] X. Armangue, J.Salvi, “Overall view regarding fundamental matrix estimation”, Image and vision computing, vol. 21, pp.205-220, 2003.
- [6] Stephen M. Smith, J.Michael Brady, “SUSAN- A new approach to low level image processing”, International Journal of Computer Vision, October 1995.
- [7] D.M.Gavrila, L.S.Davis, “Fast Correlation Matching in Large(Edge) Image Databases”, Computer Vision Laboratory, Center for Automation Research, University of Maryland, 2000

- [8] J. Salvi, X. Armangué, J. Batlle. A Comparative Review of Camera Calibrating Methods with Accuracy Evaluation. *Pattern Recognition* 35(7), pp 1617-1635, July 2002.
  
- [9] Luke Fletcher, “An Introduction to Computer Vision”;  
[http://users.rsise.anu.edu.au/~luke/cvcourse\\_files/online\\_notes/lectures\\_RV\\_1\\_dmaps\\_6up.pdf](http://users.rsise.anu.edu.au/~luke/cvcourse_files/online_notes/lectures_RV_1_dmaps_6up.pdf).
  
- [10] R. Hartley and A. Zisserman, “Multiple view Geometry in Computer Vision”, Cambridge, UK; New York: Cambridge University Press, 2000.
  
- [11] David Nister, “An Efficient Solution to the Five-Point Relative Pose Problem”, *CVPR*, pp. 195-202, Princeton, 2003.

## **APPENDICES**

**APPENDIX I: GANTT CHART**

No	Detail/ Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Construct User Interface	■	■												
2	Integrate Point Matching			■	■	■									
3	Integrate Triangulation					■	■	■							
4	Integrate 3D Plotting							■	■	■					
5	Validate System									■	■	■			
6	Fine Tuning											■	■	■	
7	Report Writing													■	■

 Progress/ Milestone