TRIANGULATION METHOD AND MORPHOLOGICAL OPERATIONS FOR MEASUREMENTS OF UNDERWATER OBJECTS USING STEREO VISION

By

ANITH NURAINI BT ABD RASHID

A Dissertation submitted for partial fulfillment of the requirement for the degree of Master of Science

Jan 2015

Acknowledgment

In the name of Allah, the most compassionate, the most merciful. I thank Allah for blessing and granting me strength to endure the journey in completing this thesis.

Firstly, I would like to offer my sincere gratitude to my supervisor, Dr Anwar Hasni Nayan @ Abu Hassan for his patience, support and guidance along the way to complete this thesis. His suggestive and constructive comments have contributes to my success.

It is also pleasure to offer my regards to my friends Siti Juliana and Nurul Farhana who helped me with their supporting advices during the research writing. I would also like to express my sincere thanks to Pn Jamaliah who always support and give advices in completing this thesis.

Last, but not least, my special thanks to my beloved husband, my mother and my mother in law for always trying to understand and never ending encouragement during this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

KONSEP SEGITIGA DAN OPERASI MORFOLOGI UNTUK PENGUKURAN OBJEK DALAM AIR DENGAN MENGGUNAKAN PENGLIHATAN STEREO

Abstrak

Tesis ini memperbaiki ketepatan dalam mengesan dan menganggarkan jarak dan saiz sesuatu objek di dalam air dengan menggunakan penglihatan stereo. Masakini, terdapat dua jenis kamera yang dihasilkan oleh industri dengan tujuan untuk merakam imej atau video di dalam air. Jenis yang pertama adalah kamera yang dihasilkan dengan fungsi kalis air manakala jenis yang ke dua adalah kamera yang diletakkan didalam bekas perlindungan kamera kalis air yang lutsinar. Oleh itu, ujikaji ini dilakukan berdasarkan dua situasi bagi menyamai situasi seperti kedua - dua jenis kamera tersebut. Konsep segitiga dipilih untuk digunakan dalam membantu menganggarkan jarak objek di lokasi berbeza. Konsep memproses imej pula digunakan bagi menganggarkan saiz objek dengan pengesan ukur lilit Sobel and operasi morfologi telah dipilih. Daripada hasil kajian, didapati bahawa purata ralat konsep segitiga dan pemprosesan imej bertambah baik setelah pembetulan di buat ke atas ukuran di bawah pengaruh pembiasan. Penurunan ralat ditunjukkan pada situasi pertama adalah lebih baik dengan pembiasan berlaku daripada air ke udara. Anggaran ukur lilit dan keluasan yang jitu pula banyak di pengaruhi oleh pemilihan nilai faktor semasa mengesan ukur lilit objek di mana pengesan ukur lilit Sobel dan operasi morfologi adalah sensitif terhadap piksel.

TRIANGULATION METHOD AND MORPHOLOGICAL OPERATIONS FOR MEASUREMENTS OF UNDERWATER OBJECTS USING STEREO VISION

Abstract

This thesis improves the accuracy in detection and range estimation of any objects underwater by using stereo vision. Nowadays there are two major types of camera produced by the industry with the purposes of capturing underwater image or video. One is with built in waterproof camera while the other type of camera converted into waterproof camera with a waterproof housing. Therefore investigation is done on two scenario setups which built up to mimic these two camera condition. In the experiment, triangulation method is selected in order to evaluate the object distance of multiple locations. Image processing methods are used to measure size of the objects which are Sobel edge detector and morphological operations. From experimental results, it is observed that the average error of the triangulation methods and image processing improved with the correction made to the measurement results under the error of refraction. The most error improvement acquired is the first scenario which refracts from water into the air. Perimeter and area measurement accuracy influenced the most by the factor chose during edge detection where the pixel is sensitive in Sobel operator and morphological operations.

CHAPTER 1

INTRODUCTION

1.1 Background Study

Nowadays, there are many products or applications that require detection of an object and measure the object either its size or its distance from any particular point in any environment condition. Environment condition here means in air or underwater. For many years, the main interest is to find the size and distance of an object lying underwater (Yamashita et al., 2008) (Treibitz et al., 2012) (Andono et al., 2012). There are various combination techniques and methods used for this purposes.

There are two types of imaging systems which are the single viewpoint (SVP) and a non SVP system (i.e., multiple viewpoints system). The non SVP commonly related to stereoscopic vision which is similar to human vision. It can be constructed using a single pin hole camera with a light that produced geometry sensed in the system (Treibitz et al., 2012). A popular non SVP system is developed using stereo vision technology with the use of two similar cameras, for example two cameras which are place on a horizontal plane with a distance between each other (Kumar & Ramakanth, 2011).

Stereo vision technology is one of the techniques that are widely used in assisting the application to acquire images underwater. Stereo vision technology is popular because it embraces visual distance estimation which is the fundamental part of human acuity. Human sight not only based on several depth cues but also the parallax phenomenon which refers to the differences in the apparent position of a visual object from different line of sight. This phenomenon has been applied in the binocular products where the visual output from this binocular is constructed by analyzing left and right eye view. Therefore, stereo vision is also known as the depth perception technique which is reliable and robust (Holzmann & Hochgatterer, 2012). There are other methods accessible for the depth perception, however stereo vision benefits users more for its accuracy and passiveness (Kheng et al., 2011).

Each image of stereo vision is analyzed to give particular information of the object intended for measurement, for example, intrinsic and parasitic parameters of a camera. Camera calibration is the process of computation these parameters. Image processing of an object, either in air or underwater, normally is involved in filtering out noise of the captured image, segmenting the image and also selecting particular image intended for the measurement. The obtained pixel information from the image is processed in order to calculate the distance or size of any objects required by the system (Mustafah et al., 2012). One of the most popular methods in distance calculation is the triangulation method.

In underwater environment, stereo vision technology was further challenged with the visibility issue of the image, which will impact the accuracy of the end result. Above water light source comes naturally from the sun, while underwater has the light limitation due to the light attenuation (Kumar & Ramakanth, 2011).

1.2 Problem Statement

Navigation and localization of an autonomous system is an important system that is used at underwater application. The distance and size measurement is a part of useful information to the system (Mustafah et al., 2012). The system is deployed in remotely operated vehicle (ROV) to serves several purpose such as to explore and monitor underwater eco-system and size estimation of an organism structures. This vehicle is attached with stereo vision technologies in order to capture images. However there are a few limitation of stereo vision usage at underwater. One of it is light refraction effect where it occurs when camera is placed at a different environment (i.e., different refractive index). It is a common practice that a camera is not exposed directly to the water or it is covered with a waterproof housing to shield water from entering camera lens. At either way, there is a different in refractive index prevent an accurate measurement. The other limitation effect is light attenuation where natural light source (i.e., sun) degraded as the objects goes deeper underwater (Yamashita et al., 2008). This creates difficulties in recognizing the objects edges during image processing.

Most of the researcher proposed their work by focusing on either distance measurement, for example, the work from Treibitz et al. (2012) or size measurement, for example, work from Yamashita et al. (2008). Thus initiates this research in considering the accuracy of the stereo vision on both distance and size measurements of objects underwater .

1.3 Objective and Scope of Study

The aim of the research is to compare and discuss the accuracy of two measurements describes in problem statement on multiple size and location of objects underwater. To realize this aim, the following objectives are adopted:

i. To investigate and evaluate the effectiveness of the triangulation method in distance measurement of underwater objects at both scenario conditions in terms of accuracy.

ii. To find an accurate parameter value to the multiple image processing algorithm in order to calculate the object size at both scenario conditions.

1.4 Research Scope

This research discussed on the stereo vision implementation on a range estimation system of underwater objects. In this research, the experiment activities are a lab base setup. For the stereo vision setup, two web cameras are used instead of high performance camera used in the industry or stereo camera. Underwater environment is created by using an aquarium filled with tap water. In this research, the stereo vision experiments are divided into two scenario applications. The web cameras are placed at two locations. The first location is at the top of the aquarium while the second location is located at the side of the aquarium where the glass of the aquarium represents the waterproof camera housing. The selected objects are pasted at the bottom of the aquarium and also pasted at a flat plastic holder in order to hold the objects vertically in the aquarium. In a real environment light source comes from the sun, however in this experiment the light source is merely from the room ceiling light.

1.5 Thesis Outline

This thesis is divided into six chapters. Chapter 1 introduces the overview of the underwater imaging system, research objective and research scope. Chapter 2 indicates the reviews of the researchers on stereo vision system, light refraction, camera calibration and image recognition. Chapter 3 outlines the methodology of this research and also demonstrates the experiment setup and the calculations used in the experiment. Results and discussions are given in Chapter 4 and lastly the conclusion and suggestion were presented in Chapter 5.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter discusses the literature review of this research. Section 2.2 describes the stereo vision system and explains the stereo vision implementation. Section 2.3 explains the calibration methodology that is commonly used by researcher while section 2.4 explains the theoretical behind the refraction phenomenon and describes the relationship of refraction to optics. In section 2.5, reviews on the image rectification and image segmentation with a discussion on the image processing methods that involved in both rectification and segmentation. Lastly, the summary of the Chapter 2 is given in section 2.6.

2.2 Stereo Vision

Stereo vision technology was developed with the intention to mimic the human vision and the way human brain processing the image received through their vision. Human vision requires a light source in order to view the image received. Therefore, in a dark environment human eyes only see a black image. It is the same requirement for a stereo vision where image captured can only be recognized with the present of light source (Borangiu & Dumitrache, 2010).

Human vision also described as stereoscopic vision which refer to the ability of having two images from left and right eye and further infer the information on a 3D structure and perceive the distance or depth of any points accurately. In order for the brain to structure the 3D of the image, the difference of the retinal location is identified. This difference in location is called the disparity (Trucco & Verri, 1998).

Following sub sections briefly explain on the stereo implementation and systems lie behind this technology.

2.2.1 Stereo Vision System Principle

The basic property of the stereoscopic acquisition systems is when two cameras positioned at different viewpoints and observe the same object or sight. There are two possible way to position the camera which either in an epipolar geometry or canonical stereoscopic system. However these systems share the same method where both system right and left centre points lie on the baseline (Cyganek & Siebert, 2009).

In figure 2-1, it is clearly showed the setup defined by the epipolar geometry. The reference, O_R and the target, O_T represent the left and right optical centre point and y is the baseline. For epipolar geometry system the left epipolar line, Π_l and right epipolar line, Π_r is perpendicular to the base line. By considering point P and Q on the same line from the reference point, the projection is on the left and right image represent by p and q on the left and p' and q' on the right image. Figure 2-1 also explains the constraint of the epipolar system where image on the same line seen as only one object instead of two objects (Mattoccia, 2013).

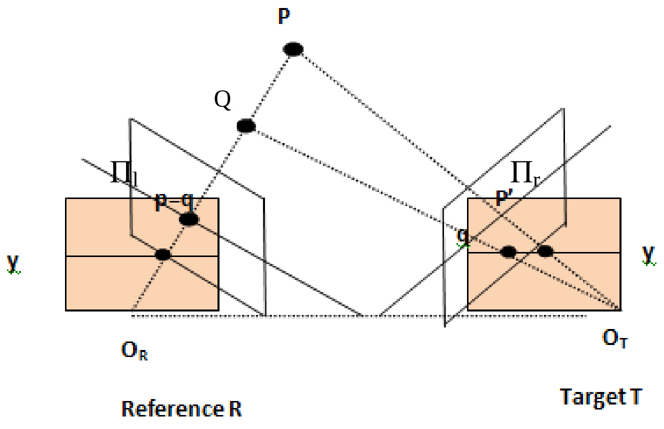


Figure 2-1: Stereo vision epipolar geometry (Mattoccia, 2013)

In figure 2-2, it showed the canonical system where two cameras angle positioned on the base line at a relative distance. It also explains that the image plane is parallel to the base line.

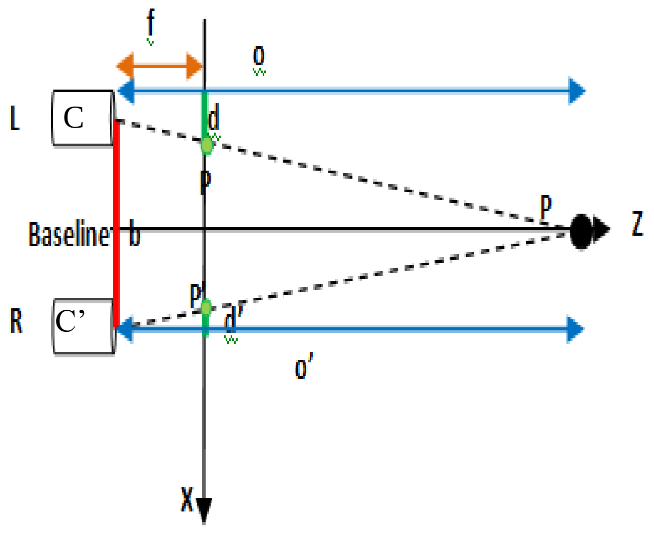


Figure 2-2: Standard canonical system (Holzmann & Hochgatterer, 2012)

2.2.2 Triangulation

From the observation of both stereoscopic acquisition systems, it is visibly seen that the image projections related to triangle. The triangulation method has a depth perception. A common method to calculate depth or the distance of stereo vision system is called triangulation. In triangulation, the depth, Z can be estimated from the disparity of the corresponding points (Trucco & Verri, 1998). Apart from that information, triangulation also needs the knowledge of calibration data such as focal length and centre of projection (Cyganek & Siebert, 2009).

In figure 2-2, C and C' are the center of projection of the stereo vision system while b is baseline which is the distance between left and right centre of projection. p and p' is the image projected from the object P. The distance between centre of projection and x axis is the focal length, f. The x axis here is predefined by the used camera through calibration (Holzmann & Hochgatterer, 2012).

The triangulation equation defines from the objective of recovering the position of a single point which is P in this case from its projections p and p'. Let d and d' be the coordinates of p and p'. The coordinates is respected to the principal point, focal length, f and the depth, Z. Z is the distance between point P and the baseline.

From the similar triangle (p,P,p') and (C,P,C'), equation (2-1) is formed (Trucco & Verri, 1998).

$$\frac{b}{Z} = \frac{b-(d+d')}{Z-f} \tag{2-1}$$

Equation 2-2 obtained by solving equation (2-1) where disparity, D = d+d’ (Holzmann & Hochgatterer, 2012).

$$Z = \frac{f.b}{D} \tag{2-2}$$

2.2.3 Stereo Vision Implementation

There are many researcher who has the interest in stereo vision technology adapted into their research experiments. For example Calin & Roda (2007) and Mustafah et al. (2012) where both of them using stereo vision on real time application. Calin & Roda (2007) focused on disparity map extraction targeting of implementation on programmable logic, where pipelined structures and condensed logic were used. On the other hand Mustafah et al. (2012) focused on real time object distance and size measurement. Apart from that, the stereo vision is also adapted into underwater research experiment. For example Yamashita et al. (2008) used stereo vision in determining the unknown refractive index by using the image of water surface and restore the images. While Kumar & Ramakanth (2011) used stereo vision technique in developing autonomous system to build 3-D Model of underwater objects with improvement in post processing stage.

2.3 Light Refraction

One of the problems that held human behind from observing the underwater environments is the light refraction phenomenon. In this section, the Snell’s Law and the effect of the refraction will be further discussed.

2.3.1 Snell's Law

The Snell's Law is the law of refraction that explain the relationships between the angle of incident and angle of refraction when a light passing through different medium with different refraction index. The equation is shown in equation (2-3), where η is the refraction index and ɵ is the angle of refraction or angle of incident (Jahne, 1997).

$$\eta_1 \sin\theta_1 = \eta_2 \sin\theta_2 \qquad (2\text{-}3)$$

Figure 2-3 shows an example of an underwater condition with ray refracted through multiple medium. The rays start its refraction from a liquid into a flat surface which in this situation is a glass and then again refracted from the glass into the air. In this condition the axis z and axis r are radial symmetrical. Axis z is perpendicular through the centre of the lens. Equation (2-4) shows the application of the Snell's Law into this underwater condition (Treibitz, et al., 2012).

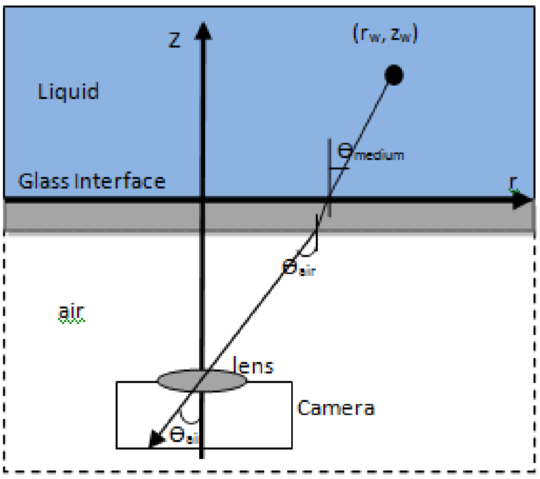


Figure 2-3: Diagram of ray refraction in underwater condition (Treibitz, et al., 2012)

$$\eta \sin \theta_{medium} = \eta_{glass} \sin \theta_{glass} = \sin \theta_{air} \tag{2-4}$$

2.3.2 Effect of Light Refraction Underwater

Figure 2-4 shows the effect of the light refraction where the size of the objects is getting bigger when underwater or any medium that is larger in refraction index compared to vacuum or air. Apart from the refraction index between air and water, the enlargement of the object also contributed by the refraction of the interface in between camera and water (Yamashita, et al., 2008).

Figure 2-4: Light refraction effect (Yamashita, et al., 2008)

Camera needs protection case when submerge underwater to prevent water from entering the camera lenses. Therefore this interface has to be considered when making error calculation. The ray path will change according to the thickness of the interface which is clearly showed in figure 2-5. It shows multiple ray projected from the objects changed the distance seen by human and also means the objects size will also varies (Telem & Filin, 2010).

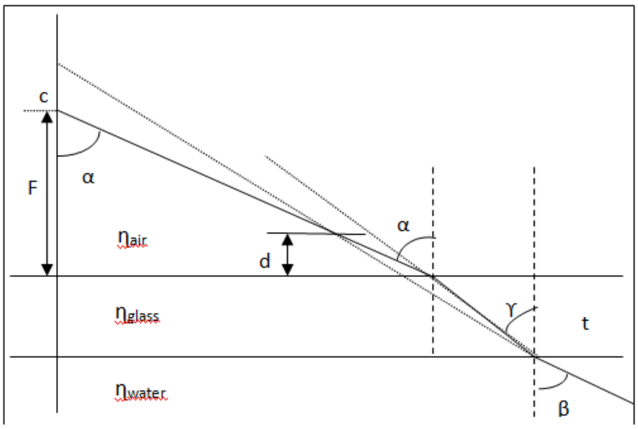


Figure 2-5: A ray path diagram of the interface thickness effect on refraction (Telem & Filin, 2010)

Talem et al. (2010) also stated that the distance between the perspective centre and the interface, F changed as the interface thickness and refractive indices ratios changes. This shows from the equation (2-5) which is the paraxial optics and then applied with the Snell's law to form equation (2-6).

$$d = \frac{\tan\beta - \tan\gamma}{\tan\alpha - \tan\beta} \,.\, t \qquad (2\text{-}5)$$

$$d \cong \frac{\left(\frac{\eta_{air}}{\eta_{water}} - \frac{\eta_{air}}{\eta_{glass}}\right)}{\left(1 - \frac{\eta_{air}}{\eta_{water}}\right)} \,.\, t = K.t \qquad (2\text{-}6)$$

Kondo et al. (2004) stated that underwater measurements can be discovered through calibration table with condition that the stereo vision is properly calibrated. This calibration table shows the relationship between distances and pixel positions, thus any distance can be easily acquired with knowing only the pixel position of the object.

However Yamashita et al. (2008) disagrees with the table usage as it limited to a specific condition of liquid. If environment changes such as from sea water to river water, the table that is base from sea water no longer useful. Therefore knowing the refractive index is crucial in determining more accurate underwater measurements.

Treibitz et al. (2012) also gave attention to the important role of the refractive index in determining the distance or the depth of an object underwater. With the used of Fermat's principle, equation (2-7) is formed in order to calculate the optical path length, L between the object underwater to the lens (Treibitz et al., 2012).

The equation is referred to figure 2-3 where η is refractive index, d is the distance between lens and glass surface along Z axis, r_w is distance between objects and centre of lens along r axis and r_g is the distance between lens and the point light refracted from glass to air.

$$\boldsymbol{L = \eta\sqrt{\left(r_w - r_g\right)^2 + z_w{}^2} + \sqrt{r_g{}^2 + d^2}} \qquad (2\text{-}7)$$

2.4 Calibration

Calibration is an important procedure in any stereo vision application before capturing and processing any image or video. Important parameter information obtained through calibration are the intrinsic parameter and the extrinsic parameter. These parameters is useful in calculating depth or distance or any measurements underwater (Bruno et al., 2011).

Intrinsic parameters explain the relationships between pixel coordinates and camera coordinates in each camera. Intrinsic parameter comprises of focal point, principal point, radial and tangential distortions and pixel size. Extrinsic parameters describe the camera position and orientation of the two cameras in the world. It is presented in a rotation and translation vector (Elgammal, 2014)

There are various calibration techniques being use by researchers and Zhang's Technique has been widely used. Most of the algorithms created are using Zhang's method as their basis. In calibration technique there are 4 types of technique that can be selected. They are either using the 3D calibration objects, 2D planar pattern, 1D object (line-based calibration) or self calibration. The more dimension used, the higher the accuracy. However the more works are needed. Figure 2-6 shows the 3D objects used for calibrating cameras using the 3D calibration technique (Zhang, 2000).

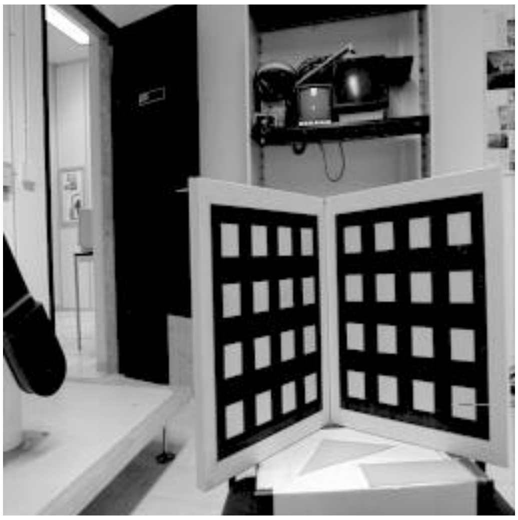


Figure 2-6: 3D object for calibrating cameras (Zhang, 2000)

In 2000, Zhang has created the calibration technique using 2D planar pattern. The image used during the calibration experiment is as shown in figure 2-7 (Zhang, 2000). However in 2002, Zhang has presented a new calibration technique using 1D object for camera calibration. Zhang proves that all four techniques can be used in calibrating cameras (Zhang, 2002).

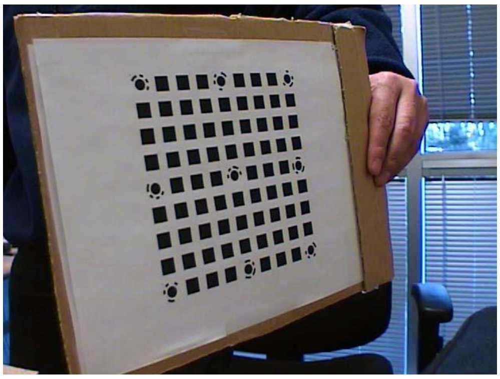


Figure 2-7: Sample image of planar pattern (Zhang, 2002)

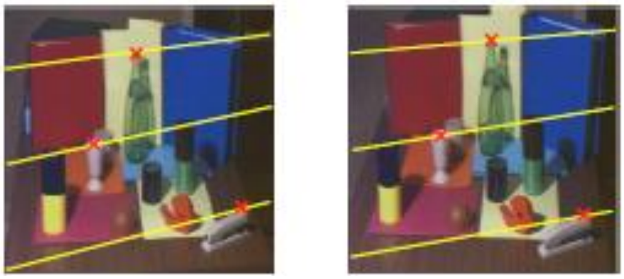
Despite all of the techniques available there is always an ultimate way of calibrating the camera. Since each technique has pro and cons, it is good to have the technique that compromise both accuracy and load of work took in calibration. Therefore 2D planar pattern serve the best purpose of having good accuracy level with a simple experimental setup.

In MATLAB, these techniques algorithm has been incorporated into their Camera Calibration Tools to ease user to start camera calibration and even exploring the calibration algorithm. Bouquet has created the toolbox with reference on Zhang technique (Bouguet, 2013).

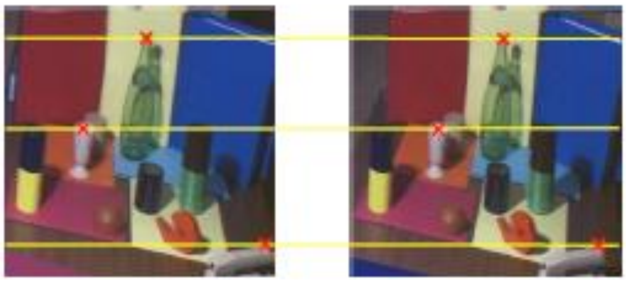
2.5 Image Recognition

2.5.1 Image Rectification

Stereo matching process is required to recover information from two images acquired at the same time but from a different location. The simplest division of stereo methods is based on the disparity map and mostly desirable is the dense disparity maps. While sparse disparity maps determined only for selected objects such as corners or edges. These methods are faster but limited to a certain application due to missing value that has to be interpolated (Cyganek & Siebert, 2009).



(a)



(b)

Figure 2-8: (a) Original stereo pair (b) The rectified pair. Left image plot to the corresponding point marked at the right image (Trucco & Verri, 1998)

In rectification of stereo, the disparity maps displacement horizontally so that the epipolar lines become collinear and parallel to the horizontal axes. Rectification is important in order to tackle correspondence problem involves in 2D in general. By looking along single scan line, a corresponding point can be determined. For a simple technique of rectification, it is necessary to know the intrinsic and extrinsic parameters. The rectified image can be assumed as having a new stereo rig by rotating the original camera around their optical centers. Figure 2-8 shows the rectified image from a stereo image where it's epipolar lines are in parallel at both left and right rectified image at the correspond point marked on the right image (Trucco & Verri, 1998).

2.5.2 Image Segmentation

Image segmentation is a necessary preliminary step in most of real time pattern recognition and scene analysis problems. Segmentation subdivides the object or a region of an image individually. The level of subdivision depends on the problem statement of a research. Segmentation is finalized once the object or region of interest extracted out from other objects of an image. However the accuracy differs especially in a case of non trivial images, which is the most difficult tasks in image processing (Gonzales & Woods, 2002). This will be one of the challenges when processing an image of underwater objects due to light limitation and caused image blurry.

Segmentation can be divided into three major types of algorithm which are thresholding, edge-based segmentation and region based segmentation (Sonka, Hlavac, & Boyle, 2008). In thresholding, the simplest technique is by portioning the image histogram using single global threshold. The final segmentation acquire by scanning and labeling each pixel as objects or background. The indicator is whether the gray level of the pixel compared to the value of T. This method depends entirely to the histogram portioning (Gonzales & Woods, 2002).

Edge-based detection relies on a set of connected pixels within the boundary between two regions. Compared to the region based which is defined with more global idea, edge is defined with a local concept. Edge detection algorithm typically requires edge pixel to be assembled into meaningful edges (Gonzales & Woods, 2002). There are three popular edge detectors which are Canny Edge Detector, Roberts Edge Detector and Sobel Edge Detector. Canny Edge Detector involves in multiple mathematical formula in its algorithm which includes Canny Enhancer, non maximum suppression and hysterisis threshold. Suppression and thresholding is the edge localization in edge detection. Canny enhancer applies Gaussian smoothing to eliminate noise in the image. The output image may contain wide edge which is then suppress by non maximum suppression to a 1 pixel wide edges. Even with the suppression and filter, noise of the image may still occurred. Therefore thresholding is also applied to the image to get rid the unnecessary edges. Hysteresis thresholding is selected to avoid false contours being discards where edge pixel must be in between lower threshold and high threshold. Roberts Edge Detector and Sobel Edge Detector are much simpler to be implemented as compared to Canny Edge Detector. Gaussian smoothing is also used to eliminate noise in both detectors. Then followed with linear filter with set the mask of each detector shows in figure 2-9. Next is the gradient magnitude estimation where equations (2-8) show the gradient magnitude, G equation use in both Robert and Sobel Edge. Followed by marking all edges where each pixel must meet the requirement state in equation (2-9) (Trucco & Verri, 1998).

-1	-2	-1
0	0	0
1	2	1

-1	0	1
-2	0	2
-1	0	1

(a)

-1	0
0	1

0	-1
1	0

(b)

Figure 2-9: Mask to compute gradient (a) Sobel (b) Roberts

$$\nabla_f = \text{mag}\,(\nabla \mathrm{f}) = \left[\mathrm{G}^2{}_{\mathrm{x}} + \mathrm{G}^2{}_{\mathrm{y}}\right]^{\frac{1}{2}} \qquad (2\text{-}8)$$

$$\boldsymbol{G}(\boldsymbol{x},\boldsymbol{y}) > \tau \qquad (2\text{-}9)$$

Where τ is the threshold value.

Region-based segmentation requires an image to be partitioned into regions. It is accomplish through boundaries finding between regions based on discontinuities in gray levels (Gonzales & Woods, 2002). In Nasrul & Akmal (2011) work, local region-based active contour is used for image segmentation in measuring area of a thyroid in ultrasound image. Since it is for medical purposes local based region is more suitable than other method such as edge based segmentation. In edge detection, utilization of a gradient is used in order to identify the boundaries. Thus increase the sensitivity towards the image noise.

2.6 Summary

In this chapter an understanding acquire on stereo vision principle and how triangulation method relates to stereo vision. The relationship enables the calculation of depth or distance of object underwater. Since this project interest is on underwater objects, an understanding was also gathered on the effect of refraction towards the object distance, depth and also the size. Apart from that, introduction to image processing method used in preliminary step of object recognition is also concluded.

CHAPTER 3

METHODOLOGY AND EXPERIMENTAL SETUP

3.1 Introduction

This chapter describes every process and method involved in this research activity. Section 3.2 will present the general flowchart of this research framework. Section 3.3 explains the stereo vision system development. Section 3.4 and section 3.5 presents the process involves in rectifying and segmenting images. Section 3.6 explains the steps involved in estimating the distance and size of objects. Lastly section 3.7 presents the experiment setup for this research and section 3.8 presents the calculation in estimating the distance and size of the objects.

3.2 Flowchart of Research Activity

Figure 3-1 displays the proposed process flow for the stereo vision system in measuring the depth, distance and size of any objects. This proposal flowchart is with the reference from the paper from Nasrul & Akmal (2011) where he measures the size of a thyroid ultrasound image and also from the stereo vision system overviewed by Mattoccia (2013).

To begin the research activity the camera has to be calibrated properly as it generates important parameters which will impact the measurement throughout the process. The next step is acquiring the image of objects required for the measurement from both left and right camera. The image size used for this system is 640x480 pixels in RGB. Next process is image rectification which will ensure both images meet the specification of parallel epipolar line. Then the rectified images is segmented before objects centroid is obtained to further process and calculate the depth and distance of the object using triangulation method. Once the distance and depth calculation are completed, the image is further processed in order to measure the size using the image processing tools. Lastly all the results will be displayed through the system display screen and further analyzed depends on the scenario interested.

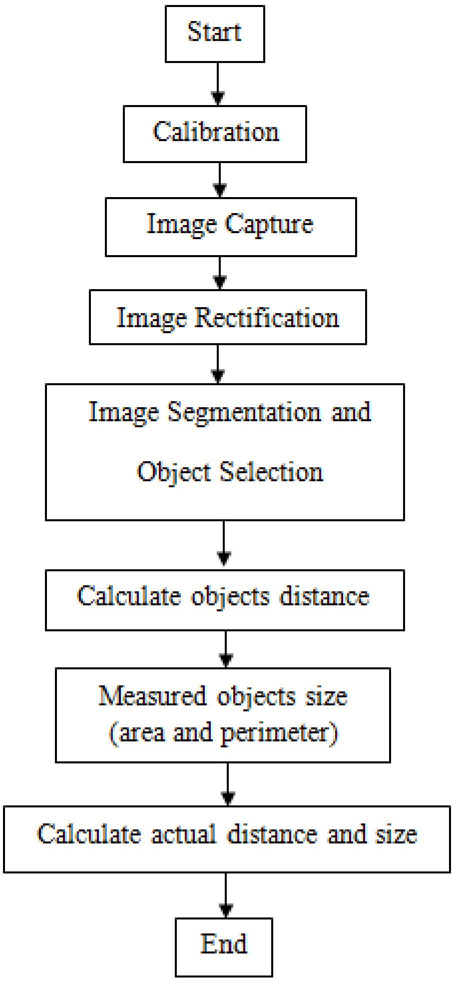


Figure 3-1: The research activity flowchart

3.3 Stereo Vision Methodology

In the experiment the stereo vision is set up with the use of two cameras arranged in parallel of its epipolar line. Figure 3-2 shows the geometry plan of the camera where the objects axis is perpendicular to the axis of the camera. The objects location presents in x and y coordinate where the centre of coordinate (0,0) is at the corner of the image frame captured by the stereo camera.

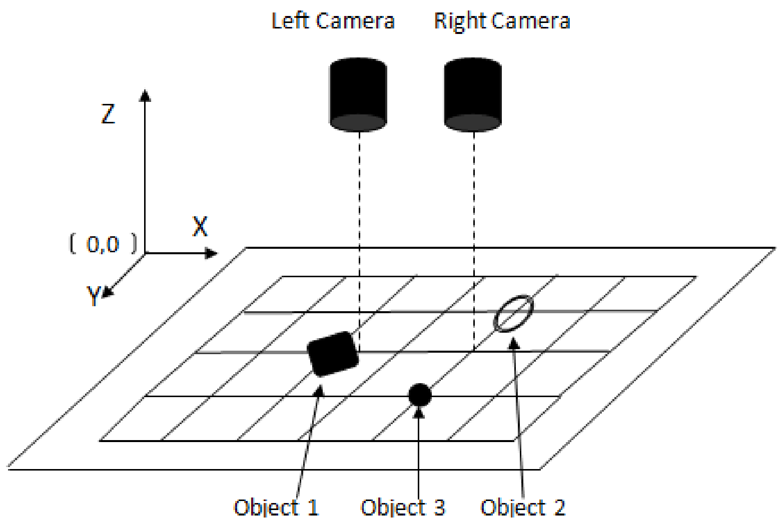


Figure 3-2: Perspective view of the workspace; location of the stereo cameras and objects

The objective of the experiments is to estimate the objects distance and size at underwater. Therefore to mimic the underwater scene, an aquarium filed with water is use and objects locate at the bottom of the aquarium. Figure 3.3 illustrate the perspective view of the workspace with aquarium and also includes in details description of the important distance required in this experiment. This includes in the depth of the objects from camera, the baseline and also the height of the water level.