# UNIVERSITI TEKNIKAL MALAYSIA MELAKA

**Projek Sarjana Muda**

**3D Reconstruction Using Silhouette Coherence for Image Acquisition under Circular Motion**

Thesis submitted in accordance with the partial requirement of Universiti Teknikal Malaysia Melaka for the Bachelor of Electronic and Computer Engineering (Computer Engineering)


**By**

**NORLIANA BINTI AZHAR**

**B021110215**

**4BENC**

**SUPERVISOR**

**DR WIRA HIDAYAT BIN MOHD SAAD**

Faculty of Electronic and Computer Engineering (Computer Engineering)


8JUN 2015.

**UNIVERSTI TEKNIKAL MALAYSIA MELAKA**
FAKULTI KEJURUTERAAN ELEKTRONIK DAN KEJURUTERAAN KOMPUTER

**BORANG PENGESAHAN STATUS LAPORAN**
**PROJEK SARJANA MUDA II**

| | | |
|---|---|---|
| **Tajuk Projek** | : | 3D Reconstruction Using Silhouette Coherence for Camera Acquisition Under Circular Motion |
| **Sesi Pengajian** | : | 2/ \| 2 \| 0 \| 1 \| 5 |

Saya NORLIANA BINTI AZHAR

mengaku membenarkan Laporan Projek Sarjana Muda ini disimpan di Perpustakaan dengan syarat-syarat kegunaan seperti berikut:

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

☐ **SULIT*** — *(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

☐ **TERHAD**** — **(Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

☐ **TIDAK TERHAD**

Disahkan oleh:

(TANDATANGAN PENULIS) (COP DAN TANDATANGAN PENYELIA)

Tarikh: 8/6/2015 Tarikh: 8/6/2015

# DECLARATION

I hereby, declared this report entitled "3D Reconstruction Using Silhouette Coherence for Camera Acquisition under Circular Motion" is the results of my own research except as cited in references

Signature :

Author's Name : Norliana Binti Azhar

Date : 8 JUNE 2015

**APPROVAL**

"I / we hereby declare that I have read this in my / our work is adequate in terms of scope and quality for the award of a Bachelor of Electronic Engineering (Computer Engineering)."

Signature : ……………………………………………….

Supervisor's Name : Dr.Wira Hidayat Bin Mohd Saad

Date : 8 JUNE 2015

## DEDICATION

Dr.Wira Hidayat Bin Mohd Saad

Azhar Bin Mat Isa

Rohani Binti Mahat

Norsyafiqah Binti Mohd Ismail

# ACKNOWLEDGEMENT

I take this opportunity to express gratitude and appreciation to my supervisor, Dr. Wira Hidayat Bin Mohd Saad for his supervision, encouragement and assistance through the research. I also thank my parents and colleagues for the unceasing encouragement, support and attention.

Norliana Binti Azhar

**ABSTRACT**

This thesis presents the reconstruction for 3D image by using silhouette algorithm and camera calibration in order to enhance the 3D reconstruction. This technique is useful to reconstruct an object's geometry taken under circular motion or turn table. Once the image acquisition steps is finish, the camera calibration is done in order to recover the extrinsic and intrinsic parameters of camera. The silhouette-based approach is used in this thesis. Silhouette coherence is a generalization of epipolar tangency constraint used to calculate the motion from silhouette or the outline of the object. This approach is capable of exploiting the information in the silhouette compared to the earlier silhouette-based approach that only depends on the epipolar tangency. With the flexibility of this technique makes it usable in a situation where outer epipolar tangents or point correspondent are not available. This thesis also presents the algorithm that used for exploiting silhouette coherence for camera motion estimation. The proposed algorithm is also used to reconstruct a high quality 3D model under circular motion sequence even when the silhouettes are truncated or the epipolar tangency points are not available. The sequence of images used in this thesis were obtained by using turn table. Visual evaluation is done for the final result of the 3D image for different camera angle and different number of sequence. All the factors that cause effect to the final 3D image are evaluated in order to achieve better 3D reconstruction result.

# ABSTRAK

Tesis ini membentangkan pembinaan semula imej 3D dengan menggunakan algoritma bayang dan kalibrasi kamera untuk menambahbaik pembinaan semula 3D. Teknik ini berguna untuk membina semula geometri objek yang diambil secara gerakan membulat atau di sekeliling objek. Setelah langkah-langkah pengambilalihan imej akhir, kalibrasi kamera dilakukan dalam rangka untuk mendapatkan kembali ekstrinsik dan intrinsik parameter kamera. Pendekatan berasaskan bayang digunakan dalam tesis ini. Kepaduan bayang adalah gambaran umum kekangan tangen epipolar yang digunakan untuk mengira gerakan daripada bayang atau garis objek. Pendekatan ini mampu mengeksploitasi maklumat dalam bayang berbanding dengan pendekatan berdasarkan bayang sebelumnya yang hanya bergantung kepada tangen epipolar. Dengan fleksibiliti teknik ini menjadikannya sesuai digunakan dalam keadaan di mana tangen epipolar luar atau titik tertentu yang tidak boleh didapati. Tesis ini juga membentangkan algoritma yang digunakan untuk mengeksploitasi kepaduan bayang untuk anggaran gerakan kamera. Algoritma yang dicadangkan juga digunakan untuk membina semula model 3D yang berkualiti tinggi di bawah urutan gerakan membulat walaupun bayang yang dipangkas atau garis tangen epipolar tidak boleh didapati. Penilaian visual dilakukan untuk keputusan akhir imej 3D untuk sudut kamera yang berbeza dan jumlah yang berbeza daripada urutan. Semua faktor yang mempengaruhi hasil imej 3D akhir dinilai untuk mencapai hasil yang lebih baik bagi pembinaan semula 3D.

# TABLE OF CONTENTS

# LIST OF FIGURES

# CHAPTER 1

# INTRODUCTION

## 1.1 Project Introduction

The 3D reconstruction is known as a process of capturing the shape and appearance of real objects in computer graphic and computer vision. By using multiple images, the 3D reconstruction is actually the creation of three-dimensional models from a set of images and can be considered as the reverse process of obtaining 2D images from 3D scenes. In this thesis, the image acquisition is done by capturing image surrounding the object that need to be constructed. With fixed camera position, the object is placed on a turn table that is then rotated and the image are taken in circular motion for every rotation. Once the image have been obtained, the object's silhouette must be extracted [20]. There are several techniques introduced to convert the RGB image to binary image or silhouette. Image segmentation is one of the technique that can be used where a digital image is divided into multiple segments. This technique is suitable to locate objects and boundaries in images. Other than that, the binary image can also be produced based on their HSV (hue, saturation, and value). Silhouette algorithm is used to enhance the 3D reconstruction imaging method by exploiting silhouette coherence for reliable and efficient camera motion estimation and also for reconstruction of very high quality 3D models [2].

In order to get better results, camera calibration is an important process that need to be done in order to retrieve the camera extrinsic parameter and recover camera intrinsic parameter. These parameters are position of the camera and its pointing direction, focal length, format size, principle point, and lens distortion. Pattern-based calibration, structure from motion [5], genetic algorithm [3], feature-based and silhouette-based are the example of camera calibration approach from non-calibrated image sequences [6].

In feature-based approach, the information of the camera and the 3D structure of the object are determined concurrently from the feature by structure from motion (SfM) algorithm. Meanwhile, in silhouette-based approach, the cameras need to be calibrated first [6], followed by the volumetric description, intersection [4] or the image-based visual hull technique for object modelling. In this thesis the technique use the silhouettes in order to recover the extrinsic and intrinsic camera parameter under circular motion [10]. The proposed technique is an improved from the previous silhouette-based method which can deal with truncated or partial silhouette. Silhouette coherence has the ability to exploits the information at the epipolar tangency as well as in the silhouette. This advantage makes this technique available to be used in situation where the outer epipolar tangents or point correspondences are unavailable [2].

The visual hull of the 3D object is the closest approximation that can be obtained by volume intersection. When the camera capturing the image, it's actually back projecting a silhouette cone or also known as visual cone through the silhouette. The algorithm makes use of this silhouette cones to reconstruct the real object in 3D. The intersection from all camera views will form a region that is called a visual hull. The precision of the reconstructed 3D image becomes higher as the number of silhouette cone increase [2]. However the 3D reconstruction based on volume intersection algorithm are not able to exactly reconstruct the object especially the object with flat surface such as cube, cylinder and prism. Since the visual hull produced is only an approximation of the real object geometry, the final 3D image will not having the same geometry as the real object. Thus it is important to evaluate all the factors that contribute to the problem.

Voxel is technically similar to pixel in term of their functions but different in dimension space on a regular grid. Both of them consist data of an image. Pixel is a sample of an original image which has variable of intensity for each pixel. More samples or more number of pixels provide more accurate representations of the original. Pixel is represents data in 2-D dimension meanwhile the voxel is in 3-D dimension. The process of adding depth to an image using a set of cross-sectional images known as volumetric dataset. Cross-sectional images or slice are made up of pixel where the space between two pixels in one slice is represents a real-world distance (interpixel distance). While the distance of any two slices represents a real-world depth (interslice distance). Opacity transformation gives voxels different opacity values. This is important to expose interior details of an image of a 3D image model.

Space carving method using an algorithm to reconstruct the 3D model from multiple images. This method is used to carve the voxels that lie outside the silhouette. A regular 3D grid of voxel elements is created for the carving process. The visual hull produced from intersection of silhouette cones is projected onto the 3D grid. Space carving algorithm will carve every voxels that lie outside the mask, leaving only the voxels inside the mask. Thus, the left over is the 3D shape of the real object.

## 1.2 Objectives

The aim of the study is to optimize the 3D reconstruction imaging method by using Silhouette algorithm for the sequence of images of the turn table. In order to achieve that, below objectives need to be accomplished:

1) To model a simple geometry block by using silhouette algorithm of 3D reconstruction method.
2) To calibrate the imaging sequence to be used with silhouette algorithm of 3D reconstruction method.
3) To do a visual evaluation of 3D reconstruction of silhouette algorithm for different camera angle and different number of sequence.

## 1.3 Problem Statement

Silhouette coherence is an improved technique of previous silhouette-based approach where it can deal with partial or truncated silhouette. This technique can be used to exploits all the information not only at the epipolar tangency points but also in the silhouette. Although it is better than other silhouette-based technique this approach facing a major problem in producing the final result of 3D image. The silhouette calculation is relative sensitive for error such as wrong camera calibration. It will cause problem for the intersection of the silhouette cones thus produces bad results for the 3D shapes [18]. Some of the factor that contribute to this issues are on the techniques of 2D images sequence, background color, and calibration technique.

Besides, the sampling of the image contour may cause discretization of the silhouette contours. Discretization is a process where each pixel of an image is converted into the black foreground or the white background and it is used for space carving. The unwanted voxel lies outside the silhouette will be carved away and if the discretization is uncontrolling it will affect the accuracy of the 3D reconstruction model.

## 1.4 Scope of Project

In this project, the image acquisition is done by retrieving the image by using camera of the turn table. To avoid getting poor results or no results at all the extrinsic and intrinsic camera parameters need to be recovered. In this thesis the silhouette coherence technique is used for camera calibration where it exploits all the information in the silhouette to estimate the camera motion. The silhouette algorithm is used to enhance the 3D reconstruction imaging method to produce very high quality of 3D models [2]. The 3D model from multiple images is then reconstructed by using space carving technique. The retrieved images are initially converted into silhouettes and will be carved away, leaving only the silhouette parts. These steps are repeated and lastly a complete 3D model is formed after several carvings. The visual evaluation of 3D reconstruction is done for different camera angle and different number of sequence. The position of the camera is located differently for each view and the number of sequence used in this thesis is 36 and 72.

## 1.5 Outline of Chapter

This paper is organized as follows: Chapter 1 is about introduction and overview of the project which consist of background of project, objectives, problem statement and scope of the project. In Chapter 2, the literature review consist of summary and analysis of the related past work and also about the approach used in this project. Methodology and the step of this project will be defined in Chapter 3. Chapter 4 will present the obtained result and Chapter 5 will conclude the whole research and proposes the future work that can be done.

# CHAPTER II

# LITERATURE REVIEW

There are various approaches for 3D image reconstruction such as projective geometry, feature points-based method and multi view geometry. These approaches use different ways of technique for data collection for 3D reconstruction [20]. In this thesis, the multi view geometry method is used .This method use multiple of silhouettes to reconstruct the 3D image of the real object [20] which the binary classification of each image into object and background to compute a bounding volume for an object. This algorithm enables the user to obtain the rough model of the real object [28].

The reconstruction of an object can be performed by using volume intersection approach where the volumetric description of an object is recovered through multiple silhouette [2]. A region of cone form due to volume intersection technique can be obtained by back projecting [18] from a view point of the corresponding silhouette for perspective projections through the silhouette [2]. For different camera view of each silhouette, a cone called visual cone form in their projection [18]. The first region provides an outer bound which has infinite volume for the object's geometry. Thus the intersection of all those regions form an initial approximation of the real object's shape and can also determine the finite volume bound [13] or also known as visual hull. The visual hull provides an upper bound of an object and the bound gets tighter as the number of different silhouette images increases. Thus the outer bound for the object's geometry [13] is reconstructed with higher precision and more accurate.

Camera calibration is the biggest challenge in image acquisition pipeline. There are so many approaches presented in order to recover the camera parameters which are important for the accuracy of 3D reconstruction model. Silhouette-based approach was studied by [10, 15, 17, and 8]. In [8], the authors introduced the concept of mirror symmetry property which derives a common homography that relates silhouette with epipoles under turntable motion. The rest of the camera parameters can be retrieved by using this approach and the epipoles can be located directly from the silhouette. Unlike the other silhouette-based method, this proposed algorithm can work well even with only few views are available. This technique is useful especially for texture-less, translucent, transparent or reflective objects in order to get accurate feature correspondences.

An accurate camera parameter can be achieved by minimizing the difference between the projections of reconstructed visual hull and the silhouette image using graphic hardware. In the thesis presented by [17], camera parameters were calibrated directly from the information of silhouette using the shape from silhouette and optimization. The proposed calibration method can work robustly for different cases and fit for visual hull reconstruction. Since the method is the combination of camera calibration and shape from silhouettes, the accurate camera parameters and object's visual hull can be estimated together. The authors had reported that their method was fully automatic and doesn't need for any calibration pattern during image capture. Besides, no correspondent feature point detection is needed since silhouette based approach is used [8]. Although this approach is robust and can handle different objects, the limitation that it faced is that it only works under turntable motion constraints.

Point-based methods are the most popular among the available technique. These techniques rely on the presence of feature points on object surface but unfortunately these points are not always available and reliable. Alternatively, the silhouette is used to exploit all the notion of epipolar and frontier points [10]. The limitation faced is that the epipolar tangency points are limited and sometimes not available per pair of images. A very good quality of silhouette is needed for this method thus making their integration difficult and also this method unable to calibrate the camera's intrinsic parameters [17].

A research performed by [15] has presented a new approach to recover camera parameters by using genetic algorithm. The camera parameters are obtained by maximizing the total coherence between all silhouettes taken under circular motion. The proposed algorithm can correctly find the optimal solution without initial values of parameters. The main problem that affects the performance of this approach is focal length. Only small variations of the silhouette coherence criterion can be produced by large variations of the focal length.

In this thesis, the silhouette coherence technique is used for camera calibration where all the information in the silhouettes is exploited to recover the camera poses and focal length. This technique extends the previous silhouette-based approach and can deal with partial or truncated silhouette and provide more generic method to compute contour coherence under circular motion. This algorithm can be used in situation where outer epipolar tangents or point correspondences are unavailable [10]. A known 3D model corresponding to the real object is not required as it is implicitly reconstructed from the silhouettes at the same time as the camera calibration by a visual hull method.

Voxel represents a value on a regular grid in 3D space. In this thesis the voxel array is created in a regular 3D grid ready to carve away. The process of adding depth to an image using a set of cross-sectional images known as volumetric dataset. Cross-sectional images or slice are made up of pixels where the space between two pixels in one slice is represents a real-world distance (interpixel distance). While the distance of any two slices represents a real-world depth (interslice distance). Opacity transformation gives voxels different opacity values. This is important to expose interior details of an image of a 3D image model.

Space carving method using an algorithm to reconstruct the 3D model from multiple images. This method is used instead of dense 3D reconstruction which is extremely slow and tedious. Dense 3D reconstruction also require dense correspondences and unrealistic assumptions such as pose, lighting and scene geometry [16]. Space carving algorithm is based on the volume intersection idea. The visual hull produced from intersection of silhouette cones is projected onto the 3D grid. Space carving algorithm will carve every voxels that lie outside the mask, leaving only the voxels inside the mask. Thus, the left over is the 3D model of the real object.

# CHAPTER III

# METHODOLOGY

## 3.1 Introduction

In this chapter, the methodology includes the theory of camera calibration, image acquisition and space carving for 3D image reconstruction. All the information and theories regarding 3D image reconstruction steps are based on related journals and also through previous project.

## 3.2 Simulation Using Matlab Software

The 3D image reconstruction is done by using the image reconstruction toolbox of MATLAB. This toolbox consists of collection of open source algorithm for image reconstruction written in Mathwork's Matlab language. The algorithm is used for image acquisition, camera calibration and also space carving.