Silhouette-based approach of 3D image reconstruction for automated image acquisition using robotic arm

N. Azhar, W. H. M. Saad*, N. A. Manap, N. M. Saad and A. R. Syafeeza
Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia

E-mail: *wira_yugi@utem.edu.my

Abstract. This study presents the approach of 3D image reconstruction using an autonomous robotic arm for the image acquisition process. A low cost of the automated imaging platform is created using a pair of G15 servo motor connected in series to an Arduino UNO as a main microcontroller. Two sets of sequential images were obtained using different projection angle of the camera. The silhouette-based approach is used in this study for 3D reconstruction from the sequential images captured from several different angles of the object. Other than that, an analysis based on the effect of different number of sequential images on the accuracy of 3D model reconstruction was also carried out with a fixed projection angle of the camera. The effecting elements in the 3D reconstruction are discussed and the overall result of the analysis is concluded according to the prototype of imaging platform.

1. Introduction
Three-dimensional (3D) reconstruction is a process to build the 3D image by capturing the surface shape of real object [1]. It has different modalities in order to create 3D models for example, using computer vision, radio waves and magnetic field, high frequency sound waves and gamma rays. The 3D reconstruction process can be accomplished either by active techniques or passive techniques [2]. It can be further classified as uni-directional and multi-directional. With active techniques the light sources are specially controlled as part of the strategy to arrive at the 3D information. Unlike active techniques, passive techniques don’t need specially controlled light and can work with ambient light available and also require simple instrumentation. From the computational point of view, active techniques are less demanding compared to the passive techniques as their applicability is restricted to the environment where the special illumination techniques can be applied. In this study the 3D reconstruction using computer vision is done by using passive techniques for multi-directional points [3]. A few sets of sequential images of objects are taken by using a robotic arm attached to a camera surrounding the object.

There is a lot of methods have been proposed in the robotic arm control system. A simple way to design the robotic arm, which can pick up, the object is using a freedom servo motor [4]. Microcontroller (ATMEGA-328) is used as the hardware interface platform to control the motion of robotic arm [5]. The potentiometer is fitted with the remote in the design while the electric pulse is produced by remote motion of the potentiometer. Servo motor will receive the digital pulse after the signal is converted by the Arduino board. In this study, the robotic arm that capable to handle the medium weight of the object is implemented. DC servo motor is an important device which is widely used in many applications. In robotic application, servo motor normally has priority to use in moving the robotic arm to a relevant position in the application which require high-speed control accuracy and high performance dynamic respond [6].
In this study, the shape from silhouette technique is used for the reconstruction of 3D shape. The silhouettes are computed for every 2D image captured surrounding the object. The computed silhouettes of every image along the center of the corresponding camera are used to define a volume. A back projected volume in 3D space can be assumed to bind the object. The intersection of these volumes associated with the set of acquired images yields a reasonable approximation of the real object. The intersection volume is known as visual hull and described as the maximal object that gives the same silhouette with the real object from any possible viewpoint [7].

2. Research Methodology
The methodology in this study is divided into three parts, image acquisition, camera calibration and 3D reconstruction.

2.1. Image Acquisition
In the block diagram of Figure 1 Arduino UNO microcontroller and G15 Shield are used to control the servo motor in order to rotate the camera. For motion system, two Cytron servo motors and two U-shaped holders are connected in series, which can rotate accordingly in a time. A USB webcam with the resolution of 0.3 megapixels act as the recognition system which is used to capture the image surrounding the object (Figure 2(a)).

![Figure 1. Hardware block diagram.](image)

![Figure 2. (a) Project subsystem and (b) The robotic arm.](image)

The sequence of images of the checkerboard pattern and the object is taken by using the autonomous robotic arm attached with a USB webcam. The checkerboard pattern and the object are placed on the platform and the autonomous robotic arm will capture the image surrounding it (Figure 2 (b)). The angular position of the robotic arm is initialized to zero degree and angle interval required which is 5° and 10°.
2.2. *Camera Calibration*

Camera calibration is a process to recover the camera extrinsic and intrinsic parameters that can affect the imaging process. The checkerboard pattern (Figure 3) is used for camera calibration due to its robustness and easy to implement. The alternately bright and dark grids as well as the grid corner features can be a very strong feature to detect and recognize. The Matlab Camera Calibrator app is used to estimate camera extrinsic, intrinsic and lens distortion parameters for a single camera. These camera parameters will be used to remove lens distortion effect from an image, measure the planar objects, reconstruct 3D scenes from multiple cameras, and perform other computer vision application.

![Figure 3. Checkerboard pattern used for camera calibration.](image)

2.3. *3D Reconstruction*

Space carving is a simple and a fast technique to recover the 3D structure from multiple images. The 3D shape is reconstructed by removing the background scene which are not consistent with the input images [8]. This process can be summaries of the following steps: load the camera and image data, convert the image into silhouette, create a voxel array, and carve the voxels using the camera images. In this method, a structure of voxels is projected into the 2D views of the scene and carved if the voxels lie outside the silhouette [8].

The sequential images of objects are converted into binary images before proceeding to the voxel creation. The process is done using a threshold conversion technique based on the color threshold values as shown in Figure 4. However, this technique faced a problem when the color of the object and platform are almost similar. This problem affects the accuracy of the silhouette image. Thus the RGB images are converted into HSV images before applying the threshold technique. Since R, G, and B components in digital image are correlated with the amount of the light hitting the object, image descriptions in terms of hue/lightness/saturation or hue/lightness/chroma are more relevant.

Silhouette image will be projected onto a regular grid in 3-dimensional space called a voxel. Along with the camera viewing parameters, the silhouette defines a back-projected cone that contains the actual object. This cone is called silhouette cones. The intersection of the cones is called a visual hull which is a bounding geometry of the actual 3D object [7].

![Figure 4. Threshold value of colour.](image)
3. 3D Reconstruction Result and Analysis

3.1. Different image sequence

Figure 5 shows the final 3D model by using an increasing number of silhouette images, 36 and 72 images. In terms of shape produced, there is not much different with a higher or lower number of images used. However, the surface of the 3D model in figure 5 (b) is much smoother compared to the 3D model in Figure 5 (a). According to [9], the higher number of image sequence holds more information for the 3D reconstruction thus it will produce more accurate final 3D model.

![Figure 5. Final 3D model using (a) 36 image sequence, (b) 72 image sequence using an autonomous robotic arm.](image)

3.2. Different camera position

The position of the camera in Position 1 is closer to the object compared in Position 2. In terms of shape, the final 3D model shape in Figure 6 (b) is closer to the real object compared to 3D model Position 1 in Figure 6 (a). The deformation cone at the bottom edges of the final 3D model in Figure 6 (b) is smaller compared to the result in Figure 6 (a).

![Figure 6. Final 3D model of (a) Position 1, (b) Position 2.](image)

Based on the analysis of camera calibration, it can be concluded that the higher the projection angle of the camera, the smaller the reprojection error. However, both Position 1 and Position 2 still produce a deformation cone like shape at the bottom of the 3D model produced. This is because the higher the camera position of the object causing the silhouette cones projected through the silhouette higher as well. The shape of the hidden parts of the camera will be estimated thus the intersection of the silhouette cones result in poor 3D reconstruction model.

4. Conclusion

The silhouette-based approach has been successfully tested for the sequence of object image taken by using the autonomous robotic arm. Other than that, the constructed autonomous robotic arm is successfully working to capture the sequence of object image surrounding the object for a different angle interval as well as different camera position. However, the 3D models produced are not accurately constructed as the original object as the models having deformation shape formed at the bottom of the 3D models. The positioning of the camera, the number of 2D image sequence, background colour and calibration technique can significantly influence the computed model. Apart from that, the technique of 3D reconstruction using Shape-from-Silhouette is prone to error. Other
than time consuming steps, the silhouette calculation is relatively sensitive for error like noise or wrong calibration.

**Acknowledgement**

The authors would like to thank to the Rehabilitation Engineering & Assistive Technology (REAT) research group under Center of Robotics & Industrial Automation (CeRIA) and Machine Learning & Signal Processing (MLSP) research group under Center for Telecommunication Research and Innovation (CeTRI) of Universiti Teknikal Malaysia Melaka (UTeM), Faculty of Electronics and Computer Engineering (FKEKK), Universiti Teknikal Malaysia Melaka (UTeM) for the funding provided under ‘UTeM Zamalah Scheme’ and Ministry of Higher Education (MOHE), Malaysia for sponsoring this work under project RAGS/1/2014/ICT06/FKEKK/B00065 and the use of the existing facilities to complete this project.

**References**

[1] Notano L G and Minghim R 2001 *Computer Graphic Forum* **20** 161-174
[2] Bianco G, Gallo A 2011 *3D-ARCH 2011*
[3] Moons T and Gool L V 2009 *Found. and Trends in Comp. Graphic and Vision* **4** 287-404
[4] Cytron Technologies, G15 Servo Motor User’s manual. Retrieved from [www.cytron.com.my](http://www.cytron.com.my)
[5] Krishna A R and Bala G S 2012 *Int. J. of Eng. Res. and Applic.* **2** 3098-3103
[6] Perera K N D, Fernando S R M. et al 2013 *Computer controlled DC servo motor, Working paper* 1-3
[7] Laurentini A 2014 *IEEE Transactions on pattern analysis and machine intelligence* **16** 150-162
[8] Perez M C and Casamitjana M 2009 *Lect. Notes in Computer Science (LNCS)* **5524** 112-119
[9] Fredriksson L 2011 *Bachelor Thesis*, University of Gavle, Sweden