3D Model Construction of Turntable and Robotic Arm Sequence

Norliana Binti Azhar[1], Wira Hidayat Bin Mohd Saad[1], Saiful Anuar bin Abd. Karim[2], Norhashimah binti Mohd Saad[1] and Nurulfajar bin Abd Manap[1]
[1]*Faculty of Electronic and Computer Engineering, Universiti Teknikal Malaysia Melaka, Durian Tunggal, Melaka, Malaysia*
[2]*Department of Electrical Engineering, Politeknik Ibrahim Sultan, 81700 Pasir Gudang. Johor, Malaysia*
wira_yugi@utem.edu.my

***Abstract*—This study is done to evaluate the factors affecting the 3D reconstruction model by using a different image acquisition technique; an autonomous rotatable robotic arm and a turntable. 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 of several different angles of the object. Other than that, an analysis based on the effect of a different number of sequential images to 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. It shows that, by using a robotic arm resulting with better final image reconstruction.**

***Index Terms*— Autonomous robotic arm, turntable, image acquisition, silhouette-based approach, 3D model reconstruction and accuracy.**

I. INTRODUCTION

The use of computer visualization technology today has ever-expanding applications in science, education, engineering, interactive multimedia, medicine, and etc. The major visualization problem in many application fields is the construction of three-dimensional (3D) models from two-dimensional data (2D). Three-dimensional (3D) reconstruction is a process to build the 3D image by capturing the surface shape of real objects [1]. In this study, multiple calibrated camera views are used for 3D reconstruction from a 2D image. There are many techniques introduced in previous studies to capture the sequence of the 2D images. Some of the techniques include using a turntable, stereo vision camera and non-calibrated camera.

Turntable method has been numerously used in graphic and computer vision to compute the 3D reconstruction models from multiple views. Most authors use projective geometry properties and multi-views relations to perform the 3D reconstruction [2]. Basically, for this technique, the real object is placed on the turntable and precisely rotated in front of a stationary camera. However the application of this system is restricted by several issues. The camera is put in a fixed position with respect to the rotational plane of the turntable. The camera movement is restricted and can only focus on the fixed camera view.

The 3D shape extraction technique can be classified into two major types, active and passive method. The active method generally employs structure illumination which is not desirable in many applications due to the restriction of the environment where special illumination technique can be applied. On the other hand, with the passive method, the light is not controlled or only with respect to the image quality and have the advantage of simplicity and applicability since this technique requires simple instrumentation. However, the major issue with this technique is poor reconstruction quality due to the difficulty in finding the accurate correspondence between images [3]. In this study, the passive approach applies to the image acquisition technique.

The acquisition of the camera views system using autonomous robotic arm is an alternative to the turntable technique [4]. In this study, the robotic arm is used to capture the sequence of the 2D image of the real object. A camera is attached to the robotic arm to allow a free movement of the camera around the object and as an independent choice of the focus for the camera view. The approach of using a robotic arm literally increase the accuracy and efficiency of 3D reconstruction.

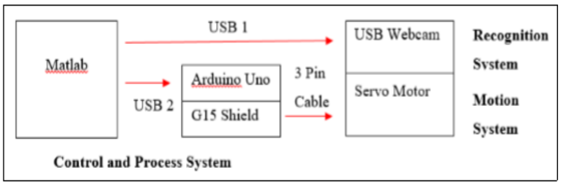
In this study, the shape from silhouette technique is used for the reconstruction of 3D shape. The silhouettes are computed for every 2D image captured surrounding the object. The computed silhouettes of every image along the center of the corresponding camera are used to define a volume. A back projected 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 [5].

II. METHOD

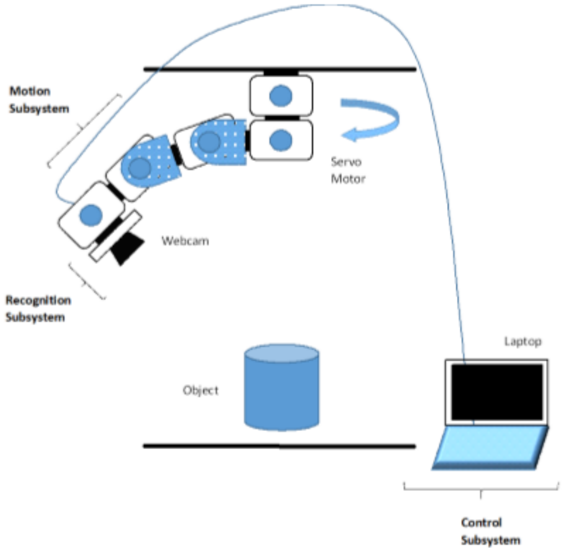
In this section, the methodology of this study is divided into three major parts, image acquisition, camera calibration and 3D reconstruction.

A. Image Acquisition

An autonomous robotic arm attached to the camera is used to capture the 2D image of the real object. The object is placed on the platform and the robotic arm will capture the image surrounding the object (Figure 3). The image is taken with an equal angle interval in between each interval. In control and process system as shown in Figure 1, the Arduino UNO microcontroller and G15 shield are used to control the servo motor in order to rotate the camera. For motion system, two servo motors and two U-shaped holders are connected in series which can rotate accordingly in time [6]. The USB webcam with a resolution of 0.3 Megapixels acts as the recognition system which is used to capture the surrounding images of the object [7] as shown in Figure 3. Matlab software is the control center of the overall system which is used to synchronize the image acquisition system and motion system by using the serial communication line.



(a)



(b)

Figure 1. (a) The hardware block diagram, (b) Project subsystem.

The flow chart in Figure 2 summarizes the whole process of image acquisition.

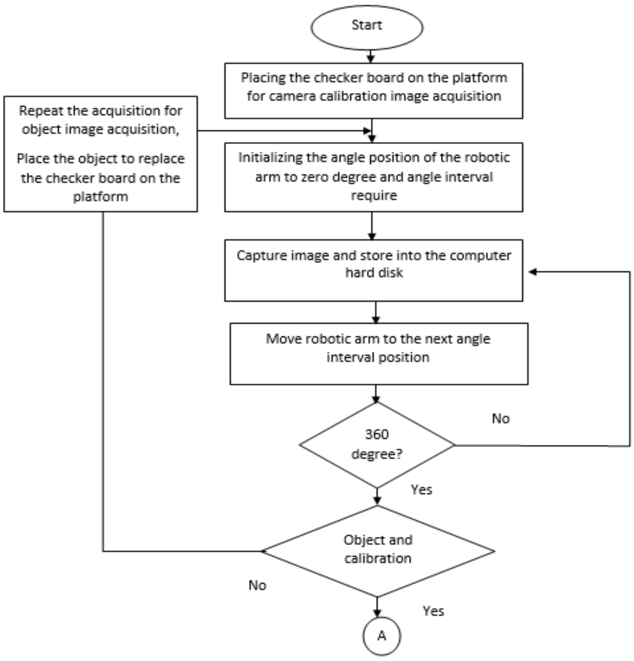


Figure 2: The flow chart of image acquisition process.

B. Camera Calibration

Camera calibration is an important step in 3D image reconstruction in order to recover the camera extrinsic and intrinsic parameters. The image of the checkerboard pattern is taken under a similar position as the object as shown in Figure 3. By using the camera calibration toolbox for MATLAB [8], four extreme corners of the calibration grid are manually identified in each image. It is used to identify features that are visible in all images and for which the coordinates are known [8]. The grid corners of the checkerboard images are then extracted and the tool will automatically count the number of squares in the pattern. After the corner extraction, the main calibration process is done to extract the intrinsic and extrinsic parameters of the camera. The flow chart in Figure 4 represents the steps of the calibration for the camera.

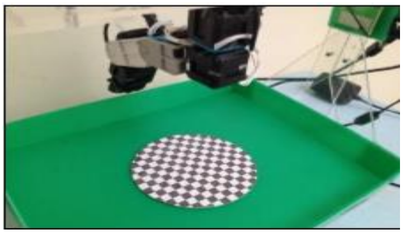


Figure 3: Camera calibration using checkerboard.

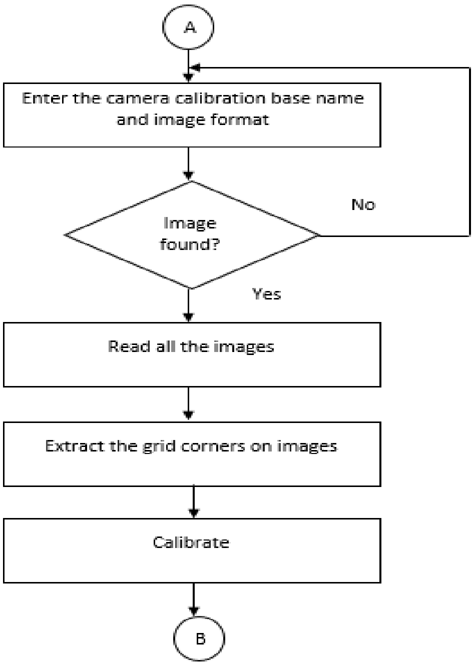


Figure 4: Flow chart of camera calibration.

Both turntable and robotic arm motion are in single axis motion where the relative motion between scene and camera can be described by rotations about a single fixed axis. In this case, there is a zero translation along the screw axis, which is used to represent 3D motion and describe the rotation about an axis and translation along an axis [9]. As the object is fixed, thus the camera internal parameters are also fixed.

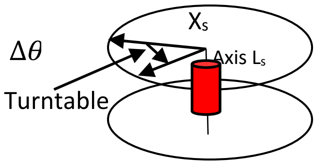


Figure 5: 3D geometry.

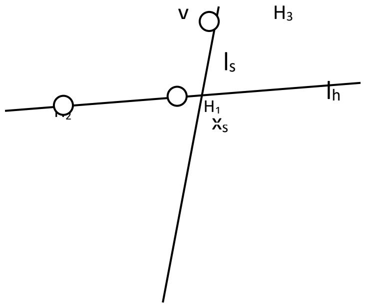


Figure 6: Fixed image entities over the sequence and their relation to the columns hi of H.

Referring to the Figure 5 and Figure 6, the line ls is the image of the rotation axis Ls. Since points on Ls are fixed under the motion, their images are also fixed under the motion. The line l_h in which πh intersects each image plane is the vanishing line of πh. The point xs is the image of the fixed point Xs and point v is the vanishing point of the rotation axis [9].

Camera matrix represents the 3x4 matrix which describes the mapping of a pinhole camera from 3D points in the world to 2D points in an image. For multi-view images, the first camera center is at position t on the X axis. It can be written:

$$P_0 = H\,[I \mid t] \quad (1)$$

where H is the homography representing the camera internal parameters and rotation about the camera centre, and t=(t,0,0)T . The operation of rotation of the camera by ɵ about the z axis can be represented as 4x4 matrices

$$\begin{bmatrix} Rz\,(ɵ) & 0 \\ 0^T & 1 \end{bmatrix} \quad (2)$$

yielding the camera Pɵ = H[Rz(ɵ) | t]. In detail, with hi the columns of H:

$$P_ɵ = [\,h_1 \quad h_2 \quad h_3] \begin{bmatrix} cos\theta & sin\theta & 0 & t \\ -sin\theta & cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (3)$$

this means that H and t are fixed over the sequence and only the angle of rotation, θi varies about the z axis for each camera Pi .

C. 3D Reconstruction

The process of recovering the 3D structure can be summaries by 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 as shown in Figure 7. 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 [10].

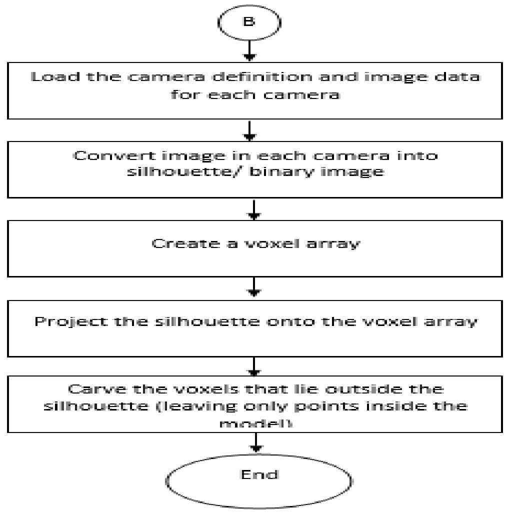


Figure 7: The flow chart of 3D reconstruction process.

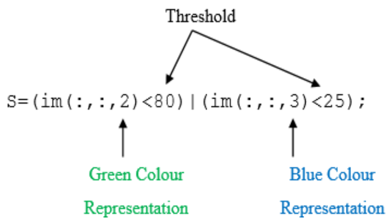


Figure 8: Threshold value of color.

The thresholding technique (Figure 8) is applied to the sequential images in order to convert those images into a binary image before it can proceed to the voxel creation. This technique is good for separating the object from the background. However, this technique faced a problem when the color of the object and the background platform are almost similar. Other than that, the reflection of light may cause the difficulty to determine the RGB color value. To compensate for this problem, the RGB images are converted into HSV images first before apply the thresholding technique.

For 3D reconstruction, the shape-from-silhouettes (SfS) technique is employed. Multiview silhouettes of the object are projected to a 3D space and their intersection is called visual hull which is a bounding geometry of the actual 3D object [11]. SfS is widely used for shape reconstruction because it can obtain dense and smooth surface points and silhouette extraction is easier in an experimental environment [12]. Despite the advantages, the visual hull might include false-positives in the concave regions and also regions that are hidden from camera views of an object shape (Figure 9). The false-positive or also called as ghost volumes can be reduced as the number of the camera grows, however, it is impossible to remove them in the concave and hidden regions of the object.

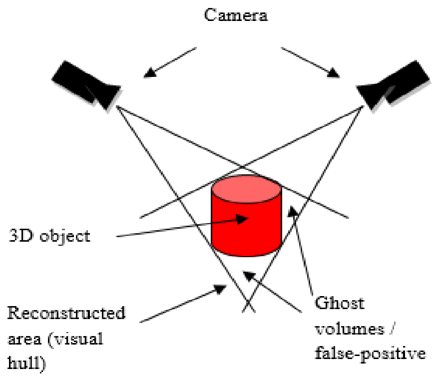
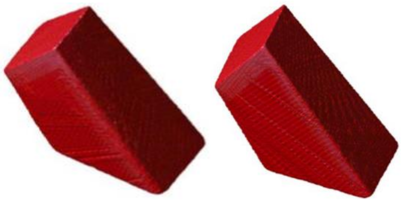


Figure 9: Shape from silhouettes

Space carving approach is used for refining a visual hull. The visual hull which is the initial shape is carved until photo-consistency is satisfied between multiple views.

III. RESULT AND ANALYSIS

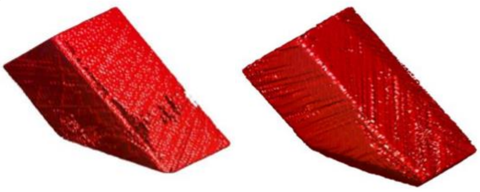
A. Autonomous Robotic Arm



(a) (b)

Figure 10: (a) Final 3D model using 36 images, (b) Final 3D model using 72 images using an autonomous robotic arm

The 3D models shown in Figure 10 are the final 3D model using 36 and 72 image sequence. In the term 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 10 (b) is much smoother compared to the 3D model in Figure 10 (a). According to [12], the higher number of images captured holds more information for the 3D reconstruction compared to the lower number of image sequence thus it will produce the final 3D model with higher accuracy.

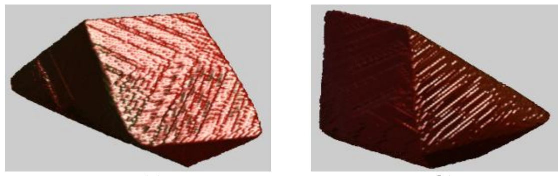


(a) (b)

Figure 11: Final 3D model of Position 1(a) and final 3D model of Position 2(b).

The position of the camera in Position 1 is put closer to the object compared in Position 2. In terms of shape, the final 3D model in Figure 11 (b) is closer to the real object compared to the 3D model Position 1 in Figure 11 (a). Despite closer to the real object, the deformation cone still formed at the bottom edges due to the projection of the silhouette cones.

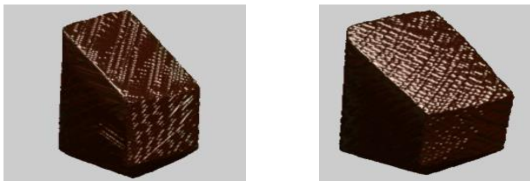
B. Turntable



(a) (b)

Figure 12: Final 3D model for (a) 36 image sequence and (b) 72 image sequence using turntable technique.

Figure 12 (a) shows the final 3D model using 36 image sequence and Figure 12 (b) shows the final 3D model using 72 image sequences by using the turntable. The final 3D model for both image sequences is not much different from the final 3D model by using the autonomous robotic arm. In Figure 12 (b), the 3D model surface is much smoother compared to the model in Figure 12 (a).



(a) (b)

Figure 13: 3D model for Position 1(a) and 3D model for Position 2(b).

As the position of the camera using the robotic arm, the camera in Position 1 (Figure 13(a)) is placed closer to the object compared to in Position 2 (Figure 13(b)). Both Position 1 and Position 2 still produce a deformation cone-like shape at the bottom of the 3D model produced.

From the result of the final 3D models in Figure 10, 11, 12 and 13 it can be concluded that the 3D model produced by using robotic arm is more accurate to the original object compared to using a turntable. The advantage of using a robotic arm over turntable is that the camera can freely move surrounding the object and the independence choice of the focus of the camera view [13]. The approach of using a robotic arm literally increase the accuracy and efficiency of 3D reconstruction. Despite using the robotic arm to capture the image, the final 3D models produced are still having the deformation cone-like shape as by using turntable method. The reason for this limitation is that the technique of 3D reconstruction using silhouette algorithm is insufficient to provide the information to compute the 3D model of non-convex objects. For the concavities problem in the object geometry, additional information must be provided otherwise it will result in self-occluded areas for the object that cannot be resolved from any viewpoint [14]. In this case the bottom part of the object where the deformation shape formed.

IV. CONCLUSION

In conclusion, the silhouette-based approach has been successfully tested for the sequence of the image surrounding the object taken by using an autonomous robotic arm and turntable. From the analysis done on the final 3D models, it can be concluded that the accuracy of the 3D reconstruction by using the autonomous robotic arm is rather higher compared to when using turntable technique. However, the 3D models produced by using a robotic arm still having the deformation shape formed at the bottom of the final 3D model. The positioning of the camera can significantly influence the computed model as well as the number of the 2D image sequence, background color and also the calibration technique. However, for the 3D reconstruction technique itself which using the Shape-From-Silhouette technique is prone to error. Apart from time consuming testing steps, the silhouette calculation is relatively sensitive for error like noise or wrong camera calibration. In future studies, the degree of freedom of the robotic arm will be increased in order to provide more accurate 2D information for better 3D reconstruction.

ACKNOWLEDGMENT

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

REFERENCES

[1] L .Notano, R.Minghim, M.Oliveira, "A Novel Approach for Delaunay 3D Reconstruction With a Comparative Analysis In The Light of Applications," Vol. 20, Issue 2, pp. 161-174, 2001.

[2] V.Fremont, R..Chellali, "Turntable-based 3D Object Reconstruction," Vol.2, pp. 1277-1282, 2004.

[3] T.Moons, L.V Gool, M.Vergauwen, "3D Reconstruction From Multiple Images," Vol. 4, Issue 4, pp. 287-398, 2009.

[4] J.Arnabat, S.Casanovas, G.Medioni, 3D Modeling from Turntable Sequences Using Dense Stereo Carving and Multiview Consistency, University of Southern California, 2004.

[5] A.Laurentini, "The Visual Hull Concept for Silhouette-Based image Understanding," IEEE Trans.PAMI, 2004, pp. 150-162.

[6] Cytron Technologies, G15 Servo Motor User's manual. 2012. Retrived from, www.cytron.com.my

[7] S.Izadi et al., "KinectFusion: Real-time 3D Reconstruction and Interaction Using a Moving Depth Webcam," UIST'11, USA, 2011.

[8] E.Carney ,Computational Modeling of the Helical Axis in the Human Knee-Detecting Changes in the Transverse Plane. MSc Thesis, Virginia Commonwealth Univ. U.S. 2009.

[9] A.Fitzgibbon, G.Cross, A.Zisserman, " Automatic 3D Model Construction for Turn-table Sequences," Proceedings of the European Workshop on 3D Structure from Multiple Images of Large-Scale Environments, pp.155-170, 1998.

[10] M.C Perez, M.Casamitjana, E.X Martin, "Space Carving Acceleration Using Uncertainty Measurements," Vol. 5524, pp. 112-119, 2009.

[11] A.Bottino, A.Laurentini, "The Visual Hull of Smooth Curved Objects," IEEE Trans. Patt. Anal. Mach. Intell, 2004,. 26(12),1622–1632.

[12] N.Ukita, K.Matsuda, "3D Reconstruction with Globally-Optimized Point Selection," IEICE Trans.INF.& SYST.,Vol. E95-D, NO.12, 2012.

[13] N.B Azhar, W.H.B.M Saad, N.B.A Manap, N.B.M Saad and A.R Syafeeza, "Evaluation of 3D Image Reconstruction Using Silhouette Algorithm," Research Journal of Applied Sciences, 11(11), pp.1195-1198, 2016.

[14] N. Azhar, W.H.M Saad, N.A., Manap ,N.M. Saad, and A.R Syafeeza, "Silhouette-based approach of 3D image reconstruction for automated image acquisition using robotic arm," IOP Conf.Series :Materials Science and Engineering, 210 012049, 2017.