Jurnal
Teknologi

Full Paper

3D SURFACE RECONSTRUCTION FROM A SINGLE UNCALIBRATED 2D IMAGE

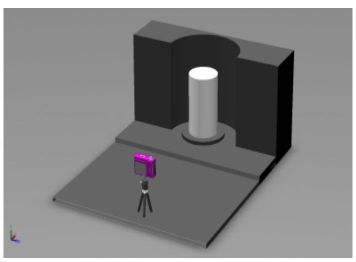
Siti Syazalina Mohd Sobani[a], Nasrul Humaimi Mahmood[b*], Nor Aini Zakaria[b]

[a]Faculty of Biosciences & Medical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia
[b]Faculty of Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Johor, Malaysia

Article history
Received
30 October 2015
Received in revised form
10 March 2016
Accepted
28 March 2016

*Corresponding author
nasrul@fke.utm.my

Graphical abstract



Abstract

This paper described a simple computation to reconstruct 3D surface using a single uncalibrated 2D image from a digital camera as an image acquisition device that also focused on fast processing. An object is placed on a table with black background for the digital camera to shoot an image of the object. Image segmentation methods are applied in order to obtain the shape of the object from silhouette. The concept Radon transform is adopted to generate sinograms of the object and it is then inverse Radon transform is used to construct 2D cross-section of the object layer by layer. Canny edge detection helps to get the outline of each cross-section and coordinate points are extracted forming 3D point cloud from the image slices. 3D surface of the object is then reconstructed using Delaunay triangulation to connect each point with another. The results obtained from this project are satisfying regarding the processing time with recognizable shape and also strengthened with considerably low percentage error in the calculation for all six objects used in the experiment.

Keywords: 3D reconstruction, uncalibrated image

Abstrak

Kertas kajian ini menghuraikan kaedah pengiraan yang ringkas untuk mengkonstruk semula permukaan tiga dimensi (3D) daripada satu imej dua dimensi (2D) dengan menggunakan sebuah kamera digital di mana pemprosesan yang pantas untuk keseluruhan kajian juga diberi tumpuan. Objek diletakkan di atas sebuah meja yang berlatar belakangkan dengan warna hitam untuk proses penggambaran dan mendapatkan imej objek tersebut. Beberapa kaedah imej segmentasi telah digunakan ke atas imej-imej tersebut bertujuan untuk memperoleh bentuk object daripada siluet. Konsep Radon transform membantu dalam menghasilkan sinogram objek dan Radon transform sonsang pula akan menghasilkan keratan rentas 2D objek lapisan demi lapisan. Pengesanan sisi Canny seterusnya mengekstrak titik-titik kordinat daripada imej-imej keratan rentas yang diperoleh dan membentuk kumpulan data 3D. Kemudian, Delaunay triangulasi akan menyambung titik-titik yang wujud dan membina permukaan 3D objek. Hasil kajian yang diperoleh adalah memuaskan dari segi masa pemprosesan yang pendek untuk membina permukaan 3D yang dapat dikenalpasti dan juga dikukuhkan lagi dengan peratus ralat yang rendah dalam pengiraan untuk kesemua enam objek-objek yang digunakan dalam kajian ini.

Kata kunci: Pembinaan semula 3D, imej tanpa penentukuran

1.0 INTRODUCTION

3D reconstruction is a process of creating shape, surface and texture of an object in virtual 3D space. There are many alternatives available to reconstruct 3D images such as scanning by utilizing laser scanner, range sensor, or multiple-camera system which require highly precise calibrations, expensive setup devices and equipment, plus time consuming [1, 2]. Reconstructing 3D object from 2D image is a challenging task in solving many stages of problems in order to obtain a desired result since it is ill-posed problem [3-5]. In recent years, 3D reconstruction has been sophisticated that some of the research can even reconstruct 3D image from uncalibrated image or video sequences [6]. However, this method still incomplete and has weaknesses such as accuracy and consistency in the reconstruction that must be improve and also overcome for any possible constraints.

This paper highlights fast processing to reconstruction 3D surface from a single 2D image with simple experimental set up. All 2D images are resized into smaller resolution to ease the computation and reduce the time consumption for the whole processing. Few image segmentation methods are applied in order to obtain the shape of the object from silhouette. The concept Radon transform is adopted to generate sinograms of the object and inverse Radon transform is then used to construct 2D cross-sections of the object layer by layer. Canny edge detection helps to get the outline of each cross-section and coordinate points are extracted forming 3D point cloud from the image slices. 3D surface of the object is then reconstructed using Delaunay triangulation to connect each point with another. Figure 1 shows an example of a reconstructed 3D model of an object involved in this project.

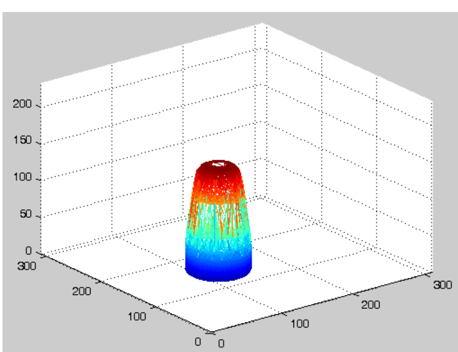


Figure 1 Reconstructed 3D surface of a model

The results obtained from this project are satisfying with recognizable shape and the whole processing time can be done in less than 5 minutes. Next section is focused on existing research related to 3D reconstruction from 2D images and uncalibrated camera methods. The whole processing for the proposed method is briefly described in the third section [7, 8]. Then, the reconstructed 3D surfaces and error calculation has been done is presented and in results and discussions section, followed with conclusion and future plan for this project.

2.0 RELATED RESEARCH

Many research have been conducted globally that are related to 3D imaging system and it is even possible to reconstruct 3D images without any calibration done to the input data. Basically, the procedure starts with a set of photos is shot around an object to be reconstructed which is vital for every project aiming to reconstruct 3D image from 2D images. But then, how to reconstruct the 3D image is depend on the proposed algorithm and it's efficiency. For example, L. Zhao *et al*. chose four coplanar points to get specific coordinates in both image and the real world object space. Then, an image is used as a reference image and warp together with the other images by using homography transformation. Layer by layer cross-sections of the object are obtained and arranged in 3D space forming the 3D object by subjected to the reference height [9]. Besides that, P. Kamencay *et al*. reconstruct 3D image from CT slices based on combination of the SURF (Speeded-Up Robust Features) descriptor and SSD (Sum of Squared Differences) matching algorithm using image segmentation. Any noise present in the image is filtered and it is split into segments using Mean - Shift segmentation algorithm. Corresponding points are then identified after edge detection with Canny operator is applied on the segments [10]. 3D reconstruction also can be done from shape from shading (SfS). Gu *et al.* [3] proposed a modified SfS algorithm for 3D reconstruction and Lambertain illumination model is used to obtain clearer and smooth silhouette 2D image. Edge detection is used to separate out the background from the object [3].

However, D. Liu *et al*. [11] reconstruct 3D object by a direct projective reconstruction obtained from two neighbouring images. This is because neighbouring images always share one common view. The corresponding points from the two images are combined and merged together with other corresponding points into projective reconstruction of 3D object. This method restricting that only two neighbouring images will be used to find the corresponding points in more than two images or computing the projective depths from fundamental matrices and epipoles [11]. Some of the research focused on cost saving, but simple and effective reconstruction like Ming *et al.* [12] have successfully reconstructed 3D face based on a 3D morphable face model from a single image and pose estimation which can automatically locate facial feature points. The pre-processing of 2D image involves illumination compensation, feature point extraction and face detection. 3D face shape is reconstructed with a linear combination according to the feature points and a 3D face database [12]. [13] have presented a new technique for simultaneous reconstruction of 3D articulated poses and camera parameters from 2D point correspondences obtained uncalibrated monocular images where the algorithm is formulated in a nonlinear optimization framework by maximizing the consistency between 2D point correspondences and the reconstructed 3D poses. The advantage of this proposed algorithm is that the system does not require someone to have deep knowledge on skeletal length or 3D poses and only needs at least five single view images to reconstruct 3D human pose of an unknown symmetric human skeletal model accurately from various sources of image such as photos uploaded in the internet and monocular video streams [13].

3.0 METHODOLOGY

Image acquisition involves a very simple setup since this project is done by using only one digital camera [14]. It does not require any sophisticated mechanical scanning system or expensive equipment which may be hard to setup. Figure 2 depicts the experimental setup for this project, in which a digital camera is mounted on a mini tripod stand aiming on about $605\,mm$ away from a cylinder shaped object. Black color background is used to create a better contrast scene which can bold up the object over the background in the image going to be captured. There are four different objects used in the experiment which are the models of amputated lower limbs at the knee as shown in Figure 3. The images are then transferred from the digital camera to the computer as input data to reconstruct 3D surface of the object.

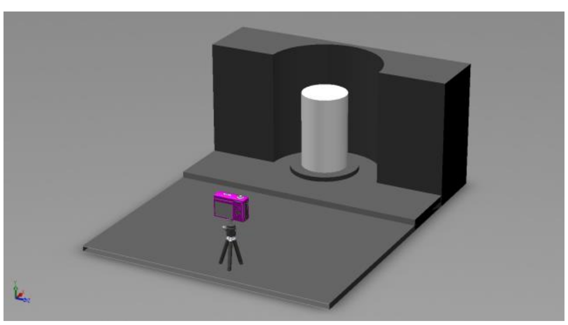
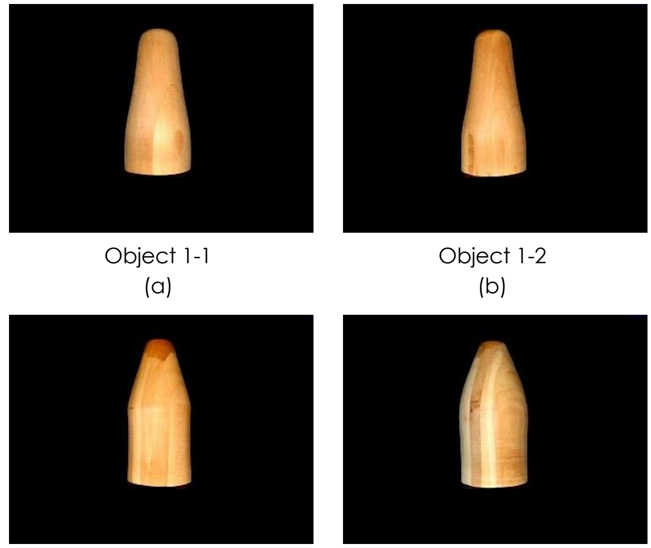


Figure 2 Experimental set up

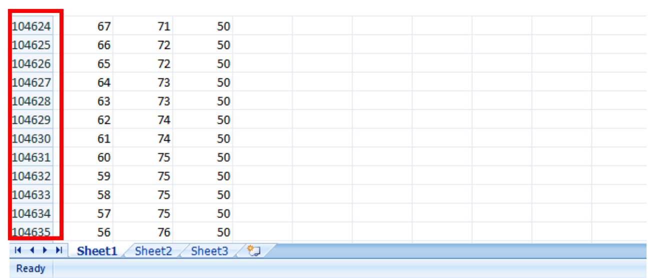
3D reconstruction is known as a process in creating shape, and surface of object that involves many stages of problems needed to be solved in order to obtain the desired result and suitable algorithm is applied on the 2D images to extract, match up, and estimate correspondence feature points to compose a 3D point. As for the first problem arise, 8-bit 20.1 Megapixels RGB image is considered as high resolution and computationally intensive. Plus, the whole reconstruction process is an iteration calculation and involved hundred thousands of data for an object Figure 4.



Object 1-1 (a) | Object 1-2 (b)

Object 2-1 (c) | Object 2-2 (d)

Figure 3 Four objects used in the experiment



104624	67	71	50
104625	66	72	50
104626	65	72	50
104627	64	73	50
104628	63	73	50
104629	62	74	50
104630	61	74	50
104631	60	75	50
104632	59	75	50
104633	58	75	50
104634	57	75	50
104635	56	76	50

Figure 4 Data of 3D point cloud generated

This study aims to use low cost equipment with relatively fast reconstruction, therefore the 2D images are resized from 5152 × 3864 pixels to 309 × 232 pixels for the purpose of easing the later tasks with faster computation. However, RGB image contains extra information that are not needed and cause the processing time increases exponentially with the three channels of red, green and blue colors, thus, segmentation of the silhouettes from the rest of the image approach is adopted in the processing. A silhouette displays an object as binary values in a 2D image where white represents a pixel occupied by an object and black is not occupied or simply the background. In order to generate the silhouette of an object, the resampled RGB image is converted to grayscale color and then Otsu thresholding is applied to extract the foreground and the background of the scene.

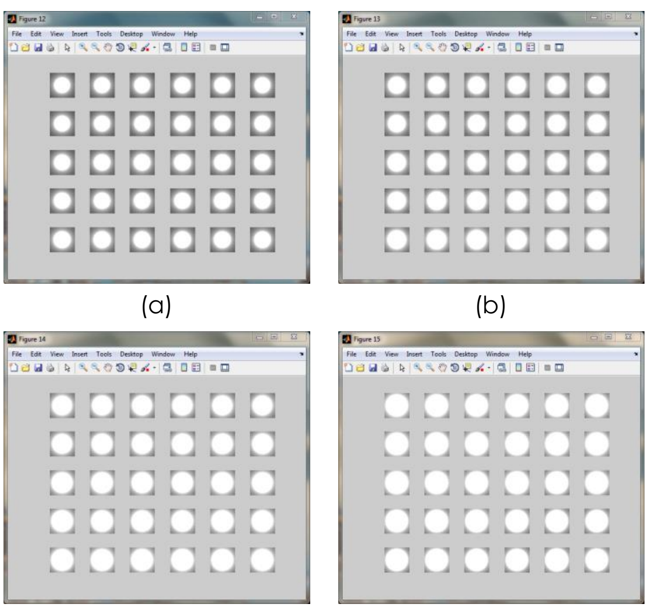
Radon Transform has been widely used in x-ray tomography imaging to produce sinogram that visualize the electron absorbed into the part of body that has been scanned [15]. In this study, sinogram of cylindrical based object are generated based on a single silhouette image for different angles of rotation since the view of the object is the same at any angle. Assume there is 36 projections, 309 × 232 pixels image is read row by row which means each row denote a single slice of the object to determine the position of the object at a certain slice. Each pixel either 0 or 1 in the specified row is substituted into R matrix iteratively by respecting to the 36 projections forming a rectangle matrix that also known as sinogram. The task is done based on the following line expression where rho is the shortest distance to the origin and theta is the angle of view.

$$\rho = x\cos\theta + y\sin\theta \tag{1}$$

Therefore, the Radon transform is a line integral for a set of parameters (ρ, θ) through an image $f(x, y)$. By using delta function, the relationship can be written as follows.

$$P(\rho, \theta) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\delta(\rho - x\cos\theta - y\sin\theta)\ dx\, dy \tag{2}$$

By utilizing the sinograms, inverse Radon transform based on parallel beam projection together with filtered back-projection algorithm to recover and reconstruct a cross-section image of the object layer by layer. (Figure 5) shows the examples of generated cross-section for Object 1-1 arranged slice by slice from Slice 31 up to Slice 150.



(a) (b) (c) (d)

Figure 5 2D slices of object (Object1) generated (a) Slice 31-60 (b) Slice 61-90 (c) Slice 91-120 and (d) Slice 121-150

From the output, edge detection is used to identify the edge of the cross-section, if only present in the specified slice, if none, means the slice only contain background. Edge detection is very useful in image recognition applications that it helps to identify points in digital images where the pixel intensities undergoes abrupt change or discontinuities that can provide strong visual clues on the shape and features of the scene in the image captured. The approach of edge detection used in this computation is the Canny operator, this is because the operator has better noise reduction in an image, therefore, a clearer boundary between the edge of object and the background can be obtained Figure 6.

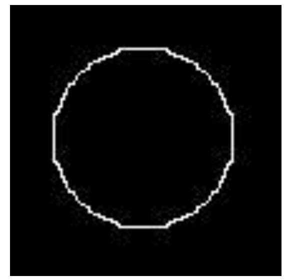


Figure 6 The cross-section image after edge detection is applied

The boundary of the edges are then traced and extracted which helps to specify the coordinates for every single point in the outline. All row and column coordinates obtained are recorded in terms of (x, y) points where z forming the layers and plotted in 3D space as shown in Figure 7 for Object1-1 Delaunay triangulation is used to create triangle and connect the x, y, z data with each other and construct 3D meshes. The results of the whole processing of 3D surface reconstruction for six different object used in the experiment are shown and discussed in the next section.

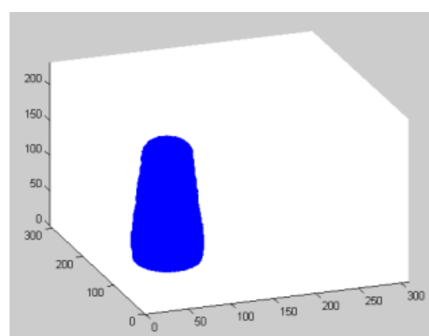


Figure 7 3D mapping of Object 1-1

4.0 EXPERIMENTAL RESULTS

The proposed algorithm has been evaluated on six different objects; Object 1-1, Object 1-2, Object 2-1, Object 2-2, Object 3-1, and Object 3-1 which are carved models of amputated lower limbs at the knee and they are the same at every angle of rotation like a cylinder. All objects are reconstructed and few calculations have been done on the height and diameters of the objects from the reconstructed 3D model to compare with the real measurement and identify the error. The calculation is done based on the mathematical concepts of right angle triangle (3) drew on the scene as shown in Figure 8 for the height where d is the distance between the camera lens and the centre of the object, and f is the focal length of the camera lens. The object's height can be determined as described in (4). Basically the reconstruction could simply being done using images captured by any digital camera and even phone camera. But it is important to know the specification of the digital camera used to capture the image in order to obtain the information on the size of camera sensor which that little information is not stated for phone camera and usually the buyer never care to know. The sensor's height, SH only involved in the calculation of the height while the sensor's width, SW only involved in the calculation of the diameter (width at the centre of the object).

$$\frac{Object\ height\ (mm)}{d\ (mm)} = \frac{Object\ height\ on\ sensor\ (mm)}{f\ (mm)} \tag{3}$$

Where,

$$Object\ height\ on\ sensor\ (mm) = \frac{SH\ (mm) \times Object\ height\ (pixel)}{SH\ (pixel)} \tag{4}$$

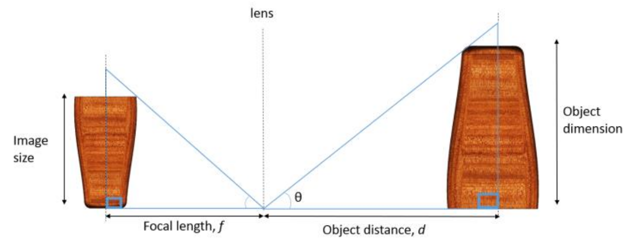
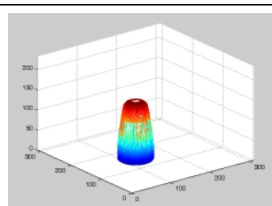
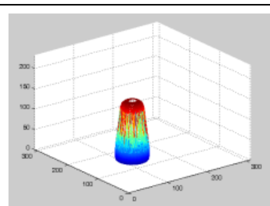
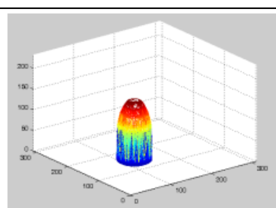
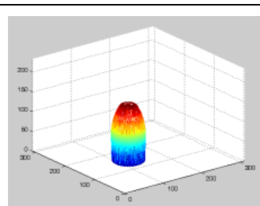


Figure 8 Two right angle of triangles

Table 1 describes the results of the reconstructions and the calculations have been done on the height and the diameters of the object models from the reconstructed 3D models. Few diameters are calculated at certain height where B1 is at the base of the object, H2 is at $\frac{1}{3}$ of the total height of the object, H3 is at $\frac{2}{3}$ of the total height of the object, and T4 is at the top of the object. The resuls are then compared and the percentage error is also calculated.

Table 1 The summary of results for all six models

Model	Reconstructed 3D Model	Position measured	Real Object measurement (mm)	3D Model measurement (mm)	Error (%)
1 - 1		B1	198	189.2826	4.4027
		H2	186	177.9821	4.3107
		H3	140	149.7310	6.9507
		T4	70	64.9776	7.1749
		Height	410	413.1752	0.7744
1 - 2		B1	192	182.8498	4.7657
		H2	183	174.4106	4.6937
		H3	135	126.5883	6.2309
		T4	67	61.8876	7.6304
		Height	410	411.4145	0.3450
2 - 1		B1	190	191.0162	0.5348
		H2	185	182.5899	1.3033
		H3	161	171.3528	6.4303
		T4	70	64.6084	7.7023
		Height	410	410.8267	0.2019
2 - 2		B1	198	194.928	1.6143
		H2	195	192.1077	1.4832
		H3	183	177.9821	2.7420
		T4	72	67.8027	5.8296
		Height	410	410.3834	0.0935

5.0 CONCLUSION

Multiple views approach is usually used to overcome the limitation of 2D image by gathering enough information of the object in the scene [16, 17]. However, from the images shown and the results obtained, it is clearly seen that the objective, which is to reconstruct 3D images surface of a simple object from a single uncalibrated 2D image, has been examined and successfully achieved. Time consumption for the entire processing is in range between two to three minutes, and the results obtained are also satisfying. The advantage of this project is that it can be done with simple experimental setup, as the devices used are usually owned by most people like phone camera unless it require to obtain exact measurement of the real object. Therefore, digital camera with thorough specification is used but still, it does not require any sophisticated mechanical scanning system or expensive equipment which may be hard to setup.

However, this project still has few weaknesses in the reconstruction accuracy for the irregular side curve shape of the model towards the top. This is because the 2D images used are resampled to low resolution to simplify the whole processing and lower the time consumption. A low resolution image provides less information on the curve of the object which means there is slight error on the measured diameter of the reconstructed 3D model. Same applies to the height of the 3D model that causes the inconsistence of the errors. The images are resampled so a low specification laptop is able to reconstruct 3D image without any problem in computing all the data involve. Furthermore, this system is focused more on the reconstruction for objects with cylindrical based design which it looks all the same at every angle of rotation that ease the processing without needing to gather too much of image to reconstruct the 3D surfaces. Thus, in order to reconstruct a model with shape such as a cube orcuboid correctly, taking more 2D images and adding more angles of views to collect enough information on the object or use video sequences is suggested for future research on this method.

Acknowledgement

The authors are grateful for a research grant from eScienceFund (Vote Number: 4S027) for this project, which has been provided by the Ministry of Science Technology and Innovation (MOSTI), as well as a scholarship from the Ministry of Higher Education of Malaysia.

References

[1] C. Yang, J. Chen, C. Xia, J. Liu and Guangda Su. 2013. A Sfm-Based Sparse to Dense 3D Face Reconstruction Method Robust to Feature Tracking Errors. *20th IEEE International Conference on Image Processing (ICIP)*. 3617-3621.

[2] C. Y. Ren, V. Prisacariu, D. Murray, and I. Reid. 2013. STAR3D: Simultaneous Tracking and Reconstruction of 3D Objects Using RGB-D Data. *IEEE International Conference on Computer Vision*. 1561-1568.

[3] Y. Gu, J. Yao, L. Li, and H. Wu. 2012. 3D Reconstruction From Single 2D Image Based On Silhouette Optimization. *2012 IEEE International Conference on Information Science and Technology*. 332-336.

[4] D. Rother and R. Vidal. 2011. A Hypothesize-And-Bound Algorithm For Simultaneous Object Classification, Pose Estimation And 3D Reconstruction From A Single 2D Image. *2011 IEEE International Conference on Computer Vision Workshops*. 553-560.

[5] B. Wandt, H. Ackermann, and B. Rosenhahn. 2015. 3D Human Motion Capture from Monocular Image Sequences. *2015 IEEE International Conference on Computer Vision and Pattern Recognition Workshops*. 1-8.

[6] L. Ling, I. S. Burnett, S. M. Lee, and E. Cheng. 2012. A Dense 3D Reconstruction Approach from Uncalibrated Video Sequences. *2012 IEEE International Conference on Multimedia and Expo Workshops.* 2-7.

[7] S. S. M. Sobani, S. A. Daud, and N. H. Mahmood. 2013. Implementation of 3D Object Reconstruction Using Radon Transform. *2013 1st International Conference on Artificial Intelligence, Modeling and Simulation*. 284-288.

[8] S. S. M. Sobani, N. H. Mahmood, N. A. Zakaria, and I. Ariffin. 2015. 3D Model Reconstruction from Multi-views of 2D Images using Radon Transform. *Jurnal Teknologi Special Edition.* 74(6): 21-26.

[9] L. Zhao, S. Liu, J. Li, H. Xu, S. Liu, and L. Zhao. 2010. An Approach for 3D Reconstruction from Uncalibrated Images. *2010 International Conference on Challenges in Environmental Science and Computer Engineering*. 390-393.

[10] P. Kamencay, M. Zachariasova, R. Hudec, M. Benco, and R. Radil. 2014. 3D image reconstruction from 2D CT slices. *2014 3DTV-Conference True Vision - Capture, Transmission and Display of 3D Video*. 1-4.

[11] D. Liu, X. Wang, S. Liu, and Q. Guan. 2010. A Practical Method for 3D Reconstruction with Uncalibrated Image Sequences. *2010 Chinese Conference on Pattern Recognition*. 1-5.

[12] Y. Ming, Q. Ruan. 2010. 3D Face Reconstruction using a Single 2D Face Image. *2010 International Conference on Educational and Information Technology.* 32-36.

[13] X. K. Wei, J. Chai .2009. Modeling 3D Human Poses from Uncalibrated Monocular Images. *2009 IEEE 12th International Conference on Computer Vision (ICCV)*. 1873-1880.

[14] T. Duckworth and D. J. Roberts. 2011. Camera Image Synchronization in Multiple-camera Real-Time 3D Reconstruction of Moving Humans. *IEEE/ACM 15th International Symposium on Distributed Simulation of Real Time Application.* 138-144.

[15] Z. Zakaria, N. H. Jaafar, N. Aqma, M. Yazid, M. Saiful, B. Mansor, and M. Hafiz. 2010. Sinogram Concept Approach in Image Reconstruction Algorithm of a Computed. *2010 International Conference on Computer Applications and Industrial Electronics.* 500-505.

[16] N.H. Mahmood, C. Omar, T. Tjahjadi. 2012. Multiview Reconstruction for Prosthetic Design. *The International Arab Journal of Information Technology*. 9(1): 49-55.

[17] N.H. Mahmood, C. Omar and I. Ariffin. 2011. Surface Reconstruction Using Reference Model For Future Prosthetic Design. *Jurnal Teknologi Special Edition*. 54: 165-179.