

3D Model Reconstruction from Multi-views of 2D Images using Radon Transform

Siti Syazalina Mohd. Sobani^a, Nasrul Humaimi Mahmood^{b*}, Nor Aini Zakaria^b, Ismail Ariffin^b

^aFaculty of Biosciences & Medical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia

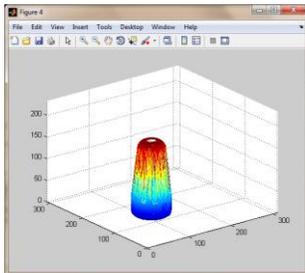
^bFaculty Electrical Engineering, Universiti Teknologi Malaysia, 81310 UTM Johor Bahru, Malaysia

*Corresponding author: nasrul@fke.utm.my

Article history

Received : 25 March 2015
Received in revised form :
11 April 2015
Accepted : 13 April 2015

Graphical abstract



Abstract

This paper presents a simple computation method to reconstruct 3-dimensional (3D) model from a sequence of 2-dimensional (2D) images using a multiple-view camera setup. The 3D model is acquired by applying several images processing on few 2D images captured by digital camera with different angle of views. The setup for this study consisted of a digital camera mounted on a tripod stand focusing at a block of model object on a turntable with black floor and background. 36 different angles are used to capture the images where every view angle differs by ten degree (10°) with another view in a fixed sequence. The image processing applied on all 2D images to be reconstructed as 3D surface are image segmentation, Radon transform (RT), image filtering, morphological operation, edge detection, and boundary extraction. The results for 3D model reconstruction shows it is well reconstructed, with a smooth texture obtained using 3D mesh and Delaunay triangulation, while the shape is nearly identical to the original model while the remaining are distinguishable.

Keywords: 3D reconstruction; multiple views; edge detection; image segmentation; Radon transform

Abstrak

Kertas kerja ini membentangkan kaedah pengiraan mudah untuk membina semula 3 dimensi model (3D) daripada urutan 2 dimensi (2D) imej dengan menggunakan susunan kamera berbilang pandangan. Model 3D diperolehi dengan menggunakan pemrosesan beberapa imej kepada beberapa imej 2D yang diambil oleh kamera digital dengan sudut pandangan yang berbeza. Persiapan kajian ini terdiri daripada kamera digital dipasang pada tripod pendirian memberi tumpuan di satu blok objek model pada meja boleh-pusing dengan lantai dan latar belakang berwarna hitam. 36 sudut yang berbeza digunakan untuk menangkap imej-imej di mana setiap sudut pandangan berbeza dengan sepuluh darjah (10°) dengan pandangan lain dalam turutan yang tetap. Pemrosesan imej digunakan pada semua imej 2D untuk dibina semula sebagai permukaan 3D adalah segmentasi imej, mengubah Radon (RT), penapisan imej, operasi morfologi, pengesanan sisi, dan pengekstrakan sempadan. Keputusan bagi pembinaan semula model 3D menunjukkan ia dibina semula dengan baik, dengan tekstur licin diperolehi dengan menggunakan mesh 3D dan *Delaunay* triangulasi, manakala bentuknya hampir sama dengan model asal manakala bakinya adalah boleh dibezakan.

Kata kunci: pembinaan semula 3D; pelbagai pandangan; pengesanan sisi; segmentasi imej; pengubahan Radon

© 2015 Penerbit UTM Press. All rights reserved.

1.0 INTRODUCTION

The development of 3D reconstruction researches has grown rapidly for applications such as computer vision, navigation system, computer-aided design (CAD), reverse engineering, virtual reality, site survey, and many more. It has been divided into two main methods for 3D reconstruction, which are the active and passive method. Active method basically interferes with the object on target to be reconstructed either mechanically or

radiometrically, while the passive method works in the opposite manner. 3D reconstruction based on active methods is the most effective approaches for accurate reconstruction. This involves 3D scanning by utilizing laser scanner, range sensor, or multiple-camera system which require highly precise calibrations, expensive setup devices and also time consuming [1, 2]. On the other hand, passive methods are capable of overcoming the weaknesses of active methods with slightly reduced of accuracy. One of the possible approaches for passive methods is

reconstruction from 2D images with multiple views of a targeted object or scene.

There are few important steps which need to be followed for 3D reconstruction from multiple views of 2D images, in which a sequence of images is shot on the target object or scene in different orientations or projections. It is suggested to set specific angle for each orientation. Then, suitable and matching algorithms designated for chosen approach are applied on the acquired 2D images to extract, match up, and estimate correspondence feature points that will form a bulk of coordinate points. The bulk of feature points obtained compose a 3D point cloud that contains information on 3D geometries and spatial location; however, the area of the surface and detail appearance is unknown but only shows approximate shape of the object. This problem is solved by creating 3D meshes, which is composed of connecting triangles from 3D point clouds. The Delaunay triangulation helps to construct surface in 3D space [3-5]. Figure 1 shows an example of a reconstructed 3D model of an object.

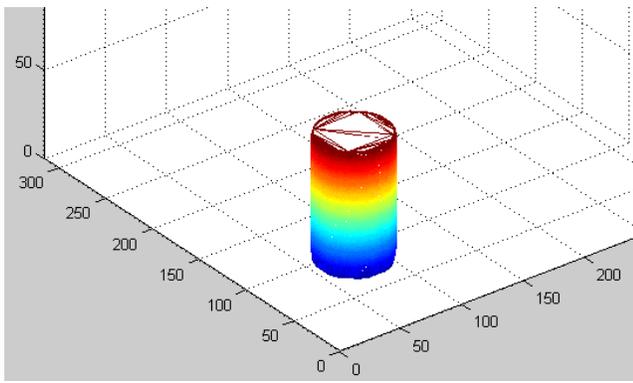


Figure 1 Reconstructed 3D surface of a model

This paper highlighting fast computing process to create 3D model reconstruction with simple experimental set up which can be made ready anytime easily. The fundamental of reconstructing 3D shape from 2D image is an ill-posed problem. A sufficient number of 2D images must be obtained, suitable algorithms need to be constructed and specified for this reconstruction process which digital camera and turntable setup for the data acquisition. For a starting point of target, measurement details of the models such as perimeter, area and volume are not being considered yet.

The results obtained from this project are strengthened with smooth surface 3D reconstructed model by using Delaunay triangulation. Next section is focused on literature reviews for this project based on researches has been conducted on 3D reconstruction from multiple views. Then, the 3D reconstruction for this study is described and explained how it was done in methodology section. The fourth section is for results and discussion, followed by conclusion and future plan for this study which is suitable with the chosen method.

2.0 RELATED RESEARCH

Research have been conducted globally and many are related to 3D imaging system. Nevertheless, all the outcomes acquired still has millions of chances to be upgraded, improve and also be used to any other suitable applications. Some examples of the research will follow. Motomura *et al.* [6] have investigated for probability and possibility to create tomographic 3D imaging using semiconductor. Both single and multiple semiconductors

Compton camera units set up are tested. The strength is highlighted where γ -ray source distribution is able to obtain in multiple directions of projection just by using a fixed-angle imaging with only a single Compton camera unit [6]. Mai and Hung [7] have conducted research to reconstruct 3D curves by exploitation of point along curves method. Multiple 2D images are taken by few uncalibrated cameras because the reconstruction does not require any 2D or 3D parameter and also additional information of point features. Even so, it is able to handle images with occluded and/or partially visible curves and can reconstruct both planar and non-planar type of curves [7]. Lee *et al.* [8] designed a system based on scene geometry by structure from motion (SfM) method to reconstruct 3D model. A single uncalibrated camera was used to capture multiple 2D images of model object for data acquisition. Iterative closest point algorithm helps register two target postures of the model object captured by patch-based multiple views stereopsis (PMVS) in order to acquire information on dense 3D points or point clouds [8].

Then, Radwan *et al.* [9] proposed an automatic approach to reconstruct 3D poses from a single 2D image. The algorithm is based on imposing both orientation and kinematic constraints in order to reduce ambiguity on motion and shading which is evaluated on difficult scenario of human poses and available datasets from public. The ambiguousness of a pose is caused by articulation of human body, cluttered backgrounds and also certain parts that hallucinated [9]. Paul *et al.* [10] presented a technique to estimate 3D motion and 3D coding to solve existing multiple views video coding (MVC) technologies problems. The problems are poor image quality, low interactive exploitation agility and inefficient computational time. The experiment was done by using multiple-cameras to capture a scene being studied from different angles of views and depth which helps to provide necessary interactive 3D spaces and able to improve distortion rate performance [10]. Alazawi *et al.* [11] focused on conversion of directional orthographic sets into perspective projection geometry perspective formula to reconstruct 3D scene from a holoscopic 3D image via multiple views extractions with high resolution viewpoint images (VPIs). This is done to overcome VPIs features which are poor in term of quality. As the outcome of the research, a high quality depth map can be built efficiently for natural 3D visualization without any eye-wear devices [11].

3.0 EXPERIMENTAL

Multiple views approach is used to overcome the limitation of passive methods that do not interfere with targeted object [12]. It is stated in previous section where highly concave shape cannot be reconstructed correctly, this is because the concave side of an object is a blind spot to the camera view. Compare with active methods the concave side of object can be determined or measure in term of distance and depth which helps in accuracy of the 3D reconstruction. In order to solve this problem, it is vital to get enough data and information of the object on the unseen sides from a single side of camera view and this is the reason why multiple views of 2D images are needed. Since this study aimed for low cost research, it is done by using a single camera only. Image acquisition is done by moving the object by using a turntable relative to that single camera [4].

Figure 2 shows the experimental setup for this study, in which a digital camera has been statically mounted on a mini tripod stand focusing on the targeted object on a turntable. Black background and floor has been chosen for a better contrast between object and background in the image going to be shot. Turntable based 3D object reconstruction is constrained to motion of axial rotation of the targeted object. However, it is possible to go outside the limit

of the constrained point if the axis of rotation does not go through the camera's optical center [1].

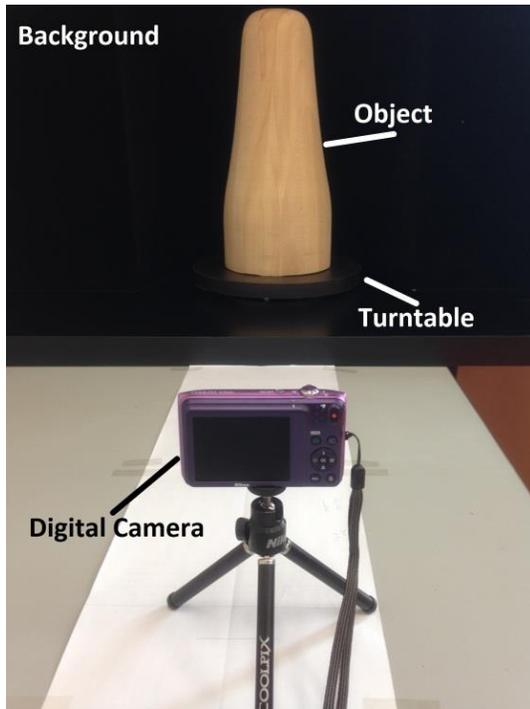


Figure 2 Experimental set up

The floor where the object is placed must not be made by any material that is shiny and can reflect light or the object placed on it, so it can save time and trouble to generate silhouette later. The turntable rotates along 360° with respect to counterclockwise. However, it is not a continuous or automatically rotating turntable. A reference mark is used to help on turntable rotation, where it should stop for the camera to take image of the object. There are three different objects used for the test which is Object1, Object2, and Object3 as shown in Figure 3. The object is a static rigid body made from blocks of wood. It is compact and heavy enough so it will not be easily move or overturn when turntable is rotated. All three objects are cylinder based shape then the sides are carved along 360° with possible curves and shapes amputated limb. It was decided to use 36 fixed spots where the turntable should be rotated and stop for the image acquisition in this study. The camera captures an image at every 10° angle starting from 0°, 10°, 20°, 30°, up to 360°, so the total number of images reaches 36. Then, the images from the camera will be transferred manually to computer in order to compute 3D reconstruction via USB cable or memory card reader. Each image will go through a few images processing steps, including generation of silhouette, computations of RT, edge detections, boundary extraction, 3D mapping, and Delaunay triangulation.

A. Generation of Silhouette

Silhouette is the image of the object as a single color solid shape subject, where the edges of the shape match with the outline of the subject and it is also featureless. The actual definition of silhouette is a solid shape subject presented as black or dark color with a light background and usually a white color. Since the silhouette color generated is the reversed and by considering the definition, this

silhouette is negative color. In order to generate the silhouette of an object, digital image thresholding is applied on the image.

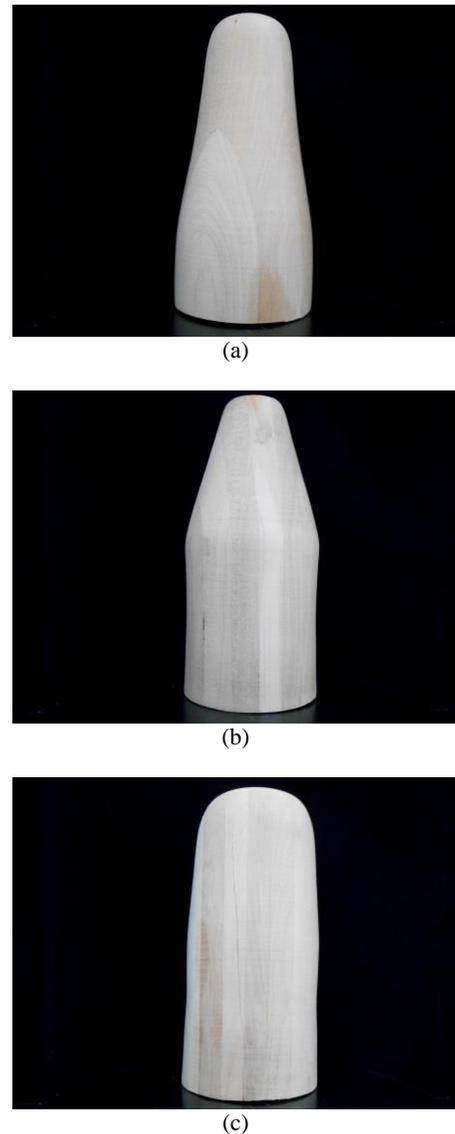


Figure 3 (a) Object1, (b) Object2, and (c) Object3

The image must be converted from red-green-blue (RGB) color to grayscale. Image analysis is done to extract intensities or color difference and determine grayscale level which is vital for thresholding. This is because image thresholding is dependent to the grayscale level, x stands for the reference value of luminance. The whole image will be ranged from 0 to 1 of the whole image. Equation 1 describes how thresholding works. Based on x value, the regions for foreground and background can be decided where foreground is the luminance greater than x and replace all the pixel values in the range with 1 while the other pixels left with 0 as the background [13].

$$\text{background} < x \leq \text{foreground} \quad (1)$$

where 0 = background (black)
1 = foreground (white)

As a result, the image will be in two colors, black and white which is also known as binary image. Figure 4 shows original image of Object1 and its nicely generated silhouette after thresholding is applied. The original image of Object1 is properly contrasted, and it is easy to differentiate between the object and the background. The object which is actually wood color turns to a white solid shape after thresholding. As defined, it is featureless and the edge of the shape is matching with the outline of the object. The faint reflection of the object on the floor in the real image has been converted to black, which is considered the background because the faint color is out of chosen x value range. All 36 images taken from image acquisition system will undergo thresholding one-by-one before computation of RT, because it helps to remove unwanted features and noise from the image.



Figure 4 (a) Original image, and (b) silhouette image

B. Computation of Radon Transform

Radon Transform has been widely used in tomography imaging based on cross-sectional scans of object and projection data. Mathematically, it is a type of linear transform and can be defined as a series of straight line integrals represented by $f(x,y)$. 2D image reconstruction is done by computing all integrals along the straight line parallel to y -axis at different offsets from the origin [14-17] as describe in Equation (2). RT is used to create sinogram for the object layer-by-layer based on all 36 silhouettes generated. Sinogram is output data from RT which is supported on sine wave graphs. It is mainly consisted of black-and-white data where the black region indicates zero and lighter regions indicate higher function values.

$$R = \int_{-\infty}^{\infty} (x - y_{min} + 1, y - x_{min} + 1) dy \quad (2)$$

Then, inverse radon transform associated with filtered-back projection is applied to form proper cross-section image and recover actual rotation which cannot be seen from 2D images acquired. The projection data obtained from all 36 different angle of view corresponding to the targeted object will form cross-sections of the object in layers from top to bottom. Figure 5 shows 2D slices represent the cross-sections of Object1 in layers from top to bottom. As can be seen in original image of Object1, it is small at the top and bigger down to bottom, and the same goes to the cross-sectional layers generated by computation of IRT as the circle shape is small as it is started, and grows bigger as it goes down the object.

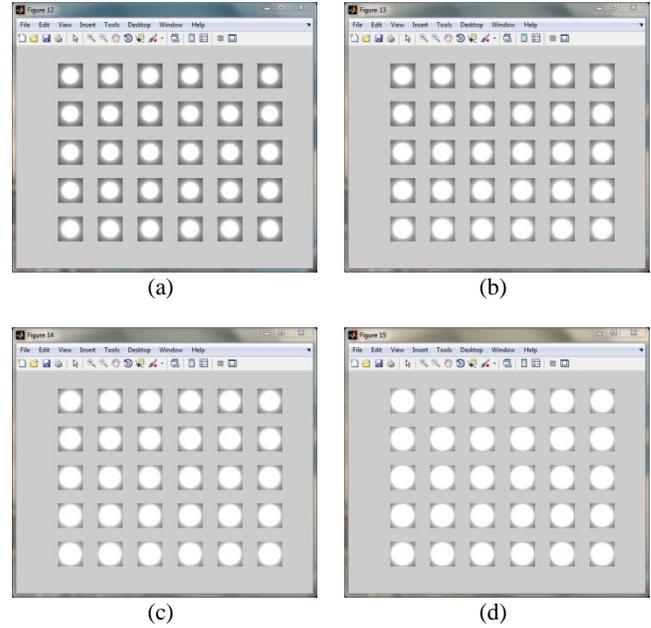


Figure 5 2D layers of object (Object1) generated using RT. (a) Layer1 (b) Layer 2 (c) Layer 3 and (d) Layer 4

C. Edge Detection and Boundary Extraction

Edge detection is very important in image recognition applications aiming to identify points in digital images where the pixel intensities and brightness undergo an abrupt change or has discontinuities. This method helps in extracting the shape and features in an image that can provide strong visual clues of the scene in the image captured. A simple edge model can be described per Equation (3) of a one-dimensional image (1D), f which is a binary image that has exactly one edge located at x . The edge is consisted of right and left sided with certain intensity values [18-19].

$$f(x) = \frac{I_r I_l}{2} \left(\operatorname{erf} \left(\frac{x}{\sqrt{2} \sigma} \right) + 1 \right) + I_l \quad (3)$$

where right edge, $I_r = \lim_{x \rightarrow \infty} f(x)$

left edge, $I_l = \lim_{x \rightarrow -\infty} f(x)$

Canny operator is chosen out of several existing edge detection techniques for example, Roberts operator, Prewitt operator, and Sobel operator. This is because Canny operator has better tolerance and good in terms of noise reduction in an image, so a clearer boundary between the edge of object and the background can be obtained.

Figure 6 shows binary image for the 30th layer of Object1 before and after edge detection is applied an outline of the round shape, Figure 6(b). Then, boundary tracing will help to specify the row and column coordinates of the points in the outline.

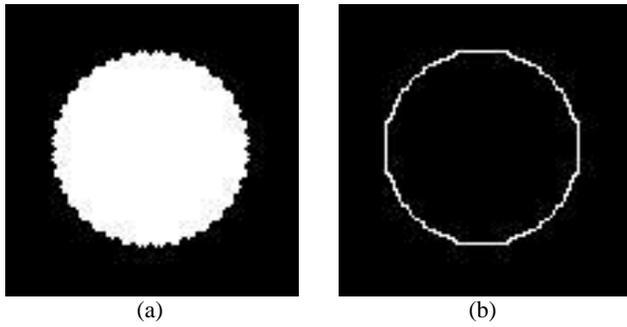


Figure 6 (a) Binary image, and (b) image after edge detection

D. 3D Mapping and Delaunay Triangulation

All row and column coordinates obtained are recorded in terms of (x, y) points where the number layers as z forming 3D point cloud. It will be plotted in 3D space as shown in (Figure 7a) together with three other views of 3D plotted Object1 which are xy -view, xz -view, and yz -view in Figure 7b, Figure 7c, and Figure 7d respectively. Then, Delaunay triangulation $DT(P)$ is computed using the x, y, z data such that no point in the 3D point cloud P is inside circumcircle of any triangle. In another important point, there is only one edge at a vertex and no intersection of edge will exist outside the set of vertices, V region in a set of edges, E . The results for this whole process of 3D model reconstruction are shown and discussed in the next section.

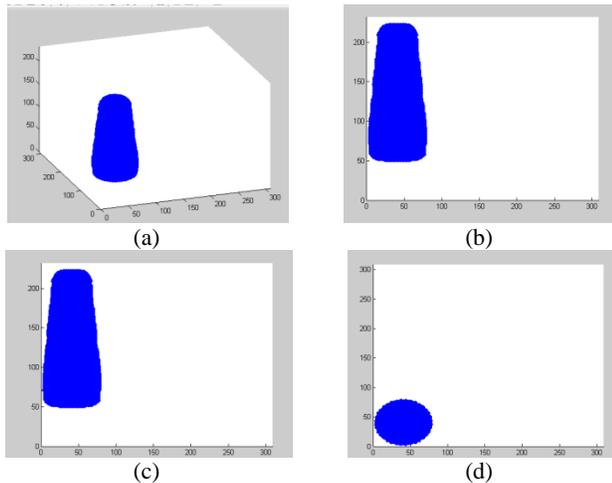


Figure 7 3D mapping of Object1

4.0 RESULTS & DISCUSSION

The entire process only takes around ten to fifteen minutes, starting from image acquisition to the 3D model reconstruction output and analysis. The device used for the reconstruction process is a laptop with Windows 7 operating system, 2.40 GHz Intel Core i5 processor including memory of 4GB RAM, and NVIDIA GeForce graphics card with CUDA. Figure 8 shows the images of the real models and results of 3D meshes for all three models being tested which are Object1, Object2 and Object3 respectively. The 3D meshes are constructed by computing the Delaunay triangulation together with 3D point cloud. From the images shown it is clearly seen that the reconstruction of the 3D models are successfully done.

The advantage of this proposed study is that it can be done with simple experimental setup, as the devices used are usually owned by most people and easily can be found. It does not require any sophisticated mechanical scanning system, which may be hard to handle and is slow. The time consumption to reconstruct 3D surface from the 2D images acquired is less than five minutes since the images taken are from an uncalibrated camera and does there are no complicated problems for the camera parameters. However, this study still has limitation in term of accuracy for the irregular side curve shape of the model. This is because the low resolution of the 2D images is used to avoid calculation problem during processing. Another limitation is that the system cannot reconstruct a model with shape such as a cube or cuboid correctly because the calculation is based on radius and it tends to form curve at the corner side of those models. This can be solved by taking more 2D images and adding more angles of views to collect enough data for that purpose, for future research on this method.

5.0 CONCLUSION

The fundamental of reconstructing 3D shape from 2D image is an ill-posed problem which causes the 3D reconstructed model less accurate. Therefore, a sufficient number of 2D images must be obtained at least 36 different angles views of the single object. Radon transform, 3D point clouds, Delaunay triangulation and few other digital image processing methods such as thresholding, edge detection, boundary extraction, are specified for this reconstruction process, which uses 2D images captured by digital camera and turntable setup as input data. The aim and objective, which is to reconstruct 3D images of simple object, has been examined and successfully achieved. Time consumption for the entire processing process is short enough (lower than fifteen minutes), and the results obtained are also considerably strong. Future works should find a better solution to overcome the weaknesses of the chosen method in order to obtain desired and more accurate result. An automatic image acquisition system is suggested to simplify the entire processing process and it can avoid any unintended movement on the object that will cause inaccuracy of the input data. It would also be good to measure the size of the reconstructed model, so it can be compared with the real object for better accuracy of the reconstruction.

Acknowledgement

The authors are grateful for a research grant from eScienceFund (Vote Number: 4S027) for this study, 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.

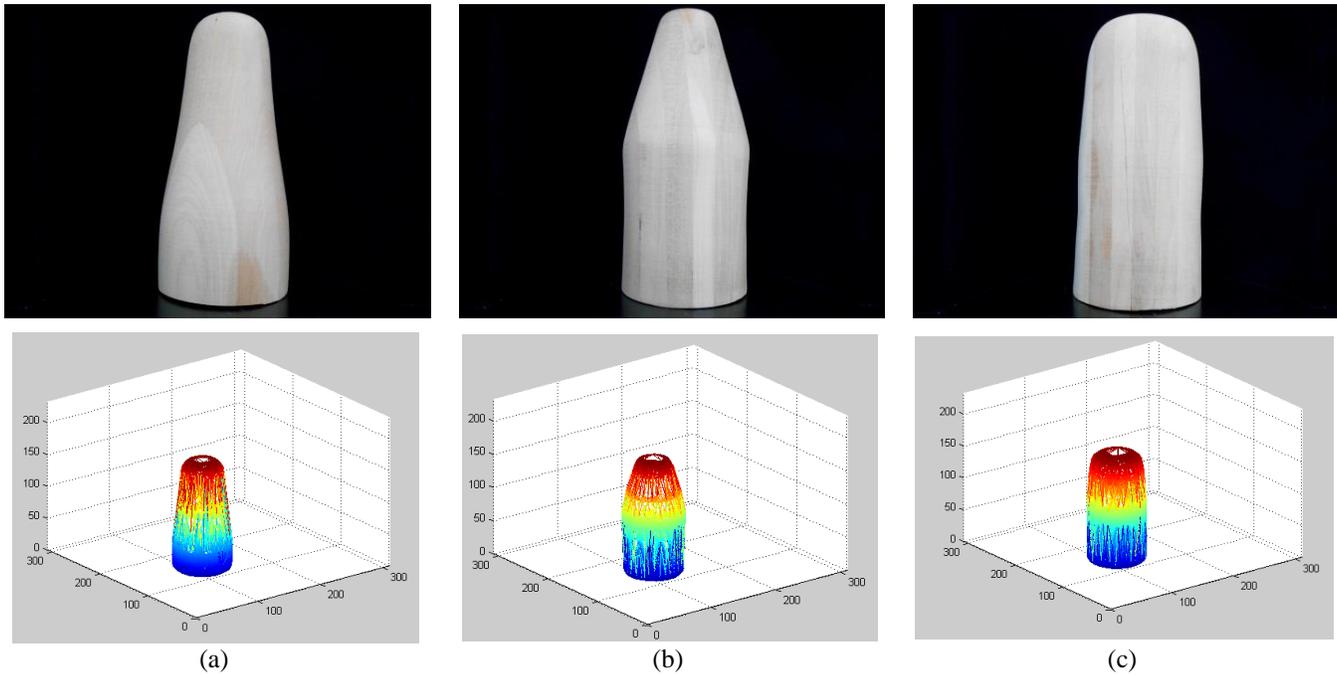


Figure 8 The original models and 3D meshes: (a) Object1, (b) Object2, and (c) Object3

References

- [1] C. Yang, J. Chen, C. Xia, J. Liu and G. 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] N. Wongwaen, S. Tiendee and C. Sinthanayothin. 2007. Method of 3D Mesh Reconstruction from Point Cloud Using Elementary Vector and Geometry Analysis, 8th International Conference on Information Science and Digital Content Technology (ICIDT). 1: 156–159.
- [4] 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.
- [5] C. Yang, F. Zhou and X. Bai. 2013. 3D Reconstruction through Measure Based Image Selection, 9th International Conference on Computer Intelligence Security. 377–381.
- [6] S. Motomura, T. Fukuchi, Y. Kanayama, H. Haba and Y. Watanabe. 2009. Three-Dimensional Tomographic Imaging by Semiconductor Compton Camera GREI for Multiple Molecular Simultaneous Imaging, IEEE Nuclear Science Symposium Conference Record (NSS/MIC). 3330–3332.
- [7] F. Mai and Y. S. Hung. 2012. Three-Dimensional Curve Reconstruction from Multiple Images. *IET Computer Vision*. 6(4): 273–284.
- [8] P.-H. Lee, J.-W. Huang and H.-Y. Lin. 2012. 3D model Reconstruction Based on Multiple View Image Capture. International Symposium on Intelligence Signal Processing and Communication System. 58–63.
- [9] I. Radwan, A. Dhall and R. Goecke. 2013. Monocular Image 3D Human Pose Estimation under Self-Occlusion, IEEE International Conference on Computer Vision. 1888–1895.
- [10] M. Paul, C. J. Evans, and M. Murshed. 2013. Disparity-Adjusted 3D Multi-View Video Coding with Dynamic Background Modelling 20th IEEE International Conference on Image Processing (ICIP). 1719–1723.
- [11] E. Alazawi, A. Aggoun, M. Abbod, M. R. Swash, O. A. Fatah and J. Fernandez. 2013. Scene Depth Extraction from Holographic Imaging Technology, *3DTV-Conference: The True Vision Capture, Transmission and Display of 3D Video (3DTV-CON)*. 1–4.
- [12] 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.
- [13] M. S. Nixon and A. S. Aguado. 2008. Feature Extraction & Image Processing, Second Edition, Academic Press, United Kingdom. 183–185.
- [14] S. Chandra and I. Svalbe. 2009. A Fast Number Theoretic Finite Radon Transform, *Digital Image Computing: Techniques and Applications (DICTA)*. 361–368.
- [15] V. Venkatraghavan, S. Rekha, J. Chatterjee and A. K. Ray. 2011. Modified Radon Transform for Texture Analysis, Annual IEEE of India Conference. 2: 1–4.
- [16] B. Kaur and M. K. Majumder. 2012. Novel VLSI Architecture for Two-Dimensional Radon Transform computations, 1st International Conference on Recent Advances in Information Technology. 570–575.
- [17] N.H.Mahmood, C. Omar, and T. Tjahjadi. 2012. Multiview Reconstruction for Prosthetic Design. *The International Arab Journal of Information Technology*. 9(1): 49–55.
- [18] B. Kaur and A. Garg. 2011. Mathematical morphological edge detection for remote sensing images, 3rd International Conference on Electronics Computer Technology (ICECT). 5: 324–327.
- [19] W. Wang and H. Xu. 2011. Edge detection of SAR images based on edge localization with optical images, 3rd International Asia-Pacific Conference on Synthetic Aperture Radar (APSAR). 2–5.