Depth Estimation and Shape Reconstruction of a 2D Image Using N.N. and Bézier Surface Interpolation

Luma Issa Abdul–Kreem
Control and Systems Eng. Dept, University of Technology, Baghdad, Iraq
e-mail: 60021@uotechnology.edu.iq

Abstract- *Inferring 3D image from 2D image is an advance topic in computer vision. This article considers a 2D image depth estimation of an object and reconstructs it into a 3D object image. The 2D image is defined by slices, where each slice contains a set of points that are located along the object contour and within the object body. The depths of these slices are estimated using the neural network technique (N.N.), where five factors (slice length, angle of the incident light and illumination of some of points that located along the 2D object, namely control points) are used as inputs to the network. The estimated depths of the slices are mapped into a 3D surface using the interpolation technique of the Bezier Spline surface. Our model was tested and evaluated using different objects with different and complex shapes. The results showed an effective performance of the proposed approach.*

Index Terms— *Object depth estimation; Bezier Spline surface interpolation; mapping 2D curved object into 3D curved object.*

I. INTRODUCTION

Object depth estimation from 2D image is one of the most active research topics because it is the basic problem in computer vision and it has important applications in robotics, pattern recognition, graphic and machine vision. 3D image construction is a challenging problem and many researches have been performed to resolve it. In [1] the authors have proposed a supervised learning approach for depth estimation using a 3D scanner to collect the training data that were used to model a conditional distribution of depths that was given a monocular image features. Other investigators [2] have developed a simultaneous phase-shifting technique using an innovating color fringe pattern with a multiple triangular modulation for 3D vision system.

Shape and depth from shading techniques for 3D surface reconstruction were presented in [3] and [4]. However, shading techniques are only valid to acquire object height information for the directions associated with the incident light and the generated object shadows.

Many works on 3D construction have focused on 2D images. In [5] the object was projected through a grid of pseudorandom encoded structured light to determine a set of reference pixels in two simultaneous views of the same object using two cameras. While in [6] the models of 3D objects that were obtained from 2D images are based on the human-computer interaction. The human was provided with much visual assistance as possible to make a correct input and verify it.

Constructing a 3D graphical image models was proposed in [7], where six-surface including front, back, left, right, top and bottom view images were used to form the model and the color matrixes.

Reconstruction method from contours lines was provided in [8]. In the model, Zhong and coauthors suggested to rid some of redundant points on every contour and interpolate them by using cubic Bézier spline curve. In [9] the authors, suggested to represent a geometric 3D shape as a probability distribution of binary variables on a 3D voxel grid, using a Convolutional Deep Belief Network. Generating a 3D image from a consecutive of 2D images has been investigated by [10]. 3D face reconstruction has been

Received 13 April 2017; Accepted 3 October 2017

 ISSN 1811-9212

introduced in [11], where the authors integrated variant face pose in order to reconstruct a 3D face model.

II. SUGGESTED MODEL

The suggested model is different from others by proposing a new methodology to estimate the depth of an object. The model takes advantage of light distribution over an object, where this distribution alters based on object depth as well as the angle of the incident light (see [12] for more details). In our model, a set of points is assigned in the 2D image that lays inside the object and along the contour. We referred to these points as control points. Neural network technique is used to estimate the depth, where object width illuminations and angle (θ) of the incident light are considered. The 2D object image is reconstructed into 3D object using Bézier surface. The following subsections introduce the suggested model in details.

A. Camera-Object Setup

In order to get 3D complex objects with different depth, AutoCAD software has been utilized to form various shapes with variant depth. The 2D images are captured by using camera with a target point light in which both are located perpendicular to the object at (150) cm. Fig. (1) represents a sketch of the camera-object setup structure whereas; L and Rd are the camera height and the depth of the object, respectively.

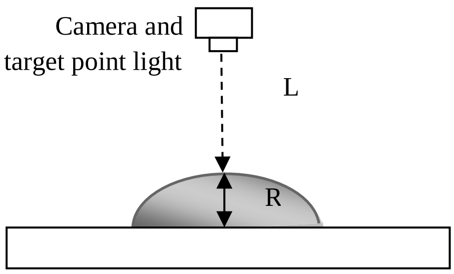

FIG. 1. THE IMPLEMENTING STRUCTURE

The resolution of the captured image is 640×480. A set of image preprocessing techniques is applied including: digital image denoising, thresholding and edge detection that were presented in [13], and the contour tracing to extract the edge as a set of connected pixels as presented in [14].

In this paper, a cylinder object with a radius of (15) cm and a height of (40) cm is used as an experimental sample to illustrate how to construct the 3-D object image through applying the proposed stages of the methodology, where Fig.s (2-a) and (2-b) represent the 2-D image of the used cylinder and its contour, respectively.

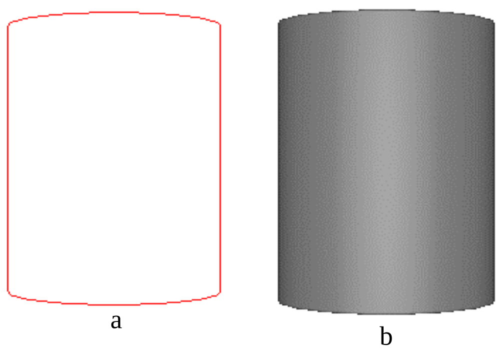

FIG. 2. CYLINDER OBJECT. (A) THE 2-D IMAGE (B) CONTOUR OF THE CYLINDER

Received 13 April 2017; Accepted 3 October 2017

ISSN 1811-9212

B. Depth Estimation Methodology

The proposed method is divided into three main parts, setting of control points, neural network technique, and construction of 3-D object via Bézier surface interpolation. These parts are described in the following.

1. Setting the Control Points

The control points setting is a substantial issue of determining the object depth, it consists from two steps. In the first step, the object contour is divided into halves, right and left. Fig. (3) shows the halves of the cylinder's contour, where each one includes two types of points, the edge and the halve control points. The edge control points are located along the object contour while the halve control points are placed in the middle of the halves. The spaces between the edge control points are equal to the spacing of the halve control points. Experimentally, (10) pixels are a convenient distance between the control points where it gives an accurate trace for object shape.

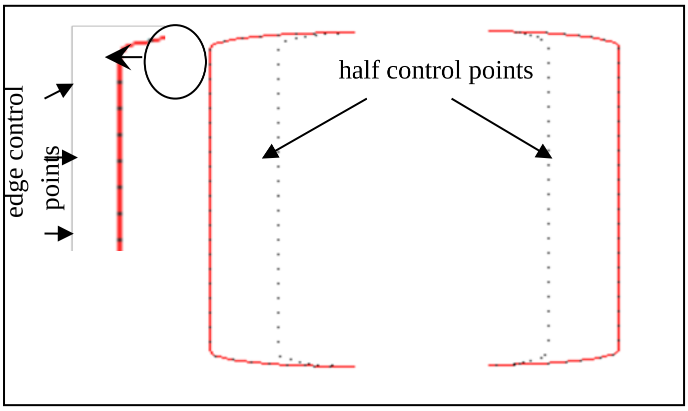

FIG.3. THE TWO HALVES CONTOUR OF THE CYLINDER

The right and left side of the object contour can be represented by two matrixes Rs and Ls respectively, where each element contains the values of (x,y) coordinates for the control points.

The right half side

$$R_s = \left[r_{s_{ij}}\right] \quad i: 1 \rightarrow n,\ j: 1 \rightarrow m. \tag{1}$$

The left half side

$$L_s = \left[l_{s_{ij}}\right] \quad i: 1 \rightarrow n,\ j: 1 \rightarrow m. \tag{2}$$

Where n is the number of the control points while m is the type number of these points and it equals to 2, edge and half. In the next step of the setting method, RS and LS are combined in one matrix TS as shown in Eq. (3).

$$T_S = [R_S \quad L_S]_{n\times k}\,. \tag{3}$$

k is the number of the whole types of the control points and it equals to 4.

In our method, a slice will be referred to any set of points that have the same level where each slice has four control points. These points are connected one to another to construct the slice shape, where Fig. (4) shows the slices shape of a cylinder that has (33) slices for length of (228) pixels.

The cylinder's ends are demonstrated with (12) slices while the cylinder's body is represented with (21) slices. Obviously, the slices that constituents the cylinder's body are in equal width while the slices that describe the cylinder's ends start with small width and grow into the body of the cylinder. On the other hand, the slices of the cylinder's ends are very close compared with the slices of the body; although

Received 13 April 2017; Accepted 3 October 2017

 ISSN 1811-9212

all the distances between the cylinder's slices are equal to 10 pixels as mention previously. This is due to the convergence between contour's sides (left and right).

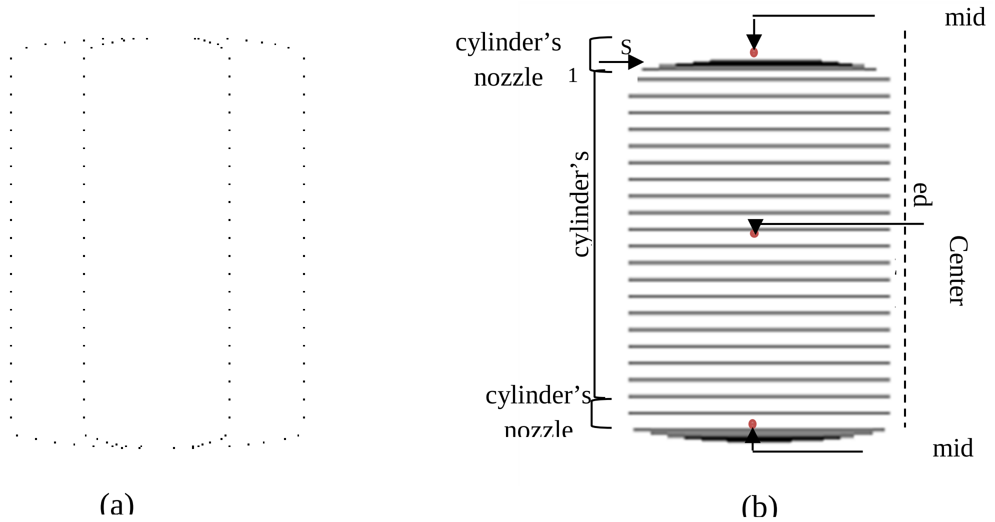

FIG. 4. CYLINDER SHAPE. (A) THE SHAPE OF THE OBJECT AS A SET OF CONTROL POINTS.

(B) THE OBJECT SHAPE AS SLICES.

2. Neural network technique.

The neural networks represent a powerful data processing technique that has reached maturity and broad application [15]. In this paper, the neural network technique is used to estimate the value of the depth. Here, a network with five inputs (slice length, angle of the incident light and illumination of some of points that located along the 2D object, namely control points), two hidden layers and one output is proposed. The hidden layers are composed form eight neurons (see [16] for learning architecture that utilizes the same generic principles). The length of the slice, the illumination of the slice midpoint, two half control points illuminations and the incident light angle (θ) represent the five inputs of the network while the depth represents the output. The first input portrays the number of the pixels in the slice while the intensity values of the midpoint and the two half control points are used as the second, third and fourth inputs respectively. The fifth input (θ), that is shown in Fig. (5), can be defined as the angle between the perpendicular incident light on the centered midpoint and the incident light on the required midpoint as in Eq.(4).

$$\theta_k = \tan^{-1}\left(\frac{y_k - y_{\frac{n}{2}}}{L}\right) \quad k:1\rightarrow n. \tag{4}$$

Where ($y_k - y_{n/2}$) is the vertical distance between the slice midpoint and the centered midpoint and L is the perpendicular distance of the incident light on the centered midpoint.

The proposed network is verified using the Matlab program, where backpropagation trainlm learning algorithm is used with 0.05 learning rate. All neurons are set to a sigmoid activation function. The network is trained using 300 sets that are selected to cover many depths cases for different curved shapes.

Received 13 April 2017; Accepted 3 October 2017
 ISSN 1811-9212

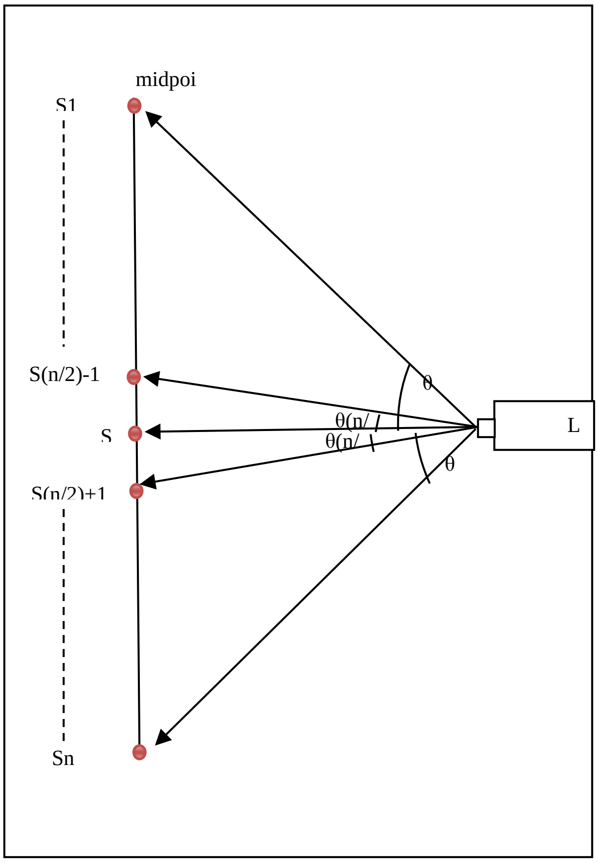

FIG. 5. THE ANGLE OF THE MIDPOINT INCIDENT LIGHT

3. Construction of the 3D shape

Bézier surface is a species of mathematical spline used in computer graphics, computer-aided design and finite element modeling. It is defined by a rectangular grid of control points; it is anchored at the four corner points and employs the other grid points to determine its shape [16].

In this work, the Bézier surface interpolation is utilized to construct the 3D shape where the value of the estimated depth to each slice will be used as the z value for the half control points of the same slice while the depth of the edge control points will be assigned to zero. Eq. (3) is divided into three (4×n) sub-matrices, where each one has a single component of coordinates. These matrices are C_x, C_y and C_z which refer to x, y and z components respectively. Eq. (5) shows the matrix form of x components, where y and z components can be represented in the same way.

$$\begin{array}{c} \begin{array}{cccc} sec_1 & sec_2 & sec_3 & sec_n \end{array} \\ \begin{bmatrix} x(1.1) & x(1.2) & x(1.3) & x(1.n) \\ x(2.1) & x(2.2) & x(2.3) & x(2.n) \\ x(3.1) & x(3.2) & x(3.3) & x(3.n) \\ x(4.1) & x(4.2) & x(4.3) & x(4.n) \end{bmatrix} \end{array} \tag{5}$$

The 3D shape is constructed using a bicubic Bézier surface. Bézier surface is widely utilized in computer graphics to model smooth surfaces. The Bézier curve uses four control points in which curve passes through first and fourth points and approximates the two other points. These points are used to manipulate the curve intuitively through affine transformation such as translation and rotation (see [17] and [18] for the underlying principles of Bézier curve).

In our model, the bicubic Bézier matrix is given by a grid of (4×4) control points. Obviously, each four slices represent 16 control points or one Bézier patch so, for an object shape that has more than four slices, a single surface patch is not enough to cover all object details. Hence, the object surface will

Received 13 April 2017; Accepted 3 October 2017

 ISSN 1811-9212

be described through several patches joined together through continuity patches to ensure the smoothness of the surfaces. The main matrixes C_x,C_y and C_z are divided into sub-matrixes where each of them has four slice (16 control points). The number of sub-matrixes depends on the number of the slices. The bicubic Bézier surface can be shown in matrix form through eq.(6):

$$S(u,v) = [U][N][CP][N][V]^T. \qquad (6)$$

Here, the 4×4 matrix $[CP]$ stores the control points, $[N]$ contains the Bernstein polynomial coefficients, $[U] = [u_3\ u_2\ u_1]$ and $[V] = [v_3\ \ v_2\ v_1]$.

The construction of 3-D shape through a set of Bézier patches and continuity patches is clarified in Fig. (6), where the cylinder shape with 33 slices is portrayed as a wire-frame object that consists from six Bézier patches and five continuity patches. The overall block diagram of our model can be shown in Fig. (7).

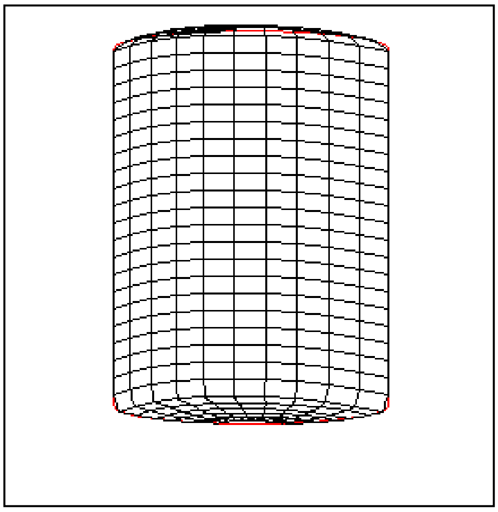

FIG. 6. BÉZIER SURFACE OF THE CYLINDER

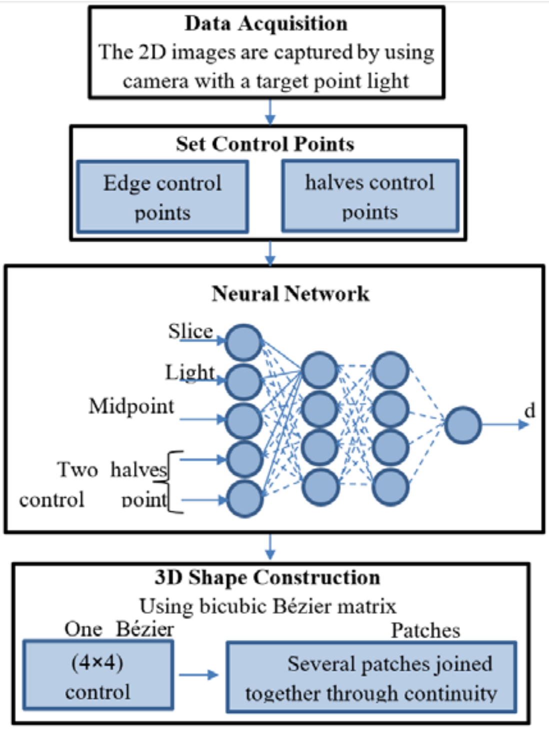

FIG. 7. THE BLOCK DIAGRAM OF THE PROPOSED APPROACH FOR DEPTH ESTIMATION AND SHAPE. RECONSTRUCTION.

Received 13 April 2017; Accepted 3 October 2017

 ISSN 1811-9212

III. EXPERIMENTAL RESULTS

In order to evaluate the performance of our methodology we selected four 2-D images of different shapes including cone, bowl, vase and sphere. These images contain variety of curves and depth which make the construction of 3D image more complex and challenging. Fig. (8) represents the tested objects where the pictures of the left column are the original 2-D object images while those of right column are the constructed 3D object shape images. The cone and bowl images are represented in the Fig.s (8-aL) and (8-bL) respectively, each of them has different depth values and a total length of 40 cm. In the first object, the depth of the upper end starts from (25) cm and grows down to (50) cm at the lower end, while in the second object, the depth starts from (30) cm where it is enlarging to reach (50) cm and then decreases until (15) cm at the end of the object. The cone and bowl shapes are constructed from (64) slices that are translated into a wire-frame Bézier surface with (11) Bézier patches connected with (10) continuity patches as shown in Fig.s (8-aR) and (8-bR), respectively.

Fig. (8-cL) shows the image of the vase object, it has disparity values of depth between (20-45) cm with height of 30 cm, this leads to construct a shape with (56) slices. These slices are translated into (10) Bézier patches connected by (10) continuity patches. The whole patches produced a wire-frame Bézier surface of the vase object with various depths as illustrated in Fig. (8-cR).

Finally, the 2-d image of a sphere object with (15) cm radius is shown in Fig. (8-dL). It is transformed to a shape of (22) slices where these slices have formed a surface with (4) Bézier patches connected by (3) continuity patches as represented in Fig. (8-dR).

The proposed methodology is evaluated through the experimental results that indicate verity values of error depending on the complexity of the tested shape. The mean error value E is calculated through Eq.(7).

$$\mathrm{E} = \frac{1}{\mathrm{N}} \sum_{i=0}^{\mathrm{N}} \mathcal{E}_i \tag{7}$$

Where N is the number of the samples and $\mathcal{E}$ is the error value of the estimated depth that is calculated for each sample as illustrated in Eq.(8).

$$\mathcal{E} = \sum_{\mathrm{j}=0}^{\mathrm{M}} \frac{Rd_{\mathrm{j}} - Ed_{\mathrm{j}}}{Rd_{\mathrm{j}}} \tag{8}$$

Where M is the number of slices in the object shape, *Rd* and *Ed* are the real depth and the estimated depth of the object, respectively. For the five tested samples including cylinder, cone, bowl, vase and sphere objects the mean error values don't exceed (3%).

IV. CONCLUSIONS

In this paper, the depth of an object in a 2-D image is estimated and utilized in the 3-D shape construction. A still camera with a target point light is used to capture the 2-D image that is processed using a set of image processing techniques to get the contour of the object. The object contour is divided in two parts (left and right) sides. Two types of control points named edge control point and half control points are located on both sides. The control points help in determining the slices shape of the object. The neural network technique is used to get an estimation value for object depth based on a set of inputs values including the length of the slice, midpoint illumination, two half control points' illuminations and the incident light angle. The Bézier surface is used to constructe the 3-D shape of the 2-D object based on the estimated depth values.

In order to evaluate our proposed model, we selected four different objects (cone, bowl, vase and sphere) where these objects are characterized by varying depth along the object shape. As a consequence, the 3D shape estimation for these objects is a challenge for our proposed model. The results reveal a beneficial estimation in which the mean error of the whole selected objects is 3%. Our

Received 13 April 2017; Accepted 3 October 2017

 ISSN 1811-9212

model has the potential to be used in robot-grasping tasks in which the geometry of 3D object is reconstructed and then the grasping position and orientation could be determined. Another potential is by implementing the model in industrial-CNC machines. This could provide the machine with image-based depth estimation to manufacture a 3D object from a 2D image.

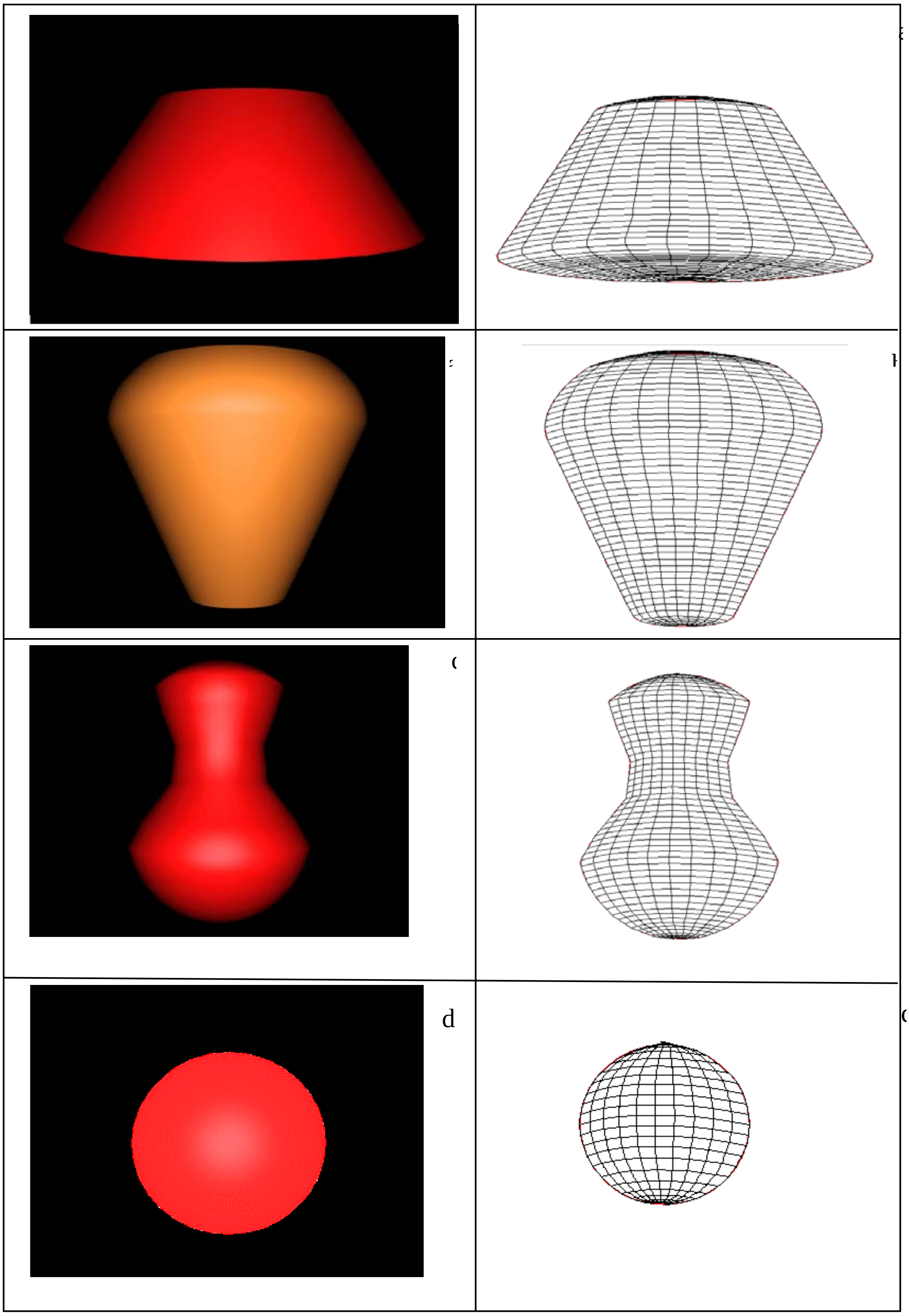

FIGURE (8) THE 2-D IMAGES OF CONE, BOWL, VASE AND SPHERE IMAGES AND THEIR 3-D SURFACES

Received 13 April 2017; Accepted 3 October 2017

 ISSN 1811-9212

REFERENCES

[1] S.Ashutosh , H. Sung ,Y. Andrew, " 3-D Depth Reconstruction from a Single Still Image," springer Int. J. Comput Vis.,2007.

[2] C.Liang and N.Xuan,"Real-Time 3- D Robot Vision Employing Novel ColorFringe Projection,"in:Proceedings of the 4th international conference on Autonomous Robots and Agents, 2009.

[3] G.Al-Kindi and K.Ali " Investigation of Object Shadows Utilization in 3d Shape Re-construction Using Inexpensive Equipment," Transaction in Utility and Image Recognition,V.1, 2010.

[4] R.Fabio, S.El-Hakim, G.Armin, Z.Li,"Turning Images into 3-d Models," IEEE Signal Processing Magazine, 2008.

[5] L.Philippe, I.Dan and M.petriu,"3-D Object Model Recovery Form 2-D Image Using Structured Light," IEEE Transaction on Instrumentation and Measurement, Vol.53, No.2, 2004.

[6] M.Yuichi, K. Avinash,"An Interactive Framework for Acquiring Vision Models of 3-D Objects from 2-D Image,"IEEE Transaction on System, Man, And cybernetics, Vol.34, No.1, 2004.

[7] T.Kuo-Yang, L.Huan-Yu and L.Tsu-Tian,"Rapidly Construction a Simple 3-D Graphic Model Using a Consumer-User Digital Still Camera,"IEEE Transaction on Consumer Electronics, Vol.52, No.3, 2006.

[8] L. Zhong, M. Li-zhuang, T. Wu-zheng and Z. Ming-xi," Reconstruction from contour lines based on bi-cubic Bézier spline surface ", Journal of Zhejiang University - Science A,Vol. 7, No. 7, pp. 1241-1246, 2006.

[9] Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, J. Xiao "3D ShapeNets:A Deep Representation for Volumetric Shapes" The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 1912-1920.

[10] D. A. Spooner, S. Chan, C. L. Simmons "Methodology for 3D scene reconstruction from 2D image sequences". US, Patent, 2014.

[11] D. Jianga,Y. Huc, S. Yand, L. Zhangc, H. Zhangc, W. Gaoa, "Efficient 3D reconstruction for face recognition" Pattern Recognition, Vol. 38, No. 6, PP. 787–798, 2005,

[12] L.G.Shapiro and G.Stockman,"Computer Vision" Prentice Hall, pp. 1-40, 2001.

[13] G. Rafael and W. Richard," Digital Image Processing," Third Edition, Pearson Education, Inc., 2008

[14] G. Al-Kindi, R. Baul, and K. Gill, “An Example of Two Dimensional Component Inspection using Computer Vision,” Proc. Instn. Mech. Engrs., Vol. 205, Part B, pp.71-83, 1991.

[15] D.Gérard,"Neural Networks Methodology and applications," Springer-Verlag Berlin Heidelberg, 2005.

[16] R. Rojas," Neural Networks, A Systematic Introduction " Springer Verlag, Berlin, 1996.

[17] S. David,"Curves and Surfaces for Computer Graphics," Springer Science+Business Media, Inc., 2006.

[18] A. K. Peters, " Fundamentals of Computer Graphics", CRC Press, fourth edition, 2015.

Received 13 April 2017; Accepted 3 October 2017

 ISSN 1811-9212