2011 2nd International Conference on Networking and Information Technology
IPCSIT vol.17 (2011) © (2011) IACSIT Press, Singapore

3D Reconstruction of Polyhedral Objects by Single Picture using Reference Plane

Thantita Chirasophon, Pichit Reaksanandana and Arbtip Dheeravongkit

Institute of Field roBOtics , King Mongkut's University of Technology Thonburi, Bangkok, Thailand
e-mail: pepzei_5@hotmail.com, e-mail: pichit@fibo.kmutt.ac.th, e-mail: arbtip@fibo.kmutt.ac.th

Abstract—This article explains the technique to perform a 3D model reconstruction of a polyhedral object from a single picture using the referenced grid plane theory. The most important part of the proposed method is to find and relate parameters in 2D and 3D spaces. The proposed method consists of three parts. Part 1 Camera Calibration: Adjust the camera to get the most accurate picture by correcting the lens parameters. Part 2: Calculate the positions of the picture obtained from part 1 on the 2D reference plane by using Homography and map the parameters onto the 3D space. Part 3: Place the object of interest on the reference plane and apply the homographic theory as explained in part 2. Mark positions of the object on the 2D picture, then map the position to the 3D coordinate plane using perspective projection theory. The accuracy of this technique depends on the camera calibration and the positions marked on the 2D picture. The advantages of the proposed method are that it can reduce cost, does not require any equipment installation and does not require prior knowledge of camera parameters.

Keywords-3D Reconstruction; Camera calibrate; Perspective projection; Polyhedral object

1. Introduction

This paper presents a method of 3D model reconstruction of a polyhedral object from a single 2D picture. There are several existing techniques to obtain the position of an object in 3-dimensional space. The first method by S. Watanabe and M. Yoneyama is to observe the time of the wave that is reflected off the object back to the receiver. This can be used to calculate the position of the object in 3D space [1]. The second method by A. Prokos, G. Karras and E. Petsa casts a laser beam onto the object, uses two cameras to capture images of the object, then utilizes trigonometry to calculate the position of the object (Stereo Technique) [2]. The third method by S. Rusinkiewicz, O. Hall-Holt and M. Levoy uses a single camera to capture images of the object. Because the data obtained from one camera is not sufficient for 3D model reconstruction, additional information is needed to help calculate the position of the object. Such additional information can be obtained by casting black and white lights onto the object [3], or by rotating the camera to capture the picture of the object from several views [4].The advantage of the third method is that it is cheaper and requires less time to set up the equipment. For this reason in this research only one single camera will be used to find the position of the object in 3D space.

2. System Overview

Fig. 1 shows the system overview and equipment setup for this research study. A grid plane is placed on the floor where the capture view of camera can cover both the grid plane and the object by using a picture size of 640 x 480 pixels.

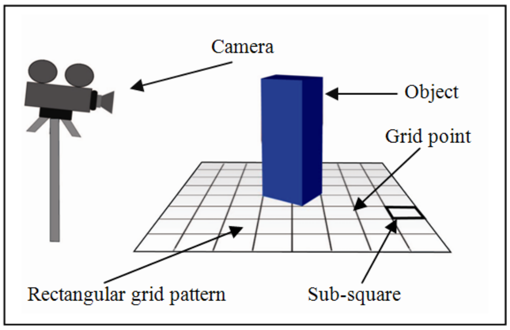

Figure 1. The system set up.

The algorithm is illustrated as a diagram on Fig. 2. The details of each step will be explained in the following sections.

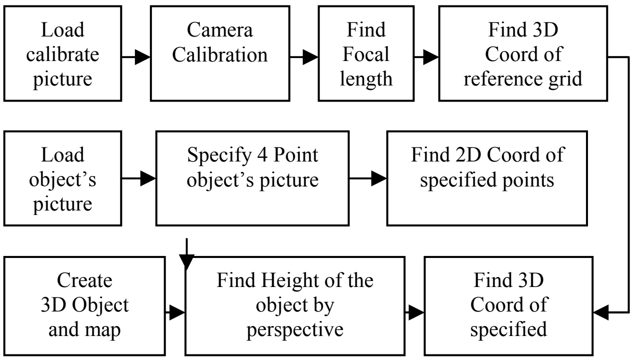

Figure 2. The algorithm.

3. Camera Model

In general, a camera uses Pinhole Camera Model to capture the 2D picture. Fig. 3 shows the camera model, where point O is the Center of Projection or Center of Camera. The Camera Coordinate frame is located at this point. The Screen Coordinate frame is located on the Image Plane, which is obtained when the picture is taken. And the coordinate of any point in 3D space is reference to the World Coordinate. These three coordinate frames relate to each other in the form of Homogenous Coordinates, as following:

$$\lambda x = PX \tag{1}$$

$$\lambda \begin{bmatrix} x_{pi} \\ y_{pi} \\ z_{pi} \end{bmatrix} = K\begin{bmatrix} R & T \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{2}$$

$$K = \begin{bmatrix} \alpha & s & x_0 \\ 0 & \beta & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

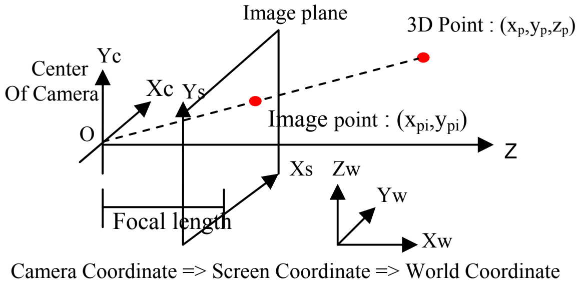

Figure 3. Camera Model and all Coordinate Frames.

From (1)-(3).

λ = An arbitrary scalar.

x = A 2D point or vector which presenting the position of the point in 3D that projects towards the image plane

X = A 3D point.

P = Camera Matrix.

K = The camera intrinsic matrix. Matrix value that can be obtained from camera calibration.

α, β = The scale factors in image x and y axes.

s = The parameter describing the skewness of the two image axes.

x_0, y_0 = The coordinates of the principal point.

$[R \quad T]$ = The extrinsic parameters, including the rotation and translation components, which relate the world coordinate system to the camera coordinate system.

4. Camera Calibration

Before doing the experiment, all pictures need to be adjusted in order to have the least error from camera's parameters. In this experiment 10 pictures are used.

For each picture, the 4 extreme corners are specified as rectangular grid pattern. The specification must be square or rectangle only. The boundary of the calibration grid is then shown in Fig. 4.

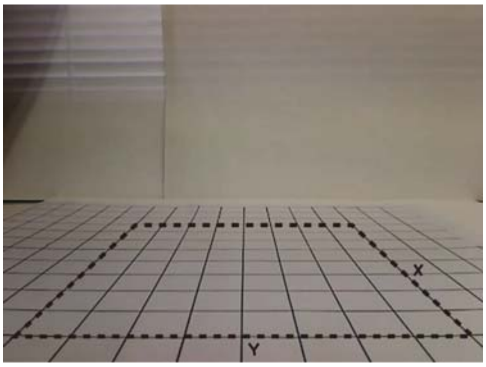

Figure 4. The 4 extreme corners specified in the picture.

The position values of the 4 specified points are referenced to, the screen coordinate frame. By using these 4 points, we can find grid points by using SUSAN Corner Detector, and calculate grid points inside the picture, by using Homography. We can compare the grid points by Susan Corner Detector and using Homography to adjust the value to get less error.

Homography is the method to present the relationship between one plane and the other [5] having the equation as follows:

$$m = HM \tag{4}$$

$$H = K[r_1 \quad r_2 \quad t] \tag{5}$$

From (4)-(5).

m = An image point m.

M = A model point M.

H = Matrix presenting relationship between m and M . M = Relationship between a model point M and an image point m .

r = Rotation matrix.

t = Translation matrix.

K = The camera's intrinsic matrix obtained from the camera calibration.

To find sub-squares in the rectangle grid pattern using Homography, we specify m as the 4 extreme corners indicated earlier, specify M as another matrix below.

$$M = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

Next, we specify the number of sub-squares in the boundary of the calibration grid. Let *a* be the number of rows and *b* be the number of columns of sub-square, then we create a scaling matrix of size $3\times(a\times b)$ as seen (6).

$$Pls = \begin{bmatrix} 0 & \frac{1}{(a\times b)} & \frac{2}{(a\times b)} & \ldots & 1 & 0 & \frac{1}{(a\times b)} & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 & \frac{1}{(a\times b)} & \frac{1}{(a\times b)} & \ldots & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}_{3\times(a\times b)} \quad (6)$$

By multiplying the matrix with Homography (*H*), we shall then have all position values of sub-square in the picture.

$$Spoint = HPls \quad (7)$$

Then, we adjust lens distortion by varying the values of matrix *K* in (5), and recalculate the position of sub-square using Homography. By keep adjusting the matrix *K* and compare the result with those obtained by SUSAN Corner Detector for all 10 pictures, we obtain the following lens distortion matrix, which yields the most accurate result:

$$K = \begin{bmatrix} -0.00148 & 0.00139 & 320 \\ 0.000680 & -0.00082 & 240 \\ 0 & 0 & 1 \end{bmatrix}$$

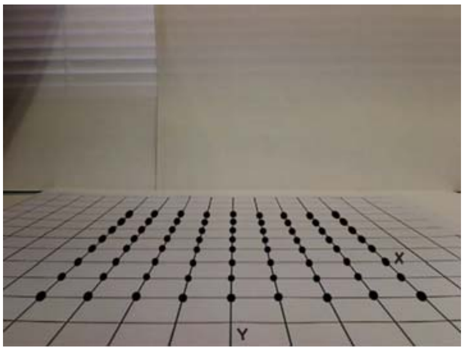

Figure 5. Grid point positions.

The error of grid point positions finding from 10 pictures can be seen on Fig. 6. The average error values on X and Y axis are 1.14791 and 1.71723 pixels.

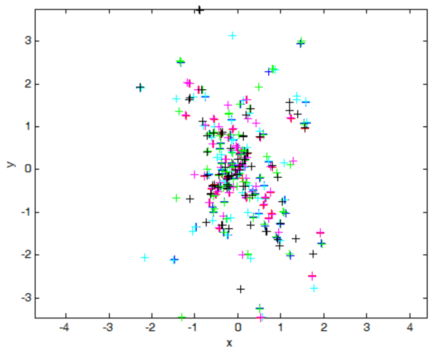

Figure 6. The error of sup-square positions.

5. Find Focal Length

The focal length is the distance measuring from the optical center of a lens to the image plane. It can be found from finding Homography between sub-square positions obtain from camera calibrating and the size of each sub-square using (4). We shall then get Homography of every picture, and use it to calculate the focal length by using Vanishing Points as (8)-(14).

$$V_{horizontal} = H(:,1) \tag{8}$$

$$V_{vertical} = H(:,2) \tag{9}$$

$$V_{diagonal1} = \frac{V_{hirozontal} + V_{vertical}}{2} \tag{10}$$

$$V_{diagonal2} = \frac{V_{hirozontal} - V_{vertical}}{2} \tag{11}$$

$$\begin{bmatrix} a1 & b1 & c1 \\ a2 & b2 & c2 \\ a3 & b3 & c3 \\ a4 & b4 & c4 \end{bmatrix} = \begin{bmatrix} \frac{V_{horizontal}}{norm(V_{horizontal})} \\ \frac{V_{vertical}}{norm(V_{vertical})} \\ \frac{V_{dicgonal1}}{norm(V_{dicgonal1})} \\ \frac{V_{dicgonal2}}{norm(V_{dicgonal2})} \end{bmatrix} \tag{12}$$

$$A_{kk} = \left[(a1 \times a2) \quad (b1 \times b2) \quad (a3 \times a4) \quad (b3 \times b4) \right] \tag{13}$$

$$B_{kk} = \begin{bmatrix} c1 \times c2 \\ c3 \times c4 \end{bmatrix} \tag{14}$$

Calculated A_{kk} and B_{kk} of all picture, we obtain.

$$A = \begin{bmatrix} A_{kk1} \\ A_{kk2} \\ A_{kk3} \\ \vdots \end{bmatrix} \text{ และ } B = \begin{bmatrix} B_{kk1} \\ B_{kk2} \\ B_{kk3} \\ \vdots \end{bmatrix} \tag{15}$$

Then using A and B to find focal length.

$$F = \sqrt{\frac{B' \times (\Sigma A')}{B' \times B}} \tag{16}$$

From the experiment, F= 528.64260 and from (12) norm is magnitude of matrix. Using the focal length value, we can calculate orientation and translation of rectangle grid pattern in the boundary of the calibrate grid. Next, using the positions of sub-square from each calibrated picture to find the actual 3D position in Camera Coordinate as seen in Fig. 7.

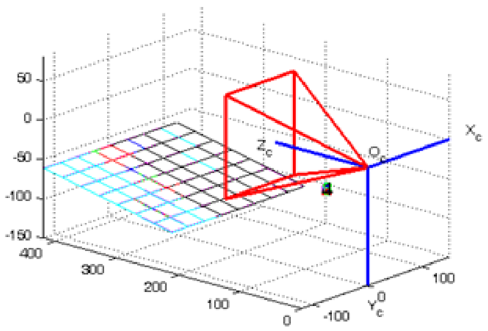

Figure 7. The actual 3D position of rectangular grid plane and camera in 3D space.

6. Comparison between 2D and 3D Rectangular Plane

From the previous calculation, we have found the 2D position of sub-square by averaging values from all 10 pictures in order to get the most accurate result. To find the position of object, in 3D space 4 additional points must be defined on the image, where 3 of these lying on the rectangular grid pattern as shown in Fig. 8 (left). Next, we use each 2D position to find actual position in 3D by finding the nearest sub-square on the rectangular grid pattern. Then we recalculate sub-squares inside that found sub-square. The size of sub-squares will be reduced to 1 millimeter as shown in Fig. 8 (right).

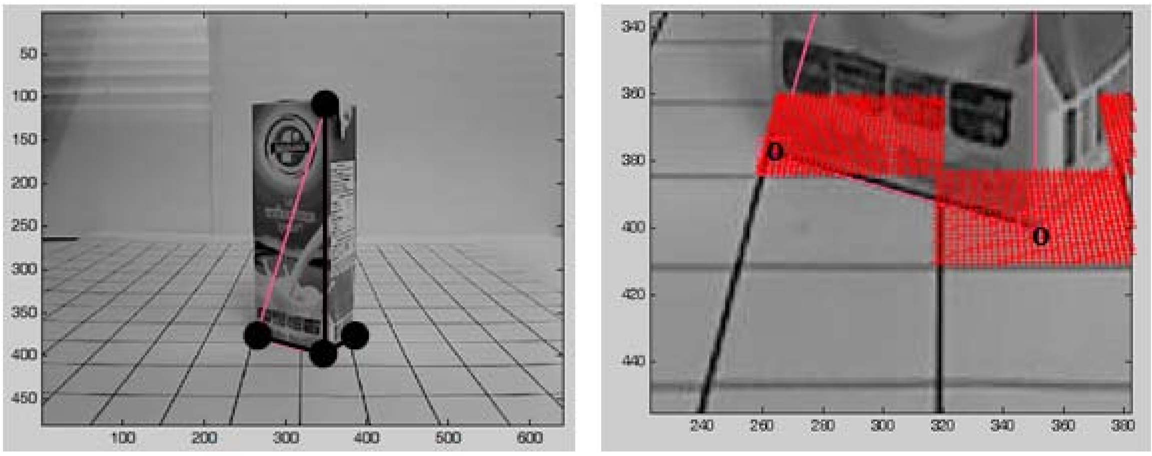

Figure 8. Left: The points specified on the image. Right: Sub-squares of the original sub-square nearest to the specified point.

For each of the 3 points defined on reference grid pattern, we calculate the closest grid points in the 1 millimeter sub-square, and compare them with the 3D position of grid points obtain camera calibrate as seen in Fig. 9 (right).

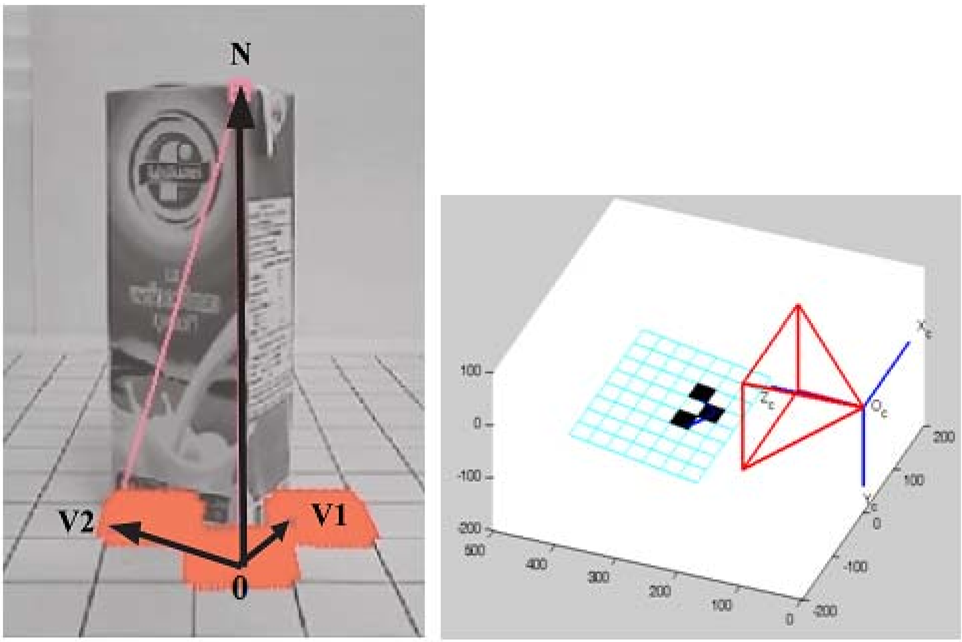

Figure 9. Left: Sub-square position of the specify point and relative vector. Right: Sub-square position of the specify point in 3D space.

7. Calculate Height of the Object

The height of the object can be found using the vector relation between the three points specified in the previous section. Let *V1* and *V2* be the relative position vectors among the three points as shown in Fig. 9(Left), the normal vector *N* of the plane *V1V2* is the height vector of the object.

$$N = V1 \times V2 \tag{17}$$

Then we can form linear equation as in (18)-(19), in order to find any position (x,y,z) on vector N in 3D space.

$$\frac{x-o_x}{N_x-o_x}, \frac{y-o_y}{N_y-o_y}, \frac{z-o_z}{N_z-o_z} \tag{18}$$

$$\begin{aligned} x &= o_x + (N_x - o_x)t \\ y &= o_y + (N_y - o_y)t \\ z &= o_z + (N_z - o_z)t \end{aligned} \tag{19}$$

By adjusting the variance *t*, we shall then get 3D points on normal vector. Then we can map those points on to the image plane using (20)-(24).

$$C = MW \tag{20}$$

$$M = \begin{bmatrix} R_{3\times3} & T_{3\times1} \\ O_{3\times1} & 1 \end{bmatrix} \tag{21}$$

$$W = \begin{bmatrix} X_W & Y_W & Z_W & 1 \end{bmatrix}^T \tag{22}$$

$$U = F \times \frac{C_x}{C_z} \tag{23}$$

$$V = F \times \frac{C_y}{C_z} \tag{24}$$

From (20)-(24).

M = Rotation and translation matrix of camera.

W = A point in 3D space.

C = Camera Coordinate of *W*.

F = The focal length.

U, V = The position of 3D point project on image plane.

Then we can compare the found position (U,V) with the last point specified in section VI, and adjust the *t* value to reduce error, if the least error. The resulting (x,y,z) will be the position indicating the height of the object.

8. Results

To test our algorithm, we set up an experiment to measure ten boxes with different sizes. The result of the experiment displayed in table 1. And the 3D reconstruction and texture mapping of the object is shown in Fig. 10.

TABLE I. THE RESULTS OF MEASUREMENT.

No	Real size (mm.)	Size of measurement (mm.)	Error (mm.)	Percentage of error (%)
1	15.0000	15.0333	0.0333	0.2200
2	29.0000	31.4006	2.4006	8.2779
3	30.0000	30.0355	0.0355	0.1183
4	45.0000	43.0116	1.9884	4.4187
5	70.0000	68.6003	1.3997	1.9996

6	77.0000	78.4347	1.4347	1.8632
7	90.0000	93.0054	3.0054	3.3393
8	115.0000	107.0420	7.9580	6.9200
9	150.0000	150.0000	0.0000	0.0000
10	180.0000	181.2236	1.2236	0.6798

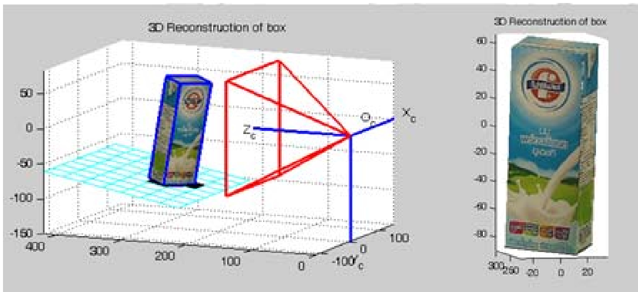

Figure 10. The 3D reconstruction and texture mapping of the object.

9. Conclusions

From the experiment, we have found that the percentage of error of the proposed method to measure the object's size is 2.78%, which resulted from the camera calibration and the error in the measuring. If the object is placed far away from the camera, the error will increase respectively, As seen in the case of object No.2 which is placed farthermost from the camera the error thus become highest. Therefore it can be concluded that the accuracy of the proposed method to measure the size of 3D object from 2D image is dependant upon the calibration process and the position of measuring. This method can reconstruct 3D objects by using a single image, without any requirement to specify position or direction of the camera. However, it is necessary that the grid plane must be within the scope of camera and the object to be reconstructed must be within the scope of referred rectangular grid pattern.

10. Acknowledgment

This research work is supported financially by the National Science and Technology Development Agency, Thailand.

11. References

[1] S. Watanabe and M. Yoneyama, "An Ultrasonic Visual Sensor for Three-Dimensional Object Recognition Using Neural Networks," IEEE Trans. Conf. Robotics Automate, vol. 8, pp. 240-249, 1992.

[2] A. Prokos, G. Karras and E. Petsa, "AUTOMATIC 3D SURFACE RECONSTRUCTION BY COMBINING STEREOVISION WITH THE SLIT-SCANNER APPROACH," ISPRS Conf. WG V/1 - Vision metrology - best practice systems and applications, vol. 38, pp. 505-509, 2010,

[3] S. Rusinkiewicz, O. Hall-Holt and M. Levoy, "Real-Time 3D Model Acquistion," ACM Transactions on Graphics (TOG), vol. 21, pp. 438-446, 2002.

[4] Y. Matsumoto, H. Terasaki, K. Sugimoto and T. Arakaway, "A Portable Three-dimensional Digitizer," IEEE Trans. Conf. 3-D Digital Imaging and Modeling, pp. 197-204, 1997.

[5] Zh. Zhang, "Flexible Camera Calibration By Viewing a Plane From Unknown Orientations," Microsoft Research, One MicrosoftWay, Redmond,WA 98052-6399, USA, Nov 1999.