Shahrul Azman Noah Azizi Abdullah
Haslina Arshad Azuraliza Abu Bakar
Zulaiha Ali Othman Shahnorbanun Sahran
Nazlia Omar Zalinda Othman (Eds.)

Communications in Computer and Information Science 378

Soft Computing Applications and Intelligent Systems

Second International Multi-Conference on Artificial Intelligence Technology, M-CAIT 2013
Shah Alam, August 2013, Proceedings

Springer

Communications in Computer and Information Science 378

Table of Contents

Real World Coordinate from Image Coordinate Using Single Calibrated Camera Based on Analytic Geometry

Joko Siswantoro[1,2], Anton Satria Prabuwono[1], and Azizi Abdullah[1]

[1] Center For Artificial Intelligence Technology, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor D.E., Malaysia
[2] Department of Mathematics and Sciences, Universitas Surabaya, Jl. Kali Rungkut Tengilis, Surabaya, 60293, Indonesia
joko_siswantoro@ubaya.ac.id
{antonsatria,azizi}@ftsm.ukm.my

Abstract. The determination of real world coordinate from image coordinate has many applications in computer vision. This paper proposes the algorithm for determination of real world coordinate of a point on a plane from its image coordinate using single calibrated camera based on simple analytic geometry. Experiment has been done using the image of chessboard pattern taken from five different views. The experiment result shows that exact real world coordinate and its approximation lie on the same plane and there are no significant difference between exact real world coordinate and its approximation.

Keywords: real world coordinate, image coordinate, analytic geometry.

1 Introduction

The determination of image coordinate of a point in real word coordinate system can be easily calculated using a transformation after camera parameters that are used in image acquisition are known [1],[2] and [3]. Generally the reverse of this problem is cannot be performed, since the transformation from real world coordinate system to image coordinate system is not invertible. Information about the depth of position loses during the transformation. But under certain condition, such as the point in real world coordinate system lies on a ground plane, determination of world coordinate from a point in image coordinate system still can be performed [4]. The determination of real world coordinate has many applications in computer vision including robot positioning [5], object reconstruction [6], and measurement [4],[7] and [8]. Therefore, determination of the real world coordinate of a point from image coordinate is challenging problem in computer vision.

Common method to determine real world coordinate of a point from image coordinate is triangulation. Triangulation is problem of determining the real world coordinate of a point from a set of corresponding image locations and known camera parameters [3]. Mohamed et al [9] and Zhang [10] have used triangulation to determine the real world coordinate of a point using two corresponding image

S.A. Noah et al. (Eds.): M-CAIT 2013, CCIS 378, pp. 1–11, 2013.

coordinate acquired by two cameras in order to measure the accuracy of their proposed camera calibration method.

Many methods have been proposed to determine real word coordinate of a point in image coordinate system. Some of them do not apply camera parameters in determination of real world coordinate. Bucher [4] has proposed a decomposable image to world mapping where the transformation of vertical coordinate is independent from the horizontal position. Memony et al. [11] have proposed a multilayer artificial neural network model (ANN) to determine real world coordinate from matched pair of images. Vilaça et al. [8] have proposed method for determination of real world coordinate in a plane using two cameras and laser line. Polynomial was used to relate between image coordinate and real world coordinate. Xiaobo et al. [12] have used direct linear transformation and back propagation neural network in the determination of real world coordinate in a plane from image coordinate.

Currently the determination of camera parameters is not a hard problem. Using established camera calibration method, such as method proposed by Tsai [13] and Zhang [10], and supported by computer vision library that provides functions for camera calibration such as OpenCV [1] and camera calibration Toolbox for Matlab [14], camera parameters can be easily estimated. Therefore using estimated camera parameters and analytic geometry, the real world coordinate of a point on a plane can be easily obtained from image coordinate. Furthermore previews method can only be used to re-project a point in image coordinate system to its original position in real world coordinate system. They cannot be used to re-project a point to a plane that is different from its original position. This paper proposes the algorithm for determination of real world coordinate of a point on a plane using single calibrated camera base on analytic geometry. The algorithm can also be used to re-project a point in image coordinate system to any plane in real world coordinate system which is not perpendicular to image plane

2 Camera Model

Camera model is usually derived from simple pinhole camera based on collinearity principle. The origin of camera coordinate system is projection center. Each point in real world coordinate system is projected into image plane system by a line through projection center. The z-axis of the camera coordinate system is principle axis. This axis is perpendicular to image plane and intersects image plane at $z = f$, where f is the focal length of camera, as shown in Fig. 1. Projection from real world coordinate system onto image coordinate system in image plane consists of two transformations. The first one is transformation from real world coordinate system to camera coordinate system and the second one is transformation from camera coordinate system to image coordinate system.

Camera parameters are needed to construct the transformation. Camera parameters consist of extrinsic and intrinsic parameters. Extrinsic parameters are used to transform point in real world coordinate system to camera coordinate system. The parameters consist of rotation matrix **R** and translation vector **T** as the following form.

$$\mathbf{R} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{33} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}, \mathbf{T} = \begin{pmatrix} t_{11} \\ t_{21} \\ t_{31} \end{pmatrix}$$

Rotation matrix is the product of three rotation matrices $\mathbf{R}_x(\psi), \mathbf{R}_y(\varphi)$ and $\mathbf{R}_z(\theta)$. Where $\mathbf{R}_x(\psi), \mathbf{R}_y(\varphi)$ and $\mathbf{R}_z(\theta)$ are rotation matrices around *x*-, *y*-, and *z*- axis with respective rotation angles ψ, φ and θ. Translation vector is a shift from real world coordinate system to camera coordinate system

Intrinsic parameters are used to transform point in camera coordinate system to image coordinate system. The parameters consist of the focal length of camera *f* and the center of image plane coordinates $\mathbf{C}(c_x, c_y)$. All camera parameters are obtained from camera calibration process.

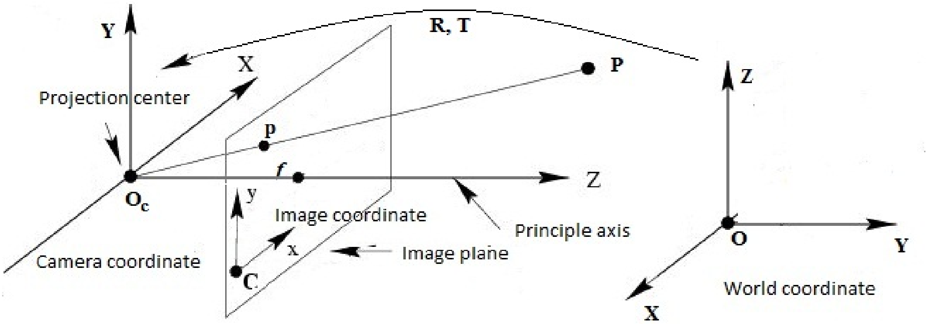

Fig. 1. The geometry of camera model based on pinhole camera model

Suppose $\mathbf{P}(x_w, y_w, z_w)$, $\mathbf{P_c}(x_c, y_c, z_c)$, and $\mathbf{p}(x_{im}, y_{imw})$ are coordinate of a point in real world coordinate system, camera coordinate system and image coordinate system respectively. By assuming that camera used in image acquisition has very small distortion such that its distortion coefficients can be neglected, then transformation from real world coordinate system to image coordinate system can be describe in the following equations [1],[2] and [3].

$$\begin{pmatrix} x_c \\ y_c \\ z_c \end{pmatrix} = \mathbf{R} \begin{pmatrix} x_w \\ y_w \\ z_w \end{pmatrix} + \mathbf{T} \tag{1}$$

$$\begin{pmatrix} x_{im} \\ y_{im} \\ 1 \end{pmatrix} = \begin{pmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_c / z_c \\ y_c / z_c \\ 1 \end{pmatrix} \tag{2}$$

Where $f_x = s_x f$, $f_y = s_y f$ are focal length in x , y direction respectively and s_x, s_y are the size of individual imager elements in x , y direction respectively.

3 Real World Coordinate from Image Coordinate

In order to determine the real world coordinate of point $\mathbf{P}(x_w, y_w, z_w)$ from its image coordinate $\mathbf{p}(x_{im}, y_{imw})$, camera calibration is firstly performed to estimate camera parameters. Suppose **P** is located on plane **A** or we want to re-project point **p** to point **P** on plane **A**, where **A** is not parallel to line l through projection center $\mathbf{O_c}$ and image point **p**. Under this assumption, the coordinate of **P** is intersection of plane **A** and line l, as shown in Fig. 2. Therefore real world coordinate of $\mathbf{O_c}$ and **p** must be determined firstly. From Eq. (1) the real world coordinate of $\mathbf{O_c}$ can be expressed as the following equation.

$$\mathbf{O_c} = \begin{pmatrix} o_{c1} \\ o_{c2} \\ o_{c3} \end{pmatrix} = \mathbf{R}^{-1}\mathbf{T} \tag{3}$$

From Eq. (2) and the fact that **p** lies on plane $z = f$ in camera coordinate system, the coordinate of $\mathbf{p}(x_{imc}, y_{imc}, z_{imc})$ in camera coordinate system can be expressed as follow

$$x_{imc} = \frac{x_{im} - c_x}{s_x} \tag{4}$$

$$y_{imc} = \frac{y_{im} - c_y}{s_y} \tag{5}$$

$$z_{imc} = f \tag{6}$$

Using Eq. (1), (4), (5), and (6), the coordinate of $\mathbf{p}(x_{imw}, y_{imw}, z_{imw})$ in real world coordinate can be expressed as follow

$$\begin{pmatrix} x_{imw} \\ y_{imw} \\ z_{imw} \end{pmatrix} = \mathbf{R}^{-1}\left(\begin{pmatrix} x_{imc} \\ y_{imc} \\ z_{imc} \end{pmatrix} - \mathbf{T}\right) \tag{7}$$

According to geometry analytic [15], a line in 3D space is determined by a point and a direction vector of the line. Direction vector of a line is a vector that parallel to the line, as shown in Fig. 2. The vector

$$\mathbf{v} = \overrightarrow{\mathbf{O_c p}} = \left(x_{imw} - o_{c1}\right)\mathbf{i} + \left(y_{imw} - o_{c2}\right)\mathbf{j} + \left(z_{imw} - o_{c3}\right)\mathbf{k} \tag{8}$$

is parallel to line *l*. Therefore the equation of line *l* is given by

$$x = o_{c1} + t\left(x_{imw} - o_{c1}\right),\ y = o_{c2} + t\left(y_{imw} - o_{c2}\right),\ z = o_{c3} + t\left(z_{imw} - o_{c3}\right),\ t \in \mathbb{R} \tag{9}$$

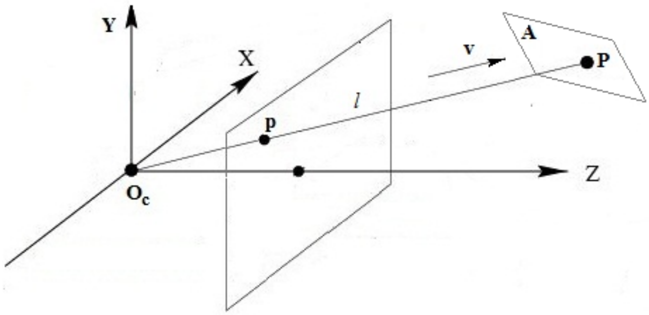

Fig. 2. Line *l* is line through $\mathbf{O_c}$ and **p**. Vector **v** is direction vector of line *l*. The real coordinate of **P** is intersection between plane **A** and line *l*.

Suppose the equation of plane **A** is given by following equation

$$ax + by + cz = d \tag{10}$$

where $a\left(x_{imw} - o_{c1}\right) + b\left(y_{imw} - o_{c2}\right) + c\left(z_{imw} - o_{c3}\right) \neq 0$. This condition will guarantee that plan **A** and line *l* are not parallel and have an intersection point. The point intersection of plane **A** and line *l* is obtained by substituting Eq. (9) to Eq. (10) and found

$$t = \frac{d - \left(ao_{c1} + bo_{c2} + co_{c3}\right)}{a\left(x_{imw} - o_{c1}\right) + b\left(y_{imw} - o_{c2}\right) + c\left(z_{imw} - o_{c3}\right)}. \tag{11}$$

Substitute back t in Eq. (11) to Eq. (9) and the coordinate of **P** in real world coordinate system is approximated by

$$x_{wapp} = o_{c1} + \frac{\left[d - \left(ao_{c1} + bo_{c2} + co_{c3}\right)\right]\left(x_{imw} - o_{c1}\right)}{a\left(x_{imw} - o_{c1}\right) + b\left(y_{imw} - o_{c2}\right) + c\left(z_{imw} - o_{c3}\right)} \quad (12)$$

$$y_{wapp} = o_{c2} + \frac{\left[d - \left(ao_{c1} + bo_{c2} + co_{c3}\right)\right]\left(y_{imw} - o_{c2}\right)}{a\left(x_{imw} - o_{c1}\right) + b\left(y_{imw} - o_{c2}\right) + c\left(z_{imw} - o_{c3}\right)} \quad (13)$$

$$z_{wapp} = o_{c3} + \frac{\left[d - \left(ao_{c1} + bo_{c2} + co_{c3}\right)\right]\left(z_{imw} - o_{c3}\right)}{a\left(x_{imw} - o_{c1}\right) + b\left(y_{imw} - o_{c2}\right) + c\left(z_{imw} - o_{c3}\right)} \quad (14)$$

From the above explanation, we propose the algorithm for real world coordinate determination of point $\mathbf{P}\left(x_w, y_w, z_w\right)$ on a plane from its image point $\mathbf{p}\left(x_{im}, y_{imw}\right)$ as follow:

Step 1. Perform camera calibration to obtain extrinsic camera parameters **R**, **T** and intrinsic camera parameters f_x, f_y, c_x, c_y .

Step 2. Find the coordinate of point $\mathbf{p}\left(x_{im}, y_{im}\right)$ in image coordinate system.

Step 3. Find the coordinate of projection center $\mathbf{O_c}$ in real world coordinate system using Eq. (3).

Step 4. Find the coordinate of $\mathbf{p}\left(x_{imw}, y_{imw}, z_{imw}\right)$ in real world coordinate system using Eq. (7).

Step 5. Approximate the coordinate of **P** in real world coordinate system using Eq. (12), (13), and (14).

4 Experiment and Result

Experiment was performed in the laboratory to validate proposed algorithm. Proposed algorithm was implemented in C++ using OpenCV library. The methodology used in this experiment consists of camera calibration, real world coordinate system construction, image acquisition, corner detection, real world coordinate approximation, and error analysis. Fig. 3 shows the methodology used in the experiment.

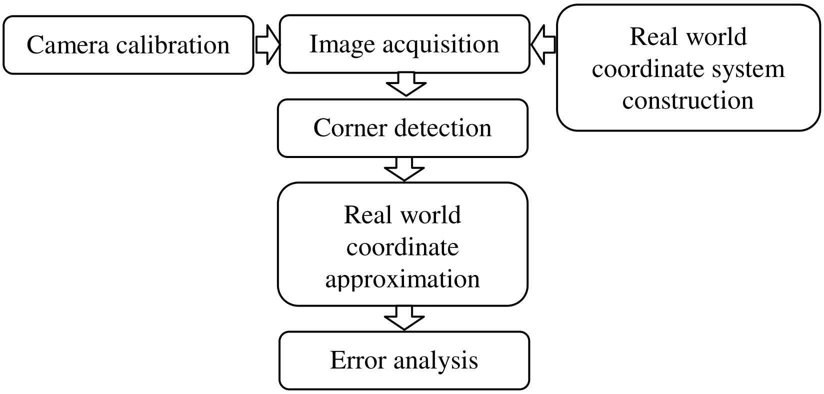

Fig. 3. Methodology used in the experiment

Web camera Logitech C270 was used for image acquisition in the experiment. The camera acquired image with dimension 640 × 480 pixels and resolution 96 dpi in both vertical and horizontal directions. The camera has very small distortion both in radial and tangential distortions. Therefore the distortion coefficients can be neglected. Camera calibration was performed base on Zhang's method [10] using OpenCV library. A 9 × 6 corners flat chessboard pattern with 24.65 mm × 24.65 mm in each square was used in camera calibration, as shown in Fig. 4. Ten views of the pattern were acquired to estimate intrinsic camera parameters and one view was acquired to estimate extrinsic camera parameters.

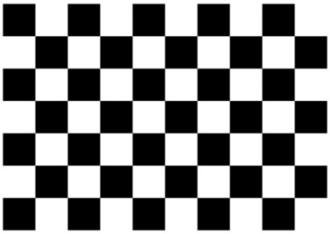

Fig. 4. Chessboard pattern used in camera calibration

Fifty four inner corners of chessboard pattern were used as points in real world coordinate system by assuming the points lie on plane $z = 0$ for simplicity. Therefore the real world coordinate of the corners have the form $(x_i, y_i, 0)$, $i = 1, 2, \ldots, 54$. For the construction of real world coordinate system, the following assumptions were used. The center of real world coordinate system is located at top left corner of chessboard pattern. Positive x-axis lies along top left corner to bottom left corner and positive y-axis lies along top left corner to top right corner. Positive z axis is perpendicular to x axis and y axis according to right hand rule. Fig. 5 shows constructed real world coordinate system used in the experiment.

The images of chessboard pattern were acquired from five different views which are three from top view with an angle of approximately 90° in different distance and two from side view with an angle of approximately 45° in x and y direction, as shown in Fig. 6. The corners of chessboard pattern were located to determine the image coordinate of the corners using sub pixel detection technique [1]. After all image coordinate of corners were calculated, these coordinates were then used to approximate its real world coordinate on plane using proposed algorithm.

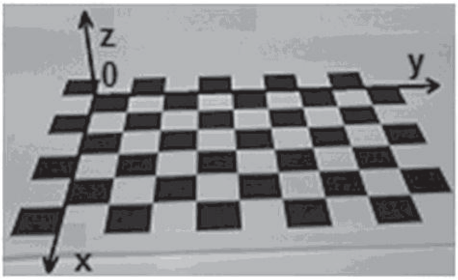

Fig. 5. Constructed real world coordinate system used in the experiment

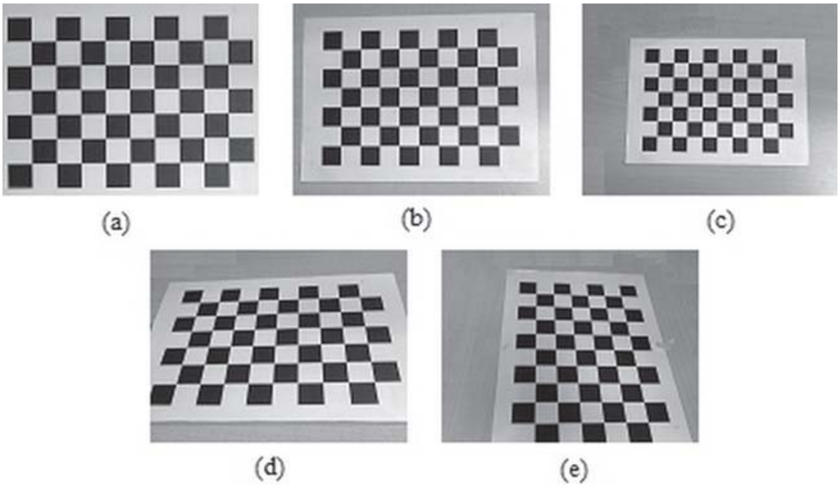

Fig. 6. The images of chessboard pattern from five different views: (a), (b), and (c) top view with an angle of approximately 90° and different distance (d) side view with an angle of approximately 45° in x direction (e) side view with an angle of approximately 45° in y direction.

The accuracy of proposed algorithm was measured using absolute error between exact real world coordinate and its approximation in x, y, and z direction using the following equation.

$$E_x = \left| x_w - x_{wapp} \right| \tag{15}$$

$$E_y = \left| y_w - y_{wapp} \right| \tag{16}$$

$$E_z = \left| z_w - z_{wapp} \right| \tag{17}$$

Where $X_w = \left(x_w, y_w, z_w\right)$ and $X_{west} = \left(x_{wapp}, y_{wapp}, z_{wapp}\right)$ are exact real world coordinate and its approximation respectively.

The result of experiment shows that the approximations of real world coordinate of 54 corners are also on plane $z = 0$. The exact coordinate of 54 corners and its approximation from image coordinate using proposed algorithm on plane $z = 0$ are shown in Fig. 7. From Fig. 7, it can be seen that there are no significant different between the exact coordinate of 54 corners and its approximations. The absolute errors between real world exact coordinate and its approximation are summarized in Table 1.

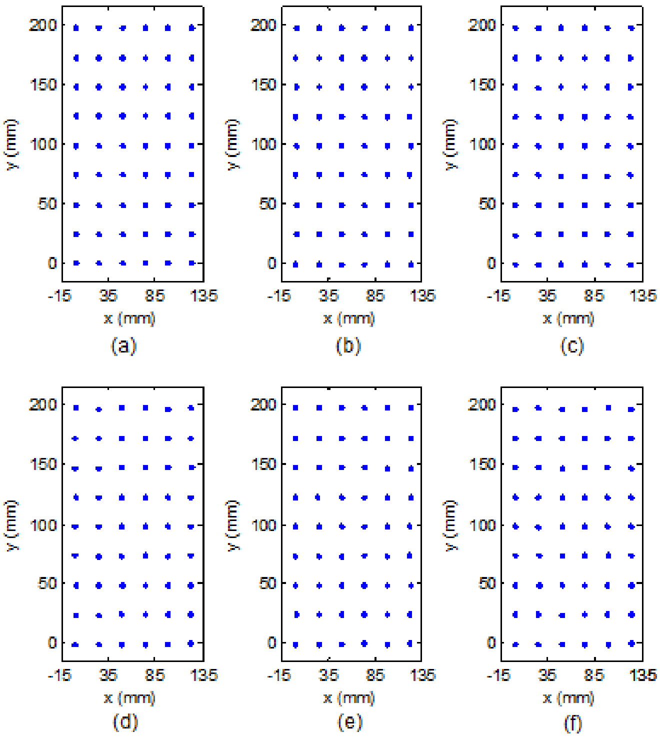

Fig. 7. (a) The exact coordinate of 54 corners on plane z = 0 and its approximations from image coordinate using proposed algorithm: (b) – (f) from image in Fig. 5 (a) – (e) respectively

Table 1. Absolute error between exact real world coordinate and its approximation in *x*, *y*, and *z* direction (in mm)

View	Mean E_x	Std. dev. E_x	Mean E_y	Std. dev. E_y	Mean E_z	Std. dev. E_z
View a	0.2188	0.1456	0.2117	0.1371	0.0000	0.0000
View b	0.2839	0.1990	0.2874	0.1732	0.0000	0.0000
View c	0.3271	0.2265	0.3616	0.2168	0.0000	0.0000
View d	0.3206	0.2052	0.2786	0.1879	0.0000	0.0000
View e	0.2399	0.1673	0.3381	0.2730	0.0000	0.0000

Since the approximation of real world coordinates of 54 corners lie on $z = 0$, therefore the absolute errors in z direction are zeros. In *x* and *y* directions the largest absolute error means are 0.3271 mm and 0.3616 mm respectively occurred in approximation using image from view c. The standard deviation of absolute error is less than 0.27 mm for all views and all direction. It indicates that the entire absolute error tend to be very close to the mean. Therefore proposed algorithm gives a good approximation for real world coordinate from image coordinate.

In this experiment error can be occurred due to error in camera parameters estimation and inaccuracy in locating the image coordinate of corners. Absolute error mean increases with increasing distance between camera and plane, as shown in view a, b, and c. In addition, decreasing angle between camera and plane $z = 0$ also has impact on increasing absolute error mean. In both case, increasing absolute error mean may be occurred due to inaccuracy in determining the image coordinate of corner when the distance from camera to image increases. Therefore camera parameters estimation and locating the image coordinate of corners play an important role in reducing the error of real world coordinate approximation.

5 Conclusion

In this paper the algorithm for determination of real world coordinate of a point on a plane using single calibrated camera base on analytic geometry was proposed. Camera calibration was performed firstly to estimate extrinsic and intrinsic camera parameters including rotation matrix, translation vector, focal length, and the center of image plane coordinate. These parameters together with camera model and simple analytic geometry were used to approximate the real world coordinate of a point on a plane from its image coordinate. The experiment result shows the absolute errors have mean less than 0.37 mm and low standard deviation. It can be inferred that proposed algorithm gives a good approximation for real world coordinate from image coordinate.

Acknowledgment. The authors would like to thanks Ministry of Higher Education Malaysia and Universiti Kebangsaan Malaysia for providing facilities and financial support under Grants No. FRGS/1/2012/SG05/UKM/02/12 and PTS-2011-054.

References

1. Bradski, G., Kaehler, A.: Learning OpenCV: Computer Vision with the OpenCV Library. O'Reilly Media, Inc. (2008)
2. Heikkila, J., Silven, O.: A four-step camera calibration procedure with implicit image correction. In: Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition, pp. 1106–1112 (1997)
3. Szeliski, R.: Computer Vision Algorithms and Applications. Springer (2011)
4. Bucher, T.: Measurement of Distance and Height in Images based on easy attainable Calibration Parameters. In: Proc. IEEE Intelligent Vehicles Symposium, pp. 314–319 (2000)
5. Neves, A.J.R., Martins, D.A., Pinho, A.J.: Obtaining the distance map for perspective vision systems. In: Proc. of the ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing (2009)
6. Mulayim, A.Y., Yilmaz, U., Atalay, V.: Silhouette-based 3-D model reconstruction from multiple images. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 33, 582–591 (2003)
7. Lee, D.J., Xu, X., Eifert, J., Zhan, P.: Area and volume measurements of objects with irregular shapes using multiple silhouettes. Optical Engineering 45,027202 (2006)
8. Vilaça, J.L., Fonseca, J.C., Pinho, A.M.: Calibration procedure for 3D measurement systems using two cameras and a laser line. Optics & Laser Technology 41, 112–119 (2009)
9. Mohamed, R., Ahmed, A., Eid, A., Farag, A.: Support Vector Machines for Camera Calibration Problem. In: 2006 IEEE International Conference on Image Processing, pp. 1029–1032 (2006)
10. Zhang, Z.: A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22, 1330–1334 (2000)
11. Memony, Q., Khanz, S.: Camera calibration and three-dimensional world reconstruction of stereo-vision using neural networks. International Journal of Systems Science 32, 1155–1159 (2001)
12. Xiaobo, C., Hanghang, F., Yinghua, Y., Shukai, Q.: The research of camera distortion correction basing on neural network. In: 2011 Chinese Control and Decision Conference (CCDC), pp. 596–601 (2011)
13. Tsai, R.Y.: A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off the shelf TV cameras and lenses. IEEE Joumal of Robotics and Automation RA-3, 323–344 (1987)
14. Bouguet, J.-Y.: `http://www.vision.caltech.edu/bouguetj/calib_doc/`
15. Weir, M.D., Thomas, G.B., Hass, J.: Thomas' Calculus. Addison-Wesley (2010)