

# Study on Target Positioning Systems Based on Image Geometry

Jieh-Shian YOUNG<sup>1,a</sup>, Yu-Wei HUANG<sup>1,b</sup>

<sup>1</sup>Institute of Vehicle Engineering, National Changhua University of Education,  
Changhua 500 Taiwan

<sup>a</sup>email: jsyoung@cc.ncue.edu.tw, <sup>b</sup>email: huangyw@cc.ncue.edu.tw

**Keywords:** CCD array, coordinate transformation, image geometry, target positioning.

**Abstract.** This study aims at developing an approach to the target positioning systems from images captured by camera based on its image geometry. This approach consists of the coordinate transformation between camera film coordinate system and fixed coordinate system, image geometry analysis, distribution analysis of CCD array, least mean square error (LMS) algorithm, etc. This paper also provides an associated algorithm to facilitate the proposed approach. A down scaled experiment will be set up to verify the results through the proposed approach. The results show that the proposed approach is feasible in achieving target positioning.

## Introduction

The object positioning by computer visual information has been studied throughout in numerous applications. Various approaches and techniques have been developed for this kind of problems both in applications and theories. Some used more than one camera so as to establish the position information of an object [1-4]. Luna et al studied a system to measure the 2-D position of an object that intercepts a plane in space [5]. Zhou proposed an algorithm to the ground-based object positioning from video stream by single camera equipped in unmanned aerial vehicle (UAV).

Calibration of cameras is one of the significant issues in computer vision. The camera's parameters that require calibrating may include intrinsic and extrinsic ones [6]. The 3D position and orientation of the camera frame relative to a certain world coordinate system are extrinsic while the internal camera geometric and optical characteristics are intrinsic. Although not all the techniques need any calibration object, the calibration of a camera is indispensable for some techniques of computer vision applications.

A sensor system for positioning objects by visual information with a single charge-coupled device (CCD) camera is presented in this paper. This paper originally employs the properties of the CCD array in a camera to evaluate the object positions with regular geometric lines, curves, or surfaces. The proposed technique does not have to apply any calibration object and can be regarded as the self-calibration, or the so-called 0D approach [7]. An overview of this approach can refer to [8] and the references therein. In this paper, we consider the radial distortion since it contributes the major errors to the distortions of the camera. This paper also elaborates the mathematical model of the camera systems in regard to the coordinate transformations between fixed coordinate and chip cell coordinate as well as constructs the computational procedure for the position evaluation system.

The notations in this paper are as follows:  $[\cdot]$  denotes a matrix;  $\vec{P}$  denotes the position vector of point  $P$  while  $\vec{O}_A$  denotes the origin position vector of  $A$ -coordinate;  $(x_A \ y_A \ z_A)_A^T$  denotes the  $x$ ,  $y$ , and  $z$  components of a position vector in  $A$  coordinate, that is,

$$\begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix}_A = \vec{O}_A + [\vec{i}_A \ \vec{j}_A \ \vec{k}_A] \begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix} = \vec{O}_A + x_A \vec{i}_A + y_A \vec{j}_A + z_A \vec{k}_A$$

where the unit vectors  $\vec{i}_A$ ,  $\vec{j}_A$ , and  $\vec{k}_A$  are the orthogonal bases of  $A$  coordinate in  $x$ ,  $y$ , and  $z$  axes, respectively. Hereafter in this paper, the position vector is in the fixed coordinate as the symbol  $A$  is omitted.

## Mathematical Preliminary

A spatial object can project onto the corresponding chip cell of the CCD array, or the pixel in an image taken by a camera. From the pinhole phenomenon, the object, the mapped pixel on the CCD array, and the lens center should be collinear after the calibration for the camera distortion. Two of them can formulate the line equation in algebra directly and the other satisfies the line equation. In general, we can use the fixed coordinate to represent these position vectors. The coordinate transformation is crucial from CCD array coordinate, or the film chip coordinate, to the fixed coordinate. Figure 1 shows the definitions of the fixed coordinate and the film chip coordinate for a camera, where  $\vec{O}$  and  $\vec{O}_F$  are the origins of the fixed coordinate and the film chip (CCD array) coordinate, respectively.

Let  $(x \ y \ z)^T$  and  $(\bar{x} \ \bar{y} \ \bar{z})_F^T$  denote any position represented in the fixed coordinate and the film chip coordinate, respectively,  $\vec{O}_M = (\vec{0} \ \bar{y}_M \ \bar{z}_M)_F^T$  denotes the mounted position of the camera, where  $\bar{y}_M$  and  $\bar{z}_M$  can be obtained from the specifications of the camera.  $\vec{O}_M$  can also be assigned above the origin of the fixed coordinate with zero  $x$ -axis and zero  $y$ -axis components, i.e.,  $\vec{O}_M = (\vec{0} \ 0 \ H)^T$  where  $H$  is the height of  $\vec{O}_M$ . According to the coordinate transformation,

$$\begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \\ 0 \\ H \end{bmatrix} = \begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_{F_0} \\ y_{F_0} \\ z_{F_0} \end{bmatrix} + \begin{bmatrix} \vec{i}_F & \vec{j}_F & \vec{k}_F \\ 0 \\ \bar{y}_M \\ \bar{z}_M \end{bmatrix}.$$

Figure 1. Definitions of the fixed coordinate and the film chip coordinate.

The origin of the film chip coordinate represented in the fixed coordinate can be

$$\vec{O}_F = \begin{pmatrix} x_{F_0} \\ y_{F_0} \\ z_{F_0} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ H \end{pmatrix} - T \begin{pmatrix} 0 \\ \bar{y}_M \\ \bar{z}_M \end{pmatrix},$$

where  $\begin{bmatrix} \vec{i}_F & \vec{j}_F & \vec{k}_F \end{bmatrix} = \begin{bmatrix} \vec{i} & \vec{j} & \vec{k} \end{bmatrix} \cdot T$ , and

$$T = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

and  $\psi$ ,  $\theta$ , and  $\phi$  denote the Euler angles of the camera for the yaw, pitch, and roll angles, respectively. Therefore, the coordinate transformation from the film chip coordinate to the fixed coordinate becomes

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_{F_0} \\ y_{F_0} \\ z_{F_0} \end{bmatrix} + T \begin{bmatrix} \bar{x} \\ \bar{y} \\ \bar{z} \end{bmatrix}.$$

In Figure 1, let  $\vec{L}$  represent the lens center position. From definition,

$$\vec{L} = \begin{pmatrix} 0 \\ 0 \\ f \end{pmatrix}_F = \begin{pmatrix} x_L \\ y_L \\ z_L \end{pmatrix} = \begin{pmatrix} x_{F_0} \\ y_{F_0} \\ z_{F_0} \end{pmatrix} + T \begin{pmatrix} 0 \\ 0 \\ f \end{pmatrix}. \quad (1)$$

The CCD array, or the film chip of a camera, consists of the CCD cells that sense the light energy to make an image collected by the corresponding pixels.  $U$  is the width of the CCD array,  $V$  is the

height of the CCD array,  $U_c$  is the width of the CCD cell,  $V_c$  is the height of the CCD cell, and  $N_u$  and  $N_v$  are the numbers of the CCD cells on the film chip in width and height directions, respectively. Let  $(P_x \ P_y)^T$  denote the index of the CCD cell on the chip where  $P_x = 1, 2, \dots, N_u$  and  $P_y = 1, 2, \dots, N_v$ .  $(P_x \ P_y)^T$  is the same as the pixel index of the image. Let  $\vec{P}_{P_y} \equiv (x_{P_y} \ y_{P_y} \ z_{P_y})^T$  be the position vector at the center of the chip cell with the pixel indexed by  $P_x$  and  $P_y$ . Then, the chip cell center position can be as follows.

$$\vec{P}_{P_y} = \begin{pmatrix} x_{P_y} \\ y_{P_y} \\ z_{P_y} \end{pmatrix} = \begin{pmatrix} \frac{(U-U_c)(2P_x-N_u-1)}{2(N_u-1)} \\ \frac{(V-V_c)(2P_y-N_v-1)}{2(N_v-1)} \\ 0 \end{pmatrix}_F = \begin{pmatrix} x_{F_0} \\ y_{F_0} \\ z_{F_0} \end{pmatrix} + T \begin{pmatrix} \frac{(U-U_c)(2P_x-N_u-1)}{2(N_u-1)} \\ \frac{(V-V_c)(2P_y-N_v-1)}{2(N_v-1)} \\ 0 \end{pmatrix}. \quad (2)$$

From (2), it can obtain the chip cell center position indexed by  $(P_x \ P_y)^T$  which is the pixel index of the image. Based on the pinhole model of a camera, an object from the scene projecting onto the CCD array, or the image, can be formulated as a line equation that makes 2 equations in 3-dimension (3-D) space. There should be  $n$  lines, or  $2n$  equations if there are  $n$  objects. In case that a set of the evaluated objects is in a regular curve, e.g., a line, a circle, etc,  $n$  objects will increase  $n+2$  unknowns including 2 regular curve parameters. Theoretically, the position evaluations for the objects are feasible since the equations increases are more than the unknowns increases.

### The Evaluations of Camera Parameters and Positions of Objects

A point  $\vec{P}$  of the objects which are collinear with the equation  $y = mx + y_0$  on the ground ( $z=0$ ) maps onto the CCD chip cell at a pixel indexed by  $(P_x \ P_y)^T$ . Any point  $\vec{P} = (x \ mx + y_0 \ 0)^T$  which is collinear with  $\vec{L}$  satisfies the line equation (2 plane equations) as follows.

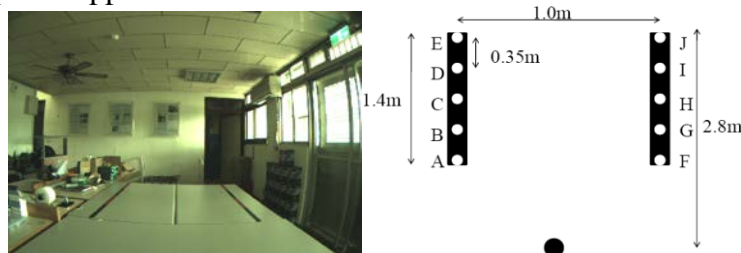
$$(x - x_L)(y_{P_y} - y_L) - (mx + y_0 - y_L)(x_{P_y} - x_L) = 0, \quad (3)$$

$$(x - x_L)(z_{P_y} - z_L) + z_L(x_{P_y} - x_L) = 0, \quad (4)$$

where  $x_L$ ,  $y_L$ ,  $z_L$ ,  $x_{P_y}$ ,  $y_{P_y}$ , and  $z_{P_y}$  are defined in (1) and (2). A set of points  $\vec{P}_i = (x_i \ mx_i + y_0 \ 0)^T$  ( $i=1, \dots, n$ ) projects onto a set of pixels  $(P_{x,i} \ P_{y,i})^T$  in the image, or the CCD array. There will be  $2n$  equations according to (3) and (4).  $\psi$ ,  $\theta$ ,  $\phi$ ,  $H$ , and  $y_0$  are the fixed variables for an image while the line parameter,  $m$ , is also a fixed variable. The least square approach can be defined as  $\min J$ , where  $J$  is the quadratic performance index and

$$J = \sum_{j=1}^{N_L} \sum_{i=1}^{N_j} [(x_i - x_L)(y_{P_{y,i}} - y_L) - (m_j x_i + y_{0j} - y_L)(x_{P_{y,i}} - x_L)]^2 + [(x_i - x_L)(z_{P_{y,i}} - z_L) + z_L(x_{P_{y,i}} - x_L)]^2 \quad (5)$$

In this research, the specifications of this camera (DH-HV2003UC) are as follows.  $U = 6.4\text{mm}$ ,  $V = 4.8\text{mm}$ ,  $U_c = 4.2\mu\text{m}$ ,  $V_c = 4.2\mu\text{m}$ ,  $N_u = 800$ ,  $N_v = 600$ ,  $\bar{y}_M = -27.25\text{mm}$ , and  $\bar{z}_M = 2.53\text{mm}$ . Based on the images taken by this camera, the radial distortion of this camera belongs to barrel type with negative distortion constant equal to  $-8.3042 \times 10^{-7}$  in pixel. The position evaluation of the lane marks is a practical example. Figure 2 shows the downscaled simulation scenario for the position evaluation of the lane marks while Figure 2(b) also illustrates the geometry of the platform. In this example,  $m_1 = m_2 = 0$ . There are 2 cases used to verify the feasibility of the proposed approach.



(a) the scenario platform. (b) the geometry for the scenario platform.

Figure 2. The downscaled simulations for the position evaluation of the lane marks.

Figure 3 shows the pictures taken by the camera for these 2 different cases. We can pick out the marked points as in Figure 2(b). Figure 3 also shows the evaluation results of the simulations in this

case studies. The evaluated positions of those marks can be solved by numerical methods recursively. These results are applicable although there are still errors between the exact positions and evaluated ones. The errors may come from the accuracy of the camera geometry, the aspheric lens of camera, the accuracy of the pixel indexes, etc.

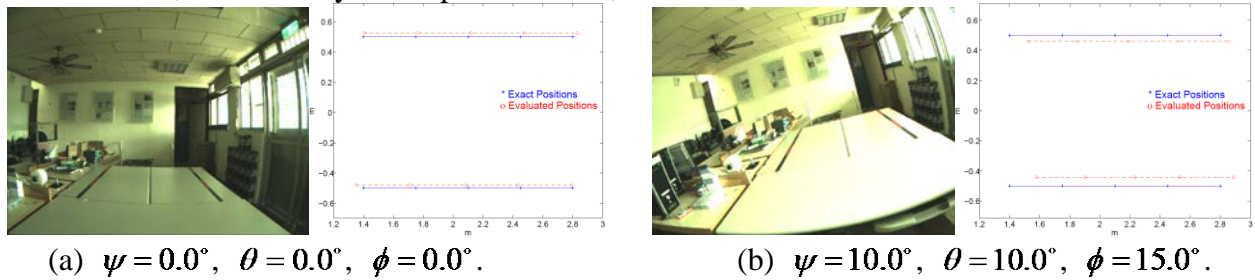


Figure 3. The result pictures and the evaluation results for 2 different cases.

## Conclusions

This paper presents an approach to evaluating the positions of objects by using a single camera. An image can provide the information of the objects captured on the CCD chip cells. For instance, a pixel can produce two equations and an additional variable. Some variables for the camera are fixed. If the objects are in regular geometry, a limit number of pixels information can form enough equations to solve the position evaluation problems. The properties of the CCD array serve the main reference dimensions for evaluating the positions of objects in an image. The accuracy of the position evaluations depends on the pixels of objects selected in an image while it is not easy to discern the exact pixels of objects in the image sometimes. That is, if the pixel indices cannot be exactly marked in an image for points, the results depending on its image coordinate will cause erroneous in objects positioning. The initial guess of the evaluated variables in (5) is also crucial and will affect the approached solutions of these variables. The future study can focus on the number of pixels selected in an image that can abate the accuracy of the initial guess in position evaluations of objects.

## Acknowledgement

This research work is sponsored by Ministry of Science and Technology under Grand NSC 100-2221-E-018-017 and MOST 103-2221-E-018-23.

## References

- [1] A. L. Rankin, A. Huertas, and L. H. Matthies, "Stereo vision based terrain mapping for off-road autonomous navigation " presented at the Proceedings of SPIE 2009, 2009.
- [2] M.-H. Chiang, H.-T. Lin, and C.-L. Hou, "Development of a stereo vision measurement system for a 3D three-axial pneumatic parallel mechanism robot arm," *Sensors*, vol. 11, pp. 2257-2281, 2011.
- [3] R. Richa, P. Poignet, and C. Liu, "Three-dimensional motion tracking for beating heart surgery using a thin-plate spline deformable model " *International Journal of Robotics Research*, vol. 29, pp. 218-230, 2010.
- [4] F. Dornaika and C. R. Chung, "Stereo Geometry From 3-D Ego-Motion Streams," *IEEE Transactions on Systems, Man, and Cybernetics* vol. 33, pp. 308-323, 2003.
- [5] C. A. Luna, J. L. Lázaro, M. Mazo, and A. Cano, "Sensor for high speed, high precision measurement of 2-D positions," *Sensors*, vol. 9, pp. 8810-8823, 2009.
- [6] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. 3, pp. 323 - 343, 1987.
- [7] Z. Zhang, "Camera calibration with one-dimensional objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, 2004.
- [8] A. Heyden and M. Pollefeys, "Multiple View Geometry," in *Emerging Topics in Computer Vision*, G. Medioni and S. B. Kang, Eds., ed: Prentice Hall, 2003, pp. 45-108.