

# GEOMETRIC DEEP LEARNED FEATURE CLASSIFICATION BASED CAMERA CALIBRATION

Cheolhyeong Park, Jisu Kim and Deekwoo Lee

Department of Computer Engineering, Keimyung University, Daegu, 42601,  
Republic of Korea

## ABSTRACT

*This paper chiefly focuses on calibration of depth camera system, particularly on stereo camera. Owing to complexity of parameter estimation of camera, i.e., it is an inverse problem the calibration is still challenging problem in computer vision. As similar to the previous method of the calibration, checkerboard is used in this work. However, corner detection is carried out by employing the concept of neural network. Since the corner detection of the previous work depends on the exterior environment such as ambient light, quality of the checkerboard itself, etc., learning of the geometric characteristics of the corners are conducted. The pro-posed method detects a region of checkboard from the captured images (a pair of images), and the corners are detected. Detection accuracy is increased by calculating the weights of the deep neural network. The procedure of the detection is de-tailed in this paper. The quantitative evaluation of the method is shown by calculating the re-projection error. Comparison is performed with the most popular method, Zhang's calibration one. The experimental results not only validate the accuracy of the calibration, but also shows the efficiency of the calibration.*

## KEYWORDS

*Calibration, Neural network, Deep learning, Re-projection error, Depth camera*

## 1. INTRODUCTION

This document describes, and is written to conform to, author guidelines for the journals of AIRCC series. It is prepared in Microsoft Word as a .doc document. Although other means of preparation are acceptable, final, camera-ready versions must conform to this layout. Microsoft Word terminology is used where appropriate in this document. Although formatting instructions may often appear daunting, the simplest approach is to use this template and insert headings and text into it as appropriate. In the past a few years, practical applications of computer vision have been increase in the areas of automotive engineering, medicine, security, etc[1]. Estimation of 3D coordinates of the 3D real world scene from multiple 2D images is the central part of computer vision technology. Quality of the estimation of 3D coordinates (or depth) highly depends on establishing the relationship between 3D coordinates and 2D coordinates and estimating distortion parameters of lenses of cameras. These procedures are called camera calibration. This paper proposes the approach to increasing accuracy of the calibration. In particular, the procedure of corner detection of checkerborard, applies deep neural network. In the past decades, one of the most popular calibration algorithms is Zhang's method, and the crucial part is calculating homography matrix[2]. Homography matrix can be defined as four categories. The first one is that homography is established by calculating a relationship between 2D and 3D coordinates. The second one is that the homography is established by calculating a relationship between 2D coordinates of images. The third and the fourth one is established using the relationships of straight lines contained in a pair of images, and these are detailed in Zhang's

Dhinaharan Nagamalai et al. (Eds) : ACSIT, SIPM, ICITE, ITCA - 2019  
pp. 43-49, 2019. © CS & IT-CSCP 2019

DOI: 10.5121/csit.2019.90305

paper [2]. The most popularly used method of calibration is using objects that have known patterns, e.g., grid pattern, circular pattern, etc. In general, multiple captures of the objects from different locations are carried out followed by fully exploiting geometric relationship between patterns in images. Contrary to the method explained above, calibration can be performed without specially designed objects and the method is called self-calibration. Since the calibration using the patterned objects shows limitations in terms of computation time. Experimental environment also affects to the calibration accuracy. However, self-calibration can alleviate the limitations of the existed calibration methods because it does not depend on the experimental conditions. Concerning all above, calibration can be categorized into two methods, conventional calibration and self-calibration. Self-calibration has brought great attention in the past a few years because many practical applications such as vehicles, robotics and medicines desire real-time and efficient calibration system. Conventional calibration has limitations as follows. In the first, multiple captures in good experimental condition are required. Cameras capture patterned objects at least 5 times from different locations (usually more than 10 captures). In the second, calibration needs iterative optimization procedures. Since the optimization procedure is basically inverse problems, the result can be overfitting leading to unexpected wrong results. In general, self-calibration chiefly deals with epipolar constraints defined from geometric relationship between 3D objects and 2D images. From the geometric relationship, fundamental matrix is estimated, and the matrix is one of the criterion decides the accuracy of calibration accuracy [3,4,5,6]. The conventional calibration still shows higher accuracy because the algorithm has been already stabilized and the calibration is usually carried out in almost ideal conditions with the known patterns such as square, circular (periodic and aperiodic) ones [7].

In this paper, we propose the approach to increasing the accuracy of camera calibration. Conventional approach, using patterned objects, with a concept of deep neural network (deep learning) is employed. Deep learning is applied in the course of corner detection. Although checkerboard is used in this work, by applying deep neural network, successful trained parameters of the network lead to fewer number of captures of checkerboard images. In particular, this paper shows that only two captures maintained or outperforms the existed calibration method in terms of re-projection error. Training of the corners are carried out in 3 hidden layers each of which contains 5 steps, Prepare, FAPL, PAMFG, Reconstruction and PADCROP [8]. Evaluation result is compared to the Zhang's method and the criterion for comparison is re-projection error. The rest of the paper is organized as follows. Section 2 introduces deep learning based calibration, the most contribution of this paper, followed by substantiating the proposed algorithm by showing experimental results (including comparison results) in section 3. Section 4 concludes this paper.

## 2. INTRODUCTION TO PROJECTION MATRIX

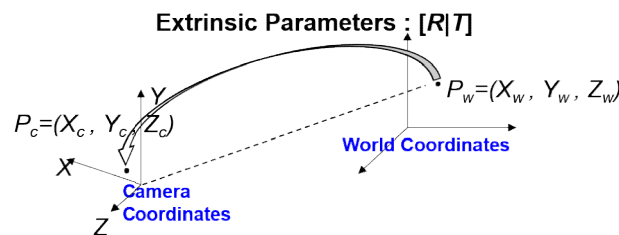


Figure 1. Extrinsic parameters define relationship between camera and real world coordinates

Camera calibration provides geometric relationship between coordinates of real world and ones of camera domain, or between the cameras themselves, and it also provides internal characteristics of the cameras. The former ones are called the extrinsic parameters and the latter ones are called

the intrinsic parameters. The extrinsic parameters are composed of a rotation matrix( $R$ ) and a translation vector( $T$ ).  $R$  and  $T$  are represented using  $3 \times 3$  matrix and  $3 \times 1$  vector, respectively (Figure. 1). The intrinsic parameters are composed of focal length, skew factor and principal point. Intrinsic parameters also deal with lens distortion. Concerning the extrinsic and the intrinsic parameters, the relationship between 2D and 3D coordinates can be written as follows. Please note that the first paragraph of a section or subsection is not indented. The first paragraphs that follows a table, figure, equation etc. does not have an indent, either. Subsequent paragraphs, however, are indented.

$$sm = PM, \quad (1)$$

where  $s$  is a scaling factor, and  $m$  and  $M$  are 2D and 3D coordinates (homogeneous), respectively.  $P$  is composed of the intrinsic parameters  $A$  and the extrinsic parameters  $[R|T]$ , i.e.,  $P=A[R|T]$ . The intrinsic parameters are composed of the focal length, skew factor and principal point's coordinates in image plane. Skew factor describes geometric distortion of the axes of sensors, i.e., if there is no distortion skew factor is zero. Once we have the extrinsic parameters, relative distances of real world scene of interest can be estimated. If the absolute distance value is required, the intrinsic parameters are needed in addition to the extrinsic parameters. The absolute value of depth is written as

$$z = bf/d, \quad (2)$$

where  $z$  is a distance between the optical center and the 3D point of interest,  $b$  is a distance between optical centers of cameras,  $d$  is a disparity value resulted from parallax of two view points and  $f$  is the focal length of cameras (focal length of cameras are assumed to be equivalent here). In other words, camera calibration is a process of estimating the values in  $P$ . Cameras are assumed to be identical, but normalized coordinates are sometimes needed if different cameras are used [9].

### 3. INTRODUCTION TO PROJECTION MATRIX

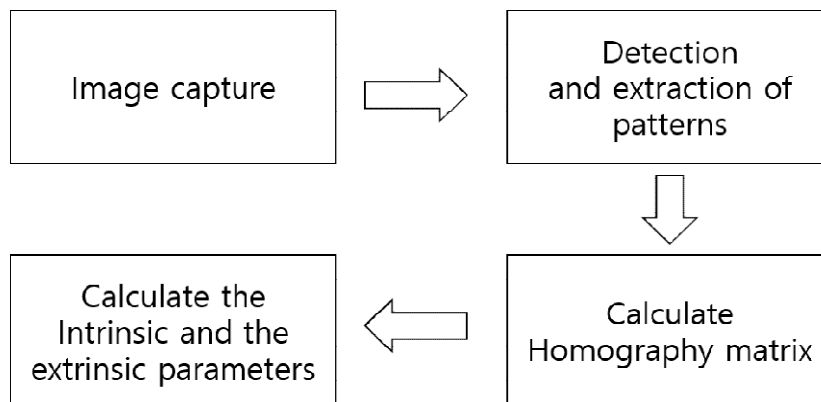


Figure 2. Conventional approach for camera calibration.

Figure 2 describes overall flow of conventional calibration of depth camera. More de-tails about conventional calibration methods are explained in [2, 7, 10, 11]. Among the methods proposed in the past few decades, Zhang's method is considered the most reliable in practice. In the method, Homography estimation is also one of the most important procedures. Homography matrix defines the geometric relationship between a pair of images. To achieve successful estimation of Homography matrix, feature extraction from the images is important. Homography between two images can be represented as follows.

$$\begin{pmatrix} \mathbf{x}_L \\ \mathbf{y}_L \\ \mathbf{w}_L \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \\ H_{31} & H_{32} & H_{33} \end{pmatrix} \begin{pmatrix} \mathbf{x}_R \\ \mathbf{y}_R \\ \mathbf{w}_R \end{pmatrix}, \quad (3)$$

where  $(x_L \ y_L \ w_L)^T$  and  $(x_R \ y_R \ w_R)^T$  are coordinates of 2D images from left and right camera, respectively. Since homogeneous coordinates system is used,  $w_L$  and  $w_R$  can be omitted or set as a constant number. Once homography is estimated, nonlinear optimization is applied to calculate calibration parameters. Calibration accuracy is evaluated using re-projection error. In this section, corner detection from checkerboard is performed based on deep neural network. In the first, checkerboard from whole image is extracted and the corner is detected from learned and weighted parameters.

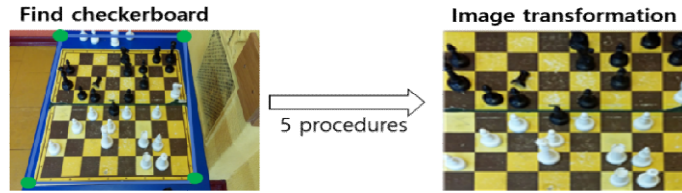


Figure 3. Learning based detection of chessboard and geometric transformation of an image.

Detection of a chessboard and of patterns lead to learning of the detection of corners in checkerboard for calibration. In Figure 3, geometric image transformation is carried out to accurately find corners. Prior to the transformation, a number of procedures is conducted. The procedures (5 procedures in Figure 3) are depicted in Figure 4.

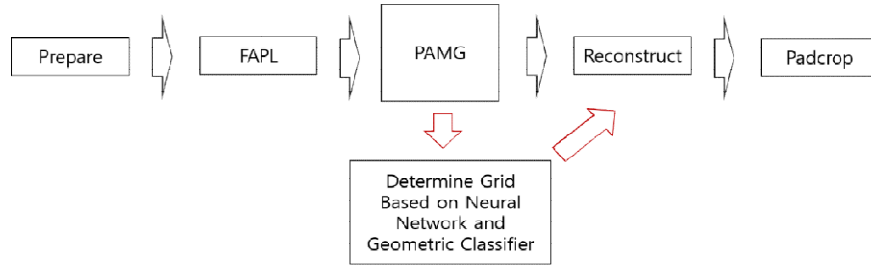


Figure 4. Learning procedures for corner detection.

As shown in Figure 3, 4 detection and extraction of corners from images employ deep learning algorithm. In the neural network in this paper, 3 hidden layers are used. Each layer processes Prepare, FAPL, PAMG, Reconstruct and Padcrop. Preapre is a preprocess to find straight lines in an image using retinex algorithm by subtracting unnecessary parts from the image [12]. FAPL extracts straight lines using retinex from the preprocessed image. In this procedure, extraction of the lines is carried out using image processing techniques such as “overexpose”, “blur” and “tint”. PAMG detects corners from the image, and the geometric image transformation can be done using Padcrop. In PAMG, neural network and geometric classifier achieves high accuracy of corner detection. By learning the structure of corners, the detection is success-fully accomplished from warped images. The activate function generate two outputs, the one is success and the other one is failure of detection. In the process of Reconstruct, geometric transformation of grid patterns is conducted as shown in the Figure 3 (right). To detect grid patterns, grid density formula is used, written as where  $p$  is a number of corners in images,  $n$  is an area and  $S$  represents density. Detection algorithm select corners that have the higher density. In the next, experiments of the proposed approach are provided.

$$S = \frac{p^3}{n \log(n)}, \quad (4)$$

## 4. EXPERIMENTS

This section provides the experimental results and the comparison results with Zhang's calibration method. Since a size of image affects to the computational complexity of an algorithm, in Table 1, size of images used in this work is provided.

Table 1. Image size and a number of squares in images

Status of images	Size	Number of corners
Before image transformation (before Padcrop)	1200 × 1200	8 × 7
After image transformation (after Padcrop)	1200 × 1200	8 × 7

Detection of chessboard that is used in the process of learning employs neural network algorithm. In total, 9,664 images are used as training set, and the learning is conducted in the cases of existence of corners and non-existence of corners. The former case is defined as "Train\_ok" and the latter one is defined as "Train\_no". There are 4,732 are used as "Train\_ok" and are 4,932 are used as "Train\_no". The results of learning are shown in Figure 5.

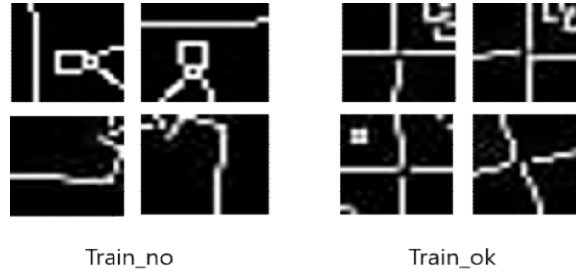


Figure 5. Training data for the cases of existence and non-existence of corners for calibration Table 2 provides the results of calibration using re-projection error when using a pair of images, i.e., stereo camera system. The proposed method and the existed method are compared. As shown in the Table 2, the proposed approach achieves the lower re-projection error.

Table 2. Heading and text fonts

Algorithm	Re-projection error
Zhang & Bouguet [2, 13]	1.9583
Proposed method	1.2735

In Table 3, the accuracy of calibration results by increasing iteration for optimization is provided. As shown in Table 3, accuracy is slightly increased.

Table 3. Accuracy (true-positive rate) with a number of iterations.

Number of iterations	Accuracy (True-positive)
50	0.9993
100	0.9996

## 5. CONCLUSION

In this paper, we proposed the approach to camera calibration by employing the concept of deep neural network. To increase performance of calibration, deep learning is applied to the procedure of corner detection when checkerboard is used for the calibration work. The proposed approach not only detects the region of the board, but also extracts corners even if the images are warped. In particular, the proposed approach is promising because re-projection error has been decreased even though a number of captures has been significantly decreased. This work can contribute to simple and efficient calibration algorithm in diverse applications that require real-time computer vision techniques.

## ACKNOWLEDGEMENTS

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (2016-0-00564, Development of Intelligent Interaction Technology Based on Context Awareness and Human Intention Understanding)

## REFERENCES

- [1] Liu, Y., Cheng, Y. and Wang, W., (GIIC 2018) "A survey of the application of deep learning in computer vision." In: Proceedings Volume 10835, Global Intelligence Industry Conference, SPIE, Beijing, China.
- [2] Zhang, Z.: A Flexible New Technique for Camera Calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(11), 1330-1334 (2000).
- [3] Kim, S-Y. and Han, J-H.: A Constrained Self - Calibration Technique. *Journal of KISS Software and Applications* 28(4), 358-368 (2001).
- [4] Zhao, B. and Hu, Z.: Camera self-calibration from translation by referring to a known camera. *Applied Optics* 54(25), 7789-7798 (2015).
- [5] Duan, S., Zang, H., Xu, M., Zhang, X., Gong, Q., Tian, Y., Liang, E. and Liu, X.: Camera self-calibration method based on two vanishing points. In: Proc. SPIE 9675, AOPC 2015: Image Processing and Analysis, LNCS, pp. 96752, Beijing, China (2015)
- [6] Luong, Q.-T. and Faugeras, O.D.: Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices. *International Journal of Computer Vision* 22(3), 261-289 (1997).
- [7] Heikkila, J.: Geometric camera calibration using circular control points. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(10), 1066-1077 (2000).
- [8] Czyzewski, M.A.: An Extremely Efficient Chess-board Detection for Non-trivial Photos. arXiv:1708.03898v1 (2018).
- [9] Kim, D-G., *OpenCV Computer Vision Programming*. 1st edn. Kame Publishing, Seoul (2014).
- [10] Tsai, R.: A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal on Robotics and Automation* 3(4) 323-344 (1987).
- [11] Kim, T. and Park, T.: Calibration Method between Two 3D LIDARs for Autonomous Vehicle. In: 2012 KSAE Annual Conference Proceedings, LNCS, pp. 827-827, Gyeonggi Province, South Korea (2012).

- [12] Provenzi, E., Carli, L., Rizzi, A. and Marini, D.: Mathematical definition and analysis of the Retinex algorithm. *Journal of the Optical Society of America* 22(12), 2613-2621 (2005).
- [13] Leiner Barba, J. Lorena Vargas, Cesar Torres, Q.M. and Lorenzo Mattos, V.: Three-Dimensional Reconstruction Optical System Using Shadows Triangulation. *AIP Conference Proceedings, LNCS, Vol 992*, pp. 1073-1077, Campinas, Brazil (2008).

### Author

**Cheolhyeong Park** is in department of computer engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on geometric image analysis, computer vision, computer graphics and machine learning. He is in the course of integrated B.S and M.S degree in computer engineering.



**Jisu Kim** is in department of computer engineering, Keimyung University, Daegu, Republic of Korea. He is currently working on image processing, computer vision, signal processing and machine learning. He is currently pursuing his M.S degree in computer engineering.



**Dr. Deokwoo Lee** is an Assistant Professor in the department of computer engineering at Keimyung University. Dr. Lee has received B.S degree in electrical engineering from Kyungpook National University, Daegu, Republic of Korea, and M.S and Ph.D degree from North Carolina State University, Raleigh, NC, USA, respectively. He has been working on the areas of computer vision, image processing, signal processing and machine learning. In particular, he has been conducting camera calibration, bio-signal analysis and image denoising.

