International Journal of Fuzzy Logic and Intelligent Systems, vol. 7, no. 3, September 2007 pp. 216-220

Robust Camera Calibration using TSK Fuzzy Modeling

Heesung Lee, Sungjun Hong and Euntai Kim

C613, School of Electrical and Electronic Engineering, Yonsei University
134 Shinchon-dong, Seodaemun-gu, Seoul, Korea, 120-749
(TEL:02-2123-2863, FAX:02-313-2879, E-mail:etkim@yonsei.ac.kr)

Abstract

Camera calibration in machine vision is the process of determining the intrinsic camera parameters and the three-dimensional (3D) position and orientation of the camera frame relative to a certain world coordinate system. On the other hand, Takagi-Sugeno-Kang (TSK) fuzzy system is a very popular fuzzy system and approximates any nonlinear function to arbitrary accuracy with only a small number of fuzzy rules. It demonstrates not only nonlinear behavior but also transparent structure. In this paper, we present a novel and simple technique for camera calibration for machine vision using TSK fuzzy model. The proposed method divides the world into some regions according to camera view and uses the clustered 3D geometric knowledge. TSK fuzzy system is employed to estimate the camera parameters by combining partial information into complete 3D information. The experiments are performed to verify the proposed camera calibration.

Key Words : Camera Calibration, TSK fuzzy model, Machine-vision

1. Introduction

In machine vision, it is important for the sensory system to passively sense the three-dimensional (3-D) structure of its surrounding environment. A common method is through disparity analysis using two images and stereo vision. However, the difficult problem in using disparity is to determine the correspondence of features between two images. Once sufficient correspondence is known, depth information of objects in the scene can be computed by measuring the spatial disparity of image features acquired by two calibrated cameras [1]. Therefore, Camera calibration is a preliminary step toward machine vision in order to extract metric information from 2D image.

Camera calibration is necessary step in computer vision in order to extract metric information from 2D images. Much work has been done, starting in the photogrammetry community and more recently, in computer vision and various methods for calibrating cameras can be found from the literatures [1-4, 8-12]. However, to our knowledge, there does not exist any calibration technique reported in the literatures which use Takagi-Sugeno-Kang (TSK) fuzzy system and this is the topic we will investigate in this paper. TSK fuzzy system is a very popular fuzzy system and approximates any nonlinear function to arbitrary accuracy with only a small number of fuzzy rules [5]. It demonstrates not only nonlinear behavior but also transparent structure. In this paper, we present a novel and simple technique for camera calibration for machine vision using TSK fuzzy model. The proposed method divides the world into some regions according to camera view and uses the clustered 3D geometric knowledge. TSK fuzzy system is employed to estimate the camera parameters by combining partial information into complete 3D information. The rest of this paper is organized as follows. In Section 2. we give some background including the camera calibration and TSK fuzzy modeling. In Section 3, New camera calibration method using TSK approach is proposed, In Section 4, experimental results are given to show the performance of the proposed method. Finally, the conclusion is drawn in Section 5.

Manuscript received Nov. 29, 2006; revised Aug. 31, 2007.
This work was supported by the Ministry of Commerce, Industry and Energy of Korea (HISP).

2. Background

2.1 Camera Calibration

The direct linear transformation (DLT) facilitates a perspective transformation between two-dimensional image space data and three-dimensional object space. The DLT combined into a single linear model the two-dimensional affine transformation from picture reader to image coordinates and the transformation form image to three-dimensional object space coordinates via the collinearity model. The basic projective equations of the DLT are as follows:

$$u+\triangle u=\frac{P_{11}X+P_{12}Y+P_{13}Z+P_{14}}{P_{31}X+P_{32}Y+P_{33}Z+1} \quad (1)$$

and

$$v+\triangle v=\frac{P_{21}X+P_{22}Y+P_{23}Z+P_{24}}{P_{31}X+P_{32}Y+P_{33}Z+1} \quad (2)$$

where u and v are the image coordinates or pixel coordinates. The 11 parameters $P_{11}, P_{12}, \ldots, P_{33}$ can be physically interpreted in therms of the interior and exterior orientation of the image, though the parameters are not strictly equivalent to the perspective parameters of the collinearity equations. Moreover, linear dependencies exist between the 11 parameters and a factor which is taken into account an alternative DLT formulation in which two orthogonality constraints are imposed in the 11 parameter transformation.

Application of the DLT has proved popular for the restitution of non-metric photography since no a priori knowledge of the interior orientation elements is required. In the digital camera context, the DLT can give two advantages. Firstly, a non-iterative, direct solution is achieved, and this offers the fast computation. Secondly, the affine coordinate correction implicit is quire appropriate for CCD [6].

If M is the world coordinates (X, Y, Z) and projects onto a point m that is the pixel coordinates (u, v), we assume a number of point correspondences $M_i \leftrightarrow m_i$ $((X, Y, Z)_i \leftrightarrow (u, v)_i)$ between 3D world points (X, Y, Z) and 2D image points (u, v) are given. We are required to find a camera matrix P, namely a 3×4 matrix such that $m_i = PM_i$ $((u, v)_i = P(X, Y, Z)_i)$. For each correspondence $M_i \leftrightarrow m_i$ $((X, Y, Z)_i \leftrightarrow (u, v)_i)$, we can derive a relationship

$$\begin{pmatrix} 0^T & -h_i(X, Y, Z)_i^T & v_i(X, Y, Z)_i^T \\ h_i(X, Y, Z)_i^T & 0^T & u_i(X, Y, Z)_i^T \\ -v_i(X, Y, Z)_i^T & u_i(X, Y, Z)_i^T & 0^T \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = 0 \quad (3)$$

or

$$\begin{pmatrix} 0^T & -h_i M_i^T & v_i M_i^T \\ h_i M_i^T & 0^T & u_i M_i^T \\ -v_i M_i^T & u_i M_i^T & 0^T \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = 0 \quad (4)$$

where each P_i is a 4-vector, the ith row of P. Alternatively, one may choose to use only the first two equations:

$$\begin{pmatrix} 0^T & -h_i(X, Y, Z)_i^T & v_i(X, Y, Z)_i^T \\ h_i(X, Y, Z)_i^T & 0^T & u_i(X, Y, Z)_i^T \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = 0 \quad (5)$$

or

$$\begin{pmatrix} 0^T & -h_i M_i^T & v_i M_i^T \\ h_i M_i^T & 0^T & u_i M_i^T \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = 0 \quad (6)$$

since the three equations of (3) are linearly dependent. From a set of n point correspondences, we obtain $2n \times 12$ matrix A by stacking up the equations (5) for each correspondence. The projection matrix P is computed by solving the set of equations $AP = 0$, where P is the vector containing the entries of the matrix P. The algorithm for estimating P from world to image point correspondences is given as follows [7].

(1) Linear Solution: Compute an initial estimate of P using a linear method

(1.1) Normalization: Use a similarity transformation (T) to normalize the image points and a second similarity transformation (U) to normalize the space points.

(1.2) DLT: From the $2n \times 12$ matrix A by stacking the (5) generated by each correspondence. Write P_i for the vector containing the entries of the matrix $\hat{P}$. A solution of $AP = 0$, subject to $\| P \| = 1$, is obtained from the unit singular vector of A corresponding to the smallest singular value.

(2) Minimize geometric error: Using the linear estimate as a starting point minimize the geometric error $\sum_i d(T(u, v)_i, \hat{P}U(X, Y, Z)_i)$.

(3) Denormalization: The camera matrix for the original coordinates is obtained from $\hat{P}$ as $P = T^{-1}\hat{P}U$.

2.2 TSK Fuzzy Modeling

The fuzzy model suggested by Takagi and Sugeno in 1985 can represent or model a general class of static or dynamic nonlinear system. It is based on fuzzy partition of input space and it can be viewed as the expansion of piecewise linear partition. It is constructed from the following rules:

$$\begin{aligned} &IF\ x_1\ is\ C_1^l \text{ and} \ldots \text{and} x_n\ is\ C_n^l, \\ &THEN\ y^l = f_l(x_1, x_2, \ldots, x_n; c^l) \\ &\qquad = c_0^l + c_1^l x_1 + \ldots + c_n^l x_n \end{aligned} \quad (7)$$

for $l = 1, 2, \ldots, M$, where M is the number of rules, C_i^l is the fuzzy set and c_i^l is the parameter set in the consequent. That is, the IF parts of the rules are the same as in the ordinary fuzzy IF_THEN rules, but the THEN parts are linear combinations of the input variables.

The predicted output of the fuzzy model is computed as the weighted average of the y^ls in (8), that is

$$\bar{y} = \frac{\sum_{l=1}^{M} y^l w^l}{\sum_{l=1}^{M} w^l} = \frac{\sum_{l=1}^{M} (c_0^l + c_1^l x_1 + \ldots + c_n^l x_n) w^l}{\sum_{l=1}^{M} w^l} \quad (8)$$

where y^l is the output of the lth rule, w^l is the lth rule's firing strength, which is obtained as the minimum of the fuzzy membership degrees of all fuzzy variables.

The physical meaning of the rule (7) is that when is constrained to the fuzzy range characterized by the IF part of the rule, the output is a linear function of the input variable. Therefore, the TSK fuzzy system can be viewed as a somewhat piece-wise linear function, where the change from one piece to the other is smooth rather than abrupt [5]. In this paper, the local regions divided according to camera view are constrained to the fuzzy range and complete 3D information is achieved by combining of the partial information. The detailed algorithms are introduced in the following section.

3. TSK Approach to Camera Calibration

The camera model that we consider is the perspective projection model based on the pinhole model. If M has world coordinates (X,Y,Z) and projects onto a point m that has pixel coordinates (u,v), the operation can be described, in homogeneous coordinates, by the following equation:

$$\begin{pmatrix} hm \\ h \end{pmatrix} = \begin{pmatrix} hu \\ hv \\ h \end{pmatrix} \tag{9}$$

or

$$P\begin{pmatrix} M \\ 1 \end{pmatrix} = P\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} P_1^T P_{14} \\ P_2^T P_{24} \\ P_3^T P_{34} \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \tag{10}$$

where matrix P is commonly referred to as perspective projection matrix and decomposed into two matrices: $P=BD$ where

$$D=\begin{pmatrix} R & t \\ 0_3^T & 1 \end{pmatrix} = \begin{pmatrix} \alpha_u & -\alpha_u/\tan\theta & u_0 & 0 \\ 0 & \alpha_v/\sin\theta & v_0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{11}$$

The 4×4 matrix D represents the mapping from world coordinates to camera coordinates and accounts for six extrinsic parameters of the camera: three for the rotation R which is normally specified by three rotation angles and three for the translation t. the 3x4 matrix B represents the intrinsic parameters of the camera: the scale factors α_u and α_v, the coordinates u_0 and v_0 of the principal point and the angle θ between the image axes. The benefit from this would be that the calibration accuracy will not only be increased, but this will also allow us to maintain the simple relation in (7) thus making following vision tasks easier [8].

In this paper, the proposed method divides the world into some regions according to camera view as shown in Fig. 1 and TSK fuzzy system is employed to combine the partial perspective projection matrices in clustered regions.

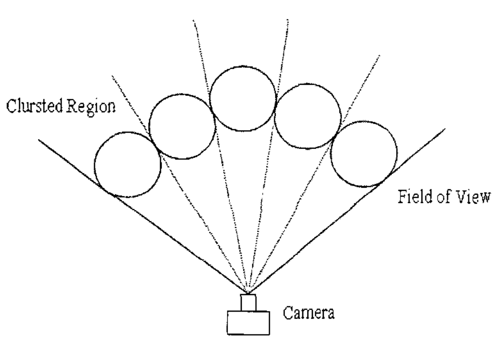

Fig. 1. The clustered regions

The TSK fuzzy model is composed of the following rules

$$\begin{aligned} &R^1: IF\ a\ is\ IC^1\ THEN\ y^1 = P^1 x^1 \\ &R^2: IF\ a\ is\ IC^2\ THEN\ y^2 = P^2 x^2 \\ &\qquad \vdots \\ &R^l: IF\ a\ is\ IC^l\ THEN\ y^l = P^l x^l \end{aligned} \tag{12}$$

where l is the numerical order of cluster regions, R^l denotes the fuzzy rule, a is an angle of the cluster region, y^l is the partial image point, x^l is the input linguistic variables to represent the world coordinates 3D point which belong to the 1-cut of IC^l and P^l is the lth clustered perspective projection matrix which is computed in the lth clustered region. Shown in Fig. 2. are the membership function of IC^l.

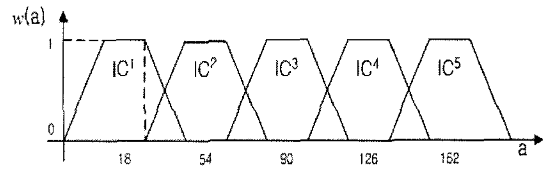

Fig. 2. The membership function

From the fuzzy model (12), we can compute the complete image point, $\bar{y}$, which combines partial 3D information is computed by the weighted average of the y^l defined as follow:

$$\bar{y} = \frac{\sum_{l=1}^{M} y^l w^l}{\sum_{l=1}^{M} w^l} = \frac{\sum_{l=1}^{M} P^l x_i^l w^l}{\sum_{l=1}^{M} w^l} \tag{13}$$

where x_i^l is the input linguistic variables to represent the world coordinates 3D point, P^l is the lth clustered perspective projection matrix which is computed in the lth clustered region and y^l is the partial image point using P^l.

4. Experimental Result

Experiments are conducted using the calibration patterns as shown in Fig. 3. The images (640 x 480) are obtained in five clustered region and consist of a box patch, with size 500mm x 500mm.

Table 1 and Fig. 4 show the perspective projection error of the proposed method. The error, root mean squared calibration error (RMSE), measures have been computed using the proposed method and by conventional direct linear transform (DLT) calibration algorithm [7] and test point pairs are not included in calibration procedure. It can be seen from this table that proposed method is better than the conventional calibration technique.

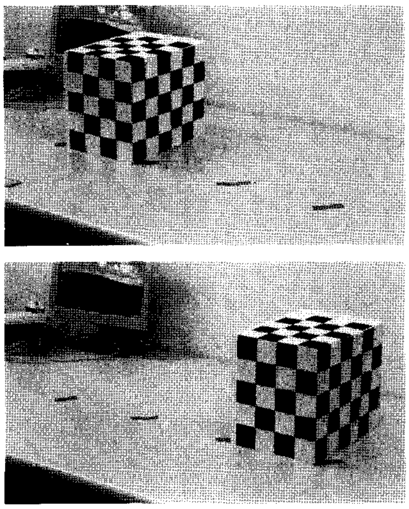

Fig. 3. The calibration patterns

Table 1. Perspective projection error (pixel)

Measure	Proposed	DLT
RMSE	3.0759	4.8201

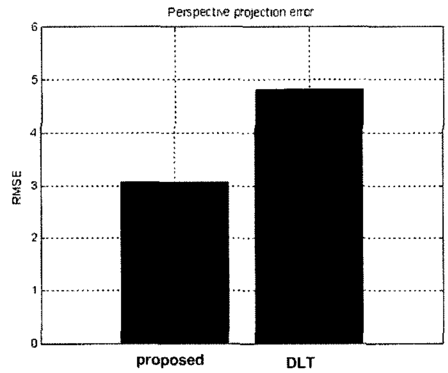

Fig. 4. Perspective projection error (pixel)

Table 2. Calibration error in reconstruction 3D points (cm)

Measure	Proposed	DLT
RMSE	1.9591	5.3177

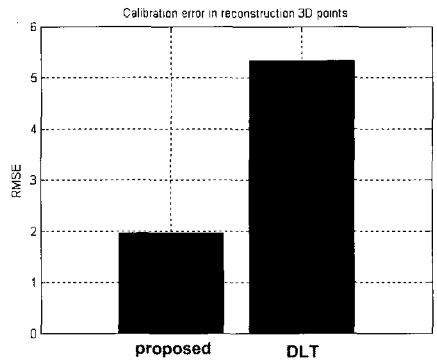

Fig. 5. Calibration error in reconstruction 3D points (cm)

Generally, with real images, the accuracy of the calibration is measured in terms of the accuracy in reconstruction 3D points through triangulation [9]. Clearly, the measure shown in Table 2 and Fig. 5 is in favor of the proposed method.

5. Conclusion

Camera calibration is necessary step in computer vision in order to extract metric information from 2D images. It is procedure of determining the internal camera geometric and optical characteristics and the 3D position and orientation of the camera frame relative to a certain world coordinate system. Much work has been done, starting in the photogrammetry community and more recently, in computer vision. In this paper, we propose novel and simple technique for camera calibration using TSK fuzzy model and describe its efficiency. There does not exist any calibration technique reported in the literatures which use TSK fuzzy system and this is the topic we will investigate in this paper. The proposed method divides the world into some regions according to camera view and TSK fuzzy system is employed to combine the partial perspective projection matrices in clustered regions. The experimental result shows that the efficiency of the proposed method.

References

[1] O. Faugeras, *Three-dimensional computer vision: a geometric viewpoint* , MIT press, 1993.

[2] Q.-T. Luong and O.D. Faugeras, "Self-Calibration of a Moving Camera from Point Correspondences and Fundamental Matrices," *Int. J. Computer Vision*, vol. 22, no. 3, pp. 261-289, 1997.

[3] Z. Zhang, "A Flexible New Technique for Camera Calibration," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330-1334, Nov. 2000.

[4] Z. Zhang, "Camera Calibration with One Dimensional Objects," *Proc. European Conf. Computer Vision*, vol. 4, pp. 161-174, May 2002.

[5] L. Wang, *A Course in fuzzy systems and control*, Prentice-Hall, 1997.

[6] A. Gruen and T. Huang, *Calibration and orientation of cameras in computer vision*, Springer, 2001.

[7] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge university press, 2004.

[8] M. Ahmed, E. Hemayed and A. Fagag, "Neurocalibration : a neural network that can tell camera calibration parameters," *Proc. of the Seventh IEEE Int. Conf.*, Vol. 1, 20-27, Sept. 1999.

[9] J. Weng, P. Cohen and M. Herniou, "Camera Calibration with distortion models and accuracy evaluation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, No. 10, Oct 1992.

[10] R. 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. RA-3, No. 4, Aug 1987.

[11] D. Choi, S. Oh, H. Chang and K. Kim, "Nonlinear camera calibration using neural networks," *Neural, parallel and scientic computations*, Vol. 2, No. 1, March 1994.

[12] M. Lee and J. Lee, " A 2-D image camera calibration using a mapping approximation of multi-layer perceptrons," *Journal of Control Automation and system engineering*, Vol. 4, pp. 487-493, 1998.

Heesung Lee
was born in Seoul, Korea, in 1979. He received the B.S. and the M.S. degrees in electrical and electronic engineering, from Yonsei University, Seoul, Korea, in 2003 and 2005, respectively. He is currently a Ph. D. candidate of Dept. of electrical and electronic engineering at Yonsei University. His current research interests include computational intelligence, pattern recognition, biometrics, and robot vision.

Sungjun Hong
was born in Jeju, Korea, in 1981. He received the B.S. degrees in electrical and electronic engineering and computer science, from Yonsei University, Seoul, Korea, in 2005. He is a graduate student of the combined master's and doctoral degree programs at Yonsei University. He has studied machine learning, biometrics and optimization.

Euntai Kim
was born in Seoul, Korea, in 1970. He received the B.S. (summa cum laude) and the M.S. and the Ph.D. degrees in electronic engineering, all from Yonsei University, Seoul, Korea, in 1992, 1994 and 1999, respectively. From 1999 to 2002, he was a full-time lecturer in the Department of Control and Instrumentation Engineering at Hankyong National University, Kyonggi-do, Korea. Since 2002, he has joined the faculties of the School of Electrical and Electronic Engineering at Yonsei University, where he is currently an associate professor. His current research interests include computational intelligence and its application to intelligent service robot and intelligent home network. Dr. Kim is an associate editor of International Journal of Control and Systems.