韓國精密工學會 '95年度 秋季學術大會論文集 pp.510~514

신경회로망을 이용한 카메라 교정과 2차원 거리 측정에 관한 연구

Neural Network Based Camera Calibration and 2-D Range Finding

정우태*, 고국원**, 조형석**

*한국원자력 연구소, **한국과학기술원

Abstract:

This paper deals with an application of neural network to camera calibration with wide angle lens and 2-D range finding. Wide angle lens has an advantage of having wide view angles for mobile environment recognition ans robot eye in hand system. But, it has severe radial distortion. Multi-layer neural network is used for the calibration of the camera considering lens distortion, and is trained it by error back-propagation method. MLP can map between camera image plane and plane the made by structured light. In experiments, Calibration of camera was executed with calibration chart whcih was printed by using laser printer with 300 d.p.i. resolution. High distortion lens, COSMICAR 4.2 mm, was used to see whether the neural network could effectively calibrate camera distortion. 2-D range of several objects will be measured with laser range finding system composed of camera, frame grabber and laser structured light. The performance of 3-D range finding system was evaluated through experiments and analysis of the results.

Keywords: Camera calibration(카메라 교정), neural network(신경회로망), 2-D range fnding (2차원 거리 측정)

1. Introduction

Laser range finder has been widely used for 2-D range finding for enviroment recogonition which has been one of very important research topics in machine vision and robotics area. Numerous researches and developments on 2-D range finding have been made, but many problems still remain[1].

There are two methods which are most frequently used for 3-D range finding. One is based on triangulation method. Stereo vision and structured light method can be categorized as this. The other is based on flight time measurement. Waves such as modulated laser or sonic wave returning by reflection from target object are detected. The distance from wave source to the object are calculated based on the measured flight time of the wave. The former is more widely used because the range finding system is easily constructed. The latter is far more expensive and bulky, but very accurate distance measurement is possible in wide range of depth. Our study is related to triangulation type laser range finder using wide view angle lense.

Wide angle lens is useful for mobile robot environment recognition[2] and robot hand eye system. The advantage of using wide view angle lens camera is that a single frame of image may contain visual information in wide area. For example, autonomous mobile robot requires range data of advancing direction in almost 180°. Several frames are needed to calculate range using a lens with 20° view angle. Only one frame of image is sufficient for acquiring range information if we use a lens having view angle of 180°. Another possible application may be found in visual dimension inspection equipment that have only small space for the installation of range finder. Generally, wide view angle have large radial distortion that the distortion should be considered in camera calibration to get accurate range finding[3-9].

Camera is the most important input device in 2-D range finding. Cameras have various distortions caused by defects in lens manufacturing, optical characteristics, and misalignment, etc. These distortions should be calibrated to get accurate range finding[3,4,6-9]. Especially when a lens with short focal length to get wide view angle is used, calibration of lens distortion is indispensable.

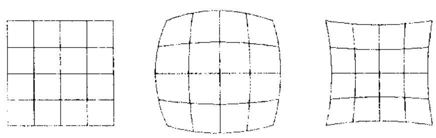
Two kinds of distortion, radial and tangential, are major concern for nonlinear camera calibration, but it should be noted that making accurate mathematical model for camera is almost impossible since there are so many kinds of distortions[8]. This is why neural network is appled for camera calibration.

Camera calibration problem can be considered as a mapping between a straight line in world coordinate system and a point in image plane[4]. Camera calibration for 2-D range finding is completed if one can get this mapping accurately. Two plane method calibrate camera by acquiring only the mapping without complex calculations to get elements in translation and rotation matrix and distortion coefficients[4].

In case where range finding is made by single camera and structured light, we should know the mapping between camera image plane and the plane made by structured light. Thus the problem of calibrating laser range finder could be treated as a plane to plane mapping. To acquire this mapping we selected multi-layer neural network. The benefits and problems of using neural network for the calibration of camera are discussed through computer simulation and 3-D range finding experiments.

2. Computer simulations

Among various distortions, radial distortion is the most dominant term. There are two types of radial distortion, positive and negative. If we look at rectangular grids as shown in Fig. 1 (a) using video camera and lens with negative or positive radial distortion, we could see images such as (b) or (c) in Fig.1 on TV monitor. In this chapter our main objective is to examine whether neural network can generate undistorted image as shown in Fig. 1 (a) from distorted image such as (b) or (c) in Fig. 1.



(a) no distortion (b) negative (c) positive

Fig. 1 *Effect of radial lens distortion*

Radial lens distortion, modeled with a 2nd order polynomial can be written as[6]

$$X_d = X_u(1+kR^2)^{-1} \quad (1)$$

$$Y_d = Y_u(1+kR^2)^{-1} \quad (2)$$

with

$$R^2 = X_u^2 + Y_u^2$$

where (X_d, Y_d) is distorted image coordinate, k is radial lens distortion coefficient, and (X_u, Y_u) is undistorted image coordinate. If k is positive, we call it pincushion or positive distortion. In case where k is negative, we say barrel or negative distortion exists.

We considered neural network as a two input two output mapper. The input of neural network is coordinate in image plane which could be acquired by using image processing. The output of neural network is two dimensional position in world coordinate system corresponding to a point in image plane. The dimension of the coordinate of a point in world coordinate system and a point in image plane is 3 and 2 respectively. To ensure plane to plane mapping, the dimension of world coordinate system should be reduced to 2. In laser range finding system, this concept is acceptable since stripes made by projected laser light is our only concern. To test neural network wheter it could learn plane to plane mapping, simple radial lens distortion model expressed by eqn (1) and eqn (2) was used.

We used multi-layer neural network having 2 hidden layers, each layer with 10 nodes, as shown in Fig. 1. Numerical teaching data defining the mapping between the distorted image coordinates and undistorted image coordinates are generated by eqn (1) and eqn (2). The input and output of neural network was distorted image coordinates and undistorted image coordinates. Radial distortion coefficient k was 0.01. After teaching neural network with the generated data, we made a file which includes coordinates of a distorted square as shown in Fig. 3. Then, we used this file as the input of trained neural network. We plotted the output of neural network in Fig. 4. From this figure, we can see positive radial distortion was corrected by trained neural network.

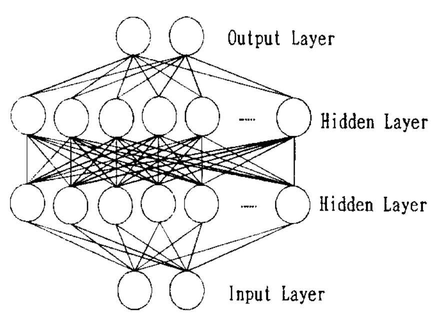


Fig. 2 *Neural network architecture for camera calibration*

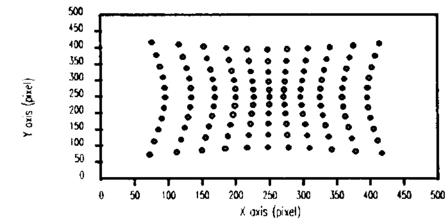


Fig. 3 *Simulation input data for neural network. Positive radial distortion included*

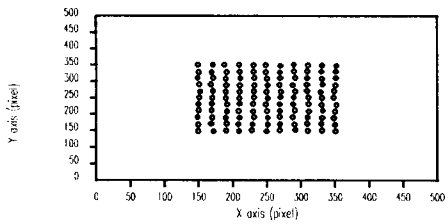


Fig. 4 *Output of neural network*

We define mapping error to be

$$E = \sqrt{(X_o - X_u)^2 + (Y_o - Y_u)^2} \qquad (3)$$

where E is mapping error, (X_o, Y_o) is the coordinate generated by trained neural network, and (X_u, Y_u) is the undistorted coordinate in eqn (1) and eqn (2). The maximum mapping error of about 5 pixels was observed by using 2 hidden layer network, each layer having 10 nodes. The number of teaching data was 600. We consider the accuracy obtained by this neural network based mapping simulation is somewhat unsatisfactory for applications requiring very high accuracy. But, for robotics or vision related applications which do not require very high accuracy measurement, 1% error would be acceptable. Anyway, we cannot say this method is resonable or not because this simulation considered only radial distortion. In chaper 3, we try to see the effectiveness of this method through real range finding experiments.

3. Range Finding experiments

We made 3-D range finding system using single camera and structured laser light as shown in Fig. 5. In front of laser, a cylindrical lens is attached to change collimated laser into structured light. Structured laser light is projected to target object. The calibration chart as shown in Fig. 7 is positioned in XY plane. Then, the image was acquired by using video camera and frame grabber. The center coordinates of fiducial marks were calculated and marked + as shown in Fig. 8. From this figure we can see the perspective relations as well as distortions are included.

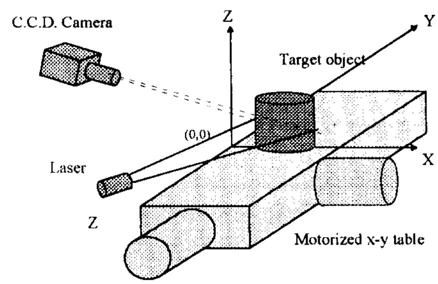


Fig. 5 *Range finding system using structured laser and video camera*

Special band pass filter is attached in front of the camera lens to cut unnecessary light. Thus, we can only observe bright stripes created by laser structured light. Fig. 6 shows image displayed on TV monitor when we measure range of triangular bar.

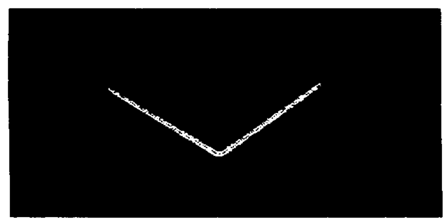


Fig. 6 *Image of Structured laser light projected to the edge of triangular bar*

We tried to train neural network to learn the mapping between the image and world coordinates of fiducial marks on calibration chart. Calibration chart is composed of 100 equidistant black fiducial marks as shown in Fig. 7. The size of the calibration chart was 124.5 mm x 124.5 mm. The center to center distance of each nearby points in horizontal and vertical direction was 13.833 mm. 300 d.p.i. laser printer was used to make this chart.

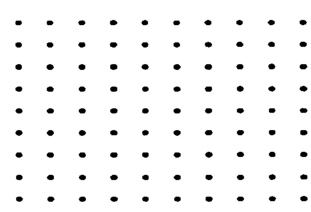


Fig. 7 *Calibration chart*

Two valued video image of the calibration chart was acquired and processed by using SONY CCD camera XC-75 with COSMICAR 4.2 mm lens and frame grabber DT 2862 of Data Translation Co. The image was reversed and labeled. Iterative algorithm of Haralick[12] was used for the labeling of each fiducial marks. The numbers shown in the Fig. 8 are labeling numbers for identifying which point in image plane is corresponds to which fiducial mark in world coordinate system.

Using the same teaching data, we trained two multi-layer neural networks. One has two hidden layers and each layer has 5 nodes. The other has two hidden layers, and each layer has 10 nodes. In each case, the number of training samples was 100.

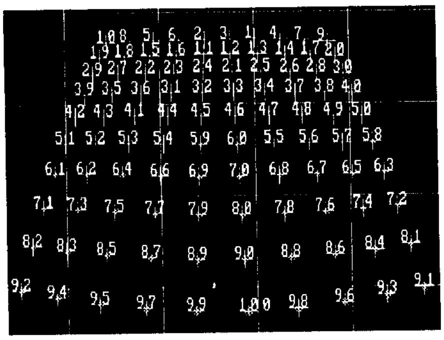


Fig. 8 *Labeled fiducial marks of calibration chart*

To test the trained neural networks, striaght laser line was generated by the projection of laser structured light to an object with flat surface. The displayed laser stripe has some thickness. Thus, for each column, center coordinates of the laser stripe are calculated. Then the input for neural network is prepared. The output of the neural network of two hidden layers and 5 nodes for each layer is shown in Fig 9. From this result we can conclude that the network approximated the range of target object, but error is somewhat large in some area. In case of two hidden layers and 10 nodes for each layer, the error was reasonably low for all area. In this case, the absolute error is less than 1mm in whole region of 120 mm x 120 mm measurement area. Several other range finding experiments were made. Fig. 9 and Fig. 10 show range data for circular cylindrical object. Fig. 11 and Fig 12 shows range data for triangular bar. For all cases, network with 2 hidden layers having 5 nodes for each layer showed better performance. With this network, range finding error of less than 1 mm was achieved for every tested target objects.

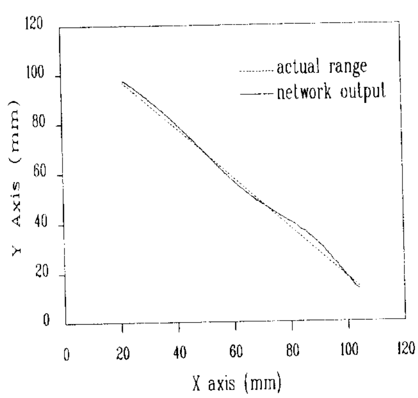


Fig. 9 *Range data for flat surface*
2 hidden layers, 5 nodes for each layer

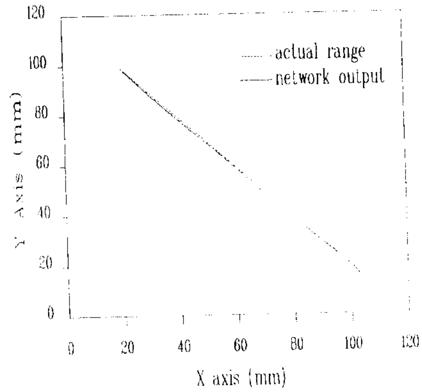


Fig. 10 *Range data for flat surface*
2 hidden layers, 10 nodes for each layer

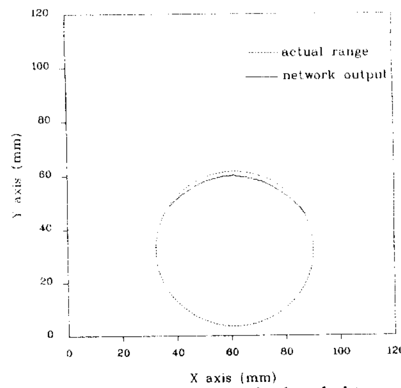


Fig. 11 *Range data for cylindrical object*
2 hidden layer, 5 nodes for each layer

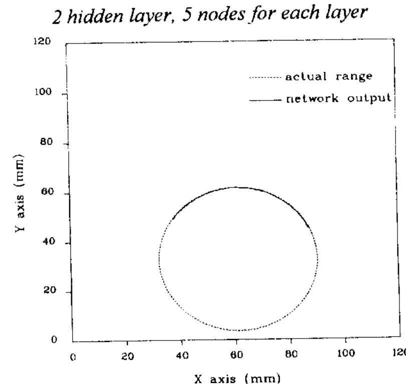


Fig. 12 *Range data for cylindrical object*
2 hidden layer, 10 nodes for each layer

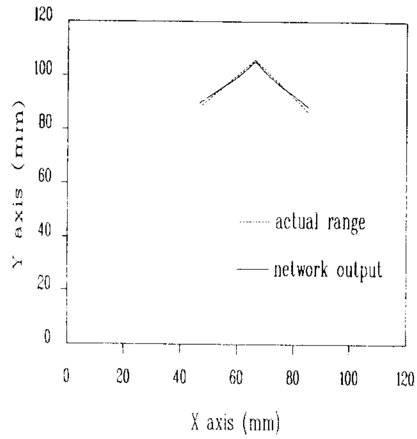


Fig. 13 *Range data for triangular bar 2 hidden layer, 10 nodes for each layer*

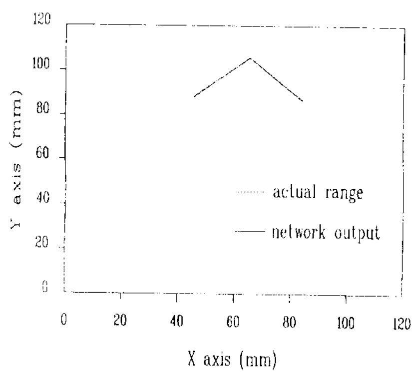


Fig. 14 *Range data for triangular bar 2 hidden layer, 10 nodes for each layer*

4. Results and conclusion

In this paper, we presented camera calibration method using neural network. With relatively small 100 training samples, lens distortions and perspective relations were trained at the same time. After training neural network, range finding experiments were done. Range finding error of less than 1 mm in 120 mm range was observed.

This calibration method was simple and reasonably accurate. With this method, direct range finding is possible without any numerical model. Application and development of more accurate network model remain for further study.

References

[1] Jarvis, R. A.: "A Perspective on Range Finding Techniques for Computer Vision", IEEE Trans. Pattern. Anal. and Mach. Intell., vol. PAMI-5, no. 2, 1983, pp. 122-139

[2] S. Yuta, S, Suziki, Y, Saito, and S, Iida, "Implementation of an Active Optical Range Sensor Using Laser Slit for In-Door Intelligent Mobile Robot," *in Proc. IROS*, pp. 415-420, 1991.

[3] Wei, Guo-Qing and Ma, Song De: "Implicit and Explicit Camera Calibration: Theory and Experiments", IEEE Trans. Pattern. Anal. and Mach. Intell., vol. 16, no. 5, May 1994, pp. 469-480

[4] Martins, H. A., Birk, J. R., and Kelly, R. B.: "Camera Models Based on Data from Two Calibration Planes", Computer Graphics and Image Processing, vol. 17, 1981 pp. 173-180

[5] Walter, J.A. and Schulten, K. J.: "Implementation of Self-Organizing Neural Networks for Visuo-Motor Control of Industrial Robot, IEEE Trans. Neural Networks, vol. 4, no. 1, Jan. 1993, pp. 86-95

[6] Lenz, R. K. and Tsai, R. Y.: "Techniques for Calibration of the Scale Factor and Image Center for High Accuracy 3-D Machine Vision Metrology", IEEE Trans. Pattern Anal. Machine Intell., vol. 10, no. 5, 1988, pp. 713-720

[7] Choi, D. H., Oh, S. Y., Chang, H. D., and Kim, K. I.: "Nonlinear Camera Calibration Using Neural Networks", Neural, Parallel & Scientific Computations, vol. 2, 1994, pp. 29-42

[8] Weng, J., Cohen, P, and Herniou, M.: "Camera Calibration with Distortion Models and Accuracy Evaluation", Trans. Pattern Anal. Machine Intell., vol. 14, no. 10, 1992, pp. 965-980

[9] Zhuang, H. and Shiu, Y. C.: "A Noise-Tolerant Algorithm for Robotic Hand-Eye Calibration with or without Sensor Orientation Measurement", IEEE Trans. Systems, Man, and Cybernetics, vol. 23, no. 4, July/Aug. 1993, pp. 1168-1175

[10] Komatsu, T., Aizawa, K., Igarashi , T., and Saito, T.: "Signal-Processing Based Method for Acquiring Very High Resolution Images with Multiple Cameras and its Theoretical Analysis", IEE Proceedings-I, vol. 140, no. 1, Feb. 1993, pp. 19-25

[11] Shiu, Y. C. and Ahmad, S.: "Calibration of Wrist-Mounted Robotic Sensors by Solving Homogeneous Transform Equations of the Form AX=XB", IEEE Trans. Robotics and Automat., vol. 5, no. 1, Feb. 1989. pp. 16-29

[12] Haralick, R. M. and Shapiro, L. G.: Computer and Robot Vision I, Addison-Wesley Publishing Co., 1992, pp.28-40