2010 IEEE/ASME International Conference on
Advanced Intelligent Mechatronics
Montréal, Canada, July 6-9, 2010

Fiducial Marker Indoor Localization with Artificial Neural Network

Gukhwan Kim and Emil. M. Petriu, *Fellow IEEE*
School of Information Technology and Engineering, University of Ottawa, Ottawa, ON, Canada

Abstract— **A vision based positioning system could be categorized into two groups. One analyzes an environment's scenery by matching the inputs with imaginary database to find the optimum result. The other uses fiduciary markers. In proposed method, the system uses fiduciary markers with a capital alphabet in it. When the known size fiduciary marker is captured by a camera, by using homography transformation, the 6-DOF camera pose with respect to the marker's local coordinate can be calculated. To recognize the character in the marker, Artificial Neural Network (ANN) with back-propagation training method is used. 12 unique features of a character are defined and used as inputs of ANN. Since more than 95% recognition rate is achieved in testing phase, the Optical Character Recognition (OCR) with ANN could be used as a marker detection method. The localization experimental result with the fiduciary marker shows that the proposed method could be a solution for indoor localization.**

I. Introduction

GPS based navigation systems have been widely used for not only military level positioning systems but also consumer grade positioning systems. However, due to its physical limitation, that is, the satellites signals are blocked inside buildings, GPSs may not be available in indoor environments [1]. Under the circumstances with the unavailability of GPS based navigation system, various types of indoor localization method using ultra-sonic, infrared, magnetic, and radio sensors have been proposed [2]-[4]. The indoor positioning system with these sensors requires an infrastructure to transmit the signal and the users are required to be equipped with receiver, amplifier, and power source. In addition, the cost to deploy the transmitters could become high. In order to reduce the cost, optical sensors, especially vision sensors, could be employed for the indoor positioning system.

The vision based positioning method could be categorized into two groups. One analyzes environmental components and matches the result with pre-stored data [5]-[7]. The other uses pre-installed markers placed in the environments [8]-[10].

In this paper, the focus is on the exploring a low-cost and efficient indoor localization method using a consumer grade web camera and fiduciary markers. Many scholars have been used fiduciary markers for indoor localization system. They are very easy to produce, manipulate, and maintain.

By using mathematical model of the camera calibration and homography transformation with known size fiduciary marker, the 6-DOF pose of a camera can be estimated. The geometrical information of 4 points (edges) of the marker gives enough information to calculate the pose of the camera. However, since it is obvious that the coordinates of the edges have some noise due to the low resolution of the image, if there are more corner points available, the distance measurement could be more accurate. Then, the information of the marker will transfer the position information with respect to the world coordinate system of the marker. In this paper, the content is an alphabet letter and it is connected to the position look-up table stored as a XML file. To recognize the alphabet in the fiduciary marker, artificial neural network is used.

Artificial Neural Networks (ANNs) are well known, trainable dynamic systems that can estimate input-output functions. The ANN is motivated by their ability to approximate an unknown non-linear input-output mapping through supervised training [11][12]. In this paper, ANN is employed for the recognition of the alphabet letter. 12 unique features of the alphabets are extracted and used for the inputs of the ANN and integer numbers from 1 to 26 are the outputs of the ANN.

This paper is organized as follows. In Section II, previous indoor navigation methods using optical markers are introduced. In Section III, mathematical model of camera calibration and localization is presented. ANN for alphabet character recognition is presented in Section IV. The experimental results are shown in Section V and the conclusion is given in Section VI.

II. Related Work

Various approaches for the reliable indoor localization system have been researched intensively for few decades, especially after GPS solved the outdoor localization problem.

In [13], the authors discuss the absolute position measurement method with pseudo- random encoding for the automated guided vehicles. This method marks the full length of the path on the ground or wall and reads the particular mark corresponding to the current position of the vehicle while the vehicle travels along the pseudorandom encoded guide-path. By using pseudorandom encoding and decoding method, the total number of bits to quantize the whole length of the path could be effectively reduced [13]-[15].

Gukhwan Kim is with the University of Ottawa, Ottawa, ON, K1N 6N5, CANADA (e-mail: gkim104@uottawa.ca).

Emil. M. Petriu is with the School of Intelligence and Technology Engineering, University of Ottawa, Ottawa, ON, K1N 6N5, CANADA (Tel:1-613-562-5800 ext. 6208).

Figure 1. Fiduciary Markers with alphabet 'A' and 'K'.

However, since the vehicle should follow the guide-path, the movement range of the vehicle is restricted. Thus a 2-D pseudorandom encoding method is introduced later in [16]. In this method, a sliding window is used to decode the pseudorandom encoded 2-D position data.

A large class of the work on indoor localization system using vision sensors has been done. Scale Invariant Feature Transform (SIFT) method uses the scale-invariant feature points as visual landmarks [17],[18]. Since these points are independent of camera view point, these are proper to be used as landmarks [19]. But this technique causes severe burden on computing time.

In order to reduce the computational cost, structured visual markers could be used in such a way as [20].The markers contain 2-D barcode that provides a unique ID and the authors demonstrate how to detect and decode such square fiduciary markers in real-time. However, this 2-D barcode shaped markers might reduce the aesthetic appeal of the room. In [21], patterns taking an aesthetic value into consideration are introduced. In the way of encoding position data into floor patterns, the proposed method is similar with [16] but the markers are designed with interior mind. The drawback of this method is that the recognition rate of the marker is decreased largely as the distance from the marker gets longer than 1,000mm.

III. Method of Pose Estimation

A. Fiduciary Marker

The fiduciary marker is consisted of a marker template, the black quadrilateral border used as a marker recognition beacon, and an alphabet letter as shown in Fig. 1. The size of the marker can be differed by a user, and here, the outer rectangle is 230mm×180mm (width × height) and the inner rectangle is 180mm×130mm. There are 8 feature points which are the edges of two rectangles and those points in the image are used for calculation of 6-DOF camera pose.

B. Image processing Algorithm

The marker detection and letter recognition processing flow is illustrated in Fig. 2. First, an image is captured by the vision sensor and the captured image is binarized with adaptive thresholding method. Then, the binarized image is color-inverted and if a white blob is bigger than certain pixels and has proper ratio *r (height / width),* then the area is labeled. As it is shown in Fig. 1, since the makers are composed of three blobs, the outer black boarder, the inner white rectangle,

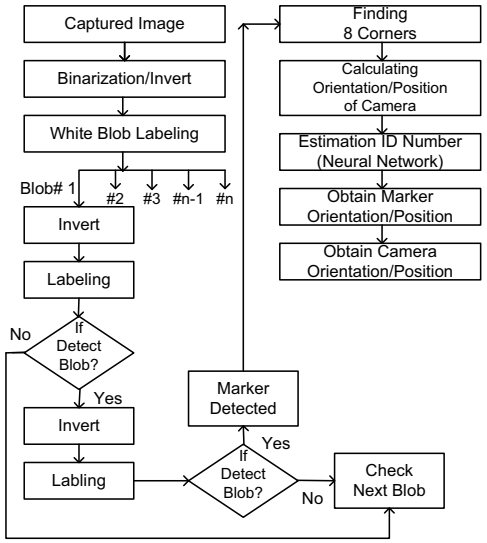


Figure 2. Fiduciary marker detection and alphabet letter recognition processing flow for robot pose estimation

and the alphabet letter, by doing blob labeling three times hierarchically, the marker could be reliably detected even in the messy indoor environment. If the marker is detected, the 8 corners of the marker are found. By using the planar homography transform technique with the 8 correspondences, the rotation vector and the translation vector are calculated and the 6-DOF pose of camera with respect to the marker coordinate is obtained. If the world coordinate of the each marker has been known, the world coordinate of the camera can be deduced.

C. Camera Pose Estimation

Since most of the consumer grade camera lens is distorted somehow, in order for the better accuracy of the pose estimation, lens correction should be done prior to any further processing. In addition, the intrinsic parameters such as lens focal lengths and projection centers offsets should be known as well. The calibration process is carried out offline and after camera calibration, the lens distortion coefficient matrix and the intrinsic matrix are obtained.

As shown in Fig. 2, the camera pose estimation is launched when the black quadrilateral border is detected. The next step of the camera pose estimation is finding 8 corners of the marker. A first corner point $c(x_0, y_0)$ is selected as a furthest point from a randomly selected point (x, y) among the contour of the outer rectangle. Then the furthest point from the first corner point $c(x_0, y_0)$ is selected as a second corner point $c(x_1, y_1)$. The third corner point $c(x_2, y_2)$ is obtained in the same way. However the fourth corner point $c(x_3, y_3)$ may not be obtained in the same way as second and third corner points. Fig. 3 shows the corner detection failure example. Since the image of the rectangle is not always shown as a rectangle, the last corner may not be chosen correctly.

Equation (1) is to solve the area of a quadrilateral by using 4 corner points of it. Among all the points of the contour of the quadrilateral, a point where the area becomes a maximum with previously obtained three corners would be the fourth corner.

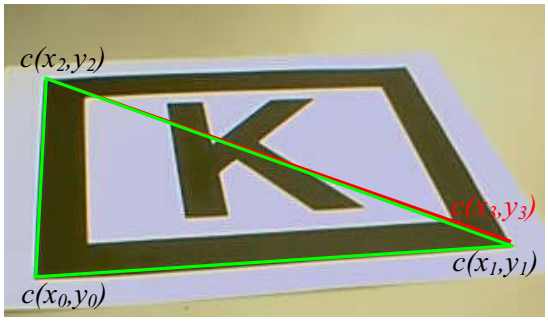


Figure 3. Corner detection failure problem encountered in affine transformed rectangle image.

$$A = \frac{1}{2}\sum_{i=0}^{3}(x_i y_i - x_{i+1} y_i) \quad (1)$$

The same method is applied to find four corners of the inner rectangle. After finding eight corners, the homography transformation is performed to find the six extrinsic parameters. Fig. 4 illustrates the geometric relationship between the image coordinates and the object coordinates in the perspective transformation. The four corners of the marker are expressed as below;

$$m_0 = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix}, m_1 = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \\ 1 \end{bmatrix}, m_2 = \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \\ 1 \end{bmatrix}, m_3 = \begin{bmatrix} X_3 \\ Y_3 \\ Z_3 \\ 1 \end{bmatrix} \quad (2)$$

$$c_0 = \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix}, c_1 = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}, c_2 = \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}, c_3 = \begin{bmatrix} x_3 \\ y_3 \\ 1 \end{bmatrix} \quad (3)$$

Those eight points could be described in the homogeneous coordinates by using homography transformation as below;

$$c = s \cdot H \cdot m = s \cdot M \cdot W \cdot m \quad (4)$$

where s is the scale factor.

An important fact is that the transformation matrix H could be divided into two parts, one is the intrinsic parameters M and the other is the physical transformation parameters. The physical transformation W is expressed as a composition of translation vector T and rotation matrix R as follows;

$$M = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}, W = [R \;\; T], \text{ where } R = [r_1 \;\; r_2 \;\; r_3] \quad (5)$$

For the convenience of the calculation, it is assumed that the object plane has $Z = 0$. Therefore, the homography transformation matrix would become simpler.

$$\begin{bmatrix} x_0 & x_1 & x_2 & x_3 \\ y_0 & y_1 & y_2 & y_3 \\ 1 & 1 & 1 & 1 \end{bmatrix} = s \cdot M \cdot [r_1 \;\; r_2 \;\; T] \begin{bmatrix} X_0 & X_1 & X_2 & X_3 \\ Y_0 & Y_1 & Y_2 & Y_3 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (6)$$

In order to calculate the homography matrix H, at least four correspondence points are required and more points would reduce errors in such a way of the Least-Squares Error. To know the world coordinate position of a camera, simple inverse-kinematics knowledge is required.

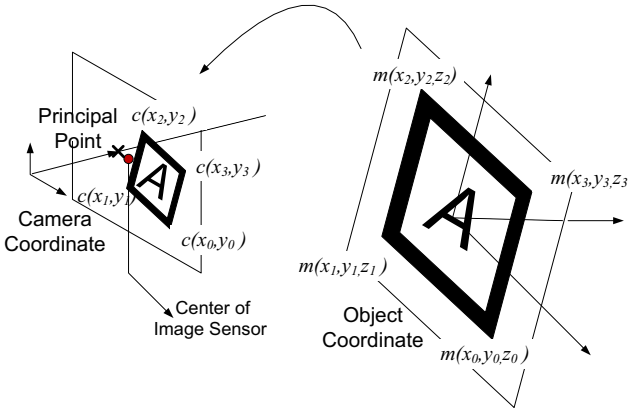


Figure 4. Perspective transformation of the correspondence points

IV. Character Recognition

A. Alphabet Feature Extraction

The computational cost of the ANN depends on the number of the nodes, i.e. the number of the connection links, in the ANN. The output nodes are specified as the total number of the English alphabets and the number of hidden layer nodes could be tuned in an empirical way. However, the number of the input nodes is required to be specified while considering the important features of the alphabet which could affect the recognition rate. In this paper, 12 features of the alphabet images among 16 features in [22] are chosen to be used for the ANN inputs. Those 12 features are enough to achieve a satisfactory result. Following descriptions are the 12 features.

1. The total number of "On" pixels in the alphabet image.
2. The mean horizontal position of all "On" pixels.
3. The mean vertical position of all "On" pixels.
4. The mean squared value of the horizontal pixel distances as measured in 2 above.
5. The mean squared value of the vertical pixel distances as measured in 3 above.
6. The mean product of the horizontal and vertical distances for each "On" pixel as measured in 2 and 3 above.
7. The mean value of the squared horizontal distance times the vertical distance for each "On" pixel.
8. The mean value of the squared vertical distance times the horizontal distance for each "On" pixel.
9. The mean number of edges encountered when making systematic scans from left to right at all vertical positions within the box.
10. The sum of the vertical positions of edges encountered as measured in 9 above.
11. The mean number of edges encountered when making systematic scans of the image from bottom to top over all horizontal positions within the box.
12. The sum of horizontal positions of edges encountered as measured in 11 above

B. Letter recognition with Neural Network

As shown in Fig. 5, the inner white area of the marker with a black alphabet letter could be found with marker detection

algorithm. Then, the smallest rectangular image which includes all 'On' pixels of the letter is extracted and resized for the letter recognition process. Subsequently, the 12 features of the alphabet image are extracted though feature extraction process. These feature values are fed into the ANN and consequently the desired output is estimated.

V. Experimental Results

In this section, the experimental results of the proposed method, which has been implemented upon OpenCV 2.0, are presented.

A. Experimental Setting

The localization experiment was done on a laptop computer (Intel Core 2 Duo 997MHz processor, 2GB of RAM) with Windows XP as the operating system and a Samsung PLEOMAX PWC-3800 USB camera connected to the laptop. The camera captures 320 × 240 pixels image sequences as the input to the pose estimation system.

B. Marker detection

Fig. 5 shows the marker detection and character extraction procedure. As shown in Fig. 5(a), using the marker detection algorithm described in Section III. B, a marker is detected. The criteria of the marker detection are listed below.

IF A blob consists of a black boarder, a white inner quadrilateral, and a black blob (alphabet) inside of it,
and *IF* $(130+26)/100 > b/2d > (130-26)/100$,
and *IF* $(180+36)/130 > (c-2d)/b > (180-36)/130$,
and *IF* $(230+46)/180 > c/a > (230-46)/180$,
Where a, b, c, and d is shown in Fig. 5(b),
THEN The blob is a marker.

Based on these criteria, the maker detection algorithm can be reliable beyond the messy environment. This algorithm with the above stated experimental setting and criteria can provide a maximum of 3500 mm detection range. From the result, it is suggested that for the continuous, stable, and constant pose estimation, the markers are required to be deployed at the interval of no more than 3000 mm.

C. Alphabet recognition rate and processing time

The ANN consists of an input layer, two hidden layer, and an output layer. 12 input nodes and 26 output nodes construct the input layer and the output layer. Each hidden layer consists of 100 nodes. In order to train the ANN, the backpropagation [11] method is used. 1,000 training images of each alphabet character, totaling 26,000 images, are generated. The 25% of the training images are randomly tilted images between 20 degree and -20 degree, another 25% of the training images are edge blended and tilted images, another 25% of the images are partially deleted and tilted images, and the remained 25% of the images are tilted, edge blended, and partially deleted images. Fig. 6 shows the some samples of the training images. As a result of the training of the ANN with these 26,000 images, the connection link weights are solved and stored in a XML file.

In Fig. 5(d), the normalized alphabet image produced by the image pre-processing step is presented. After the marker

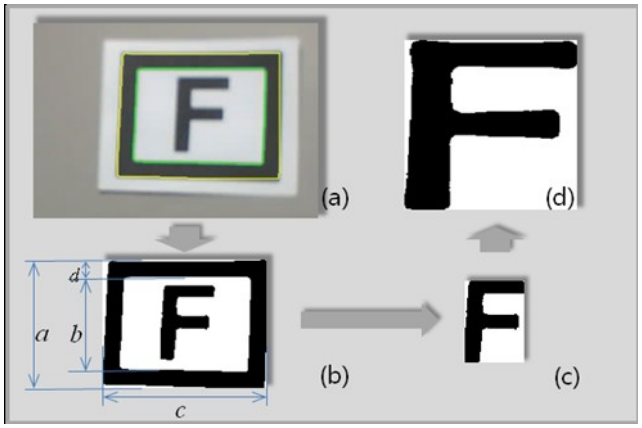


Figure 5. Marker detection and character extraction procedure
(a) Detected marker image, (b) Binarized Marker image,
(c) Extracted character image, (d) Normalized character image

detection phase, the inner quadrilateral is unwarped through the affine transformation. Then, the alphabet blob is trimmed into smallest box where the blob can fit into. The alphabet blob is then normalized as a square image. Fig. 7 illustrates the four different alphabet images normalized into 200 × 200 pixels. After the normalization, the pre-processing step would be finalized with feature extraction phase. The four normalized images in Fig. 7 have the 12 numeric features presented in Table I and these feature values go through the ANN.

The performance of the ANN alphabet recognition algorithm is evaluated in the perspective of the recognition rate and the processing time with four different size images (200 × 200, 100 × 100, 50 × 50, 25 × 25) of all alphabets.

Consequently, the average recognition rates are 99.63%, 99.10%, 98.02%, and 96.73% for each case of 200 × 200, 100 × 100, 50 × 50, and 25 × 25 pixel size images respectively. As shown in Table II, except the case of B, D, H, and N, more than 95% recognition rates are achieved in overall cases. Besides, the recognition rates are more than 95% upon the all the alphabet images when it comes to the equal or larger than 50 × 50 pixels. The reason why the recognition rates are decreased in the case of B, D, H, and N is that the binarized and normalized images of the captured scene from the camera

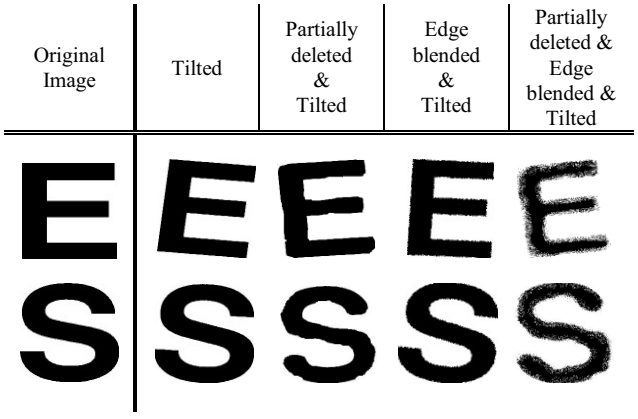


Figure 6. Some sample alphabet images for the ANN training

Figure 7. Four examples of the normalized alphabet image

are degraded thus the features of the normalized images are become similar to other alphabet image's features.

Fig. 8 shows the overall processing time of the four different size images of all alphabets. The processing rates are 15.8Hz, 17,2Hz, 18.2Hz, and 21.2Hz in the case of 200 × 200, 100 × 100, 50 × 50, and 25 × 25 pixel size images respectively. Since the size of an image affects the feature extraction time, the smaller image gives shorter processing time. However, for the reliable system the size of the normalized image is suggested to be larger than 25 × 25.

D. Position estimation errors

When the alphabet character in the marker is recognized by the ANN, the world coordinate 3-D position of the center of the marker and the rotation matrix, i.e. transformation matrix, are retrieved from the marker database stored as a XML format file. Subsequently, by doing a transformation from the marker coordinate to the world coordinate, the world coordinate camera could be calculated.

In this Section, the focus is on checking the performance of the pose estimation algorithm. The performance of this method is evaluated with the errors of the 3D translation vector elements from the origin of the marker to the origin of the camera. The marker is settled down on fixed position and the camera is set by the distance of 50 cm, 100 cm, 150 cm, 200 cm, and 250 cm from the marker. In addition, the camera and the marker are placed while the z axes of the marker and the camera are coincided for the experimental convenience. Fig. 9 depicts the 3D position estimation errors.

As shown in Fig. 9(a), the errors of all axes become larger as the distance between the camera and the marker is increased. Since the captured image from the camera gets more blurred as the distance is increased, the positions of 8 corner points are become inaccurate and consequently the transformation matrix looses the accuracy. Based on the experimental results, therefore, the marker is required to be deployed at the intervals of no more than 2 m for the better accuracy.

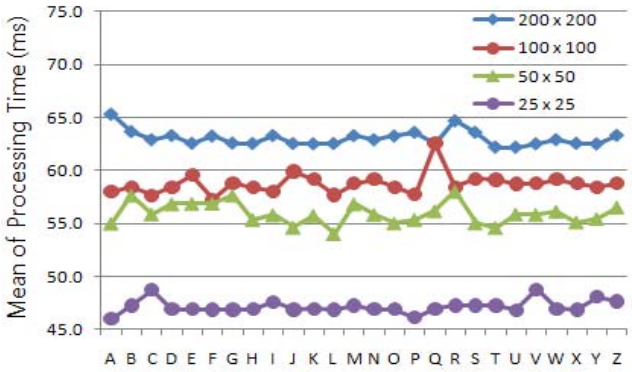


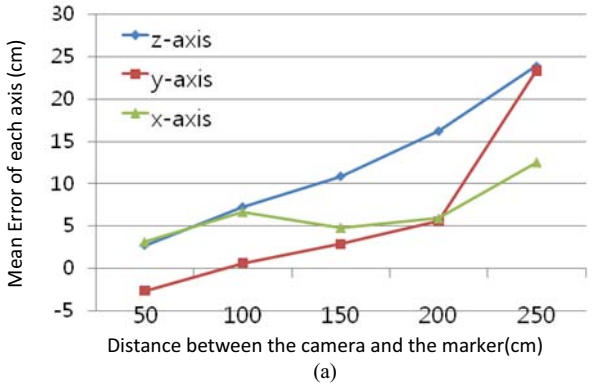
Figure 8. Alphabet recognition processing time of four different size images

TABLE I
EXAMPLES OF THE 12 NUMERICAL ALPHABET FEATURE DESCRIPTIONS

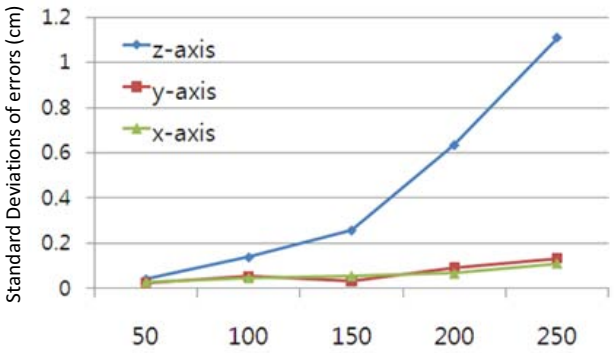
#	B	D	M	W
1	26748	23516	26852	17555
2	91	90	100	100
3	99	99	99	95
4	97011836	95208338	116203451	47151561
5	97928086	91343550	87092101	52462265
6	2094514	-136616	279419	472519
7	250168384	13054150	-414142503	-702560463
8	-468907920	-874019928	-82471225	52019225
9	989	996	1442	2351
10	1111	875	834	2068
11	149475	138512	146605	239549
12	109672	82567	81324	182596

VI. CONCLUSION

The proposed fiducial marker indoor localization system developed based on the OpenCV 2.0 and used the consumer grade web camera was described. There are numerous types of technologies for the indoor localization system but our main concerns are reducing the cost of the installation and maintenance of the system as well as the usability. In this perspective, the proposed work has its own value since it is easy to implement. We only need to deploy the markers on the hall way or the room and register the markers into the marker



(a)



(b)

Figure 9. Position estimation errors along the distances between the camera and the marker. (a)Mean of errors for 60sec, (b)STDs of errors

database. For the large size buildings, by using the combination of the alphabets inside the marker, more possible markers could be generated. As the Table II shows, the printed alphabet recognition with ANN achieved high recognition rate. This fact gives us a promising idea which is that the marker could have a form of environment-friendly shape with human recognizable signs.

As shown in Fig. 9, the accuracy of the localization is reduced as the distance increased. The 320 × 240 pixels size video images are used for the overall system and the images are blurred due to the low performance of the USB web camera. Since the blurred images cause the degradation of the binarized images, the higher definition camera could provide better accuracy of the pose estimation. Thus, the proposed system with various types of camera will be scrutinized to verify the usability of the system later on. Moreover since finding corner points' position is the most important step in the proposed method, a way to provide more accurate position with different markers will be examined.

TABLE II
ALPHABET RECOGNITION RATE WITH VARIOUS IMAGE SIZES

Alphabet	Alphabet Image size(Pixels)			
	200 × 200	100 × 100	50 × 50	25 × 25
A	100	100	98.2	97.5
B	98.6	97.4	96.3	93.2
C	99.5	99.1	98.2	96.8
D	98.4	97.5	96.1	93.6
E	100	100	98.9	97.1
F	99.8	99.2	98.0	96.9
G	99.4	98.9	98.9	96.8
H	98.3	97.1	95.9	92.7
I	100	99.9	99.5	99.1
J	100	99.7	98.2	97.9
K	100	99.9	98.3	98.0
L	100	99.5	98.7	98.1
M	99.9	99.4	97.8	96.8
N	98.0	96.5	94.9	92.1
O	99.8	99.3	98.0	97.6
P	99.7	99.0	98.6	97.2
Q	99.7	99.1	98.6	97.6
R	100	99.8	99.1	98.2
S	100	100	98.9	98.6
T	100	100	99.1	98.7
U	99.9	98.3	97.9	96.2
V	99.9	98.7	97.3	96.4
W	99.6	98.9	97.8	96.0
X	100	99.9	98.8	97.9
Y	100	99.7	98.4	97.1
Z	100	99.9	98.3	97.0
Average	99.63	99.10	98.03	96.73

REFERENCES

[1] M.S, Grewal, L.R. Weill, and A.P. Andrew, *Global positioning Systems, Inertial navigation and integration*, New Jersey: John Wiley & Sons,, 2nd ed., 2007.

[2] O. Wijk, H.I. Christensen, "Localization and navigation of a mobile robot using natural point landmarks extracted from sonar data," *Robotics and Autonomous Systems*, vol.31, issues 1-2, pp.31-42, 2000.

[3] F. Raab, E. Blood, T. Steiner, and R. Jones, "Magnetic position and orientation tracking system," *IEEE Trans. Aerospace and Electronics Systems*, vol.AES-15, No.5, pp.709-718, 1979.

[4] M. Kourogi, N. Sakata, T. Okuma, and T. Kurata. "Indoor/Outdoor Pedestrian Navigation with an Embedded GPS/RFID/Self-contained Sensor System". in *Proc.16th ICAT2006*, pp.1310-1321, 2006.

[5] D. Sticker and T. Kettenbach, "Real-time and Markerless Vision-Based Tracking for Outdoor Augmented Reality Applications," in *Proc. ISAR 2001*, pp. 189-190, 2001.

[6] L. Vacchetti, V. Lepetit, and P. Fua, "Combining edge and texture information for real-time accurate 3d camera tracking," in *Proc. ISMAR2004*, pp. 48-57, 2004.

[7] A. J. Davison, "Real-Time Simultaneous Localization and Mapping with a Single Camera," in *Proc. ICCV'03*, Vol. 2, pp. 1403-1410, 2003

[8] H. Kato and M. Billinghurst, "Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System," in *Proc. IWAR '99*, pp. 85-94, 1999.

[9] M. Kalkusch, T. Lidy, M. Knapp, G. Reitmayr, H. Kaufmann, and D. Schmalstieg, "Structured Visual Markers for Indoor Pathfinding," *In Proc. ART02*, 2002.

[10] L. Naimark and E. Foxlin, "Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker," in *Proc. ISMAR2002*, pp. 27-36, 2002.

[11] H.N. Robert, *Neural networks for perception: computation, learning, architectures*, Orlando: Harcourt Brace & Co., 1992, pp.65-93.

[12] I. A. Basheer, M. Hajmeer, "Artificial neural networks: fundamentals, computing, design, and application," *Journal of Microbiological Methods*, vol. 43, pp. 3-31. 2000.

[13] E.M. Petriu and J.S. Basran, "On the position measurement of automated guided vehicles using pseudorandom encoding," *IEEE Trans. Instrumentation and Measurement* , vol. 38, No. 3, pp. 799-803, 1989.

[14] E. M. Petriu, "Absolute-type position transducers using a pseudorandom encoding," *IEEE Trans. Instrumentation Measurements.,* vol. 1M-36, pp. 950-955, Dec. 1987

[15] E. M. Petriu, "Absolute position recovery for automated path-guided vehicles," presented at ROBOTS 12 and VISION '88 Conf. Exibit., Detroit, MI, June 5-9, 1988.

[16] E.M. Petriu, W.S. McMath, S.K. Yeung, N. Trif, and T. Bieseman, "Two-dimensional position recovery for a free-ranging automatedguided vehicle," *IEEE Trans. Instrumentation and Measurement*, vol 42, issue 3, pp. 701-706, 1993.

[17] D.G. Lowe, "Object Recognition from Local Sclae-Invariant Features," in *Proc. of the International Conference on Computer Vision*, vol.2, pp. 1150-1157, 1999

[18] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, issue 2, pp. 92-110, 2004.

[19] S.Y. Park, S.C. Jung, Y.S. Song, and H.J. Kim, "Mobile robot localization in indoor environment using Scale-Invariant visual landmarks," in *Proc. the 18th IAPR International Conference on Pattern Recognition (ICPR'06)*, pp. 159-163, 2008.

[20] A. Mullpni, D. Wagner, D. Schmalstieg, and I. Barakonyi, "Indoor positioning and navigation with camera phones," *IEEE Pervasive Computing*, vol 8. issue 2. pp. 22-31, 2009.

[21] S. Saito, A. Hiyama, T. Tanikawa, M. Hirose, "Indoor Marker-based Localization using Coded Seamless Pattern for Interior Decoration,"

[22] P.W. Frey and D. J. Slate, "Letter recognition using Holland-style adaptive classifiers," *Machine Learning,* vol. 6, no. 2, pp . 161-182, 1991.