International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol. 10, No. 7 (2017), pp. 73-84
http//dx.doi.org/10.14257/ijsip.2017.10.7.07

Robust Detection of ROI of Signboards in a Noisy Environment

Daeyeong Lim[1], Youngbaik Kim[2], Incheol Park[3], Jihoon seung[4] and Kilto Chong[5*]

[1,3,4,5] 1567 Baekje-daero, deokjin-gu, Jeonju-si, Jeollabuk-do 54896 ROK
[2]KEPCO Engineering and Construction Company, Inc., Republic of Korea
[1]tamudylim@gmail.com, [2]ybkkim@kepco-enc.com, [3]incp486@naver.com,
[4]jeehun@nate.com, [5]kitchong@jbnu.ac.kr*

Abstract

In this paper, we present the prior art to find the ROI of computer vision research in order to acquire the necessary information using images or video. We adaptively obtain the ROI from the problems of conventional method with limited size of the template method and the assumption that the signboard image is located at the center. A method for signboard detection is proposed by searching the Hough line and determining the area of interest and reconstructing it taking into consideration the warped image.

Keywords*: Sign-board image, Natural scenes, Hough transform, Geometric transformations, Image segmentation*

1. Introduction

For computer vision research to acquire the necessary information using images or video, prior art searching ROI is needed. In this paper, ROI detection method of building signboard is studied. The text in a natural image can provide the most clear and meaningful information available. Particularly, the signboard includes information such as business name and the telephone number. Since the characteristics of the building can be known through classification, it can be used as the main information for autonomous navigation of the unmanned robot.

The candidate region of the signboard first binarizes the image through preprocessing operation, and then uses point and line to find the feature. The signboard candidate region is detected using the above feature.

We focus on preprocessing work which has the greatest influence on the detection of the candidate area of the signboard and compare existing techniques with the proposed technique.

2. Signboard ROI Detection

2.1. Difficulty in Recognizing Signs in Natural Images

Similar studies on sign recognition include license plate recognition, traffic sign recognition, and container identification. The above given studies show the size and color as well as the approximate location of the attachment. On the other hand, signboards are composed of size, location, and various colors. Signboard image recognition present various obstacles due to their characteristics [11, 16].

Common difficulties in recognition are the effects of shadows and lighting as well as shape distortion due to the tilting of the camera during shooting. This difficulty can be solved through color analysis and structural analysis [2, 8, 17-18].

In particular, there are various types of obstacles in recognizing signboards. Typical examples of these obstacles are electric poles, wires, trees, cars, and doors. Since most natural images contain these obstacles, they must be discussed. Horizontal edge, vertical

ISSN: 2005-4254 IJSIP

edge, point and line information are used in order to recognize a signboard. Complex horizontal, vertical, and point information are contained in the case of trees and it is difficult to separate them into non-signboard areas. It is difficult to separate wires, cars, doors, *etc.* into non-signboard areas due to light reflection and color information.

Therefore, existing methods to solve the problems inherent in natural images is reviewed and their performance is compared with the proposed method.

2.2. Related Works

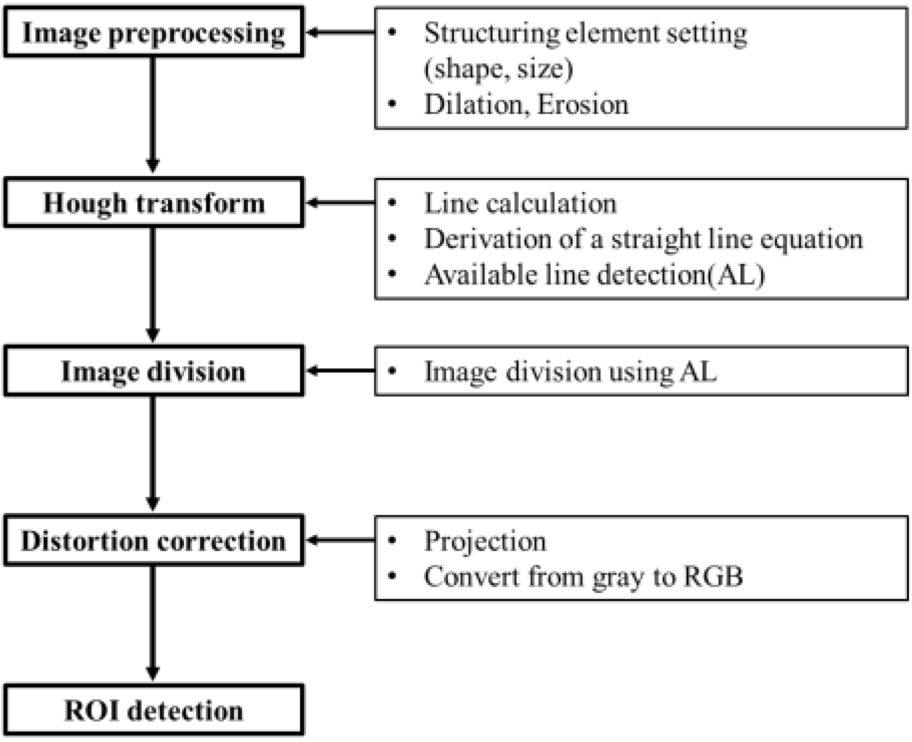


Figure 1. Block Diagram of the Proposed Method

The procedure for detecting the ROI of the proposed signboard is shown in Figure 1 above. There is an image preprocessing step for image binarization that will be introduced in Section 2.2 with the subsequent steps covered in Sections 2.3 through 2.5.

2.2.1. Horizontal Edge

The Sobel mask technique is used for the preprocessing method using the horizontal edge. Sobel is used to detect an edge as an operator for performing a derivative on an image to compute a gradient. Sobel mask is one of the first-order differential operators that extract edges in all directions. Since the projected pixel values are relatively averaged, they have a strong characteristic in noise [1, 3-4].

The horizontal and vertical operations of the Sobel mask are as follows.

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A \text{ and } G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \quad (1)$$

The above equations represent vertical and horizontal Sobel masks, respectively, and **A** represents the gray level input image. If there are no obstacles or interference in natural reminiscence, the signboard can be separated quickly and accurately using the horizontal edge of the Sobel mask. However, if there is an obstacle or interference, the Sobel mask includes edge information of the non-signboard area, which makes it difficult to separate the background from the signboard.

2.2.2. Color Distance

The color information of the image is important for sign recognition. Many studies use HSV color models. Signs can be defined into similar colors in natural images. The similarity of colors is measured by the color distance. The conversion from an RGB color image to an HSV color image is as follows [5, 14-15, 19].

$$R' = \frac{R}{255}, G' = \frac{G}{255}, B' = \frac{B}{255} \tag{2}$$

$$C_{max} = \max(R', G', B'), C_{min} = \min(R', G', B'), \Delta = C_{max} - C_{min} \tag{3}$$

Hue calculation:

$$\mathbf{H} = \begin{cases} 0' & , \Delta = 0 \\ 60' \times (\frac{G' - B'}{\Delta} mod6) & , C_{max} = R' \\ 60' \times (\frac{B' - R'}{\Delta} + 2) & , C_{max} = G' \\ 60' \times (\frac{R' - G'}{\Delta} + 4) & , C_{max} = B' \end{cases} \tag{4}$$

Saturation calculation:

$$\mathbf{S} = \begin{cases} 0 & , C_{max} = 0 \\ \frac{\Delta}{C_{max}} & , C_{max} \neq 0 \end{cases} \tag{5}$$

Value calculation: $\mathbf{V} = C_{max}$

Using the H, S, and V values, the color distance can be calculated as follows.

$$\text{CD} = \sqrt{Hue^2 + Saturation^2 + Value^2} \tag{6}$$

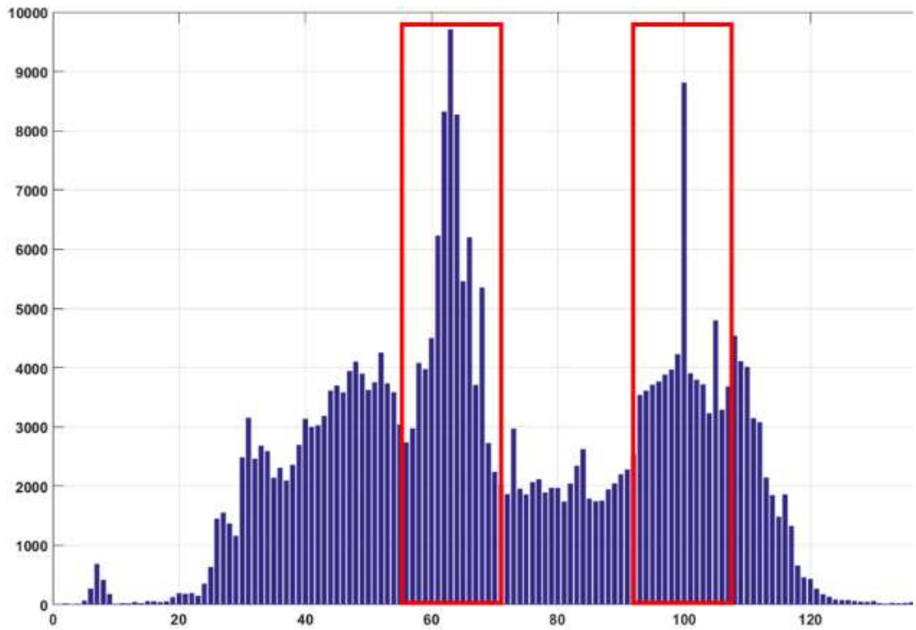


Figure 2. Histogram after CD Application

The histogram of the image is obtained by using the CD value obtained Figure 2. Based on the obtained histogram, it is possible to separate non-signboard areas by using a portion having a large amount of information. Such a method requires the signboard to be at least half the size of the image size, and only when the color distance between the background and the signboard is too long.

2.2.3. Fuzzy C-mean

The fuzzy c-mean algorithm (FCM) is used to cluster image color information using feature vectors. The FCM algorithm has the following objective function [6-7, 13].

$$J_{FCM} = (U,V) = \sum_{i=1}^{c}\sum_{k=1}^{N}u_{ik}^{p}\|x_k - v_i\|^2 \quad (7)$$

Where p is any real number that indicates fuzziness of resulting partition, u_{ik}^{p} is the membership of x_k in the cluster i, x_k is the pixel of data, v_i is the center of the cluster. In FCM, partitioning is done with an iterative optimization of the function and membership u_{ik}^{p} is updated iteratively and the cluster centers v_i are also updated along with it. $\|\cdot\|$ norm represent the distance between pixel and the centre. The performance index $J_{FCM}(U,V)$ measures the weighted sum of distance between cluster centre and elements in corresponding fuzzy clusters and it is minimized when the pixel is close to centroid of their clusters. That pixel is assigned with high membership values and the remaining pixels are with low membership value [25-26].

After performing the fuzzy c-mean, it can be seen that the histogram is classified into two cluster values as described above.

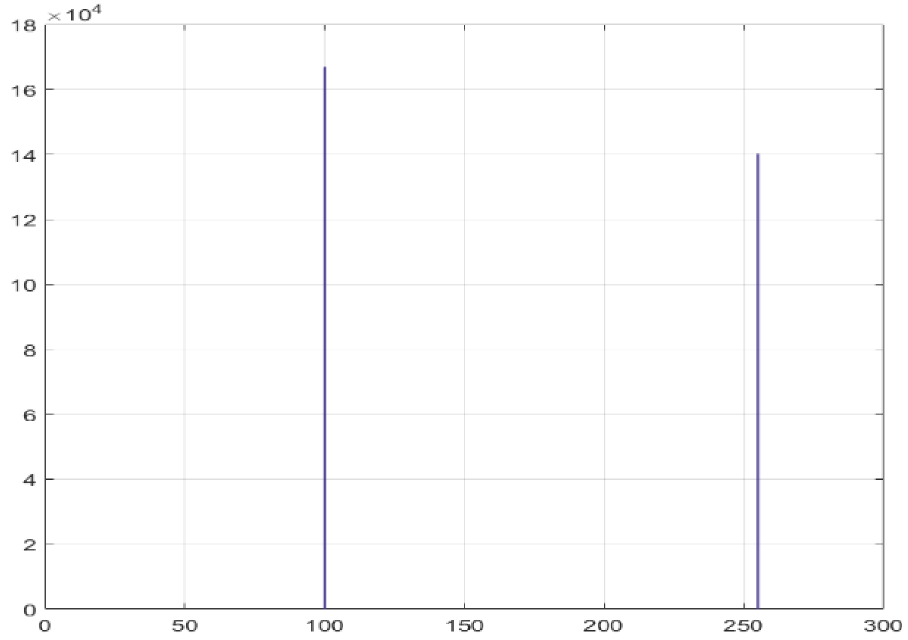


Figure 3. Histogram after Fuzzy C-Mean Application

2.3. Proposed Method

The proposed preprocessing technique is based on morphology. Morphology means analyzing the geometric shape of an image that involves erode and dilate operations.

The preprocessing process is as follows.

Erode → Splitting and Merging → Dilate → Splitting and Merging

Dilation of the gray scale image f by the structural element b is denoted by $f \oplus b$ and is defined as follows.

$$(f \oplus b)(x,y) = \max\{f(x - x', y - y') + b(x',y') | (x',y') \in D_b\} \quad (8)$$

The structural element is rotated 180 degrees about its origin. The structural element values rotated at each moved position are added to the image pixel values and the maximum value is then calculated.

Erosion of the gray scale image f by the structural element b is denoted by $f \ominus b$ and is defined as follows.

$$(f \ominus b)(x,y) = \min\{f(x + x', y + y') - b(x', y') | (x', y') \in D_b\} \quad (9)$$

Subtract the structural element value of the image pixel values at each movement position and then calculate the minimum value.

Splitting and Merging steps can be summarized as follows[22][23][24].

Step1. Split into four disjoint quadrants any region R_i for which $P(R_i) = FALSE$.

Step2. When no further splitting is possible, merge any adjacent regions R_i and R_j for which $P(R_i \cup R_j) = TRUE$.

Step3. Stop when no further merging is possible.

In step 2, the divided image can be generated by filling all the quadrant areas satisfying the attribute with 1 and filling the non-quadrant areas with 0.

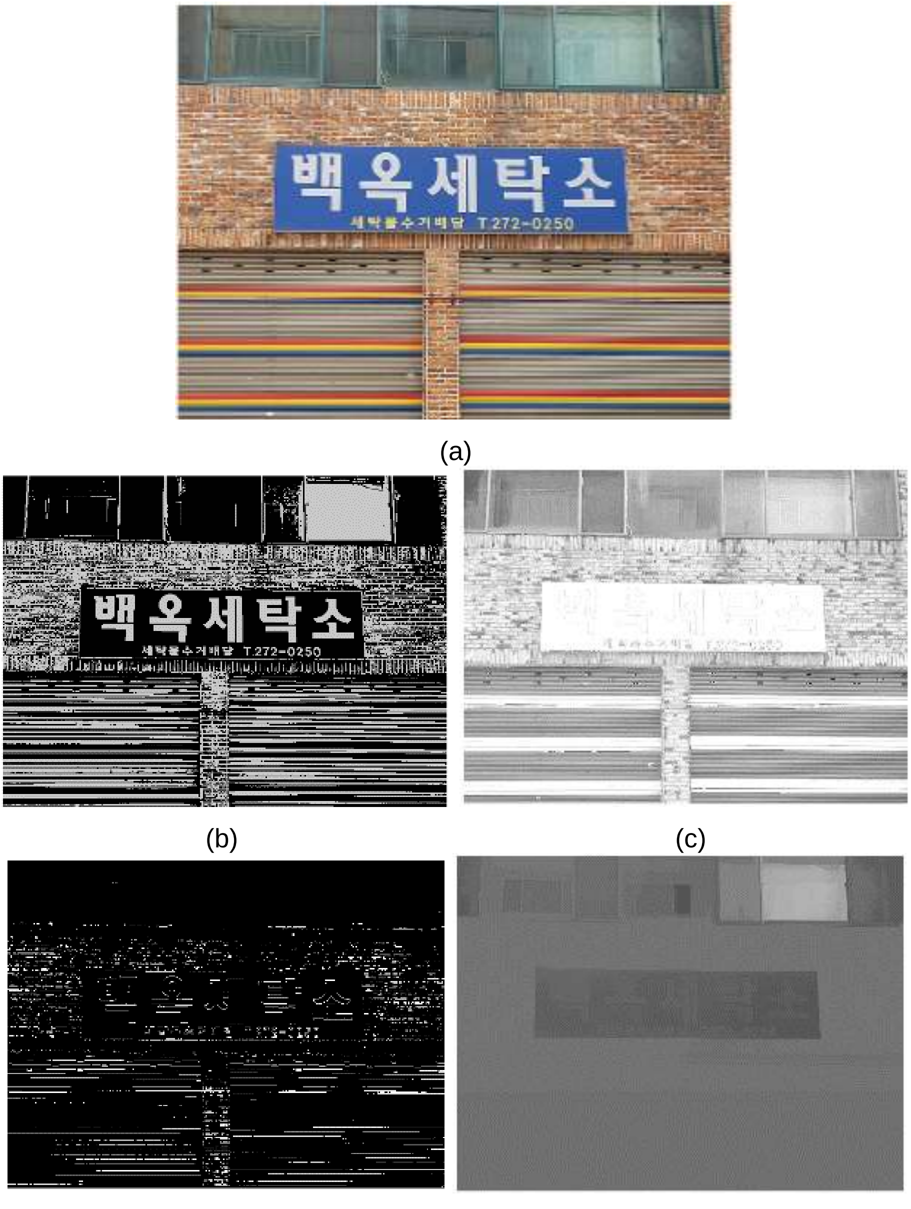


Figure 4. (a) Original Image, (b) Image Preprocessing Result using Horizontal Edge, (c) Color Distance, (d) Fuzzy C-mean, (e) Proposed Method

2.4. Hough Transform

In Hough transform, we consider point (x, y) and all the lines that pass through this point. There are infinitely many lines passing through the point (x, y) and all of the lines satisfy the slope-slicing linear formula $y = ax + b$ for certain values of a and b. If we write this equation as $b = -ax + y$ and consider the ab-plane, we have one linear equation for a fixed pair (x, y) [8, 12-13, 20].

We can draw the parameter space straight lines for all image points (x, y) in the xy-plane and identify the major straight lines in the xy-plane by identifying the intersection point of several straight lines in the parameter space. However, to avoid a (slope of a straight line) going infinitely when a straight line approaches in a vertical direction, use the following normal form of straight line.

$$x \cos\theta + y \sin\theta = \rho \tag{10}$$

For line detection through Hough transform, two pairs of points were calculated for ten line extractions, and line components less than about seventeen percent (17%) of the image width were excluded from line components. Adjacent lines and intersecting lines were removed for the detected ten line components. The two point pairs were converted into one linear equation for convenience.

The following figure show the results of applying four techniques to natural images that include many horizontal components of doors, windows, and bricks. It can be confirmed that the proposed method effectively detects the upper and lower lines of the signboard area.

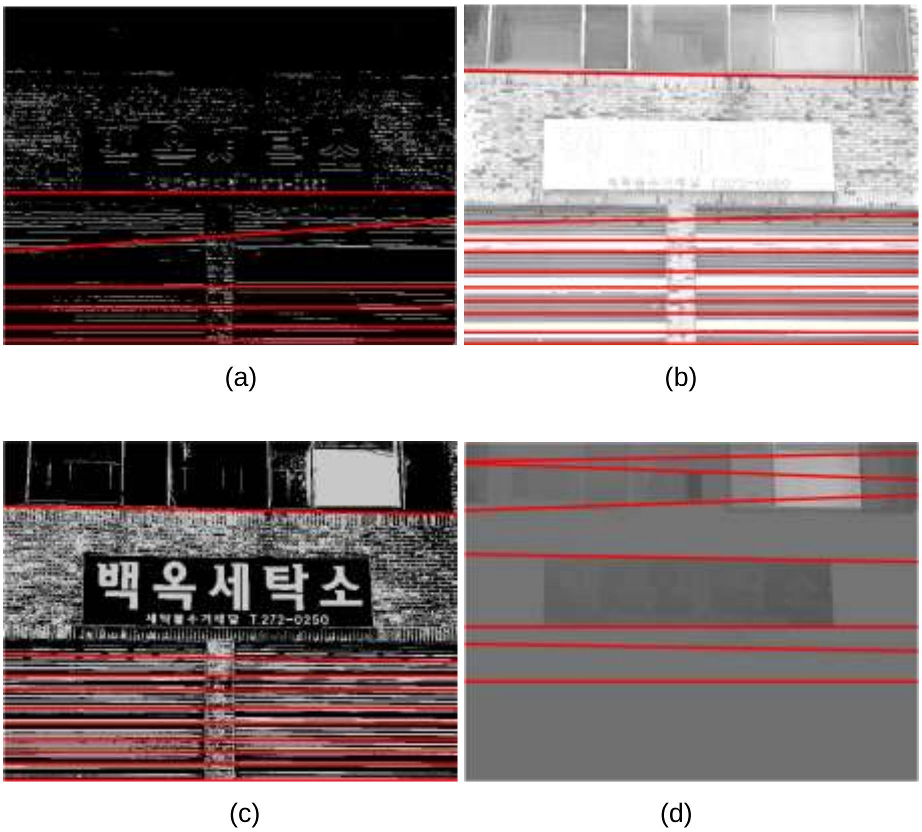


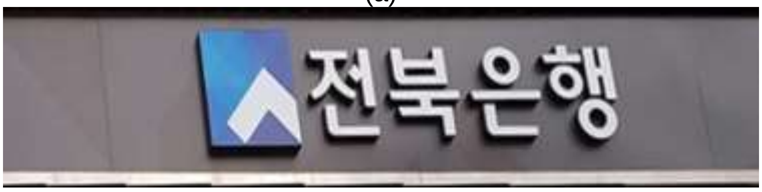
Figure 5. (a) Hough Transform Result using Horizontal Edge, (b) Color Distance, (c) Fuzzy C-mean, and (d) Proposed Method

2.5. Image Division and Distortion Correction

Image division is performed by storing the image between two straight lines using a linear equation calculated from the Hough transform. Alignment was performed based on y-intercept in order to distinguish the uppermost and lowermost straight lines.

Distortion reconstruction performed geometric transformation through projection of warped images [9-10, 21]. Four points on both sides of the image were used for image projection. In addition, the signboard part can be detected in the candidate area using the vertical axis component of the gray image [27-28].

(a)



(b)

Figure 6. Examples of (a) Warped and (b) Projection Image

3. Experimental Results

A total of 107 natural images were used in testing the proposed method which were then compared with the existing three methods. The performance index are subdivided as shown in the table below and the final results are somewhat lower because they all exhibit 0% detection except for certain conditions.

Table 1. The Performance Index.

Accuracy	Performance index
100% (case1)	Signboard detected correctly.
70%~90% (case2)	Analyze performance according to degree when background other than signboard is included.
80%~90% (case3)	Part of the signboard is cut off, but the text is fine.
0% (case4)	When the text of the signboard is lost and the signboard is not recognized

Table 2 shows the performance of each algorithm based on the evaluation index in Table 1. Accuracy and processing speed are the average values for the entirety. Figure 6 shows an example performance index.

Figure 7. Example of Performance Index

Table 2. The Performance Analysis

Algorithm	Accuracy (%)	Processing speed (s)
Horizontal edge(HE)	49.8131	0.0153
Color distance(CD)	76.0280	0.0421
Fuzzy c-mean(FC)	63.1776	4.1671
Proposed method(PM)	87.5701	0.0428

Experimental results show that PM has the highest accuracy at 87.5701% with an average operation speed for natural images at 0.0428s. FC is considered to be inadequate in terms of real-time application due to its slower computation speed even though it has 63.1776% accuracy. The processing speed of HE is fast, but it is difficult to use it alone without the constraint that the signboard image is located at the center of the image. On the other hand, CD with similar computation speed to PM showed a negative accuracy difference of 11.5421. CD can be used to improve PM's accuracy through CM and PM fusion because five images were detected out of ten images by CD that cannot be detected by PM.

4. Conclusion

In this paper, we propose the preprocessing technique of natural image to recognize signboards through Hough transform. The effectiveness of the proposed method is verified compared with the existing methods since the preprocessing method has a great influence on the recognition of the signboard. Experimental results show that the proposed method has high accuracy and suitable processing speed in real time. The proposed method can detect the signboard area without being restricted as to the size and position of the signboard. Sign recognition technology can be used not only to provide information necessary navigation of an autonomous mobile robots but also to the users. As shown in the experimental results, performance improvement can be expected by combining CM with the PM. In the future, we plan to increase recognition rate of signboards of various natural images and also continue to study word recognition through text detection.

Acknowledgments

This research was financially supported by the Ministry of Trade, Industry and Energy(MOTIE), Korea Institute for Advancement of Technology(KIAT) through the International Cooperative R&D program. (N046200012) and the Brain Korea 21 PLUS Project and Business for Cooperative R&D between Industry, Academy, and Research Institute funded Korea Small and Medium Business Administration in 2016 (Grants No.1601001522).

References

[1] H. Shen and X. Tang, “Generic Sign Board Detection in Images”, ACM, **(2003)**, pp. 144-149.

[2] K. H. Park and G. S. Lee, “A sign board region detection using Hough Transform”, vol. 12, no. 1, **(2009)**, pp. 57-59.

[3] J. Park and G. Lee, “Automatic detection and recognition of Korean text in outdoor signboard images”, Elsevier, pattern recognition letters, vol. 31, **(2010)**, pp. 1728-1739.

[4] J. H. Park and G. S. Lee, “Text Detection and Recognition in Outdoor Korean Signboards for Mobile System Applications”, vol. 46, no. 2, **(2009)**, pp. 187-193.

[5] A. Adam and C. Ioannidis, “Automatic Road-sign detection and classification based on support vector machines and HOG descriptors”, vol. 5, **(2014)**, pp.1-7.

[6] R. Harrabi and E. Ben Braiek, “Color Image Segmentation Based on a Modified Fuzzy C-mean Technique and Statistical Features”, vol. 2, no. 1, **(2012)**, pp.120-135.

[7] S. M. Mahajan and Y. K. Dubey, “Color Image Segmentation using Kernalized Fuzzy C-means Clustering”, IEEE, **(2015)**, pp.1142-1146.

[8] M. Yu and Y. D. Kim, “An Approach to Korean License Plate Recognition Based on Vertical Edge Matching”, IEEE, **(2000)**, pp. 2975-2980.

[9] I. Milevskiy and J. Y. Ha, “A Fast Algorithm for Korean Text Extraction and Segmentation from Subway Signboard Images Utilizing Smartphone Sensor”, vol. 5, no. 3, **(2011)**, pp. 161-166.

[10] H. Watabe and T. Kawaoka, “Extraction and Revision of Signboard Images for Recognition of Character Strings”, KES 2005, LNAI 3684, **(2015)**, pp. 709-715.

[11] J. W. Bae and S. H. Jung, “Coin Calculation system Using Banalization and Hue Histogram”, KIISE, vol. 21, no. 6, **(2015)**, pp. 424-429.

[12] H. C. Lee, K. B. Kim, H. J. Park and E. Y. Cha, “An alpha cut Automatic Set based on Fuzzy Binarization Using Fuzzy Logic”, JKIICE, vol. 19, no. 12, **(2015)**, pp. 2924-2932.

[13] Y. W. Woo, “Navigational Path Detection Using Fuzzy Binarization and Hough Transform”, KSCI, vol. 19, no. 2, **(2014)**, pp. 31-37.

[14] L. Li and C. L. Tan, “Real Scene Sign Recognition”, GREC 2009, LNCS 6020, **(2010)**, pp. 175-186.

[15] T. N. Dinh, J. Park and G. S. Lee, “Korean Text Detection and Binarization in Color Signboards”, IEEE, DOI 10.1109, **(2008)**, pp. 235-240.

[16] J. H. Yun, M. R. Choi and S. S. Lee, “An image enhancement algorithm for detecting the license plate region using the image of the car personal recorder”, Journal of the Korea Academia-Industrial, vol. 17, no. 3, **(2016)**, pp. 1-8.

[17] J. H. Heo and S. B. Lee, “A Study on Generation of Reverse Parking Guideline Reflecting Position of Camera”, Journal of the Korea Academia-Industrial, vol. 17, no. 3, **(2016)**, pp. 591-598.

[18] Y. C. Kim, Y. M. Kim, S. G. Kim, H. B. Kim and M. T. Cho, “Development of the Mechanical System and Vision Algorithm for the External Appearance Test Using Vision Image Processing”, Journal of the Korea Academia-Industrial, vol. 17, no. 2, **(2016)**, pp. 202-208.

[19] M. J. Kwon, “Study of Traffic Sign Auto-Recognition”, Journal of the Korea Academia-Industrial, vol. 15, no. 9, **(2014)**, pp. 5446-5451.

[20] S. W. Jang, Y. J. Park and M. H. Huh, “Character Region Extraction Based on Texture and Depth Features”, Journal of the Korea Academia-Industrial, vol. 14, no. 2, **(2013)**, pp. 885-892.

[21] C. D. Scott and R. E. Smalley, “Diagnostic Ultrasound: Principles and Instruments”, Journal of Nanosci. Nanotechnology, vol. 3, no. 2, **(2003)**, pp. 75-80.

[22] R. C. Gonzalez, “Digital Image Processing using MATLAB”, 2^{nd} edition, **(2015)**.

[23] R. Laganiere, “OpenCV 2 Computer Vision Application Programming Cookbook”, **(2012)**.

[24] S. Kim, “3D Computer Graphics”, **(2013)**.

[25] I. Pitas, “Digital Image Processing Algorithms and Applications”, **(2000)**.

[26] R. S. Wright, “OpenGL Superbible”, 3^{rd}, **(2005)**.

[27] D. Kim, “OpenCV Programming”, **(2012)**.

[28] D. Kim, “OpenCV Computer Vision Programming (Motion Detection/Tracking & Camera Calibration)”, **(2014)**.

Authors

Daeyeong Lim, received the B.S. and M.A. degree from the Chonbuk National University, Jeon ju, Korea, in 2010 and 2013, respectively, all in electrical engineering. He was Visiting Researcher with Texas A&M University, College Station, from 2011 to 2012. Currently, he is a pursuing the Ph.D. degree with the Division of Electronic Engineering in Chonbuk National University, South Korea. His research interests are in the areas of artificial intelligence, computer vision, neural network, deep learning, automation control, and IoT related industries.

Youngbaik Kim, received the B.S. and M.A. degree from the Chonbuk National University, Jeonju, Korea, in 1980 and 1982, respectively, all in electrical engineering. Currently, he is a principal engineer at KEPCO Engineering & Construction Company, Daejeon, Korea and is pursuing the Ph.D. degree in Chonbuk National University. His research interests are in the areas of nuclear instrumentation and control, thermal image processing, inertial navigation systems, and fire control systems.

Incheol Park, received the B.S. from the wonkwang University in 2008 and 2015. He is currently pursuing M.A. degree from the Chonbuk National University, Jeon ju, Korea, and researcher at Electronics and Telecommunications Research Institute(ETRI). His research interests are in the areas of Embedded system, Robotics, Image processing and Automation control

Jihoon Seung, received the M.A. degree in electronic engineering from Chonbuk National University, South Korea in 2013. He was Visiting Researcher with Texas A&M University, College Station, from 2011 to 2012. Currently, he is a pursuing the Ph.D. degree with the Division of Electronic Engineering in Chonbuk National University, South Korea. His research interests are in the areas of sensor fusion, navigation system, mobile robot control and battery management system.

Kilto Chong, received the Ph.D. degree in mechanical engineering from Texas A&M University, College Station, in 1995. Currently, he is a Professor at the School of Electronics and Information Engineering, Chonbuk National University, Jeonju, Korea, and Head of the Mechatronics Research Center granted from the Korea Science Foundation. His research interests are in the areas of motor fault detection, network system control, time delay system, and neural networks.

