International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:2, No:7, 2008

Binarization of Text Region based on Fuzzy Clustering and Histogram Distribution in Signboards

Jonghyun Park, Toan Nguyen Dinh, and Gueesang Lee

***Abstract*—**In this paper, we present a novel approach to accurately detect text regions including shop name in signboard images with complex background for mobile system applications. The proposed method is based on the combination of text detection using edge profile and region segmentation using fuzzy c-means method. In the first step, we perform an elaborate canny edge operator to extract all possible object edges. Then, edge profile analysis with vertical and horizontal direction is performed on these edge pixels to detect potential text region existing shop name in a signboard. The edge profile and geometrical characteristics of each object contour are carefully examined to construct candidate text regions and classify the main text region from background. Finally, the fuzzy c-means algorithm is performed to segment and detected binarize text region. Experimental results show that our proposed method is robust in text detection with respect to different character size and color and can provide reliable text binarization result.

***Keywords*—**Text detection, edge profile, signboard image, fuzzy clustering.

I. Introduction

SIGNS (or texts) are everywhere in our real world. A sign is an object that suggests the presence of a fact, condition or quality. It can be a displayed structure bearing characters or symbols, used to identify something or advertise a place of business. Especially, texts are important information embedded in natural scenes. They often carry useful information such as traffic signals, advertisement billboards, dangerous warnings, commercial contents, etc. Although text with a limited scope can be successfully detected using existing methods, it is difficult to detect texts with varying environment and embedded in real world. And then compared with other object detection applications in computer vision, many features are unavailable for a text detection. Generally text region in natural images is no motion, no approximation shape and no color reflection information.

To text detection and recognition from images, variety techniques have been researched for specific applications including document image analysis [1]-[3], content-based image/video indexing [4]-[6], assistance to visually impair person [7], and text detection, text recognition and translation system in natural scene images [8]-[12].

There are so many relevant works for detection and extraction of texts in the literature. Most of the text detection methods can be classified as edge-based method, connected component based method and texture-based methods. These methods have different advantage and disadvantage in reliability, accuracy and computation complexity.

Edge-based methods focus on the high contrast between the text and the background [8]-[14]. The edges of the text boundary are identified, and several heuristics are used to filter out the non-text regions. Generally, edge is very useful for image analysis which also can be used to find possible text areas. Text is mainly composed of the strokes in a certain direction, so it can be considered that the region with higher edge strength in the corresponding direction is the text region.

Connected component based methods use spatial structure analysis of the color connected component and can work well on most kinds of texts such as the characters on book covers, news titles and video captions [12]-[17]. These methods use a bottom-up method by grouping small components into successively larger components until all regions are identified in the image. A geometrical analysis is needed to merge the text components using the spatial arrangement of the components so as to filter out non-text components and mark the boundaries of the text regions.

Texture-based methods are based on the fact that texts in images have distinct textural properties that can be used to discriminate them from the background or other non-text regions [18][19][20][21][22]. And texture-based algorithms are more robust than the CC-based methods in dealing with complex background. X. Tang *et al*. proposed text detection using texture information of objects [20]. However, the method failed in the case of small texts. And M. Fujii *et al*. proposed the method using wavelet features from fix-size blocks of pixels and classify the feature vectors into text or non-text using neural networks [22]. However, the neural network based approach is not efficient in terms of computation cost with complex image.

There are so many approaches to perform in the case of variant text orientation, size, language, low resolution image, etc. [18][21][23][24][25][28]. Many approaches assume that

J. H. Park is with the Electronics & Computer Engineering, Chonnam National University, Gwangju, Chonnam, Korea (corresponding author to provide phone: 82-62-530-0147; fax: 82-62-530-1759; e-mail: jhpar@ chonnam.ac.kr).

T. N. Dinh was with Electronics & Computer Engineering, Chonnam National University, Gwangju, Korea (e-mail: Toan_mulmi@hotmail.com).

G. S. Lee is with the Electrics & Computer Engineering, Chonnam national University, Gwangju, Korea (e-mail: gslee@chonnam.ac.kr).

the text direction is horizontal or vertical and text font size is in limited scale. Although there are many techniques, the proposed methods were very restrictive in the type of texts they can process and therefore restrict the application field of the text detection.

In this paper, we focus on detecting and binarizing a main text on signboard images for mobile system applications. An automatic text translation system utilizes a mobile phone system to capture a signboard image in city containing shop name, detect main texts in the image and binarize texts. Automatic detection and binarization of texts from signboard images are prerequisites for an automatic text recognition task. However, due to the complexity of text appearance in signboard images, text detection is a difficult and challenging task in image processing.

We present a novel method to precisely detect text regions in signboard images to make a readable signboard for recognition part. The proposed approach is targeted towards being robust with respect to diverse kinds of handheld device, such as mobile phone, PDA and smart phone.

Since our approach analyzes text region contours directly and not the entire image, which is more stable than complete color images under varying signboard images, it is robust to varying illumination conditions. Furthermore, instead of performing a global analysis on the whole image, we consider candidate text regions separately to detect text region and perform the binarization on the text region, which makes our approach more efficient and robust in mobile system.

The rest of the paper is organized as follows: In Section II, we present the method for text detection using edge profile. We describe a segmentation and binarization using fuzzy c-means clustering and histogram in boundary of text region in Section III. In Section IV, we present some experiments on the proposed approach and evaluation results. Finally, Section V concludes the paper.

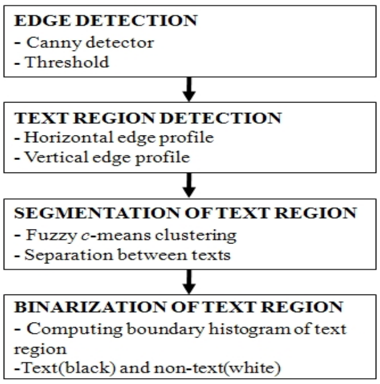


Fig. 1 The proposed method in our work

II. Text Region Detection

A. Adaptive Edge Detection

Edge detection is an important pre-processing step of our approach, although edge detection is not the focus of this paper. With good edge detection result, the time complexity of later text detection will be reduced and the detection accuracy will be improved. The canny edge operator [26] is applied on the gray-scale image to get the edged image. Generally, edge feature is robust to the size and the uneven illumination variation. We compute threshold value of edge magnitude by Eq. (1) around the center, because we assume that the interested main texts in the signboard are located on the center of the image.

$$\mathbf{TV} = \frac{\text{sum of edge magnitude in selected region around the center}}{\text{Selected region size}} + \alpha \qquad (1)$$

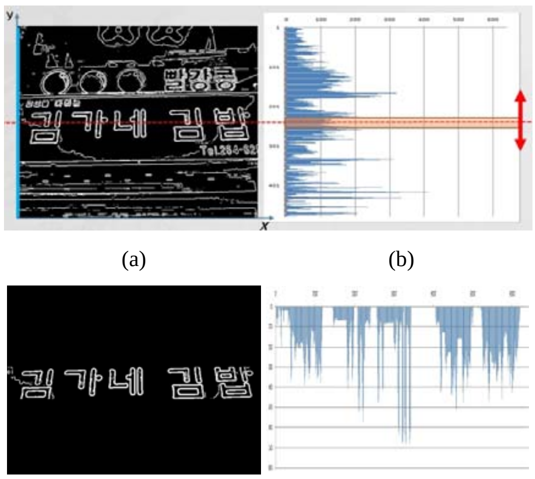
Where α is constant for detecting detail edge. In this paper for computing adaptive threshold value, we select regions around center line from edge magnitude image.

B. Text Region Detection Including Shop Name Based on Edge Profile

In order to detect text region in this step, we analyze horizontal components of texts in the edge map because the signboard text is usually arranged with horizontal direction. Then we calculate the horizontal histogram of edge image to detect the text regions. Fig. 3 shows edge profile of horizontal and vertical direction. We use Eq. (2) and (3) for detecting text region of horizontal and vertical direction. When scanned outside on the center line in image, selected region for horizontal profile is red box at Fig. 2(b) and selected region for vertical profile is width of used image. If horizontal profile is less than *HTR/3*, outside is non-text region and inside is text region. And then, we perform vertical region detection from the detected horizontal text image. To compute *HTR* and *VTR*, selected region size is 30 in *HTR* and entire region in *VTR*.

$$\mathbf{HTR} = \frac{\text{sum of horizontal edges in selected region}}{\text{selected region size}} \qquad (2)$$

$$\text{VTR} = \frac{\text{sum of vertical edges profile}}{\text{width of image}} \qquad (3)$$



(a) (b)

(c) (d)

Fig. 2 Text detection by the proposed method: (a) edge image, (b) horizontal edge profile, (c) detected text region of horizontal direction, (d) vertical edge profile

We detect the interested text region and remove other parts (background) in the edge map. The interested text region in our application is usually located around the center and the height of interested text is higher than that of other texts. Based on these characteristics, the interested text region corresponds to the largest horizontal histogram range located around the center of the image and in which the histogram values are high.

Fig. 3 The detected text region by the proposed method in Signboard image

III. Segmentation and Binarization

After we detect the main text region with edge profile property of horizontal and vertical direction, we can perform binarization of main text region and send the result to character recognition system. Although text region is perceived by the same features, the components of region may vary significantly. So, it is necessary to perform region clustering to compensate these effects.

In the detected text region, we perform segmentation with all color components into two distinctive colors to discriminate between text and other non-text region. To segment the detected region, we use a fuzzy c-mean (FCM) clustering described with Eq. (4) to depict the color distribution. The FCM has been shown to be an effective data clustering algorithm [27]. It is an iteratively optimal algorithm usually based on the least square method to make a fuzzy partition of the dataset. The object function of the traditional FCM is described by the following equation:

$$J = \sum_{k=1}^{N} \sum_{i=1}^{c} u_{ik}^{m} \left\| v_k - y_i \right\|^2 \qquad (4)$$

Where y_i, $i = 1, \cdots, N$, is the set of intensity component, $\boldsymbol{N}$ is the number of clusters, $\mathbf{m} \in [1, \infty)$ is a weighting exponent and u_{ik} is the amount of fuzziness of y_i as it approaches a clustering center v_k. The operation ||•|| is the Euclidean norm.

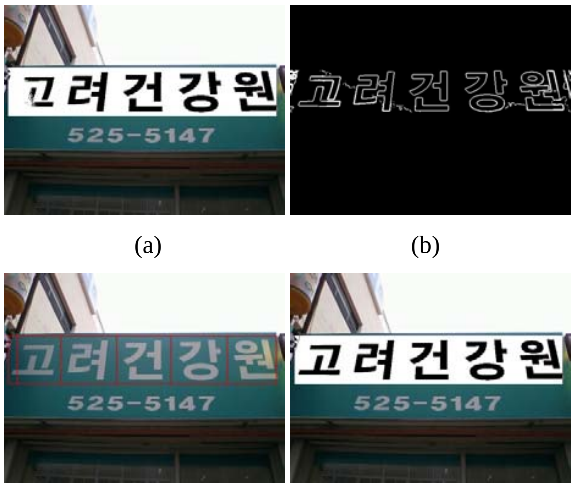
Because fuzzy c-means clustering is suited for fitting clusters, we use iterates by adjusting the parameters of the clustering model to optimize the membership function of data in dataset. The iteration stops when the difference between two successive iterations becomes negligible.

Upon convergence of the clustering algorithm, the two mean vectors can be recorded as the two dominant groups in a text region, i.e. text and non-text. Thus we can binarize the text region by measuring the boundary histogram distribution of the detected text region. The proposed method for binarization is presented in Algorithm I in detail.

ALGORITHM II. Binarization algorithm by boundary histogram distribution of detected text region.

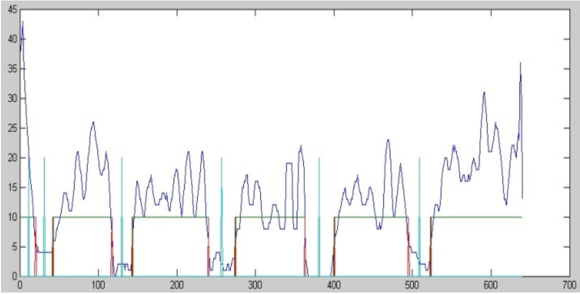
```
// It uses boundary histogram to distinguish between text and background // at the segmented text region. Then, "1" for text region and "0" for
// background are assigned by resulting histogram distribution.
for j=x_start, j<x_end, j++;
    if (seg_reg[y_start][j] == lable_value[k]), then
            label_histogram_1[k]++;
     if (seg_red[y_end][j] == label_value[k]), then
            label_histogram_2[k]++;
for i=y_start, i<y_start+x, i++;
     if seg_reg[i][y_start] == lable_value[k], then label_histogram_1[k]++;
    if seg_red[i][y_end] == label_value[k], then label_histogram_2[k]++;
for i=y_start; i<y_end; i++;
      for j=x_start; j<x_end; j++;
           if label_histogram_1[k] > label_histogram_2[k], then
                   label_histogram_1 = background;
           else  label_histogram_2 = background;
```

Fig. 5 is shown binarization result of the text region. If region segmentation is performed by entire region in Fig. 5(a), the result of "고" character is not good segmentation performing global feature because the signboard images have uneven illumination on left, middle and right. This result is provided low performance of a character recognition system. Therefore, we perform local segmentation in regions separated by considering text gap. Firstly, we perform separation using vertical edge profile in detected text region as Fig. 5 (b) and (c). In order to separate between texts, we use mean value of the edge profile. If vertical edge profile is less than mean/3, the image is separated in between texts. Then, each region of red box in Fig. 5(c) is segmented using local features based on the FCM. The result is shown in Fig. 5 (d).



(a) (b)

(c) (d)



(e)

Fig. 4 Segmentation and binarization by the proposed method: (a) segmentation result, (b) edge image in the detected text region, (c) separated image by vertical profile, (d) segmentation result by local feature, (e) vertical profile in the (b) image

IV. Experimental Results

To assess the performance of the proposed algorithm, we have conducted an experiment using signboard scenes captured by various conditions. Currently, our approach has been implemented in C++ language under Windows-XP with Intel Pentium-IV and 1GB memory. In this paper, in order to evaluate the actual performance of the proposed approach, we test 445 real signboard images which include various conditions. Text appearance varies with different colors, orientations and the character font sizes. The image resolution is 640 × 480 pixels. Used Korean signboard images are captured by Samsung mobile PDA phone in our city.

Fig. 5 Binarization result by the proposed method in Korea signboard images

Fig. 6 Binarization results by the proposed method in English signboard images

Fig. 5 and 6 show the experimental results of various kinds of images selected from Korean and English signboard images. The figure contains examples of automatic main text detection, where the rectangles give detection results. We can see that the main text region is accurately detected by the proposed method proposed in Section II.

Generally, signboard image captured from a mobile phone might have some uneven illumination because of gradation, shading and shadow in text region, which can reduce detection rate and recognition accuracy. Fig. 7 shows two sets of examples of text detection and binarization, where the left side of each image Fig. 7(a) is the binarization result with segmentation using global features, the right side of each group Fig. 7(b) is the result with segmentation using local features. Observe that only partial text can be detected from images because the vertical edge profile makes the missing text stroke on the right side.

International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:2, No:7, 2008

(a) (b)

Fig. 7 Examples of binarization in signboard images with uneven illumination: (a) binarization results by global features, (b) binarization results by local features

The evaluation result of the test image set is summarized in Table I. And used dataset is 445. In order to evaluate performance results, we classify into three groups. 1st group is failure the case cannot use next step for text recognition, 2nd group is binarized main text and background but including isolated regions and 3rd group is good without isolated regions.

TABLE I
TEXT DETECTION AND BINARIZATION RATIO IN USED DATASET

	Num. of images	Detection & Binarization ratio
1st group	57	13%
2nd group	252	57%
3rd group	136	30%

Generally, signboard images show some incorrectly result of text binarization. The errors occurred because some texts are strongly uneven-illumination and complex background in the image and some non-text objects have strong text like texture patterns.

Experimental results show that the binarization procedure with FCM using local features is quite usable. Even when the uneven-illumination inside the detected region is generated, the binarized regions become nearly coincident, which makes the binarization process more robust.

V. CONCLUSION

In this paper, we have presented an approach for automatic detection and binarization of signboard texts for application to mobile systems. The proposed method can robustly detect and binarize main texts from signboard images. Firstly, we perform detection of main text region using edge histogram method with horizontal and vertical direction in edge map image. After text region verification, detected region is segmented by fuzzy c-mean clustering and each region is distinguished as text region (black) and background region (white) by boundary histogram in detected text region. The experimental results show that the proposed method can be successfully applied to detect main text in signboard images captured under various conditions.

REFERENCES

[1] G. Nagy, "Twenty years of document image analysis" IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 38-62, 2000.

[2] H. Peng, F. Long, Z. Chi, "Document image recognition based on template matching of component block projections," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 9, pp. 1188-1192, 2003.

[3] I.S. Oh, C.Y. Suen, "Distance features for neural network-based recognition of handwritten characters," International Journal on Document Analysis and Recognition, vol. 1, pp. 73-88, 1998.

[4] H. Li, D. Doermann, O. Kia, "Automatic text detection and tracking in digital videos," IEEE Transactions on Image Processing, vol. 9, no. 1, pp. 147-156, 2000.

[5] J. Xi, X. Hua, L. Wenyin, H.J. Zhang, "A video text detection and recognition system," International Conference on Multimedia and Expo, pp. 873-876, 2001

[6] A.K. Jain, B. Yu, "Automatic text location in images and video frames," Pattern Recognition, vol. 31, no. 12, pp. 2055-2076, 1998.

[7] N. Ezaki, K. Kiyota, B.T. Minh, M. Bulacu, L. Schomaker, "Improved text-detection methods for a camera-based text reading system for blind persons," International Conference on Document Analysis and Recognition, pp. 257-261, 2005.

[8] X. Chen, J. Yang, J. Zhang, A. Waibel, "Automatic detection and recognition of signs from natural scenes," IEEE Transactions on Image Processing, vol. 13, no. 1, pp. 87-99, 2004.

[9] K. Jung, K.I. Kim, A.K. Jain, "Text information extraction in images and video: a survey," Pattern Recognition, vol. 37, pp. 977-997, 2004.

[10] J. Gao, Y. Yang, Y. Zhang, A. Waibel, "Text detection and translation from natural scenes," Tech. Report CMU-CS-01-139, 2001.

[11] Q. Ye, J. Jiao, J. Huang, H. Yu, "Text detection and restoration in natural scene images," Journal of Visual Communication and Image Representation, vol. 18, pp. 504-513, 2007.

[12] G. Obinata, A. Dutta, Vision Systems: Segmentation and Pattern Recognition, I-Tech, pp. 307-332, 2007.

[13] J. Gllavata, R. Ewerth, B. Freisleben, "A robust algorithm for text detection in images," International Symposium on Image and Signal Processing, vol. 2, pp. 611-616, 2003.

[14] W. Wu, X. Chen, J. Yang, "Detection of text on road signs from video," IEEE transaction Intelligent Transportation Systems, vol. 6, no. 4, pp. 378-390, 2005.

[15] D. Chen, H. Bourlard, J.P. Thiran, "Text identification in complex background using SVM," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 621-626, 2001.

[16] Y. Zhong, K. Karu, A.K. Jain, "Locating text in complex color images," pattern Recognition, vol. 28, no. 10, pp. 1523-1536, 1995.

[17] A.K. Jain, B. Yu, "Automatic text location in image and video frames," International Conference on Pattern Recognition, vol. 2, pp. 1497-1499, 1998.

[18] R. Lienhart, A.Wernicke, "Localizing and segmenting text in images and videos," IEEE Transactions on Circuits System, vol. 12, no. 4, pp. 256–268, 2002.

[19] Y. Zhong, H. Zhang, A.K. Jain, "Automatic caption localization in compressed video," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 4, pp. 385–392, April 2000.

[20] X. Tang, X. Gao, J. Liu, H. Zhang, "A spatial-temporal approach for video caption detection and recognition," IEEE Transactions on Neural Network, vol. 13, no. 4, pp. 961–971, 2002.

[21] K. In Kim, K. Jung, J. Hyung, "Texture-based approach for text detection

in images using support vector machines and continuously adaptive mean shift algorithm," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 12, pp. 1631–1639, 2003.

[22] M. Fujii, W.J.R. Hoefer, "Filed-singularity correction in 2-D time-domain Haar-wavelet Modeling of waveguide components," IEEE Transactions on Microwave Theory and Techniques, vol. 49, no. 4, pp. 685–691, 2001.

[23] Q. Ye, W. Gao, W. Wang, W. Zeng, "A robust text detection algorithm in images and video frames," International Conference on Information, Communications and Signal Processing, vol. 2, pp. 802-806, 2003.

[24] Y. Zheng, H. Li, D. Doermann, "Machine printed text and handwriting identification in noisy document images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 3, pp. 337-353, 2004.

[25] S. Belongie, J. Malik, J. Puzicha, "Matching shapes," International Conference on Computer Vision, vol. 1, pp. 454-461, 2001.

[26] M. Sonka, V. Hlavac, R. Boyle, Image Processing, Analysis and Machine Vision, Thomson Asia Pte Led.

[27] R.L. Cannon, J.V. Dave, J.C. Bezdek, "Efficient implementation of the fuzzy c-means clustering algorithm," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8. No. 2, pp. 248-255, 1986.

[28] Y. Liu, S. Goto, T. Ikenaga, "A contour-based robust algorithm for text detection in color images," IEICE Transaction Information and System, vol. E89-D, no. 3, pp. 1221-1230, 2006.