International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

AUTOMATIC DETECTION AND RECOGNITION OF TAMIL SHOP NAME IN OUTDOOR SIGNBOARD

R.Karthikeyan

Asst professor Dept of Computer science and Engineering BIHER, Chennai

rkarthikeyan78@yahoo.com

Abstract

In this paper, a system for automatic detection and recognition of Tamil texts shop names in outdoor signboard images is described. The system includes detection, binarization and extraction of text in a signboard image captured by a camera of a mobile phone for the recognition of the shop name. It can deal with different font styles and sizes as well as illumination changes. Individual characters detected by connected component analysis are recognized by using nonlinear mesh, in which feature vectors of vertical and horizontal components are extracted from the binarized image. Proposed methods have been applied to a Tamil text translation system, which can automatically detect and recognize Tamil texts and generate the translation result.

Key words: Text detection, Text segmentation, Recognition, signboard image.

Introduction

Vision techniques an image are more and more considered in a drastic expansion of low priced cameras such as mobile system and digital cameras. And text recognition is a fast changing field, which is included in a large spectrum, named text understanding. Interesting applications such as mobile phones operating becomes more and more powerful opening up a new range of application. With a handheld device, natural scenes in daily life may be analyzed to give user access to text, coupled with a text-to-transformation algorithm and so on. Therefore such devices are really expected [1][2].

There have been techniques developed for the detection and recognition of texts from document images[3][4] content-based image/video indexing [5][6], assistance to visually impaired person and text translation in natural scene images [7].own a mobile phone with an embedded camera, which is all that our system needs.

Various approaches have been studies in the past for detection and recognition of text in natural scenes. These methods take into consideration different properties related to text in an image such as color, intensity, connected- components, edges etc. These properties are used to distinguish text regions from their background and texts within the natural scenes. In order to extraction text regions, color clustering method was proposed by Wang and Kangas in [8].

The scene image is clustered into different color components and an intensity component. The method utilizes the fact that generally the color in texts different from the color in the background. And potential text regions are extracted by connected-component based heuristics from these layers. An aligning and merging analysis method is used for extracting texts. The approach used in which each row and column information is analyzed. The proposed method is robust in locating mostly Chinese and English in scenes; some false alarms happened because of uneven illumination or light reflection in natural scenes.

Also text detection approaches were proposed by [9] and [10]. The approaches are based on color continuity [9] and frequency, orientation and spacing of texts within an image[10]. The methods show that the algorithm is robust in most cases, however it is difficult to detect small texts. Also in the case of low contrast in the image, misclassifications occur in the texture segmentation.

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

A support vector machine classifier was proposed to segment text from non-text in an image or video frame by [11]. Text is extracted in multi-scale images using edge components, morphological operation and projection profiles. And the extracted text regions are verified using wavelet features and SVM. The proposed approach is robust with respect to variance in color and scale of texts. However the method has complexity of algorithm and processing time.

In this work, we are interested in automatic detection and recognition of texts in outdoor signboard images for mobile applications. The application scenario is as follows. A user uses a camera to capture an image or a sequence of images with text in the center of the scene. In the image, the text is supposed to be extracted for recognition and finally it can be applied to applications exploiting the text information.

In this paper, we present an approach for detecting and recognizing text in signboard images. The text includes stereotypical forms such as shop names. In order to evaluate performance of the proposed approach, our database of signboard images were taken in real environment and we assume that the text is located around the center line of the image. The proposed method consists of text detection, binarization and recognition. First, a local clustering is used to effectively handle luminance variations of the captured signboard images. Connected component analysis is carried out to obtain the bounding box of the individual character. Finally, the text is recognized using extracted direction features based on nonlinear mesh in each character region.

We have applied the proposed approach in developing a Tamil text translation system. The system has been designed for mainly foreigners, who do not have knowledge on the Tamil language, to understand the Tamil text in signboards.

Text Detection

The text in outdoor signboard images is affected by changes of lights, orientation and location of Signboards, where the orientation is decided by the view angle of the Camera. In this paper, we assume that the text region is located around the center line.

2.1 Detection of Candidate Text Region

Generally, intensity information of an image is an important feature for text detection, but it is not robustness to lighting variations. On the other hand, the edge component is less sensitive to lighting changes. Therefore, we use edge-based features in the coarse detection step. The canny edge operator is applied on the gray-scale image to get the edged image. In order to detect coarse 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. 1 shows edge profile of horizontal and vertical direction. We use Eq. (1) and (2) 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. 1(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.

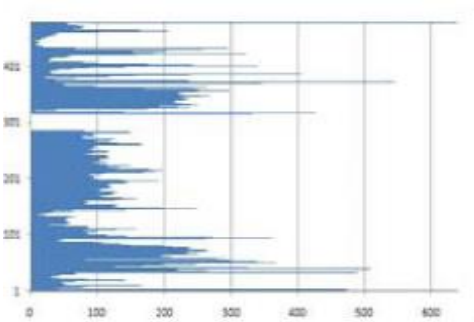
$$HTR = \frac{\text{Sum of horizontal edges in selected region}}{\text{Selected Region size}} \quad (1)$$

$$VTR = \frac{\text{Sum of vertical edges profile}}{} \quad (2)$$

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

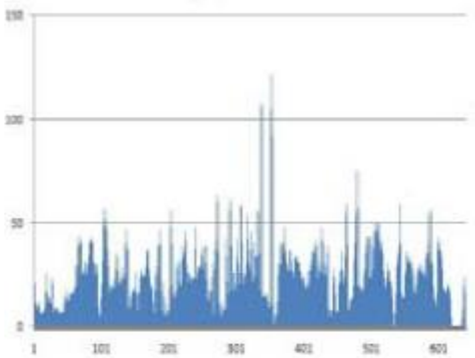
Width of image

(a)



(b)

(c)



(d)

Fig.1 Txet detection by a proposed method:

a) Edge Image b) Horizontal edge profile

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

c)Detected Text Region 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. 2 is shown the detection result in outdoor signboard image.

Fig.2 The detected text region by the proposed algorithm in signboard image

2.2 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 then send the result to character recognition system.

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. (3) to depict the color distribution. FCM clustering is an unsupervised approach that has been widely used in image segmentation [12]. It assigns pixels to each cluster by using fuzzy memberships.

Let Y ={yk |1≤k≤N,and,$y_k \in R^d$} presents the observed image, where d=I the gray level case. Mathematically, the standard FCM objective function of partitioning a dataset X into c-clusters is given by the following equation:

$$J_m = \sum_{i=1}^{c} \sum_{k=1}^{N} \pi_{ik}^{m} \left\| y_k - v_i \right\|^2 \qquad (3)$$

Where $\{v_i\}_{i=1}^{C}$ are the centroids or prototypes of the clusters, the parameter m is a weighting exponent on each fuzzy membership and determines the amount of fuzziness of the resulting classification and the array of membership function A = [Π_{ik}] is a fuzzy partition matrix satisfying

$$\Lambda \in \left\{ \pi_{ik} \in [0,1] \mid \sum_{i=1}^{c} \pi_{ik} = 1, \forall k, 0 < N, \forall i \right\}$$

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.

Fig. 3 is shown binarization result of the text region. If region segmentation is performed by entire region in Fig. 3(a), the result of the 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

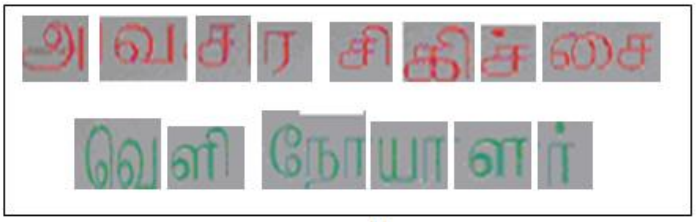
International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

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. 3(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. 3(c) is segmented using local features based on the FCM. The result is shown in Fig. 3(d).

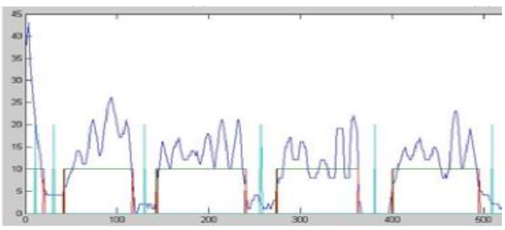
(a)

(b)



(c)

(d)



(e)

Fig.3 Segmentation and binarization by the proposed method

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

a) Segmentation result

b) Edge image in the detected text region

c) Separate image by vertical profile

d) Segmentation result by local feature

e) Vertical profile in the (b) image first word.

2.3 Layout Analysis

The objective of layout analysis in the detected text region images is to align characters in an optimal way, so characters belong to the same character will be aligned together. Tamil character layout has some unique features. How can a system know if a text region is character. In order to this work, we use layout analysis approaches, which utilize various heuristics, to deal with this problem. First of all, we perform noise and isolated regions in segmented text region because they provide cause of error in merging characters. The criterion of noise removal is ratio of vertical and horizontal of labeled objects. So, noise regions are removed that have ratio of 10:1 comparing text with noise. Therefore we perform merge of separated sub-components in a character by the following bottom procedure.

-First, we find objects by the labeling in the binarized text on the FCM. The result is shown in Fig. 3(d).

-Second, draw a horizontal line on center of text region; seek an object with gravity center below a horizontal line.

-Third, seek an object with the center point at smallest rectangle from the bottom on center line of detected text

region.

-Fourth, merge objects on vertical direction of the detected object. This is 1st merge step.

-Fifth, compute mean value of vertical length for text object.

-Sixth, merge as one character if the vertical length is less 1.5 than average length from neighbor objects. Therefore we can extract a character from text string.

Feature Extraction and Character Recognition

In this sub-section, we focus on the character only. Recognition system for the characters captured from outdoor signboard images is more difficult than that of document analysis. For a traditional document analysis task, a scanner with a stable embedded lighting system is used to obtain high quality images, which are then easily binarized. However, the noise ratio is much higher because the signboard image is captured by a camera from outdoor condition under various lighting conditions. So if we use intensity features for recognition, it is difficult to remove noises before the feature extraction for the recognition. Thus, binary images keep all the pattern structure of characters essential to the decision making stage.

We use a nonlinear mesh for feature extraction. Because of its superior properties, nonlinear mesh has been widely used for data compression, handwriting recognition and other pattern recognition in recent years. In recognition applications, nonlinear mesh approach has been applied to binary image.

In order to extract feature, we compute cell feature value in region with black pixels by vertical and horizontal ratio after divided a character image into mxn nonlinear mesh of two directions of vertical and horizontal. The proposed feature extraction is described in detail by the following:

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

First step: divide a character image into mxn nonlinear mesh on vertical and horizontal direction.

Second step: compute $RLV_{x,Y}$ and RLHx,y of the vertical and horizontal direction respectively in black pixel region.

Third step: compute contribution on DCVx,y and DCHx,y

$$DCH_{x,y} = {RLH_{x,y}} \Big/ {(RLH_{x,y} + RLV_{x,y})} \qquad (4)$$

$$DCV_{x,y} = {RLV_{x,y}} \Big/ {(RLH_{x,y} + RLV_{x,y})}, \qquad (5)$$

$$where\ 0 \leq x < W, 0 \leq y < H$$

Fourth step: computer average values of DCH and DCV in all black pixels within each cell, and make feature vectors of mxnx2 dimension.

We obtain the feature vector for used all characters, a standard for commonly used two different fonts, such as Gothic and Ming-style printing types. All of these character images are binary scale. For each type, we generate the 200 font types and each font has 800 characters. We first extract candidate group until top 5-order using Manhattan distance for compare similarity from data set. Finally, we reduce feature dimension for character recognition from the candidate group by Fisher dimension of Eq. (6).

$$F_{ij}(k) = \frac{\partial_{ij}(k)}{\partial_i(k) + \partial_j(k)} \qquad (6)$$

Where F_{ij} (k) is k-th feature value in i-th character for character recognition and j-th character for character recognition. a, (k) is variance of k-th feature value in i-th character and aj (k) is variance of k-th feature value inj th character. Finally, a, (k) is variance of k-th feature value of all sample characters. We extract DSF features of 252 from candidate character group. Character recognition of Fisher dimension is performed based on vectors of 32 dimensions extracted by Eq. (6). Therefore we can recognize one character by comparing distance with Fisher dimension defined for two characters.

Conclusion

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. Our Research is towards to solve those errors in the translation.

In this paper, we described a system for automatic detection and recognition of texts in signboard images captured by mobile phone cameras. The captured images are taken from the natural lighting environment with noises and irregularity. The proposed method can robustly detect and recognize texts from such images. The edge-based method is used for detection of the candidate text region and the text is binarized by clustering in local regions of individual characters. And nonlinear-based feature extraction is employed for character recognition. The detection and recognition technique has been implemented for a tamil translation system,which can automatically detect Tamil text input from signboard images, recognize the text,and translate the recognized Text.

International Journal of Modern Agriculture, Volume 9, No. 4, 2020
ISSN: 2305-7246

References

1. Adrian Canedo-Rodríguez,Soohyung Kim,Jung H. Kim,Yolanda Blanco-Fernández “English to Spanish Translation of Signboard Images from Mobile Phone Camera”, IEEE Conferences on SOUTHEASTCON’09,May 2009.

2. Yu Zhong, Hongjiang Zhang and Jain A.K, “Automatic caption localization in compressed video,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 22, Issue 4,April, 2000, pp. 385-392.

3. Canny, J., *A Computational Approach To Edge Detection*, IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679-714, 1986

4. Xilin Chen, *Member, IEEE*, Jie Yang, *Member, IEEE*, Jing Zhang, and Alex Waibel, “Automatic Detection and Recognition of Signs From Natural Scenes”. IEEE Transactions on Image Processing, Vol. 3, No.1, January 2004.

5. H. Li, D. Doermann, 0. Kia, "Automatic text detection and tracking in digital videos," IEEE Transactions on Image Processing, vol. 9,no.1,pp.147-156,2000.

6. 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

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. K. Wang, J.A. Kangas, "Character location in scene images from digital camera," The journal of the Pattern Recognition Society, vol. 36, pp. 2287-2299, 2003.

9. K.C. Kim, H.R. Byun, YJ. Song, YW. Choi, S.Y Chi, K.K. Kim, YK. Chung, "Scene text extraction in natural scene images using hierarchical feature combing and verification," International Conference on Pattern Recognition, pp. 679-682, 2004.

10. V. Wu, R. Manmatha E.M. Riseman, "Textfinder: An automatic system to detect and recognize text in images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 11, pp. 1224-1229, 1999.

11. R. Lienhart, A. Wernicke, "Localizing and segmenting text in images and videos," IEEE Transactions on Circuits and Systems for Video Technology," vol. 12, no. 4, pp. 256-268, 2002.

12. X. Li, T. Zhang, Z. Qu, "Image segmentation using fuzzy clustering with spatial constraints based on Markov random field via Bayesian theory," IEICE Transactions on Fundamentals, vol. E91-A, no. 3, pp.723-729, 2008.

13. Gao, J. and Yang, J., "An Adaptive Algorithm for Text Detection from Natural Scenes," Proceedings of Computer Vision and Pattem Recognition (CVPR 2001).

14. Jayasakthivel murugan K and Dr.Dorairangaswamy M A., “Translating tamil to English signboard images using mobile phone camera” Proceedings of the international conference on science and Technology (ICST 2011).