ADAPTIVE HISTOGRAM ANALYSIS FOR SCENE TEXT BINARIZATION AND RECOGNITION

M. Basavanna[1], P. Shivakumara[2], S. K. Srivatsa[3], G. Hemantha Kumar[4]

[1] Post Graduate Department of Computer Science
Government College (Autonomous), Mandya
Karnataka-India

[2] Faculty of Computer Science and Information Technology, University of Malaya (UM)
Kuala Lumpur, Malaysia

[3] St. Joseph College of Engineering
Chennai-Tamil Nadu-India

[4] Department of Studies in Computer Science
University of Mysore, Mysore
Karnataka-India.

Email: basavanna_m@yahoo.com[1], shiva@um.edu.my[2], profsks@rediffmail.com[3], ghk.2007@yahoo.com[4]

Abstract

Scene text binarization and recognition is a challenging task due to different appearance of text in clutter background and uneven illumination in natural scene images. In this paper, we present a new method based on adaptive histogram analysis for each sliding window over a word of a text line detected by the text detection method. The histogram analysis works on the basis that intensity values of text pixels in each sliding window have uniform color. The method segments the words based on region growing which studies spacing between words and characters. Then we propose to use existing OCRs such as ABBYY and Tesseract (Google) to recognize the text line at word and character levels to validate the binarization results. The method is compared with well-known global thresholding technique of binarization to show its effectiveness.

Keywords: Adaptive histogram analysis, Scene text binarization, Scene text recognition, Region growing, Word segmentation, Global thresholding.

1.0 INTRODUCTION

Scene text recognition through binarization has drawn attention of the researchers to improve the recognition rate. This is because applying existing available OCR's on scene text directly leads to poor recognition rate typically from 0% to 45% [1]. Besides, achieving good recognition rate is important for the applications, such as assisting visually impaired people to walk freely on street, assisting drivers to drive vehicle safely on road, assisting tourists to look for tourist spots with the help of GPS information and retrieving events from the large database based on semantics [2, 3]. There are two types of methods in literature for recognizing scene text: (1) methods that recognize without binarization and segmentation of words and characters [2, 3] and (2) methods that recognize through robust binarization [4-6]. The first method works well but it requires large number of features and classifiers along with multiple thresholds to achieve good recognition rate. The second method works based on foreground and background features and it requires segmentation of text lines, words and characters. These methods are good when the segmentation method separates foreground (text) and background in a proper manner. The first method requires pre-defined knowledge of data while these knowledge is not required in the second method. Therefore, we propose a new method by considering the advantage of the first method that consists of sliding window operation, foreground and background information from second method as well to overcome the problems of two methods. Overall, to the best of our knowledge, the existing methods achieve around 75% recognition rate at word and character levels [2, 3]. However, the above-mentioned applications require more than 90% recognition rate in document analysis for developing a working model. Therefore, there is a great demand and scope for discovering new method that can give a good recognition rate for scene text in natural scene images without assumptions.

2.0 RELATED WORK

It is known that text binarization of scanned document is a familiar topic in the field of document analysis. Text binarization methods were developed to recognize scanned document with plain and white background. However, these methods may not be suitable for text recognition in natural scene images as these images are complex, i.e., having clutter background and different appearance of text [1-6]. Therefore, in this work, we review both text binarization methods related to document and scene text analysis. Besides that, thresholding techniques (global and local) are quite popular in the field of document analysis. Several improvements over thresholding techniques have been proposed recently in the field of document analysis by extending them to support scene text binarization.

Otsu's method is popular for binarizing the image with less parametric as it works based on global thresholding technique. However, it gives good results when the image gives two distributions (foreground and background). Based on these distributions, the method computes global threshold [7] by minimizing the variance between the two distributions. The limitation of these distributions are resolved by Ye et al. [8] that apply Otsu method recursively until it gives a stable mode instead of two modes. In order to avoid the problems of global thresholding, Moghaddam and Cheriet[9] proposed a method based on an adaptive procedure, which is same as Otsu's method by using local information. The threshold automatically classifies text and non-text regions. Out of many adaptive threshold binarization methods, Sauvola and Pietikainen's method [10] is one of the best methods that determine threshold based on Niblack's method [11] procedure, which uses windows for text binarization. An improved version of Sauvola and Pietikainen's method has been developed by Trier and Jain [12] to classify non-text regions and their designed threshold value determination requires two parameters. The state of the art method for text binarization is proposed by Shayegan et al. [13]. This method classifies background roughly and uses the same background information for estimating threshold values. Su et al. [14, 15] proposed text binarization method which is placed first in the DIBCO'09 binarization contest done by Gatos et al. [16]. This method involves four steps: (i) background separation by polynomial fitting on the rows; (ii) stroke edge finding using Otsu's in gradient domain, (iii) thresholding using local information by computing average for the detected edge pixels over a window and, finally, (iv) post processing to improve the accuracy. Fabrizio et al. [17] proposed method for text binarization, which is placed second in the same contest based on the toggle mapping morphological operations. To solve the problem of the salt-and-pepper noise, the method ignores the pixels that are close for both erosion and dilation operations. Then the method classifies the pixels as text, background, and uncertain. Based on boundary class, the uncertain pixels are grouped as background. The method which is placed third in the same contest was proposed by Rivest-Henault et al. [18]. This method involves the level set framework to find the boundaries of text strokes and binarizes a document image. This method consists of several steps: (i) Initialization with the help of stroke map (SM) as in Farrahi-Moghaddam and Cheriet [19]; (ii) The level set framework for corrections using erosion mode and local linear models; and finally, (iii) The same level set operation with stroke Gray level force, which gives the final text regions as the interior regions.

Though the document image quality is affected by blur or degradation, there are chances of document containing a region that can be classified as text and background. Several learning methods generally estimate probable text and background regions, and study the behavior of regions to classify it as text or background. For instance, Don [20] proposed a simple thresholding technique for identifying text and background. Bernsen[21] is the first one who successfully modifies Otsu's method to become an adaptive method. However, the main weakness of this method is that finding correct values for the parameters. The second weakness of this method is that the global Otsu threshold itself. Moghaddam and Mohamed Cheriet[22] proposed AdOtsu: An adaptive and parameter less generalization of Otsu's method for document image binarization. Similar to Otsu method, this work does not depend much on parameters.

It is observed from the above methods that the document OCR engine does not work for camera based natural scene images due to the failure of text binarization in handling non-uniform background and non-illumination. Therefore, poor character recognition rate (67%) is reported in the ICDAR-2003 competition data [23]. This shows that despite high contrast of camera images, the best accuracy reported is 67% so far. It is noted that character recognition rate varies from 0% to 45% [1] for applying OCR directly on natural scene images. The experimental results of the existing baseline methods such as Niblack[11] and Sauvola et al. [10] show that thresholding techniques give poor accuracy for the scene images. An automatic text binarization method for color text areas in images and video based on convolutional neural network is proposed by Saidane and Garcia [4]. The performance of the method depends on the number of training samples. Edge based text binarization for video text image has been proposed by Zhou et al. [5] to improve the video character recognition rate. This method takes Canny of the input image as an input and applies a modified flood fill algorithm to fill the gap on the contour. This method only works well for small gaps but not for big gaps on the contours. In addition to this, the method's primary focus is on the graphics text and big font but not both graphics and scene text.

Recently, Roy et al. [6] have proposed Wavelet-Gradient-Fusion for video and image text binarization method. In this work, they propose a new method using fusion of horizontal, vertical and diagonal information (obtained by the wavelet transformation process) and the gradient on text line images to enhance the text information. They apply k-means with k=2 on row-wise and column-wise pixels separately to extract possible text information. Next, the method uses connected component analysis to merge some subcomponents based on nearest neighbor criteria. The foreground (text) and background (non-text) are separated based on new observation that the color values at edge pixel of the components are larger than the color values of the pixel inside the component. Finally, they use Google Tesseract OCR to validate their results and the results are compared with the baseline thresholding techniques to show that the proposed method is superior to existing methods in terms of recognition rate on 236 video and 258 ICDAR 2003 text lines.

From the above literature review, it is found that there is no perfect method to give perfect solution on text binarization and recognition of scene text images. Hence, we propose a new method called adaptive histogram based method to overcome the problems of the existing methods.

3.0 PROPOSED METHODOLOGY

In order to recognize the text line detected by the text detection methods, either we need to propose our own classifier or use available classifier by OCRs. In this work, we choose the second option that uses existing OCR rather than developing our own OCR or classifiers. We know that OCR accepts only binary image to recognize the text. It is also true that separating foreground (text) and background (non-text) of scene text line is challenging due to degradations, loss of information and distortions. Therefore, we propose an adaptive histogram analysis for each sliding window over a text line. Since the proposed method requires segmented text lines (i.e., detected by the text detection method) as input, we consider the height of the text block (i.e., fixed by the text detection method) as width of the sliding window and form a square window. For segmented text line image, we plot a histogram by considering pixel values in X axis and number of pixels in Y axis. Then, we choose pixels that give a highest peak in the histogram bin as text pixels and display them as white pixels and generate output results in binary. We use region growing to segment the words and characters by studying the number of regions grown in the place of space between the words and characters. Dynamically, the method chooses threshold to segment the words. The dynamic threshold is determined based on the fact that the space between the words is higher than space between the characters. The segmented words and characters are subjected to the proposed adaptive histogram analysis to obtain binary image. Then we use ABBYY and Tesseract OCR to recognize the words and characters. Our proposed method is divided into two sub-sections. First section describes words and characters segmentation and second section describes an adaptive histogram analysis for text binarization.

3.1. Segmentation of Words and Characters

We are inspired by the work presented in [24] for traversing arbitrarily orientated text in video. We propose to modify the growing method such that it studies the space between the words and characters. While growing, the method counts the number of regions grown in the space between the words and characters. Results are generated based on the distance between the words and characters. Our method applies k-means clustering with k=2 to separate low distances in a group and high distances in another group. Means are computed for both the groups. The low mean is considered for segmenting characters and high mean is considered for segmenting words. Then, we modify the region growing to determine thresholds dynamically for segmenting both words and characters. The words and characters segmentation is illustrated in Figure 1. The first column shows input text lines segmented from input image with natural scene, second column shows words segmentation done by the proposed method and third column shows character segmentation for the words in second column based on the proposed method. From Figure 1, it is observed that the proposed method works well for different fonts, font size, background and different orientation.

Input Text Line Words Segmentation Character Segmentation

Fig. 1: Sample results of the word and character segmentation

3.2. Adaptive Histogram Analysis based Method (AHA) for Text Binarization

We observe from pixel values in text lines (i.e., detected by the text detection method [24]) that the values of text pixel in each character component have uniform colors compared to whole word. Based on this observation, we propose an Adaptive Histogram Analysis based method (AHM). For each text line image, we compute Height (H) of the text block and it is considered as height of the window. This is an advantage for text line segmentation from the natural scene images that the closed bounding box for text lines is fixed by text detection methods. With this bounding box, we can easily compute height of text information. The same length of the height is considered as width of window to form a squared window. This is a simple and effective way to decide the size of the window compared to the methods [2, 3] that assume window size as a constant value or estimated from stoke width. This assumption led to major drawback of the existing methods that creates poor accuracy for recognizing text line in natural scene images. This is the major contribution of the proposed method.

Then, we move the squared window as a sliding window, pixel by pixel over a text line image. For each sliding window, we plot a histogram to choose highest peak by considering intensity values in X axis and number of pixels in Y axis. We display all pixels in the highest peak as white pixels in the separate image. This process continues until the end of the text lines. As a result, a binary image is generated for the detected text line. The sample input text line image, square window based on height of text and generated binary image are shown in Figure 2(a), (b) and (c) respectively. From Figure 2(c), it is observed that the proposed AHM preserves the shape of the characters that is required to get good recognition rate by the existing OCR after the text binarization process.

In order to show effectiveness of the proposed AHM, we compare the proposed AHM with Otsu that is a well-known global thresholding technique for text binarization. By applying Otsu method on a whole image, we get a result with poor accuracy. Figure 3 shows the input images, text binarization results based on Otsu method and recognition results within double quote. It can be seen from recognition results that Otsu method implementation on whole image does not give good results.

Input Text Line sliding window highest peak

(a). (b) (c)

Fig. 2: Illustration for sliding window

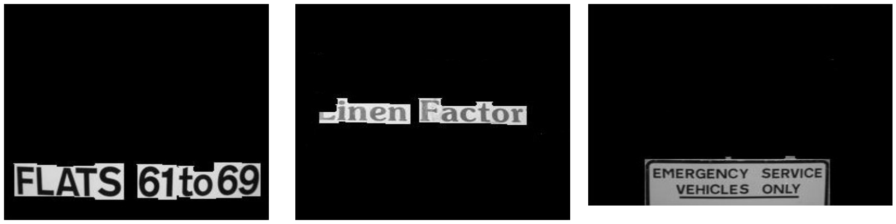
(a) Input full images

"[FLATS 61t069," "1' :actory" "LJHIIII/1 T}???. ;..~ 'g5"
1 \.m{J[L13 "MERHNCV
SERVICEVEHICLES ONLY"

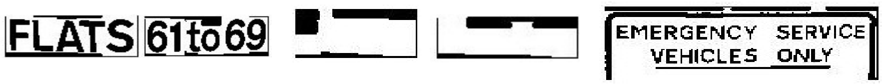
(b) Otsu thresholding

Fig. 3: Illustration of Otsu thresholding on full images

We have tested Otsu method on whole text line image and sliding window over text line for comparison in terms of effectiveness. Figure 4 illustrates effectiveness of both the proposed and Otsu thresholding method. For the input text lines images (i.e., as shown in Figure 4(a)), we apply Otsu method on whole text line images to obtain binary image (i.e., as shown in Figure 4(b)). It is observed that Otsu method does not work well for whole text line image. Then, we apply Otsu method on our proposed sliding window to obtain binary images (i.e., as shown in Figure 4(c)). Results show that Otsu method does not work well for text line image with our proposed sliding window. On the other hand, the proposed AHM gives better results for input text line images (i.e., as shown in Figure 4(d)) where the text binarization results and recognition results are better than Otsu methods. The main reason forgetting a poor accuracy by Otsu method is that Otsu method only works well when the image gives clear distinction between text and non-text information. This assumption is good for plain background images but not for the images taken from natural scene images where the background variation is unpredictable. This work shows that the proposed AHM is more effective than Otsu method.

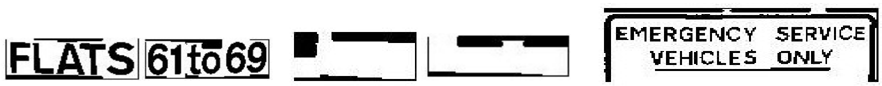


(a) Input text line images



"FES\§; " 5i._9" "" " " "EMERGENCY SERWCE
VEHICLES ONLY "

(b) Otsu method



"FES\§" " ;5i._9 " " " " " "EMERGENCY SERWCE
VEHICLES ONLY "

(c) Adaptive Otsu method

"FLATS "61 to 69" "_inen" "Factor" "EMERGENCY SERVICE
VEHICLES ONLY"

(d) Adaptive Histogram Analysis based Method (AHM)

Fig. 4: Comparison with Otsu and the Proposed Method

4.0 EXPERIMENTAL RESULTS

We consider different datasets to show that the proposed method is capable for handling different situations and diversified datasets. The proposed method is tested on 312 High Resolution Camera Images (HCI), 230 Low Resolution Mobile Camera Images (LMI), and the 210 standard dataset ICDAR-2003 competition data [25] to evaluate the performance of the proposed method. In total, the proposed method is tested on 752 images to show that the proposed method is superior to existing methods. For all three datasets, we test Otsu method on whole text lines image, adaptive Otsu method with sliding window, and the proposed Adaptive Histogram Analysis based Method (namely, Otsu, AOM and AHM respectively) to study the effectiveness of the methods. In addition, Otsu and AOM are considered as existing methods for comparative study in this work since Otsu is well-known method for document binarization. Further, character recognition rate is considered as a measurement to validate the text binarization results given by both existing and the proposed method. The results given by the text binarization methods are sent to both ABBYY [26] and Tesseract OCR [27] to obtain recognition results. The recognition results are shown in double quotation in all Figures. We conduct experiments on words and characters to test the character recognition rate by both ABBYY and Tesseract OCR. In addition, we evaluate the proposed and existing method in terms of average processing time (APT) that is defined as total processing time of the recognition of characters divided by number of text lines.

4.1. Experiments on High Resolution Camera Images (HCI)

Figure 5 shows sample results of the proposed and existing methods for the HCI data. In general, the proposed AHM gives better results compared to Otsu and AOM. In Table 1, based on the reported quantitative results by ABBYY and Tesseract OCR, the character recognition rate of the proposed AHM is better than Otsu and AOM at text line, word and character level. Since ABBYY OCR is an advanced and improved version of Tesseract, it gives better results than Tesseract OCR for all the cases in our experiment. However, in Table 1, it is observed that character recognition rate at character is lower than word and line level in contrast to our discussion on word and character segmentation. This is because when we apply sliding window operation on segmented character, the methods fail to select global parameter for Otsu method and highest peak for the proposed method correctly. The background complexity of segmented characters is lower than background complexity of words. This is not true for the words as we can see higher character recognition rate at words level than text line level by both the OCRs. In addition, the built-in language model in OCR helps in recognizing word even word misses' character information. Therefore, the proposed and existing methods give better recognition rate for words but not text lines and characters. Table 1 also show experiments of Otsu method on whole image, where it gives worst accuracy than other methods including the proposed method. Overall, we can infer that the proposed method is good for scene text recognition at word level compared to text line and character levels. It is noticed from Table 1 that the average processing time of ABBAY OCR requires more time compared to the Tesseract OCR. This is because ABBAY OCR involves several complex steps compared to Tesseract OCR.

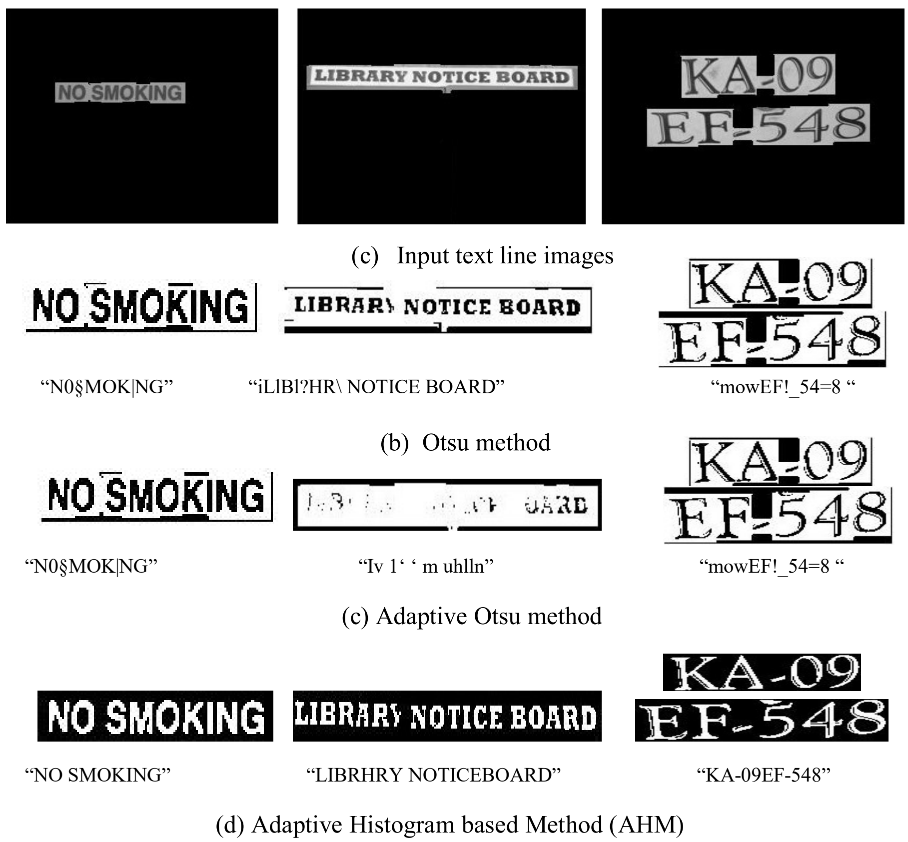


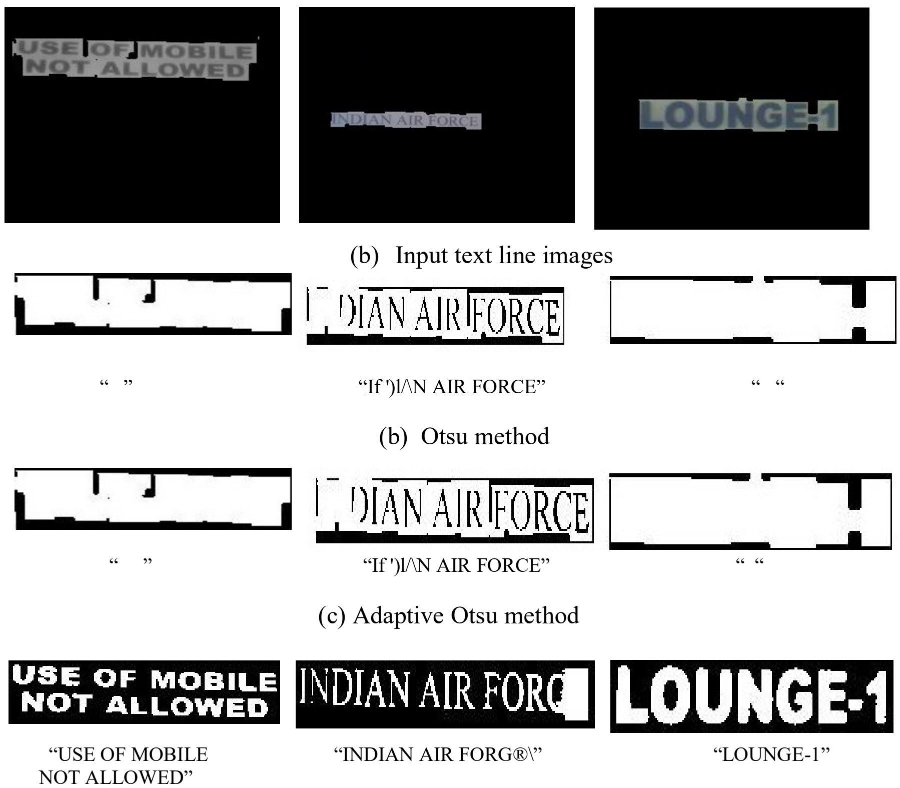
Fig. 5: Sample results of the proposed and existing methods for HCI

Table 1: Quantitative results of the proposed and existing methods for HCI (in %)

Methods	Character Recognition Rate (CRR)									
	ABBYY OCR					Tesseract OCR (Google)				
	Image	Text Line	Word	Character	APT(s)	Image	Text Line	Word	Character	APT(s)
AHM	-	89.32	91.15	88.00	1.9	-	85.43	86.34	84.54	1.89
AOM	-	42.80	39.35	39.02	1.91	-	38.52	39.00	37.06	1.90
Otsu	42.40	45.69	44.64	43.55	1.95	37.67	40.22	42.14	40.55	1.93

4.2. Experiments on Low Resolution Mobile Camera Images (LMI)

This experiment shows the proposed method works well for both low-and high-resolution text images. Sample results of the proposed and existing methods are shown in Figure 6. Based on the results, the proposed method gives better results than the existing methods. The quantitative results of the proposed method and existing methods at image, line, word and character level given by both ABBYY and Tesseract OCR are shown in Table 2. It is noticed that the proposed method at word gives better results compared to existing methods in terms of character recognition rate at all levels. The reason for poor accuracy of the existing methods is same as discussed in previous section4.1.



(d) Adaptive Histogram based Method (AHM)

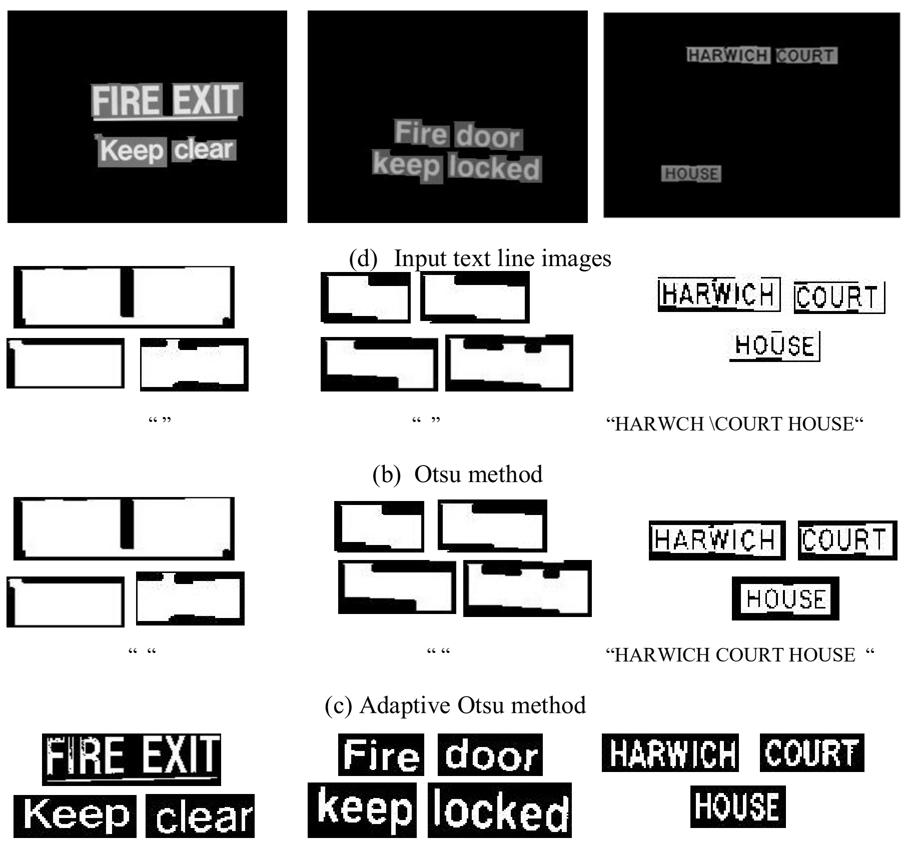
Fig. 6: Sample results of the proposed and existing methods for LMI

Table 2: Quantitative results of the proposed and existing methods for LMI (in %)

Methods	Character Recognition Rate (CRR)									
	ABBYY OCR					Tesseract OCR (Google)				
	Image	Text Line	Word	Character	APT(s)	Image	Text Line	Word	Character	APT(s)
AHM	-	86.20	87.23	85.92	2.00	-	82.49	83.72	81.66	2.00
AOM	-	43.00	47.16	43.84	2.25	-	39.19	37.04	38.43	2.11
Otsu	36.84	42,59	46.34	42.61	2.40	28.64	38.29	36.51	37.61	2.39

4.3. Experiments on ICDAR 2003 Data

ICDAR 2003dataset is a publicly available benchmark data for scene text detection. This dataset consists of complex background, non-uniform illumination and unfavorable characteristics of scene text. Our method is tested on this dataset and the results show that the proposed method is suitable for this dataset. Sample results of the proposed and existing methods are shown in Figure 7. Based on the results, the proposed method gives better results than the existing methods. The quantitative results of the proposed method and existing methods at image, line, and word and character level given by both ABBYY and Tesseract OCR are shown in Table 3. It is noticed that the proposed method at word level gives better results compared to existing methods in terms of character recognition rate at all levels. The reason for poor accuracy of the existing methods is same as discussed in previous Section 4.1.



(d) Input text line images

" " " " "HARWCH \COURT HOUSE"

(b) Otsu method

" " " " "HARWICH COURT HOUSE "

(c) Adaptive Otsu method

"FIRE EXIT Keep clear" "Fire door keep locked" "HARWICH COURT HOUSE"

(d) Adaptive Histogram based Method (AHM)

Fig. 7: Sample results of the proposed and existing methods for ICDAR 2003 data

From Table 1 to Table 3, we can assert that the character recognition rate improves when using segmented word as an input because the segmented word background complexity is reduced when compared with the text line image background complexity. The results are not same when using segmented word as an input for recognition and segmented character as an input for recognition. This is because the language model in

OCR works well for the words but not for text lines and characters. Text line image has complex background compared to the background of words and language model requires more than two characters to find semantics. The processing time of ABBAY OCR is higher than Tesseract OCR for entire Table 1 to 3. This is because ABBAY OCR involves complex steps for improving recognition results while Tesseract OCR does not. Therefore, in our experiments, character recognition rate at word level is better than text line and character levels.

Table 3: Quantitative results of the proposed and existing methods for ICDAR 2003 data (in %)

Methods		Character Recognition Rate (CRR)								
	ABBYY OCR					Tesseract OCR (Google)				
	Image	Text Line	Word	Character	APT(s)	Image	Text Line	Word	Character	APT(s)
AHM	-	89.06	90.08	88.62	2.13	-	82.36	85.32	81.11	2.19
AOM	-	39.54	40.42	37.39	2.29	-	32.54	38.72	34.06	2.33
Otsu	43.08	42.32	43.71	40.05	2.32	36.81	37.32	38.01	35.54	2.40

5.0 CONCLUSIONS AND FUTURE WORK

This paper presents new text binarization method for scene text recognition. This method explores color information of the character components where it is observed that color of pixel in each component have same values. With this intuition, we propose adaptive histogram analysis based method to choose uniform color values by performing sliding window operation over text line, words and characters. This simple idea works better than well-known Otsu method and adaptive Otsu method. We have modified the boundary growing method to segment words and characters from the text line images based on number of iterations while growing. This method also works on multi-oriented text lines. Experimental results show that character recognition rate at word level improves over character rate at text line level but not at character level. We plan to extend this method for multi-oriented text lines and curved text lines in future.

ACKNOWLEDGEMENT

This research is supported in part by BKP grant UM. TNC2/IPPP/UPGP/261/15 (BK010-2013). We thank Prof. D. Krishne gowda, Principal, Dr. H.R. Sreepad, Associate Professor and Prof. I. Priya Uthaiah, Assistant Professor, Govt. College (Autonomous), Mandya, Karnataka for their support to improve the quality of the work.

REFERENCES

[1] D. Chen and J. M. Odobez, "Video text recognition using sequential Monte Carlo and error voting methods", *Pattern Recognition Letters*, 2005, pp. 1386-1403.

[2] K. Wang, B. Babenko and S. Belongie. "End-to-end scene text recognition", *International Conference on Computer Vision,* 2011.

[3] A. Mishra, K. Alahari and C. V. Jawahar, "Top-Down and Bottom-Up Cues for Scene Text Recognition", *In Proc. International Conference on Computer Vision and Pattern Recognition,* 2012.

[4] Z. Saidane and C. Garcia, "Robust Binarization for Video Text Recognition", *In Proc. International Conference on Document Analysis and Recognition*, 2007, pp. 874-879.

[5] Z. Zhou, L. Li and C. L. Tan, "Edge based Binarization of Video Text Images", *In Proc. International Conference on Pattern Recognition*, 2010, pp. 133-136.

[6] S. Roy, P. Shivakumara, P.P. Roy and C.L. Tan, "Wavelet-gradient-fusion for video text binarization", *In Proc. International Conference on Patten Recognition*, 2012, pp. 3300-3303.

[7] N. Otsu," A threshold selection method from gray-level histograms", *IEEE Transactions on Systems, Man and Cybernetics* 9, 1979, pp. 62–66.

[8] X. Ye, M. Cheriet and C. Y Suen, "Stroke-model-based character extraction from gray-level document images", *IEEE Transactions on Image* Processing, Vol. 10 No. 8, 2001, pp. 1152–1161.

[9] R. F. Moghaddam and M. Cheriet, "A multi-scale framework for adaptive binarization of degraded document images", *Pattern* Recognition, Vol. 43 No. 6, 2010, pp. 2186–2198.

[10] J. Sauvolaand M. Pietikainen, "Adaptive document image binarization", *Pattern* Recognition, Vol. 33, No. 2, 2000, pp. 225–236.

[11] W. Niblack, "An Introduction to Digital Image Processing", *Strandberg Publishing Company*, Birkeroed, Denmark, Denmark, 1985.

[12] O. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods", *IEEE Transactions on Pattern Analysis and Machine* Intelligence, Vol. 17 No. 12, 1995, pp. 1191–1201.

[13] M. A. Shayegan, S. Aghabozorgi, and R. G. Raj, "A Novel Two-Stage Spectrum-Based Approach for Dimensionality Reduction: A Case Study on the Recognition of Handwritten Numerals," Journal of Applied Mathematics, vol. 2014, Article ID 654787, 14 pages, 2014. doi:10.1155/2014/654787.

[14] B. Su, S. Lu and C. L. Tan, "Binarization of historical document images using the local maximum and minimum", *In Proc. Document Analysis System,*2010, pp. 159–166.

[15] B. Su, S. Lu and C. L. Tan, "A self-training learning document binarization frame- work", *In Proc. International Conference on Pattern Recognition,*2010, pp. 3187–3190.

[16] B. Gatos, K. Ntirogiannis and I. Pratikakis, "DIBCO 2009 : Document Image Binarization Contest", *International Journal on Document Analysis and Recognition*, 2010, pp. 1–10.

[17] J. Fabrizio, B. Marcoteguiand and M. Cord, "Text segmentation in natural scenes using toggle-mapping", *In Proc. International Conference on Image Processing,*2009, pp. 2373–2376.

[18] D. Rivest-Henault, R. F. Moghaddam and M. Cheriet, "A local linear level set method for the binarization of degraded historical document images", *International journal on Document Analysis and Recognition*, 2011.

[19] R. F. Moghaddam and M. Cheriet, "RSLDI: restoration of single-sided low-quality document images", *Pattern Recognition,* Vol. 42 No. 12, 2009, pp. 3355–3364.

[20] H. S. Don, "A noise attribute thresholding method for document image binarization", *International Journal on Document Analysis and Recognition,* Vol. 4 No. 2, 2001, pp. 131–138.

[21] J. Bernsen, "Dynamic thresholding of grey-level image", *In Proc. International Conference on Pattern Recognition*, 1986,pp. 1251-1255

[22] R. F. Moghaddamand and M. Cheriet, "AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization", *Pattern Recognition*, Vol. 45, 2012, pp. 2419–2431.

[23] L. Neumann and J. Matas, "A Method for Text Localization and Recognition in Real-World Images", *In Proc. Asian Conference on Computer Vision*, 2011, pp. 770-783.

[24] M. Basavanna. P. Shivakumara, S. K. Srivatsa and G. Hemantha Kumar, "Multi-Oriented Text Detection in Scene Images", *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 26, No. 7, 2012, pp 1-19.

[25] S. M. Lucas, "ICDAR 2003 Robust Reading Competitions", *In Proc. International Conference on Document Analysis and Recognition*, 2003.

[26] ABBYY Fine Reader 9.0,http://www.abbyy.com/

[27] Tesseract OCR, http://code.google.com/p/tesseract-ocr