International Journal of Computer Science Trends and Technology (IJCST) – Volume 2 Issue 4, Jul-Aug 2014

RESEARCH ARTICLE OPEN ACCESS

Adaptive Contrast Binarization Using Standard Deviation for Degraded Document Images

Jeena Joy[1], Jincy Kuriakose[2]
Department of Computer Science and Engineering,
Govt. Engineering College Idukki,
Kerala-India

ABSTRACT

Document image binarization is of great significance in the document image analysis and recognition process because it affects further steps of the recognition process. The variation between the foreground text and the background of different document images is a challenging task. This paper presents a new document image binarization technique that focus on these issues using adaptive image contrast. The adaptive image contrast is a combination of the local image contrast and the local image gradient so as to tolerate the text and background variation caused by different types of document degradations. In the proposed image binarization technique, an adaptive contrast map is first constructed for an input degraded document image which is then adaptively binarized and combined with Sobel's edge detector to identify the text stroke edge pixels. The document text is further segmented using a local threshold.

Keywords:- Binarization, thresholding, adaptive image contrast, stroke width.

I. INTRODUCTION

Documents are a ubiquitous medium in our daily life. Importing documents into a computer calls for a mechanism of converting handwritten and printed characters into an electronic form. Characters are often captured optically, such as using scanners or cameras, to create images. Successfully extracting characters from an image background is a necessary first step before further analysis. This process is known as document binarization. Document image binarization is of great importance in the document image analysis and recognition pipeline since it affects further steps of the recognition process.

Document image binarization aims to divide a document image into two classes, namely, the foreground text and the document background. It is usually performed in the document preprocessing stage and is very important for ensuring document image processing tasks such as document image retrieval and optical character recognition (OCR). It converts a gray-scale document image into a binary document image and accordingly facilitates the ensuing tasks such as document skew estimation and document layout analysis. As more and more text documents are scanned, fast and accurate document image binarization is becoming ever more important.

Though document image binarization has been studied for many years, the thresholding of degraded document images is still an unsolved problem. This can be justified by the fact that the modeling of the document foreground/background is very difficult due to various types of document degradation such as image contrast variation, uneven illumination, bleeding-through, and smear as illustrated in Fig. 1. We try to develop robust and efficient document image binarization techniques which are able to produce good results for badly degraded document images.

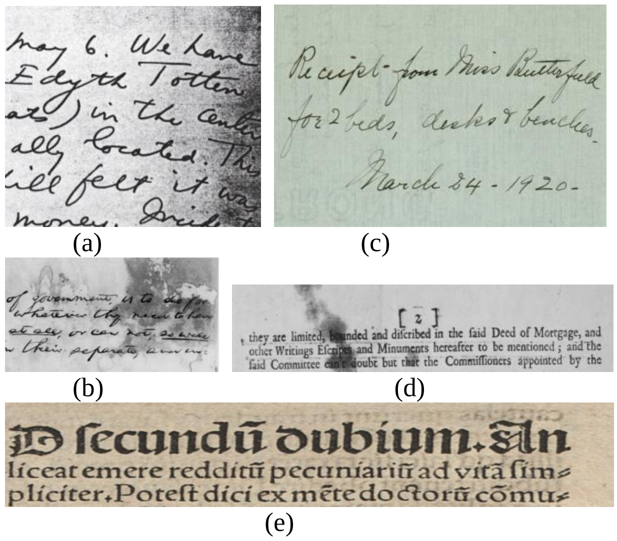


Fig.1. Five degraded document image examples (a)–(e) are taken from DIBCO series datasets

The proposed method is simple and able to handle different types of degraded document images with minimum parameter tuning. In the proposed image binarization technique, an adaptive contrast map is first constructed for an input degraded document image which is then adaptively binarized and combined with Sobel's edge detector to identify the text stroke edge pixels. The stoke width is taken as a constant value. After that the document text is segmented by a local threshold.

II. RELATED WORK

Though document image binarization has been studied for many years, the thresholding of degraded document images is still an unsolved problem due to the high inter/intra-variation between the text stroke and the document background of the different document images. The handwritten text within the degraded documents often shows a certain amount of variation in terms of the stroke brightness, stroke connection, stroke width, and document background. In addition, historical documents are usually degraded by the bleed through where the ink of the other side seeps through to the front. These different types of document degradations tend to generate the document thresholding error and make degraded document image binarization a big challenge to most state-of-the-art techniques.

Document image binarization plays a key role in document processing since its performance affects quite critically the degree of success in subsequent character segmentation and recognition. In general, approaches that deal with document image binarization are categorized in two main classes: (i) global and (ii) local. In a global approach, threshold selection results in a single threshold value for the entire image. Global thresholding [2, 10, 15] has a good performance in the case that there is a good separation between the foreground and the background. However, in the case of degradations (e.g. shadows and non-uniform illumination, ink seeping) the current trend is to use local information that guides the threshold value pixel wise in an adaptive manner. Adaptive thresholding [1, 6, 9, 16, 17, 14, 19], which estimates a local threshold for each document image pixel, is often a better approach to deal with different variations within degraded document images.

For example, the early window based adaptive thresholding techniques [17, 14] estimate the local threshold by using the mean and the standard variation of image pixels within a local neighbourhood window. The main drawback of this window-based thresholding approach is that the thresholding performance depends heavily on the window size and hence the character stroke width. Other approaches have also been reported to binarize historical document images through background subtraction [13, 8], texture analysis [12], decomposition method [4], morphological operations [11], shape based local thresholding [7], classification framework [20], and cross section sequence graph analysis [5], and so on. These approaches combine different types of image information and domain knowledge and are often complex and time consuming.

The local image contrast and the local image gradient are very useful features for segmenting the text from the document background because the document text usually has certain image contrast to the neighbouring document background. They are very effective and have been used in many document image binarization techniques [3], [1], [17], [14]. In Bernsen's paper [1], the local contrast is defined as follows:

$$C(i,j) = f_{\max}(i,j) - f_{\min}(i,j) \qquad (1)$$

where $C(i, j)$ denotes the contrast of an image pixel (i, j), $f_{max}(i, j)$ and $f_{min}(i, j)$ denote the maximum and minimum intensities within a local neighbourhood windows of (i, j), respectively. If the local contrast $C(i, j)$ is smaller than a threshold, the pixel is set as background directly. Otherwise it will be classified into text or background by comparing with the mean of $f_{max}(i, j)$ and $f_{min}(i, j)$. Bernsen's method is simple, but cannot work properly on degraded document images with a complex document background. There is an earlier document image binarization method [3] by using the local image contrast that is evaluated as follows:

$$C(i,j) = \frac{f_{\max}(i,j) - f_{\min}(i,j)}{f_{\max}(i,j) + f_{\min}(i,j) + \varepsilon} \qquad (2)$$

where ε is a positive but infinitely small number that is added in case the local maximum is equal to 0.

Compared with Bernsen's contrast in Equation 1, the local image contrast in Equation 2 introduces a normalization factor (the denominator) to compensate the image variation within the document background. However, the image contrast in Equation 2 has one typical limitation that it may not handle document images with the bright text properly. This is because a weak contrast will be calculated for stroke edges of the bright text where the denominator in Equation 2 will be large but the numerator will be small. To overcome this over-normalization problem, the local image contrast is combined with the local image gradient and derives an adaptive local image contrast [18] as follows:

$$C_a(i,j) = \alpha C(i.j) + (1-\alpha)(f_{\max}(i,j) - f_{\min}(i,j)) \quad (3)$$

where $C(i,j)$ denotes the local contrast and $(f_{\max}(i,j) - f_{\min}(i,j))$ refers to the local image gradient that is normalized to [0, 1]. The local windows size is set to 3 empirically and α is the weight between local contrast and local gradient that is controlled based on the document image statistical information.

III. PROPOSED METHOD

This section describes the proposed document image binarization techniques. Given a degraded document image, an adaptive contrast map is first constructed and the text stroke edges are then detected through the combination of the binarized adaptive contrast map and the sobel's edge map. The text is then segmented based on the local threshold that is estimated from the text stroke edge pixels.

A. Contrast Image Construction

In the proposed method the input image is subjected to an adaptive filtering for smoothing the image. The average filter is used for this purpose. The filtered image is given to a wiener filter for adaptive noise removal. The degraded document images may contain Gaussian noise. This can be effectively removed by the wiener filter. Then the contrast *C(i,j)* is calculated using the Equation 4.

$$C(i,j) = C_a * (0.01)^S \tag{4}$$

where C_a is the contrast calculated using the equation 3. In equation 3, α is the weight between local contrast and local gradient found as described in [18].

$$\alpha = \left(\frac{std}{128}\right)^{\gamma} \tag{5}$$

In the proposed method we set the value of γ in the equation of α as 1. The *S* in contrast equation is used to denote the matrix by rearranging each 3-by-3 block of the contrast image C_a into a column of a temporary matrix, and then applying the standard deviation function to this matrix . The output of this step is used to detect the text stroke edge pixels in next step.

Fig.2. Contrast image map

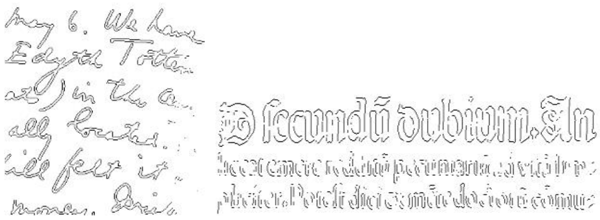
B. Text Stroke Edge Pixel Detection

The constructed contrast image has a bi-modal pattern, where the adaptive image contrast computed at text stroke edges is obviously larger than that computed within the document background. To detect the text stroke edge pixels we compute a value *T* which is used to binarize the contrast image. The mean and standard deviation of *S* is calculated and added together to get the value of *T*. Here we use an adaptive thresholding using the value of *T* which helps to extract the text stroke edge pixels properly. The binary maps are shown in Fig. 3(a).

The binary map can be further improved through the combination with the edges by Sobel's edge detector, which helps to find the edges that are close to real edge locations in the detecting image. The Sobel's edge detector help to extract a large amount of non-stroke edges as illustrated in Fig. 3(b) without tuning the parameter manually.

(a)

(b)



(c)

Fig. 3: (a) Binary contrast maps, (b) sobel edge maps, and their (c) combined edge maps of the sample document images in Fig. 1(a) and (e), respectively.

In the combined map, we keep only pixels that appear within both the high contrast image pixel map and sobel edge map. The combination helps to extract the text stroke edge pixels accurately as shown in Fig. 3(c).

C. Local Threshold Estimation

The text can then be extracted from the document background pixels once the high contrast stroke edge pixels are detected properly. Document images possess the following characteristics: First, the text pixels are close to the detected text stroke edge pixels. Second, there is a distinct intensity difference between the high contrast stroke edge pixels and the surrounding background pixels. The document image text can thus be extracted based on the detected text stroke edge pixels as follows:

$$R(x,y) = \begin{cases} 1 & f(x,y) \le E_{mean} + \dfrac{E_{std}}{5} \\ 0 & Otherwise \end{cases} \tag{6}$$

where *E*mean and *E*std are the mean and standard deviation of the intensity of the detected text stroke edge pixels within a neighbourhood window *W*, respectively. In this method the window size is taken as a constant and we used *W=6* in this method.

Fig. 4: Binarized result of the sample document images in Fig.1(a) and (e).

IV. EXPERIMENTAL MEASURES AND RESULTS

A. *Experimental Measures*

There exist several experimental measures to evaluate and compare the binarization methods.

1) *F-Measure*

F-measure (FM) is the harmonic mean of precision and recall. F-measure is calculated at the pixel level.

$$F-Measure = \frac{2 \times \mathrm{Re}call \times \mathrm{Pr}ecision}{\mathrm{Re}call + \mathrm{Pr}ecision} \qquad (7)$$

where *Recall =TP/TP+FN*, *Precision =TP/TP+FP*, TP, FP, FN denote the True positive, False positive and False Negative values, respectively.

2) *PSNR*

The peak signal to noise ratio (PSNR) is defined as,

$$PSNR = 10\log(\frac{C^2}{MSE}) \qquad (8)$$

Here, C is a constant that denotes the difference between foreground and background. This constant is set to 1. MSE is the mean square error.

3) *Negative Rate Metric (NRM)*

The negative rate metric NRM is based on the pixel-wise mismatches between the GT and prediction. It combines the false negative rate NR_{FN} and the false positive rate NR_{FP}. It is denoted as follows:

$$NRM = \frac{NR_{FN} + NR_{FP}}{2} \qquad (9)$$

where $NR_{FN} = \frac{N_{FN}}{N_{FN} + N_{TP}}$, $NR_{FP} = \frac{N_{FP}}{N_{Fp} + N_{TN}}$. N_{TP} denotes the number of true positives, N_{FP} denotes the number of false positives, N_{TN} denotes the number of true negatives, N_{FN} denotes the number of false negatives. In contrast to F-Measure and PSNR, the binarization quality is better for lower NRM.

4) *Distance Reciprocal Distortion Metric (DRD)*

The Distance Reciprocal Distortion Metric (DRD) has been used before to measure the visual distortion in binary document images [21]. It properly correlates with the human visual perception and it measures the distortion for all the *S* flipped pixels as follows:

$$DRD = \frac{\sum_{k=1}^{S} DRD_k}{NUBN} \qquad (10)$$

where *DRDk* is the distortion of the k-th flipped pixel and it is calculated using a 5x5 normalized weight matrix *WNm* as defined in [21]. *DRDk* equals to the weighted sum of the pixels in the 5x5 block of the Ground Truth *GT* that differ from the centered kth flipped pixel at (x,y) in the binarization result image *B*.

$$DRD_k = \sum_{i=-2}^{2}\sum_{j=-2}^{2} |GT_k(i,j) - B_k(x,y)| \times W_{Nm}(i,j) \qquad (11)$$

Finally, NUBN is the number of the non-uniform (not all black or white pixels) 8x8 blocks in the GT image.

5) *Misclassification penalty metric (MPM)*

$$MPM = \frac{MP_{FN} + MP_{FP}}{2} \qquad (12)$$

where $MP_{FN} = \frac{\sum_{i=1}^{N_{FN}} d_{FN}^{i}}{D}$, $MP_{FP} = \frac{\sum_{j=1}^{N_{FP}} d_{FP}^{j}}{D}$. d^i_{FN} and d^j_{FP} denote the distance of the i^{th} false negative and the j^{th} false positive pixel from the contour of the GT. D is the sum over all the pixel-to-contour distances of the GT object.

B. *Experimental Results*

We compared our proposed method with other state-of-the-art techniques on DIBCO 2009[22], H-DIBCO 2010 [23] and DIBCO 2011 [24] datasets. The methods used for comparison are Sauvola's method (SAU) [17] and the adaptive contrast binarization method (ADA) [18]. We have taken a total of 26 images from the datasets given above. Our datasets consist of four degraded handwritten documents and five degraded printed documents from DIBCO 2009 dataset, nine degraded handwritten documents from H-DIBCO 2010 dataset and eight degraded handwritten documents from DIBCO 2011 dataset. The evaluation results are shown in table.

TABLE I
EXPERIMENTAL RESULTS

Method	FM	NRM	PSNR	DRD	MPM ($x10^3$)
SAU	61.47	0.1105	11.17	21.34	16.68
ADA	51.25	0.3070	13.44	6.66	1.31
Our Method	70.86	0.1500	13.86	5.37	2.33

V. CONCLUSIONS

This paper presents document image binarization technique which is based on an adaptive image contrast using the standard deviation that is tolerant to different types of document degradation. Experimental results on the dataset show that the proposed method provides better performance than existing methods. The proposed technique is simple and executes in less time, only few parameters are involved. Moreover, it works for different kinds of degraded document images. The method is compared with two other method and the results are also given.

REFERENCES

[1] J. Bernsen. Dynamic thresholding of gray-level images. International Conference on Pattern Recognition, pages 1251–1255, October 1986.

[2] A. Brink. Thresholding of digital images using two-dimensional entropies. Pattern Recognition, 25(8):803–808, 1992.

[3] B. Su, S. Lu, and C. L. Tan, “Binarization of historical handwritten document images using local maximum and minimum filter,” in Proc. Int. Workshop Document Anal. Syst., Jun. 2010, pp. 159–166.

[4] Y. Chen and G. Leedham. Decompose algorithm for thresholding degraded historical document images. IEEE Proceedings on Vision, Image and Signal Processing, 152(6):702–714, December 2005.

[5] A. Dawoud. Iterative cross section sequence graph for handwritten character segmentation. IEEE Transactions on Image Processing, 16(8):2150–2154, August 2007.

[6] L. Eikvil, T. Taxt, and K. Moen. A fast adaptive method for binarization of document images. International Conference on Document Analysis and Recognition, pages 435–443, September 1991.

[7] Jichuan Shi, Nilanjan Ray, Hong Zhang, “Shape Based Local Thresholding for Binarization of Document Images”, Elsevier, Aug., 2011.

[8] B. Gatos, I. Pratikakis, and S. Perantonis. Adaptive degraded document image binarization. Pattern Recognition, 39(3):317–327, 2006.

[9] I.-K. Kim, D.-W. Jung, and R.-H. Park. Document image binarization based on topographic analysis using a water flow model. Pattern Recognition, 35(1):265–277, 2002.

[10] J. Kittler and J. Illingworth. On threshold selection using clustering criteria. IEEE transactions on systems, man, and cybernetics, 15:652–655, 1985.

[11] Akihiro Okamoto, Hiromi Yoshida and Naoki Tanaka, : “A Binarization Method for Degraded Document Images with Morphological Operations” IAPR International Conference on Machine Vision Applications, May. 2013.

[12] Y. Liu and S. Srihari. Document image binarization based on texture features. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(5):540–544, May 1997.

[13] S. J. Lu and C. L. Tan. Binarization of badly illuminated document images through shading estimation and compensation. International Conference on Document Analysis and Recognition, pages 312–316, 2007.

[14] W. Niblack. An Introduction to Digital Image Processing. Prentice-Hall, Englewood Cliffs, New Jersey, 1986.

[15] N. Otsu. A threshold selection method from gray level histogram. IEEE Transactions on System, Man, Cybernetics, 19(1):62–66, January 1978.

[16] J. Parker, C. Jennings, and A. Salkauskas. Thresholding using an illumination model. International Conference on Document Analysis and Recognition, pages 270–273, October 1993.

[17] J. Sauvola and M. Pietikainen. Adaptive document image binarization. Pattern Recognition, 33(2):225–236, 2000.

[18] Bolan Su, Shijian Lu and Chew Lim Tan,” Robust Document Image Binarization Technique for Degraded Document Images”, IEEE transactions on image processing, vol. 22, no. 4, april,2013.

[19] J.-D. Yang, Y.-S. Chen, and W.-H. Hsu. Adaptive thresholding algorithm and its hardware implementation. Pattern Recognition Letters, 15(2):141–150, 1994.

[20] Bolan Su, Shijian Lu and Chew Lim Tan,” Combination of Document Image Binarization Techniques”.

[21] H. Lu, A. C. Kot and Y.Q. Shi, “Distance-Reciprocal Distortion Measure for Binary Document Images”, IEEE Signal Processing Letters, vol. 11, No. 2, pp. 228-231, 2004.

[22] B. Gatos, K. Ntirogiannis, and I. Pratikakis, “ICDAR 2009 document image binarization contest (DIBCO 2009),” in Proc. Int. Conf. Document Anal. Recognit., Jul. 2009, pp. 1375–1382.

[23] I. Pratikakis, B. Gatos, and K. Ntirogiannis, “ICDAR 2011 document image binarization contest (DIBCO 2011),” in Proc. Int. Conf. Document Anal. Recognit., Sep. 2011, pp. 1506–1510.

[24] I. Pratikakis, B. Gatos, and K. Ntirogiannis, “H-DIBCO 2010 handwritten document image binarization competition,” in Proc. Int. Conf. Frontiers Handwrit. Recognit., Nov. 2010, pp. 727–732.