DOCUMENT IMAGE BINARIZATION BASED ON ADAPTIVE CONTRAST METHOD FOR DEGRADED DOCUMENT IMAGES

[1] Arathy U.P., [2] Arun P.S.

[1,2]*Computer Science & Engineering, Kerala University, (India)*

ABSTRACT

Segmentation of text from badly degraded document images is a very challenging task due to the high inter/intra variation between the document background and the foreground text of different document images. Here, we propose a novel document image binarization technique that addresses these issues by using adaptive image contrast. The adaptive image contrast is a combination of the local image contrast and the local image gradient that is tolerant to text and background variation caused by different types of document degradations.In the proposed technique, an adaptive contrast map is first constructed for an input degraded document image. The contrast map is then binarized and combined with Canny's edge map to identify the text stroke edge pixels. The document text is further segmented by a local threshold that is estimated based on the intensities of detected text stroke edge pixels within a local window. The proposed method is simple, robust, and involves minimum parameter tuning

***Keywords*: Adaptive image contrast, document analysis, document image processing, degraded document image binarization,**

I. INTRODUCTION

In document image processing, the paper documents are initially scanned and stored in the hard disk or any other required location. It is easy to define document image processing as scanning-storing-retrieving-managing. The final outcome of document image processing will be in compatible electronic format, which makes documents easier and quicker to access. Document image processing comprises of a set of simple techniques and procedures, which are used to work upon the images of documents and convert them from pixel information into a format that can be read by a computer. Document image binarization is usually performed in the pre-processing stage of different document image processing related applications such as optical character recognition (OCR) and document image retrieval. Basically, document image processing using OCR is divided into two steps. The first step captures the text, based on the information from the document. It identifies the reorientations, tables, words and their colours, font sizes and other textual matter in the file. The second step involves graphical processing which works on drawings, dividing lines between paragraphs and sections, logos and other pictorial representations. As the images are one of the most important components in the documents, it is very important to process the images rather than just locating them in the document.

Document image binarization aims to segment the foreground text from the document background. But the segmentation of text from badly degraded document images is a very challenging task due to the high inter and intra variation between the document background and the foreground text of different document images. Document image binarization converts the acquired image to binary format, the objective of binarization is to automatically choose a threshold that separates the foreground and background information. Document image binarization has been studied for many years[1]-[3], the thresholding of degraded document images is still an unsolved problem due to the high inter and intra variation between the text stroke and the document background across different document images

Fig. 1. Five degraded document image examples (a)–(d) are taken from DIBCO series datasets and (e) is taken from Bickley diary dataset

II. RELATED WORKS

A new document image binarization technique that segments the text from badly degraded historical document images. Here makes use of the image contrast that is defined by the local image maximum and minimum. Compared with the image gradient, the image contrast evaluated by the local maximum and minimum has a nice property that it is more tolerant to the uneven illumination and other types of document degradation such as smear. Given a historical document image, the technique first constructs a contrast image and then detects the high contrast image pixels which usually lie around the text stroke boundary. The document text is then segmented by using local thresholds that are estimated from the detected high contrast pixels within a local neighbourhood window.They divide this section into three subsections, which deal with the contrast image construction, the high contrast pixel detection, and the local threshold estimation, respectively

2.1Contrast Image Construction

The image gradient has been widely used in the literature for edge detection. However, the image gradient is often obtained by the absolute image difference within a local neighbourhood window, which does not incorporate the image intensity itself and is so sensitive to the image contrast/brightness variation. Take an unevenly illuminated historical document image as an example, the gradient of an image pixel (around the text stroke boundary) within bright document regions may be much higher than that within dark document regions. To detect the high contrast image pixels around the text stroke boundary properly, the image gradient needs to be normalized to compensate for the effect of the image contrast or brightness variation. At the same time, the normalization suppresses the variation within the document background as well. In the proposed technique, we suppress the background variation by using an image contrast that is calculated based on the local image maximum and minimum as follows:

$$D(x,y) = \frac{f_{max}(x,y) - f_{min}(x,y)}{f_{max}(x,y) + f_{min}(x,y) + \epsilon}$$

where fmax(x, y) and fmin(x, y) refer to the maximum and the minimum image intensities within a local neighbourhood window. In the implemented system, the local neighbourhood window is a 3 × 3 square window. The term e is a positive but infinitely small number, which is added in case the local maximum is equal to 0.The image contrast in above equation, lowers the image background and brightness variation properly. In particular, the numerator (i.e. the difference between the local maximum and the local minimum) captures the local image difference that is similar to the traditional image gradient. The denominator acts as a normalization factor that lowers the effect of the image contrast and brightness variation. For image pixels within bright regions around the text stroke boundary, the denominator is large, which neutralizes the large numerator and accordingly results in a relatively low image contrast. But for image pixels within dark regions around the text stroke boundary, the denominator is small, which compensates the small numerator and accordingly results in a relatively high image contrast. As a result, the contrasts of image pixels (lying around the text stroke boundary) within both bright and dark document regions converge close to each other and this facilitates the detection of high contrast image pixels lying around the text stroke boundary(to be described in the next subsection).

2.2 High Contrast Pixel Detection

The purpose of the contrast image construction is to detect the desired high contrast image pixels lying around the text stroke boundary. As described in the last subsection, the constructed contrast image has a clear bimodal pattern where the image contrast around the text stroke boundary varies within a small range but is obviously much larger compared with the image contrast within the document background. We therefore detect the desired high contrast image pixels (lying around the text stroke boundary) by using Otsu's global thresholding method.

2.3 Historical Document Thresholding

The text pixels can be classified from the document background pixels once the high contrast image pixels around the text stroke boundary are detected properly. The document thresholding from the detected high contrast image pixels is based on two observations. First, the text pixels should be close to the detected high contrast image pixels because most detected high contrast image pixels lie around the text stroke boundary. Second, the intensity of most text pixels should be close or lower than the average intensity of the detected high contrast image pixels within a local neighbourhood window[4]-[6]. This can be similarly explained by the fact that most detected high contrast image pixels lie around the text stroke boundary.

For each document image pixel, the number of the detected high contrast image pixels is first determined within a local neighbourhood window. The document image pixel will be considered a text pixel candidate if the number of high contrast image pixels within the neighbourhood window is larger than a threshold. The document image pixel can thus be classified based on its intensity relative to that of its neighboring high contrast image pixels as

$$R(x,y) = \begin{cases} 1 & N_e \geq N_{min} \quad \&\& \\ & I(x,y) \leq E_{mean} + E_{std}/2 \\ 0 & \text{otherwise} \end{cases}$$

where Emean and Estd are the mean and the standard deviation of the image intensity of the detected high contrast image pixels(within the original document image) within the neighbourhood window that can be evaluated as

$$E_{mean} = \frac{\sum_{neighbor} I(x,y) * (1 - E(x,y))}{N_e}$$

$$E_{std} = \sqrt{\frac{\sum_{neighbor} ((I(x,y) - E_{mean}) * (1 - E(x,y)))^2}{2}}$$

where I refers to the input document image and (x, y) denotes the position of the document image pixel under study. E refers to the binary high contrast pixel image where E(x, y) is equal to 0 if the document image pixel is detected as a high contrast pixel. Ne refers to the number of high contrast image pixels that lie within the local neighbourhood window. So if Ne is larger than Nmin and I(i, j) is smaller than Emean+ Estd/2, R(i, j) is set at 1. Otherwise(i, j) is set at 0.

Disadvantages:

- It can't handle document images with bright text properly. This is because a weak contrast will be calculated for stroke edges of the bright text where the denominator in Equation(image contrast, is calculated based on the local image maximum and minimum)will be large but the numerator will be small.
- Time consuming.

III. PROPOSED METHOD

In 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 canny edge map. The text is then segmented based on the local threshold that is estimated from the detected text stroke edge pixels. Some post-processing is further applied to improve the document binarization quality

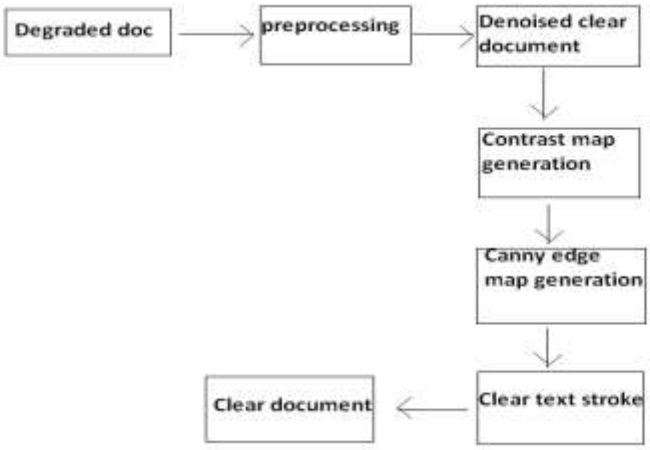

Fig2:System design of document image binarization technique

3.1 Contrast Image Construction

The image gradient has been widely used for edge detection and it can be used to detect the text stroke edges of the document images effectively that have a uniform document background. On the other hand, it often detects many non-stroke edges from the background of degraded document that often contains certain image variations due to noise, uneven lighting, bleed-through, etc. To extract only the stroke edges properly, the image gradient needs to be normalized to compensate the image variation within the document background.

In earlier method , The local contrast evaluated by the local image maximum and minimum is used to suppress the background variation as described in Equation (image contrast, is calculated based on the local image maximum and minimum). In particular, the numerator (i.e. the difference between the local maximum and the local minimum) captures the local image difference that is similar to the traditional image gradient. The denominator is a normalization factor that suppresses the image variation within the document background. For image pixels within bright regions, it will produce a large normalization factor to neutralize the numerator and accordingly result in a relatively low image contrast. For the image pixels within dark regions, it will produce a small denominator and accordingly result in a relatively high image contrast.However, the image contrast in the earlier method 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(image contrast, is calculated based on the local image maximum and minimum)will be large but the numerator will be small. To overcome this over-normalization problem,we combine the local image contrast with the local image gradient and derive an adaptive local image contrast as

$$C_a(i,j) = \alpha C(i,j) + (1-\alpha)(I_{\max}(i,j) - I_{\min}(i,j))$$

where $C(i, j)$ denotes the local contrast in the previous equation and$(I\max(i, j) - I\min(i, j))$ refers to the local image gradient that is normalized to [0, 1]. The local windows size is set to 3 empirically. α is the weight between local contrast and local gradient that is controlled based on the document image statistical information. Ideally, the image contrast will be assigned with a high weight (i.e. large α) when the document image has significant intensity variation. So that the proposed binarization technique depends more on the local image contrast that can capture the intensity variation well and hence produce good results. Otherwise, the local image gradient will be assigned with a high weight. The proposed binarization technique relies more on image gradient and avoid the over normalization problem of our previous method.We model the mapping from document image intensity variation to α by a power function as

$$\alpha = \left(\frac{Std}{128}\right)^{\gamma}.$$

where *Std* denotes the document image intensity standard deviation, and γ is a pre-defined parameter. The power function has a nice property in that it monotonically and smoothly increases from 0 to 1 and its shape can be easily controlled by different γ .γ can be selected from $[0,\infty]$, where the power function becomes a linear function when $\gamma = 1$. Therefore, the local image gradient will play the major role in $Ca(i, j)$when γ is large and the local image contrast will play the major role when γ is small.

3.2 Text Stroke Edge Pixel Detection

The purpose of the contrast image construction is to detect the stroke edge pixels of the document text properly. The constructed contrast image has a clear bi-modal pattern [7],where the adaptive image contrast computed at text stroke edges is obviously larger than that computed within the document background. We therefore detect the text stroke edge pixel candidate by using Otsu's global thresholding method.

As the local image contrast and the local image gradient are evaluated by the difference between the maximum and minimum intensity in a local window, the pixels at both sides of the text stroke will be selected as the high contrast pixels. The binary map can be further improved through the combination with the edges by Canny's edge detector, because Canny's edge detector [8]-[13] has a good localization property that it can mark the edges close to real edge locations in the detecting image. In addition, canny edge detector uses two adaptive thresholds and is more tolerant to different imaging artifacts such as shading.

3.3 Local Threshold Estimation

The text can then be extracted from the document background pixels once the high contrast stroke edge pixels are detected properly. Two characteristics can be observed from different kinds of document images: 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

$$R(x,y) = \begin{cases} 1 & I(x,y) \le E_{\text{mean}} + \frac{E_{\text{std}}}{2} \\ 0 & \text{otherwise} \end{cases}$$

where Emean and Estd are the mean and standard deviation of the intensity of the detected text stroke edge pixels within a neighbourhood window W, respectively. The neighbourhood window should be at least larger than the stroke width in order to contain stroke edge pixels.

3.4 Post-Processing

The binarization result can be further improved by incorporating certain domain knowledge as described in post processing algorithm. First, the isolated foreground pixels that do not connect with other foreground pixels are filtered out to make the edge pixel set precisely. Second, the neighbourhood pixel pair that lies on symmetric sides of a text stroke edge pixel should belong to different classes (i.e., either the document background or the foreground text). One pixel of the pixel pair is therefore labelled to the other category if both of the two pixels belong to the same class. Finally, some single-pixel artifacts along the text stroke boundaries are filtered out by using several logical operators.

Advantages:

The main advantages of the system are as follows:

- It relies more on image gradient and avoid the over normalization problem
- The proposed method simple, robust and capable of handling different types of degraded document images with minimum parameter tuning

IV. CONCLUSION

Segmentation of text from badly degraded document images is a very challenging task due to the high inter/intravariation between the document background and the foreground text of different document images. Here, propose a novel document image binarization technique that addresses these issues by using adaptive image contrast. The adaptive image contrast is a combination of the local image contrast and the local image gradient that is tolerant to text and background variation caused by different types of document degradations. In the proposed technique, an adaptive contrast map is first constructed for an input degraded document image. The contrast map is then binarized and combined with Canny's edge map to identify the text stroke edge pixels. The document text is further segmented by a local threshold that is estimated based on the intensities of detected text stroke edge pixels within a local window. This is an adaptive image contrast based document image binarization technique that is tolerant to different types of document degradation such as uneven illumination and document smear. Moreover, it works for different kinds of degraded document images. The proposed technique makes use of the local image contrast that is evaluated based on the local maximum and minimum. The proposed method is simple, robust, and involves minimum parameter tuning

V. ACKNOWLEDGEMENT

We would like to thank various technological experts who researches about degraded document image and improve the result by implementing new methods. We would also like to thank Google for providing details on different issues addressing the challenges of degraded document image binarization technique from twitter and about other related techniques.

REFERENCES

[1] 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.

[2] 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

[3] 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.

[4] S. Lu, B. Su, and C. L. Tan, "*Document image binarization using background estimation and stroke edges,*" Int. J. Document Anal. Recognit., vol. 13, no. 4, pp. 303–314, Dec. 2010.

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

[6] G. Leedham, C. Yan, K. Takru, J. Hadi, N. Tan, and L. Mian, “Comparison of some thresholding algorithms for text/background segmentation in difficult document images,” in Proc. Int. Conf. Document Anal. Recognit., vol. 13. 2003, pp. 859–864.

[7] O. D. Trier and T. Taxt, “Evaluation of binarization methods for document images,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 17,no. 3, pp. 312–315, Mar. 1995.

[8] J. Sauvola and M. Pietikainen, “Adaptive document image binarization,”Pattern Recognit., vol. 33, no. 2, pp. 225–236, 2000.

[9] J.-D. Yang, Y.-S. Chen, and W.-H. Hsu, “Adaptive thresholding algorithm and its hardware implementation,” Pattern Recognit. Lett., vol. 15,no. 2, pp. 141–150, 1994.

[10] Blayvas, A. Bruckstein, and R. Kimmel, “Efficient computation of adaptive threshold surface for image binarization,” Pattern Recognit.,vol. 39, no. 1, pp. 89–101, 2006.

[11] M. Sezgin and B. Sankur, “Survey over image thresholding techniques and quantitative performance evaluation,” J. Electron. Imag., vol. 13,no. 1, pp. 146–165, Jan. 2004

[12] L. Eikvil, T. Taxt, and K. Moen, “A fast adaptive method for binarization of document images,” in Proc. Int. Conf. Document Anal. Recognit., Sep. 1991, pp. 435–443.

[13] J. Canny, “A computational approach to edge detection,” IEEE Trans.Pattern Anal. Mach. Intell., vol. 8, no. 6, pp. 679–698, Jan. 1986