

# Contrast adaptive binarization of low quality document images

Meng-Ling Feng<sup>a)</sup> and Yap-Peng Tan<sup>b)</sup>

*School of Electrical and Electronic Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798, Singapore*

a) [FENG0010@ntu.edu.sg](mailto:FENG0010@ntu.edu.sg)

b) [eyptan@ntu.edu.sg](mailto:eyptan@ntu.edu.sg)

**Abstract:** We propose in this paper an improved method for binarizing document images by adaptively exploiting the local image contrast. The proposed method aims to overcome the common problems encountered in low quality images, such as uneven illumination, low contrast, and random noise. Experiments have been conducted and the results are presented to show the effectiveness of the proposed method.

**Keywords:** document image analysis, image binarization, optical character recognition, OCR

**Classification:** Science and engineering for electronics

## References

- [1] M. Pilu and S. Pollard, "A light-weight text image processing method for handheld embedded cameras," Hewlett-Packard Laboratories, Bristol, England, March 2002.
- [2] H.-S. Don, "A noise attribute thresholding method for document image binarization," *Third International Conference on Document Analysis and Recognition*, pp. 231–234, 1995.
- [3] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, March 1979.
- [4] O. D. Trier and T. Taxt, "Evaluation of binarization methods for document images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, no. 3, pp. 312–315, March 1995.
- [5] W. Niblack, *An introduction to digital image processing*, Englewood Cliffs, Prentice Hall, N.J., pp. 115–116, 1986.
- [6] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," *Patt. Recogn.*, vol. 33, pp. 225–236, 2000.
- [7] C. Wolf and J.-M. Jolion, "Extraction and recognition of artificial text in multimedia documents," *Patt. Anal. Appl.*, vol. 6, no. 4, pp. 309–326, Feb. 2004.
- [8] Transym Optical Character Recognition (TOCR) [Available online] <http://www.sorcery.demon.co.uk>.

## 1 Introduction

We have witnessed the migration of small and low-cost imaging devices into portable electronic gadgets, such as personal digital assistants (PDAs) and mobile phones, over the past few years [1]. This revolutionary integration has made possible a range of new and useful applications, e.g., quick capture and analysis of document images for faxing, auto-note taking, foreign text translation, etc. However, the integration also brings new technical challenges to document image analysis owing to the size and cost constraints imposed by these imaging devices.

The problems commonly seen in the document images captured by these imaging devices include: (1) poor contrast due to the lack of sufficient or controllable lighting, (2) non-uniform image background intensity due to uneven illumination, and (3) immoderate amount of random noises due to limited sensitivity of imaging sensor and lack of adjustable camera's exposure time and aperture size. To extract text information from these low quality document images, an image binarization method that is robust against these problems is therefore indispensable. This paper proposes such a binarization method.

## 2 Related Work

Image binarization, a common first step to document image analysis, converts the gray values of document images into two-level representations for text and non-text regions. A large number of binarization techniques have been proposed during the past decade [1]-[7]. These techniques can be generally classified into two categories: one is based on global thresholding; the other, local thresholding.

Global thresholding uses only a single threshold value, which is generally chosen based on some heuristics or the statistics of some global image attributes, to classify image pixels into text (foreground) and non-text (background) pixels [3]. The main drawback of global thresholding is that it cannot adapt well to uneven illumination and random noises, and hence performs unsatisfactorily for low quality document images.

In local thresholding, the threshold value is varied based on the local content of the document image. In comparison with global thresholding, local thresholding generally performs better for low quality images, especially in classifying pixels near text and object boundaries. One of the well performing local thresholding methods was proposed by Niblack [5]. The main idea of Niblack's method is to build a threshold surface, based on the local mean,  $m$ , and standard deviation,  $s$ , of gray values computed over a small neighborhood around each pixel in the form of

$$T = m + k \times s, \quad (1)$$

where  $k$  is a negative constant set to  $-0.2$ . This method, however, tends to produce a large amount of binarization noises in non-text image regions [e.g., see Fig. 3 (b)]. As a result, computationally intensive post-processing is

required to reduce or remove the noises prior to the subsequent OCR (Optical Character Recognition) analysis.

Sauvola *et al.* [6] improve Niblack's method by imposing a hypothesis on the gray values of text and non-text pixels—text pixels have near 0 gray values and non-text pixels have near 255 gray values—and compute the local threshold value as

$$T = m \times \left( 1 - k \times \left( 1 - \frac{s}{R} \right) \right), \quad (2)$$

where  $k$  is a constant set to 0.5, and  $R$  denotes the dynamic range of image gray-value standard deviation and is fixed to 128. Sauvola's method could outperform Niblack's for well-scanned document images, but it faces difficulties in dealing with images that do not agree well with the hypothesis (e.g., documents images captured at insufficient illumination, especially when the gray values of text and non-text pixels are close to each other), as illustrated in Fig. 3 (c).

To overcome this problem, Christian *et al.* [7] propose to determine the local threshold value by normalizing the contrast and the mean gray value of the image as follows

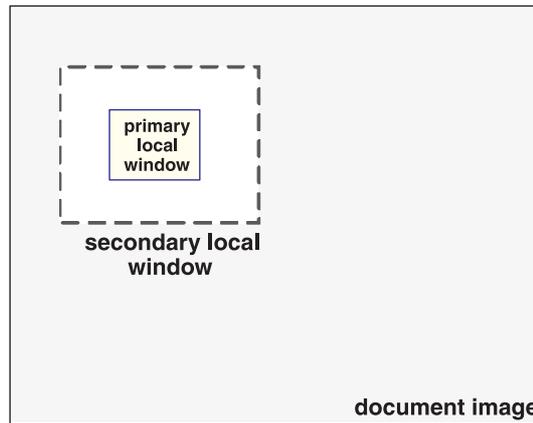
$$T = (1 - k) \times m + k \times M + k \times \frac{s}{R} \times (m - M), \quad (3)$$

where  $k$  is fixed to 0.5,  $M$  is the minimum gray value of the image, and  $R$  is set to the maximum gray-value standard deviation obtained from all local image neighborhoods. This approach achieves the best binarization results among the three local thresholding methods considered. However, the performance of Christian's method degrades considerably when there is a notable change of background's gray values across the image, as all the local threshold values are determined based on the minimum gray value and the maximum standard deviation of local gray values computed from the whole image.

### 3 Proposed Method

To extract useful text information from low quality document images, we have devised a new and reliable local thresholding method to perform binarization based on image local gray-value contrast. Specifically, to determine the local threshold value, the proposed method first computes the local mean,  $m$ , the minimum,  $M$ , and the standard deviation,  $s$ , of gray values in a primary local window (as shown in Fig. 1) with size large enough to cover 1-2 characters. To compensate for the negative effect of uneven illumination, the dynamic range of gray-value standard deviation  $R_S$  is calculated in a larger local window, the secondary local window, instead of the whole image, as also illustrated in Fig. 1. Our study shows that the size of this secondary local window can be properly set to tolerate different degrees of illumination unevenness. To this end, we propose to calculate the local threshold value as:

$$T = (1 - \alpha_1) \times m + \alpha_2 \times \left( \frac{s}{R_S} \right) \times (m - M) + \alpha_3 \times M, \quad (4)$$



**Fig. 1.** Primary local window and secondary local window of the proposed binarization method.

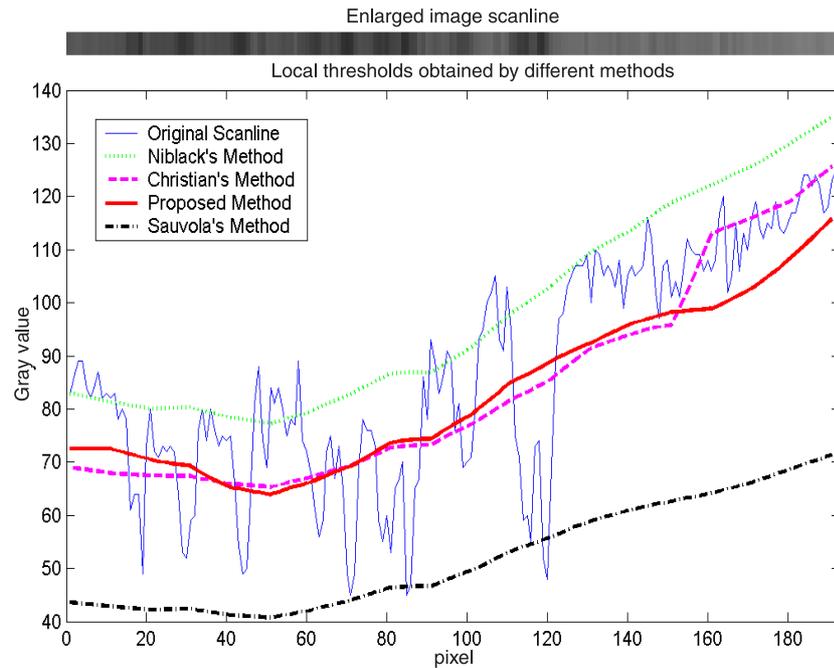
where  $\alpha_2 = k_1(\frac{s}{R_S})^\gamma$ ,  $\alpha_3 = k_2(\frac{s}{R_S})^\gamma$ , and  $\alpha_1$ ,  $\gamma$ ,  $k_1$  and  $k_2$  are positive constants.

The proposed local threshold as computed in (4) consists of three components, and three coefficients ( $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$ ) are defined to allow more flexible and adaptive weighting of these three components. Furthermore, coefficients  $\alpha_2$  and  $\alpha_3$  are adaptively set based on the normalized local standard deviation,  $s/R_S$ , for the reason that local windows consisting of texts normally assume a larger gray-value standard deviation (and hence a larger  $s/R_S$  ratio) than those without covering any texts. Consequently, setting the values of coefficients  $\alpha_2$  and  $\alpha_3$  based on the  $s/R_S$  ratio enables the local threshold so calculated to separate text regions from non-text regions more effectively. Moreover, we found from our empirical study that the power of  $s/R_S$  ratio affects, to a certain extent, the binarization performance as well. Hence, an additional exponential parameter ( $\gamma$ ) is introduced to reflect this effect.

In our implementation of the proposed method, a  $5 \times 5$  median filter is also applied to remove, as many as possible, the random noises from the image under consideration. To minimize the computational load, the local threshold values calculated at the centers of the primary local windows are bilinearly interpolated to obtain the threshold values for all the pixels. Based on our experiments, setting  $\gamma$  to 2 and the values of  $\alpha_1$ ,  $k_1$  and  $k_2$  in the ranges of 0.1-0.2, 0.15-0.25 and 0.01-0.05, respectively, can generally obtain good binarization results for a large variety of document images.

#### 4 Experimental Results

We have applied the proposed method, with parameters  $\gamma$ ,  $\alpha_1$ ,  $k_1$  and  $k_2$  set to 2, 0.12, 0.25 and 0.04, to a set of document images with various sizes and qualities. Fig. 2 shows the thresholds determined by the three existing local thresholding methods and the proposed adaptive method along a scanline of a test document image; the image bar on top of the figure shows the enlarged content of the scanline, where dark regions denote text pixels and gray regions represent non-text pixels. Compared with Niblack's and Christian's methods,



**Fig. 2.** Local thresholds determined by different binarization methods along an image scanline.

it is evident that the local thresholds obtained by our proposed method can adapt better to the unevenness of illumination. In particular, our method could successfully classify the right portion of the scanline (from pixel 140 onwards) as non-text region, while the other two wrongly classified part of it as text region. For Sauvola's method, owing to the constraint imposed by its additional hypothesis, the threshold values turned out to be near zero and failed to binarize the scanline satisfactorily.

Fig. 3 (a)-(e) show the results of applying the methods under comparison to three test images captured by a low-resolution handheld camera. The results show that Niblack's method can segment the text characters well, but it also produces excessive noises in non-text regions. On the other hand, Sauvola's method is suitable for good quality images, but it does not perform well for images that do not agree with its additional hypothesis. Both Christian's method and the proposed method produce superior results. Nevertheless, Christian's method suffers from the effect of uneven illumination, resulting in some noises in non-text regions and broken text strokes. Our proposed method is supreme in terms of robustness to uneven illumination, suppression of noises in non-text regions, and completeness of binarized text characters. More experimental results are shown in Fig. 3 (f)-(h) to illustrate the superiority of our proposed method as compared with Christian's method.

In conjunction with the Transom Optical Character Recognizer (TOCR) [8], we have also performed OCR on a large set of test document images. The binarization results of our proposed method can achieve an average 90.8% correct character recognition rate, while that of Christian's method can only attain 85.4%.



Fig. 3. Binarization results. (a) Original images. (b) Niblack's method. (c) Sauvola's method. (d) Christian's method. (e) Proposed method. (f) Additional images. (g) Christian's method. (h) Proposed method.

## 5 Conclusion

We have presented an improved local thresholding method for binarizing document images by adaptively exploiting the local image contrast. Compared with other existing local thresholding methods, our proposed method performs better for low quality document images, especially those with uneven illumination, low contrast, and random noises. Experimental results obtained from various challenging document images have also been presented to testify the effectiveness and superiority of our proposed method.