A NOVEL METHOD FOR BINARIZATION OF BADLY ILLUMINATED DOCUMENT IMAGES

Seyed Amin Tabatabaei
MSc Student, University of Tehran, Tehran, Iran
a.tabatabaei@ece.ut.ac.ir

Mehdy Bohlool
PhD Student, Florida Institute of Technology, Melbourne, Florida, USA
Mbohlool2009@my.fit.edu

ABSTRACT

This paper presents a novel document image binarization technique that separates text from background in badly illuminated document images. This technique is based on background estimation by using morphological closing operation. Binarization methods can be categorized into two categories: the global methods which do not work well in badly illuminated images, and the local methods which are usually parametric. In this paper a new local method is proposed. The most important feature of this local method is that, contrary to other common local methods, it is not dependent on any parameter. In other words it is an automatic method. Morphological closing operation is used to compensate for uneven background illumination. Closing operation is applied to remove small dark details while living the overall gray and larger dark features relatively undisturbed. Experiment results show that the proposed method offers better result for document images with bad degradation and lighting variance in comparison to former common methods.

Index Terms— binarization, thresholding, document images, morphological operations

1. INTRODUCTION

Recently, the document images have become more and more popular. Generally, the first step to process them is binarization. Binarization is always a great challenge in all image process fields, especially in the process of document images where binarization result can directly affect the Optical Character Recognition (OCR) rate directly [1].

Degradations such as lighting variance bring negative effects on document image binarization and make the difference of gray levels between the foreground pixels and their neighboring background pixels so small that local histograms of foreground and backgrounds become extremely close or even overlap [2].

There exist two types of binarization, global and local methods [3]. Otsu's method is one of the famous global methods [4]. The global thresholding algorithms choose a fixed intensity threshold value T (from 0 to 255). If the intensity value of any pixel of an input image is more than T, the pixel is set to white, otherwise it is set as black [5]. The global methods are simple and runtime saving, but they cannot adapt to lighting variance images or work well on low quality images.

Another method is the local adaptive technique, such as the one presented in [7] to [11]. In these methods different regions of an image have different thresholds [6]. Niblack's method [7] computes the threshold based on local mean and local variance of image using a move-window. The size of move-window is selected based on predefined character size assumption, but character size varies not only in different documents, but also in different areas of a single document. Thus this assumption imposes an important constraint on these approaches. Sauvola's algorithm [8] is a modification of Niblack's method, and claims to give improved performance on documents in which the background contains light texture, big variation and uneven illumination. Bukhari [9] used Sauvola's method to propose a nonparametric algorithm. On the other hand, Yanwitz's method [10] used an iterative scheme to calculate the threshold surface based on local maxima of the gradient. There are also other methods such as BST's in [11] by which the background intensity surface is computed first, and then final threshold surface is calculated based on the estimated background.

In general, local adaptive methods provide much better results than global methods, but they are often dependent on some specific parameters. This paper proposes a novel nonparametric local adaptive document image binarization method which is based on the estimation of background using closing operation. The results of this method are compared with a global method, as well as two local parametric methods and a local nonparametric method. Experiments show that the proposed method offers better result for document images with bad degradation and lighting variance.

This paper is organized as follows: in section two previous works are described. Section three describes the proposed method in detail and section four provides with experimental results. At last, section five draws a conclusion.

2. PREVIOUS WORKS

In this section two well-known binarization method are explained, a global method [4] and a local one [7].

Otsu's thresholding method is a global binarization method which involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels on each side of the threshold, i.e. the pixels that either treated as foreground or background. The aim is to find the threshold value where the sum of foreground and background spreads is at its minimum. This method gives satisfactory results when the numbers of pixels in each class are close to each other. The Otsu method still remains the reference for comparing different thresholding and binarization methods in general.

Niblack's algorithm, on the other hand, calculates a pixel-wise threshold by sliding a rectangular window over the image. The threshold T is computed by using the mean m and standard deviation s, of all pixels in the window, and is denoted as:

$$T = m + k * s \qquad (1)$$

where k is a constant which determines how much of the total print object edge is retained, and has a value between 0 and 1. The value of k and the size SW of the sliding window define the quality of binarization. Fig. 1 (a), (b) show binarization result in thick and unclear strokes with a small k value, and slim and broken strokes with a large k value, while with a small SW value noise is closer to texture as shown in Fig. 1(c).
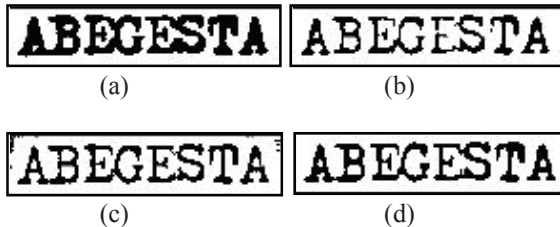

Figure 1. Binarization using Niblack's algorithm. (a) with SW=25x25 and k=0.1 (b) with SW=25x25 and k=0.9 (c) with SW=11x11 and k=0.6 (d) with SW=25x25 and k=0.6

3. THE PROPOSED METHOD

This section provides instructions on how to use closing operator for document image binarization and describes how to determine the size of closing structuring element.

3.1. Using Closing Operation to Binarize a Badly Illuminated Document Image

Uneven background illumination can be eliminated by closing operation as shown in Fig. 2.
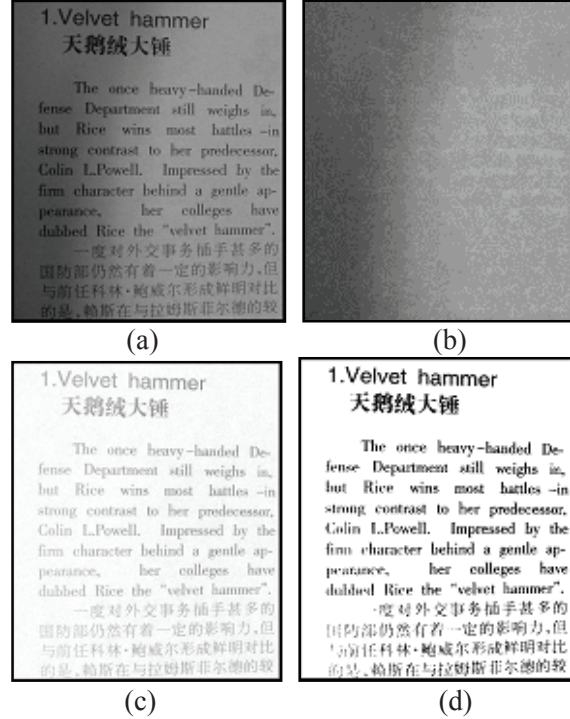

Figure 2. (a)original image; (b)result of closing on inverse of original image; (c) result of subtracting (a) from inverse of (b); (d) new thresholded image.

Fig. 2(a) shows a badly illuminated document image. The uneven illumination makes image thresholding difficult. Closing the image can produce a reasonable estimation of the background across the image as long as the structuring element is large enough so that it does not fit entirely within the text and small enough so that non-text dark region does not fit entirely within it. Fig. 2(b) shows the result of closing operation. By subtracting original image from this image, an image with a reasonably even background can be produced. Fig. 2(c) shows the result while fig. 2(d) presents the new thresholded image.

3.2. Computing size of Structuring Element

Experiments have shown that a structuring element with approximately twice the size of text font gives the best results. Thus, it is of benefit to estimate the text size in the document image by some means.

The result of global binarization method on a badly illuminated document image can be categorized in four regions:

- Region0: white pixels.

- Region1: several tiny black regions, caused by noise. Experimental results showed that this region accounts for less than two percent of the total black regions. If the sample has good quality, this region will not show up in results.

- Region2: the non-text black regions caused by shadows. The size of these areas is usually larger than the size of Region3 areas. Thus, there are some structuring elements which fit entirely in these areas but do not fit entirely in region3 areas. If we can determine the size of text regions, we can fix the size of structuring element.

- Region3: text regions which comprise more than two percent of all black pixels. Effort should be made to preserve these regions and look for other real text regions, such as the ones occurring in region0 (labeled as background by the global threshold method) or in region2 (their neighbors' pixels labeled as foreground by the global threshold method).

We assume that the size of smallest shadow element (within region2) is larger than twice of the size of the largest text element (within region3). In the next section the size of the largest text element in the image is computed based on text properties. Proposed method's steps are as follow.
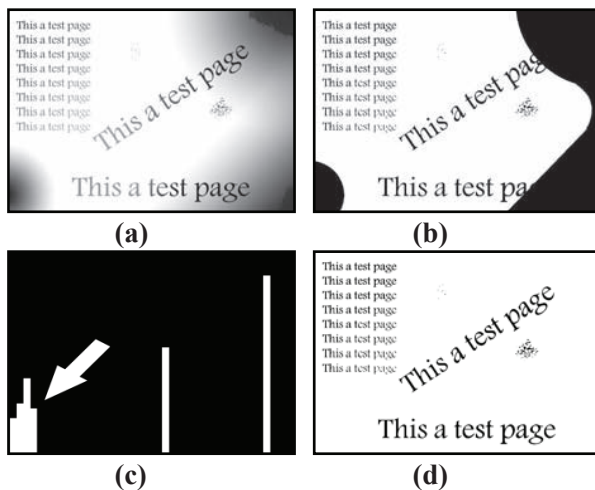

Figure 3. Proposed method steps (a) original sample; (b) result by global method; (c) S-histogram of (b); (d) result obtained from proposed implementation.

3.2.1 Global binarization

The first step consists of binarizing the original image, as shown in Fig. 3(a), by using a global method(in this case, Otsu's method), The result is shown in Fig. 3(b).

3.2.2. Finding largest black square for each pixel

During the second step, the size of the largest black square for each pixel of the binarized image is obtained, and these values are saved in a matrix M.

3.2.3 Setting one value for each connected set of black pixels in M

Following that, we compute the biggest value of M in each connected set, and assign their values t to those elements in M. An element of M is a member of a connected set if it has

non-zero value and one of its neighbors is a member of that set.

3.2.4. Forming a histogram with new definition (S-histogram)

At this step we make a S-histogram. The value of S-histogram in point I, f(I) , is equal to the number of elements of M with the value I ,as shown in Fig. 3(c).

3.2.5. Determining the size of closing structuring element

Then, we want to determine the size of closing structuring element, so we look for the largest x which:

$$\sum_{i=1}^{x} f(i) \ > \ 0/02 \cdot \sum_{i=1}^{\infty} f(i) \qquad (2)$$

$$\sum_{i=x}^{2x} f(i) \ = \ 0 \qquad (3)$$

In Fig. 3(c) an arrow points to the largest x (X_{max}) value. Finally, the size of closing structure would be $2X_{max}$ that we found in previous step. The result is shown in Fig. 3(d).

4. EXPERIMENTAL RESULTS

The proposed method has been tested on many poorly illuminated document images. Besides the shading degradation, some of the test documents are also degraded by variations in character size as illustrated in Fig. 2(a).
By means of experiments, the proposed method is compared with other three well-known techniques, namely Otsu's method [4], Niblack's method [7] and Sauvola's method [8]. In particular, the parameters of Niblack's and Sauvola's methods are set as ideal values that we want to achieve. In other words, the best results from these methods are compared with the results from the proposed method. Moreover, the proposed method is compared with a recently published nonparametric method (Bukhari [9]).

4.1. Evaluation Method

The performance evaluation is based on a well established technique [12] that counts True Positive (TP), False Positive (FP) and False Negative (FN) pixels in order to calculate Recall and Precision metrics.

- A pixel is classified as TP if it is ON in both Ground Truth (GT) and binarization result images.

- A pixel is classified as FP if it is ON only in the binarization result image.

- A pixel is classified as FN if it is ON only in the GT image.

The Recall metric shows the ratio of the number of pixels, which our method truly classified as foreground, to the number of all pixels classified as foreground from the

ground truth image. Precision metric is the ratio of the number of pixels, which our method truly classifies as foreground, to the number of all pixels which classified as foreground. Setting C_{TP} as the number of TP pixels, C_{FP} as the number of FP pixels and C_{FN} as the number of FN pixels, Recall (RC) and Precision (PR) metrics are given as follows:

$$RC = C_{TP} / (C_{FN} + C_{TP}) \qquad (4)$$

$$RC = C_{TP} / (C_{FP} + C_{TP}) \qquad (5)$$

Recall and precision metric have values between zero and one. As these metrics approach one, the results get better.
The overall metric that is used for evaluation is the F-Measure (FM) which is calculated as follows:

$$FM = (2 \times RC \times PR/(RC + PR)) \times 100\% \qquad (6)$$

4.2 Results

The evaluation methodology was applied on ten handwritten images and machine-printed documents with low quality, shadows, non-uniform illumination, presence of characters from the other side of the page and other significant artifacts. Evaluation results for Fig. 2(a) and average evaluation results for all test images are respectively shown in table 1 and table 2.

Table1 .Evaluation metrics for each binarization technique

Fig. 2(a)	Otsu	Niblack	Sauvola	Bukhari	Proposed method
Recall	0.8000	0.9700	0.9999	0.9861	0.9981
Precision	0.9951	0.9745	0.9738	0.9999	0.9961
FM	88.69%	97.22%	98.66%	99.29%	99.70%

Table 2.The average value of F-measure, Recall and Precision, for each binarization technique for all test images.

	Otsu	Niblack	Sauvola	Bukhari	Proposed method
Recall	0.8187	0.9318	0.9917	0.9657	0.9916
Precision	0.9954	0.9782	0.9914	0.9962	0.9937
FM	89.84%	95.44%	99.15%	98.07%	99.26%

As is shown in table 2, our method is not as good as Niblack method according to Recall metric, and is too close to Otsu's method according to Precision metric. But according to F-measure, that is the overall evaluation metric, our method gives the best result (99.26%) in comparison to Otsu (89.84%), Niblack (95.44%), Sauvola (99.15%) and Bukhari (98.07%).

5. CONCLUSION

This paper presents a novel document image binarization technique that is categorized as a local method, but the main difference of the proposed method and other methods is that it is not parametric due to the fact that it can compute these parameters automatically. Although the method shows good results in regular images, its main advantage is in badly illuminated document images and uneven backgrounds. Experiments show that the proposed method overcomes other local and global methods numerically, especially in bally illuminated and uneven background.

6. REFERENCES

[1] O.D. Trier, A.K. Jain, "Goal-directed Evaluation of Binarization Methods," IEEE Trans Pattern Analysis and Machine intelligence, pp. 1191-1201, 1995.

[2] M.J.Taylor, C. R. Dance, "Enhancement of Document Images from Cameras," proceeding of SPIE Conference on Document Recognition, pp. 230-241, 1998.

[3] B. Sankur, M. Sezgin, "Survey over Image Thresholding Techniques and Quantitative Performance Evaluation," journal of Electronic Imaging, pp. 146-165, 2004.

[4] N. Otsu, "A Thresholding Selection Method from Gray-level Histogram," IEEE Trans. System, Man, and Cybernetics, pp. 62-66, 1978.

[5] Q. D.M. Do, A. C. Downton, and J. H. Kim J.He, "A Comparison of Binarization Methods for Historical Archive Documents," IEEE Computer Society, vol1, pp. 538-542, 2005.

[6] C. Wang, R. I. Dai Y. Zhu, "Document Image Binarization Based on Storke Enhancement," 18th International Conference on Pattern Recognition (ICPR 2006), vol1, pp. 955-958, 2006.

[7] W. Niblack, "An Introduction to Digital Image Processing," Prentice-Hall, pp. 115-116, 1986.

[8] J. Sauvola, T. Seppanen, S.Haapakoski and M.Pietkainen, "Adaptive Document Binarization," Elsevier Science B.V., pp. 147-152, 1997.

[9] S. S. Bukhari, F. Shafait, T. M. Breuel "Adaptive Binarization of Unconstrained Hand-Held Camera-Captured Document Images", Journal of Universal Computer Science, vol. 15, no. 18 , 2009

[10] A. M. Bruckstein, S. D. Yanowitz, "A New Method for Image Segmentation," 9th International Conference on Pattern Recognition, pp. 82-95, 1988.

[11] C. R. Dance, M. Seeger, "Binarising Camera Images for OCR," 6th International Conference on Document Analysis and Recognition, pp. 54-58, 2001.

[12] C. R. Dance, M. Seegar, "On the Evaluation of Document Analysis Components by Recall, Precision, and Accuracy," Proceedings of the Fifth International Conference on Document Analysis and Recognition, pp. 713-716, 1999.