# A Novel Thresholding Method for Text Separation and Document Enhancement

Adnan Khashman [1], Boran Sekeroglu [2]

[1] Electrical & Electronic Engineering Department, [2] Computer Engineering Department, Near East University, Nicosia, Cyprus
{[1] amk, [2] bsekeroglu}@neu.edu.tr

**Abstract**

Many thresholding-based image enhancement techniques have been developed and used for document analysis, where the simplicity and efficiency of thresholding makes it ideal to use for classifying layers within documents. However, the efficiency of these enhancement techniques can be impaired by the variation of grey levels in different documents, thus causing over-thresholding or under-thresholding. This paper presents a novel global single- stage thresholding method for separating background and foreground layers in text documents. The method finds an optimum thresholding value or exact separation point for each document using the relationship between luminance value and mean intensity of the document without considering peak values in the grey level histogram. The proposed method is implemented using 50 historical documents and five specifically designed words, and then compared to three other efficient and known thresholding methods. Experimental results suggest that the proposed method performs well for text separation and enhancement of document images.

**Keywords:** Image thresholding, document enhancement, text separation

## *1. Introduction*

Analyses of documents have become more significant where efficient recognition and enhancement required for paper-based documents. Efficiency of recognition and analysis of these documents depend on the efficient separation and classification of the background and foreground layers. Thus, the initial purpose of document analysis techniques is the effective preparation and separation of various layers in documents in order to provide sufficient and clear data for recognition systems and human readers.

Historical and handwritten documents have different and various levels of noise which causes pessimistic separation of the layers because of the age, paper, pen and pencil influences on the documents [Zheng et Al. (2004)]. Age factor adds irremovable noise and meaningless random shapes on the documents which prevent efficient separation and recognition of the layers. Paper properties such as patterned

or coloured papers; add different background layers to the scanned documents. In addition, the variety of pens and pencils produces different and various foreground layers for the documents. Therefore, analyzing these degraded and multi-layered documents generally requires effective techniques.

One of the simplest methods that can be used to separate foreground and background layers from each other is thresholding. This is based on the assumption that objects and background layers in the image can be distinguished by their grey level values [Kittler and Illingworth (1986)]. Thresholding methods can be categorized into two groups as Global Thresholding and Local (Adaptive) Thresholding. Global thresholding is a simple and efficient method where a defined or computed threshold value is used to separate foreground objects from background and Local (Adaptive) Thresholding is the assigning of a value to each pixel to determine whether it is a foreground or background pixel using local information from the image.

This paper presents a novel global single-stage thresholding technique, which uses the mean intensity and the luminance value of a document image to separate the foreground from the background with minimum loss of information and minimum deviation from the ideal threshold point. The proposed novel method is implemented using 50 historical document scans and 5 specifically created greyscale words of various grey level characters. The robustness of our text separation method is evident when comparing it to Kittler and Illingworth, Otsu, and Niblack thresholding methods.

## *2. Overview of Binarization Methods*

Several thresholding methods had been developed and proposed to separate foreground layers from background layers [Kittler and Illingworth (1986)], [Solihin and Leedham (1999)], [Kapur et Al. (1985)], [Otsu (1979)], [Cheriet et. Al. (1998)], [Parker (1991)], [Niblack (1986)], [Kavallieratou and Antonopoulou (2005)]. Global thresholding methods use a defined or computed threshold value for the entire image. Most of the basic global methods are based on the brightness histogram of images which are called histogram-derived thresholds such as fixed point threshold, which is an alternative way in very high contrast images [Abutaleb (1989)], Minimum Algorithm [Brinck and Pendock (1996)] which uses three-point mean filter iteratively to smooth the histogram and to determine the threshold point, and Concavity algorithm [Doermann and Rosenfeld (1992)]. In addition to these methods, Background-symmetry algorithm is one of the basic global thresholding methods which assume a distinct and dominant peak for the background that is symmetric about its maximum.

Trier and Jain [Trier and Jain (1995)] compared the most known and successful 11 local and 4 global thresholding methods. In their comparison results, it was shown that Otsu Thresholding Technique [Otsu (1979)] is the most successful method in the

global methods and followed by the Kittler and Illingworth method [Kittler and Illingworth (1986)], Abutaleb method [Abutaleb (1989)] and Kapur Method [Kapur et Al. (1985)], respectively. Trier and Jain noticed also that, Niblack method [Niblack (1986)] showed better results than the other local thresholding methods, but they also noted that Niblack method has a disadvantage with the size of neighborhood. The size should be small enough to preserve local details and large enough to suppress noises.

Another important evaluation of thresholding methods was performed by Sezgin et al. [Sezgin and Sankur (2004)] for 40 selected thresholding methods using 40 document images and 40 nondestructive testing images. Sezgin et al. noticed that, Kittler and Illingworth minimum error thresholding method performed best results in both image sets in both global and local thresholding methods.

Otsu Thresholding was proposed in 1979 [Otsu (1979)], as a selection method which was based on image histogram. It uses discriminant analysis to divide foreground and background by maximizing the discriminant measure. The threshold operation is regarded as the partitioning of the pixels of an image into two classes such as objects and background at grey-level *t*, and an optimal threshold point can be determined by minimizing equations using within-class variance, between-class variance, and the total variance . Kittler and Illingworth method, which was proposed in 1986 [Kittler and Illingworth (1986)], starts by choosing an arbitrary initial threshold *T*, calculates error between both sides of *T* and finds the minimum error point of *T*, which is used for the threshold value. Niblack method [Niblack (1986)] is based on the varying threshold over the image, based on the local mean and local standard deviation, where the threshold at pixel (x,y) is calculated by using sample and standard deviation values in a local neighbourhood of *(x,y)*. However, determining optimum mask size is important factor and problem in Niblack's method to get well-classified images. Because small mask sizes adds additional noises, besides losing information by using large mask sizes.

These methods, proved their successes for text separation from document images, thus; Otsu, Kittler-Illingworth and Niblack methods can be considered as the most efficient document image thresholding methods, and will be used for comparison purposes with our proposed method that is described in the following section.

## *3. The Proposed Method*

Our proposed method uses a global single-stage thresholding technique that finds the optimum threshold value for document images, using luminance value (brightest point) and mean of intensities of image. The relationship between colour values of greyscale images provides a threshold point of foreground and background of image. Our proposed method differs from histogram-based global thresholding methods by using luminance value instead of peak values in histogram to provide efficient separation for both bimodal and unimodal images; and to consider the all information

that belongs to image. The luminance value can always be considered as 255, however, in document images, it is observed that only a few of the documents have a luminance value of 255; thus, the luminance value is considered as a variable in this work. The, luminance value is computed by:

$$G_{\max} = F_{\max}(g), \tag{1}$$

where $G_{max}$ denotes the Luminance value, $F_{max}$ is a function to find maximum grey point in the x-plane of brightness histogram which was represented as $g$.

The mean of intensities of an image is used to compute the initial separation point of background and foreground layers as follows:

$$M = \left( \sum_{y=0}^{\dim_y} \sum_{x=0}^{\dim_x} I[x, y] \right) / (\dim_y \times \dim_x), \tag{2}$$

where $M$ represents the mean of image, $dim_x$ and $dim_y$ denotes the x and y dimensions of image respectively, and $I[x,y]$ represents the original greyscale image.

Local Deviation is the difference between luminance value and mean of intensities, which represents the deviation of the optimum threshold point from the initial separation point; as follows:

$$D = G_{\max} - M, \tag{3}$$

where $D$ is the Local Deviation of the mean from luminance value.

Thus, the difference between mean value which is the initial separation point, and local deviation, gives exact separation point for a document as:

$$\Theta = \mathrm{M} - \mathrm{D}, \text{ and therefore } \Theta = \left| M - (G_{\max} - M) \right|, \tag{4}$$

where θ denotes the optimum threshold value of image (exact separation point). This novel method provides minimum loss of information with a minimum fuzziness of the texts which is important for further processing and analysis of documents. A general formula obtained from equations 1-4 can be written as:

$$MI[x, y] = \begin{cases} 0, & \text{if } I[x, y] \le \Theta \\ 255, & \text{else.} \end{cases}, \tag{5}$$

where $MI[x,y]$ represents the thresholded image.

## *4. Document Image Database*

The proposed thresholding method, Otsu Method, Kittler and Illingworth method, and Niblack Method were implemented using 55 documents aiming to clearly separate

and classify foregrounds that consists text from backgrounds. The implementations were carried out using two image sets of documents. The first set comprises 50 historical documents which contain a total of 2021 words with different contrasts and brightness. The second set comprises 5 created words which contain a total of 45 characters with different backgrounds and different colour to test the occurrence of the characters after thresholding.

After the implementation of the four thresholding methods, the thresholded documents were then analyzed and visually inspected by 15 independent analyzers to first set. Analyzers were also asked to consider noise occurrence and continuity of characters to determine clear characters in the thresholded words in the second set. A general comparison is then performed by combining the results using the two sets of documents. General results are categorized as: *recognized* or *unrecognized* words in Set 1, and *clear* or *unclear* characters in Set 2.

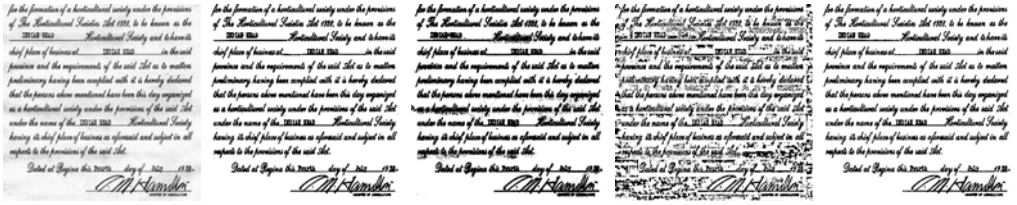
## *5. Implementation Results and Comparison*

The implementation of our proposed thresholding method, Otsu method, Kittler and Illingworth method, and Niblack method was carried out using the above described document database which comprised 50 historical documents (Set 1), and 5 created words with varying backgrounds and character colors (Set 2). The general implementation results using the first set are shown in Table 1, and examples of the enhanced images can be seen in Figure 1 and Figure 2.

Using our thresholding method for text separation from 50 documents yielded an 85.60 % recognition rate; where 1730 words out of the total 2021 words were clearly recognized by the 15 independent analyzers. Otsu thresholding method also showed good results, however it over-thresholds which causes minor loss of information. Kittler and Illingworth method produced efficient results for small documents which do not contain high level of contrasts, but in degraded historical documents, Kittler and Illingworth method adds noise to the document images. The efficiency of Niblack's method depends on the mask size; however there is no exact rule to define the mask size. Small mask size adds noise to the documents and larger mask sizes cause some loss of information.

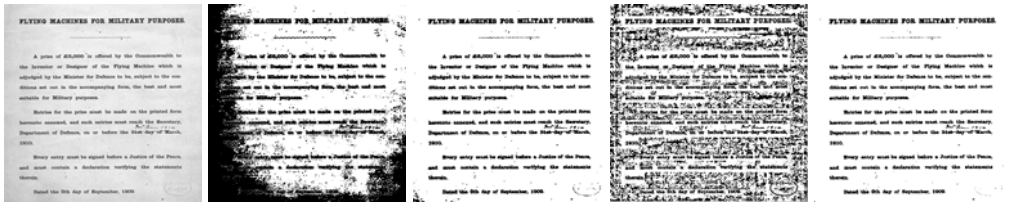
***Table 1.*** *Recognition Rates of Words in (Set 1)*

| Thresholding Method | Total Words | Recognized Words | Unrecognized Words | Recognition Rate |
|---|---|---|---|---|
| Otsu | 2021 | 1657 | 364 | 81.98 % |
| Kittler and Illingworth | 2021 | 1045 | 976 | 51.70 % |
| Niblack | 2021 | 1699 | 322 | 84.06 % |
| Proposed Method | 2021 | 1730 | 291 | 85.60 % |



(a) (b) (c) (d) (e)

***Figure 1.*** *Example of (Set 1) Implementation of 1928 Document (a) Original Image (b) Otsu Method (c) Kittler and Illingworth Method (d) Niblack Method (15x15 mask) (e) Proposed Method*



(a) (b) (c) (d) (e)

***Figure 2.*** *Example of (Set 1) implementation of 1908 Document (a) Original image (b) Otsu Method (c) Kittler and Illingworth Method (d) Niblack Method (15x15 mask) (e) Proposed Method*

The implementations using the second set investigate the performance of the four thresholding methods when characters or words of varying colors exist. All methods apart from Otsu, produced readable characters with various degrees of clearance.

In general, when combining the implementation results from both sets, our proposed thresholding method shows more superior results than the other three methods, thus it can be efficiently applied for text separation from documents with minimum noise and minimum loss of information. Table 2 list the general implementation results using the second set, and implementation examples are shown in Figure 3.

***Table 2.*** *Recognition Rates of Characters in (Set 2)*

| Thresholding Method | Total Characters | Clear Characters | Unclear Characters | Recognition Rate |
|---|---|---|---|---|
| Otsu | 45 | 40 | 5 | 88.88 % |
| Kittler and Illingworth | 45 | 45 | 0 | 100 % |
| Niblack | 45 | 45 | 0 | 100 % |
| Proposed Method | 45 | 45 | 0 | 100 % |

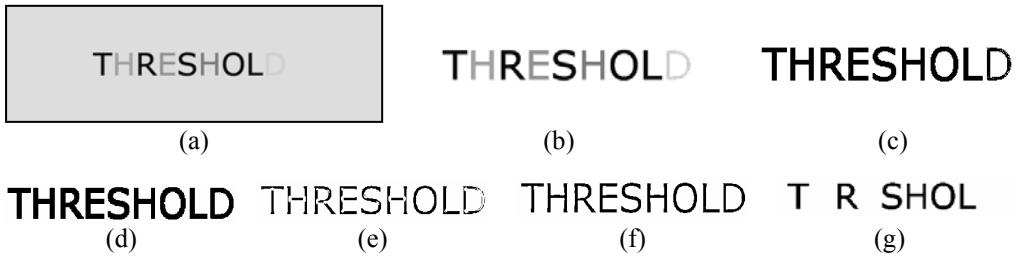



***Figure 3.*** *Example of (Set 2) Implementation (a) Original Text (b) Proposed Method –Greyscale (c) Proposed Method-Binary (d) Kittler and Illingworth (e) Niblack (9x9 Mask) (e) Niblack (15x15 Mask) (g) Otsu Method*

## *5. Conclusions*

This paper presented a novel thresholding method that can be efficiently used for foreground separation from backgrounds in document analysis. The proposed method uses the local and total deviation of the document pixels to find an optimum threshold value for the document image. Threshold values vary from document to another as they represent the color and contrast values of a document image. The difference in averaged values provides unique and optimum thresholding point for each document image, and thus, clearly, separates foregrounds from backgrounds with minimal loss of information and minimum fuzziness of the thresholded image.

The efficiency of the proposed method was demonstrated by its implementation on 55 documents that comprised 50 historical document scans and 5 created words with various character colours. A recognition rate of 85.6% was achieved using the first document set, and a 100 % recognition rate was achieved using the second set.

A comparison was also drawn between our method and three other known and efficient thresholding methods, namely Otsu, Kittler and Illingworth, and Niblack methods. The comparison results show that our method is superior to the other methods, thus it can be efficiently used for enhancing scanned documents as part of a document analysis scheme. Future work will include the analysis of thresholded images using OCR Modules, in addition to investigating the use of a neural network for determining different optimum global thresholds for different document images.

## *References*

Abutaleb, S. (1989), *Automatic Thresholding of Gray-Level Pictures Using Two Dimensional Entropy*, Computer Vision, Graphics, and Image Processing, vol. 47, pp.22-32.

Brink, A.D., Pendock, N.E. (1996), *Minimum Cross-Entropy Threshold Selection*, Pattern Recognition, vol. 29(1), pp.179-188.

Cheriet, M., Said, J.N., Suen, C.Y. (1998), *A Recursive Thresholding Technique for Image Segmentation*, IEEE Transactions on Image Processing, vol. 7(6), pp.918-921.

Doermann, D.S., Rosenfeld, A. (1992), *Recovery of Temporal Information from Static Images of Handwriting*, Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, pp.162-168.

Kapur, J.N., Sahoo, P.K., Wong, A.K.C. (1985), *A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram*, Computer Vision, Graphics, and Image Processing, vol. 29, pp.273-285.

Kavallieratou,, E., Antonopoulou, H. (2005), *Cleaning and Enhancing Historical Document Images*, Lecture Notes on Computer Science, Springer-Verlag, Berlin Heidelberg New York, vol. 3708, pp.681-688.

Kittler, J., Illingworth, J. (1986), *Minimum Error Thresholding*, Pattern Recognition, vol. 19, pp.41-47.

Niblack, W. (1986), *An Introduction to Digital Image Processing*, Prentice Hall, pp.115-116.

Otsu, N. (1976), *A Threshold Selection Method from Gray-Level Histogram*, IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, pp. 62-66.

Parker, J.R. (1991), *Gray level Thresholding in Badly Illuminated Images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 13(8), pp.813-819.

Sezgin, M., Sankur, B. (2004), *Survey Over Image Thresholding Techniques and Quantitative Performance Evaluation*, Journal of Electronic Imaging, vol. 13(1) pp.146-168.

Solihin, Y., Leedham, C.G. (1999), *Integral Ratio: A New Class of Global Thresholding Techniques for Handwriting Images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21(8), pp. 761-768.

Trier, O.D., Jain, A.K. (1995), *Goal-Directed Evaluation of Binarization Methods*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, pp.1191–1201.

Zheng, Y., Li, H., Doermann, D (2004), *Machine Printed Text and Handwriting Identification in Noisy Document Images*, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26(3), pp. 337-353.