ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Simple and Innovative Document Image Binarization Using Adaptive Image Contrast Method


Prasanth Kumar Embeti
M.Tech Student,
Dept of Computer Science and Engineering,
CMR Institute of Technology, Hyderabad.

Mr.A.Balaram
Associate Professor,
Dept of Computer Science and Engineering,
CMR Institute of Technology, Hyderabad.

Abstract:

Image segmentation is a set of segments that collectively cover the entire image, or a set of contours extracted from the image. In the process of improving degraded document images segmentation is one of the difficult task due to background and foreground variation. This paper presents a new approach for enhancement of degraded documents. It consists of an adaptive image contrast based document image binarization technique that is tolerant to different type of document degradation such as uneven illumination document smear involving smudging of text, seeping of ink to the other side of page, degradation of paper ink due to aging etc. The images i.e. scanned copies of these degraded documents are provided as an input to the system. They are processed to get the finest improved document so that the contents are visible readable. Contrast image construction can be constructed using local image gradient and local image contrast. Further edge estimation algorithm is used to identify the text stroke edge pixels .The text within the document is further segmented by a thresholding technique which is based on the height and width of letter size present in degraded document image. It works for different format of degraded document images. The method has been tested on Document Image Binarization Contest (DIBCO) experiments on Bickley diary dataset, consists of several challenging degraded document images.

Keywords:

Degradation, Equations, Histograms, Image edge detection, Image segmentation, Mathematical model, Robustness.

I.INTRODUCTION:

The Image segmentation is an essential task in the fields of image processing and computer vision. It is a process of partitioning the digital images and is used to locate the boundaries into a finite number of meaning full regions and easier to analyze. The Simplest method for image segmentation is thresholding. Thresholding is an important technique in image segmentation, enhancement and object detection. The output of the thresholding process is a binary image whose gray level value 0 (black) will indicate a pixel belonging to a print, legend, drawing, or target and a gray level value 1 (white) will indicate the background. The main complexity coupled with thresholding in documents applications happen when the associated noise process is non-stationary. The factors that make difficult thresholding action are ambient illumination, variance of gray levels within the object and the background, insufficient contrast, object shape and size non-commensurate with the spectacle. The lack of objective measures to assess the performance of thresholding algorithms is another handicap. Many methods have been reported in the literature. It can extract the object from the background by grouping the intensity values according to the thresholding value.

Thresholding divides the image into patches, and each patch is thresholding by a threshold value that depends on the patch contents. In order to decrease the effects of noise, common practice is to first smooth a boundary prior to partitioning. The Binarization technique is aimed to be used as a primary phase in various manuscript analysis, processing and retrieval tasks. So, the unique manuscript characteristics, like textual properties, graphics, line drawings and complex mixtures of the layout-semantics should be included in the requirements.

II. RELATED WORK:

Many thresholding techniques have been reported for document image binarization. is usually not a suitable approach for the degraded document binarization. Adaptive thresholding, 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 estimate the local threshold by using the mean and the standard variation of image pixels within a local neighborhood window.

The main drawback of these window-based thresholding techniques is that the thresholding performance depends heavily on the window size and hence the character stroke width. Other approaches have also been reported, including background subtraction, texture analysis, recursive method, decomposition method, contour completion, Markov Random Field, matched wavelet, cross section sequence graph analysis, self-learning, Laplacian energy user assistance and combination of binarization techniques.

These methods combine different types of image information and domain knowledge and are often complex. 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 neighboring document background.

They are very effective and have been used in many document image binarization techniques. In Bernsen's paper, the local contrast is defined as follows:

$$C\,(i,j) = I_{max}(i,j) - I_{min}(i,j)$$

Where $C\,(i,j)$denotes the contrast of an image pixel (i, j), $I_{max}(i,j)$and $I_{min}(i,j)$ denote the maximum and minimum intensities within a local neighborhood 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 $I_{max}(i,j)$ and $I_{min}(i,j)$ Barnes's method is simple, but cannot work properly on degraded document images with a complex document background. We have earlier proposed a novel document image binarization method by using the local image contrast that is evaluated as follows :

$$C\,(i,j) = \frac{I_{max}(i,j) - I_{min}(i,j)}{I_{max}(i,j) + I_{min}(i,j) + \in}$$

III. LITERATURE SURVEY

IMAGE NOISE:

The concept can be defined also for signals spread over more complicated the main source of noise in digital images arises during image acquisition (digitization) or during image transmission. The performance of image sensor is affected by variety of reasons such as environmental condition during image acquisition or by the quality of the sensing element themselves.Image noise styles may be divided differently according to different criterion. The criterions include: the causes of image noise's generation, the shape of the noise amplitude distribution over time, noise spectrum shape and the relationship between noise and signal, and so on. For example, image noise can be divided into additive noise and multiplicative noise according to the relationship between noise and signal. There are many types of image noise.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Such as additive noise, multiplicative noise, salt and pepper noise, Gaussian noise.Image noise is random (not present in the object imaged) variation of brightness or color information in images, and is usually an aspect of electronic noise. It can be produced by the sensor and circuitry of a scanner or digital camera. Image noise can also originate in film grain and in the unavoidable shot noise of an ideal photon detector. Image noise is an undesirable by-product of image capture that adds spurious and extraneous information. The original meaning of "noise" was and remains "unwanted signal"; unwanted electrical fluctuations in signals received by AM radios caused audible acoustic noise ("static"). By analogy unwanted electrical fluctuations themselves came to be known as "noise". Image noise is, of course, inaudible.

The magnitude of image noise can range from almost imperceptible specks on a digital photograph taken in good light, to optical and radio astronomical images that are almost entirely noise, from which a small amount of information can be derived by sophisticated processing (a noise level that would be totally unacceptable in a photograph since it would be impossible to determine even what the subject was) For processing of digital image, we can add Gaussian noise, Poisson noise, salt and pepper noise to the original image in the Mat lab platform. The Gaussian noise is Gaussian white noise with constant mean and variance. The probably of most frequently occurring noise is additive Gaussian noise. The PDF of a Gaussian random variable, z , is given by

$$P(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-u)^2/2\sigma^2}$$

Where z represents gray level, μ is the mean of average value of z , and σ is its standard deviation. Salt and pepper noise refers to a wide variety of processes that result in the same basic image degradation: only a few pixels are noisy, but they are very noisy. The PDF of Salt and pepper noise is given by

$$P(z) = \begin{cases} P_a & \text{for } z = a \\ b_o & \text{for } z = b \\ 0 & \text{otherwise} \end{cases}$$

IV. IMPLEMENTATION

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 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.

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 our earlier method, the local contrast evaluated by 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

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

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 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 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 follows:

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

where C(i, j) denotes the local contrast and (Imax(i, j) − Imin(i, j))refers to the local image gradient that is normalized to [0, 1]. The local windows size is set to 3 empirically. αis 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 follows:

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

Where Stddenotes 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,∞], where the power function becomes a linear function when γ = 1.

Therefore, the local image gradient will play the major role in Equation 3.1 when γ is large and the local image contrast will play the major role when γ is small.

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, 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 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.

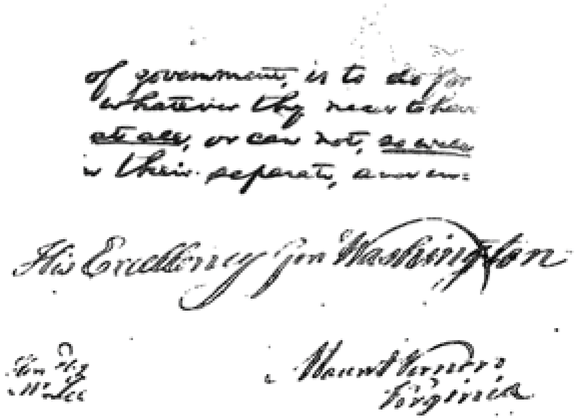


Figure 3.1: Binary contrast maps

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

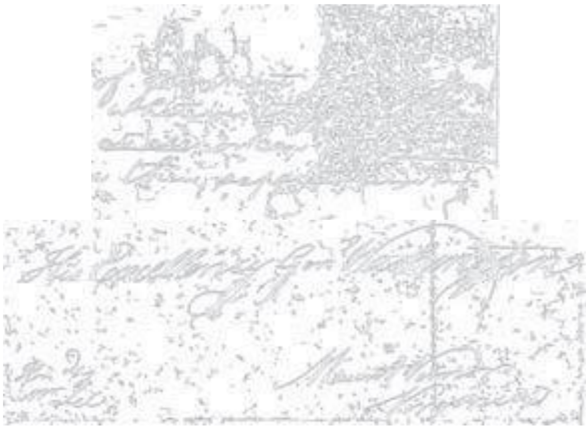


Figure : canny edge maps

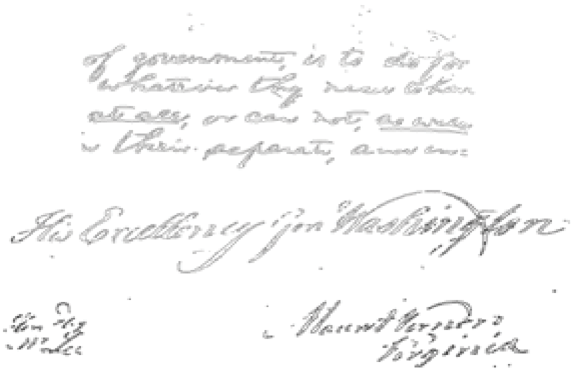


Figure: Combined edge maps of the sample document images

It should be noted that Canny's edge detector by itself often extracts a large amount of non-stroke edges as illustrated without tuning the parameter manually. In the combined map, we keep only pixels that appear within both the high contrast image pixel map and canny edge map. The combination helps to extract the text stroke edge pixels accurately.

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 follows:

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

Where Emean and Estd are the mean and standard deviation of the intensity of the detected text stroke edge pixels within a neighborhood window W, respectively. The neighborhood window should be at least larger than the stroke width in order to contain stroke edge pixels. So the size of the neighborhood window W can be set based on the stroke width of the document image under study, EW, which can be estimated from the detected stroke edges as stated in Algorithm 1. Since we do not need a precise stroke width, we just calculate the most frequently distance between two adjacent edge pixels (which denotes two sides edge of a stroke) in horizontal direction and use it as the estimated stroke width.

First the edge image is scanned horizontally row by row and the edge pixel candidates are selected as described. If the edge pixels, which are labeled 0 (background) and the pixels next to them are labeled to 1 (edge) in the edge map (Edg), are correctly detected, they should have higher intensities than the following few pixels (which should be the text stroke pixels). So those improperly detected edge pixels are removed. In the remaining edge pixels in the same row, the two adjacent edge pixels are likely the two sides of a stroke, so these two adjacent edge pixels are matched to pairs and the distance between them are calculated .

Algorithm 1 Edge Width Estimation
Require: The Input Document Image I and Corresponding Binary Text Stroke Edge Image Edg
Ensure: The Estimated Text Stroke Edge Width EW

1. Get the width and height of I
2. for Each Row i = 1 to height in Edgdo
3. Scan from left to right to find edge pixels that meet the following criteria:
    a) its label is 0 (background);
    b) The next pixel is labeled as 1(edge).
4. Examine the intensities in I of those pixels selected in Step 3, and remove those pixels

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

that have a lower intensity than the following pixel next to it in the same row of I.

5. Match the remaining adjacent pixels in the same row into pairs, and calculate the distance between the two pixels in pair.
6. end for
7. Construct a histogram of those calculated distances.
8. Use the most frequently occurring distance as the estimated stroke edge width EW.

After that a histogram is constructed that records the frequency of the distance between two adjacent candidate pixels. The stroke edge width EW can then be approximately estimated by using the most frequently occurring distances of the adjacent edge pixels.

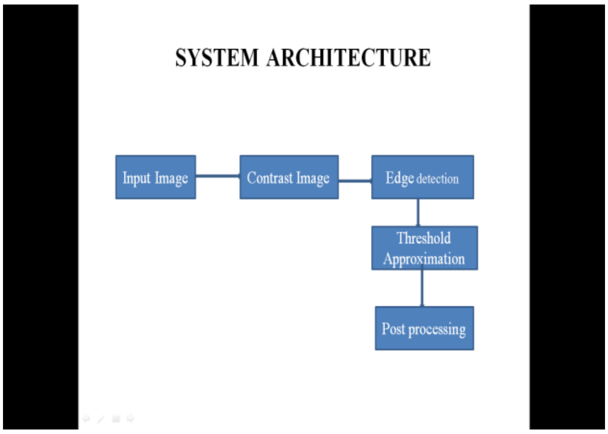
4) Post-Processing

Once the initial binarization result is derived from as described in previous subsections, the binarization result can be further improved by incorporating certain domain knowledge as described in Algorithm 2. 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 neighborhood 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 labeled 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 as described.

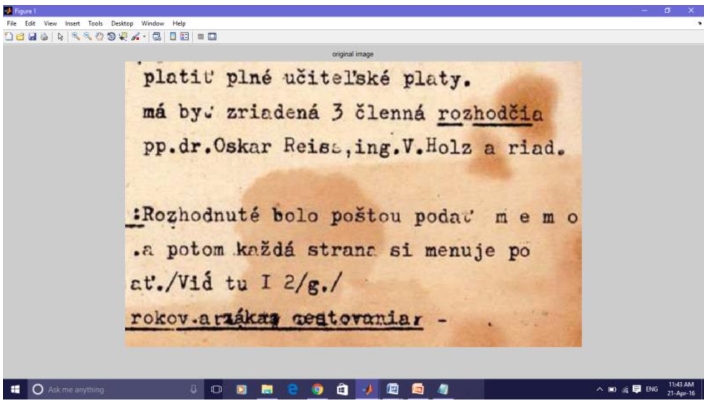
Algorithm 2 Post-Processing Procedure
Require: The Input Document Image I , Initial Binary Result B and Corresponding Binary Text Stroke Edge Image Edg
Ensure: The Final Binary Result B f

1. Find out all the connect components of the stroke edge pixels in Edg.
2. Remove those pixels that do not connect with other pixels.
3. for Each remaining edge pixels (i, j): do
4. Get its neighborhood pairs: (i − 1, j) and (i + 1, j); (i, j − 1) and (i, j + 1)
5. if The pixels in the same pairs belong to the same class (both text or background) then
6. Assign the pixel with lower intensity to foreground class (text), and the other to background class.
7. end if
8. end for
9. Remove single-pixel artifacts along the text stroke boundaries after the document thresholding.
10. Store the new binary result to B f.

System Architecture:

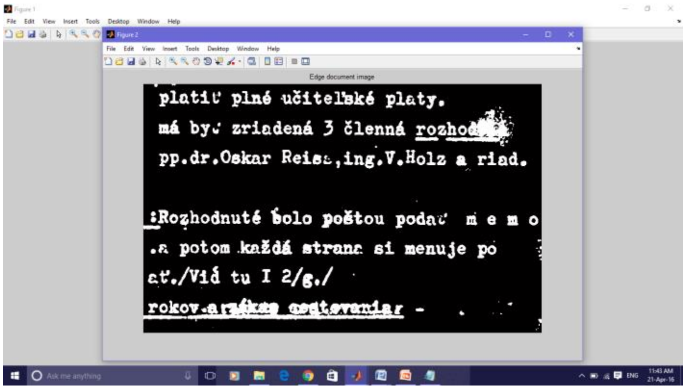


Input Image :

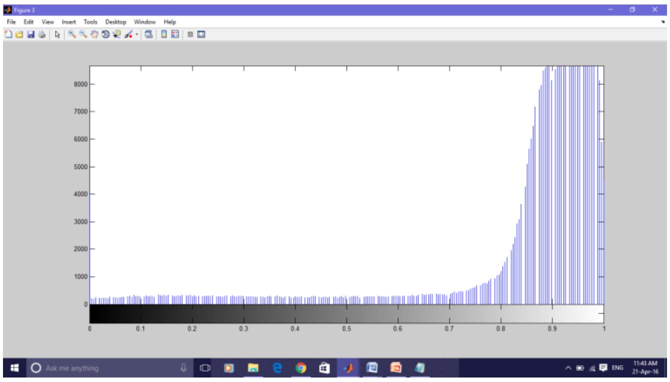


Edge Document Image:

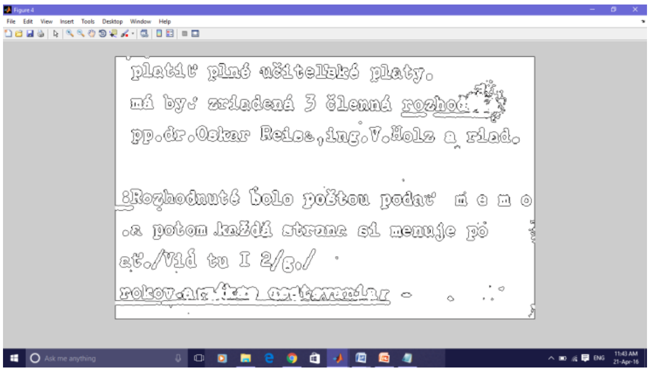
ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal



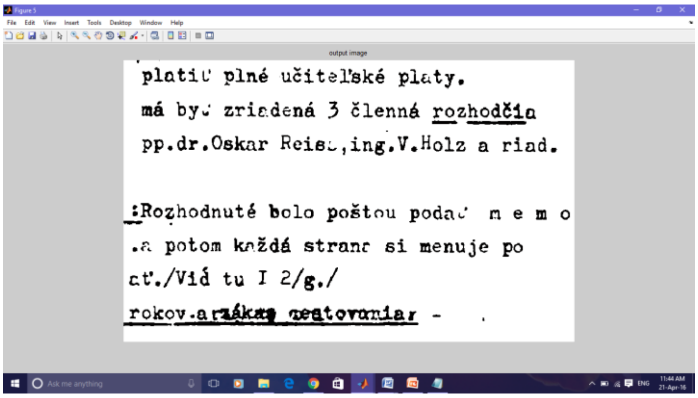
Histogram Image :



GrayscaleImage :



Output Image :



V. RESULTS & DISCUSSIONS

Table : Evaluation Results of the Dataset of DIBCO 2009

Methods	F-Measure	PF-measure	PSNR	NRM	MPM	DRD
Adaptive	66.22528	0.89023	9.401	0.292	0.052	161.99
BERN	62.15014	0.47091	6.753	0.355	0.121	298.842
SAUV	66.95062	0.91148	9.431	0.287	0.054	161.21

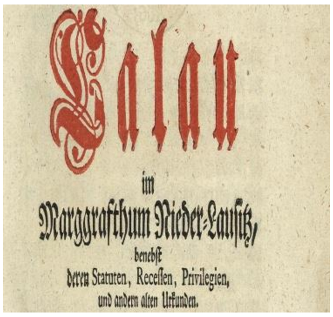


Fig : Original Image

Fig: Grayscale Image

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

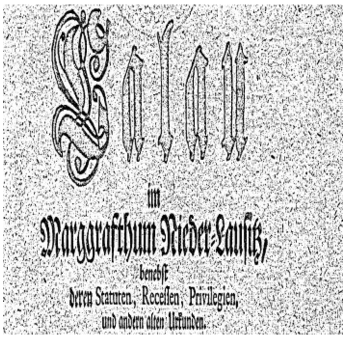


Fig :Binarized Image

VI. CONCLUSION & FUTURE SCOPE:

In this work presents 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. The proposed technique is simple and robust, only few parameters are involved. 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 has been tested on the various datasets. Experiments show that the proposed method outperforms most reported document binarization methods in term of the F-measure, pseudo F-measure, PSNR, NRM, MPM and DRD.

VII. REFERENCES

[1] [1] Bolan Su, Shijian Lu, and Chew Lim Tan, Senior Member, IEEE, “Robust Document Image Binarization Technique for Degraded Document Images”, IEEE Transactions on Image Processing, Vol. 22, No. 4, April 2013.

[2] 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.

[3] I. Pratikakis, B. Gatos, and K. Ntirogiannis, “H-DIBCO 2010 hand- written 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 back- ground 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 filter,” 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] 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.

[8] O. D. Trier and A. K. Jain, “Goal-directed evaluation of binarization methods,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 12, pp. 1191–1201, Dec. 1995.

[9] 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.

[10] A. Brink, “Thresholding of digital images using two-dimensional entropies,” Pattern Recognit., vol. 25, no. 8, pp. 803–808, 1992.

[11] J. Kittler and J. Illingworth, “On threshold selection using clustering criteria,” IEEE Trans. Syst., Man, Cybern., vol. 15, no. 5, pp. 652–655, Sep.–Oct. 1985.

[12] N. Otsu, “A threshold selection method from gray level histogram,” IEEE Trans. Syst., Man, Cybern., vol. 19, no. 1, pp. 62–66, Jan. 1979.