Performance Evaluation of Image Binarization Technique for Recognition of Ancient Historical Documents

V.Sathya Narayanan
Department of ECE
Kongu Engineering College
Erode, TamilNadu

N.Kasthuri
Department of ECE
Kongu Engineering College
Erode, TamilNadu

***Abstract*—Information acquisition from degraded historical document has always been a challenging task due to various forms of degradation. Image binarization is very much essential in the restoration of degraded historical documents. Eventhough many algorithms have been proposed, still there is a need for an effective algorithm to solve all kind of degradation problems. In this paper ,an optimum binarization technique is proposed that addresses these issues by using a combined approach of local image contrast and gradient . Firstly,the input image is binarized and then canny egde map is applied to extract text stroke edge pixels.To enhance further,morphological operations are carried out based on shapes.Finally an adaptive thresholding is applied to segment foreground and background pixels.To determine the quality, the proposed method has been tested on 3 public data sets (DIBCO 2009, 2010, 2011) that were taken from pattern recognition and image analysis(PRImA) research lab. Simulation result shows that the proposed binarization method achieves performance improvement interms of F-measure, NRM, MPM, PSNR as 88.4461, 0.0708, 0.00265 and 18.420 respectively**.

Keywords - Binarization, Canny Edge Map, Adaptive Thresholding

I. INTRODUCTION

Preservation of cultural heritage Historical ancient document is very important throughout the world. To maintain the quality of such original document, conversion of machine editable format is highly required to access that information [1]. However such ancient documents may be degraded due to various forms of degradation factors especially complex background, poor quality, overlapping of characters, document aging,etc.[2].So restoration and enhancement of degraded document is very much essential in today's digital library world.

When compared with traditional filtering methods[3],[4],[5], document image binarization takes part a major role in any kind of document analysis since it influences the subsequent segmentation, feature extraction and classification process [6]. Global and Local document image binarization techniques are most widely approaches. Global approach determines the global threshold for the entire document whereas the Local approach [7,1] determines the threshold for every pixel adaptively.Among these two methods,local document image binarization performs better than global document image binarization for the degraded documents.Adaptive binarization [8,9] also performs better to remove noise while segmenting text.

An adaptive reconstructive filter for the recognition of misrepresented text present in the document was proposed by Stubberud et al. [10]. Adrian et al. [11] proposed an adaptive mask size for connecting the broken lines to increase recognition rate. Banerjee et al. [2] proposed the content based text reconstruction for enhancing the text model from complex background and overlapping of characters in the historical documents. Lazzara et al. [12] proposed a multiscale scheme for enhancing Sauvola's method. Milyaev et al.[13] presented scene text binarization method for achieving better performance. Ntirogiannis et al. [14] introduced a binarization method using picture element based approach for the recognition of ancient document images. Punctuation marks cannot be recognized by this method.

An optimum binarization technique using modified adaptive contrast with canny's edge map method is proposed to overcome the various limitations found in the literature. The organization of the paper is framed as follows. The database details that were taken from pattern recognition and image analysis(PRImA) research lab are presented in section 2. Section 3 presents the framework of proposed technique with mathematical model. The results and the evaluation of performance metrics are described in section 4. Finally Section 5 describes the conclusion of the proposed technique.

II. DATABASE DETAILS

The DIBCO (Document Image Binarization Contest) Dataset mainly represents major European libraries. It contains most of the historical document layout which needs to be digitized to preserve the cultural heritage. The content of this dataset includes newspapers, magazines, images and layouts. Majority region covers graphic regions, textual regions with different fonts and sizes.

TABLE I SAMPLE IMAGES OF DIBCO 2009 DATASET

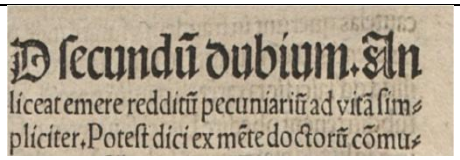
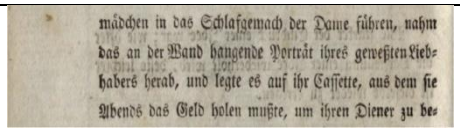
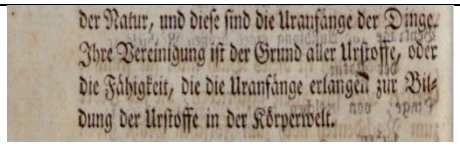
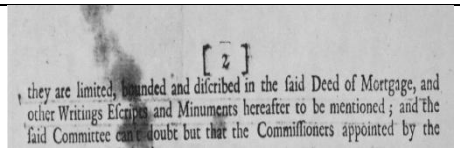
S.No	Samples
1	
2	
3	
4	
5	

TABLE II. SAMPLE IMAGES OF DIBCO 2010 DATASET

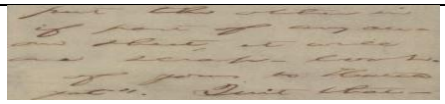
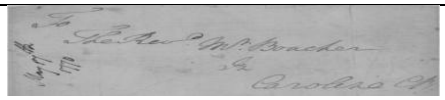
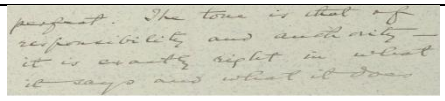
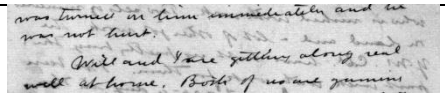
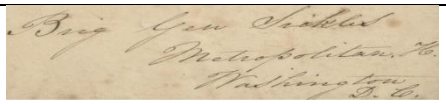
S.No	Samples
1	
2	
3	
4	
5	

TABLE III. SAMPLE IMAGES OF DIBCO 2011 DATASET

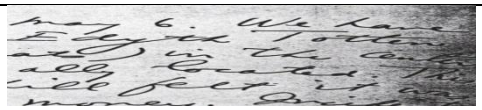
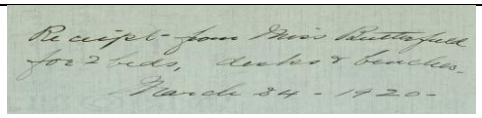
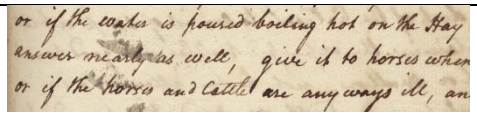
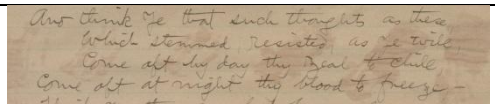
S.No	Samples
1	
2	
3	
4	
5	

Table I,Table II and Table III shows the samples images of DIBCO 2009,DIBCO 2010 and DIBCO 2011 dataset respectively.

III. PROPOSED SYSTEM

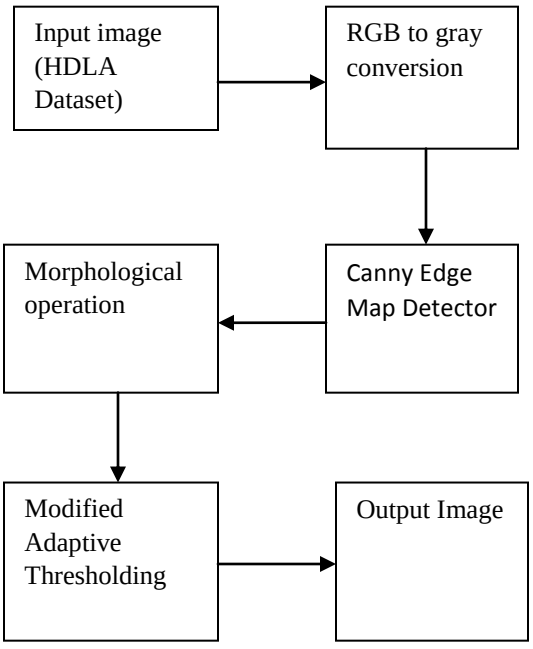


Fig.1. Proposed System Module

A. Proposed Module

The proposed method includes the following four modules

1. RGB to gray conversion
2. Canny Edge Map Detector
3. Morphological Operation
4. Adaptive Thresholding

B. Module Description

*1) RGB to Gray conversion :*The input image is always considered as RGB image for the purpose of sensing and representing in electronic systems. To preserve the brightness(luminance) of the image, the input 24 bit colour image is converted into a 8-bit image .The intensity of colours RED, GREEN and BLUE are combined in the proportion of 30% , 60% and 11% respectively to obtain a grayscale image.

2) Canny Edge Map Detector: A wide range of edges can be detected using a multistage algorithm in canny edge map detector. It is used to extract meaningful information from images. The amount of data required to process the image is also reduced by usin this operator. It provides low error rate,good localization and minimal response. It is one of the most widely used edge detection operator due to its its optimality and simplicity.

The algorithm to perform canny edge map detection is given as follows :

1.Image can be smoothened by applying Gaussian filter.

Gaussian filter kernel equation of size $(2p+1)\times(2p+1)$ is given by:

$$H_{m,n} = \frac{1}{2\pi\sigma^2}\left(\frac{-(m-(p+1))^2+(n-(p+1))^2}{2\sigma^2}\right) \quad (1)$$

2.Intensity gradients is determined by using the following equation

$$G = \sqrt{G_x^2 + G_y^2} \quad (2)$$

Where G_x,G_y represents horizontal and vertical directions respectively.

3.Estimate potential edges by means of applying double threshold.

4.Weak edges are suppressed by tracking the edges by means of hysteresis.

3) Morphological Operation : The structure and shape of objects can be described by using morphological operations that operates on Sets(objects in an image). The output of morphological operations depends on size of an original image and Structural Element.

Basically,there are four morphological operations namely Dilation,Erosion,Opening and Closing. The opening of A by B is obtained by the erosion operation followed by dilation operation. The closing of A by B is obtained by the dilation operation followed by erosion operation. Dilation is used for expansion of an image by pixel addition to the object boundaries whereas erosion is just the opposite of dilation operation. To shrink the image, erosion operation is used by eliminating the pixels on the boundaries. Addition and deletion of pixels depend on the boundary of the structuring element.

Rules for Dilation and Erosion

Dilation:

Highest pixel value out of all pixel's in the neighbourhood at the input is assigned as the output pixel in this morphological operation. For example, the pixel in the output is set to 1 if any of the pixel at the input is 1 in an binary image.

Erosion:

Lowest pixel value out of all pixel's in the neighbourhood at the input is assigned as the output pixel in this morphological operation For example, the pixel in the output is set to 0 if any of the pixel at the input is 0 in an binary image.

4) Modified Adaptive Thresholding : Thresholding is a technique commonly used to obtain an binary image from grayscale image .Adaptive thresholding changes its threshold value according to the pixel values present in the image which results a strong illuminated shadows. To segment an image, thresholding can be performed. For every pixel, the threshold value is computed. Then thresholding is done by assigning higher intensity values as foreground valus and lower intensity value as background value.Higher and lower intensity values are estimated by comparing the values with threshold.Adaptive thresholding also considers spatial variations in illumination.

The proposed approach traverses the image from left side to right side and thus covers the entire image. If the pre-processed image value $F'(x, y)$ goes beyond the threshold T, text regions are identified. The threshold T [15] is estimated based on concavities of the histogram $C(F')$.

$q(F')$ indicates the image probability mass function. The deepest concavity points become threshold once if the histogram convex hull is computed. The following sigmoid function is used if the gradients of background and the text regions are similar to exhibit the behaviour of the pixel values of background region.

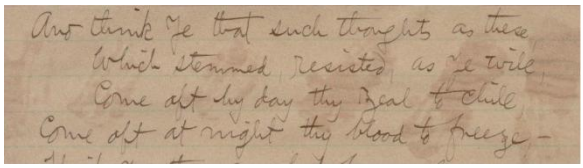
$$C\left(F'(x,y)\right) = r \propto \left(\frac{(1-q_2)}{1+\exp\left(\frac{-4F'(x,y)}{a(1-q_1)}\right)} + \frac{2(1+q_1)}{1-q_1} + q_2\right) \qquad (3)$$

where r is a weighting parameter, the difference between average black and white pixels is denoted by $\propto$ and q_1,q_2 are constants.

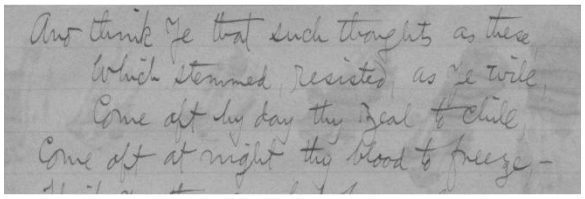
IV. Simulation And Results

A. Simulation Analysis

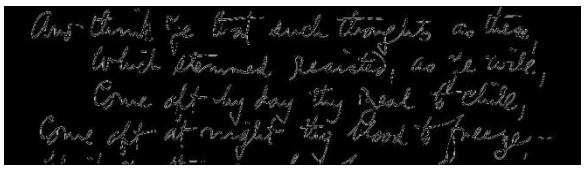
To determine the quality of the proposed method,various document images have been considered. The proposed method is tested on 150 degraded historical documents. The proposed method is compared with existing famous Otsu method.



(a)



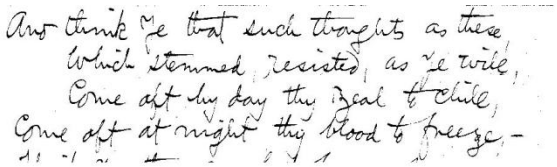
(b)



(c)

(d)

(e)



(f)

Fig.1.Binarization of Historical Document Image (a)Original document image (b)RGB to gray conversion (c)Canny edge map (d)Dilation (e)Erosion (f) Modified Adaptive Contrast output

B. PERFORMANCE METRIC ANALYSIS

1) F- measure:

Accuracy of binarization can be tested by evaluating F-measure.Its value should be higher for accurate binarization.

$$F_{measure} = \frac{2\times Precision \times Recall}{Precision + Recall} \qquad (4)$$

$$Recall = \frac{TP}{TP+FN} \qquad (5)$$

$$Precision = \frac{TP}{TP+FP} \qquad (6)$$

where FP,TP,FN indicates False Positive,True Positive and False Negative respectively.

2) PSNR (Peak Signal to Noise Ratio):

The restoration of the desired pixels is indicated by PSNR.Ideal value of PSNR lies in the range of 15 dB to 25dB.

$$PSNR = 10\log\left(\frac{d^2}{MSE}\right) \qquad (7)$$

Where Mean Square Error

$$MSE = \sum_{ij} \frac{[f_1(i,j)-f_2(i,j)]^2}{i*j} \qquad (8)$$

and d denotes a constant

3) MPM (Misclassification Penalty Metric):

It determines the prediction against ground truth(GT) for analyzing the proposed algorithm in a better manner. Low MPM value indicates better performance.

$$MPM = \frac{\sum_{m=1}^{FN} g_{FN}^{m} + \sum_{n=1}^{FP} g_{FP}^{n}}{2D} \qquad (9)$$

Where g_{FN}^{m} denotes distance of the false negative pixel. g_{FP}^{n} denotes the distance of the false positive pixel .D represents a Normalization Factor.

4) NRM (Negative Rate Metric):

It indicates better enhancement of output image with respect to ground truth image. Better performance is indicated by low NRM value.

$$NRM = \frac{\left(\frac{FN}{FN+TP}\right)+\left(\frac{FP}{FP+TN}\right)}{2} \quad (10)$$

where, TP,TN,FP,FN indicates True Positive, True Negative, False Positive and False Negative respectively.

TABLE.IV EVALUATION RESULTS OF DIBCO 2009,DIBCO2010 AND DIBCO 2011 DATASET

Perfor-mance Metric	**Existing Otsu Method**	**Proposed method for DIBCO 2009 Dataset**	**Proposed method for DIBCO 2010 Dataset**	**Proposed method for DIBCO 2011 Dataset**
FM	65.754	91.400	90.994	91.601
NRM	0.0678	0.054	0.0437	0.0463
MPM	0.0526	0.0339	0.0014	0.043
PSNR	11.815	19.480	19.367	20.677

From the above table IV, it is found that the proposed technique achieves improved results by evaluating the performance metric such as F-measure, PSNR, NRM and MPM.

V. CONCLUSION

This paper provides an optimum technique for binarization of degraded ancient historical document images using a combined approach of local image contrast and gradient. The proposed method overcomes the different degradation factors such as nonuniform illumination, complex background and so on.. The proposed technique involves modified adaptive thresholding, Canny edge map detection method and morphological operations. Samples of DIBCO 2009,DIBCO 2010 and DIBCO 2011 datasets are experimented through the proposed method by computing the performance metric such as F-measure, PSNR, NRM and MPM and compared with existing technique. Significant improvement is achieved for the binarization of degraded ancient historical documents to preserve the ancient cultural heritage.

REFERENCES

[1] E. Kavallieratou, E. Stamatatos, Improving the quality of degraded document images, in: Document Image Analysis for Libraries, 2006. DIAL'06. in: Proceedings of the Second International Conference on, IEEE, 2006, pp. 10.

[2] J. Banerjee, A. Namboodiri, C. Jawahar, Contextual restoration of severely degraded document images, in: CVPR, 2009, pp. 517–524.

[3] A. CBOVIK, T.S. HUANG, M. JR., A generalization of median filtering using linearcombinations of order statistics, IEEE Transactions on Acoustics, Speech, and Signal Processing ASSP-31 (6).

[4] R.C. Gonzalez, R.E. Woods, Pocessamento Digital de Imagens, Prentice Hall,Pearson, 2010.

[5] B. Han, Y. Zhu, D. Comaniciu, L.S. Davis, Kernel-based bayesian filtering for object tracking, in: CVPR (1), 2005, pp. 227–234.

[6] B. Gatos, I. Pratikakis, S.J. Perantonis, Adaptive degraded document image binarization, Pattern Recognit. 39 (3) (2006) 317–327.

[7] Ø.D. Trier, A.K. Jain, Goal-directed evaluation of binarization methods, IEEE Trans. Pattern Anal. Mach. Intell. 17 (12) (1995) 1191–1201.

[8] R.F. Moghaddam, M. Cheriet, Adotsu: an adaptive and parameterless generalization of otsu's method for document image binarization, Pattern Recognit. 45 (6) (2012) 2419–2431.

[9] R. Hedjam, R.F. Moghaddam, M. Cheriet, A spatially adaptive statistical method for the binarization of historical manuscripts and degraded document images, Pattern Recognit. 44 (9) (2011) 2184–2196.

[10] P. Stubberud, J. Kanai, V. Kalluri, Adaptive image restoration of text images that contain touching or broken characters, in: ICDAR, 1995, pp. 778–781.

[11] A.P. Whichello, H. Yan, Linking broken character borders with variable sized masks to improve recognition, Pattern Recognit. 29 (8) (1996) 1429–1435.

[12] G. Lazzara, T. Géraud, Efficient multiscale sauvolas binarization, Int. J. Doc. Anal.Recognit. (IJDAR) 17 (2) (2014) 105–123.

[13] S. Milyaev, O. Barinova, T. Novikova, P. Kohli, V. Lempitsky, Fast and accurate scene text understanding with image binarization and off-the-shelf ocr, Int. J. Doc. Anal. Recognit. (IJDAR) (2015) 1–14.

[14] K. Ntirogiannis, B. Gatos, I. Pratikakis, Performance evaluation methodology for historical document image binarization, IEEE Trans. Image Process. 22 (2) (2013)595–609.

[15] A. Rosenfeld, D.La. Torre, Histogram concavity analysis as an aid in thresholdselection (in image processing), IEEE Trans. Syst., Man, Cybern.13(1983)231–235.