Robust Document Image Binarization Technique for Degraded Document Images by using Morphological Operators

Rupinder Kaur
M.TECH (ECE) Student
BGIET, Sangrur

Naveen Kumar
Assistant Professor
BGIET, Sangrur

ABSTRACT
To make a robust document images from badly degraded images it is necessary to discriminating a text from background images but it is a very challenging task. There are so many binarization techniques can be used for making the document pictures reliable. But problem of thresholding and filtering cannot be solved. In the existing method, edge based segmentation can be done and Canny edge detector used. In our proposed technique, Image Binarzation for degraded document images has being use Region based segmentation. Firstly, an RGB image convert into gray image then image filtering can be done on the basis of Wiener Filtering and Gaussian filter. In the proposed method, we can modify algorithms and test degraded document images then compare the result that come from previous paper results.

Keywords
Adaptive Binarization Techniques, Document Segmentation, Image Processing, Morphological Operators and Thresholding.

1. INTRODUCTION
Document images, as a substitute of paper documents, mainly consist of common symbols such as hand written machine-printed characters, symbols and graphics. [HYPERLINK \l "Nag00" 1] In many practical applications, we only need to keep the content of the document, so it is sufficient to represent text and diagrams in binary format which will be more efficient to transmit and process instead of the original gray-scale image. It is essential to threshold the document image reliably in order to extract useful information and make further processing such as character recognition and feature extraction, especially for those poor quality document images with shadows, non-uniform illumination, low contrast, large signal-dependent noise, smear and smudge. 2] Therefore, thresholding a scanned gray-scale image into two levels is the first step and also a critical part in most document image analysis systems since any error in this stage will propagate to all later phases.

Document Image Binarization is the first process that occurs in document analysis and it is used for discriminating foreground text from document background. [HYPERLINK \l "Nag00" 1]Binarization is used as pre-processor before Optical Character Recognition. This technique converts the gray scale document image into binary document image as shown in figure (1.1).

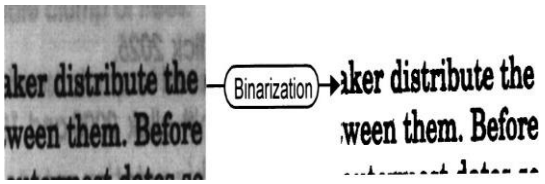

Fig 1.1: Example of Binarization

Binarization is a preprocessing stage for report examination and it is utilized to fragment the frontal area content from the archive foundation. This system guarantees quicker and exact document picture handling tasks.3] Most record investigation calculations are based in view of hidden binarized picture information. The utilization of bi-level information diminishes the computational load and aides in utilizing streamlined investigation techniques contrasted with 256 levels of dim scale or shading picture data. Document picture understanding systems oblige consistent and semantic substance conservation for thresholding. Despite the fact that record picture binarization has been considered for a long time, the thresholding of pictures is still a testing undertaking because of the high variety between the content stroke and the archive foundation. For a data picture, some handling stages ought to be utilized before the content extraction. One of the steps includes binarization. There are so many applications of document image binarization are as follows:

1. An optical character recognition (OCR) system for the handwritten numerical data is needed in order to speed up the process of database creation and extension. [HYPERLINK \l "Naf01" 4]
2. It is used in the Character recognition in video is a challenging task.
3. A New Efficient Binarization Method for MRI of Brain Image.
4. From text detection in videos to person identification.

2. RELATED WORK
A number of methodologies have been proposed by several researchers on image segmentation using binarization and its applications toward moving object detection and human gait recognition. A review of the recent research on binarization is given here.

Otsu (1979) [HYPERLINK \l "NOt79" 5] suggested a nonparametric automatic optimal threshold selection for picture segmentation in order to maximize the separability of the resultant classes in gray levels. Optimal Threshold is naturally selected based on global property but not on local property.

It can be used for picture segmentation so that maximum separation can be done between resultant classes of gray levels.

Niblack (1986)6] introduced a maximum Kapur et al (1985) and entropy algorithm that divides the histogram of the image into two probability distributions, one representing the objects and one representing the background.

Solihin, Y. and C.G. Leedham (1999) [HYPERLINK \l "Yan99" 7] proposed a class of histogram based global thresholding techniques called Integral Ratio. It is based on two level thresholding approach in which each pixel of handwritten image can be divided into three parts: foreground, background and fuzzy area between them.We can decide whether a pixel lies in foreground or background on the basis of Native IR and Quadratic IR.

Yang and Yan (2000)8] presented a logical adaptive thresholding method to binarize seriously degraded and very poor quality grayscale document images. This method can deal with complex signal-dependent noise, variable background intensity caused by non uniform illumination, shadow, smear or smudge and very low contrast. It cannot affect useful information.

Sauvola et al (2000) [HYPERLINK \l "Jaa00" 9]presented a new method for adaptive document image binarization, where the page is considered as a collection of subcomponents such as text, background and picture. e evaluation of local threshold is based on estimation of local mean and local standard deviation.

Randolph et al (2001) 10]suggested a binary domain approach that enhances fax documents by directional filter bank enabling edges and contours in the text letters to be smoothed appropriately. It can be used for enhancement in Fax documents. A directional filter bank has been used which is capable for smoothing of edges and contours.

Wu et al (2003) [HYPERLINK \l "Adn03" 11]experimented with a multi-stage global thresholding approach followed by a local spatial thresholding, which works well for simple and complex images of postal envelopes. In first stage global thresholding technique used. In second stage refinement of threshold value can be done.

Gatos et al (2004)12] proposed a digital image binarization scheme for low quality historical documents by five distinct steps: a pre-processing low-pass Wiener filter, a rough estimation of foreground regions using Niblack's approach, a background surface calculation by interpolating neighboring background intensities, a thresholding by combining the calculated background surface with the original image and finally a post-processing quality and connectivity step.

Chen et al (2005) [HYPERLINK \l "Che05" 13] compared global or local thresholding techniques for degraded historical documents images and introduced a local feature thresholding decompose algorithm or document sub-regions using quad-tree decomposition.

Badekas et al (2007) 14]suggested framework for the binarization of typical and corrupted archives for visualization and recognition of content characters by a Kohonen adaptive neural system. He displayed a method for the binarization of content pieces in shading report pictures that contain content and representation profoundly blended with the foundation, in light of a shading lessening.

Gatos, Ioannis Pratikakis et al. (2008) [HYPERLINK \l "Bas082" 15] paper presented a new adaptive approach for document image Binarization. The proposed method is mainly based on the combination of several state of the art binarization methodologies as well as on the efficient incorporation of the edge information of the grey scale source image. For image enhancement morphological operators can be used

DIBCO 200916] is the first International Document Image Binarization Contest organized in the context of ICDAR 2009 conference. Out of 43 algorithms, Institute of Infocomm Research ,Singapore was ranked first by DIBCO. This algorithm deals with background extraction, stroke edge detection, local thresholding & post processing. Second rank was given to the University Pierre et Marie Curie & CMM, France. In this toggle mapping operator concept had been used.

H. Yi. et al (2010) [HYPERLINK \l "YiM10" 17]presented a novel user-assisted approach to reduce the ink-bleed interference found in old documents. This procedure used dual layer Markov Random Field method for recognize all foreground and background pixels of the old bleed documents.

Bolan Su et al (2013)18] presented a new approach called Adaptive Contrast Mapping Method. They proposed a novel document image binarization technique that addresses these issues by using adaptive image contrast. The adaptive image contrast is a combination of the local image contrast and the local image gradient that is tolerant to text and background variation caused by different types of document degradations.

3. PROPOSED WORK

This section describes the proposed document image binarization techniques. Firstly, convert RGB image to gray scale image document and binarized that image and get binary one. Then, image enhancement can be done by using filtering technique. After that morphological operators can be implemented for region based segmentation. To achieve more accurate and better performance Otsu threshold is applied to the output of morphological operators. To maintain the quality of image binarization technique further Sauvola's thresholding is used.

Step 1. RGB to Gray Image Conversion

Step 2. Image Filtering

Step 3. Morphological Operators

Step 4. Thresholding

RGB is a device-dependent color model. The fundamental reason for the RGB shading model is for the sensing, representation and presentation of pictures in electronic systems. To form a color with RGB, three light beams (one red, one green, and one blue) must be superimposed. Each of the three beams is known as a component of that color. [HYPERLINK \l "Cin84" 19]Zero intensity for each component gives the darkest shade and full intensity of each gives a white. When all the three colors have same value for intensity then it gives a grey image as a output

Wiener Filter

Wiener filter, known as "minimum mean square error filter" or "least square error filter". 20]We can say that it is an versatile linear filter which is connected to image locally, by considering the neighborhood picture fluctuation. At the point

when the difference in a image is huge the Wiener filter brings about light neighborhood smoothing, while when the fluctuation is little, it gives an enhanced nearby smoothing. The product of complex quantity with its conjugate is equal to magnitude of complex quantity squared. This result can be described as Wiener Filter.

Gaussian Filter

Gaussian filter is a low pass filter which perform Fourier Transform. It is also known as smoothing frequency filter because they can be used to smooth edges.

Morphological Operators

The meaning of morphology is to describe the properties of the structure and shape of any objects. Morphological operations operate on Sets. In mathematically morphology, sets represent objects in the image. Morphological operations can be used both in preprocessing and final stages for image processing. The output of morphology operation relies on configuration and size of an original picture and Structural Element.

Opening and Closing derived from basic Dilation and Erosion operations. Opening is the dual of shutting (closing) i.e. opening the foreground pixels with a specific structuring element is equivalent to closing the background pixels with the same element. The opening of X by Y is obtained by the erosion of X by Y, followed by dilation of the resulting image by Y. The closing of X by Y is obtained by the dilation of X by Y, followed by erosion of the resulting structure by Y.

Otsu Thresholding

Otsu Thresholding is most successful global thresholding technique. The simplest way for implementation of thresholding is to select an intensity value as threshold level. [HYPERLINK \l "JRP91" 21] The value which is below the threshold level is treated as 0(black) and which is above the threshold level that select as 1(white).

Sauvola's Thresholding

The strategy proposed by Sauvola's et al. is local-variance based method. It is a change on the system proposed by Niblack, particularly when the background contains light composition, enormous variations, re-colored and badly degraded documents. It adapts the contribution of the local mean and standard deviation. When document is dirty or re-colored paper then threshold value is lowered.

4. RESULTS

A. Parameter used

For evaluation there are few parameters that can be use to check the Binarization performance like F-Measure, Peak Signal to Noise Ratio (PSNR), Distance Reciprocal Distortion (DRD) and Misclassification Penalty Metric (MPM).

F-Measure

It is a measure of a test's accuracy. It assumes both the precision p and the recall r of the test to compute the score.

$$F\ Measure = 2.precision.{}^{recall}/_{precision + recall}$$

p is the number of correct results divided by the number of all returned results and r is the number of correct results divided by the number of correct results that should have been returned. F measure achieves its best value at 1 and worst score at 0.

Peak signal-to-noise ratio (PSNR)

PSNR is a term for the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. Because many signals have a very wide dynamic range, PSNR is commonly described in terms of the logarithmic decibel scale. PSNR is most effectively characterized through the mean squared error (MSE).

To evaluate the PSNR, the block first calculates the mean-squared error using the following equation:

$$\mathrm{MSE} = \sum_{MN} \frac{[I_1(m,n) - I_2(m,n)]^2}{M * N}$$

$$\mathrm{PSNR} = 10 \log_{10}\left(\frac{R^2}{\mathrm{MSE}}\right)$$

Misclassification penalty metric (MPM)

The Misclassification penalty metric MPM evaluates the prediction against the Ground Truth (GT) on an object-by object basis. Misclassification pixels are punished by their distance from the ground truth object's border.

$$MPM = \frac{MP_{FN} + MP_{FP}}{2}$$

Where $MP_{FN} = \frac{\sum_{i=1}^{N_{FN}} d_{FN}^{i}}{D}$, $MP_{FP} = \frac{\sum_{j=1}^{N_{FP}} d_{FP}^{j}}{D}$

Distance Reciprocal Distortion

It is used to measure visual distortion in digital binary document images. This measure is based on the reciprocal of distance that is straightforward. This method provides an efficient way to measure distortion in binary document images. It is superior than MSE, SNR or PSNR.

B. Testing on Competition Dataset

The DIBCO 2011 dataset includes eight degraded handwritten documents and eight degraded printed documents. Therefore, total 16 document images. Table 4.1 shows the evaluation results as follows for figure 4.1(a) Binarization results of the sample document image (PR 08).

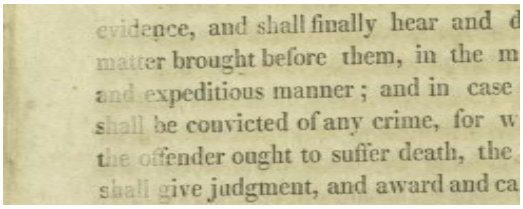

Fig 4.1(a)

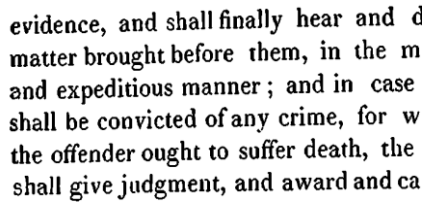

Fig 4.1(b)

Table 1. Evaluation Results Of The Dataset Of DIBCO 2011

Methods	**F-Measure (%)**	**PSNR**	**DRD**	**MPM**
OTSU	82.22	15.77	8.72	15.64
SAUV	82.54	15.78	8.09	9.20
NIBL	68.52	12.76	28.31	26.38
BERN	47.28	7.92	82.28	136.54
GATO	82.11	16.04	5.42	7.13
LMM	85.56	16.75	6.02	6.42
BE	81.67	15.59	11.24	11.40
LELO	80.86	16.13	104.48	64.43
SNUS	85.2	17.16	15.66	9.07
HOWE	88.74	17.84	5.37	8.64
BOLAN	87.8	17.56	4.84	5.17
Proposed	79.34	13.19	5.16	5.40

5. CONCLUSION

This paper describes an image Binarization technique for degraded images by using morphological operators. The proposed method is easy, more reliable and an efficient way. The proposed method makes use of morphological operators then Otsu and Sauvola's thresholding. The output can compare with DIBCO 2011 dataset on the basis of PSNR, F-measure, DRD and MPM.

6. FUTURE SCOPE

As described in previous sections, the proposed method involves several parameters, most of which can be automatically estimated based on the statistics of the input document image. The superior performance of our proposed method can be explained by several factors There are so many parameters are used in our proposed method to check the ability to remove the different kinds of degradation in an input document images. Firstly, our proposed technique makes the document images stable and noise free. Secondly, region based segmentation gives better performance instead of edge based segmentation. Third, our proposed techniques extract foreground from background by using morphological operators. The proposed method is easy, more reliable and an efficient way.

7. REFERENCES

[1] George Nagy, "Twenty years of document image analysis in PAMI," IEEE Transactions on Pattern Analysis and Machine Intelligence22.1, pp. 38-62, 2000.

[2] David L., et al. Milgram, Algorithms and hardware technology for image recognition.: MARYLAND UNIV COLLEGE PARK COMPUTER SCIENCE CENTER, 1978.

[3] Ioannis Pratikakis, and Stavros J. Perantonis. Gatos Basilios, "Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information," , 2008.

[4] and Fatos T. Yarman-Vural Arica Nafiz, "An overview of character recognition focused on off-line handwriting," in Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 31.2 ," , 2001, pp. 216-233.

[5] N. Otsu, "A threshold selection method from gray-level histograms," in IEEE Trans. Systems, Man, and Cybernetics, 1979, pp. 62-66.

[6] P.K. Sahoo, and A.K.C. Wong Kapur J., "A new method for gray-level picture. Thresholding using the Entropy of the Histogram," in Computer Vision Graphics and Image Processing vol. 29, 1985, pp. 273-285.

[7] Y. and C.G. Leedham Solihin, "Integral Ratio: A New Class of Global Thresholding Techniques for Handwriting images," in IEEE Trans. on PAMI, vol. 21, 1999, pp. 761-768.

[8] Yibing Yang and Hong Yan, "An adaptive logical method for binarization of degraded document images Pattern Recognition," in Pattern Recognition Society, Elsevier Science , vol. 33, no. 5, 2000, pp. 787-807.

[9] and Matti Pietikäinen Sauvola Jaakko, "Adaptive Document Image Binarization," in Pattern Recognition 33.2, 2000, pp. 225-236.

[10] M Randolph T. Smith, "Enhancement of fax documents using a binary angular representation," in in Proceedings of, Int Symp on Intelligent Multimedia, Video and Speech Processing, Hong Kong,Cjina, 2001.

[11] Wu Sue Adnan Amin, "Automatic Thresholding of Gray-level Using Multi-stage Approach," in in Proceedings of the 7th International Conference on Document Analysis and Recognition (ICDAR 2003),IEEE, 2003.

[12] S. Perantonis Gatos B I. Pratikakis, "An Adaptive Binarization Technique for Low Quality Historical Documents," in Document Analysis Systems VI, vol. 3163, , 2004.

[13] Chen Y. and G. Leedham, "Decompose algorithm for thresholding degraded historical document images," in IEEE Proc.-Vis. Image Signal Process vol. 152, , December 2005.

[14] Chen Y. and G. Leedham, "Document binarization using Kohonen," in IET Image Process, 2007, pp. 67-85.

[15] Ioannis Pratikakis, and Stavros J. Perantonis. Gatos Basilios, "Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information," in Pattern Recognition, vol. ICPR 2008. 19th International Conference on. IEEE, 2008, 2008.

[16] Konstantinos Ntirogiannis, and Ioannis Pratikakis Gatos Basilios, "ICDAR 2009 Document Image Binarization Contest (DIBCO 2009)," in ICDAR, vol. 9, 2009.

[17] Michael S. Brown, and Dong Xu. Huang Yi, "User-assisted ink-bleed reduction," in Image Processing, IEEE Transactions , oct 2010, pp. 2646-2658.

[18] Shijian Lu, and Chew Lim Tan Su Bolan, "Robust document image binarization technique for degraded

document images," in Image Processing, IEEE Transactions on 22.4 , 2013, pp. 1408-1417.

[19] et al. Goral Cindy M., "Modeling the interaction of light between diffuse surfaces," in ACM SIGGRAPH Computer Graphics, vol. 18, 1984.

[20] and Rendong Zhang Ke Li, "Multiscale wiener filtering method for low-dose CT images," in Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference , vol. 1 IEEE.

[21] J. R. Parker, "Gray level thresholding in badly illuminated images," in IEEE Transactions on Pattern Analysis and Machine Intelligence 13.8 , 1991, pp. 813-819.