International Journal of Emerging Trend in Engineering and Basic Sciences (IJEEBS)
ISSN (Online) 2349-6967
Volume 2 , Issue 1(Jan-Feb 2015), PP314-318

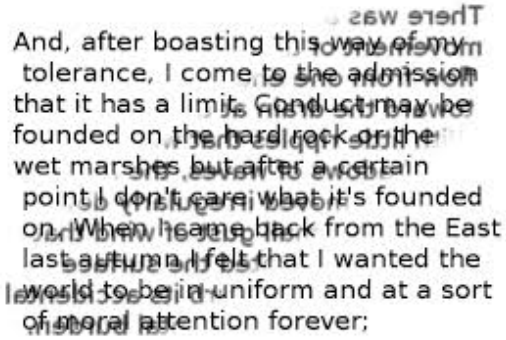
A Survey on Document Image Binarization Techniques

Ms.Supriya Sunil Lokhande[1], Prof.N.A.Dawande[2]
[1]*(Department of E&TC, Dr.D.Y.Patil College of Engineering, India)*
[2]*(Department of E&TC, Dr.D.Y.Patil College of Engineering, India)*

Abstract :- In this paper a comprehensive survey has been conducted on document image binarization techniques. The main objective of this paper is to evaluate recent developments in document image binarization for degraded document images. This paper describes some state -of -the -art image binarization techniques, compares them with the help of evaluation performance measures which are widely used for document image analysis and recognition. The principal aim of this paper is to find the best method among some eminent existing methods.
Keywords: - background, binarization, foreground, image contrast, segmentation

I. INTRODUCTION

Document image binarization is performed in the preprocessing stage for document analysis. It is used to perform optical character recognition. Several historical documents suffer from degradation as they begin to age over a period of time. Document images suffer from uneven illumination, image contrast variation bleed-through and smear. Two of these effects have been illustrated in Fig 1.and 2 below



.

Fig. 1.Documents suffering from bleed through

As illustrated in Fig.1 the text suffers from bleed through where the ink from the other side seeps through the front.

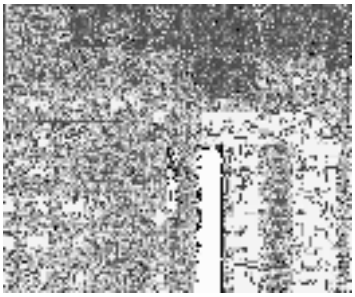


Fig. 2 the document image is degraded by smear making the text in it almost indecipherable

Thus, these different types of document degradation result in document thresholding error and makes degraded document image binarization a formidable challenge for many state of the art techniques.

The main objective of any document image binarization technique is to segment the foreground text from the background text. The thresholding of degraded documents is a major challenge.Thresholding refers to the conversion of gray-scale image to a binary image .Some thresholding techniques have been elucidated below.

II. REVIEW OF DOCUMENT IMAGE BINARIZATION

2.1 OTSU'S METHOD N.Otsu(1979)[8]

In Otsu's method clustering-based image thresholding, is used for, the reduction of a gray level image to a binary image. This algorithm tries to reduce combined spread (intra class variance) by assuming that the image contains two classes of pixels. It assumes that an image follows a bimodal histogram i.e. it contains foreground and background pixels .It then calculates the optimum threshold separating the two classes to ensure that so its combined spread is minimal. This method gives acceptable results when the pixels in each class are close to each other.The limitations of this method are that many degraded documents do not have clear bimodal patterns. Also another limitation is that minimization of intra class variances maximizes between class scatter. [3]

2.2. Niblack's Method W.Niblack (1986) [4]

By sliding a rectangular window over the grey level image Niblack's algorithm calculates a pixel wise threshold.

The threshold T is computed by using the mean m and standard deviation s, of all the pixels in the window, and is denoted as:

$$T = m + k \times s$$

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 defines the quality of binarization [9].The limitation of Niblack's method is that the resulting binary image suffers from a great amount of background noise especially in areas without text. [10]

2.3. Sauvola's Method J. Sauvola and M. Pietikainen (2000) [5]

In this method the page is considered as a collection of subcomponents such as text, background and picture. To define a threshold for each pixel of the background and pictures a soft decision method is used. To define a threshold for each pixel of textual and line drawing areas a text binarization method is used. Finally the results of these algorithms are combined. [5].Although this method solves the problem posed by Niblack's approach but in many cases the characters become extremely thinned and broken. [10]

2.4. Bernsen's Method J. Bernsen (1986)

In this method 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 greater than a threshold, the pixel will be classified into text or background by comparing with the mean of $I_{max}(i,j)$,and $I_{min}(i,j)$.If the local contrast C(i,j) is smaller than the threshold then the pixel is set as background. This method is simple. But the limitation of this method is that it does not work properly on degraded document images with a complex background. [2]

2.5 Background Estimation Method

This method is an improvement over BERN'S Method and handles the documents with a complex background well. In this method the local image contrast introduces a normalization factor. This normalization factor compensates for the image variation within the document background. Here the local image contrast is evaluated as follows:-

$$C(i, j) = \frac{I_{max}(i, j) - I_{min}(i, j)}{I_{max}(i, j) + I_{min}(i, j) + €} \qquad (1)$$

where € is a positive but infinitely small number that is added in case the local maximum is equal to 0.[2] 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 acts as a normalization

factor that lowers the effect of the image contrast and brightness variation. For image pixels within bright regions around the text stroke boundary, the denominator is large, which neutralizes the large numerator and accordingly results in a relatively low image contrast. But for image pixels within dark regions around the text stroke boundary, the denominator is small, which compensates the small numerator and accordingly results in a relatively high image contrast. [7]

As a result, the contrasts of image pixels (lying around the text stroke boundary) within both bright and dark document regions converge close to each other and this facilitates the detection of high contrast image pixels lying around the text stroke boundary.[7]

III. PERFORMANCE METRICS

2.1 F-Measure

$$\text{F-Measure}=\frac{2\times \text{Recall}\times \text{Precision}}{\text{Recall}+\text{Precision}}$$

$$\text{where Recall}=\frac{TP}{TP+FN},\quad \text{Precision}=\frac{TP}{TP+FN} \qquad (2)$$

TP, FP, FN denote the True positive, False positive and False Negative values, respectively. [1]

2.2 PSNR

$$\text{PSNR}=10\log\frac{C^2}{\text{MSE}} \qquad (3)$$

$$\text{Where MSE}=\frac{\sum_{X=1}^{M}\sum_{y=1}^{N}(I(x,y)-I'(x,y)^2)}{MN}$$

PSNR is a measure of how close is an image to another. Therefore, the higher the value of PSNR, the higher is the similarity of the two images. .C is the difference between foreground and background [1]

2.2 . Distance Reciprocal Distortion Metric (DRD)

The Distance Reciprocal Distortion Metric (DRD) has been used before to measure the visual distortion in binary document images [1]. It properly correlates with the human visual perception and it measures the distortion for all the ***S*** flipped pixels as follows

$$\text{DRD}=\frac{\sum_{k=1}^{S}\text{DRD}_k}{\text{NUBN}} \qquad (4)$$

where DRD_k is the distortion of the kth flipped pixel and it is calculated using a 5x5 normalized weight matrix W_{Nm} as defined in [13]. *DRDk* equals to the weighted sum of the pixels in the 5x5 block of the Ground Truth *GT* that differ from the centered k^{th}flipped pixel at (x,y) in the Binarization result image *B* (Eq. 5).

$$\text{DRD}_k=\sum_{i=-2}^{2}\sum_{j=-2}^{2}\left|\text{GT}_k(i,j)-\text{B}_k(x,y)\right|\times \text{W}_{Nm}(i,j) \qquad (5)$$

Finally, NUBN is the number of the non-uniform (not all black or white pixels) 8x8 blocks in the GT image.[1]

2.3. 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 penalized by their distance from the ground truth object's border

$$MPM = \frac{MPFN + MPFP}{2} \qquad (6)$$

where $MPFN = \sum_{i=1}^{NFN} d_{FN}^{i} / D$, $MPFP = d_{FP}^{i} / D$

d_{FN}^{i} and d_{FP}^{i} denote the distance of the i th false negative and the j th false positive pixel from the contour of the GT segmentation. The normalization factor D is the sum over all the pixel-to-contour distances of the GT object. A low MPM score denotes that the algorithm is good at identifying an object's boundary. [1]

IV.RESULTS OBTAINED AFTER TESTING ON DATASET OF DIBCO 2011

TABLE I
EVALUATION RESUTS FOR DATSET OF DIBCO 2011

Methods	F-Measure %	PSNR	DRD	MPM
Otsu [8]	82.22	15.77	8.72	15.64
Sauvola [5]	82.54	15.78	8.09	9.20
Niblack [4]	68.52	12.76	28.31	26.38
Bernsen [6]	47.28	7.92	82.28	136.54
BE [7]	81.67	15.59	11.24	11.40
AC [2]	87.8	17.56	4.84	5.17

From the above table it can be inferred that the last method has the highest value of F-Measure and PSNR. This method also has the lowest value of DRD and MPM.

V.CONCLUSION

This paper presents a survey of some of the widely used document image binarization technique .It compares the various existing techniques with each other using evaluation measures such as F-Measure, PSNR.DRD, MPM.The inference that can be drawn from this comparison is that all the methods listed above have some advantages and some drawbacks. The adaptive contrast method outperforms most of the document image binarization methods in terms of F-measure, PSNR, DRD and MPM.

VI. ACKNOWLEDGEMENTS

I would like to thank Prof.N.A Dawande for his valuable guidance and deep insight.

REFERENCES

Journal Papers:

[1] 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.*

[2] Bolan Su, Shijian Lu, and Chew Lim Tan, Robust Document Image Binarization Technique for Degraded Document Images, *IEEE transactions on image processing, vol. 22, no. 4, April 2013*

[3] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electron. Imag, vol. 13,*

[4] W. Niblack, An Introduction to Digital Image Processing. Englewood Cliffs*, NJ: Prentice –Hall, 1986 no 1, pp. 146-165, Jan 2004*

[5] J. Sauvola and M. Pietikainen, *"Adaptive document image binarization,"Pattern Recognit., vol. 33, no. 2, pp. 225–236, 2000.*

[6] J. Bernsen, "Dynamic thresholding of gray-level images," in *Proc. Int. Conf. Pattern Recognit., Oct. 1986, pp. 1251–125.*

[7] S. Lu, B. Su, and C. L. Tan, "Document image binarization using background estimation and stroke edges," *Int. J. Document Anal. Recognit., vol. 13, no. 4, pp. 303–314, Dec. 2010.*

[8] N. Otsu, ''A threshold selection method from gray level histograms,*'' IEEE Trans. Syst. Man Cybern.* ***SMC-9****, 62–66 1979!*

[9] B.Gatos, I. Pratikakis, S.Perantonis, 'An Adaptive Binarization Technique for Low Quality Historical Documents ', *Springer-Verlag pp102-113, 2004*

[10] B. Gatos, I. Pratikakis, and S. Perantonis, "Adaptive degraded document image binarization," *Pattern Recognit., vol. 39, no. 3, pp 317–327, 2006*