International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-04, Issue-10, Jan 2019

Performance Analysis of Degraded Document Image Binarization Techniques

Usha Rani, Assistant Professor, University College Ghanaur, Patiala, Punjab, India, usha_gupta7@yahoo.co.in

Dr. Amandeep Kaur, Associate Professor , Central University of Punjab, Bathinda, India, aman_k2007@hotmail.com

Dr. Gurpreet Singh Josan, Associate Professor Department Of Computer Science, Punjabi University, Patiala, India, josangurpreet@pbi.ac.in

Abstract - The archival documents are very important for historical study. Since the original manuscripts are very rare and are not available in every library or museum, so such documents are scanned and digitized using Document Analysis System to increase their availability and to facilitate and encourage historical research. Binarization is the first and most important stage of the Document Analysis System that converts the gray or colored images into bilevel form that is the text is represented by black color pixels and background by white colored pixels or vice versa. The results of the subsequent stages of the Document Analysis System totally depend upon the binarization result. But the original historical documents face several kinds of degradations due to aging, chemical procedure of paper fabrication and storage conditions over the years. Due to the presence of degradations, binarization of degraded documents is not so easy task because these images possess extensive background noise, nonuniform illumination, smear, stains etc. This paper discusses the various existing document binarization techniques and presents the comparison between them.

Keywords- Binarization, Degraded Documents, Document image Processing, Thresholding.

Introduction

Document image binarization is the important pre-processing stage in the document image analysis and recognition pipeline, results of which seriously affects the accuracy of other phases of OCR (Optical Character Recognition). The main objective of any binarization technique is to segment the foreground text from the background. It converts the document image into bi-level form in such a way that the foreground information is represented by black pixels and the background by white ones. The main purpose of the binarization is to decrease the computation load and increase in the efficiency of the system.

Though document image binarization has been studied for many years and several binarization techniques have been proposed and developed in the literature. But selection of the binarization method is still a challenging task because document images suffer from various kinds of degradations such as ink bleed through, uneven illumination, smear, strain, contrast variation, degradations due to humidity etc.

Sezgin et al. [12] classified the binarization techniques into six categories based on the type of threshold technique used: histogram based, clustering –based entropy based, attribute based, spatial binarization, and local adaptive methods. While Wen et al. [14] divided these methods into three major categories: clustering based, threshold based, and hybrid methods.

Binarization techniques based on threshold are mainly categorized into global and local methods. The global binarization methods [1], [2], [3], [4], [5] use single threshold value for the entire image. The global methods are fast and perform well in case of good quality and well contrasted images. But in case of images with degradations such as nonuniform illumination, stains and smear global methods do not produce good results. Local binarization methods [6], [7], [8], [9], [10] estimate a different threshold value for each image pixel depending upon features of its neighborhood. These methods are able to extract text even from severely degraded images, but are slow because computation of image features from local neighborhood is done for each pixel in the image.

The rest of the paper is organized as follows: Section 2 discusses the existing binarization methods, Section 3

 DOI : 10.18231/2454-9150.2018.1339

describes the experimental set up and evaluation measures. The results are presented and discussed in section 4 and conclusions are drawn in section 5.

II. EXISTING BINARIZATION METHODS CONSIDERED FOR THIS WORK

The binarization produce bi-level image of the gray level or colored document image i.e. only two intensity levels having black (or white) text on white (or black) back-ground. The main objective of binarization is to reduce unnecessary information leaving only the useful data for further processing. The threshold based binarization techniques categorized as global and local are discussed below:

A. Global Binarization Techniques

Global binarization techniques use a single threshold value for the entire image. Otsu [1], Kapur et al. [2], Kittler and Illingworth [3], Brink and Pendock [5] are well known global binarization methods which are briefly discussed.

Otsu [1] is a histogram based global binarization method that separates the gray level histogram into two classes: foreground and background. Then it estimates the optimum threshold value for these two classes such that inter class variance is minimal or intra class variance is maximal.

Kapur et al. [2] method of binarization is based on entropy concept. It divides the probability distribution of gray levels into two classes, foreground and background such that sum of entropies associated with these distributions is maximized to obtain the maximum information between object and background in the image. The discrete value that maximizes the sum of entropies is the threshold value.

Kittler and Illingworth [3] presented the binarization algorithm based on fitting of the mixture of Gaussian distributions and it changes the binarization problem to a minimum error Gaussian density fitting problem.

Brink and Pendock [5] used the cross entropy to compute the optimal threshold.

Among all the global binarization techniques, Otsu's method of binarization is most successful for the OCR systems due to its computational efficiency and effectiveness. But in case of documents containing degradations such as non-uniform illumination, stains and smear, none of the existing global binarization method gives good result. Fig. 1 shows the binarization results for the good quality image and Fig. 2 and Fig. 3 shows the binarization results of the degraded images with both global as well as with local methods.

B. Local Binarization Techniques

Local binarization techniques use a different threshold value for each image pixel depending upon the features of its neighborhood. These methods produce better results in case of document images containing non-uniform illumination and other degradations such as smear, stains, bleed through, discoloration, poor contrast etc. as these can handle intensity variation more effectively than the global methods. Most of the local binarizati+on methods in the literature are sliding window based. In sliding window methods, a fixed size window is moved, pixel to pixel over the image. Then threshold value for the center pixel is computed using the features of the pixels within the window. Bernsen [6], Niblack [7], Sauvola and Pietaksinen [8], Wolf and Jolion [9] and Khurshid et al. [10] are some of the existing sliding window based local binarization methods which are discussed below:

Bernsen et al. [6] is the local contrast method that uses the mean and contrast information to compute threshold value over a local region. In this if the local contrast (I_{max}-I_{min}) is above or equal to the user provided contrast threshold, then the threshold is set at the mid- grey value (i.e. the mean of minimum and maximum grey values in the local window). If local contrast is less than the user provided contrast threshold the neighborhood is considered to belong only to single class (i.e. foreground or background depending on the mid-grey value). Authors recommended the window size equal to 15. This method is fairly fast because it do not need any complex computations.

Niblack [7] method calculates the threshold for each pixel by siding the rectangular window over the gray-level document image. The threshold value T for the center pixel of the window is computed by local mean and local standard deviation of the gray values in the window as follows:

$$T(x,y) = m(x,y) + k \times \sigma(x,y) \ldots\ldots (1)$$

where, m (x, y) and σ (x, y) are the local mean and local standard deviation of the pixels within the local window. The value of 'k ' controls the amount of text region inside the local window. Varying illumination level can be handled effectively by using small window size but if the window size is too small, it will not be able to eliminate noise in the gray image. Value of 'k' and size of the window are the parameters that need to be decided based on the document image. Recommended size of the window is 15×15 and k=-0.2.

Sauvola and Pietaksinen [8] proposed a modification of Niblack's algorithm using dynamic range of image gray value standard deviation, 'R' to compute threshold as follows:

DOI : 10.18231/2454-9150.2018.1339

$$T(x,y) = m(x,y)\left[1 + k\left(\frac{s(x,y)}{R} - 1\right)\right] \ldots\ldots(2)$$

Here k=0.5 and R=128

Wolf and Jolion [9] addressed the issues in Sauvola's algorithm. It proposes to normalize the contrast and the mean gray value of the image. The threshold is computed as:

$$T(x,y) = (1-k) \times m(x,y) + k \times M + k\frac{s(x,y)}{R}(m(x,y) - M) \ldots(3)$$

where k=0.5, M is the minimum gray value of image and R is set to the maximum gray value standard deviation obtained over all the local neighborhood (windows).

Rais et al. [10] like the Sauvola's method this method is based on the Niblack image thresholding method, which use the global and local characteristics adaptively. It does not entirely depend upon the local statistical characteristics like the Niblack's method, but also considers the global statistics. It calculates the value of parameter 'k' at run time for each pixel and binarization is done using the Niblack formula.

$$k = -0.3 \times \frac{m_g(x,y) \times s_g(x,y) - m_l(x,y) \times s_l(x,y)}{max(m_g(x,y) \times s_g(x,y), m_l(x,y) \times s_l(x,y))} \ldots\ldots\ldots(4)$$

Here, m_g (x, y) and s_g (x, y) are the global mean and global standard deviation. While m_l (x, y) and s_l (x, y) are the local mean and local standard deviation respectively. In this method value of k will remain in the range of 0.3 and -0.3.

Khurshid et al. [11] proposed an algorithm called NICK which drives its thresholding formula from the basic Niblack algorithm.

$$T(x,y) = m(x,y) + k\sqrt{\frac{(\sum[I(x,y)]^2 - [m(x,y)]^2}{NP}} \ldots\ldots\ldots(5)$$

Where, k is the Niblack factor, m (x, y) local mean gray value, I (x, y) is the intensity value of center pixel in the window, NP is the number of pixels. This method uses two parameters k and window size 'w', which needs to be determined manually. Value of k is recommended as -0.2 and window size is optimally set to 19×19.

III. EXPERIMENTAL SETUP

The above discussed global and local binarization techniques are implemented and tested on images taken from the Internet and DIBCO datasets.

To compare the results of different binarization techniques quantitatively, evaluation metrics F-measure, Peak to signal ratio (PSNR) and Negative rate metric (NRM) adopted from DIBCO contest are used. These metrics are ground truth image based and measures how well an algorithm can retrieve the desired pixels. All the DIBCO datasets include the variety of real degraded document images and their corresponding semi-automatically generated ground truth images, so to evaluate the results quantitatively, test images from DIBCO datasets are used.

The F-measure (FM) is the harmonic mean of precision and recall. It represents the overall binarization accuracy. Precision represent the binarization noise, and recall represents text body accuracy. The high value of precision means binarized image is less noisy and high recall value indicates that desired pixels are retrieved more efficiently.

$$FM = \frac{2 \times Recall \times Precision}{Recall + Precision},$$

$$Precision = \frac{C_{TP}}{C_{FP} + C_{TP}}, \quad Recall = \frac{C_{TP}}{C_{FN} + C_{TP}}$$

C_{TP} the number of true positive pixels, C_{FP} the number of false positive pixels and C_{FN} the number of false negative pixels.

PSNR measures the closeness of one image to another. A higher value of this indicates the greater similarity between two images.

$$PSNR = 10 \times log\left(\frac{C^2}{MSE}\right),$$

where C is the difference between foreground and background and its value is set equal to one for binary images. MSE denotes the mean square error.

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

I (x, y), the binarized image pixel value compared with I'(x, y) which is the ground truth pixel value at the same pixel coordinates.

NRM measures the pixel mismatch rate between the ground truth and the resulting image.

$$NRM = \frac{A+B}{2},$$

$$A = \frac{C_{FN}}{C_{FN}+C_{TP}}, B = \frac{C_{FP}}{C_{FP}+C_{TN}}$$

Here C_{FN}, C_{TP}, C_{FP}, C_{TN} denote the number of false negative, true positive, false positive and true negative pixels respectively. The binarization quality is better for lower NRM.

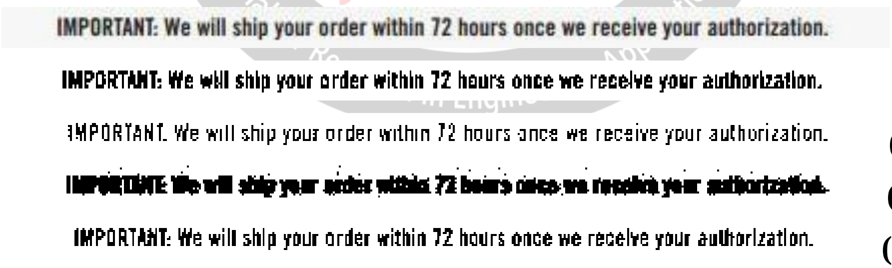
IV. OBSERVATIONS

The above discussed binarization methods are tested on images from the Internet and the DIBCO datasets. Since it is not possible to include all the visual results here, we have shown some results in figures 1 to 3. Quantitative

 DOI : 10.18231/2454-9150.2018.1339

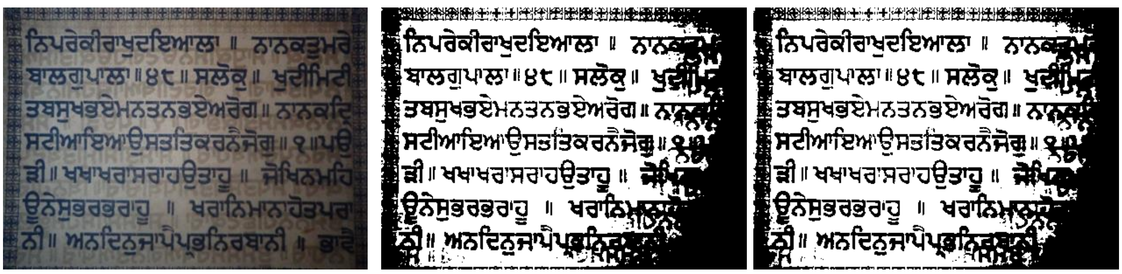
results of the image in figure 3 using all above discussed methods are tabulated in table 1. Summary of the results is given as follows:

1) Fig. 1 shows the binarization results of a good quality image using global methods. It is clear that global methods are good for images in good condition.
2) The results of Niblack and Rais et al. methods are almost same because both use the same formula to compute threshold value. But the execution time of Rais et al. method is more compared to Niblack because it computes value of parameter 'k' at run time for every pixel in the image, while value of 'k' is fixed in Niblack's method.
3) Fig. 2 and Fig. 3 shows the binarization results of degraded images using all methods global as well as local. It shows that all the local methods are good in extracting text from the degraded images compare to global techniques. Local methods Bernsen, Niblack, Rais et al. and Wolf and Jolion produce noise in non-text regions which can be reduced by setting window size parameter and using noise elimination techniques.
4) Sauvola's and NICK methods eliminate the black noise problem and produce better results compared to all methods discussed here.
5) In case of light and low contrast images, Sauvola's method results are fainted with broken text compared to NICK.
6) NICK overcomes the black noise problem of all methods and also low contrast problem of the Sauvola's method.
7) Table 1. shows the quantitative results of input image shown in Fig. 3 with all binarization techniques. The high value of precision means less noisy results and high value of recall means text is extracted properly. As F-measure is harmonic mean of recall and precision, so for high value of F-measure means better binarization result.
8) It shows that recall value is good in case of all the methods except Kapur's method. Kapur's method produced almost blank results that is why precision is 100% and recall is only 0.03 %.
9) In case of all methods except Sauvola's and NICK the value of precision is very low, which means that methods extract background pixels as foreground and so cause noise. In Sauvola's and NICK methods, the values of both precision as well as recall are high, so they produced better results compared to other methods.
10) Table 1. shows that the value of F-measure and PSNR is the highest and the value of NRM is the lowest in case of NICK method, so the NICK method produced the best binarization result for image in Fig. 3.
11) In all local methods, the results will be very much dependent upon the window size. To get the best results, the size of the window needs to be set for document image under consideration which is a tedious task.
12) As the document images can contain the variable text sizes, the fixed window size will not work well.



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

Fig. 1. Binarization results of Global methods with input image without degradations (a) Input image, (b) Otsu, (c) Kapur's , (d) Kittler, (e) Brink and Pendock.



(a) (b) (c)

 DOI : 10.18231/2454-9150.2018.1339

Fig. 2. Binarization results of different methods with input image with degradations (a) Input image, (b) Otsu, (c) Kapur's, (d) Kittler, (e) Brink and Pendock, (f) Bernsen, (g) Niblack, (h) Sauvola, (i) Rais, (j) Wolf, (k) NICK

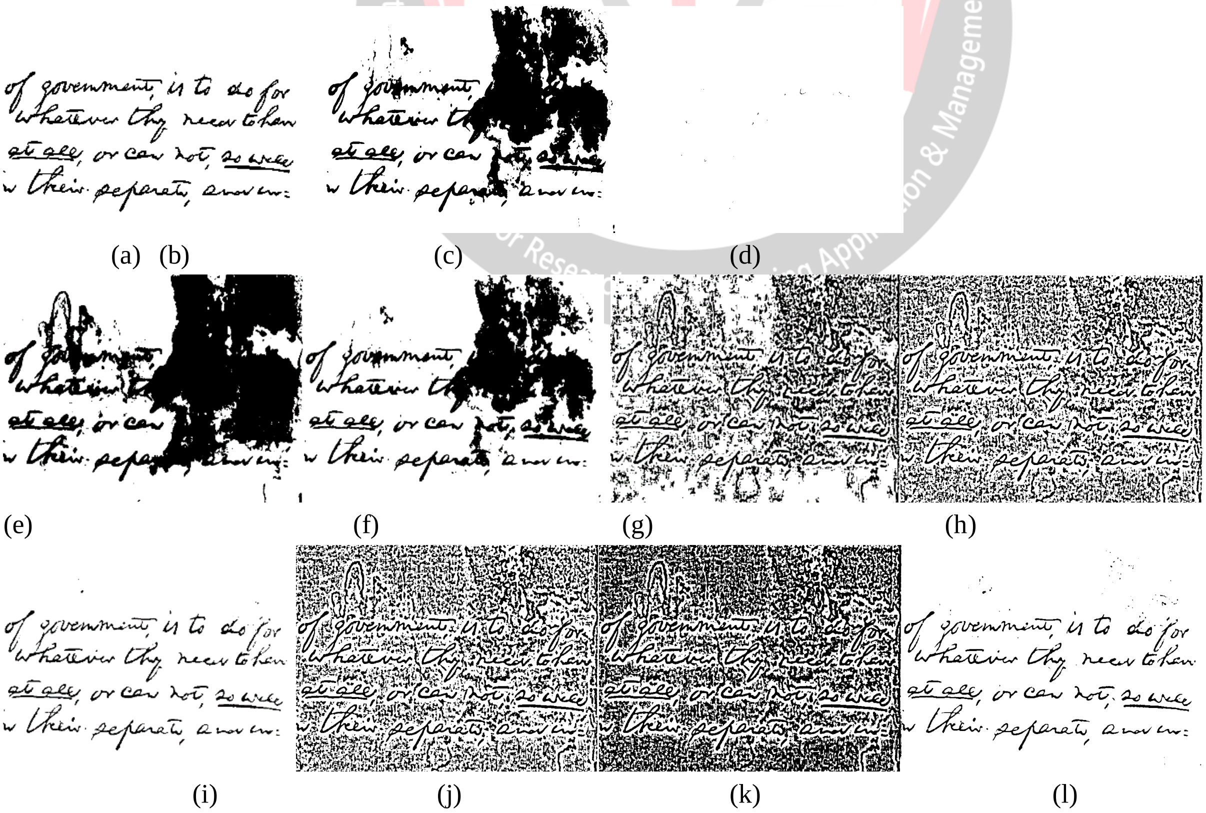


Fig. 3. Binarization results of different methods with input image with degradations (a) Input image, (b) Ground truth image, (c) Otsu, (d) Kapur's, (e) Kittler, (f) Brink and Pendock, (g) Bernsen, (h) Niblack, (i) Sauvola, (j) Rais, (k) Wolf, (l) NICK

 DOI : 10.18231/2454-9150.2018.1339

Table1. Evaluation parameters for the image in Fig. 3

Method	Precision (%)	Recall (%)	F-measure (%)	PSNR	NRM ($\times 10^{-2}$)
Otsu	25.52	98.71	40.56	6.73	14.5
Kapur	100	0.3	0.6	11.36	49.8
Kittler	17.65	99.95	30.00	4.66	18.49
Brink&Pendock	28.91	97.39	44.59	7.5	13.78
Bernsen	21.6	96.4	35.3	5.86	15.64
Niblack	19.05	91.4	31.52	5.4	19.7
Sauvola	98.5	72.9	83.79	16.84	13.6
Wolf	13.9	96.7	24.36	3.6	25.3
Rais	19.03	90.4	30.21	5.3	19.9
NICK	96.9	74.54	**84.28**	**16.9**	**12.8**

V. CONCLUSIONS

This paper presents the comparative analysis of existing global and local binarization techniques for degraded document images. The results are compared visually and quantitatively in terms of F-measure, PSNR and NRM. The local binarization techniques are found to be better than the global techniques. Nick method outperforms among local binarization techniques. The drawback of local binarization techniques is that the results are very much dependent upon the window size. Also, in case of document images with variable character sizes, fixed window size will not work well.

REFERENCES

[1] N. Otsu, “A threshold selection method from gray level histograms”, *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62- 66, 1979.

[2] J.N. Kapur, P.K. Sahoo and, A.K.C. Wong, “A new method for gray-level picture thresholding using the entropy of the histogram”, *Graphical Models and Image Processing*, vol. 29, no. 3, pp. 273- 285, 1985.

[3] J. Kittler, J. Illingworth, “Minimum error thresholding”, *Pattern Recognition*, vol. 19, no. 1, pp. 41-47, 1986.

[4] A. S. Abutaleb, “Automatic thresholding of gray-level pictures using two-dimensional entropy”, *Computer Vision Graphics and Image Processing*, vol. 47, no. 1, pp. 22-32, 1989.

[5] D. Brink, N.E. Pendock, “Minimum cross-entropy threshold selection”, *Pattern Recognition*, vol. 29, no. 1, pp. 179-188, January1996.

[6] J. Bernsen, “Dynamic thresholding of gray level images”, *Proceedings of IEEE International Conference on Pattern Recognition*, pp. 1251-1255, 1986.

[7] W. Niblack, *An Introduction to Image Processing*. Prentice-Hall, Englewood Cliffs, NJ, pp. 115-116, 1986.

[8] J. Sauvola, M. Pietaksinen, “Adaptive document image binarization”, *Pattern Recogntion Letters*, vol. 33, no. 1, pp. 225-236, 2000.

[9] C. Wolf, J.M. Jolion, “Extraction and recognition of artificial text in multimedia documents”, *Pattern Analysis and Applications*, vol. 6, no. 4, pp. 309-326, 2003.

[10] N. Rais, M.S. Hanif, I.A. Taj, “Adaptive thresholding technique for document image analysis”, *8th IEEE International Multi-topic Conference*, pp.61-66, 2004.

[11] K. Khurshid, I. Siddiqi, C. Faure, N. Vincent, “Comparison of niblack inspired binarization methods for ancient document”, *16th IEEE International Conference on Document Recognition and Retrieval*, pp. 1-10, 2009.

[12] M. Sezgin, B. Sankur, “Survey over image thresholding techniques and quantitative performance evaluation”, *Electronics Imaging*, vol. 13, no.1, pp.146–165, 2004.

[13] S. Chen, D. Li, “Image binarization focusing on objects”, *Neuro computing”,* vol. 69, pp. 2411-2415, 2006.

[14] J. Wen, S. Li, J. Sun, “A new binarization method for nonuniform illuminated documented images”, *Pattern Recognition*, vol. 46, pp. 1670-1690, 2012.

 DOI : 10.18231/2454-9150.2018.1339