

Comparative Study and Image Analysis of Local Adaptive Thresholding Techniques

M.Chandrakala

Department of ECE, MGIT, Hyderabad, Telangana, India

ABSTRACT: Thresholding is a simple but effective technique for image segmentation. In this paper, a general locally adaptive thresholding methods using neighborhood processing is presented. Local adaptive techniques are more effective in eliminating both uneven lighting disturbance, noise and ghost objects. In order to demonstrate the effectiveness, locally adaptive thresholding methods namely Niblack, Sauvola, Wolf's, Darek Bradley, Nick's thresholding had been implemented with real world images, printed text document and hand written text document images. Threshold based segmentation methods had been analyzed quantitatively and qualitatively.

KEYWORDS: Image thresholding, Image segmentation, window size, Misclassification Error, False Positive Rate, False Negative Rate.

I. INTRODUCTION

Thresholding is an important technique in image segmentation and machine vision applications. A survey of thresholding methods and their applications exist in literature [1]. Thresholding techniques can be divided into global and local thresholding depending on the number of thresholds required to be detected. Global thresholding selects a single threshold value from the histogram of the entire image. The Otsu [2] method is good for thresholding a histogram with bimodal or multimodal distribution. The Otsu method, however, fails if the histogram is unimodal. The Otsu method works for images whose histograms show clear bimodal distributions. Local method is an adaptive selection method in which threshold value is determined over a small region.

Local threshold method performs better in case of badly illuminated images and document image. In local adaptive thresholding processing time is dependent on region statistics. Niblack's [3] method calculates the mean and standard deviation over a small window. Sauvola's algorithm [4] claims to improve Niblack's method by computing the threshold using the dynamic range of image gray-value standard deviation. Wolf [5] proposed to normalize the contrast and the image gray-level mean. Darek Bradley et al. [6] proposed a modifying thresholding technique to consider spatial variation in illumination, Khurshid [7] proposed modified version to Niblack's method is known as Nick's algorithm. It shifts the thresholding value down to overcome the issue of black noise in Niblack method and to give solution to

low contrast image difficulty in Sauvola's method. Thresholding methods can remove noise but also preserve edges[8]

The paper is structured as follows. In Section 2, provides a short review of a few thresholding methods that are selected for comparison purpose namely the ones developed by Sauvola's algorithm [4], Wolf [5], Niblack [3], Darek Bradley [6] and Nick's method. Section 3 discusses experimental results for the selected adaptive threshold methods. Finally, Conclusion is drawn in section 4.

II. REVIEW OF SELECTED THRESHOLDING TECHNIQUES

In this section, we briefly review the Otsu method for selecting optimal image threshold. Since the thresholding is done once for the whole image, one may lose certain local characteristics. Locally adaptive threshold based methods namely Niblack [3], Sauvola's algorithm [4], Wolf's [5], Darek Bradley method [6] and Nick estimate a different threshold for each pixel according to the grayscale information of the neighbouring pixels. In local thresholding, the thresholding value is varied based on local content of image.

A. Niblack's Technique

Niblack[3] is a local thresholding algorithm calculates the threshold value to the local mean and the local standard deviation by sliding window around each pixel location. The local threshold at any pixel (x, y) is calculated as

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

Here, $m(x, y)$, $\sigma(x, y)$ are the local mean and local standard deviation respectively. The quality of thresholded image determines with the value of 'k' and the size of the sliding window. The size of the local region (window) is dependent upon the application. Segmentation result of binary image depending on size of window. Document images window size is required to change depending on character size. Some images have local neighbourhood contrast quite low. In that case threshold value $T(x, y)$ goes below the mean value

thereby removing dark regions of the background. The value of the weight 'k' is used to control and adjust the effect of standard deviation due to objects features. value of k gives slim and broken strokes. Niblack algorithm suggests the value of 'k' to be -0.2. Niblack fails to adapt large variation in illumination, especially in the document images.

B. Sauvola Technique

Sauvola's [4] proposed a method by computing the threshold value to the local mean and the local standard deviation by sliding window around each pixel location. The local threshold at any pixel (x, y) is calculated as

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

Where, m (x, y), σ (x, y) are the local mean and local standard deviation respectively. The typical suggested value for

k = 0.5 and R = 128. The value of k and window size gives large effect on quality of image. Result of thresholded image degraded gradually, when gray values of foreground and background pixels are close to each other. Sauvola method performs very well compared to Niblack algorithm in images where foreground text pixel have near '0' gray value and back ground non text pixels have approximately '255' gray value. Background noise problem in Niblack method is solved by Sauvola approach.

C. Wolf's Technique

Christian wolf [5] binarization technique calculates image contrast, gray level mean and standard deviation with in local window and over whole image. The shifting window is placed over each pixel of the image and neighboring pixels values are considered for calculating m (mean) and σ (standard deviation). The local threshold at any pixel (x, y) is calculated as

$$T(x, y) = (1 - k) * m + k * M + k * \frac{\sigma}{R} (m - M) \quad (3)$$

Where 'k' =0.5, 'M' is the minimum gray amount of the image, 'R' is highest gray value of global standard deviation 'σ' is local standard deviation. The calculation of local threshold value is done using minimum gray value and maximum standard deviation of local gray values of whole document image. Wolf's algorithm normalizes the contrast and the mean gray value of the image as compared to Sauvola algorithm. Thresholded result of image is degraded if there is a sharp change in background intensity values across the image. Small noise in image influences the values of influence 'M' and 'R' values.

Here, 'k' is a constant has a value between 0 and 1. With small value of k binarization gives thick and blurry strokes, and with large

D. Darek Bradley Technique

Darek Bradley technique [6] is well-suited for scenes with strong spatial changes in illumination. Temporal variations in illumination are also handled automatically, which is not the case for global thresholding methods. First integral image is form, as an input image. To calculate the integral image, save at each location, I(x, y), the total of all f (x, y) terms to the left and above the pixel (x, y). This is accomplished in linear time using the following equation for each pixel.

$$I(x, y) = (x, y) + I(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1) \quad (4)$$

After integral image calculated, the sum of the function for any rectangle with upper left corner (x₁, y₁) and lower right corner (x₂, y₂) can be computed in constant time using the following equation(5). The main drawback of this method is that image processed twice.

$$\sum_{x=x_1}^{x_2} \sum_{y=y_1}^{y_2} f(x, y) = I(x_2, y_2) - I(x_2, y_1 - 1) - I(x_1 - 1, y_2) + I(x_1 - 1, y_1 - 1) \quad (5)$$

E. Nick's Technique

The method proposed by Khurshid [7] is known as NICK's method. This method is a modified version of Niblack's algorithm. Nick algorithm solves the problem of Niblack algorithm and low contrast image difficulty in Sauvola algorithm.

$$T(x, y) = m + k \sqrt{\frac{\sum p_i^2 - m^2}{NP}} \quad (6)$$

Here, 'k' is the Niblack factor and vary between - 0.1 and -0.2 according to the application need, m is the average gray-level, p_i is the gray-level of pixel and NP is the total number of pixels. Khurshid suggested that for document images, the value of k must be set at -0.1 and in applications where we don't desire any noise, k

should be -0.2 . This method eliminates the problem of black-noise and also performs

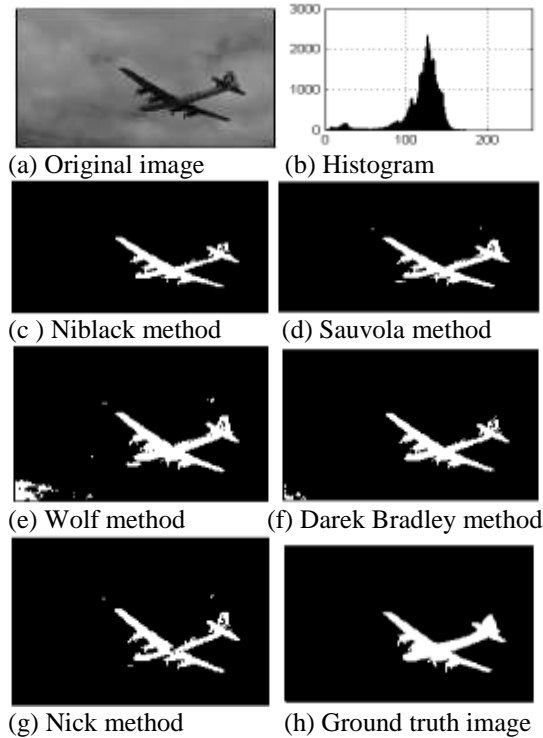


Fig.1 (a) Original image of airplane, (b) Histogram (c) Niblack's method, (d) Sauvola's method (e) Wolf method, (f) Darek Bradley method (g) Nick method, (h) Ground truth image

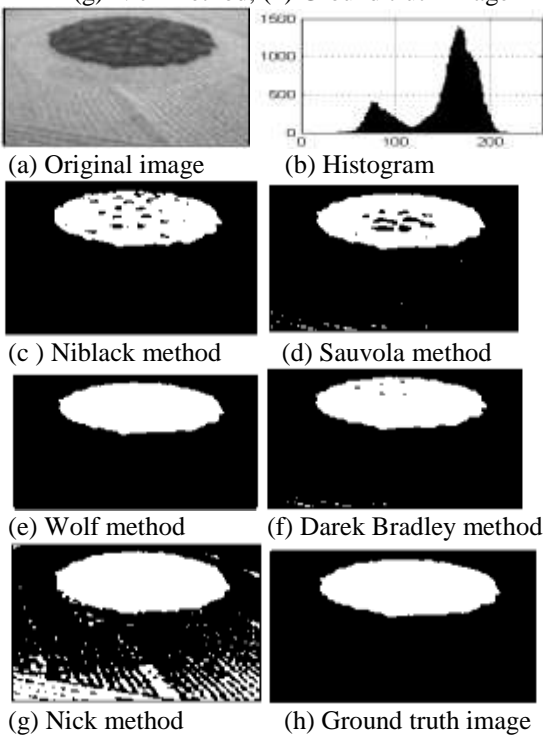


Fig.2 (a) Original image of field, (b) Histogram (c) Niblack's method, (d) Sauvola's method (e) Wolf method, (f) Darek Bradley method (g) Nick method, (h) Ground truth image

effectively well in case of low contrast.

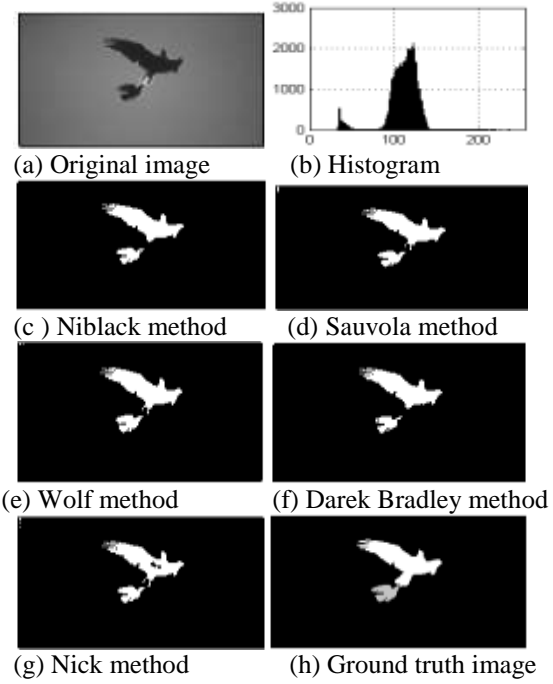


Fig.3 (a) Original image of bird, (b) Histogram (c) Niblack's method, (d) Sauvola's method (e) Wolf method, (f) Darek Bradley method (g) Nick method, (h) Ground truth image

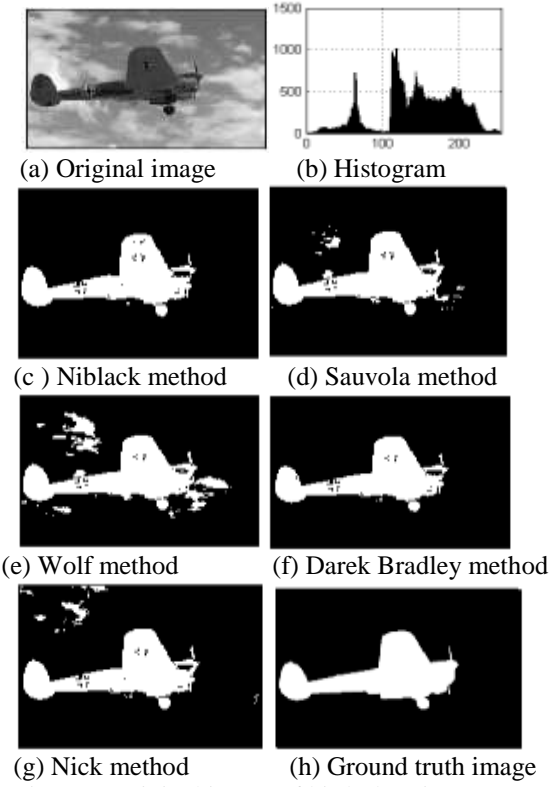
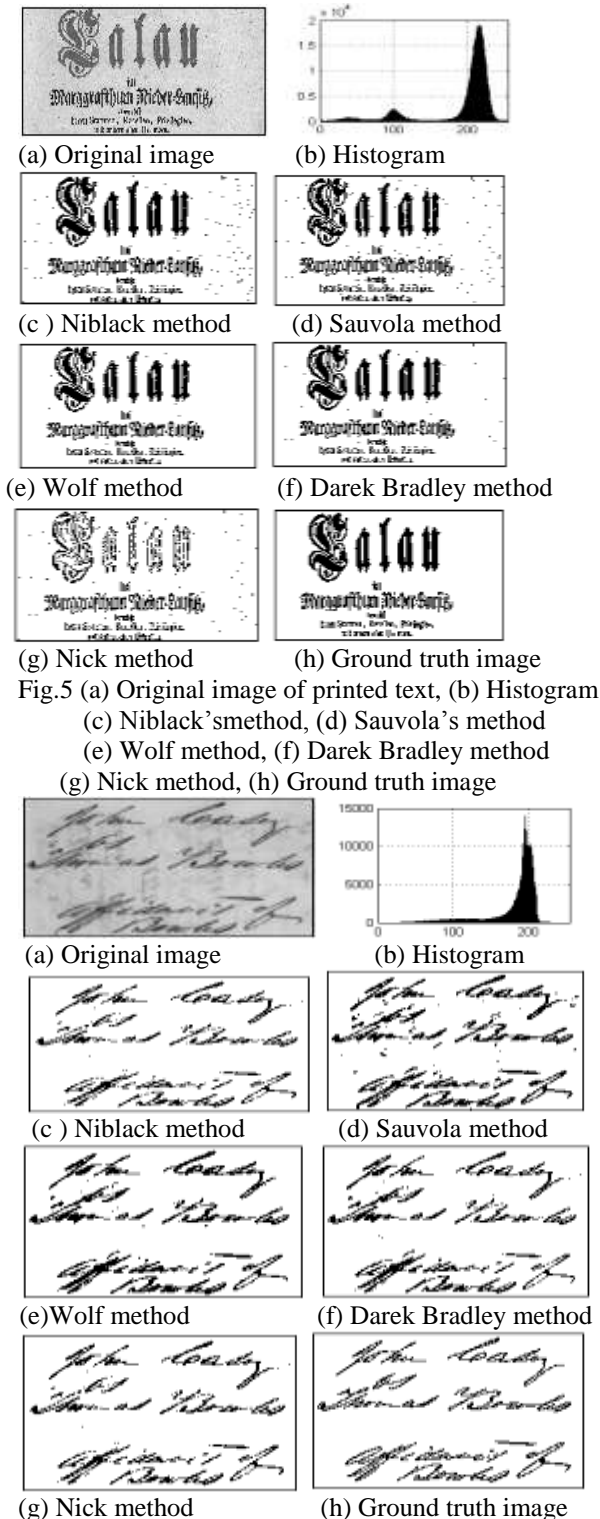


Fig.4 (a) Original image of airplane, (b) Histogram (c) Niblack's method, (d) Sauvola's method (e) Wolf method, (f) Darek Bradley method (g) Nick method, (h) Ground truth image



III. EXPERIMENTAL RESULTS AND ANALYSIS

In order to demonstrate the effectiveness of local adaptive thresholding techniques, we used four real world images (airplane1 image, field image, bird image and airplane2 image) from Berkeley segmentation data set and two images (printed text

document image and hand written text document image) from DIBCO -2009 data set. We have considered six images with non-uniform illumination condition having different sizes, airplane1 image(256×256), field image(256×256), bird image(256×256), airplane2 image(256×256) printed text document image(493x1153) and hand written text document(493x1153) images are shown in Fig. 1(a) to 6(a). The corresponding ground truth images from Berkeley segmentation data set (real world images) and DIBCO -2009 data set (Printed text document image, hand written text document image) is shown in Fig. 1(h) to 6(h). All the experiments are performed in MATLAB 14.0. As shown in fig. 1(b) to 6(b), the histograms of original images are not bimodal. Clearly, it is difficult to find a threshold that can separate the objects from the background.

Therefore, the use of an adaptive local thresholding is necessary to solve the problem. Using local adaptive thresholding, each pixel in the image will have its own threshold value to segment the object from the background. Five local adaptive thresholding techniques had been experimented with namely the Niblack's thresholding, Sauvola's thresholding, Wolf's thresholding, Darek Bradley and Nick's threshold based segmentation methods







Experimental result obtained by Niblack thresholding is shown in Fig.1(c) to 6(c). Result obtained by Sauvola's local thresholding is shown in Fig.1(d) to 6(d). Result obtained by Wolf's local thresholding is shown in Fig.1 (e) to 6(e). Result obtained by Darek Bradley local thresholding technique is shown in Fig.1 (f) to 6(f). Experimental results obtained by Nick's local thresholding are shown in Fig.1 (g) to 6(g). In our experiments neighborhood size and coefficient K varied from image to image in all local adaptive thresholding techniques. In the experiments we tested the performance Niblack's thresholding, Sauvola's thresholding, Wolf's thresholding, Darek Bradley and Nick's threshold based segmentation methods quantitatively and qualitatively.

For each experiment quality of thresholded images quantitatively analyzed using misclassification error [9], false positive rate and false negative rate [10]. Misclassification error (ME) estimates the percentage of wrongly classified pixels. Smaller the value of ME indicates the better the segmentation accuracy. ME value varies between '0' and '1'. ME value '0' means segmented accurately and '1' means totally erroneous result. Here, B_G and F_G denote background and foreground pixels of ground truth images respectively. B_S and F_S denote background and foreground pixel of segmented images using local adaptive thresholding.

$$ME = \frac{|B_G \cap B_S| + |F_G \cap F_S|}{|B_G + |F_G|} \quad (7)$$

False positive rate (FPR) and false negative rate (FNR) provide more information about misclassified pixels in comparison with misclassification error.

Table 1: Quantitative comparison on ME, FPR and FNR for the images

Images	Performance measures	Niblack	Sauvola	Wolf	Darek Bradley	Nick
	ME FPR FNR	0.0211945 0.0000029 0.9969790	0.0163879 0.0000478 0.9966265	0.0250702 0.0008729 0.9963570	0.0206909 0.0001523 0.9967974	0.0223389 0.0000322 0.9970229
	ME FPR FNR	0.0231476 0.0000003 0.9965429	0.0185242 0.0001011 0.9963493	0.0039825 0.0000800 0.9960784	0.0090942 0.0000316 0.9962295	0.1458588 0.0029286 0.9960784
	ME FPR FNR	0.0115356 0.0000131 0.9963582	0.0113220 0.0000024 0.9963523	0.0113831 0.0000869 0.9962726	0.0146484 0.0000202 0.9965938	0.0134430 0.0000143 0.9965058
	ME FPR FNR	0.0226593 0.0000004 0.9964736	0.0237122 0.0001846 0.9963158	0.0533142 0.0009914 0.9962519	0.0204926 0.0000000 0.9964196	0.0294800 0.0002696 0.9963755
	ME FPR FNR	0.0161199 0.0000430 0.9961117	0.0201327 0.0000583 0.9961154	0.0145119 0.0000584 0.996088	0.0101948 0.0000302 0.9960964	0.0645463 0.0002869 0.9960968
	ME FPR FNR	0.0239479 0.0000091 0.9962549	0.0229479 0.0000091 0.9962549	0.0334563 0.0000167 0.9962071	0.0259443 0.0000327 0.9961584	0.0243309 0.0000684 0.9961157

For atwo-class segmentation, FPR and FNR can be respectively formulated as

$$FPR = \frac{|F_G \cap B_S|}{|F_G|} \quad (8)$$

$$FNR = \frac{|B_G \cap F_S|}{|B_S|} \quad (9)$$

FPR and FNR value varies between ‘0’ and ‘1’. FPR indicates over-segmentation. FNR indicates under segmentation. Higher the value of FPR indicates serious over segmentation. Higher the value of FNR indicates serious under segmentation.

Experimental results are listed in Table 1. In addition, the Table1 compares ME for six different images with, Niblack, Sauvola, Wolf’s, Darek Bradley and Nick’s thresholding techniques. A smaller value of ME indicates better segmentation. More over by analyzing the results reported in Table 1, airplane1 image ME is less in Sauvola algorithm than other methods. Field image ME is less in wolf thresholding method than other methods. Bird image ME is less in Sauvola technique than other methods. Airplane2 image ME is less in Darek bradley algorithm than other methods. Printed text document image ME is less in Darek bradley algorithm than other methods. Handwritten text document image ME is less in Sauvola techniques than other methods. After analysis we concluded that Misclassification error is less in airplane2 and printed document image in Darek

Bradley technique as compared to remaining techniques. Therefore it can be concluded that Misclassification error is less in Airplane1 image, bird image and hand written text document image Sauvola technique than other methods.







Experimental results in Table 1, compares FPR for six different images with thresholding techniques. A larger value of FPR indicates over segmentation. Moreover by analyzing the results reported in Table 1, airplane1 image FPR is more in Wolf’s algorithm than other methods. In Field image FPR is more in Nick’s method than other methods. Bird image FPR is more in Wolf’s technique than other methods. Airplane2 image FPR is more in Wolf’s technique than other methods. Printed text document image FPR is more in Nick’s algorithm than other methods. Hand written text document image FPR is more in Nick’s techniques than other methods. After analysis we conclude that FPR is more for field image,. After analysis we concluded that FPR is more in Nick’s and Wolf’s thresholding techniques as compared to Niblack, Sauvola, Darek Bradley thresholding techniques. Therefore, it can be concluded that Wolf’s thresholding and Nick’s thresholding techniques are over segmented than other techniques.

Quantitative comparison on false negative rate for the images as shown in table 1, In addition, the Table 1, compares FNR for six different images with, Niblack, Sauvola, Wolf’s, Darek Bradley and Nick’s thresholding techniques. A larger value of

FNR indicates under segmentation. Moreover by analyzing the results reported in Table 1, airplane1 image FNR is more in Nick’s method than other methods. In Field image FNR is more in Niblack algorithm than other methods. Bird image FNR is

more in Darek Bradley algorithm than other methods. Airplane2 image FNR is more in Niblack algorithm than other methods.

Table 2: Qualitative comparison on local adaptive thresholding techniques for the images

Images	Performance Measures	Niblack	Sauvola	Wolf	Darek Bradley	Nick
	Time MSE PSNR	0.729943 0.262518 53.97319	0.741791 0.264792 53.99574	0.813653 0.271546 53.82635	0.888439 0.264966 53.93288	0.695260 0.261848 53.98429
	Time MSE PSNR	0.580734 0.432470 51.80523	0.608455 0.432338 51.80656	0.714018 0.435017 51.87972	0.765833 0.435077 51.77713	0.678862 0.418980 51.94285
	Time MSE PSNR	0.718230 0.2228243 54.685173	1.259651 0.2230423 54.780926	0.867814 0.2235470 54.671110	0.864138 0.2215864 54.709368	0.713496 0.2216801 54.707531
	Time MSE PSNR	0.677206 0.442246 51.70815	0.761272 0.443166 51.69913	0.826692 0.446922 51.70152	0.766528 0.442961 51.70913	0.781373 0.444110 51.68989
	Time MSE PSNR	1.018833 0.047454 61.40201	0.888833 0.047923 61.35931	0.991219 0.045804 61.55572	1.032198 0.046560 61.58462	0.912751 0.053817 60.85558
	Time MSE PSNR	0.852916 0.074288 59.45556	0.827642 0.075524 59.15981	0.839937 0.076540 59.32590	0.801489 0.074410 59.44845	0.871837 0.075990 59.40896

Printed document image FNR is more in Sauvola algorithm than other methods. Hand written text document image FNR is more in Niblack and Sauvola technique than other methods After analysis we conclude that FNR is less only in Wolf’s thresholding than other thresholding techniques. Therefore, it can be concluded that Wolf’s thresholding technique is less under segmented than other techniques.

The experimental results in terms of qualitative measures like Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR) and computational time by applying various methods to different images are shown in table 2. Experimental results are listed in Table 2. In addition, the table 2, compares processing time for various threshold based segmentation techniques, qualitative measurement evaluation consists of the Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR).To Compare computational efficiency of various thresholding methods, we specifically recorded their computational time on the different images with different window sizes.

Adaptive thresholding methods, computational time depend on window size. Performance varies at different window size. For certain images, these local adaptive techniques are not suitable for smaller window size. For certain images,

these local adaptive techniques are not suitable for larger window. Moreover by analyzing the results reported in Table 2, airplane1 image processing time is less in Nick’s method than other methods. In Field image processing time is less in Niblack algorithm than other methods. Bird image processing time is less in Nick’s algorithm than other methods. Airplane2 image processing time is less in Niblack

algorithm than other methods. Printed document image processing time is less in Sauvola algorithm than other methods. Hand written text document image processing time is less in Darek Bradley technique than other methods After analysis we conclude that processing time is less in Nick’s and Niblack thresholding methods than other thresholding methods.

Qualitative comparison on MSE for the images as shown in table 2, In addition, the Table 2, compares MSE for six different images with various thresholding techniques. The large value of MSE means that image is poor quality. Moreover by analyzing the results reported in Table 2, airplane1 image MSE is more in Wolf’s method than other methods. In Field image MSE is more in Darek bradley algorithm than other methods. Bird image MSE is more in Wolf’s algorithm than other methods. Airplane2 image MSE is more in Wolf’s algorithm than other methods. Printed document image MSE is more in Nick’s algorithm than other methods. Hand

written text document image MSE is more in Wolf's technique than other methods. After analysis we conclude that MSE is more in Wolf's thresholding and Nick's thresholding techniques than other thresholding techniques.

The PSNR measurement denotes how much a given original image is similar to thresholded image. The higher the value of PSNR is, the more the similarity between the thresholded image and original image. Moreover by analyzing the results reported in Table 2, airplane1 image PSNR is more in Sauvola algorithm than other methods. Field image PSNR is more in Wolf thresholding method than other methods. Bird image PSNR is more in Sauvola technique than other methods. Airplane2 image PSNR is more in Darek Bradley algorithm than other methods. Printed text document image PSNR is more in Darek Bradley algorithm than other methods. Handwritten text document image PSNR is more in Niblack than other methods. After analysis we conclude that PSNR is more in case of airplane2 and printed document image in Darek Bradley technique as compared to remaining techniques. PSNR is more in for airplane1 image, bird image in Sauvola technique as compared to remaining techniques.

IV. CONCLUSION

In this paper, we present the qualitative and quantitative analysis of Niblack's, Sauvola, Wolf's, Darek Bradley and Nick's methods for comparison. We evaluated the performance varies methods by adaptive window size selection has been tested by uneven lighting images. In our experiments neighborhood size and coefficient K varied from image to image in all local adaptive thresholding techniques. For document images, the window size is required to change depending on the character size. Local thresholding methods, computational time depend on window size. Experimental results on variety of real world images and text images demonstrate the effectiveness of the thresholding techniques. Adaptive thresholding techniques applied to images depending on contrast and illumination.

REFERENCES

- [1]. M. Sezgin, B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", *J. Electronic Imaging*, Vol. 13, pp. 146–165, 2004.
- [2]. Otsu, N., 1979. "A threshold selection method from gray-level histograms", *IEEE Transactions on Systems Man Cybernet SMC-9* (1), pp.62–66.
- [3]. W. Niblack, "An Introduction to Digital Image Processing", Engle wood Cliffs, N.J. Prentice Hall, pp. 115-116, 1956.
- [4]. Sauvola, J., Seppanen, T., Haapakoski, S., and Pietikainen, M, "Adaptive document binarization", *Proc. 4th Int. Conf. On Document Analysis and Recognition*, pp. 147–152, Ulm Germany, 1997.
- [5]. C. Wolf, J-M. Jolion, "Extraction and Recognition of Artificial Text in Multimedia Documents", *Pattern Analysis and Applications*, 6(4), pp. 309- 326, (2003).
- [6]. Bradley, D., Roth, G, "Adaptive thresholding using integral imag", *J. Graph. Tools* 12(2), pp.13–21, 2007.
- [7]. K. Khurshid, I. Siddiqi, C. Faure, and N. Vincent, "Comparison of Niblack inspired binarization methods for ancient documents", in 16th International conference on Document Recognition and Retrieval, Proceedings of SPIE, San Jose, Calif. USA, 2009.
- [8]. Rimpay Garg, Prabhneet Kaur Sandhu, "A New Approach for Image Enhancement using Hybrid Threshold", *International Journal of Engineering Trends and Technology*, Vol. 4, Issue 7, pp. 3229-3232, July 2013.
- [9]. Yasnoff, W.A., Mui, J.K., Bacus, J.W., 1977. "Error measures for scene segmentation", *Pattern Recognition* 9, pp.217–231.
- [10]. Hu Q, Hou Z, Nowinski WL. "Supervised range constrained thresholding", *IEEE Trans. Image. Process*, 15(1), 228-240, Jan 2006