Enhancement of Document Images Resolution by Binarization Produced by Camera

Anshu Pundhir[1], Suresh Gawande[2]

[1]M. Tech. Student, Department of Electronics & Communication, B.E.R.I., Bhopal, Madhya Pradesh, India

[2]Assistant Professor Department of Electronics & Communication, B.E.R.I., Bhopal, Madhya Pradesh, India

Abstract: ***The process of separation of pixel values into two groups, white as background and black as foreground is called Image Binarization. A major role plays by thresholding in binarization of images. In this paper, we propose a novel binarization method for document images produced by cameras. We have done the comparison study between the various algorithms based on Binarization algorithms. We have achieved better resolution of the image by using the Binarization method of optimum thresholding techniques.***

Keywords: Binarization, Thresholding, Optimum Threshold.

1. Introduction

The accumulating availability of low-priced, high performance, portable digital imaging appliances like camera has constructed a great option for supplementing usual scanning for document image accession. Cameras are fast, flexible and non-contact document imaging then scanner. The use of cameras has enabled human interaction with any type of document and has greatly eased document acquisition. Its capability to seize non paper chronicle images like scene text has several applications like road sign recognition, licence plate recognition, document archiving and wearable computing. Documents captured with hand-held camera are typically done in an uncontrolled environment. Several types of distortions induce in the captured image. Some of major distortions are: motion blur, low resolution, uneven light shading, under-exposure, over-exposure, perspective distortion and non-planar shape. A binarization process precedes the analysis and recognition procedures in most document processing systems. Thresholding plays a major in binarization of images. Thresholding can be categorized into global thresholding and local thresholding. Information and character stroke width. If a binarization method computes a threshold for an entire image, it is called a global method. Global thresholding techniques are generally based on histogram analysis. It works well for images with well separated foreground and background intensities. However, most of the document images do not meet this condition and hence the application of global thresholding methods is limited. It is difficult to control the imaging environment, so camera-captured images often exhibit non-uniform brightness. As such, global binarization methods are not suitable for camera images. On the other hand, if a method computes a threshold for the neighbourhood around each pixel or for each designated block in the image, it is called a local method. Local methods use a dynamic threshold across the image according to the local information. These approaches are generally window-based and the local threshold for a pixel is computed from the gray values of the pixels within a window centred at that particular pixel. Text is the most important information in a document. Segmentation of text regions involves the detection of text and then its extraction given the viewpoint.

Different binarization methods have been analyzed in for different types of documents. Here we present an analysis of eleven locally adaptive binarization methods for gray scale images with low contrast, variable background intensity and noise. In that analysis, Niblack's method was found to be the better of them all. Different improvements have since been made to the original Niblack's method to improve the results. These include Sauvola's algorithm [5], Wolf's work [6] and Feng's method [7]. In this paper, we focus on the binarization of grayscale documents using local adaptive thresholding technique

2. Related Work

The number of survey was done on thresholding. Comparison of different techniques for document binarization has received some attention in the past. The problem is that in every case, they try to use results from ensuing tasks in document processing hierarchy, in order to estimate the performance of the binarization algorithm. In case of historical documents where their quality obstructs the recognition, and sometimes the word segmentation as well, this method of evaluation can be proved problematic. Trier et al. computed eleven different locally adaptive binariza- tion methods for gray scale images with low contrast, variable background intensity and noise. Their evaluation proved that Niblack's method performed better than other local thresholding methods. In the following section, some of the well known sliding window binarization methods. We will limit our discussion only to the simple sliding-window thresholding methods, high- lighting their pros and cons and finally suggestion our improved method.

2.1 Local Adaptive Thresholding

A technique used for segmentation is Thresholding, which separates an image into two meaningful regions: foreground and background, through a selected threshold value T. If the image is a grey image, T is a positive real number.

Consider a greyscale document image in which $g(x, y) \varepsilon [0, 255]$ be the intensity of a pixel at location (x, y). In local adaptive thresholding techniques, the main aim is to compute a threshold $T(x, y)$ for each pixel such that

$$o(x,y) = \begin{cases} 0 & if\ g(x,y) \le T(x,y) \\ 255 & otherwise \end{cases} \quad (1)$$

A. Niblack's Method

In Niblack's binarization method, the threshold $t(x, y)$ is calculated using the mean $m(x, y)$ and standard deviation $\delta(x, y)$ of the pixel intensities in a $w \times w$ window cantered around the pixel (x, y):

$$T(x, y) = m(x, y) + k * \delta(x, y) \quad (2)$$

where $\boldsymbol{m(x, y)}$ and $\boldsymbol{\delta (x, y)}$ are the local mean and standard deviation of the pixels inside the local window and $\boldsymbol{k}$ is a bias. The result is adequate at k is fixed to -0.2 and $\boldsymbol{w=15}$.

The local mean $\boldsymbol{m(x, y)}$ and standard deviation $\boldsymbol{\delta(x, y)}$ adjust the value of the threshold according to the contrast in the local neighbourhood of the pixel. **The** bias $\boldsymbol{k}$ controls the level of reconciling varying the threshold value. Advantage of Niblack is that it always distinguishes the text regions correctly as foreground but on the other hand tends to produce a large amount of binarization noise in non-text regions also.

B. Sauvola's Algorithm

Sauvola's algorithm [5] claims to improve Niblack's method by computing the thres -hold using the dynamic range of image gray value standard deviation,

$$T(x,y) = m(x,y)\left[1 + k\left(\frac{\delta(x,y)}{R} - 1\right)\right] \quad (3)$$

where k is set to 0.5 and R to 128. This method outperforms Niblack's algorithm in images where the text pixels have near 0 gray-value and the background pixels have near 255 gray-values. However, in images where the gray values of text and non-text pixels are close to each other, the results degrade significantly.

3. Integral Sum Image for Calculating Local Means and Variances

The notion of integral images was popular- ized in computer innovation by Viola and Jones based on previous work in graphics. An integral image $\boldsymbol{I}$ of an input image $\boldsymbol{g}$ is described as the image in which the intensity at a pixel position is equal to the sum of the intensities to the left of that position and of all the pixels above in the original image. So the intensity at position (x, y) can be written as:

$$I(x,y) = \sum_{i=1}^{x} \sum_{j=1}^{y} g(x,y) \quad (4)$$

The integral sum image of any gray scale image can be efficiently determined in entire pixels at $\boldsymbol{(x, y)}$.

$$I(x,y) = g(x,y) + I(x,y-1) + I(x-1,y) - I(x-1,y-1), \quad (5)$$
$$x = 2 \ldots p, y = 2 \ldots q$$

Thus the integral sum image $\boldsymbol{I}$ can be determined in a single pass. The local sum $\boldsymbol{s(x, y)}$ at $\boldsymbol{(x, y)}$ which is the centre of the local window of size $\boldsymbol{w \times w}$ of an image I is the sum of all the pixel intensities within the local window The sum $\boldsymbol{s(x, y)}$ can be computed in two passes as:

$$s(x,y) = \sum_{i=x-b}^{x+b} \sum_{j=y-b}^{y+b} g(x,y) \quad (6)$$

The local sum $s(x, y)$ for any window size can be determined simply by using one subtraction and two addition operations instead of the summation over all pixel values within that window:

$$s(x,y) = (g(x+w/2, y+w/2) + g(x-w/2, y-w/2) - g(x+w/2, y-w/2) - g(x-w/2, y+w/2)) \quad (7)$$

the local mean $m(x, y)$ for any window size can be determined

$$m(x,y) = \frac{s(x,y)}{w^2} \quad (8)$$

Similarly, if we consider the calculation of the local variance

$$s^2(x,y) = \frac{1}{w^2} \sum_{i=x-p}^{x+p} \sum_{j=y-p}^{y+p} g^2(i,j) - m^2(x,y) \quad (9)$$

4. Our Approach

We approach an efficient way of computing local threshold, to minimise the computation –al time of local thresholding calculation. Binarization can be speed up considerably as a result of using the integral sum image to calculate the local sum in where computational time does not based on the window dimension.In other techniques like Niblack and Sauvola , local mean m(x, y) and standard deviation δ(x, y) are necessary to calculate threshold for each pixal. In this technique we are using local mean and mean deviation for calculating threshold.

$$t(x,y) = m(x,y)\left[1 + k\left(\frac{\partial(x,y)}{1-\partial(x,y)} - 1\right)\right] \quad (10)$$

Where $\boldsymbol{\partial(x, y) = I(x, y) - m(x, y)}$

$\boldsymbol{\partial(x, y)}$ is the local mean deviation and $\boldsymbol{k}$ is a bias which can control the level of adaptation varying threshold value. Its range is [0, 1] only.

If local sum is available, using simple arithmetic operation mean can be calculated without depending on the window size. Integral sum image is computed as a prior process for calculating the local sum easily. The value of $\boldsymbol{k}$ plays a major role in determining threshold value.

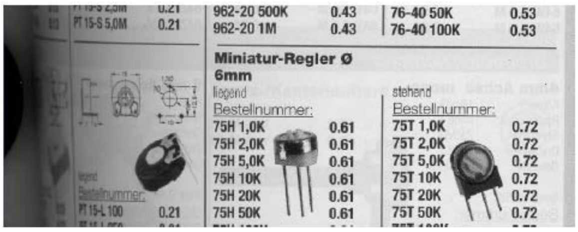


Figure 1: Original Images

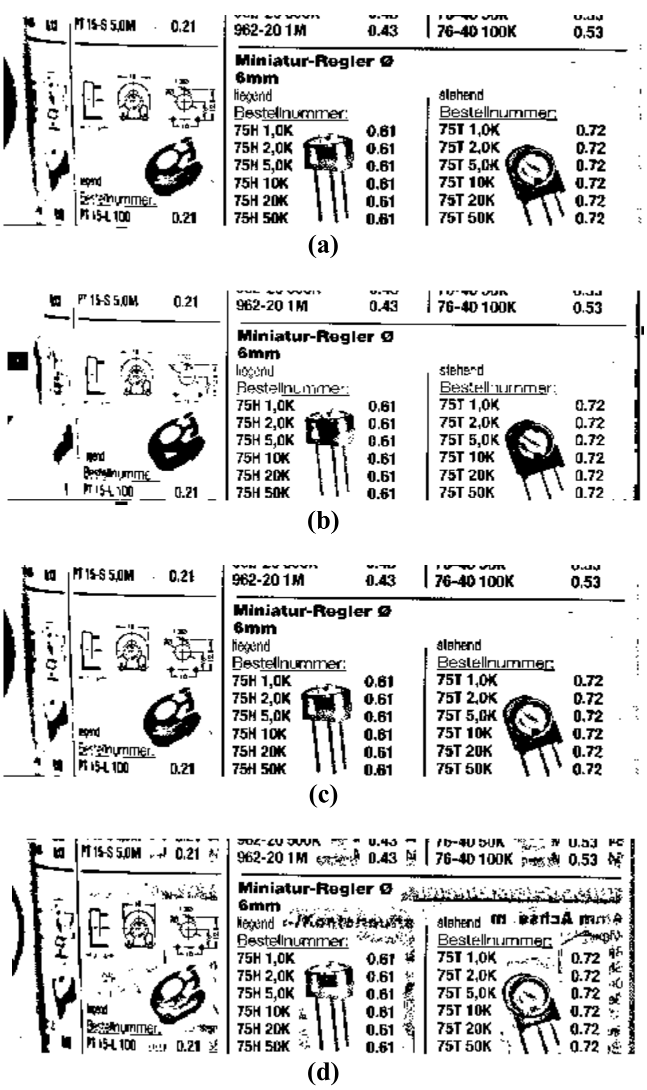
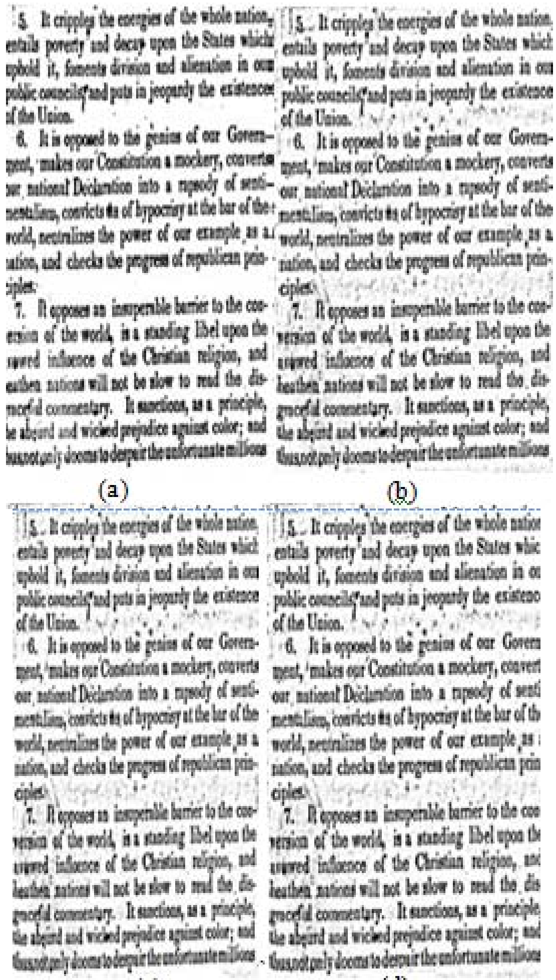
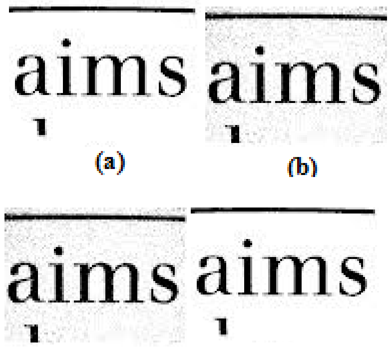


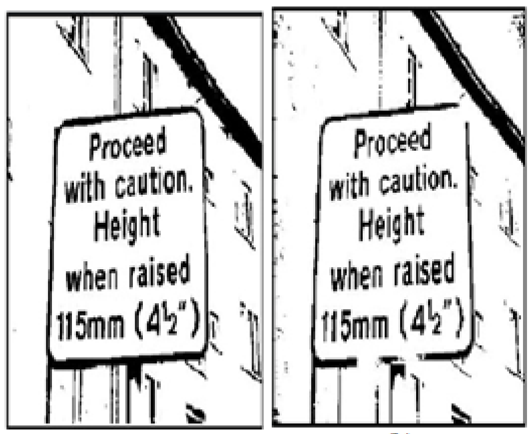
Figure 2: Comparison on scanned document: (a) Proposed (b) Bernsen's.(c) Sauvola's (d). Niblack's at k=-2



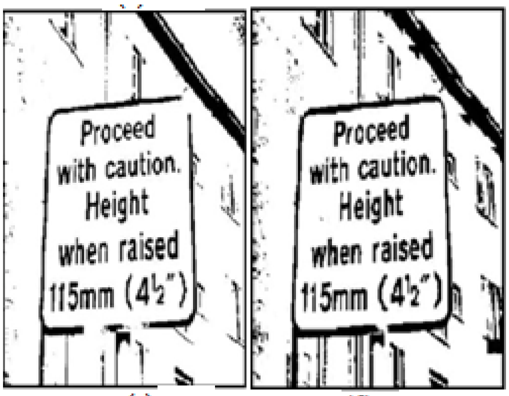
(a) (b) (c) (d)



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



(a) (b)



(c) (d)

Figure 3: Comparison on scanned document at size: (a) Proposed (b) Bernsen's.(c) Sauvola's (d). Niblack's Method

Table1: Calculational Time Comparison on image (in Sec)

Window Size	*Proposed*	*Bernsen*	*Niblack*	*Sauvola*
3	0.252	0.8121	7.286	7.1546
7	0.242	1.3678	7.4312	7.4165
11	0.242	2.2145	7.1234	7.9754
15	0.242	3.3919	8.5177	8.6743
23	0.242	6.3455	10.1654	10.0542
27	0.2213	7.9876	10.9876	11.0607
31	0.2034	9.9879	12.3245	12.1457
35	0.1868	11.1676	12.9867	13.3126

5. Experiments and Results

The experiment for the preferred technique was carried out using MATLAB 7.11 (R2010b) on a PC with the following configuration: Intel(R) Core(TM) i5-2430M CPU @ 2.40 GHz, 4GB RAM, 64 bit OS (Windows7). Qualitative analysis provides a set of images for differentiation to the reader for personal analysis. Qualitative analysis gives a set of readings. Qualitative and quantitative both analyses are achieved in comparison of the preferred technique with other applicable techniques. Comparison of calculational times of the proposed method with other local adaptive methods is shown in Table 1.

6. Conclusion

This paper recommends a threshold technique based on Locally Adaptive Binarization method. We use integral image for calculating mean and variance in local window. So that running time does not depend on the local window size. It's running time close to global binarization method. This method is compared with other techniques and got to be better than other, both in terms of quality and quantity. It is faster than other methods.

References

[1] Cao, L., Shi, Z.K., and Cheng, E.K.W.: „Fast automatic multilevel thresholding‟, Electron. Lett., 2002, 38, pp.868–870

[2] Chang, C.C., Chang, C.H., and Hwang, S.Y.: „A connectionist approach for thresholding‟. Proc. Int.Conf. Pattern Recognit., 1992, vol. 11, pp. 522–525

[3] Gonzalez and Woods, Digital image processing, 2nd Edition, prentice hall, 2002

[4] Haralick, R.M., and Shapiro, L.G.: „Computer and robot vision, vol. 1‟ (Addison-Wesley, 1992)

[5] Kamel M. and Zhao A., ''Extraction of binary character/graphics images from grayscale document images,'' Graph. Models Image Process.55 (3), 203–217(1993).

[6] Khurshid1 Khurram, Siddiqi1 Imran , Claudie Faure2, and Nicole Vincent1, Comparison of Niblack inspired Binarization methods for ancient documents ,1Laboratoire CRIP5 – SIP, Université Paris Descartes, 2006

[7] Leedham G., Varma S., Patankar A., Govindaraju V.,"Separating Text and Background in Degraded Document Images" Proc. 8th IWFHR, pp. 244-249,September, 2002

[8] Lee, S.U., Chung, S.Y., and Park, R.H.: „A comparative performance study of several global thresholding techniques for segmentation‟, Comput. Vis. Graph Image Process. 1990, 52, pp. 171–190

[9] Niblack W., *An Introduction to Image Processing*, Prentice-Hall, Englewood Cliffs, NJ, 1986.

[10] Otsu, N.: A thresholding selection method from graylevel histogram‟, IEEE Trans. Syst. Man Cybern.,1979, 9, (1), pp. 62–66

[11] O.D.Trier, A.K. Jain, Goal-directed evaluation of binarization methods, IEEE Tran. Pattern Analysis and Machine Intelligence, PAMI-17 (1995) 1191-1201