Local Contrast and Mean based Thresholding Technique in Image Binarization

O. Imocha Singh
Department of Computer Science, Manipur University
Canchipur, Imphal-795003, Manipur, India.

Tejmani Sinam
Department of Computer Science, Manipur University
Canchipur, Imphal-795003, Manipur, India

O. James
DOEACC, Imphal Centre,
Imphal-795008, India

T.Romen Singh
Department of I.T, School of Technology, Assam University,
Silchar – 788011, Assam, India

ABSTRACT

Binarization is a process of separation of pixel values of an input image into two pixel values like white as background and black as foreground. It is an important part in image processing and it is the first step in many document analysis and OCR processes. Most of the binarization techniques associate a certain intensity value called threshold which separate the pixel values of the concerned input grayscale image into two classes like background and foreground. Each and every pixel should be compared with the threshold and transformed to its respective class according to the threshold value. Thus threshold takes a major role in binarization. Hence determination of proper threshold value in binarization is a major factor of being a good binarised image and it can be approached in two categories like global thresholding and local thresholding techniques. In uniform contrast distribution of background and foreground documents, global thesholding is more suitable than that of local thresholding one. In degraded documents, where considerable background noise or variation in contrast and illumination exists, local technique is more suitable than that of global one. In this paper a local thresholding technique using local contrast and mean is described. Local adaptation is carried out with the local contrast and mean.

General Terms

Image binarization using local minimum and maximum pixel value within a local window w×w.

Keywords

Binarization, local thresholding, local min, local max and local mean.

1. INTRODUCTION

In general, scanned documents include text, line-drawings and graphic regions. It can also be considered as mixed type documents. In many practical applications, we need to recognize or improve the content of the document. In such cases, it is preferable to convert the documents into a binary form. Thus Image binarization plays a key role in the field of Image Processing. Binarization is a process of transforming a gray scale image to a binary image which contain only two classes of pixels white as background and black as foreground. Classification is carried out with a separation intensity value called threshold. Threshold plays a major role in binarization and choosing of an appropriate threshold value is an important one. There are many techniques of determining an appropriate threshold value in binarization

techniques. Document binarization is the first step in most document analysis systems. The goal of document binarization is to convert the given input grayscale or color document into a bi-level representation. This representation is particularly convenient because most of the documents that occur in practice have one colour for text (e.g. black), and a different colour (e.g. white) for background.

Thresholding then becomes a simple but effective tool to separate objects from the background. Examples of thresholding applications are document image analysis, where the goal is to extract printed characters,[2,3] logos, graphical content, or musical scores: map processing, where lines, legends, and characters are to be found[4] : scene processing, where a target is to be detected[5] and quality inspection of materials,[6,7] where defective parts must be delineated.

The output of the thresholding operation is a binary image where one state will indicate the foreground objects, that is, printed text, a legend, a target, defective part of a material, etc., while the complementary state will correspond to the background. Depending on the application, the foreground can be represented by gray-level 0, that is, black as for text, and the background by the highest luminance for document paper, that is 255 in 8-bit images, or conversely the foreground by white and the background by black.

The binarization techniques for grayscale documents can be grouped into two broad categories: global thresholding binarization and local thresholding binarization. Global methods like that of Otsu[1] try to find a single threshold value for the whole document. Then each pixel is assigned to page foreground or background based on its gray value comparing with the threshold value. Global methods are very fast and they give good results for typical scanned documents. For many years, the binarization of grayscale documents was based on the global thresholding statistical algorithms [8–13]. These statistical methods, which can be considered as clustering approaches, are suitable for converting any grayscale image into a binary form but are inappropriate for complex documents, and even more, for degraded documents. If the illumination over the document is not uniform, for instance in the case of scanned book pages or camera-captured documents, global binarization methods tend to produce marginal noise along the page borders. To overcome these complexities, local thresholding techniques have been proposed for document binarization. These techniques estimate a different threshold for each pixel according to the grayscale information of the neighboring pixels. The

techniques of Bernsen [13], Chow and Kaneko [14], Eikvil [15], Mardia and Hainsworth [16], Niblack [17], Taxt [18], Yanowitz and Bruckstein [19], J.Sauvola & Pietikainen [8] and TR Singh *et al* [20][38] belong to this category. The hybrid techniques : L.O'Gorman [22] and Liu [21]., which combine information of global and local thresholds belong to another category.

In this paper we focus on the binarization of grayscale documents using local adaptive thresholding technique, because in most cases color documents can be converted to grayscale without losing much information as far as distinction between page foreground and background is concerned. This methods can achieve good results even on severely degraded documents, but it is slow since the computation of local mean, max and min from the local neighborhood is to be done for each image pixel.

2. THESHOLDING

A threshold value $T(x,y)$ is a value such that

$$b(x,y) = \begin{cases} 0 & if\ I(x,y) \leq T(x,y) \\ 1 & otherwise \end{cases} \qquad (1)$$

where $b(x,y)$ is the binarized image and $I(x,y) \in [0,1]$ be the intensity of a pixel at location (x,y) of an image I. Image binarization is the process of separation of pixel values into two groups, white as background and black as foreground. Thresholding is a process of finding an appropriate threshold value for binarization. It plays a major role in binarization of images. Thresholding can be categorized into Global thresholding and Local thresholding.

2.1 Global Thresholding

In the global thresholding technique, a unique threshold value $T(x,y)$ is selected for the entire input image and each pixel is compared with the selected threshold value at the time of transformation. Global thresholding [2] is an efficient, less time consuming and region independent one. It is more appropriate in the uniformly illuminated images than the badly illuminated images. In images with uniform contrast distribution of background and foreground like document images, global thresholding is more appropriate. But in degraded document images, where considerable background noise or variation in contrast and illumination exists, there exists many pixels that cannot be easily classified as foreground or background. In such cases, binarization with Global thresholding is not appropriate

2.2 Local Thresholding

In local thresholding technique, a threshold $T(x,y)$ is calculated for each pixel, based on some local statistics such as range, variance, or surface-fitting parameters of the neighborhood pixels within a local block of size $w \times w$.

Consider a grayscale document image in which $I(x,y) \in [0,1]$ be the intensity of a pixel at location (x,y). In local adaptive thresholding techniques, the aim is to compute a local threshold $T(x,y)$ for each pixel such that

$$b(x,y) = \begin{cases} 0 & if\ I(x,y) \leq T(x,y) \\ 1 & otherwise \end{cases} \qquad (2)$$

where $b(x,y)$ is the binarized image.

In locally adaptive thresholding technique of algorithms, a threshold is calculated at each pixel. It can be approached as local variance method and local contrast method. Nakagawa and Rosenfeld[30] and Deravi and Pal[31] were the early users of adaptive techniques for thresholding. Niblack[17], Sauvola & Pietaksinen[8], T.R.Singh *et al*[20] and LAAB[38] use the local variance. The method from Bernsen [13] adapts the threshold according to the local variance technique.

2.3 Niblack's Technique:

In this method the local threshold value $T(x,y)$ at (x,y) is calculated with a window size of $w \times w$ as

$$T(x,y) = m(x,y) + k * \delta(x,y) \qquad (3)$$

where $m(x,y)$ and $\delta(x,y)$ are the local mean and standard deviation of the local window $w \times w$, k is a bias and set as k =-0.2 and local window size is $w = 15$.

The method from Niblack [17] adapts the threshold according to the local mean $m(x,y)$ and standard deviation $\delta(x,y)$ and calculated within a window of size $w \times w$. Sauvola and Pietaksinen proposed an improved technique.

2.4 Sauvola's Technique

In Sauvola's technique[8], the threshold $T(x,y)$ is computed using the mean $m(x,y)$ and standard deviation $\delta(x,y)$ of the pixel intensities in a $w \times w$ window centered around the pixel at (x,y) and express as:

$$T(x,y) = m(x,y)\left[1 + k\left(\frac{\delta(x,y)}{R} - 1\right)\right] \qquad (4)$$

where R is the maximum value of the standard deviation ($R = 128$ for a grayscale document), and k is a parameter which takes positive values in the range [0.2, 0.5].

The local mean $m(x,y)$ and standard deviation $\delta(x,y)$ adapt the value of the threshold according to the contrast in the local neighborhood of the pixel.

When there is high contrast in some region of the image, $\delta(x,y) \sim R$ which results in $T(x,y) \sim m(x,y)$. This is the same result as in Niblack's method. However, the difference comes in when the contrast in the local neighborhood is quite low. In that case the threshold $T(x,y)$ goes below the mean value thereby successfully removing the relatively dark regions of the background.

The parameter k controls the value of the threshold in the local window such that the higher the value of k, the lower the threshold from the local mean $m(x,y)$. A value of k = 0.5 was used by Sauvola1 and Sezgin. Badekas *et al*. experimented with different values and found that k = 0.34 gives the best result. In general, the algorithm is not very sensitive to the value of k used. The statistical constraint in Eqn (4) gives very good result even for severely degraded document. In order to compute the threshold $T(x,y)$, local mean and standard deviation have to be computed for each pixel its computational complexity is $O(n^2 \times w^2)$ in a naive way for an image of size $n \times n$. It means that its computational complexity is window size dependent. T.R Singh proposed window size independent technique of thersholding using integral sum image as prior process.

2.5 T.R Singh's Technique

To minimize the computational time of local thresholding calculation, T.R Singh *et al* [20] propose an efficient way of determining local threshold using integral sum image as prior process for determining local sum. It uses only local mean and it is very convenient to use integral sum while other techniques like Sauvola's and Niblack's. technique are not much convenient because of using both use local mean and standard deviation. Due to uses of integral sum image this technique is local window size independent. This technique is expressed as

$$T(x,y) = m(x,y)\left[1 + k\left(\frac{\partial(x,y)}{1-\partial(x,y)} - 1\right)\right] \quad (5)$$

where $\partial(x,y) = I(x,y) - m(x,y)$ is the local mean deviation and $k \in [0,1]$ is a bias which can control the level of adaptation varying threshold value. The value of k takes a major role in determining threshold value. The lower value of k makes the threshold value higher and vice versa.

2.6 LAAB(Local Adaptive Automatic Binarization)

Automatic binarization, proposed by T.R singh *et al* [38], is a process of transforming a gray scale image $I(x,y)$ to a binary image $b(x,y)$ automatically without using any threshold $T(x,y)$ by adapting the pixels within a local region environment. It is an automatic binarization with local adaptation. Local adaptation is carried out within a local window of size $w \times w$ with the help of local mean $m(x,y)$ of pixel intensity values of pixels within the local region.

The automatic binarization is designed as:

$$b(x,y) = \frac{|1-2v|-(1-2v)}{2|1-2v|} \quad (6)$$

Where $v = \frac{k(1+\partial)}{(1-\partial)}$, k is a bias such that $0.5 < k < 0.6$, $\partial = \{g(x,y) - m(x,y)\}\{1 - m(x,y)\}$ and $g(x,y)$ is the original pixel at (x,y). The bias k control the level of adaptation of pixels within the local region at the time of transformation to the binarised image $b(x,y)$. The greater the value of k, the more area of background and less area of foreground and vice versa. .

2.7 Bernsen's Technique

This technique, proposed by Bernsen [13], is a local binarization technique, which uses local contrast value to determine local threshold value. The local threshold value for each pixel (x,y) is calculated by the relation

$$T(x,y) = \frac{I_{max}+I_{min}}{2} \quad (7)$$

where I_{max} and I_{min} are the maximum and minimum gray level value in a $w \times w$ window centered at (x,y) respectively. But the threshold assignment is based on local contrast value and hence it can be expressed as

$$T(x,y) = \begin{cases} \frac{I_{max}+I_{min}}{2} & if \ I_{max} - I_{min} > L \\ GT & if \ I_{max} - I_{min} < L \end{cases} \quad (8)$$

where L is a contrast threshold and GT is a global threshold value calculated from the application of the technique of Otsu to the entire image.

3. PROPOSED ALGORITHM

Thresholding takes a major role in binarization. Determination of thresholding is major issue in binarization. Based on the technique of determination of threshold there are many techniques of binarization. It can be approached broadly in two ways like global and local technique. In this paper only local techniques are considered. Local technique is normally local window block size dependent both in result and time complexity. Many techniques which associates local mean, use integral image as a prior process to determine the local mean. It reduce time complexity. The technique of TR Singh *et al* use local mean only while that of Souvola's and Niblack's technique use standard deviation also. Bernsen use local contrast pixel values to determine the threshold value. Local contrast is the difference of the local maximum and minimum pixel values.

This proposed technique of binarization is carried out with a local thresholding technique which uses local contrast and mean.

It is expressed as below:

$$T(x,y) = k[m(x,y) + (I_{max} - I_{min})(1 - I(x,y)] \quad (9)$$

where $k \in (0,1)$ is a bias constant, $m(x,y)$ is local mean, I_{max} and I_{min} are the local maximum and minimum pixel values within the local window of size $w \times w$ and $I(x,y)$ is the concerned pixel value.

In this technique local maximum and local minimum pixel values are associated. As a result of using local minimum and maximum this technique is not suitable for using integral image which is suitable for finding local mean with no time. Hence it works like the other local techniques whose computational time is local window size dependent.

The bias constant k takes a major role to control the level of binarization by controlling the threshold value. The higher the value of k will result the higher value of threshold resulting more in foreground (black) and like this less the value of k will result the less in foreground (black). If the local region is uniform ie. $I_{max} - I_{min} \sim 0$, then $T(x,y) \sim k * m(x,y)$. Thus the value of $T(x,y)$ depends on the value of k. In this case the region becomes background for the other techniques while this technique, $T(x,y)$ is under controlled. In this way this technique gives unlike result of the other local techniques in a better way regarding results.

4. EXPERIMENTAL RESULT

The proposed technique Image Binarization by Adaptive Thresholding Technique using Local contrast and mean is tested comparing with other techniques on many categories of images like medical image, Scanned document images, Camera captured document images and non document images.
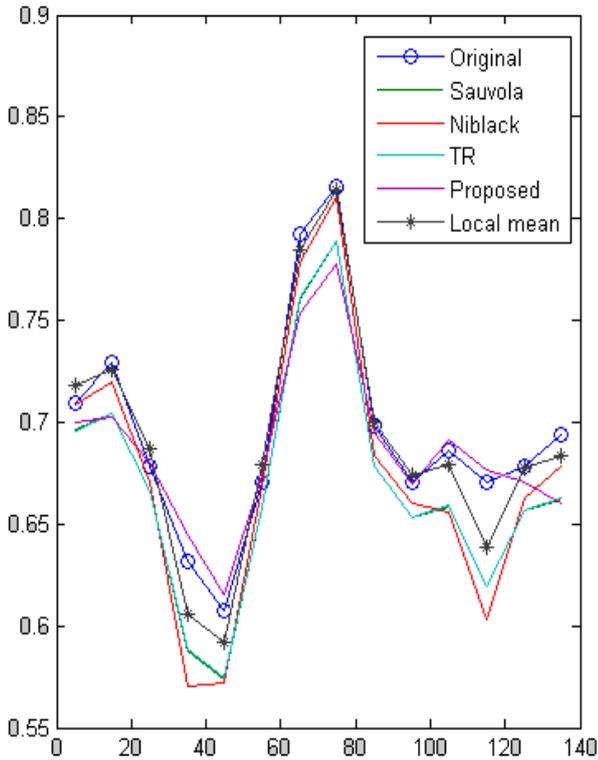
Figure 1. Nature of threshold values of different techniques with a window of size w=5 based on the result at Figure 2.

Figures 2-5 show some of the results comparing with others and found to be better than others. The experiment was carried out using MATLAB 7.3 (R2006b) on a PC with the following configuration: Intel® Core™2Duo CPU E6550,GHz 2.33 GHz, 2GB RAM, 32 bit OS(Windows Vista). Figure 1 shows the graphical representation of different levels of the threshold values of the different techniques including the proposed one based on the result shown in Figure 2. From the experimental result we find that TR Singh and Sauvola give similar result. But their result depend on the local window size. Sauvola requires a window of size $w = 15$ to get the high quality result while Niblack required a window size w=31 for a quality result. The result shown in fig. 2-5 are tested on a window size of $w = 5$. Even the widow size is less, the proposed technique give better result for large character size while the other detect only the character boundary like in fig 4. This new technique can give better result even at a small window size say $w = 3$. Since the other local technique remove the background, continuous contrast image like Figure 3, can not be binarized while this new technique can binaries efficiently like global one.

Due to uses of integral images, in LAAB and TR Singh *et al*, the computational time is independent of local window size w and time complexity is $O(n^2)$ while other techniques depend on local window size w with a complexity of $O(n^2 \times w^2)$. Integral image can provide the local sum. But this technique uses local contrast and mean value, it does not require integral image. Hence its time complexity is window size dependent and its time complexity is $O(n^2 \times w^2)$. Based on the experimental results this new technique is better than the other techniques.

Figure 2. Binarised result of human face by different techniques with local window size w=5 : (a) original image (b) LAAB at k=0.502, (c) T.R Singh at k=0.03, (d) Sauvola at k=0.03. (e) Bernsen at c=0.06 (f) Niblack at k=-0.5 and (g) Proposed at k=0.95.
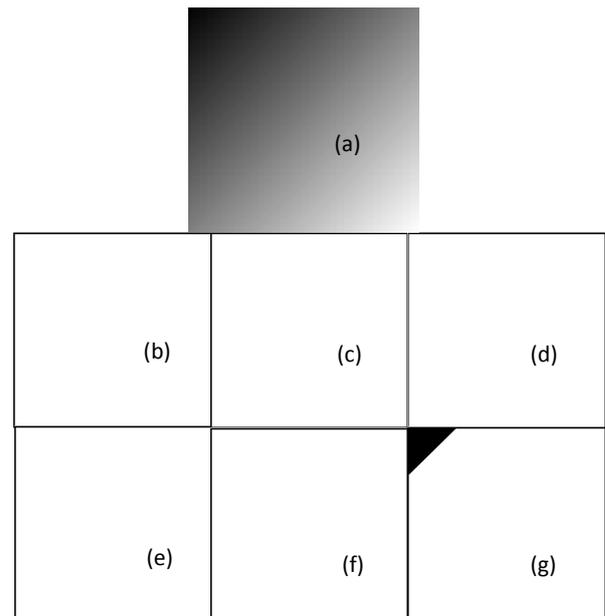
Figure 3. Binarised scaned image by different techniques with local window size w=5 : (a) original image (b) LAAB at k=0.502, (c) T.R Singh at k=0.02, (d) Sauvola at k=0.02. (e) Bernsen at c=0.06 (f) Niblack at k=-0.3 and (g) Proposed at k=0.9.

5. CONCLUSION

This paper presents a new way of image binarization using a local threshold value which is determined using contrast pixel value within a local window size w for local adaptation. We compare the result with other local techniques like Sauvola, Niblack, TR Singh *et al*, LAAB and Bernsen techniques.
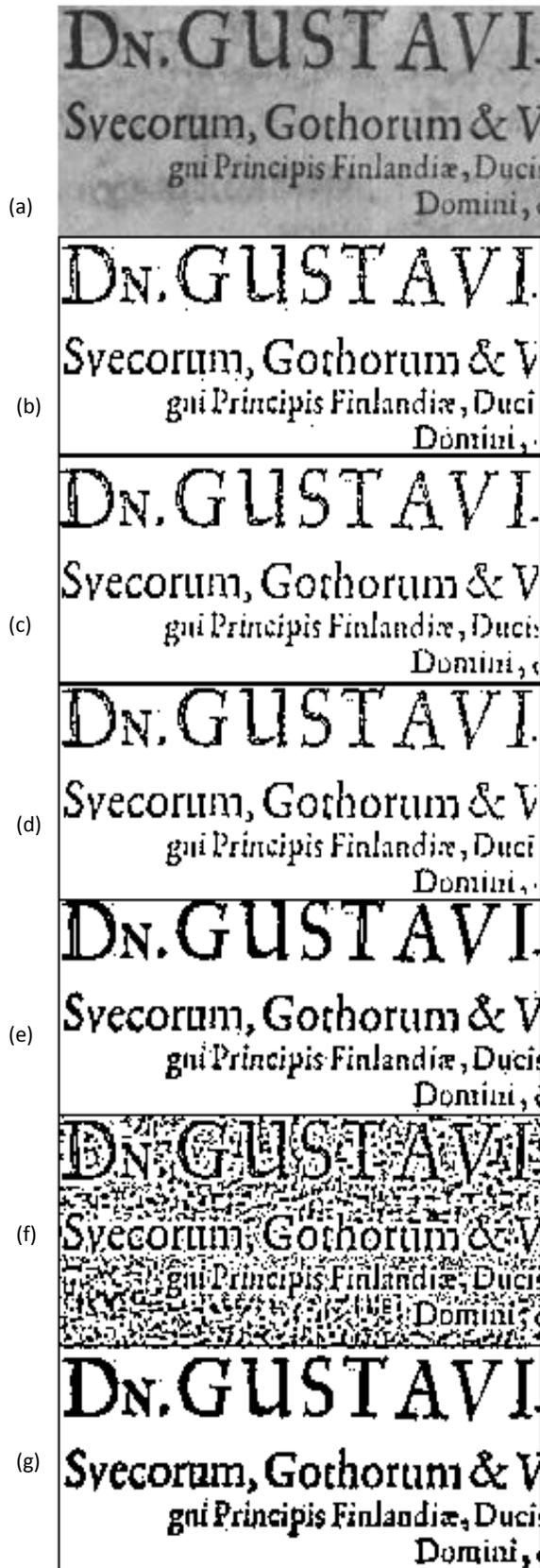
Figure 4. Binarised scaned image by different techniques with local window size w=5 :(a) original image (b) LAAB at k=0.54, (c) T.R Singh at k=0.2, (d) Sauvola at k=0.2. (e) Bernsen at c=0.3 (f) Niblack at k=-0.3 and (g) Proposed at k=0.67.
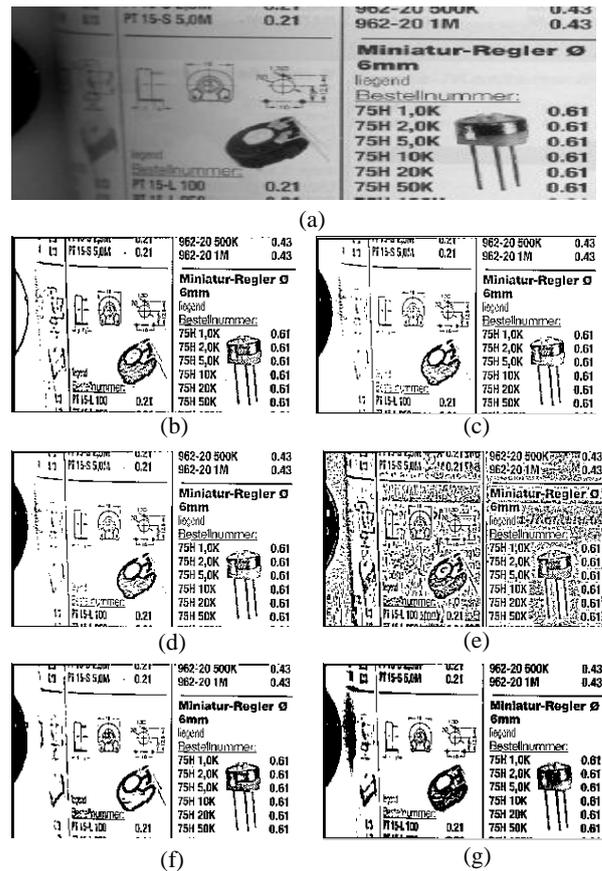
Figure 5. Binarised scaned image by different techniques with local window size w=5 : (a) original image (b) LAAB at k=0.54, (c) T.R Singh at k=0.2, (d) Sauvola at k=0.2. (e) Bernsen at c=0.3 (f) Niblack at k=-0.3 and (g) Proposed at k=0.67.

6. REFERENCES

[1] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Systems, Man, and Cybernetics 9(1), pp.62-66 , 1979.

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

[3] T. Abak, U. Baris,, and B. Sankur, ''The performance of thresholding algorithms for optical character recognition,'' *Intl. Conf. Document Anal. Recog. ICDAR'97,* pp. 697–700 (1997).

[4] O. D. Trier and A. K. Jain, ''Goal-directed evaluation of binarization methods,'' *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-17, 1191– 1201(1995).

[5] B. Bhanu, ''Automatic target recognition: state of the art survey,'' *IEEE Trans. Aerosp. Electron. Syst.* AES-22, 364–379 (1986).

[6] M. Sezgin and R. Tasaltin, ''A new dichotomization technique to multilevel thresholding devoted to inspection applications,'' *Pattern Recogn. Lett.* 21, 151–161 (2000).

[7] M. Sezgin and B. Sankur, ''Comparison of thresholding methods for non-destructive testing applications,'' *IEEE*

ICIP'2001, Intl. Conf. Image Process., pp. 764–767 (2001).

[8] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," Pattern Recognition 33(2), pp. 225–236, 2000.

[9] P. Viola and M. J. Jones, "Robust real-time face detection," Int. Journal of Computer Vision 57(2), pp. 137– 154, 2004.

[10] R. Cattoni, T. Coianiz, S. Messelodi, and C. M.Modena, "Geometric layout analysis techniques for document image understanding: a review," tech. rep., IRST, Trento, Italy, 1998.

[11] F. Shafait, D. Keysers, and T. M. Breuel, "Performance comparison of six algorithms for page segmentation," in 7th IAPR Workshop on Document Analysis Systems, pp. 368–379, (Nelson, New Zealand), Feb. 2006.

[12] J. M. White and G. D. Rohrer, "Image thresholding for optical character recognition and other applications requiring character image extraction," IBM Journal of Research and Development 27, pp. 400–411, July 1983.

[13] Bernsen, J.: 'Dynamic thresholding of gray-level images'. Proc. 8th Int. Conf. on Pattern Recognition, Paris, 1986, pp. 1251–1255

[14] Chow, C.K., and Kaneko, T.: 'Automatic detection of the left ventricle from cineangiograms', Comput. Biomed. Res., 1972, 5, pp. 388–410

[15] Eikvil, L., Taxt, T., and Moen, K.: 'A fast adaptive method for binarization of document images'. Proc. ICDAR, France, 1991, pp. 435–443

[16] Mardia, K.V., and Hainsworth, T.J.: 'A spatial thresholding method for image segmentation', IEEE Trans. Pattern Anal. Mach. Intell., 1988, 10, (8), pp. 919–927

[17] Niblack, W.: 'An introduction to digital image processing' (Prentice- Hall, Englewood Cliffs, NJ, 1986), pp. 115–116

[18] Taxt, T., Flynn, P.J., and Jain, A.K.: 'Segmentation of document images', IEEE Trans. Pattern Anal. Mach. Intell., 1989, 11, (12), pp. 1322–1329

[19] Yanowitz, S.D., and Bruckstein, A.M.: 'A new method for image segmentation', Comput. Vis. Graph. Image Process., 1989, 46, (1), pp. 82–95

[20] T.Romen Singh, Sudipta Roy, O.Imocha Singh, Tejmani Sinam and Kh.Manglem Singh," A New local Adaptive Thresholding Technique in Binarization", IJCSI-Vol 8, issue 6 No. 2 pp. 271-277 (Nov, 2011).

[21] Liu, Y., and Srihari, S.N.: 'Document image binarization based on texture features', IEEE Pattern Anal. Mach. Intell., 1997, 19, (5), pp. 540–544

[22] L. O'Gorman, "Binarization and multithresholding of document images using connectivity," Graphical Model and Image Processing 56, pp. 494–506, Nov. 1994.

[23] K. Sobottka, H. Kronenberg, T. Perroud, and H. Bunke, "Text extraction from colored book and journal covers,"

Int. Jour. on Document Analysis and Recognition 2, pp. 163–176, June 2000.

[24] C. M. Tsai and H. J. Lee, "Binarization of color document images via luminance and saturation color features," IEEE Trans. on Image Processing 11, pp. 434–451, April 2002.

[25] E. Badekas, N. Nikolaou, and N. Papamarkos, "Text binarization in color documents," Int. Jour. of Imaging Systems and Technology 16(6), pp. 262–274, 2006.

[26] F. Shafait, J. van Beusekom, D. Keysers, and T. M. Breuel, "Page frame detection for marginal noise removal from scanned documents," in 15th Scandinavian Conference on Image Analysis, pp. 651–660, (Aalborg, Denmark), June 2007.

[27] P. Viola and M. J. Jones, "Robust real-time face detection," Int. Journal of Computer Vision 57(2), pp. 137– 154, 2004.

[28] O. D. Trier and T. Taxt, "Evaluation of binarization methods for document images," IEEE Trans. On Pattern Analysis and Machine Intelligence 17, pp. 312–315, March 1995

[29] F. C. Crow, "Summed-area tables for texture mapping," Computer Graphics - Proceedings of SIGGRAPH' 84 18(3), pp. 207–212, 1984.

[30] Y. Nakagawa and A. Rosenfeld, ''Some experiments on variable thresholding,'' *Pattern Recogn.* 11~3!, 191–204 ~1979!.

[31] F. Deravi and S. K. Pal, ''Gray level thresholding using second-order statistics,'' *Pattern Recogn. Lett.* 1, 417–422 ~1983!.

[32] W. A. Yasnoff, J. K. Mui, and J. W. Bacus, ''Error measures for scene segmentation,'' *Pattern Recogn.* 9, 217–231 (1977).

[33] M. Sezgin and B. Sankur, ''Comparison of thresholding methods for non-destructive testing applications,'' *IEEE ICIP'2001, Intl. Conf. Image Process.*, pp. 764–767 (2001).

[34] M. D. Levine and A. M. Nazif, ''Dynamic measurement of computer generated image segmentations,'' *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-7, 155–164 (1985).

[35] Y. J. Zhang, ''A survey on evaluation methods for image segmentation,'' *Pattern Recogn.* 29, 1335–1346 (1996).

[36] E. Badekas and N. Papamarkos, "Document binarization using Kohonen SOM", IET Image Process., Vol. 1, No. 1, March 2007

[37] Mehmet Sezgin and Bu¨ lent Sankur "Survey over image thresholding techniques and quantitative performance evaluation", Journal of Electronic Imaging 13(1), 146–165 (January 2004).

[38] T.Romen Singh, Sudipta Roy, and Kh.Manglem Singh," Local Adaptive Automatic Binarization(LAAB" ,International Journal of Computer Applications (0975 – 8887) Volume 40– No.6, February 2012