

Document Image Binarization Techniques, Developments and Related Issues: A Review

Rashmi Saini

Assistant Professor, Govind Ballabh Pant
Engineering College
Pauri Garhwal, 246001, Uttarakhand, India

ABSTRACT

Image binarization is still a relevant research area due to its wide range of applications in the field of document analysis and recognition. Accuracy of binarization methods affected by many factors such as shadows non-uniform illumination, low contrast, large signal-dependent noise etc. This paper provides a comprehensive survey of major binarization techniques. We also emphasis on the problems being encountered and the related issues in the research area of document image binarization. In addition, some important issues affecting the performance of image binarization methods are also discussed. This literature review suggests that designing a suitable document image binarization method is a prerequisite for a successful document image analysis and recognition.

General Terms

Image Processing, Document image analysis.

Keywords

Image binarization; Thresholding; Segmentation; Optical Character Recognition(OCR).

1. INTRODUCTION

Image binarization is the initial step in many document analysis and recognition system, its objective is to extract the text from image background. This Bi-level representation is preferred for document analysis and recognition. Usually, it is the initial step in major document analysis problem and the performance of subsequent character segmentation and recognition system depends on the accuracy of binarization method. Kefali et al. [28] have claimed that the binarization process aims to decrease the presence of unwanted data and preserve only the desired data in the document images. Complexity reduction achieved by the binarization process is a major motivation for including it in virtually all documents processing area Moghaddam et al. [10]. A fast and accurate document image binarization technique is important for the ensuing document image processing tasks such as optical character recognition (OCR) Bolan Su et al. [22]. Sezgin and Sankur [4] classified the image binarization methods into six categories: entropy-based methods histogram-based methods, foreground attribute-based methods, spatial binarization methods clustering-based methods, and locally adaptive methods. Thresholding is one of the popular technique for image binarization and large number of methods are proposed in literature based on this technique. Thresholding is performed usually in two ways: either globally or locally. In global thresholding, one threshold value is computed to segment the image whereas in local thresholding many different threshold values are computed to divide the image into object and background. Global thresholding methods has a good performance in the case that there is a clear separation between the foreground and the background[14].

In case of document image binarization, global thresholding methods are not sufficient since document images usually are degraded and have poor quality including shadows, non-uniform illumination, smear and strains, low contrast, large signal-dependent noise B. Gatos [14]. To address the such issues local image statistics may guide threshold value in adaptive thresholding techniques in order to preserve textual information for better segmentation result. One of the powerful global thresholding method is Otsu's method but it is non-adaptive in nature. This method determines the global threshold value based on maximum of the between-class variances. Although there are many locally adaptive methods are in literature that computes threshold value pixelwise and having parameters to estimate in order to deal with a high degree of variability. Development of a universally acceptable, high-performance image binarization technique is a challenging task. A large number of document image binarization techniques have been proposed in literature. The majority of binarization methods are complex and are compounded from filters and existing operations [7]. However, the simple binarization methods are either suffers from the poor segmentation or need to set the parameters manually. In this paper, we discuss several image binarization technique, concepts, issues, advantages, disadvantages, challenge of different binarization techniques. The paper is organized as follows: Section 2 include different document image binarization techniques with an emphasis on advantages and disadvantages of various methods. Conclusions are drawn in Section 4.

2. IMAGE BINARIZATION TECHNIQUES

2.1 Niblack's Method

Niblack [16] proposed an algorithm that calculates a threshold value for each pixel by shifting a rectangular window across the image. The threshold T for the center pixel of the window is computed using the mean and the variance of the intensity values within the window

$$T = m + k\sigma \quad \dots (2.1)$$

where k is a constant (value between 0 to 1) set to -0.2 by Niblack, window size may be determined by the user; for example, the window size is defined as 15×15 , m is the mean value and σ is the standard deviation value of the pixels inside the window Sezgin and Sankur[4]. The quality of binarization depends on the parameter k and size of sliding window. This method produces thick and unclear strokes with a small k value, and slim and broken strokes with a large k value.

Niblack's method is one of the most commonly used methods and it is the basic principle behind several other method proposed in literature. This method detect the text body effectively in case of low contrast too. One main limitation of

this method is that it produces a large amount of noise in the empty windows. On the other hand it also requires the window size and values of factors k to be determined manually.

2.2 Sauvola's Method

Sauvola [3] proposed a method that was inspired from Niblack's method. This method aims to solve the problem of noise depending on the impact on the standard deviation value by using a range of gray-level values in the images. The threshold value is computed by following equation

$$T = m \left(1 - k \left(1 - \frac{\sigma}{R} \right) \right) \quad \dots (2.2)$$

where R is the dynamics of the standard deviation fixed to 128 and k is a factor takes on positive values (usually set to 0.5). This method solve the problem of noise up to great level. However, it fails if the contrast between the foreground and background is low or if the text is in form of thin pen stroke. Sauvola's Method also requires the values of the factors k and the window size to be determined manually.

2.3 Bernsen's Method

Bernsen [11] proposed a method that is based on image contrast. The threshold value is calculated at midrange that is mean of minimum $I_{min}(i, j)$ and maximum gray value $I_{max}(i, j)$ in the local window (window size =31). If the contrast $C(i, j) = I_{max}(i, j) - I_{min}(i, j)$ is below a certain contrast threshold value, the pixels within the window may be set to background or to foreground according to the class that most suitably describes the window [11].

This binarization method is dependent on the value of k and also on the size of the window.

$$T(i, j) = 0.5 \{ \max_w [I(i + m, j + n)] \min_w [I(i + m, j + n)] \} \quad \dots (2.3.1)$$

Where window size is 31

$$\text{contrast} = C(i, j) = I_{max}(i, j) - I_{min}(i, j) \geq 15 \quad \dots (2.3.2)$$

2.4 Wolf's Method

Wolf et al. [26] proposed a method to address the problems of Sauvola Method [3], the proposed algorithm normalize the contrast and the mean value of the image. Threshold value is computed by the following equation:

$$T_{wolf} = (1 - k) * m + k * M + k * \frac{s}{R} (m - M) \quad \dots (2.4)$$

where k is fixed to 0.5, M is the minimum gray value of the image and R is set to the maximum gray-value standard deviation obtained over all the local neighborhoods (windows) [26]. Wolf's method perform well in most of the cases, but it has been observed that the performance of this method is degraded if there is a sharp change in background intensity values across the image. The reason of this is values of the parameter M and R are computed from whole image. So even a small noisy patch could significantly influence the value of M and R , thus eventually calculating misleading binarization thresholds [25].

2.5 Feng's Method

Feng et al. [27] proposed a method that is based on the concept of two local window: one contained within other. Wolf's method calculates dynamic range of gray-value standard deviation from the whole image, while this method calculate it locally. The values of standard deviation s , the minimum gray-level M and local mean m are calculated in the

primary local window while the dynamic range standard deviation R_s is calculated in the larger window that is secondary window [27]. Threshold value is calculated by using following equation

$$T = (1 - \alpha_1)m + \alpha_2 \left(\frac{s}{R_s} \right) (m - M) + \alpha_3 M \quad \dots (2.5)$$

Where $\alpha_2 = k_1 (s/R_s)^\gamma$ and $\alpha_3 = k_2 (s/R_s)^\gamma$. Author set the value of γ is 2. Values of other parameters α_1 , k_1 and k_2 are proposed to be in the ranges 0.1-0.2, 0.15-0.25 and 0.01-0.05 respectively. Feng's method addresses the problem of Wolf's algorithm. However this method introduced three parameters, determined empirically, leaves the robustness of this method questionable [25].

2.6 Gatos' Method

Gatos' et al. [14] proposed a method that is based on local adaptive approach for image binarization and enhancement of degraded documents. This method does not require any parameter tuning by the user and can deal with degradations which occur due to shadows, large signal-dependent noise, non-uniform illumination, smear and strain and low contrast. This method consist the following distinct steps. Step 1: Preprocessing stage of is required for the elimination of noisy areas, smoothing of background texture and contrast enhancement. Author used low-pass Wiener filter for this purpose. Pre-processing stage of this method use an adaptive Wiener filter based on statistics estimated from a local neighborhood. Step2: A rough estimation of foreground regions, in order to proceed to an initial segmentation of foreground and background regions that provide a superset of the correct set of foreground pixels (Sauvola's method [3]). Step3: A background surface calculation by interpolating neighboring background intensities. Step 4: A thresholding by combining the calculated background surface with the original image. For better quality binary image this method incorporate an efficient upsampling technique (bi-cubic interpolation) that provides satisfactory results. Last step: Perform to post-processing (shrink and swell filtering) of the resulting binary image in order to eliminate noise, improve the quality of text regions and preserve stroke connectivity by isolated pixel removal and filling of possible breaks, gaps or holes [14]. Gatos's method demonstrated superior performance against well-known techniques specially on degraded document images.

2.7 Darek bradley's Method

Darek Bradley et al. [12] proposed a method for adaptive thresholding using the integral image of the input. Adaptive thresholding is used in order to take into account spatial variations in illumination. Proposed method is an extension of a previous method (Wellner' method [29]). However, this technique is more robust to illumination changes in the image as well as simple and easy to implement. To compute the integral image, store at each location, $I(x, y)$, the sum 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 [12].

$$I(x, y) = f(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1) \dots (2.7.1)$$

After converting the given image into integral image, sum of the function for any rectangle with upper left corner (x_1, y_1) and lower right corner (x_2, y_2) can be computed in constant time using the following equation,

$$\sum_{x=x_1}^{x_2} \sum_{y=y_1}^{y_2} f(x, y) = I(x_2, y_2) - I(x_2, y_1 - 1) - I(x_1 - 1, y_2) + I(x_1 - 1, y_1 - 1) \quad \dots (2.7.2)$$

This thresholding method is an extension of Wellner's method [29]. The main concept of Wellner's method is that each pixel is compared to moving average of last s pixel seen average of the surrounding pixels. Main advantage of this technique is that only a single pass through the image is required. One major problem with this method is that binarization result is dependent on the scanning order of the pixels. Further, the moving average is not a good representation of the surrounding pixels. Proposed technique compute the average of an $s \times s$ window of pixels centered around each pixel. This is a better average for comparison since it considers neighboring pixels on all sides. The average computation is accomplished in linear time by using the integral image [12]. In first pass this method calculate the integral image through the input image. In a second pass, compute the $s \times s$ average using the integral image for each pixel in constant time and then perform the comparison to assign a pixel in either foreground or background category. Bradley's method of image binarization perform well for uneven illumination images and suitable for processing live video streams at a real-time frame-rate, it a valuable tool for interactive applications such as augmented reality [12].

2.8 NICK's Method

The method developed by Khurshid et al. [26] is known as NICK's method. This method is a modified version Niblack's method. The goal of this method is to solve the problem of black noise in Niblack's method and the low contrast problem of Sauvola's method by shifting the thresholding value downward [7].

The thresholding value is calculated by following equation

$$T = m + k \sqrt{\frac{\sum pi^2 - m^2}{N}} \quad \dots(2.8)$$

Where k is a parameter with value in the range $[-0.2, -0.1]$, P_i is the pixel value of the gray-scale levels in the image, and N is the total number of pixels in the image and window size is 19×19 .

Parameter k value close to -0.2 make sure that noise is eliminated but characters can break a little bit, while with values close to -0.1 , some noise pixels can be left but the text will be extracted crisply and unbroken [25]. Khurshid suggested that for an OCR application, 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 solves the problem of black-noise and also perform well in case of low contrast. However, it still fail when the contrast is too small or the text is in thin pen stroke text [7]. This method also suffer from the problem of determining the factors manually.

2.9 Bilal's Method

Bilal Bataineh et al [7] proposed a method is to address all binarization challenges and to overcome the weaknesses of the major existing methods. In line with Niblack's, Sauvola's and NICK's inspirations, author proposed a thresholding equation used the mean value and the standard deviation as well as effective factors that help to address the major challenges of binarization [7]. Proposed method is particularly effective in solving the problems of thin pen stroke text and low contrast between the foreground and background. This method adopted the local approach and use two elements: a dynamic method to divide the image into the windows based on the characteristics of the image and another method to determine the appropriate threshold value within each window in order to binarize the image. This method proposed a

thresholding equation used the mean value and the standard deviation as well as effective factors that help to address the various challenges of binarization. The proposed adaptive binarization method is as follows [7]: Step 1. Apply thresholding confusion equation onto global image Step 2. Classify the global image into three levels black, red or white pixels Step 3. Determine primary window size based on probability of red over black pixels Step 4. Check appropriateness of each primary window size. If population of red pixels higher than black pixels then create secondary window size Step 5. Apply thresholding window equation for each adaptive window size or local image

Bilal's method gave the best performance compared to Niblack's method, Sauvola's method, and the NICK's method. This method also solved the problem of identifying the factors and size of windows while in earlier methods these factors are either fixed by author or decided by user. It is highly adapted to deal with many binarization problems, such as thin pen stroke and low-contrast images.

3. CONCLUSION

Binarization is an essential step for document image analysis and recognition system. In general, a large number of binarization techniques are proposed to address different types of binarization problems. The global thresholding techniques do not provide satisfactory binarization result for document images that usually suffer from poor quality including shadows, non-uniform illumination, low contrast and signal-dependent noise etc. Adaptive thresholding techniques are used to address such issues and to preserve textual information for better segmentation result. Most of the image binarization techniques proposed in literature are complex and are compounded from filters and existing operations [7]. However, the simple binarization methods are either suffer from the poor segmentation or need to set the parameters manually (window size, factor k). In this paper, we have discussed eight image binarization techniques, concepts, issues, advantages, disadvantages, challenge of different binarization methods. Following conclusions are drawn.

- Niblack's approach, the resulting binary image generally suffers from a great amount of background noise especially in case of empty window as well as user have to set parameters manually.
- Sauvola's method. solves the problem of background noise but in many cases characters become extremely thinned and broken.
- Gatos's method provide superior performance compared with all other methods (Niblack, Sauvola) and performs well for noisy and degraded document image. At the same time this method does not require any parameter tuning by the user.
- Bradley's method of image binarization perform well for uneven illumination images and suitable for processing live video streams at a real-time frame-rate, it is suitable for interactive applications [12].
- Wolf's algorithm perform well for many images, however there are occasions when the characters disappear or break if intensity variations are very low.
- Feng's method generally works very well but the main drawback remains its susceptibility to the

empirically determined parameter values. A small change in parameter values could drastically affect the binarization results.

- NICK's Method perform well as compared to the other methods, and it performs better specially when the images have extremely low intensity variations and for whiter images[25]. NICK is much more efficient as compared to Wolf's and Feng's method.
- This mehod outperforms the other sliding-window thresholding methods in efficiency and robustness for low quality document images.
- Bilal's method gave the best performance as compared to Niblack, Sauvola, and NICK's method.
- One major problem of the earlier method is some parameter (window size) are either fixed by author or decided by user. Bilal's method solve such problems and highly adapted to deal with many binarization problems, such as thin pen stroke and low-contrast images.

4. REFERENCES

- [1] N. Otsu, "A threshold selection method from gray-level histograms", IEEE Transaction on Systems Man and Cybernetics (SMC), Vol. 9, pp. 62-66, 1979.
- [2] J. Kittler, M. Hatef, R. Duin, J. Matas, On combining classifiers, IEEE Trans. Pattern Anal. Mach. Intell. 20 (3) (1998) 226-239.
- [3] J. Sauvola, M. Pietikainen, "Adaptive Document Image Binarization", Pattern Recognition, Vol. 33, pp. 225- 236, 2000.
- [4] Mehmet Sezgin, B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", International Journal of Electronic Imaging(IJEI) Vol. 13 No. 1, pp.146-165, 2004.
- [5] Graham Leedham, Chen Yan, Kalyan Takru, Joie Hadi Nata Tan and Li Mian, "Comparison of Some Thresholding Algorithms for Text/Background Segmentation in Difficult Document Images", IEEE International conference on Document Analysis and Recognition (ICDAR), pp.859-864, 2003.
- [6] Naveed Bin Rais, M. Shehzad Hanif and rmtiaz A. Taj, "Adaptive Thresholding Technique for Document Image Analysis", 8th IEEE International Multitopic Conference, Vol. 3, pp. 61-66, 2004.
- [7] Bilal Bataineh, Siti Norul Huda, Sheikh Abdullah, Khairuddin Omar, "An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows", Pattern Recognition Letters Vol. 32, pp. 1805-1813, 2011.
- [8] Pierre D. Wellner "Adaptive thresholding for digital desk" Technical Report(EPC) pp. 93-110, 1993.
- [9] W. A. Yasnoff, J. K. Mui, and J. W. Bacus, "Error measures for scene segmentation," Pattern Recognition, vol. 9(4), pp.217-231, 1977
- [10] Moghaddam, R.F., Cheriet, M.: AdOtsu: an adaptive and parameter less generalization of Otsu's method for document image binarization. Pattern Recogn. 45(6), 2419-2431 (2012)
- [11] Bernsen, J.: Dynamic thresholding of gray level images. In: Proceedings of International Conference on Pattern Recognition (ICPR), pp. 1251-1255 (1986)
- [12] Bradley, D., Roth, G.: Adaptive thresholding using the integral image. J. Graph. Tools 12(2), 13-21 (2007)
- [13] Anjos, A., Shahbazkia, H.: Bi-level image thresholding—a fast method. Biosignals 2, 70-76(2008)
- [14] Gatos, B., Pratikakis, I., Perantonis, S.J.: Adaptive degraded document image binarization. Pattern Recogn. 39(3), 317-327 (2006)
- [15] Kapur, N.J., Sahoo, P.K., Wong, C.K.A.: A new method for gray-level picture thresholding using the entropy of the histogram. J. Comput. Vis. Graph. Image Process. 29(3), 273-285 (1985)
- [16] Niblack, W.: An introduction to digital image processing, pp. 115-116. Prentice Hall, Eaglewood Cliffs (1986)
- [17] Ntirogiannis, K., Gatos, B., Pratikakis, I.: A combined approach for the binarization of handwritten document images. Pattern Recogn. Lett. 35, 3-15 (2014).
- [18] Stathis, P., Kavallieratou, E., Papamarkos, N.: An evaluation technique for binarization algorithms.J. Univers. Comput. Sci. 14(18), 3011-3030 (2008)
- [19] Ntirogiannis, K., Gatos, B., Pratikakis, I.: An objective evaluation methodology for document image binarization techniques. In: 8th IAPR Workshop on Document Analysis Systems (2008)
- [20] Zhang, C., Yang, J., 2010. Binarization of document images with complex background. In: Proc. 6th Internat. Conf. in Wireless Communications Networking and Mobile Computing (WiCOM), pp. 1-4. 2010
- [21] P.K. Sahoo, S. Soltani, "A survey of thresholding techniques", Computer Vision Graphics Image Processing(CVGIP). Vol. 41, pp. 233-260. 1988.
- [22] Su, B., Lu, S., Tan, C.L.: Robust document image binarization technique for degraded document images. IEEE Trans. Image Process. 22(4), 1408-1417 (2013)
- [23] Mehmet Sezgin, B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", International Journal of Electronic Imaging(IJEI) Vol. 13 No. 1, pp.146-165, 2004.
- [24] Nikhil R.Pal and Sankar K.Pal, "A review on Image segmentation techniques", Pattern Recognition Vol. 26 No. 9 pp. 1277-1294, 1993.
- [25] .Wolf, J-M. Jolion, "Extraction and Recognition of Artificial Text in Multimedia Documents", *Pattern Analysis and Applications*, 6(4):309-326, (2003).
- [26] Meng-Ling Feng and Yap-Peng Tan, "Contrast adaptive binarization of low quality document images", IEICE Electron. Express, Vol. 1, No. 16, pp.501-506, (2004).
- [27] Kefali, A., Sari, T., Sellami, M., Evaluation of several binarization techniques for old Arabic documents images. In: The First Internat. Symp. on Modeling and Implementing Complex Systems (MISC'2010), Algeria, pp. 88-99 , 2010
- [28] Wellner, P. D. Adaptive thresholding for the digitaldesk. Tech. Rep. EPC-93-110, EuroPARC, 1993