IEEE International Conference on Advances in Engineering &Technology Research (ICAETR - 2014), August 01-02, 2014, Dr. Virendra Swarup Group of Institutions, Unnao, India

A Proposed Method For Document Image Binarization Based on Bit Plane Slicing

Karthika M
Department of Computer Science and Engineering,
Government Engineering College Thrissur,
Kerala, India 680009
minika89@gmail.com

Ajay James
Department of Computer Science and Engineering,
Government Engineering College Thrissur,
Kerala, India 680009
ajayjames80@gmail.com

***Abstract*— Document image binarization has not been an easy task in the area of historical document restoration accounting to the difficulty in choosing a threshold value when the document has noises due to say, its degradation, ageing of the paper, other unwanted impurities, background images, bleed-through, and so on. Also, in case of handwritten documents there could be inconsistency in the darkness of the letters. Though many algorithms have been proposed for Document image binarization, there has not been an effective algorithm which can efficiently binarize all kinds of documents. In this paper we propose a method to perform document image binarization using bit-plane slicing. The proposed method is deployed based on the rule of divide-and-conquer. As per this technique, the 8 bit planes of the gray image are extracted and processed separately, at the end of which, the results are combined to give the final binarized output.**

Keywords*— *Document image binarization; Bit plane slicing; Thresholding

I. INTRODUCTION

Historical document restoration and translation are gaining importance nowadays. Together with the up rise of their importance, the challenges in these processes are also increasing. The earlier methods to restore documents were to directly process the document sheets through chemical processing. However, this increases the risk of damaging the historical asset. Document Image Analysis and processing emerged as an outcome of this.

The ultimate aim of document image binarization, as the name suggests, is to convert the document image into an image with just 2 colors (bi-colored) such that the foreground text pixels are black and the background is white. Though it is only required to threshold the document image, due to various reasons this has not been an easy task. The reasons include improper illumination, coloring due to aging, inter-intra intensity variations in text pixels and background, seep-through, faded ink, and diversity in the type of document to be processed. Due to these facts this has turned to a research problem [2].

II. APPROACHES

A. Global Thresholding

Global thresholding methods find a threshold from the information of an entire image [15]. Otsu's method [1] was one of the primitives of global thesholding methods. The Otsu's algorithm assumes that the image to be thresholded contains two classes of pixels or bi-modal histogram then calculates the optimum threshold separating those two classes so that their combined spread (intra-class variance) is minimal [1]. Global thresholding methods, though produce good results when the document background is invariant, these methods fail considerably when the image background is complex. This is because no simple global threshold value can fit for all characters accounting to the fact that the intensity of noise at one region may be the same as the intensity of some text pixels in another region.

B. Adaptive Thresholding

Adaptive thresholding estimates a local threshold for each document image pixel [8] [13] [4] [14]. A main drawback of such approaches is the time taken to produce the results. Window Based Adaptive Thresholding techniques [19] estimate the local threshold by using mean and the standard variation of image pixels within a local neighbourhood window. The drawback of these techniques is that they heavily rely on window size and hence character stroke width [4].

C. Others

Some others such as [2] use Laplacian energy or gradient as an approach for binarization. Though the methodology used in [2] produces results with better clarity compared to that by [4], it fails in identifying faint ink strokes in case of handwritten documents and some letters in case of printed documents. Some other approaches such as in [15], image color features such as saturation and hue are considered so as to perform binarization. But such algorithms have the limitation that they

978-1-4799-6393-5/14/$31.00 ©2014 IEEE

can be applied only to color images with reasonable hue and saturation values. They also fail in processing documents with half-tone colors and decaying background. Fuzzy and neaural network approaches are among the other approaches used to binarize document images.

III. Literature Review

Thresholding algorithms can be categorized into two basic classes: global and local. Global thresholding techniques use a single threshold value for the whole image, whereas local methods have threshold values for a set of pixels that are dependent on some characteristics of the pixels in the neighbourhood.

Among the global methods, Otsu's algorithm [1] is accepted as a benchmark. As per this method, the image pixels are classified into two classes, the foreground and the background, such that the inter-class variance between them is the maximum and the intra-class variance is the minimum.

Nina *et.al* [30] propose a recursive version of Otsu method for document image binarization. Ying Liu and Sargur N. Srihari [20] propose to perform Otsu's method iteratively on the gray-scale image to obtain candidate threshold values, from which texture features associated with each candidate threshold are extracted from the run-length histogram of the accordingly binarized image. The optimal threshold is selected so that desirable document texture features are preserved.

Srinivasa Rao *et.al* [22] propose a method to deal with complex backgrounds. The idea is to use two threshold values chosen heuristically and a combination of these thresholds, instead of choosing a single threshold value. If the two threshold values chosen are T1 and T2, the pixels whose intensities are greater than T1 are eliminated. The pixels with intensities less than T2 are preserved and the ones with a value between T1 and T2 are preserved if they are located near a pixel with intensity less than T2, and deleted otherwise.

Global thresholding performs well in the case of the presence of a good separation between the foreground and the background. However, very often, document images are exposed to degradations that account for loss of such a separation.

Adnan and Boran [23] propose a global binarization algorithm based on neural networks. As per this method, thresholding is first applied using Mass-Difference thresholding to obtain many local threshold values. The neural network is then trained using these values as input and a single global threshold value as output. This method has a disadvantage that the results obtained have several false positives.

Niblack [24] uses an adaptive binarization method in which the threshold value varies over the image based on the local mean, m and local standard deviation, s, in a small neighbourhood window of each pixel. The threshold, T, is calculated as: T=m+k*s, where k is a user-defined parameter and may assume negative values. However, this method fails when the document image consists of background containing light texture. Sauvola [19] uses a modified version of Niblack algorithm, in which, the threshold, T, is computed as: T=m*[l+k*(s/R-I)], where R is the dynamic range of standard deviation, and parameter k gets positive values.

One of the main issues concerned with adaptive binarization is the choice of window size. In [25], an approach to choose the window size is mentioned. The idea is based on the theory of wavelet transforms. Wavelet transforms is used for segmenting coarse objects from document images. The scales associated with them can also be calculated. [25] suggests to group and classify pixels that express the same scale features. Using scale as the determinant, the neighborhood size is determined.

Ying Liu *et.al* [21] discusses a hybrid method involving both global and local thresholds. As per this method, a set of global thresholds are chosen based on some global characteristics. Among these global threshold values, a single value is chosen based on local image features. Though this method works better than global thresholding methods, this method fails if it fails to properly identify the set of global threshold values.

Another hybrid approach is suggested by Ergina and Stamatatos [26] in which a global threshold is first applied after which areas still contain noise. These areas are separately processed for obtaining better results. However, as with global thresholding methods, handling of text stroke pixels and noise pixels which are of the same intensity are not handled properly.

Gatos *et.al* [12] proposes an adaptive document image binarization method which involves several distinct steps: a pre-processing procedure using a low-passWiener filter, a rough estimation of foreground regions, a background surface calculation by interpolating neighboring background intensities, a thresholding by combining the calculated background surface with the original image while incorporating image up-sampling and finally a post-processing step in order to improve the quality of text regions and preserve stroke connectivity.

Pratikakis [8] proposes a hybrid binarization algorithm that combines some well performing document image binarization algorithms and a proper edge detection algorithm. Some morphological operations are performed on the resultant image to further improvise the results.

Thibault Lelore and Fr´ed´eric Bouchara [27] propose a super-resolution method in which the algorithm tries to super resolve the likelihood of text instead of the gray value of pixels. The computation of the likelihood is based on a rough estimation of the text position based on the modified version of the Canny's algorithm. A super-resolved version of this image is then computed by using a projective model. Finally, the super resolved binarized image is computed.

Hoang *et.al* [28] propose a nonparametric local adaptive binarization method for poorly degraded document images. It firstly estimates the contrast image through a morphological closing operator. Then, two threshold values are used to classify the pixels into three regions corresponding to text,

978-1-4799-6393-5/14/$31.00 ©2014 IEEE

near-text, and non-text regions. The entropy of the contrast image is the maximized. Then, a local thresholding is applied on the near-text areas to best choose the text pixels.

Jung *et.al* [29] propose an adaptive binarization algorithm for document images suffering from uneven light conditions. However, the output produced by this algorithm blurred characters thereby making it useless for OCR.

As per the method suggested by Su-Lu-Tan [4], an adaptive contrast map is first constructed after which, the text stroke edges are detected through the combination of the binarized adaptive contrast map and the Canny's edge detection technique. The image is then binarized based on the local threshold, which is estimated from the detected text stroke edge pixels. Some post-processing is done so as to improve the document binarization quality.

Local (adaptive) thresholding techniques have been widely used in document image binarization because they have exhibit a better performance in extracting the character strokes from a document image that contains spatially uneven gray levels due to degradations. However, a common drawback with these approaches is their dependency on the size of neighbourhood window chosen.

Marco Block *et.al* [3] and Howe [2] uses Laplacian energy to binarize document images. The former method gives good results compared to any other global thresholding methods, even in case of document images containing shadows, while it fails to properly binarize document images suffering from poor contrast. Howe's method uses Laplacian energy to identify text edge pixels, which is then combined with Canny's edge detection method to further improvise the results. Though Howe's method works better than any other global and adaptive binarization methods, it fails in identifying faint text strokes.

Thibault Lelore [11] and *Cao et.al* [10] presents an approach for document image binarization based on Markov Random Field (MRF) model of the document. Though this technique exhibits good adaptation into different defect types such as illumination and noise, this technique too fails in identifying faint text strokes.

The idea suggested in [9] is to apply histogram equalization method over the different image partitions at each hierarchical level. A membership value to each pixel at each level of histogram equilization is calculated using some scheme. The individual membership values are then linearly merged to obtain the net membership value of each pixel. Binarization is done based on these membership values.

Lee *et.al* [5], Kavitha *et.al* [6] and Kita *et.al* [7] propose binarization methods based on image color features such as luminance and saturation. Though these give reasonable outputs, these methods do not serve good for practical purposes accounting to the fact that most of the historical documents do not possess contents in multiple colors. The text in such documents would be in a dark color such as blue, brown or black, written on a light background.

IV. Proposed Method

The usual techniques used to binarize a degraded document rely on applying the technique to the document image as a whole. But in the proposed method, the rule of divide and conquer is made use of. The division is with respect to the bit planes that a gray scale digital image is constituted of. In a gray scale digital image, in which the intensity ranges from 0 to 255, a pixel value can be represented using 8 bits. This makes it possible for the whole image to be divided into 8 constituent bit plane images, where Plane 1 contains the lowest order bit of all the pixels in the image and plane 8 contains the highest order bits of all the pixels in the image. The reconstruction of the image from its 8 bit planes is done by multiplying the pixels of the n^{th} plane by the constant 2^{n-1}. Bit plane slicing, which is a variant of intensity level slicing, yields the result of contribution of each bit of the 8-bits of every pixel that represents the intensity in a gray-level image. The technique of bit plane slicing has been employed in [16] for removal of Gaussian noise from high contrast images, and in [17] for restoration of damaged digital images.

To binarize an image, processing is done on each of these bit planes and their results are combined to produce the binarized version of the input document image. This processing includes calculation of inverted image of each of the bit planes produced, done by including a bias value. These inverted images would be mostly constituted of background and noise pixels. After this, the difference of the gray-scale image and the image obtained through bit plane slicing is calculated. Together with the image difference calculated so, it is required to have another image difference which is theoretically the inverse of the image just obtained. Once the difference images are formed, the edge maps of all these 16 images are created using Canny's edge detection technique. The edge map formed using Canny's algorithm has the advantage that it detects even faint ink strokes but the disadvantage is that most of noise pixels are also considered in this process, resulting in the edge map to have many false positives. To overcome this, we also calculate the Laplacian of these images. This will improvise the binarization results, as the Laplacian measures the divergence of the gradient which helps to separate low intensity ink pixels from the high intensity background pixels.

Binarized images are then constructed from these edge maps using the method mentioned in [2], which uses graph cut algorithm proposed in [18]. These two binarized images are combined to enhance the binarized results. The result is enhanced by adding the edge information from the difference images and the original gray scale image.

Once the binarized images are formed in this way for all the bit planes, the results are combined to give the final binarized image.

Fig. 1 shows the pictorial representation of the proposed method.

978-1-4799-6393-5/14/$31.00 ©2014 IEEE

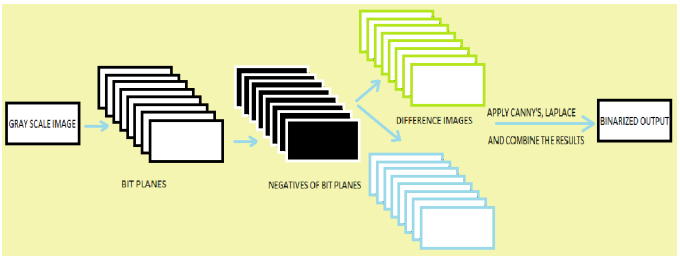

Fig. 1. Proposed Method.

V. Conclusion

A new methodology for document image binarization, based on an image enhancement technique, viz., bit-plane slicing was proposed. This method is expected to be easy to implement and to exhibit good performance. The algorithm shall be capable of handling different kinds of images of documents, both printed and handwritten, with different kinds of noises.

Acknowledgment

The authors would like to thank all friends and colleagues for having extended their help and cooperation during the preparation of the paper.

References

[1] NOBUYUKI OTSU, A threshold selection method from gray level histogram, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS 9 (1978), no. 1, pp. 62-66.

[2] Howe, Nicholas R, A Laplacian energy for document binarization, IEEE 2011 International Conference on Document Analysis and Recognition (ICDAR), 2011, pp. 6-10.

[3] Block, Marco and Rojas, Ra´ul, Local Contrast Segmentation to Binarize Images, ICDS 9, Vol. 9, 2009, pp. 294-299.

[4] Su, Bolan and Tan, C and Lu, S, A robust document image binarization technique for degraded document images, IEEE Transactions on Image Processing, 2013, Vol 22, pp.1408-1417.

[5] Tsai,C. M. and Lee, H. J., Binarization of color document images via luminance and saturation color features, IEEE Transactions on Image Processing, 2002, Vol 11, pp. 434-451.

[6] Madhubalan, Kavitha and Lee, Gueesang, Binarization of Degraded Characters Using Tensor Voting Based Color Clustering, 2011 IEEE 11th International Conference on Computer and Information Technology (CIT), 2011, pp. 299-305.

[7] Kita, Kohei and Wakahara, Toru, Binarization of color characters in scene images using k-means clustering and support vector machines, 2010 20th International Conference on Pattern Recognition (ICPR), 2010, pp. 3183-3186.

[8] Gatos, Basilios and Pratikakis, Ioannis and Perantonis, Stavros J, Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information, 19th International Conference on Pattern Recognition, 2008. ICPR 2008, pp. 1-4.

[9] Satapathy, SureshChandra and Avadhani, P S and Abraham, Ajith, Binarization of Document Images Using Hierarchical Histogram Equalization Technique with Linearly Merged Membership Function, Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2012, 2012, Vol 132, pp. 639-647.

[10] Cao, Huaigu and Govindaraju, Venu, Preprocessing of low-quality handwritten documents using markov random fields, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, Vol 31, pp. 1184-1194.

[11] Lelore, Thibault and Bouchara, Fr´ed´eric, Document image binarisation using markov field model, IEEE 10th International Conference on Document Analysis and Recognition, 2009, pp. 551-555.

[12] Gatos, Basilios and Pratikakis, Ioannis and Perantonis, Stavros J, Adaptive degraded document image binarization, Pattern recognition 39, Elseiver, 2006, Vol 39, pp. 317-327.

[13] Su, Bolan and Lu, Shijian and Tan, Chew Lim, Binarization of historical document images using the local maximum and minimum, Proceedings of ACM International Workshop on Document Analysis Systems, 2010, pp. 251-260.

[14] S. Lu, B. Su, and C. L. Tan, Document image binarization using background estimation and stroke edges, International Journal on Document Analysis and Recognition (IJDAR) 13 (2010), no. 4, pp. 303-314.

[15] C. M. Tsai and H. J. Lee, Binarization of color document images via luminance and saturation color features, IEEE Transactions on Image Processing, 11 (2002), no. 4, pp. 434-451.

[16] R. Agarwal, Bit plane average filtering to remove gaussian noise from high contrast images, International Conference on Computer Communication and Informatics (ICCCI), 2012 , Jan 2012, pp. 1-5.

[17] E. Ardizzone, H. Dindo, and G. Mazzola, Restoration of digitized damaged photos using bit-plane slicing, 2007 IEEE International Conference on Multimedia and Expo, July 2007, pp. 1643-1646.

[18] Boykov, Y., Kolmogorov, V., An experimental comparison of mincut/maxflow algorithms for energy minimization in vision, IEEE Trans. on Pattern Analysis and Machine Intelligence 26 (9), pp. 1124-1137 (2004).

[19] Sauvola, Jaakko, and Pietikinen. Matti, Adaptive document image binarization, Pattern Recognition 33, Elsevier (2000), no. 3, pp. 225-236.

[20] Liu, Ying and Srihari, Sargur N, Document image binarization based on texture features, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, Vol 19, pp. 540-544.

[21] Liu, Ying and Feinrich, R and Srihari, Sargur N, An object attribute thresholding algorithm for document image binarization, Proceedings of the IEEE Second International Conference on Document Analysis and Recognition, 1993, pp. 278-281.

[22] Rao, A and Sreenivasu, Tinnathi and Rao, Nenkkanti Venkata and Sastry, ASCS and Reddy, L Pratap and Prabhu, TS Krishna, Binarization of Documents with complex Backgrounds, IEEE Second International Conference on Machine Vision, 2009, ICMV'09, 2009, pp. 8-12.

[23] Khashman, Adnan and Sekeroglu, Boran, Global binarization of document images using a neural network, Third International IEEE Conference on Signal-Image Technologies and Internet-Based System, 2007. SITIS'07, 2007, pp. 665-672.

[24] Niblack, Wayne, An introduction to digital image processing, Strandberg Publishing Company, 1985.

[25] Chang, Fu and Liang, Kung-Hao and Tan, Tzu-Ming and Hwan, Wen-Liang, Binarization of document images using Hadamard multiresolution analysis, Proceedings of the Fifth International Conference on Document Analysis and Recognition, 1999. ICDAR'99, 1999, pp. 157-160.

[26] Kavallieratou, Ergina and Stathis, Stamatatos, Adaptive binarization of historical document images, 18th International Conference on Pattern Recognition, 2006. ICPR 2006, 2006, Vol 3, pp. 742-745.

[27] Thibault Lelore and Fr´ed´eric Bouchara, Super-resolved binarization of text based on the fair algorithm, 2011 International Conference on Document Analysis and Recognition (ICDAR), 2011, pp. 839-843.

[28] Le, T Hoang Ngan and Bui, Tien D and Suen, Ching Y, Ternary entropybased binarization of degraded document images using morphological operators, 2011 International Conference on Document Analysis and Recognition (ICDAR), 2011, pp. 114-118.

[29] Kuk, Jung Gap and Cho, Nam Ik, Feature based binarization of document images degraded by uneven light condition, 10th International Conference on Document Analysis and Recognition, 2009. ICDAR'09, 2009, pp. 748-752.

[30] Nina, Oliver and Morse, Bryan and Barrett, William, A recursive Otsu thresholding method for scanned document binarization, 2011 IEEE Workshop on Applications of Computer Vision (WACV), 2011, pp. 307-314.

978-1-4799-6393-5/14/$31.00 ©2014 IEEE