International Journal of Latest Trends in Engineering and Technology
Special Issue SACAIM 2017, pp. 315-318
e-ISSN:2278-621X

A STUDY ON IMAGE BINARIZATION TECHNIQUE FOR DEGRADED DOCUMENTS

Thouseef Ahammad[1], Mohammad Naushad Ali[2] & Mr. Roshan Suvaris[3]

Abstract- Image binarization is a process of converting the text image into binary image containing text as foreground and white as background.Segmentation of the corrupted archive images should be possible by utilizing image binarization techniques to obtain a reasonable image which will be like its original image report. Document images may be degraded due to fading, document aging poor quality of paper, ink plot, Visibility of document degrades due to the scanning and printing that means it is very difficult to understand them. Degraded document image restoration plays a very important role in enhancing the degraded noise in images. Document Image Binarization is the major technique for the segmentation of the text values from the background images. Separation of the pixel values into black as foreground and white as background can be acquire to extract the clear image. Thresholding process is the well-known technique used for the binarization of the document images. Global and local thresholding technique are the two important types of the thresholding process. There are many techniques have been proposed for this binary documentation process. Thus this work presents a survey about the document image binarization techniques.

Keywords – Thresholding, binarization, segmentation, contrast enhancement, Hybrid, Morphological, Stroke Edges

1. INTRODUCTION

Image binarization tries to remove just the content stroke pixels from the gray scale images, and is normally performed in the record preprocessing stage [1].Binarization is important for the ensuing document image processing tasks such as optical character recognition and document layout analysis. Text Image binarization could be a crucial step in document image process. For the purpose of optical character recognitions, when we working on RGB images, we will get a big range of colors. RGB images contain various backgrounds with various foregrounds and so when we performing OCR on an RGB image we may miss some foreground data. Similarly in grayscale converted images, they will also have a range of pixels. All the images foreground data will not be of the same color and hence grayscale images will also have some losses during OCR. To overcome this problem in OCR, binarization method are used. Text Image binarization involves a task that automatically converts the document images from a grayscale or color image into a binary image in a manner that foreground info is represented by black ones and background info by white pixels. This process of thresholding applies to allow document to be recognized and retrieved more efficiently [2].

2. METHODOLOGY

2.1 Binarization of degraded document based on contrast enhancement –

A binarization of degraded document based on contrast enhancement was proposed in this paper which is focuses on the contrast all the time. For easy distinguished areas, contrast enhancement with a smaller degree is used to magnify the difference between foreground and background.

Meanwhile, contrast enhancement with a smaller degree has a great effect on restraining noise. For hard distinguished areas, contrast enhancement with a relatively larger degree is deserve to use to adjust gray values so that it can easily distinguish foreground and background. Hence, no matter what kind of the area is, there is always an exclusive method to gain a satisfactory result. What's more, the proposed method is particularly significant for degraded document images with bleeding-through and severly uneven illumination.

In this paper, assume that A, C, and D are the target areas with characters. For degraded document images, there may also be a significant difference in gray contrast among the rest regions. Hence, further subdivision needs to be done for the rest regions. For instance subregion A, it will be divided for the second time after coarse division. If the maximum grayscale contrast in subregion AB is less than 1 k times maximum grayscale contrast of the former division(such as formula 8), it means there is no significant variance in this subregion. So this subregion is also background and outputted directly.

$$CABmax(x,y) \leq k1 .CAmax(x,y) \quad (8)$$

If the maximum grayscale contrast in subregion AA is more than 2 k times maximum grayscale contrast of the former division (such as formula 9), it means this subregion has a significant variance. So week contrast enhancement

[1] MCA V Semester, St Aloysius College, AIMIT, Mangaluru, Karnataka, India
[2] MCA V Semester, St Aloysius College, AIMIT, Mangaluru, Karnataka, India
[3] Asst. Professor, Department of MCA, St Aloysius College, AIMIT, Mangaluru, Karnataka, India

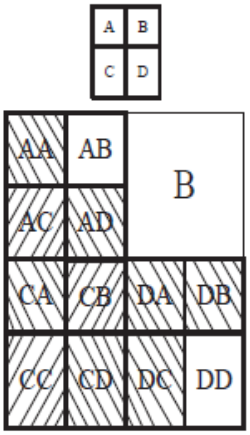
will be used in this subregion.

$$CAAmax(x,y) \geq k2 \cdot CAmax(x,y) \quad (9)$$

If the maximum grayscale contrast in subregion AC is between 1 k and 2 k times maximum grayscale contrast of the former division (such as formula 10), it means that its grayscale variance is comparatively significant. Strong contrast enhancement will be used in this subregion.

$$k \cdot CAmax(x,y) \leq CACmax(x,y) \leq k2.CAmax(x,y)$$

where CABmax(x,y).CAamax (x ,y). and CAAmax (x,y) are the maximum grayscale contrasts of subreions AB, AA, and AC after the second division.CAmax(x, y) is the maximum grayscale contrast of subregion A after the first division, and 2 k is the partition coefficient between the significant and comparatively significant areas.



Background Area with significant grayscale contrast , Areas with comparatively significant grayscale contrast

A large number of experiments showed that the ranges of values for k1 and k2 were k2 ∈[0,0.4] and k2 ∈ [0.7,1] . Twice division is the optimal segmentation number to determine the property of grayscale variance empirically. Too many divisions will lead to large amounts of calculation, and make mistakes between noise and target, and cannot handle noise well either. At the same time, too few divisions will also lose details.

The experimental results and comparisons with otsu, niblack, bernsen, sauvola, singh classical algorithms are also illustrate that the proposed method has a great segmentation quality and short execution time for severely degraded documents[3].

2.2 Hybrid Binarization Technique for Degraded Document Images –

This paper introduces an efficient method to enhance the degraded documents which having various degradedness, like low contrast, uneven illumination, noise and broken letters. The method uses combination of Laplacian operator, Gaussian Blur, Adaptive Bilateral filter and gray scale morphological operations to enhance the low contrast suppressed noise as well as eliminate background and enhance the image respectively. During the first stage of this method Laplacian operator is used to highlights the edges because degraded documents have very low contrast edges. After applying Laplacian operator Gaussian blur smoothing filter is used to blur entire image by using Gaussian function. The Gaussian filter is based on the following equation:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Where x and y are the pixel value, σ controls the level of smoothing The result of Gaussian filter and Laplacian operator are subtracted by using Back ground elimination to eliminating the dark background and noise completely. The result of the previous step is applied to bilateral filter which is a non-linear, edge-preserving, and noise-reducing smoothing filter for images. Bilateral filter is given by the equation

$$I = \frac{1}{C} \sum_{y \epsilon N(x)} e^{\frac{-\|y-x\|^2}{2\sigma_d^2}} e^{\frac{-|I(y)-I(x)|^2}{2\sigma_r^2}} I(y)$$

Where C is

$$C = \sum_{y \epsilon N(x)} e^{\frac{-||y-x||^2}{2\sigma_d^2}} e^{\frac{-|I(y)-I(x)|^2}{2\sigma_r^2}}$$

C is the normalization constant, y-x is the spatial distance between the center pixel and other pixel, and I(y) and I(x) is the intensity difference between those pixels. Then Morphological operations can be applied gray scale image which is the result of bilateral filter. Finally applied dilation followed by erosion to enhance the character. This method experimentation was done in more than 500 different degraded historical palm leaf and paper documents. Results was well suited for enhancing degraded documents with low contrast and noise[4].

2.3 Binarization of Degraded Handwritten Documents based on Morphological Contrast intensification –

This paper presents a simple, but efficient binarization technique that can handle different types of degradation such as faint characters, bleeding-through, large background ink stains and the noise introduced in the scanning process (block noise) which are normally present at the border of images. The first task towards the processing of the documents starts usually with binarization of the scanned image as a pre-processing step. The objective of the binarization is to clearly divide the pixels into two classes, namely, the foreground and the background. The accuracy of binarization is pivotal for the success of the subsequent steps for further processing. Handwritten documents, particularly the historical documents are difficult to binarize due to lack of standardization of the input and degradation of all sorts due to ageing and other factors. In historical handwritten documents the degradation includes faint characters, bleeding through and ink stains. Some noise may also be introduced in the scanning process in the form of black patches or back impressions [5].

2.4 Text Extraction and Document Image Segmentation Using Matched Wavelets and MRF Model.--

Matched wavelet estimated from a given image is the optimized solution for that image only and, thus, may not represent the complete class of images. Further, for our problem, nontext regions and text regions are expected to have a variety of textural properties. For example, nontext region in an image could correspond to a landscape, background, traffic scene, sketch or computer generated texture. Hence, we need a methodology for designing a set of global filters suited for a variety of scenes and text regions.

The problem of correcting the misclassification belongs to very general class of problems in vision and can be formulated in terms of energy minimization. Every pixel must be assigned a label in the set {text, picture, background}. refers to a particular labeling of the pixels and refers to the value of the label of a particular pixel. We consider the first order MRF model. This simplifies the energy function to the following form:

$$E(f) = \sum_{\{p,q\} \in N} V_{p,q}(f_p, f_q) + \sum_{p \in P} D_p(f_p)$$

Where N is the set of interacting pair of pixels. We consider N to be eight neighborhood of a pixel. The first and second terms in the above equation are referred to as Esmooth(interaction energy) and Edata (energy corresponding to the data term) in the literature [6]. Dp (data term) is the measure of how well a label fp fits the pixel given the observed data. Esmooth makes F smooth everywhere. Although smoothing is required for removing misclassification, yet oversmoothing can lead to poor results at the object boundaries. Energy functions that do not have this problem are called as discontinuity preserving. One example of discontinuity preserving energy function is Potts interaction penalty. The major difficulty with MRF energy minimization lies in the enormous computational costs. Even with the simple Potts energy model computing the global minima is NP-hard [6]. For energy minimization, we use the –expansion algorithm proposed in [6] and [7]. In this algorithm, we model the image pixels as weighted graph depending on their labels and use min graph cut algorithm to associate each pixel to a label that results in minimization of the energy.

Analyzed several types of images so as to demonstrate the performance of the algorithm. The images are taken from scanned pages and different websites.

Applied algorithm on several structured and highly unstructured images with complex backgrounds and obtained encouraging results. The results indicate that GMWs have the efficacy to discriminate between textures, and can be effectively applied for text identification. The method can be tuned for specific classes of texts depending on the dataset used for estimating GMWs.

Segmentation of document image as a three class problem, which provides a new vision for automatic document image understanding. Thus, it opens up new research directions in the document image analysis which can provide an avenue for the development of many new applications based on document image analysis.

2.5 Document Image Binarization Using Background Estimation and Stroke Edges.—

Document image binarization is frequently performed in the preprocessing stage of different document image processing related applications such as optical character recognition (OCR) and document image retrieval. It converts a gray-scale document image into binary document image and facilitates the resultant tasks such as document skew estimation and document layout analysis [8]. This paper introduces proposed method which is the combination of polynomial smoothing, document background estimation, text stroke edge detection, local threshold estimation, and post-processing. This technique makes use of a document background surface that is estimated through an iterative polynomial smoothing procedure. Hence it

introduces a polynomial smoothing. Smoothing is a process by which signals are weighted within a local neighborhood window. Polynomial smoothing has been used in many different applications for the background surface estimation. Here polynomial smoothing is used to estimate the fingerprint background where a local two-dimensional polynomial surface is fitted by using fingerprint pixels within a sliding window. And local polynomial smoothing to estimate the background surface of the pre-detected document text regions. The stroke edge is a strong text indicator which has been used for document image thresholding. Stroke edges may not be detected properly in degraded document images due to various types of document degradation. Here text stroke edges detected based on the local image variation. After text stroke edges are detected the document text can be obtained based on the observation such as the document text is surrounded by text stroke edges and also has a lower intensity level compared with the detected stroke edge pixels. The document text is extracted based on the detected text stroke edges [8]. A certain amount of error are introduced in Document image thresholding that can be corrected through a series of post-processing operations. It uses 3 operations to remove the errors. First it removes the text components of a very small size that often result from image noise such as salt and pepper noise. Next, we remove the falsely detected text components that have a relatively large size. Last, document image thresholding often introduces a certain amount of single-pixel holes, concavities, and convexities along the text stroke boundary these single pixel defects are actually artifacts, which can be removed by using certain logical operators that can be simply set according to their neighborhood patterns [8].
This paper all about the technique which first implements an iterative polynomial smoothing procedure to estimate a document background Surface efficiently. After the stroke edges are detected based on the local image variation within the compensated document image by using estimated document background surface. Finally, the local threshold is estimated based on the detected stroke edge pixels within local neighborhood window.

3. CONCLUSION

Presentation of this paper is a survey of document image binarization techniques that improves the performance of document image binarization methods. This paper studies different binarization method which helps to restore the degraded documents contents. By the help of this paper new method can be implements by combining above methods to provide better binarization techniques.

4. REFERENCES

[1] K.Ajitha, Prof. P.Kannan, “A servey on Degraded document Image binarization techniques”, PET Engineering college, vol 3,issue 11,November 2014.

[2] Saranya Kanagaraj, Manoj Ravindra Phirke, "Binarization of multibackground text images for OCR" Imaging Tech Lab, HCL Technologies, Bangalore, India Recognition Letters, vol.3,issue 10,october 2015.

[3] LU Di, HUANG Xin, LIU Changyuan, LIN Xue, ZHANG Huayu, YAN Jun, "Binarization of degraded document image based on contrast enhancement" Proceedings of the 35th Chinese Control Conference July 27-29, 2016, Chengdu, China

[4] Ranganatha D, Ganga Holi, " Hybrid Binarization Technique for Degraded Document Images" Department of ISE PES IT, Bengaluru Visvesvaraya Technological University, India.

[5] Sekhar MandaIl, Sugata Das, Amrit Agarwae, Bhabatosh Chanda , " Binarization of Degraded Handwritten Documents based on Morphological Contrast intensification " Department of Computer Science and Technology Indian Institute of Engineer Science and Technology, Shibpur, India Electronics and Communication Sciences Unit Indian Statistical Institute, Kolkata, India

[6] Y. Boykov, O. Veksler, and R. Zabih, “Fast approximate energy minimization via graph cuts,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 11, pp. 1222–1239, Nov. 2004.

[7] V. Kolmogorov and R. Zabih, “What energy functions can be minimized via graph cuts,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 2, pp. 147–159, Feb. 2004.

[8] Shijian Lu, Bolan Su, Chew Lim Tan, " Document Image Binarization Using Background Estimation and Stroke Edges "