Ancient Document Images Enhancement Using Phase Based Binarization

J.Rajalakshmi

M.E student, Department of Computer Science, Arasu Engineering College, Kumbakonam, India.

S.Venkateswaran

Assistant Professor, Department of Computer Science, Arasu Engineering College, Kumbakonam, India.

Abstract – In this paper, we present a phase-based binarization model for degraded document images, also a post processing method that can improve any binarization method and a ground truth generation tool. Usually, many binarization techniques are implemented in the literature for different types of binarization problems. It include an adaptive image contrast based document image binarization technique that is tolerant to different type of document degradation such as uneven illumination document smear involving smudging of text, seeping of ink to the back side of page, degradation of paper ink because of aging and so on similar reasons. An objective evaluation based methodology for handwritten document image binarization techniques that aims to reduce the human involvement in the ground truth construction and consecutive testing Image binarization is the method of separation of pixel values into dual collections, foreground for black and background for white. In images the grayscale and color images into black and white images. Ancient and degraded document improvement using image processing is attracting many researchers in the recent period. Binarization is very popular cleaning the document for further processing.

Index Terms – Phase-based binarization model, degraded document images, grayscale images, binarization method, ground truth construction tool.

1. INTRODUCTION

Automated processing of documents for the purpose of extracting information from degraded documents is a challenging task and open issue for research. Generally, the main form of transmission and storage of information has been done by paper documents. These documents include many common types: historical and ancient documents, engineering drawing and maps, technical manuals, music notations etc. The preservation of historical documents is a main task of the National Archives department. Also everyone has the requirement to preserve the documents of importance for the future. However, many environmental factors, improper handling and poor quality of materials used in the creation of documents causes a high degree of degradation of these documents.

In document image processing, the paper documents are initially scanned and stored in the hard disk or any other required location for further processing. The final outcome of document image processing has to be in compatible electronic format, which enables the documents easier and quicker to access. Document image processing comprises of a set of simple techniques and procedures, which are used to work on the images of documents and convert them from pixel information into a format that can be read by a computer. Many research studies have been carried out to solve the problems that arise in the binarization of document images characterized by degradations including shadows, non-uniform illumination, low contrast, large signal-dependent noise, smear, strain, faded ink, bleed-through, show-through, image contrast, and deterioration of the cellulose structure. In this paper, a robust phase-based binarization method is proposed for the binarization and enhancement of hand held device captured documents and manuscripts. The three main steps in the proposed method are: preprocessing, binarization and post-processing. The paper is organized as follows: Section-II presents the previous work in this area. In section-III the proposed document image enhancement technique is presented. Section-IV presents comparison with other enhancement techniques and finally section-V presents the conclusion.

2. PREVIOUS WORK

Some selected previous works that perform binarization of ancient documents are given below. Given the binarization a technique of some reported methods, the common procedure is to divide the document image pixels into three sets namely foreground pixels, background pixels and uncertain pixels. Then a classifier is applied to those uncertain pixels that classifies into foreground and background, based on the pre-selected foreground and background sets.

Gatos et al. [1] proposed an adaptive approach based on rough estimation of foreground regions and background surface. By combining the calculated background surface with the original image by image up-sampling and through a post-processing step a good quality of text regions have been obtained that preserves stroke connectivity. Rachid Hedjam [2] also proposed a probability model based on maximum likelihood (ML) classification. It uses a priori information and

the spatial relationship of the pixels in the image. In this method, a local estimation of text and background features are obtained using grid-based modeling and in-painting techniques. Binarization is obtained through ML classification that classifies into black and white pixels. Valizadeh [3] etc. map input images into a two-dimensional feature space in which the foreground and background regions can be distinguished. By partition this feature space into several small regions, which are classified into text and background based by applying Niblack's method [21].
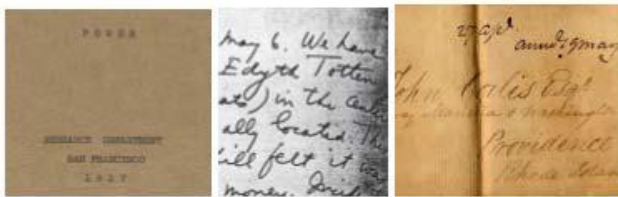

Fig 1: Three degraded document image samples from DIBCO'11

Lu [4] proposed a binarization method by estimating the background stroke width. The L1-norm gradient image is used for binarization of strokes and sub-strokes. Su [5] and Su and Lu [6] make use of local contrast by detecting the high contrast image pixels that lie around the text stroke boundary. The document text is then segmented by using local thresholds that are estimated from the detected high contrast pixels within a local neighborhood window. Farrahi Moghaddam [7] proposed a multi-scale binarization framework, in which the input document is binarized several times using different scales and then combined to form the final output image. Gatos et.al [8] combines several state-of-the-art binarization methodologies as well as the efficient incorporation of the edge information of the gray scale source image. An enhancement step based on mathematical morphology operations is also involved in order to produce a high quality result while preserving stroke information. Gatos [9] proposed a combination of global and local adaptive binarization method that is limited to binarizing handwritten document images only.

However [10] presents a Markov Random Field (MRF) model approach for the binarization of seriously degraded manuscript. Depending on the available information, the model parameters (clique potentials) are learned from training data. Expectation maximization (EM) algorithm is used to extract text and background features. However, many state-of-the-art methods are particularly suitable for the document images that suffer from certain specific type of image degradation or have certain specific type of image characteristics. To the best of our knowledge, none of the proposed methods can deal with all types of documents and degradation.

3. PROPOSED WORK

To overcome degradations in document images, the binarization along with different noise removal techniques are used in this paper. The preprocessing step mainly involves image denoising with phase preservation, followed by some morphological operations. We incorporate the canny edge detector and a de-noised image to obtain a binarized image. Then we use the phase congruency features for the main binarization step. Phase congruency is used in the machine vision literature such as palm-print verification, object detection, finger knuckle-print recognition, and biomedical applications. This process is clearly shown in Figure.1.

A. Image Denoising

An image denoising method proposed by Peter Kovesi [20] is used in this paper. This method considers that phase information is the most important feature of images. This method uses non-orthogonal, complex valued log-Gabor wavelets, which extract the local phase and amplitude information at each point in the image. The denoising process consists of determining a noise threshold at each scale and shrinking the magnitudes of the filter response vector appropriately without changing the value of phase. The main drawback of this document binarization method is that it misses weak strokes and sub-strokes. To address this problem, we combine this binarized image with an edge map obtained using the Canny edge detection method [22]. This preserves the connectivity losses between the text information. At the end of this step, the structure of foreground text is determined. The denoised image is shown in the Figure.3. However, the image is still noisy, and the strokes and sub-strokes have not been accurately binarized and also the binarization output is still affected by some type of degradations. We consider removing such effects in the following steps.
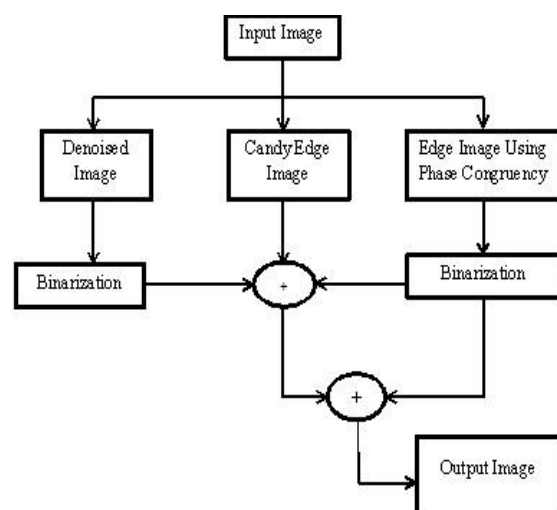

Fig.2. Block Diagram of our proposed work

B. Binarization using phase congruency

In this section, phase congruency-based feature map is extracted from input image is discussed. This feature map is based on the Peter Kovesi [19] phase congruency model. Phase congruency is a robust method to detect edges and corners. Phase congruency's robustness to image variations stems from the multi-scale and multi-orientation approach. A number of parameters impact the phase congruency output. It uses number of 2D log-Gabor filters with different scales. Then a conversation to binary form is performed. We choose a global threshold for this purpose. This global threshold has been obtained from Otsu global thresholding method.

The phase congruency map is obtained using kovesi's model. Let M_-^e and M_-^o denotes even and odd symmetric log - Gabor wavelet at a scale o, multiplying those wavelets with the image.

$$[e_\rho(x), o_\rho(x)] = [f(x) * M_\rho^e, f(x) * M_\rho^o] \quad (1)$$

Where $e_n(x)$ and $o_n(x)$ are the real and imaginary parts of a complex valued wavelet response. Then the local phase $\varphi_o(x)$ and local amplitude $A_n(x)$ is computed by using the wavelet responses.

$$\varphi_\rho(x) = arctan2(o_\rho(x), e_\rho(x)) \quad (2)$$

$$A_\rho(x) = \sqrt{e_\rho(x)^2 + o_\rho(x)^2} \quad (3)$$

The weighting mean function W(x) is defined as

$$W(x) = \frac{1}{1 + e^{\gamma(c - s(x))}} \quad (4)$$

$$s(x) = \frac{1}{N}\left(\frac{\sum_\rho A_\rho(x)}{A_{max}(x)}\right) \quad (5)$$

Where s(x) is the fractional measure of spread, N denotes the total number of filter scales; A_{max} denotes the amplitude of the filter pair with maximum response; C is the cut-off value and ⋎ is the gain factor. The phase deviation function is given by

$$\Delta\varphi_\rho = \cos(\varphi_\rho(x) - \bar{\varphi}(x)) - |\sin(\varphi_\rho(x) - \bar{\varphi}(x))| \quad (6)$$
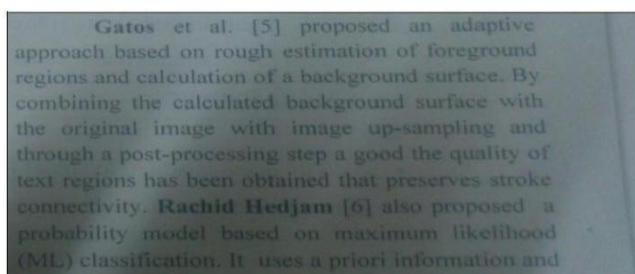

Fig.3. Hand held device captured document image
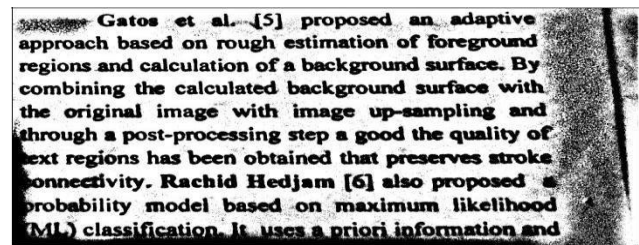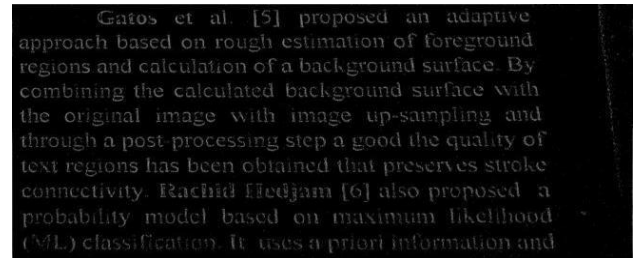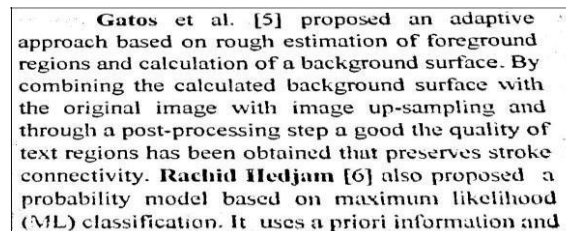

Fig.4. Denoised image

Fig.5. Phase congruency map

Fig.5. Result of binarization using phase congruency

Where $\varphi_\rho(x) - \bar{\varphi}(x)$ is the phase deviation at scale is ρ, $\bar{\varphi}(x)$ indicates the mean phase angle. Then the phase congruency is given by

$$PC(x) = \frac{\sum_\rho W(x)|A_\rho(x)\Delta\varphi_\rho(x)|}{\sum_\rho A_\rho(x)} \quad (7)$$

The phase congruency map is obtained by using the equation (7) is shown in the Fig.4. In this map the pixels of lower values are removed in order to remove the unwanted lines as well as strokes. By doing this all the unwanted lines except text information are removed, but the connectivity between the text information are lost. These discontinuities and some gaps in the text information can be filled in the post processing steps.

C. Post processing

In this section the quality of the document images are improved by using morphological operations is shown in Figure.5. It performs morphological opening on the gray scale or binary image with the structuring element. The argument structuring element must be a single structuring

element object, as opposed to an array of objects. The radius of the structuring element is chosen with respect to the document image. If more number of gaps in the resultant image, radius will be 3 or 4, for moderate quality image, radius should be less than 3. The morphological open operation is erosion followed by dilation that uses the same structuring element.

4. PERFORMANCE EVALUATION AND COMPARISION

The proposed method is tested with the image captured using hand held device and resized into 862x1574 pixels for simulation analysis as shown in Figure.1. It is seen that the captured image contain different types of degradations such as shadows, poor contrast and non-uniform illumination. The performance of the proposed method is evaluated and compared against the results of the different state-of-art methods using the same input image.

A. Otsu's method

The result obtained by using Otsu's method [18] is shown in the Figure.6. Because of selecting threshold for the whole image the most of the degraded parts are set as zero, the text information in those places are lost, also this method is not suitable for some types of degradation.

B. Niblack's method

This method assumes that if the pixel value is more than that of the threshold defined below, the pixels are considered as object and the rest of the pixels are background [21].

$$T=\mu+K\sigma \qquad (8)$$

Where μ is mean, σ is standard deviation and K is constant assumed as -0.1.

C. Sauvola method

The threshold value used in this method [17] is defined by

$$T=\mu(1+K\sigma/(R+1)) \qquad (9)$$

Where μ is mean, σ is standard deviation and K is constant assumed as -0.005

D. Bradley's method

This method is same as that of the Sauvola's method, it computes only the mean. The variance is calculated by using the formulae.

$$Variance(X) = E(X^2) - (E(X))^2 \qquad (10)$$

It saves almost 50% of time compared to the other methods listed above and provides better results compared to others as shown in Figure.7. But none of these methods are able to remove all types of degradations; our proposed method provides better results even when the document image is

affected by degradations such as non uniform illumination, shadows, poor contrast, bleed through and show through. A subjective evaluation of the resulting images are shown in Table-I.
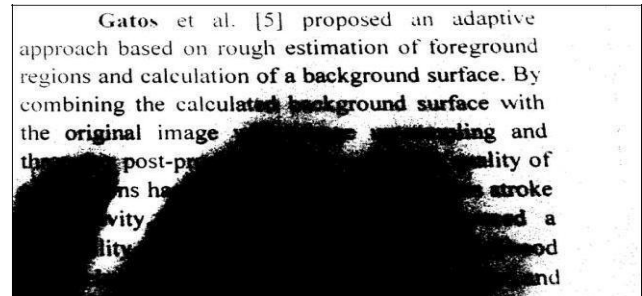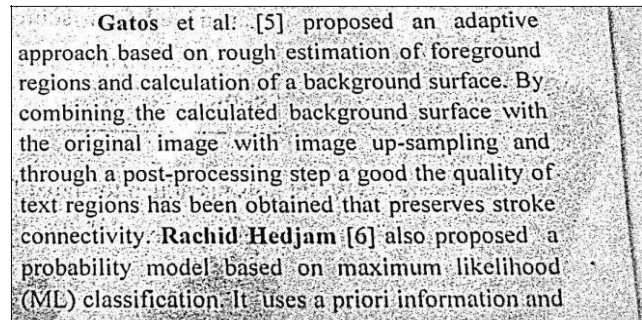
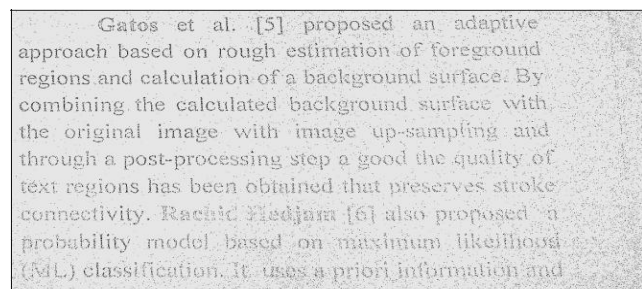
Fig.6. Otsu's method

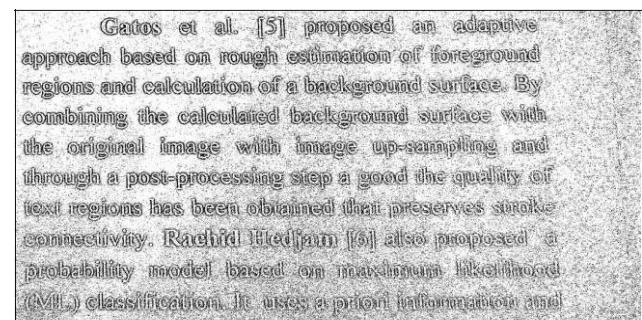Fig.7. Bradley's method

Fig.8. Niblack's method

Fig.9. Sauvola method

Gatos et al. [5] proposed an adaptive approach based on rough estimation of foreground regions and calculation of a background surface. By combining the calculated background surface with the original image with image up-sampling and through a post-processing step a good the quality of text regions has been obtained that preserves stroke connectivity. **Rachid Hedjam [6]** also proposed a probability model based on maximum likelihood (ML) classification. It uses a priori information and

Fig.10. Proposed method

TABLE I. COMPARISON OF THE PERFORMANCE OF OTHER METHOD USING PROPOSED RESULT

	Binarization Methods	Subjective
1	Otsu's method	Poor
2	Bradley's method	Better
3	Niblack's method	Poor
4	Sauvola's method	Poor
5	Proposed method	Good

5. CONCLUSION

In this, we have introduced an image binarization method that uses the phase information of the input image, and robust phase-based features extracted from that image are used to build a model for the binarization of ancient manuscripts. Phase-preserving denoising followed by morphological operations are used to preprocess the input image. Then, two phase congruency features, the maximum moment of phase congruency covariance and the locally weighted mean phase angle, are used to perform the main binarization[1][7]. For post-processing, we have proposed a few steps to filter various types of degradation; in particular, a median filter has been used to reject noise, unwanted lines, and interfering patterns.

In future work, we plan to expand the application of phasederived features, which ensures the stable behavior of document images, to other cultural heritage fields, such as microfilm analysis and multispectral imaging.

REFERENCES

[1] B. Gatos, I. Pratikakis, and S. Perantonis, "Adaptive degraded document image binarization," Pattern Recognit., vol. 39, no. 3, pp. 317–327, 2006.

[2] R. Hedjam, R. F. Moghaddam, and M. Cheriet, "A spatially adap-tive statistical method for the binarization of historical manuscripts and degraded document images," Pattern Recognit., vol. 44, no. 9, pp. 2184–2196, 2011.

[3] M. Valizadeh and E. Kabir, "Binarization of degraded document image based on feature space partitioning and classification," Int. J. Document Anal. Recognit., vol. 15, no. 1, pp. 57–69, 2010.

[4] S. Lu, B. Su, and C. Tan, "Document image binarization using background estimation and stroke edges," Int. J. Document Anal. Recognit., vol. 13, no. 4, pp. 303–314, 2010.

[5] B. Su, S. Lu, and C. Tan, "Binarization of historical document images using the local maximum and minimum," in Proc. 9th IAPR Int. Workshop DAS, 2010, pp. 159–166.

[6] B. Su, S. Lu, and C. L. Tan, "Robust document image binarization technique for degraded document images," IEEE Trans. Image Process., vol. 22, no. 4, pp. 1408–1417, Apr. 2013.

[7] R. F. Moghaddam and M. Cheriet, "A multi-scale framework for adaptive binarization of degraded document images," Pattern Recognit., vol. 43, no. 6, pp. 2186–2198, 2010.

[8] B. Gatos, I. Pratikakis, and S. Perantonis, "Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information," in Proc. 19th ICPR, Dec. 2008, pp. 1–4.

[9] K. Ntirogiannis, B. Gatos, and I. Pratikakis, "A combined approach for the binarization of handwritten document images,Pattern Recognit. Lett., vol. 35, pp. 3–15, Jan. 2014

[10] N. Howe, "Document image binarization using Markov field model," in Proc. ICDAR, 2011, pp. 6–10.

[11] H.Z. Nafchi, R.F. Moghaddam, and M. Cheriet, "Phase-Based Binarization of Ancient Document Images: Model and Application," IEEE Trans. Image Process. vol. 23, no. 7, pp.2916 –2930, july 2014. [12]

[12] R. F. Moghaddam and M. Cheriet, "AdOtsu: An adaptive and parameter-less generalization of Otsu's method for document image binarization," Pattern Recognit., vol. 45, no. 6, pp. 2419–2431, 2012.

[13] H. Z. Nafchi, R. F. Moghaddam, and M. Cheriet, "Historical document binarization based on phase information of images," in Proc. ACCV, 2012, pp. 1–12.

[14] H. Z. Nafchi and H. R. Kanan, "A phase congruency based document binarization," in Proc. IAPR Int. Conf. Image Signal Process., 2012, pp. 113–121.

[15] B. Su, S. Lu, and C. L. Tan, "Combination of document image binarization techniques," in Proc. ICDAR, Sep. 2011, pp. 22–26

[16] X Peng, S Setlur, V Govindaraju and R Sitaram, "Markov Random Field Based Binarization for Hand-held Devices Captured Document Images," in Proc. ICVGIP'10, pp. 71–76.

[17] J. Sauvola and M. Pietikinen, "Adaptive document image binarization," Pattern Recognit., vol. 33, no. 2, pp. 225–236, 2000

[18] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Syst., Man, Cybern.,Syst., vol. 9, no. 1, pp. 62–66, Jan. 1979.

[19] P. Kovesi, "Image features from phase congruency," Videre, J. Comput. Vis. Res., vol. 1, no. 3, pp. 1–26, 1999.

[20] P. Kovesi, "Phase preserving denoising of images," in Proc. Int. Conf. Digital Image Comput., Techn. Appl., 1999.

[21] W. Niblack, An Introduction to Digital Image Processing. Englewood Cliffs, NJ, USA: Prentice-Hall, 1986.

[22] J. Canny, "A computational approach to edge detection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 8, no. 6, pp. 679–698, Nov. 198.