International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 2 pp. 24-32 June 2018-Dec 2018 www.csjournals.com

Adaptive Thresholding to Robust Image Binarization for Degraded Document Images

Prashant Devidas Ingle, Ms. Parminder Kaur
PG scholar, Dept of CSE JNEC, Aurangabad, MH, India.
Assistant Professor, Dept of CSE, JNEC, Aurangabad, MH, India.

Abstract: Segmentation of text from badly degraded document images is a tough task owing to the high inter/intra-variation between the document background and additionally the foreground text of varied document pictures. A document image binarization technique is proposed that addresses these issues by using adaptive image distinction. The adaptive image distinction is a mixture of the native image distinction and additionally the native image gradient that is tolerant to text and background variation caused by differing kinds of document degradations. In the proposed method, associate adaptive distinction map is first created for degraded document image. The distinction map is then binarized and combined with canny's edge map to identify the text stroke edge pixels. The document text is additional segmented by a local threshold that's estimated supported the intensities of detected text stroke edge pixels inside a local window . The proposed method is straightforward, robust, and involves minimum parameter standardization. The experimentation is performed on DIBCO 2011 dataset and additionally the proposed method achieved higher accuracy than the existing methods.
Keywords:- document image process, degraded document image binarization, document analysis, document segmentation, image thresholding.

1. INTRODUCTION

Binarization is that the strategy of changing an element image to a binary image. Few years earlier binarization was vital for sending faxes. Nowadays it's necessary for things like digitizing text.. Image binarization is the method of separation of element values into twin collections, black as foreground and white as background. Thresholding is found to be popular technique used for binarization of document images. Document image binarization is often performed among the preprocessing stage of varied document image processing related applications like optical character recognition (OCR) and document image retrieval. It converts a gray-scale document image into a binary document image and consequently facilitates the subsequent tasks like document skew estimation and document layout analysis. As plenty of text documents are scanned, fast and proper document image binarization is becoming very important [1].

Document image binarization [1-15] tries to extract exclusively the text stroke pixels from the gray-scale document pictures, and is usually performed as part of document preprocessing stage. Even though document image binarization has been studied for many years, the thresholding of degraded document images continues to be unresolved. This might be explained by the actual fact that the modeling of the document foreground/background is extremely troublesome because of varied types of document degradation like uneven illumination, image contrast variation, bleeding-through, and smear.

In this paper, the input images are passed through the contrast enhancement technique to smooth the degraded input image. Then, contrast enhanced image is given for edge detection to search out the helpful pixels. based on the edge pixels, the contrast enhanced image is segmented using thresholding operation. The thresholded image is post processed to get the binary image. The paper is organized as follows: Section 2 presents the literature review. Section 3 presents the proposed method for document binarization. Section 4 discusses the results and conclusion is given in section 5.

2. LITERATURE REVIEW

An image binarization technique is proposed in [1] for a degraded document image that takes into accounts the adaptive image distinction. Associate adaptive distinction map is first created for associate input-degraded document image. The distinction map is then binarized and combined with Canny's edge map to identify the text stroke edge pixels. The document text is any segmental by an area threshold that's computable supported the intensities of detected text stroke edge pixels among an area window. it's been tested on 3 public datasets achieving accuracies of around ninety. Another associate adaptive water flow model for the binarization of degraded document pictures is given in [2]. During this approach, the image surface is considered a three-dimensional terrain and water is poured

on it. The water finds the valleys and fills them. The algorithmic rule controls the rainfall method, pouring the water, in such a way that the water fills up to 1/2 the valley depth. once stopping the rainfall, each wet region represents one character or a noisy component. To segment each character, the wet regions area unit labeled and thought to be blobs. variety of the blobs represent noisy elements. A multilayer perceptron is trained to label each blob as either text or non-text. The algorithm is shown to preserve stroke property. Experimental verification shows superior performance against six well-known algorithms on three sets of degraded document pictures with uneven illumination.

In [3] a multi-scale binarization, framework is introduced, which can be utilized in conjunction with any accommodative threshold-based binarization technique. This framework increases the binarization results and revives weak connections and strokes, particularly among the case of degraded historical documents. Another document image binarization technique is given in [4]. Here, native accommodative binarization is employed as a guide to accommodative stroke dimension detection. The skeleton and the contour points of the binarization output area unit is combined to spot the stroke dimension. Additionally, an accommodative native parameter is outlined that enhances the characters and improves the performance achieving further correct binarization results for every written and written documents with specific concentration on degraded historical documents. The document image binarization using native maximum and minimum technique [5] makes use of the image contrast that is defined by the local image maximum and minimum. Compared with the image gradient, the image contrast evaluated by the local maximum and minimum encompasses a nice property that it's extra tolerant to the uneven illumination and different kinds of document degradation like smear. The technique has been tested over the dataset that is utilized within the Document Image Binarization Contest (DIBCO) 2009. Experiments show its superior performance.

3. PROPOSED METHOD: accommodative thresholding to sturdy image binarization for degraded document pictures

Given a degraded document image, associate accommodative distinction map is first created and conjointly the text stroke edges are then detected through the mixture of the binarized accommodative distinction map and conjointly the clever edge map. The text is then segmental supported the native threshold that is calculated from the detected text stroke edge pixels. Some post-processing is applied to improve the document binarization quality. Figure1 shows the diagram of the proposed method.

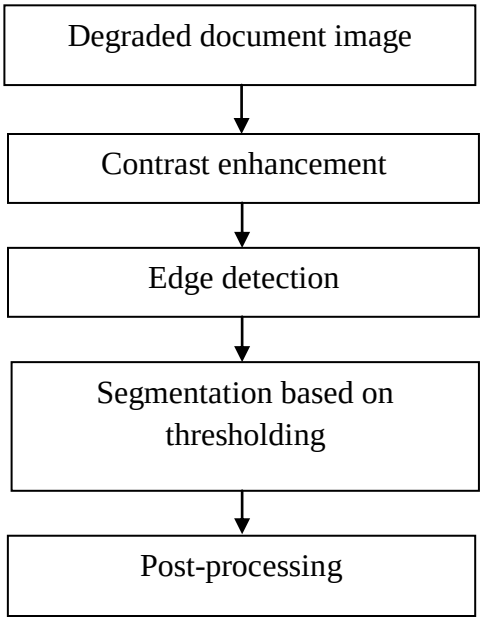


Figure 1. Block diagram of the proposed method

International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 2 pp. 18-23 June 2018-Dec 2018 www.csjournals.com

A. CONTRAST ENHANCEMENT

The image gradient has been wide used for edge detection and it may be used to detect the text stroke edges of the document images effectively that have a regular document background. On the opposite hand, it typically detects several non-stroke edges from the background of degraded document that always contains sure image variations because of noise, uneven lighting, bleed-through, etc. To extract solely the stroke edges properly, the image gradient must be normalized to compensate the image variation inside the document background. In [5], the local contrast evaluated by the local image maximum and minimum is used to suppress the background variation. This technique has one typical limitation that it's going to not handle document images with the intense text properly. This is often as a result of a weak contrast are calculated for stroke edges of the intense text. To beat this over-normalization drawback, we tend to mix the local image contrast with the local image gradient and derive an adaptive native image contrast as follows:

$$C_a(i,j) = \alpha C(i,j) + (1-\alpha)(I_{\max}(i,j) - I_{\min}(i,j)) \qquad (1)$$

where $C(i,j)$ denotes the local contrast and $(I_{\max}(i,j) - I_{\min}(i,j))$refers to the local image gradient that is normalized to [0, 1]. The local windows size is set to 3 empirically. α is the weight between local contrast and local gradient that is controlled based on the document image statistical information. Ideally, the image distinction are going to be assigned with a high weight (i.e. massive α) once the document image has vital intensity variation. so the planned binarization technique depends a lot on the native image distinction which will capture the intensity variation well and then produce proper results. Otherwise, the local image gradient are assigned with a high weight. The projected binarization technique depends additional on image gradient and avoid the over normalization drawback. The mapping from document image intensity variation to α by a power function is modeled as follows:

$$\alpha = \left(\frac{std}{128}\right)^{\gamma} \qquad (2)$$

where std denotes the document image intensity standard deviation, and γ is a pre-defined parameter. The power function has a nice property, it monotonically and smoothly increases from 0 to 1 and its shape can be easily controlled by different γ .γ can be selected from [0, ∞], where the power function becomes a linear function when γ = 1. Therefore, the local image gradient will play the major role in equation 1 when γ is large and the local image contrast will play the major role when γ is small.

Histogram based contrast enhancement: Histogram equalization could be a method in image processing of contrast adjustment using the image's histogram. This method usually increases the global contrast of the many images, especially once the usable data of the image is described by close contrast values. Through this adjustment, the intensities may be better distributed on the histogram. this allows for areas of lower local contrast to gain a higher contrast. histogram equalization accomplishes this by effectively spreading out the foremost frequent intensity values. the method is useful in images with backgrounds and foregrounds that are both bright or both dark. in particular, the method will lead to better views of bone structure in x-ray images, and to better detail in pictures that are over or under-exposed. A key advantage of the method is that it's a reasonably straightforward technique and an invertible operator. therefore in theory, if the histogram equalization function is known, then the original histogram can be recovered. The calculation is not computationally intensive.

Filter based contrast enhancement: Contrast enhancement filter enhances the edge contrast of an image or video in an effort to improve its acutance (apparent sharpness). The filter works by identifying sharp edge boundaries within the image, like the edge between a subject matter and a background of a contrasting color, and increasing the image contrast within the area immediately round the edge. This has the effect of creating subtle bright and dark highlights on either side of any edges within the image, referred to as overshoot and undershoots, leading the edge to seem additional outlined once viewed from a typical viewing distance.

B. EDGE DETECTION

Edge detection is a well-developed field on its own in image processing. Region boundaries and edge area units are closely connected, since there is usually a point adjustment in intensity at the region boundaries. Edge detection

techniques have thus been used as another segmentation technique. Segmentation methods can be applied to edges obtained from edge detectors. The aim of the distinction image construction is to find the stroke edge pixels of the document text properly. The created distinction image contains a transparent bi-modal pattern, where the adaptive image distinction computed at text stroke edges is clearly larger than that computed among the document background. Therefore the text stroke edge element candidate is detected by exploitation Otsu's global thresholding method**.** because the local image contrast and also the local image gradient are evaluated by the difference between the maximum and minimum intensity in a very local window, the pixels at both sides of the text stroke will be selected because the high contrast pixels.

Canny's edge detector: The binary map can be additional improved through the combination with the edges by Canny's edge detector, as a result of Canny's edge detector features a smart localization property that it will mark the perimeters near real edge locations within the sleuthing image. additionally, canny edge detector uses two adaptive thresholds and is a lot of tolerant to completely different imaging artifacts like shading. It ought to be noted that Canny's edge detector by itself typically extracts an oversized quantity of non-stroke edges as while not calibration the parameter manually. within the combined map, the sole picture elements that seem among each the high distinction image pixel map and clever edge map area unit unbroken. the mixture helps to extract the text stroke edge pixels accurately.

Sobel edge detector: The Sobel operator, generally referred to as the Sobel–Feldman operator or Sobel filter, is employed in image process and pc vision, notably among edge detection algorithms where it creates an image emphasizing edges. Technically, it's a separate differentiation operator, computing an approximation of the gradient of the image intensity function. At each point in the image, the results of the Sobel–Feldman operator is either the corresponding gradient vector or the norm of this vector. The Sobel–Feldman operator is predicated on convolving the image with a tiny low, separable, and integer-valued filter within the horizontal and vertical directions and is thus comparatively cheap in terms of computations. On the opposite hand, the gradient approximation that it produces is comparatively crude, above all for high-frequency variations within the image.

C. THRESHOLDING

The text can be extracted from the document background pixels once the high distinction stroke edge pixels are detected properly. The characteristics are determined from all totally different styles of document Image. First, the text pixels are close to the detected text stroke edge pixels. Second, there is a distinct intensity difference between the high contrast stroke edge pixels and the surrounding background pixels. The document image text can thus be extracted based on the detected text stroke edge pixels as follows:

$$R(x,y)=\begin{cases}1 & I(x,y)\le E_{mean}+\dfrac{E_{std}}{2}\\ 0 & otherwise\end{cases} \qquad (3)$$

where E_{mean} and E_{std} are the mean and standard deviation of the intensity of the detected text stroke edge pixels within a neighborhood window *W*, respectively.

The neighborhood window ought to be a minimum of larger than the stroke breadth so as to contain stroke edge pixels, therefore the size of the neighborhood window W are often set supporting the stroke breadth of the document image. *EW*, which can be estimated from the detected stroke edges. Since a precise stroke width is not required, just calculate the most frequently distance between two adjacent edge pixels (which denotes two sides edge of a stroke) in horizontal direction is calculated and use it to calculate stroke breadth. First the edge image is scanned horizontally row by row and the edge pixel candidates are selected. If the edge pixels, which are labeled 0 (background) and the pixels next to them are labeled to 1 (edge) in the edge map, are correctly detected, they should have higher intensities than the following few pixels (which should be the text stroke pixels). So those improperly detected edge pixels are removed. In the remaining edge pixels in the same row, the two adjacent edge pixels are likely the two sides of a stroke, so these two adjacent edge pixels are matched to pairs and the distance between them are calculated. After that a histogram is constructed that records the frequency of the distance between two adjacent candidate pixels. The stroke edge width *EW* can then be approximately estimated by using the most frequently occurring distances of the adjacent edge pixels.

International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 2 pp. 18-23 June 2018-Dec 2018 www.csjournals.com

Local thresholding: In native or adaptative thresholding, totally different threshold values for various native areas are used. global thresholding: In global thresholding, one threshold for all the image pixels is used. once the pixel values of the parts which of background ar fairly consistent in their individual values over the whole image, international thresholding may be used. OTSU rule may be an international thresholding algorithm. It stores the intensities of the pixels in an array. The threshold is calculated by using total mean and variance. Based on this threshold value each pixel is set to either 0 or 1. i.e. background or foreground. Thus here the change of image takes place only once.

D. POST-PROCESSING

Once the initial binarization result is derived, the binarization results are improved by incorporating bound domain data. First, the isolated foreground pixels that do not connect with other foreground pixels are filtered out to make the edge pixel set precisely. Second, the neighborhood elements are combined that lies on isobilateral sides of a text stroke edge element belonging to completely totally different categories (i.e., either the document background or the foreground text). One pixel of the pixel pair is therefore labeled to the other category if both of the two pixels belong to the same class. Finally, some single-pixel artifacts along the text stroke boundaries are filtered out by using several logical operators.

4. EXPERIMENTAL RESULTS

A. OTSU'S Method (1979)

In Otsu's method clustering-based image thresholding is used for the reduction of a gray level image to a binary image. This algorithmic program tries to reduce combined spread (intra class variance) by assuming that the image contains two classes of pixels. It assumes that an image follows a bimodal histogram i.e. it contains foreground and background pixels .It then calculates the optimum threshold separating the two classes to confirm that thus its combined spread is minimal. This methodology gives acceptable results once the pixels in every class are close to each other. The limitations of this method are that many degraded documents don't have clear bimodal patterns. Also another limitation is that minimization of intra class variances maximizes between class scatter.

B. Nib lack's Method [4]

By sliding a rectangular window over the grey level image Nib lack's algorithm calculates a pixel wise threshold. The threshold T is computed by using the mean m and standard deviation s, of all the pixels in the window, and is denoted as:T = m + k x s. Where k is a constant, which determines how much of the total print object edge is retained, and has a value between 0 and 1. The value of k and the size SW of the sliding defines the quality of binarization [9].The limitation of Niblack's method is that the resulting binary image suffers from a great amount of background noise especially in areas without text.

C. Sauvola's Method (2000) [5]

In this method the page is considered as a collection of subcomponents such as text, background and image. To define a threshold for each pixel of the background and imagea soft decision method is used. To define a threshold for each pixel of textual and line drawing areas a text binarization method is used. Finally the results of these algorithms are combined. [5].Although this method solves the problem posed by Niblack's approach but in many cases the characters become extremely thinned and broken. [10]

D. Bernsen's Method (1986)

In this method the local contrast is defined as follows C (i, j) =Imax (i, j)-Imin (i, j). Where C (i, j) denotes the contrast of an image pixel (i, j). Imax(i,j) ,and Imin(i,j) denote the maximum and minimum intensities within a local neighborhood windows of (i,j)respectively. If the local contrast C(i,j) is greater than a threshold, the pixel will be classified into text or background by comparing with the mean of Imax(i,j) ,and Imin(i,j).If the local contrast C(i,j) is smaller than the threshold then the pixel is set as background. This method is simple. But the limitation of this method is that it does not work properly on degraded document images with a complex background. [2]

International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 2 pp. 18-23 June 2018-Dec 2018 www.csjournals.com

E. Background Estimation Method

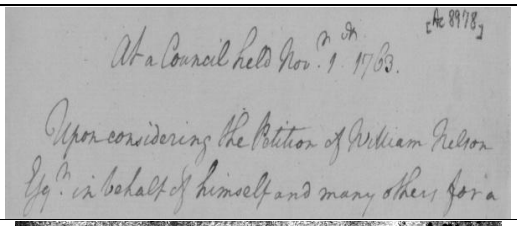
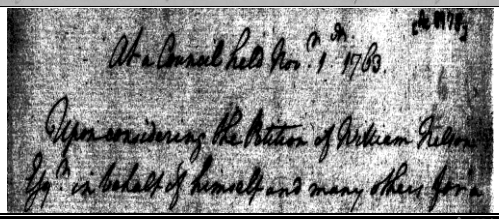
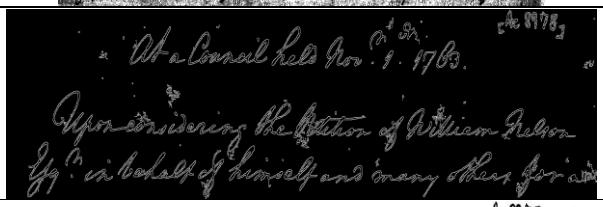
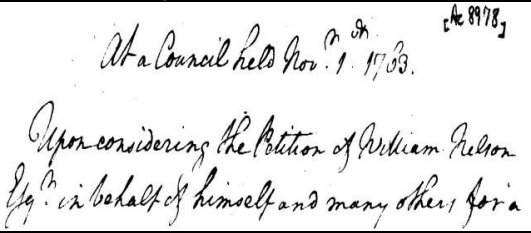
This method is an improvement over BERN'S Method and handles the documents with a complex background well. In this method the local image contrast introduces a normalization factor. This normalization factor compensates for the image variation within the document background. Here the local image contrast is evaluated as follows

$$C(i,j)=\frac{I_{max}(i,j)-I_{min}(i,j)}{I_{max}(i,j)+I_{min}(i,j)+€}$$

Where € is a positive but infinitely small number that is added in case the local maximum is equal to 0.[2] In particular, the numerator (i.e. The difference between the local maximum and the local minimum) captures the local image difference that is similar to the traditional image gradient. The denominator acts as a normalization factor that lowers the effect of the image contrast and brightness variation. For image pixels within bright regions around the text stroke boundary, the denominator is large, which neutralizes the large numerator and accordingly results in a relatively low image contrast. But for image pixels within dark regions around the text stroke boundary, the denominator is small, which compensates the small numerator and accordingly results in a relatively high image contrast. [7] As a result, the contrasts of image pixels (lying around the text stroke boundary) within both bright and dark document regions converge close to each other and this facilitates the detection of high contrast image pixels lying around the text stroke boundary.

EXAMPLE 1:

Table 1. Image 1 DIBCO Dataset 2009

Image 1 DIBCO Dataset 2009	
Input Image	
Contrast Image	
Edge Detection Image	
Output Image	

International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 2 pp. 18-23 June 2018-Dec 2018 www.csjournals.com

Table 2. Accuracy values for image 1 DIBCO dataset 2009

		Contrast enhancement method	Edge detection method	Thresholding method	Post-processing channel
Existing method		Filter	Canny	Local	Red
	Accuracy	96		99	97
Proposed Methods		Histogram	Canny	Local	Red
	Accuracy	98		97	99
		Histogram	Canny	Global	Green
	Accuracy	98		99	99
		Histogram	Sobel	Local	Red
	Accuracy	96		99	97
		Histogram	Sobel	Global	Green
	Accuracy	96		97	99

Results and Discussion

A few experiments are designed to demonstrate the effectiveness and robustness of the proposed methodology. We tend to initial analyze the performance of the proposed technique. The proposed technique is then tested and compared with existing methods [1, 5]. The binarization performance is evaluated by using accuracy, F-measure and PSNR.

For the given degraded document image, the contrast is enhanced initial by histogram based} methodology and filter based methodology. Then, the text stroke edges are then detected by Canny's edge detector and Sobel edge detector. The text is then segmented based on the local threshold and global threshold that's estimated from the detected text stroke edge pixels. Some post-processing is further applied to improve the document binarization quality.

Table 3. F-measure values for image 1 DIBCO 2009

		Contrast enhancement method	Edge detection method	Thresholding method	Post-processing channel
Existing method		Filter	Canny	Local	Red
	F-measure	98		97	97
Proposed Methods		Histogram	Canny	Local	Red
	F-measure	96		99	97
		Histogram	Canny	Global	Green
	F-measure	98		99	97

International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 2 pp. 18-23 June 2018-Dec 2018 www.csjournals.com

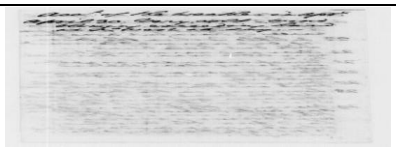
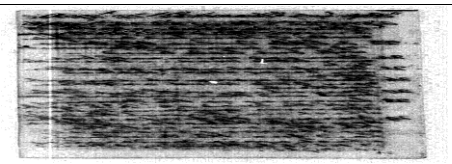
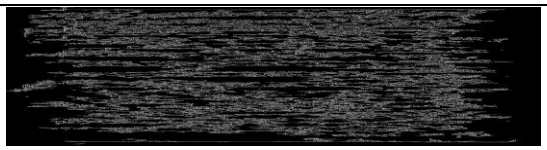
Table 4. PSNR values for image 1 DIBCO 2009

		Contrast enhancement method	Edge detection method	Thresholding method	Post-processing channel
Existing method		Filter	Canny	Local	Red
	PSNR	20		28	28
Proposed Methods		Histogram	Canny	Local	Red
	PSNR	20		30	30
		Histogram	Canny	Global	Green
	PSNR	18		30	28

5. Conclusion

This paper presents an adaptive image contrast based document image binarization technique that's tolerant to different types of document degradation such as uneven illumination and document smear. The proposed technique is simple and robust, only few parameters are concerned. Moreover, it works for different types of degraded document images. The proposed technique makes use of the local image contrast that's evaluated based on the local maximum and minimum. The proposed algorithms were tested with images including differing kinds of document components and degradations. The results were compared with two best-known techniques within the literature. The benchmarking results show that the method adapts and performs well in each case qualitatively and quantitatively. In future, the proposed may be extended to search out the optimal threshold for segmentation.

EXAMPLE 2:

Image 2 DBCO Dataset 2009	
Input Image	
Contrast Image	
Edge Detection Image	
Output Image	

References

1. 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).
2. Morteza, V., Ehsanollah, K.: An adaptive water flow model for binarization of degraded document images. Int. J. Doc. Analysis Recogn. (IJDAR) 16(2), 165–176 (2013).
3. Moghaddam, R.F., Cheriet, M.: A multi-scale framework for adaptive binarization of degraded document images. Pattern Recogn. 43(6), 2186–2198 (2010).
4. Ntirogiannis, K., Gatos, B., Pratikakis, I.: A modified adaptive logical level binarization technique for historical document images. In: Proceedings of 10th International Conference on Document Analysis and Recognition, pp. 1171–1175 (2009)
5. Su B, S Lu and C L Tan, Binarization of historical document images using the local maximum and minimum, International Workshop on Document Analysis Systems, 9-11 June 2010, Boston, MA, USA.
6. Sehad, Abdenour, et al. "Ancient degraded document image binarization based on texture features." Image and Signal Processing and Analysis (ISPA), 2013 8th International Symposium on.IEEE, 2013.
7. Nafchi, HosseinZiaei, Reza FarrahiMoghaddam, and Mohamed Cheriet. "Application of Phase-Based Features and Denoising in Postprocessing and Binarization of Historical Document Images." Document Analysis and Recognition (ICDAR), 2013 12th International Conference on.IEEE, 2013.
8. Parker, Jon, OphirFrieder, and Gideon Frieder. "Automatic Enhancement and Binarization of Degraded Document Images." Document Analysis and Recognition (ICDAR), 2013 12th International Conference on.IEEE, 2013.
9. Ntirogiannis, Konstantinos, Basilios Gatos, and IoannisPratikakis. "Performance Evaluation Methodology for Historical Document Image Binarization." Image Processing, IEEE Transactions on 22.2 (2013): 595-609.
10. Rabeux, Vincent, et al. "Quality evaluation of ancient digitized documents for binarization prediction." Document Analysis and Recognition (ICDAR), 2013 12th International Conference on.IEEE, 2013.
11. Gaceb, Djamel, Frank Lebourgeois, and Jean Duong. "Adaptative Smart-Binarization Method: For Images of Business Documents." Document Analysis and Recognition (ICDAR), 2013 12th International Conference on.IEEE, 2013.
12. Wagdy, M., Ibrahima Faye, and DayangRohaya. "Fast and efficient document image clean up and binarization based on retinex theory." Signal Processing and its Applications (CSPA), 2013 IEEE 9th International Colloquium on.IEEE, 2013.
13. Papavassiliou, Vassilis, et al. "A Morphology Based Approach for Binarization of Handwritten Documents." Frontiers in Handwriting Recognition (ICFHR), 2012 International Conference on.IEEE, 2012.
14. Su, Bolan, Shijian Lu, and Chew Lim Tan. "A learning framework for degraded document image binarization using Markov random field." Pattern Recognition (ICPR), 2012 21st International Conference on.IEEE, 2012.
15. Patvardhan, C., A. K. Verma, and C. Vasantha Lakshmi. "Document image denoising and binarization using Curvelet transform for OCR applications."Engineering (NUiCONE), 2012 Nirma University International Conference on. IEEE, 2012.