International Journal of Applied Engineering Research
ISSN 0973-4562 Volume 10, Number 10 (2015) pp. 26509-26519

http://www.ripublication.com

Degraded Document Image Restoration

Navami.A.P
M Tech Student, ECE Department
Amrita Vishwa Vidyapeetham University,
apnavami91@gmail.com
Dr.S.Ravishankar
Professor, ECE Department,
Amrita Vishwa Vidyapeetham University,
s_ravishankar@blr.amrita.edu

Abstract

Recovering of text details from degraded images is a difficult task due to high inter or intra variations between the foreground text and the background. The goal of document image restoration is to recover an actual image from the degraded image. In the proposed work, document image restoration based on Background Estimation and Expectation-Maximization algorithm is implemented using Matlab 2008b. Expectation Maximization algorithm estimates the parameters of the Gaussian Mixture Model and then extract the text data from the background based on Maximum Likelihood Estimation. The method shows less sensitivity to background noise and more sensitivity to the text details. The proposed method is performed and evaluated on various types of degraded document images with considerable background noise and different illumination conditions. The results of the proposed method show better performance over similar works reported in the current literature.

Keywords: Degraded Document Image; Binarization; Expectation-Maximization (EM) Algorithm; Maximum Likelihood Estimation(MLE).

Introduction

Document Image restoration is the method of estimating the original image from corrupted or degrade image. Degradation of document image is due to improper illumination, aging, inter-intra intensity variations in text pixels and background pixels, seep through, faded ink etc. Historical documents related to art and architecture, printed documents, manuscripts are still available for reference today which are preserved by universities and libraries all around the world. Most of these documents are facing degradation due to aging and unscientific method of archiving etc. These dilapidated documents or manuscripts can be restored back close to its original form by using image processing techniques such as image enhancement and restoration.

Image enhancement technique deals with accentuation or sharpening of image features, such as contrast, edges, boundaries etc, whereas image restoration is used to make an image resembling the true image from its degraded form. Generally document restoration process consists of three steps: pre-processing, binarization and post processing. Pre-processing deals with the noise removal present in the image, binarization transforms the grayscale image into the binary image and post-processing enhances the quality of image [1-6]. Thus image restoration improves the human readability and can be used in many applications such as document image analysis and resolution improvement, automatic detection of license plate of vehicles etc.

Rest of the paper is sectioned as follows. The section II explains the current works in document restoration. Section III gives the detailed description of the proposed method, section IV and V deals with the simulation results and comparison to the previous works in document image restoration and section VI concludes the proposed work and future scope for the work to proceed.

Related Work

Document binarization is one of the most important step in restoration process which separates foreground text details and background of documents images using a threshold value. The methods employed for thresholding in binarization can be either global or local. Global thresholding uses a single threshold value for the image pixel classification. Otsu et al. proposed a global thresholding approach for partitioning the given image into foreground and background with minimum intra-class variance and maximum inter-class variance [1]. This method yields better results when the document background is homogeneous. Therefore global thresholding approach does not produce good result for the document image containing non homogeneous background or document with high range of contrast.

In local or adaptive thresholding approach, different threshold values are used for binarization. Bernsen et al proposed a binarization method using maximal and minimal values of a local window [2]. Niblack et al. and sauvola et al. used standard deviation and mean values for the threshold calculation [3-4]. These two approaches show improvement in performance for images that contain uneven gray levels due to degradation. The above algorithms are computationally not efficient.

Gatos et al. proposed an adaptive binarization technique based on background estimation [5]. Chen et al. improvised the Gatos et al. method by using double threshold binarization based on canny edge detection [6]. In order to enhance the separation of objects or foreground from the background Chen et al. used a new edge connection method by considering distance and the edge orientation.

Uneven illuminated background is one of the main problem during thresholding. Tabatabaei et al. introduced a local adaptive binarization method based on the estimation of degraded background using morphological operation [7]. Nina et al. proposed a recursive version of global thresholding for degraded document binarization with an iterative background estimation using large median filter [8]. These two approaches helps to separate the text details from the uneven background.

Lelore.T et al. and F,Bouchara et al. proposed a technique to restore the document by computing the maximum likelihood by estimating the text position using canny edge detection algorithm [9-10]. These two methods are not suitable under a severe intensity variation within the document background.

This paper overcomes the above cited limitations by including a powerful background estimation technique using morphological filtering and text detection and extraction using non parametric approach under uneven illuminated background condition. The results of the proposed approach have shown considerable improvement in the quality of the image restored as compared to the methods described in [1],[2],[3],[10] and [11].

Proposed Methodology

The proposed method is described in Fig .1. In this approach background estimation of grayscale document images are performed by morphological filtering under uneven illuminated background conditions. After removal of background the text is extracted using Maximum Likelihood (ML) approach. The parameters of the mixtures of likelihood functions are estimated using EM algorithm. The initial estimation of the parameters used in the EM algorithm is calculated by k-means algorithm. The text information is recovered by using Maximum Likelihood (ML) Estimation.

\

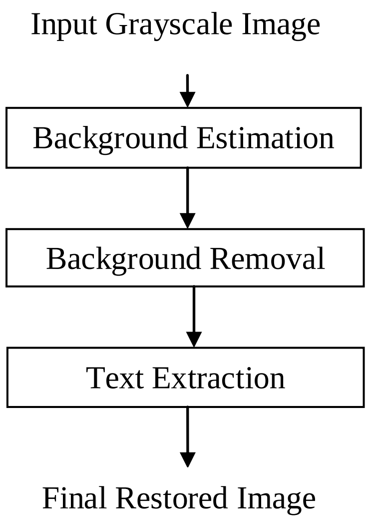


Figure 1: Block diagram of proposed method

Background Estimation

Unevenly illuminated background of document images due to poor lighting makes the image restoration very difficult. The method used for solving this problem is estimation of background surface to remove the non-uniform background. In the proposed work, the first step estimates the background of text documents using morphological filter that uses dilation and erosion.

Dilation and erosion are mathematical operations. In dilation operation the pixels are added to the boundaries of objects present in the image while in erosion operation pixel is removed along the object boundaries. The holes and gaps present in the image become smaller in dilation. Erosion shrinks an image by stripping away a layer of pixels from both the outer and inner boundaries of object regions. The number of image pixel intensity values removed or added from the objects boundaries depends on the shape and size of the structuring element used for processing the image. The rule used for the determination of dilation process states that the output pixel value is maximum of the input pixel values and for the erosion process it states that the output pixel value is minimum of the input pixels in the neighborhood region.

The input image used for dilation process is represented as F, the structuring element as S and the output image obtained as G which is determined using equation 1.

$$\mathrm{G} = \mathrm{D}(\mathrm{F},\mathrm{S}) = \mathrm{F} \oplus \mathrm{S} = \cup_{\beta \epsilon \mathrm{S}}(\mathrm{F} + \beta) \tag{1}$$

The input image used for erosion process is represented as F, the structuring element as S and the output image obtained as G which is determined using equation 2.

$$\mathrm{G} = \mathrm{E}(\mathrm{F},\mathrm{S}) = \mathrm{F} \ominus \mathrm{S} = \cup_{\beta \epsilon \mathrm{S}}(\mathrm{F} - \beta) \tag{2}$$

For the estimation of unevenly illuminated background, the shape and size of structuring element depends on the particle present in the images.

In this proposed method, the image is dilated and eroded morphologically in order to estimate the background. By doing so a rough estimation of the background without foreground text is obtained. The best results of morphological operation can be achieved by choosing the size of the structuring element as twice the size of text font. The low pass filtered image gives a smoothed version of the background. $\bar{I}(x,y)$ represents estimation of background and $I(x,y)$ represents the input image.

Background Removal.

The estimated background $\bar{I}(x,y)$ is removed by subtracting the input image $I(x,y)$ from it. The difference image is computed using equation 3 which can maintain the original polarity of darker text on lighter background [7].

$$\hat{I}(x,y) = (255 - \max([\bar{I}(x,y) - I(x,y)], 0)\ ) \tag{3}$$

Text Extraction

After non-uniform illuminated background removal, the resulting image has the problem of lower brightness than the true image. In order to avoid this problem , extract the text data from the background on the basis of MLE [10, 11].The EM algorithm is used to improve the estimation of parameters. Then the segmentation of background and text is performed by MLE method.

Initial estimation of θ1 and θ2 and the corresponding mean vector and standard deviation vector (μ0(0),σ0(0),μ1(0),σ1(0)) for EM algorithm is computed by k-means algorithm as described in [12]. The EM algorithm is an iteratively generated algorithm for the calculation of maximum-likelihood when the observation consists of incomplete data sets [13]. Each iteration in EM algorithm consists of two steps, an expectation step or E-step followed by a maximization step or M-step. The observed image takes into account a mixture of two Gaussian densities which represents the background and foreground components of document images.

EM Algorithm

EM algorithm is an iterative process [12-13]. At each iteration of the algorithm, a new label for each pixel is computed. Input k is taken as 2, representing the number of classes (foreground and background).

$\theta^{(0)}=[\pi0^{(0)}, \pi1^{(0)}, \mu0^{(0)}, \sigma0^{(0)}, \mu1^{(0)}, \sigma1^{(0)}]$, estimated using k means algorithm, where μ represents mean, σ represents standard deviation and $(\pi0^{(0)}, \pi1^{(0)})=(1/2, 1/2)$.

H represents previously defined histogram vector and 'e' is the threshold limit prescribed for the algorithm convergence

A). Expectation Step:

Computation of posterior probabilities ($\hat{Z}^{(t)}_{(i,k)}$) performed.

$$\hat{Z}^{(t)}_{(i,k)} = \frac{\hat{\pi}^{(t)}_{k} f(H_i;\hat{\theta}k^{(t)})}{\sum_{l=0}^{k} \hat{\pi}^{(t)}_{l} f(H_i;\hat{\theta}l^{(t)})} \tag{4}$$

b). Maximization Step:

Estimation of π_k , μ_k ,σ_k by maximizing $Q(\theta;\hat{\theta}^{(t)})$,

$$\hat{\pi}^{(t+1)}_{k} = \frac{\sum_{i+1}^{n} \hat{z}^{(t)}_{k}}{n}; \tag{5}$$

Computation of probability $\hat{\pi}_k$ of k[th] density is performed.

$$\hat{\mu}^{(t+1)}_{k} = \frac{\sum_{i=1}^{n} \hat{z}^{(t)}_{i,k} H_i}{\sum_{i=1}^{n} \hat{z}^{(t)}_{i,k}}; \tag{6}$$

Computation of mean $\hat{\mu}_k$ of k[th] density is performed.

$$= \frac{\sum_{i=1}^{n} \hat{z}_{i,k}^{(t)}(H_i - \hat{\mu}_k^{(t+1)})^2}{\sum_{i=1}^{n} \hat{z}_{i,k}^{(t)}}; \quad (7)$$

Estimates covariance matrix k of k^{th} density of the mixture.

$$\text{Until } \left\| Q(\theta; \hat{\theta}^{(t)}) - Q(\theta; \hat{\theta}^{(t-1)}) \right\| < e \quad (8)$$

Maximum Likelihood Estimation:

The background and foreground can be separated by ML estimation by using $\theta^{(EM)}$. The MLE approach computes the probability of pixel belonging to a corresponding class which can be text region or background and assigns it to the class when its probability is maximum. The probability distribution is expressed using Raleigh law.

$$P_{k=1,2} = \frac{1}{\hat{\mu}_k\sqrt{\frac{2}{\pi}}} \exp\left(-\frac{(\hat{I}(x,y))^2}{2(\hat{\mu}_k\sqrt{\frac{2}{\pi}})^2}\right) \quad (9)$$

Pixel (x, y) in $\hat{I}$ is labeled as 1 or 0 according to the following conditions.

$$\hat{I}(x,y) = \begin{cases} 1 \text{ for } \max(P1, P2) \\ 0 \text{ otherwise} \end{cases} \quad (10)$$

Evaluation Measures

For comparing the results of different restoration approaches some measurements of the image quality are to be computed.

F- Measure:

The harmonic mean of recall and precision is known as Fmeasure which is represented in equation 11,

$$\text{F-Measure} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}}; \quad (11)$$

Where,

$$\text{Precision} = \frac{TP}{TP+FP}; \text{ Recall} = \frac{TP}{TP+FN}; \quad (12)$$

The number of pixels which are correctly labeled as background and foreground is called True Positive (TP). The number of pixels actually belonging to the background region but detected as foreground is known as False Positive (FP). False Negative (FN) is the number of pixels labeled as background but actually belonging to the foreground region. High values of F-measure indicates that the restored image and the ground truth image are approximately equivalent.

Peak Signal To Noise Ratio (PSNR):

The closeness between two images is measured using a parameter called PSNR. As the PSNR value increases, similarity between the images also increases. The difference between the foreground intensity pixel value and the background intensity pixel value is denoted as C.

$$PSNR = 20 \log \frac{C^2}{MSE} \text{ in dB.} \tag{13}$$

where,

$$MSE = \frac{\sum_{x=1}^{m} \sum_{y=1}^{n} (I(x,y) - \hat{I}(x,y))^2}{m*n} \tag{14}$$

I(x,y) represents the restored pixel intensity value and $\hat{I}(x, y)$ is the pixel intensity value for a ground truth image at (x, y). The size of image is represented as m*n .

Structural Similarity (SSIM) Index:

Structural similarity index is a quality factor that gives similarity between the two images. It measures the image fidelity by comparing it with the initial distortion-free reference image.

Results and Analysis

This section presents discussion and comparison of the proposed method with several algorithms from the image restoration background. The proposed method is implemented in MATLAB 2008b and is compared with the previous techniques as given in [1],[2],[3],[11] and [10] in terms of SSIM, PSNR and F measure.

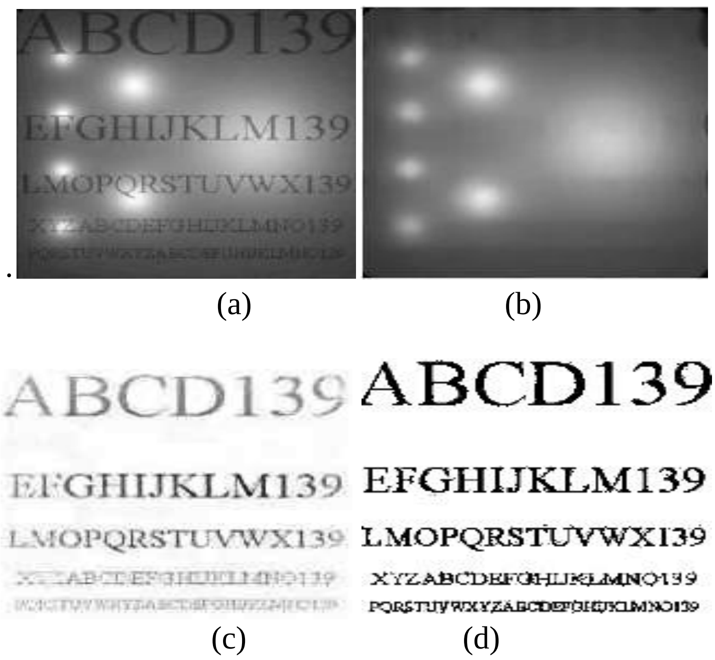


Figure 2: Proposed method : (a) Input image, (b) Background estimated image, (c)Image after background removal, (d)Restored image.

The steps described in the proposed method is shown in fig.2. In the first step, background of the grayscale document image is estimated by morphological operations such as dilation and erosion using a square structuring element with a size of 10 as shown in Fig.2(b). Then background is removed by subtracting input image from the background estimated image as shown in Fig.2(c), Fig.2(d) shows the final binarized image using probabilistic method.

The evaluation methodology was applied on faded documents, handwritten and machine-printed documents under uneven background condition. The table.1, table.2 and table.3 summarize the results obtained from some of the previous techniques in document restoration. Results in Tables shows that the proposed method provides a near optimum solution to enhance the old degraded images. Fig.3 and fig.4 shows the different binarization approaches performed on various images.

Table 1: Summarization of various methods on printed document under uneven background.

Method	PSNR	SSIM	F-Measure (%)
Otsu et al.[1]	5.7976	0.5403	83.2499
Bernsen et al.[2]	5.4956	0.4989	85.5891
Niblack et al.[3]	9.4147	0.7289	93.8085
Boussellaa et al.[11]	7.2731	0.6149	89.0794
Lelores et al.[10]	15.6732	0.9160	97.7592
Proposed method	18.9828	0.9517	98.0245

Table 2: Summarization of various methods on handwritten document under uneven background.

Method	PSNR	SSIM	F-Measure (%)
Otsu et al.[1]	10.2940	0.7920	95.1936
Bernsen et al.[2]	12.5956	0.8371	97.1936
Niblack et al.[3]	11.6384	0.7721	96.4844
Boussellaa et al.[11]	8.4458	0.7259	91.8483
Lelores et al.[10]	13.6248	0.8376	97.8544
Proposed method	19.6805	0.8875	98.3906

Table 3: Summarization of various methods on faded document.

Method	PSNR	SSIM	F-Measure (%)
Otsu et al.[1]	14.4923	0.8476	97.6599
Bernsen et al.[2]	12.2492	0.7673	97.8273
Niblack et al.[3]	14.0241	0.8207	96.6442
Boussellaa et al.[11]	15.2736	0.8230	97.2625
Lelores et al.[10]	14.0491	0.8327	96.8535
Proposed method	17.5663	0.8554	97.2409

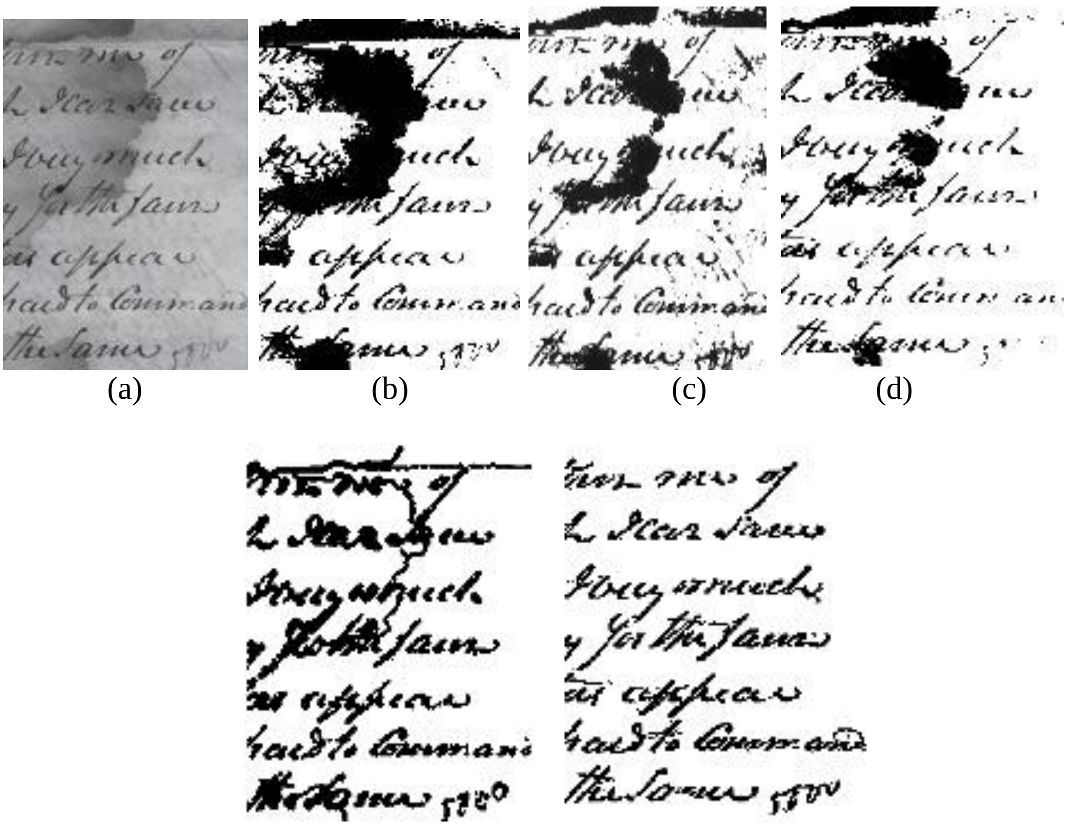
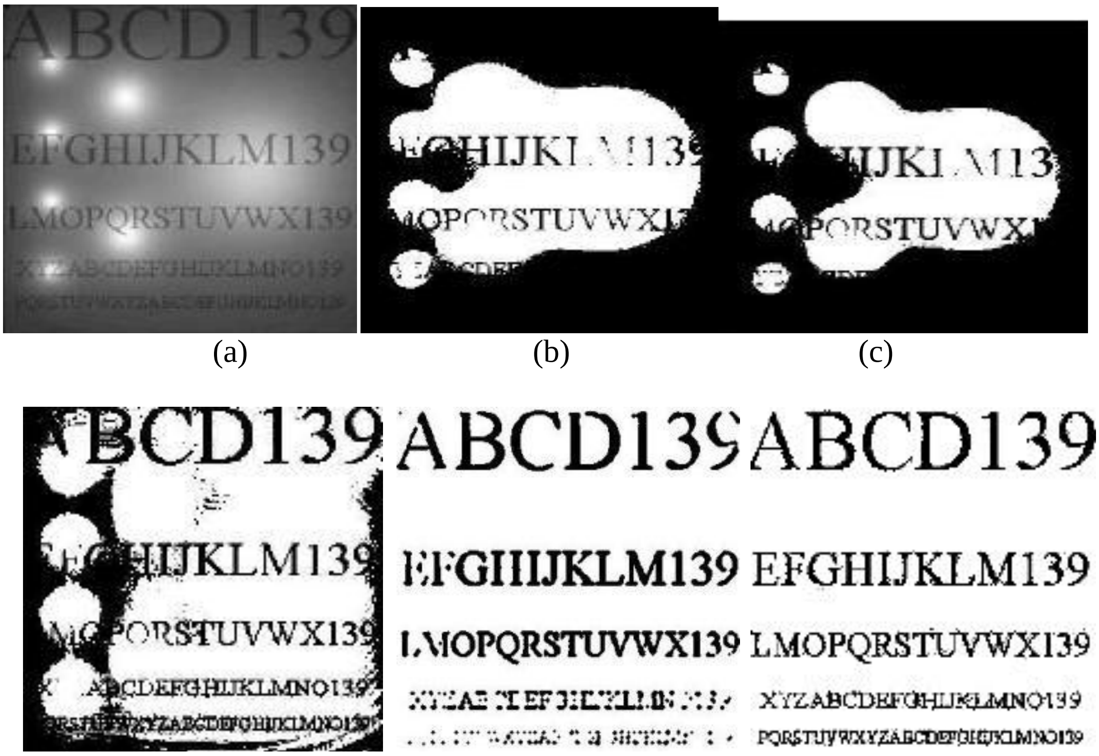

(a) (b) (c) (d) (e) (f)

Figure 3: (a)Input image, (b)Otsu et al. method[1], (c)Niblack et al. method[3], (d)Bernsen et al. method[2] ,(e)Lelores et al. method[10],(f)The proposed method



(a) (b) (c)

(d) (e) (f)

Figure 4: (a)Input image, (b)Otsu et al. method[1], (c) Bernsen et al. method[2], (d)Niblack et al. method[3], (e)Lelores et al. method[10], (f)The proposed method

Conclusions

This paper presents a novel document restoration technique based on background estimation and Maximum likelihood estimation. The simulation results performed in Matlab 2008b clearly indicates improvement in the proposed work when compared to the previous algorithms in terms of SSIM, PSNR and F-Measure. The proposed method is simple, robust and can withstand different types of degradations such as documents which are faded and unevenly illuminated.. Further work will be oriented on design and implementation of proposed method on FPGA for fast document image analysis.

References

[1] Otsu, “A threshold selection method from gray level histogram”, IEEE Trans in Sys Man, and Cybernetics, pp:62-66, vo-9, 1979.

[2] Bernsen.J, “Dynamic Thresholding of Gray-Level Images,” Proc. of 8-th Int. Conf. Pattern Recognition, pp: 1251-1255.

[3] Niblack.W “An Introduction to Digital Image Processing “, P-Hall, N-Jersey, Englewood Cliffs; pp:115–116,1986.

[4] Pietikainen , Sauvola, “Adaptive document image binarization”, Pattern Recognition;pp:225–236;2000.

[5] Gatos.B, S.J. Perantonis, Pratikakis.I, “Adaptive Degraded Document Image Binarization,” Pattern Recognition;2006;vol-39;no-3, pp:317-327.

[6] D.-s. Xia, Chen.Q, Heng.P.A, Q.-s. Sun, “A Double-Threshold Image Binarization Method Based on Edge Detector,” Pattern Recognition; pp:1254-1267;2008, vol-41, no-4.

[7] S A Tabatabaei, Bohlool, “A novel method for binarization of badly illuminated document images”, Proc. of 17-th International Conf. on Image Processing, 2010-Sept. 26-29,IEEE, Hong Kong, pp:3573 – 3576.

[8] O.Nina, Barrett.W, Morse, “ A Recursive Otsu Thresholding Method for Scanned Document Binarization” Appli. of Comp. Vision (WACV), 2011; IEEE Workshop,Kona, HI,pp:307-314

[9] F. Bouchara, Thibault .L “Super-resolved binarization of text based on the fair algorithm”, Intern’l Conf. on Document Analysis and Recognition (ICDAR),2011; pp:839-843.

[10] Thibault Lelore and Bouchara.F,. “FAIR: A Fast Algorithm for Document Image Restoration” IEEE Trans on pattern analysis and machine intelligence,aug 2013,vol. 35, no. 8, page no:2039-2048.

[11] W. Boussellaa, Bougacha.A, A. Alimi, Zahour.A, Abed.H.EL, “Enhanced Text Extraction from Arabic Degraded Document Images using EM Algorithm”, 10-th Int. Conf. on Document Analysis and Recognition,2009, Barcelona,pp- 743 – 747

[12] W. Boussellaa, A. Alimi, A. Zahour ,“A methodology for the separation of foreground/background in Arabic historical manuscripts using hybrid methods”, J. of Universal Comp. Science, 14(2) pp:284–298,2008.

[13] N. Laird, Dempster.A.P., Rubin.D.B, “Maximum Likelihood from Incomplete Data via the EM Algorithm”, J. Royal Stat. Soc., Series-B;pp:1-38; no-39;vol-1;1997.