A Multistage Binarization Technique for the Degraded Document Images

Usama W. A. Mousa, Hossam E. H. Abd El Munim, Mahmoud I. Khalil
Computer and Systems Eng. Department, Faculty of Engineering, Ain Shams University, Cairo, Egypt
{9607, hossameldin.hassan, mahmoud.khalil}@eng.asu.edu.eg

***Abstract*— This paper presents a multistage binarization technique for the degraded document images. The proposed technique can deal with many types of the degradations, where we propose edge detection methods to find the edges of the objects in the image even if this image suffers from inhomogeneities. Then, to find the rest of the objects contents, we propose a combination of Niblack's method, integral image, and a machine learning technique, where Markov random field is applied in an energy minimization framework using graph cut. Also we propose new formulas to measure the binarization efficiently and a new method for the binarization measurement. The results of extensive experiments on many datasets show the robustness of the proposed technique on various types of degradations in the document images where our technique demonstrates superior performance against also many other methods.**

***Keywords*— Binarization; Degraded document image; Edge detection; Error measurement; MRF; Restoring text.**

I. INTRODUCTION

A. Problem Definitions and Importance

The binarization is the process of converting a document image to a binary image which has only two possible values (black or white) for each pixel. This process is the most important process in many applications as well as the optical character recognition (OCR) [1]. In addition, it is very important process to save and use the historical documents which are usually degraded. This process is often required to obtain a good restoring text from a degraded document image. The binarization is an automatic process which is what the machine learning provides, so the best way for the binarization is to use the tools of the probability theory [2]. The binarization process is a challenging and a very difficult task because these document images may be suffering from aging, leaked ink, bleeding through, stains, smudge, spots, image contrast variation, non-uniform illumination, heat damage, etc. The binarized image may be suffering from both false positive (false text) and false negative (false background).

B. Who Did What and What Is Missing?

The researchers proposed many algorithms to binarize document images. The intense activity in binarization has made the Document Image Binarization Contest (DIBCO) series one of the most popular competitions in this field [3-10], where it gives the degraded document images and its ground truth images. An example is shown in Fig. 1.a. and Fig. 1.b. The ground truth image in DIBCO is used to measure the binarization but it is often (itself) suffering from a false positive where the text may be wider in the thickness than the original as shown in Fig. 2, which reduces the results of the contests.

We participated in DIBCO 2018 contest [10] and they gave us a rank number one in the evaluation results on DIBCO 2017 dataset.

Some binarization algorithms apply global thresholding where they select a single threshold for the entire image to separate between the foreground and the background. The method proposed by Otsu [11] is the most successful global thresholding method and widely used in many text segmentation applications. This method gives satisfactory results when the numbers of pixels in the foreground and the background are close to each other where the histogram becomes bimodal, otherwise it gives false positive and false negative as usual in the degraded document images as shown in Fig. 1.c where the histogram for the image in Fig. 1.a is shown in Fig. 4.a which shows clearly that it does not have a bimodal shape and it is appeared the difficulty to find a threshold to separate between the foreground and the background. Some binarization algorithms apply local thresholding where they select a special threshold for each pixel depends on some local statistics such as the mean or the variance of the pixel's neighbourhood. Sauvola's method [12] and Niblack's method [13] are proved to be of the best local threshold methods. Fig. 1.d shows the binarized image of Fig. 1.a by using Sauvola's method. These techniques were used but it usually produces a false positive and a false negative because of the degradations in the document image. To overcome the false negative and false positive, some methods come with pre-processing and post-processing steps. A method makes that [1] where it consists of four steps. In one step, they depend on Sauvola's method and accept its results as it is. This step produces false positive and false negative, which forced them to use three post-processing procedures which may produce more false positive. Some researchers make edge detection of the text and apply the energy minimization using graph cut [2, 14]. They depend on creating foreground seeds and background seeds which have different weights and probabilities for each pixel in the image. Then they create and optimize an energy function that consists of a data part and a smoothing part, then minimizing this energy by using graph cut method. A method made that [2] where they used this method but they take all the output of the energy minimization as it is which may contains false positive. Another method [14] used the graph cut in the text

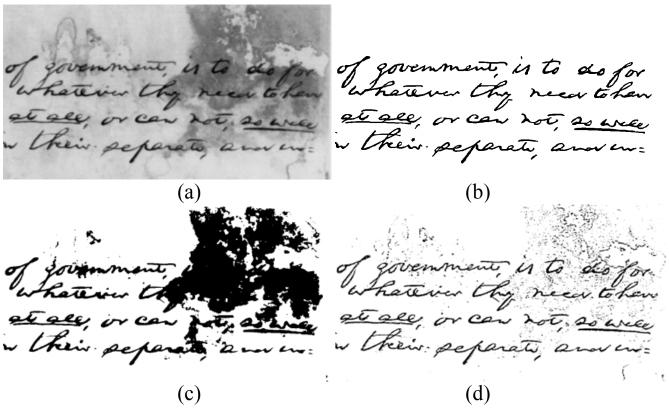


Fig. 1 (a) A degraded document image (DIBCO 2009 – H4), (b) The ground truth, (c) Otsu's method, and (d) Sauvola's method.

Fig. 2 The ground truth (lower images as examples) is wider in the thickness of the text than the original (upper images) in DIBCO contests.

Segmentation but when selecting a background seed they go faraway five pixels in the opposite direction of a foreground seed which may produce a false positive.

C. Contributions

In this paper we take all the previous into consideration and we propose a multistage binarization technique which binarizes and enhances poor quality and degraded documents images, and which can achieve excellent results on a wide variety of Challenging, this is by using the Efficient methods in this field. The rest of this paper is organized as follow: methodology, experimental results, and conclusion.

II. Methodology

A. The Technique

The overall pipeline of the proposed multistage binarization technique is presented in Fig. 3.a, where the proposed technique begins with the grayscale image instead of the color image as the source image and consists of the next stages:

1) Niblack's Method:

This stage binarizes the original image by combining Niblack's method [13] as a local threshold, and Otsu's method [11] as a global threshold, where the local threshold is computed as follows:

$$T(x,y) = m(x,y) - k * \sigma(x,y) \quad (1)$$

Where the better value of $k = 0.2$ for the window size 30*30, $m(x,y)$ is the local mean, and σ is the local standard deviation.

2) The First Proposed Edge Detection:

This stage proposes the first edge detection, where the proposed formula is:

$$I(x,y)_{new1} = \left(\frac{I(x,y)_{max} - I(x,y)}{I(x,y)_{max} + I(x,y) + \epsilon}\right)^2 \quad (2)$$

Where ϵ is a small value to prevent dividing by zero, $I(x,y)$ is the current original pixel, we obtain $I(x,y)_{max}$ as follows:

$$B = \{I(x+2,y), I(x-2,y), I(x,y+2), I(x,y-2)\} \quad (3)$$

Then we find the maximum value of this set:

$$I(x,y)_{max} = Max(B) \quad (4)$$

Then by using Otsu's method [11] we can find the edge pixels. Equation 2 eliminates a great amount of noise and gives the details of the text from the degraded document image almost without noise. Its form may be near from Van Herk and Bolan's method where the difference in the numerator is as Canny's method [15]. If $I(x,y)_{max} \gg I(x,y)$ then $I(x,y)_{new1}$ will be large, If $I(x,y)_{max}$and $I(x,y)$ are great then $I(x,y)_{new1}$ will be very small, and If $I(x,y)_{max}$and $I(x,y)$ are small then $I(x,y)_{new1}$ will be small. The power two is more useful than power one where it eliminates the noise. Equation 3 uses $x + 2$ which is more useful than $x + 1$ or $x + 5$.

3) The Second Proposed Edge Detection:

We propose the second edge detection by using the same contents and steps in the first edge detection:

$$I(x,y)_{new2} = \frac{(I(x,y)_{max} - I(x,y))^2}{I(x,y)_{max} + I(x,y) + \epsilon} \quad (5)$$

The power in the numerator is the difference between (2) and (5).

4) Selecting Seeds:

This stage applies the next algorithm: every black pixel in the first edge detection image that has gray value in the source image greater than the average value of all seeds (T1) then it becomes white. So by this procedure, we can remove noise seeds. The probabilistic model for this stage is:

$$p\left(I(x,y)_{black\ pixel\ in\ Ie1} = white \middle| I(x,y) > T1\right) = 1 \quad (6)$$

Where $T1 = \sum_N I(x,y)/(N+\epsilon)$, $I(x,y)_{black\ pixel}$ is a black pixel in the first edge detection image, and N is the double of the number of all black pixels. Equation 6 means that the probability that the black pixel becomes white pixel equals 100% if the gray value of this black pixel is greater than *T1*.

5) Machine Learning to Find the Inner Part of the Objects:

In this stage, we use the machine learning to produce the machine learning binarized image by using a tool of the probability theory: Markov random field (MRF), which also called an undirected graphical model (UGM), which is much more natural for the binarization. A graph $G = (V,N)$ is created for the gray image where V stands for the set of nodes and N represents the set of edges connecting two adjacent nodes, and we will assign each node i in V a unique label x_i where $x_i \in \{0,1\}$. We apply the energy minimization by using graph cut [2] to find the text pixels. The key that lets us solve the problem efficiently is the reducing of the energy minimization problem to the well-known s-t min-cut problem. Fig. 3.b shows a small min-cut problem, where there are two possible s-t cuts: $S = \{s,p\}$ or $\{s\}$, one of the cuts has minimal cost $w(S) = 10$, so the minimum energy is equivalent to the minimum weight, where the energy equation is:

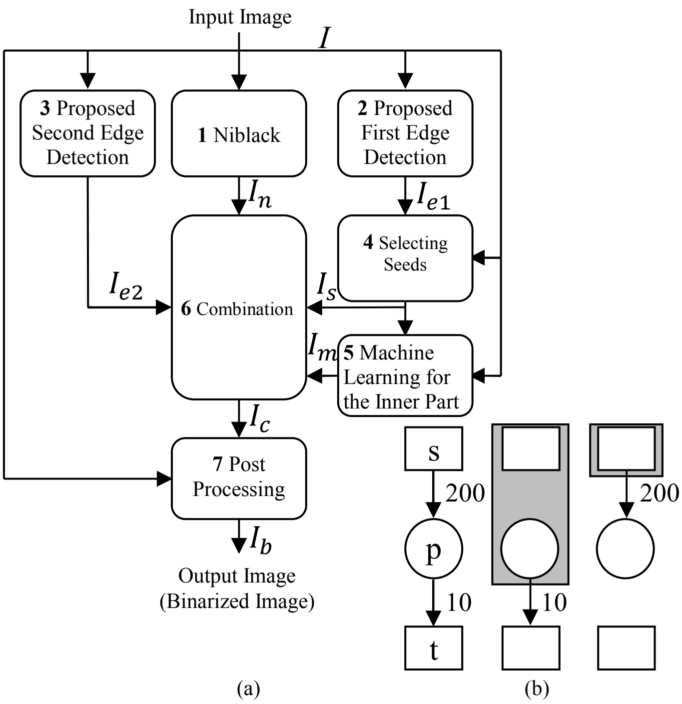


Fig. 3 (a) The block diagram of the proposed technique, (b) A simple s-t min-cut problem with two arcs, where the machine learning stage uses this key idea to reduce the energy minimization problem to s-t min-cut problem.

$$E(x) = \sum_{i\in V} E_1(x_i) + \sum_{(i,j)\in N} E_2(x_i, x_j) \qquad (7)$$

Where $X = \{x_i\}$ is a set of binary labels for each pixel in the image. This energy equation consists of two parts: the data part: $\sum_{i\in V} E_1(x_i)$ and the smoothing part: $\sum_{(i,j)\in N} E_2(x_i, x_j)$. The data part represents the relation between the seeds and the pixel, and the smoothing part represents the relation between the pixel and its neighbourhood. So this technique depends on a set of black seeds, a set of white seeds, the neighbourhood that surrounds the required pixel, and the new results. Then it finds the global minima x^* as follows:

$$x^* = \arg\min E(x) \qquad (8)$$

It can be efficiently computed by using graph cut [2], where for each unlabeled pixel x_i :

$$E_1(x_i) = \begin{cases} \frac{C(i)-K}{C(i)+K} & if\ x_i\ is\ black \\ \frac{K-C(i)}{C(i)+K} & if\ x_i\ is\ white \end{cases}, \quad E_2(x_i, x_j) = \frac{|x_i - x_j|}{ex1+\epsilon},$$

K is the average value of all gray value in the source image for the seeds, $C(i)$ is the gray value in the source image for node i where $C(i)$ in the graph is equivalent to $I(x,y)$ in the image, and $ex1 = IIF(((C(i) > C(j)\ and\ x_j = white)\ or$
$(C(i) < C(j)\ and\ x_j = black)),0,|C(i) - C(j)|)$

Where $E_1(x_i)$ is the data energy that represents the unary cost when node i takes label x_i, and $E_2(x_i, x_j)$ is the smooth energy that encodes the pair-wise cost when neighbouring nodes i and j take label x_i and x_j. For pixel i, the data energy tends to assign a pixel with label as text if it has similar intensity with text seed pixels and vice versa. Smoothness energy prefers giving two neighbouring pixels the same label for similar intensity. As example, $E_2(x_i, x_j)$ will be large if ($C(i) > C(j)\ and\ x_j = white$, $C(i) < C(j)\ and\ x_j = black$, or $C(i) = C(j)$) and x_i not equals x_j.

6) Combination to Create the Combination Image:
By using the second edge detection, Niblack, and machine learning images, this stage converts any black pixel from those images attachment to a black seed in the selecting seeds image to become a black seed and it repeats that until there is no other seed to test. We note that Niblack's image has a good property that there are white holes between the text and its noise which prevent to add noise there.

7) Post Processing:
By using the combination image and the original image, we propose the next technique, where any white pixel attachment to a black pixel with gray value less than the gray value of the black pixel in the source image will be black to produce the binarized image. The probabilistic model for this stage is:

$$p\big(x_i = black \,\big|\, C(i) < C(j), x_j = black\big) = 1 \qquad (9)$$

Note that we always use the integral image technique [16] whenever possible to make all stages and algorithms very fast. Where the integral image at location x, y equals:

$$J(x,y) = I(x,y) + J(x-1,y) + J(x,y-1) - J(x-1,y-1) \qquad (10)$$

Where I is the original image. Then any rectangular sum can be computed in four array references.

B. The Measurements

DIBCO events use many formulas for computing, as F-measure (FM) and $PSNR$ [3-10]. But if the true positive = $TP = 0$ then $Recall = \frac{TP}{TP+FN} = 0$, $Precision = \frac{TP}{TP+FP} = 0$, and then $FM = \frac{2*Recall*Precision}{Recall+Precision} = 0/0$ = Indeterminate value, where FP = *false positive* and FN = *false negative*. So we propose another formula for FM, where we define the next variables: GB = the number of the black pixels in the ground truth image, BB = the number of the black pixels in the binarized image, and $CB = TP$ = the number of the common black pixels. Then, $Recall = CB/GB$, $Precision = CB/BB$, and FM:

$$FM = \frac{2*CB}{GB+BB} \qquad (11)$$

And when $TP = 0$ that $CB = 0$ then $FM = 0$. We propose also another measurement formula that is the error measure which we will denote to it by: EM, where:

$$EM = FP + FN = GB + BB - 2 * CB = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(I(x,y) - I'(x,y))^2}{C^2} \qquad (12)$$

EM is simple and gives an accurate result. Also we can now write the $PSNR$ [3-10] as follows:

$$PSNR = 10\log\left(\frac{C^2}{MSE}\right) = 10\log\left(\frac{MN}{EM}\right) \qquad (13)$$

Where $MSE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(I(x,y) - I'(x,y))^2}{MN}$, c is the difference between foreground and background, and MN is the size of the image. I and I' are the binarized image and the ground truth image respectively. Some researchers use the Jaccard index: J, as in DIBCO 2016 (H-KWS 2016) [17] where they called it the intersection over union (IoU) and the formula is:

$$J = IoU = \frac{intersection}{union} = \frac{CB}{GB+BB-CB} \qquad (14)$$

F-measure (FM) and Jaccard index (J) are inaccurate, as shown in the next example which is illustrated in Table 1: an image has a size = 600 and GB = 200 is binarized by two methods. The first method produces CB = 100 and BB = 100.

TABLE 1

THE DATA OF THE JACCARD EXAMPLE

Method	Size	*GB*	*BB*	*CB*	*FM%*	*J%*	*EM*	*PSNR*
1	600	200	100	100	67	50	100	7.78
2			390	200	68	51	190	4.99

The second method produces *CB* = 200 and *BB* = 390. The result is shown in Table 1, where the error (*EM*) in the second method is almost double the error in the first method although *FM* and *J* said the opposite, where they cannot find the best method, but *EM* and *PSNR* can do that.

EM is always accurate, but *PSNR* depends on the image's size. Also in DIBCO, some degraded images produce a huge amount of noise as a collapse or a breakdown in the binarization, but in the total result, we note that *FM* and *PSNR* approximately will hide this collapse. So *FM* and *PSNR* are unable to express precisely the results of different methods as in the next example. Example 2: Four images, each of them has a size = 100000 are binarized by two methods. The result is shown in Table 2, where the *EM* and *PSNR* give opposite results to each other. So *PSNR* is not always good. Also when *GB = BB = CB = 0* then the *FM* and *PSNR* fail to find the result but *EM* = 0. In addition, we can take the summation of *EM* and sizes for all images to find the total of *EM* and *PSNR*. Also, as we said we saw that the ground truth image in DIBCO is often wider in the thickness of the text than the original image as shown in Fig. 2. In DIBCO 2010 [4] until now, they introduced the pseudo F-Measure (*p-FM*), so we propose a method such as that and we will calculate the recall by a new way, where we will neglect any ground truth black pixel that attaches to a white pixel from calculating of the recall, where in this case we compute pseudo F-Measure (*p-FM*), pseudo PSNR (*p- PSNR*), and pseudo EM (*p- EM*) as the next equations:

$$p\text{-}PSNR = 10 \log\left(\frac{M*N}{p-EM}\right) \quad (15)$$

$$p\text{-}EM = FP + p\text{-}FN = GB2 + BB - (CB + CB2) \quad (16)$$

$$p\text{-}FM = \frac{2*Recall2*Precision}{Recall2 + Precision} = \frac{2*CB*CB2}{CB*GB2 + CB2*BB} \quad (17)$$

Where *CB2* and *GB2* are counters as *CB* and *GB* but for the black pixels in the ground truth image those have not any attachment white pixels. *Recall2 = CB2/GB2*. In the next experimental results we use *FM = p-FM*, *EM = p-EM* and *PSNR = p-PSNR*.

DIBCO contests used other measures like *NRM* until DIBCO 2010 [4] and *MPM* until DIBCO 2011 [5]. Also they used Distance Reciprocal Distortion Metric (*DRD*) measure from DIBCO 2011 [5] until now, where *DRD* whenever it is less then it is better. But *DRD* is inaccurate, as example, in the result of DIBCO 2012 [6] the *DRD* for rank number one is greater than the *DRD* for rank number two.

III. EXPERIMENTAL RESULTS

A. DIBCO Datasets

DIBCO datasets [3-9] from 2009 to 2016 are used for the evaluation. To prove the effectiveness of our proposed method, we binarized these images and tested the results with the ground truth images. Fig.4.b shows the binarization for the image in Fig 1.a by using our proposed method where it is great although the histogram for the image in Fig. 1.a which is shown in Fig. 4.a shows clearly that it does not have a bimodal shape. Our computer has Ram = 2 gigabytes and CPU = 3 GHz. Table 3 shows the results from this testing. The average time for any image equals 18 seconds and as we see the final result of *FM* is equals 94.06 %, so our technique appears promising.

TABLE 2

THE DATA OF EXAMPLE TWO THAT COMPARES BETWEEN *EM* AND *PSNR*

Image	Size	First Method		Second Method	
		EM	*PSNR*	*EM*	*PSNR*
1	100000	1000	20	70000	1.55
2		1000	20	100	30
3		2000	16.99	200	26.99
4		2000	16.99	200	26.99
Average		1500	18.49	17625	21.38

1) *Method's Comparison*

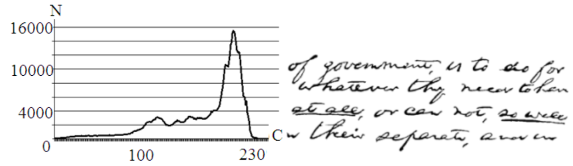
Table 4 compares the final results from our proposed method in Table 3 with Otsu's method [11] and Sauvola's method [12] in *FM*, *PSNR*, and *EM*. The difference between these methods in *FM* and *PSNR* is not great, so as we said, *FM* and *PSNR* in the total are not suitable for this comparing, where in some images Otsu and Sauvola achieve poor results as breakdown but *FM* and *PSNR* hide this breakdown, so the effect of the breakdown disappears and not clear in *FM* and *PSNR* in the total.

2) *EM Importance*

The solution of this problem is the using of *EM* as shown in Table 4 and Fig. 5.b, where they show that the error in Otsu and Sauvola are more six and four times than the error in our proposed method respectively, so *EM* is the best expression to compare between different methods of binarization. Fig. 5.a shows the *EM* of Otsu method and our proposed method for all DIBCO images in Table 3, and it is clear the instability of Otsu method in comparing to our proposed method which has small errors and very stable near the zero errors. Table 5 shows the seven DIBCO contests [3-9] and a comparison between the results of the proposed method and DIBCO's winners in *FM*, *PSNR*, and the time. We see that the proposed method is better in the time and the total of *FM* and *PSNR*. Table 6 shows the *EM* of Otsu method and our proposed method for the three images in Fig. 6, which are DIBCO: 2012-5, 2013-H5, and 2013-P4 respectively. Here we can note that the *EM* for our proposed method is almost negligible for these three images but it is very high and breakdown in Otsu method. This shows that our proposed method is an efficient technique. *EM* is an accurate binarization measurement formula.

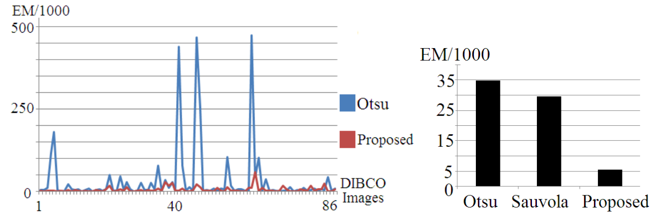
B. PHIBC 12 Dataset

We also used PHIBC 12 dataset [23] for the evaluation. Table 7 shows a comparison between the results of the proposed method and PHIBC_12'S winner in *FM* and *PSNR*. Also the total for the proposed method is greater than the total of the winner in the *FM* and *PSNR*. All these results show the efficient of our proposed technique.



(a) (b)

Fig. 4 (a) The histogram for the image in Fig 1.a, (b) The binarized image of Fig. 1.a by using the proposed method.



(a) (b)

Fig. 5 (a) The results of Otsu and the proposed methods for the 86 DIBCO images by using *EM*, (b) The results of different methods by using *EM*/1000.

TABLE 3

EVALUATION ON DIBCO 2009, H-DIBCO 2010, DIBCO 2011, H-DIBCO 2012, DIBCO 2013, H-DIBCO 2014, AND H-DIBCO 2016 FOR THE PROPOSED METHOD

Image	*FM* %	*PSNR*	*EM*/1000	Time (sec)
2009-H1	97.78	27.186	1.649	12
2009-H2	98.86	34.429	0.466	17
2009-H3	95.68	21.505	2.025	4
2009-H4	96.09	24.356	2.325	11
2009-H5	93.23	25.202	2.886	15
2009-P1	98.52	26.072	0.824	5
2009-P2	98.86	23.624	1.646	6
2009-P3	98.83	24.997	1.799	9
2009-P4	97.69	23.910	2.683	11
2009-P5	97.32	24.048	1.242	5
2010-1	94.41	21.001	4.494	8
2010-2	94.63	25.258	3.935	19
2010-3	98.31	30.235	0.337	4
2010-4	91.03	22.171	3.046	6
2010-5	99.11	31.390	0.490	9
2010-6	94.93	28.229	0.520	5
2010-7	92.41	23.097	3.987	12
2010-8	95.02	25.123	2.285	10
2010-9	94.95	29.525	0.823	9
2010-10	88.70	21.888	7.143	17
2011-H1	84.53	13.820	19.888	7
2011-H2	96.02	28.448	1.360	12
2011-H3	97.03	27.330	1.767	15
2011-H4	85.48	16.847	5.787	5
2011-H5	98.66	26.566	0.934	6
2011-H6	68.63	16.703	11.551	8
2011-H7	97.76	31.081	0.503	9
2011-H8	98.96	33.866	0.168	6
2011-P1	98.70	24.426	1.834	8
2011-P2	93.57	18.722	5.876	8
2011-P3	99.12	25.617	1.198	7
2011-P4	98.35	25.809	3.850	23
2011-P5	98.28	25.216	1.416	7
2011-P6	97.08	26.894	2.874	23
2011-P7	68.00	17.062	6.656	5
2011-P8	89.85	19.518	3.100	4
2012-1	86.55	20.306	27.557	49
2012-2	78.92	17.762	18.136	16
2012-3	90.67	20.330	21.714	41
2012-4	96.08	27.028	1.627	10
2012-5	98.24	27.615	1.821	22
2012-6	92.18	22.272	8.342	22
2012-7	90.59	25.513	1.019	5
2012-8	99.43	31.690	0.505	11
2012-9	96.37	23.191	3.511	11
2012-10	89.35	18.952	22.551	36
2012-11	92.76	20.142	14.843	29
2012-12	94.68	25.685	2.153	12
2012-13	93.56	26.487	3.239	23
2012-14	97.56	29.794	2.255	33
2013-H1	92.71	26.258	10.332	70
2013-H2	98.98	31.574	0.442	9
2013-H3	76.63	20.483	10.326	17
2013-H4	99.08	31.934	2.317	59
2013-H5	98.92	32.738	1.282	40
2013-H6	96.07	25.671	9.639	58
2013-H7	92.55	29.561	1.699	22
2013-H8	98.62	29.233	2.959	40
2013-P1	98.45	30.110	1.465	24
2013-P2	99.48	31.031	0.992	19
2013-P3	99.40	32.281	0.674	17
2013-P4	97.36	22.432	11.320	41
2013-P5	97.52	20.615	8.382	16
2013-P6	84.93	14.315	62.077	33
2013-P7	99.37	27.355	0.591	5
2013-P8	92.13	19.336	9.270	13
2014-1	97.32	29.353	1.445	19
2014-2	95.35	24.929	2.751	12
2014-3	99.58	31.599	2.323	57
2014-4	99.34	28.634	0.513	5
2014-5	98.35	24.681	1.291	5
2014-6	69.54	13.104	17.444	5
2014-7	94.51	20.946	5.081	9
2014-8	99.55	36.581	0.291	20
2014-9	99.61	32.583	0.392	10
2014-10	99.56	31.723	0.410	9
2016-1	96.25	24.042	6.352	27
2016-2	92.03	27.508	4.102	41
2016-3	98.58	30.754	2.162	44
2016-4	83.18	19.646	15.766	24
2016-5	98.66	29.557	2.523	40
2016-6	93.77	22.959	5.438	17
2016-7	85.88	19.641	6.861	10
2016-8	87.54	14.322	22.003	11
2016-9	98.91	27.214	0.768	7
2016-10	95.90	20.452	1.073	2
Average	94.06	25.152	5.528	18

IV. CONCLUSIONS

This paper has proposed a robust multistage binarization technique for the degraded document images, where we

Proposed to use a machine learning technique which depends on the probability theory where we used Markov random field (MRF) and integral image. This is by applying the energy minimization using graph cut in a proposed combination algorithm with proposed edge detection methods and Niblack's method. There are a proposed algorithm for the post processing stage, a proposed algorithm for the binarization measurement, a proposed formula for *FM*, and a proposed error measurement formula where we called it: *EM*, which gives an accurate result for the binarization measurement. All of these formulas, stages, and methods construct our works and our proposed technique.

DIBCO 2018 contest gave us a rank number one in the evaluation results on DIBCO 2017 dataset. This brilliant result motivates us to continue to develop. The proposed technique gives good results with DIBCO and PHIBC_12 datasets and we expect that this technique can achieve promising results.

TABLE 4

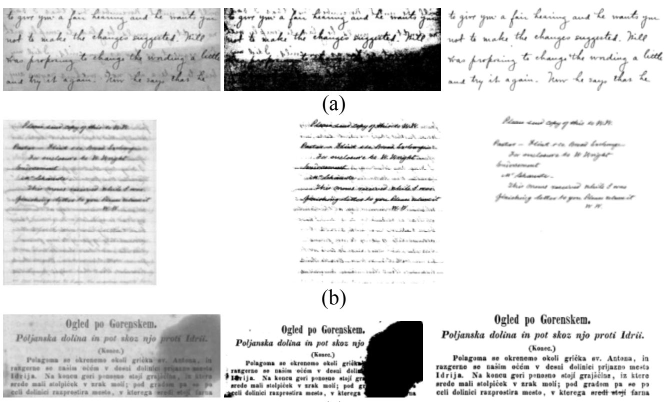
THE RESULTS OF DIFFERENT METHODS

Method	*FM* %	*PSNR*	*EM*/1000
Otsu [13]	86.72	20.361	34.926
Sauvola [14]	80.10	18.061	29.727
Proposed Method	94.06	25.152	5.528

TABLE 5

DIBCO'S WINNERS [3-9] AND THE PROPOSED METHOD

Year	DIBCO's winners			The proposed method		
	FM %	PSNR	Time	*FM* %	*PSNR*	Time (sec)
2009	91.24	18.66		97.29	25.53	10
2010	91.50	19.78		94.35	25.79	10
2011	80.86	16.13		91.88	23.62	10
2012	89.47	21.80	29.7 m	92.64	24.06	23
2013	92.12	20.68		95.14	26.56	31
2014	96.88	22.66	17.43 s	95.27	27.41	16
2016	87.61	18.11		93.07	23.61	23
Total	89.95	19.69	14.9 m	94.06	25.15	18



(a) (b) (c)

Fig. 6 (a) A degraded document image DIBCO 2012–5 (From left to right: Original, Otsu, and Proposed), (b) DIBCO 2013–H5 (Original, Otsu, and Proposed), (c) DIBCO 2013–P4 (Original, Otsu, and Proposed).

TABLE 6

OTSU AND THE PROPOSED METHODS FOR THE IMAGES IN FIG. 6

Image	*EM*/1000	
	Otsu [13]	Proposed Method
2012-5	440.357	1.821
2013-H5	104.171	1.282
2013-P4	475.424	11.320

TABLE 7

PHIBC'S WINNER [18] AND THE PROPOSED METHOD

The Method	*FM* %	*PSNR*	*EM*/1000	Time (sec)
PHIBC_12'S winner	88.55	18.66		
The proposed method	93.29	23.001	11.490	17

REFERENCES

[1] Gatos, B., Pratikakis, I., Perantonis, S.J., (2006): *Adaptive degraded document image binarization*. Pattern Recognition 39(3), 317-327.

[2] N. R. Howe, (2011), "*A laplacian energy for document binarization*," in Document Analysis and Recognition (ICDAR), International Conference on, pp. 6–10.

[3] Gatos, B., Ntirogiannis, K., Pratikakis, I., (2009): *ICDAR 2009 Document Image Binarization Contest -DIBCO 2009)*. In: Proc. Int. Conf. on Document Analysis and Recognition 14(1) 35-44.

[4] Pratikakis, I., Gatos, B., Ntirogiannis, K., (2010): *H-DIBCO 2010 – Handwritten Document Image Binarization Competition*. In: 12th ICFHR. 727-732.

[5] Pratikakis, I., Gatos, B., Ntirogiannis, K., (2011): *ICDAR 2011 Document Image Binarization Contest (DIBCO 2011)*. In: Proc. Int. Conf. on Document Analysis and Recognition. 1506-1510.

[6] Nratikakis, I., Gatos, B., K.,(2012):H-DIBCO 2012 *Handwritten Document Image Binarization Competition*. In: Proc. Int. Conf. on Frontiers in Handwriting Recognition. 813-818.

[7] I. Pratikakis, B. Gatos and K. Ntirogiannis, (2013), "*ICDAR 2013 Document Image Binarization Contest (DIBCO 2013)*", 12th International Conference on Document Analysis and Recognition (ICDAR 2013), pp. 1471 - 476, Washington, DC, USA.

[8] Ntirogiannis, K., Gatos, B., Pratikakis, I., (2014): *ICFHR 2014 Competition on Handwritten Document Image Binarization (H-DIBCO 2014*). 2167-6445, IEEEDOI -10.1109/ICFHR.141.

[9] I. Pratikakis, K. Zagoris, G. Barlas, and B. Gatos, (2016), "*ICFHR 2016 handwritten document image binarization contest* (H-DIBCO 2016)," in ICFHR. IEEE, pp. 619-623.

[10] I. Pratikakis, K. Zagoris, P. Kaddas, and B. Gatos, (2018),"ICFHR 2018 competition on handwritten document image binarization (H-DIBCO 2018)," in 16th ICFHR. IEEE.

[11] Otsu, N., (1978): *A threshold selection method from gray level histogram*. IEEE Trans. on System, Man, Cybernetics 19(1), 62-66.

[12] Sauvola, N., Pietikainen, M., (2000): *Adaptive document image binarization*. Pattern Recognition 33(2), 225-236.

[13] W. Niblack. , (1985), *An Introduction to Digital Image Processing*. Strandberg Publishing Company.

[14] Shangxuan Tian , Shijian Lu y, Bolan Su y and Chew Lim Tan, (2015), "*Robust Text Segmentation using Graph Cut*", **,** International Conference on Document Analysis and Recognition (ICDAR) .

[15] Su, B., Lu, S., & Tan, C. L., (2010), *Binarization of Historical Document Images Using the Local Max and Min. In Workshop on Document Analysis Systems (DAS) (pp. 159-166)*. Boston, USA.

[16] P. Viola and M. Jones, (2001), "*Rapid object detection using a boosted cascade of simple features*," CVPR, pp. 511– 518.

[17] I. Pratikakis, Joan P., A. H. Toselli, E. Vidal, and B. Gatos, (2017), *Handwritten Keyword Spotting Competition* (H-KWS 2016).

[18] S. M. Ayatollah and H. Z. Nafchi, (2013), "*Persian heritage image binarization competition (PHIBC 2012)*," in Proc. Iranian Conf. Pattern Recognition. Image Anal.