ICFHR 2012 Competition on Handwritten Document Image Binarization (H-DIBCO 2012)

Ioannis Pratikakis[1], Basilis Gatos[2] and Konstantinos Ntirogiannis[2,3]

[1]Department of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, Greece
e-mail: ipratika@ee.duth.gr

[2]Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, NCSR "Demokritos", Athens, Greece,
e-mail: bgat@iit.demokritos.gr

[3]Department of Informatics and Telecommunications, National and Kapodistrian University of Athens, Athens, Greece
e-mail: kntir@di.uoa.gr

Abstract— **H-DIBCO 2012 is the International Document Image Binarization Competition which is dedicated to handwritten document images organized in conjunction with ICFHR 2012 conference. The objective of the contest is to identify current advances in handwritten document image binarization using meaningful evaluation performance measures. This paper reports on the contest details including the evaluation measures used as well as the performance of the 24 submitted methods along with a short description of each method.**

Keywords: handwritten document image, binarization, performance evaluation

I. INTRODUCTION

Handwritten document image binarization contributes significantly to the success of the handwritten document image recognition challenging task. This makes it imperative to create a framework for benchmarking purposes, i.e. a benchmarking dataset along with an objective evaluation methodology in order to capture the efficiency of current image binarization practices for handwritten document images. To this end, following the success of DIBCO series competitions (DIBCO 2009 [1] organised in conjunction with ICDAR'09, H-DIBCO 2010 [2] organized in conjunction with ICFHR 2010 and DIBCO 2011 [3] organised in conjunction with ICDAR'11) the follow-up of this contest is organised in conjuction with ICFHR 2012. In H-DIBCO 2012 (Handwritten Document Image Binarization Contest), the general objective is to record recent advances in handwritten document image binarization using established evaluation performance measures. In this contest, we focused on the evaluation of document image binarization methods using a variety of scanned handwritten documents for which the corresponding binary ground truth image has been created. The authors of submitted methods had initially registered to the competition and downloaded representative document image samples along with the corresponding ground truth. At a next step, all registered participants were required to submit their binarization executable. After the evaluation of all candidate methods, the testing dataset (14 handwritten images with the associated ground truth) along with the evaluation software has been released as publicly available in the following link:

(http://utopia.duth.gr/~ipratika/HDIBCO2012/benchmark).

The remainder of the paper is structured as follows: Each of the methods submitted to the competition is briefly described in Section II. The evaluation measures are detailed in Section III. Experimental results are shown in Section IV while in Section V conclusions are drawn.

II. METHODS AND PARTICIPANTS

Nineteen (19) distinct research groups have participated in the competition with twenty four (24) different algorithms because certain participating research groups have submitted more than one algorithm. Brief descriptions of the methods are given in the following (the order of appearance reflects the chronological order of expressing an interest to participate in the competition).

1) LabGED Laboratory, Badji Mokhtar – Annaba University, Algeria (A. Kefali, T. Sari and H. Bahi): The submitted algorithm uses a feed-forward artificial neural network to classify the pixels into two classes: foreground and background. The neural network has a single hidden layer, 27 inputs and a single output neuron. To assign a new value (black or white) to a pixel, the neural network takes as input a vector of 27 values; the first 25 ones are the grayscale values of a 5*5 window. The last two values are the average gray levels of the two parts of the histogram which is obtained by an Otsu's threshold. All parameters (neural network configuration and statistical measures) were chosen after several experimentations.

2) Centre for Development of Advanced Computing, Kolkata, India (Yumnam Kirani Singh): This research group has submitted two algorithms : **(a)** The algorithm first finds the three cluster centers of a gray image using k-means. To get consistent cluster centers for a given image, we derive the initial three cluster centers from the pixel range of the image. The cluster centers are updated continuously as there is significant change in successive

978-0-7695-4774-9/12 $26.00 © 2012 IEEE
DOI 10.1109/ICFHR.2012.216

updates. When there is no change in cluster centers between two successive updates, the corresponding cluster centers are taken as final cluster centers. Then, the threshold for binarization is determined from the three cluster centers. The threshold is the average of the first cluster center and the average of the second and the third cluster center. Once the threshold is determined, the document image in gray is binarized by setting all pixels above the threshold value to ones and the remaining pixels to zeros. This algorithm works better in the case that document images with uniform background are binarized; **(b)** Sometimes, background noise in a document image may not be uniform. In such case, finding proper threshold for document image for binarization is difficult. To distinguish the printed or handwritten characters from the background pixels, a suitable edge enhancement filter is applied to the gray image. The resulting image is then binarized in the same way as in (a).

3) Institute of Software, Chinese Academy of Sciences, China (Longlong Ma, Haiyang Xu, Jian Wu): The submitted algorithm consists of the following five steps: (i) pre-processing using Wiener filter; (ii) estimating relative foreground based on the difference between the pre-processed image and estimated background surface as in [4]; (iii) the edge map of the image are also extracted using improved Canny edge detection based on the pre-processed image; (iv) the document image is binarized combining global with local threshold methods based on relative foreground and edge map; and (v) Heuristics rules are applied as a post- processing.

4) Institute for Infocomm Research, & School of Computing, National University of Singapore (Bolan Su, Shijian Lu, Chew Lim Tan): This research group has submitted two algorithms : **(a)** This method consists of four main steps. First, local image contrast which is evaluated by local maximum and minimum and local image gradient are combined using an exponential function with an adaptive factor. Second, the local image contrast is combined with the edge map to extract an accurate text character edge image. Third, the document image is binarized by a local threshold which is decided based on the constructed edge map and estimated stroke width. At last, some post-processing work is applied to produce better results. **(b)** The submitted algorithm is first using the local maximum minimum method to obtain the initial binarization results, then the Markov Random Field model is applied to generate the final binary results incorporating the edge information.

5) Graduate School of Maritime Sciences, Kobe University, Japan (Akihiro Okamoto, yuichi Nakata, Hiromi Yoshida, Naoki Tanaka): This research group has submitted **two** algorithms with the same underlying idea but with slightly different parameters such as the size of structuring element. Initially, the input image is converted into a grayscale image for which a median filtering is applied. Then, 'MaskImageI' is constructed by applying a morphological black top-hat operation and a binarization by the Otsu method. Also, a marker image J is generated by considering edge elements acquired by a morphological gradient and Canny edge detector. The final outcome is a reconstruction which is implemented via a morphological conditional dilation operation based on the 'MaskImageI' with marker J.

6) Smith College, Department of Computer Science, Northampton (MA), USA (Nicholas R. Howe): The submitted method is based on the ICDAR 2011 paper [5] which optimizes a global energy function based on the Laplacian image. It improves on that work by dynamically setting the regularization coefficient and Canny thresholds for each image. This is done using a stability criterion on the final result.

7) Tlemcen University & Qatar University, (Yazid Hassaïne Yazid, Abdelaali Hassaine and Somaya Al Maadeed): This method was adapted from a technique for restoration of optical soundtracks of old movies [6] in which the text part is considered as the opaque region of the optical soundtrack.

8) Institut für Nachrichtentechnik, Technische Universität Braunschweig (Germany), University of Guadalajara (Mexico), Institut für Informatik, Freie Universität Berlin (Germany) (Marte A. Ramírez-Ortegón, Volker Märgner, Erik Cuevas, Rául Rojas): In the submitted method, the region of interest (ROI) is detected by a novel technique based on gray-intensity variances. Then, an improved version of the transition method [7][8] is computed subject to the ROI. This generates a first binary image which is post-processed by several algorithms: edge restoration, bleed trough removal, object dilation, filling operators (to fill unwanted foreground holes), hole detection (to restore character holes), and smooth contour operators. Most of these operators are based on transition sets theory and techniques of expectation / maximization.

9) Hochschule Hannover - University of Applied Sciences and Arts, Hannover, Germany (Prof. Dr. Karl-Heinz Steinke): The submitted method is a combination of an adaptive threshold and Otsu approach, where we search for the threshold that minimizes the intra-class variance defined as a weighted sum of variances of the two classes. First, the image is convolved with the mean of a large pixel neighborhood (31x31). The original image is substracted from the convolved image. Then, the difference image is thresholded with a constant (C). The resulting image is compared with the Otsu-Image and a decision is made according to contained blobs (size and number).

10) Military College of Signals (MCS), National University of Sciences and Technology (NUST), Islamabad, Pakistan (Syed Ahsen Raza): The algorithm is based on three processing steps, namely, preprocessing, thresholding and post-processing. In preprocessing, conditional noise removal and edge based processing is performed. Thresholding step involves a computation of final threshold for background and text segmentation based on an average value computed through multiple thresholds (based on 4 different Niblack inspired thresholding formulas). In the final step of post processing again conditional noise removal and constrained morphological operations are performed to get the final binarized image.

11) LSIS, UMR CNRS 6168, South University of Toulon-Var, La Garde, France (Thibault Lelore and Frédéric Bouchara): The method first uses a median filtering of the input image and then upscale it using linear interpolation. The binarization is then computed through four different steps. First, text position is roughly estimated due to an edge detection approach. The edges previously detected are then filtered using a clustering algorithm, thanks to a method based on the size and the gradient magnitude of each edge. Then, another clustering algorithm is applied to pixel values close to the remaining edges which produces a three classes image (Text, Background and Unknown). Finally, a post-processing step assigns a class to 'Unknown' pixels using heuristic rules. We rescale the binarized image back to the correct resolution using bi-cubic interpolation. The final black and white image is obtained by using a global threshold.

12) Center of Informatics, Federal University of Pernambuco, Recife, PE, Brazil (Rafael Galvão de Mesquita, Carlos Alexandre Barros de Mello): This algorithm is based on the method proposed in [9], which simulates the information perceived by an observer by distance. Herein, we propose a new scheme to estimate the stroke width using a vertical Sobel´s filter. The estimated width is used to define (i) the size of the structuring elements of the morphological closing operations proposed in [9] and (ii) the simulated distance from the observer to the document, based on the Minimum Angle of Resolution used on Snellen visual acuity test. In addition, we use a combination of Otsu´s thresholding algorithm and k-means clustering algorithm (with 3 classes – one for background, one for foreground and one for intermediate colors) to perform the final binarization step.

13) University of Central Florida, USA, (Oliver Nina): This algorithm is an improvement of the Lu [10] algorithm. The main steps improved were the background approximation in which not only polynomial fitting is used for approximating the background but also a median filter of the original image with a large kernel is used. An average of both approximations is used to obtain the final background estimation. Also, the edge detection step was improved by applying a median filter on the compensated image and calculating the L_2 norm gradient magnitude on it. The final edge detection is obtained by using an average of the Otsu [11] and the minimum error threshold [12] as the final threshold for the high gradient values.

14) Dept. of Computer Science and Engineering, Jadavpur University, Kolkata, India (Ayatullah Faruk Mollah, Subhadip Basu, Mita Nasipuri): In the submitted algorithm, the input image is, at first, partitioned into blocks. Based on gray level intensity range of the blocks, they are assigned fuzzy membership values to denote the possibilities of being foreground blocks. Then, a neighborhood analysis is performed with the fuzzy values for finer classification of the blocks as part of foreground or background. Noisy blocks are expected to fall into background and thus, they get removed. After the classification, adjacent foreground blocks form text regions that are binarized with any suitable technique.

15) Synchromedia Lab, ETS, University of Quebec, Montreal, Quebec, Canada, (Reza Farrahi Moghaddam and Mohamed Cheriet): This method is an adaptive and parameterless generalization of Otsu's method based on using the estimated background. The adaptiveness is obtained by combining grid-based modeling and the estimated background map. The parameterless behavior is achieved by automatically estimating the document parameters, such as the average stroke width, the average line height, and the average text height. For initialization, a rough binarization map is used generated using a double-scale variation of the grid-based Sauvola method powered by a gray-scale nonlinear transformation. For more details, please see [13].

16) REGIM (ENIS-sfax-TUNISIA) (Nizar Zaghden, Badreddine Khelifi, Mohamed Adel Alimi): The submitted system consists of analysing the image and applying a variable threshold in every portion of the image. This means separating the pixels of images in two classes, the first one having a maximal level (typically 255) and assists a minimal level (0). This method of binarization requires beforehand the division of portions of the images that don't have a big difference in intensity values. Then, the separation in two classes is made in every region of the image. The system is a priori efficient for images having grey levels or even for colored images

17) REGIM – ENIS, University of Sfax, Tunisia (Mohamed Benjlaiel, Adel M. Alimi): The submitted method consists of three steps: (i) preprocessing step including adaptive histogram equalization, contrast adjustment and small noise abject removal using morphological operations; (ii) an adaptive Otsu's N-thresholding method [11] is applied for separating the pixels into foreground and background; (iii) Heuristics rules are applied as a post processing to remove remaining small connected components.

18) Department of Electrical Engineering, Bu-Ali Sina University, Hamedan, Iran and Synchromedia Lab, École de technologie supérieure, University of Quebec, Montreal, Quebec, Canada (Hossein Ziaei Nafchi and Hamidreza Rashidy Kanan and Reza Farrahi Moghaddam): This research group has submitted **three** algorithms : **(a)** This method consists of three steps: preprocessing, main binarization and posprocessing. Phase congruency, a robust edge detector, and morphological operators are used for preprocessing, by selecting the foreground candidate pixels [14]. Then, a variation of Otsu's method is used to further improves the binarization results. Locally-weighed mean phase and a noise reduction step, based on the phase congruency, are applied for further improving the results. Finally, the standard deviation of objects in the denoised image [15] is used to remove the background objects. During binarization, three snapshots of the binarized image are generated which are the intermediate output of the proposed method and only include the high-probability foreground pixels.
(b) This method consists of three steps: preprocessing, main binarization and posprocessing. Similar to (a) foreground candidates are selected [14]. An image denoising algorithm [15] and adaptive Gaussian smoothing filter are used for the main binarization. An algorithm is used to estimate the stroke width for each object. By considering the stroke width, a set of post processing steps, based on the phase congrunecy and the denoised image, are applied to further improves the binarization results.
(c) This method consists of three steps: preprocessing, main binarization and posprocessing. Similar to (a) foreground candidates are selected [14]. Then, a variation of Otsu's method is used to further improves the binarization results. Locally-weighed mean phase and a noise reduction step, based on the phase congruency, are applied for further improving the results. Finally, the standard deviation of objects in the denoised image [15] is used to remove the background objects. Similar to (a) three snapshots of the binarized image are generated which include the high-probability foreground pixels.

19) Institute for Language and Speech Processing, Athena - Research and Innovation Center in Information, Communication and Knowledge Technologies, Athens, Greece (Vassilis Papavassiliou, Fotini Simistira, Vassilis Katsouros): In the submitted algorithm the gray image is first filtered by applying median order filtering and the top-hat-by-reconstruction procedure to suppress probable specks in the intensities and compensate for non-uniform background illumination. Considering seed points, the pixels with values lower than 0.9*T (T is the Otsu's threshold), we apply the reconstruction transformation to provide a binary image (BW1) that includes the initial estimation of text regions. Then, we produce a binary image (BW2) with 1s at the locations where the Laplacian of the filtered image has positive values. By using BW1∩ BW2 as a marker and their union as a mask, the binary reconstruction transformation results to a binary image R with the text parts and some coarse regions due to the "noisy" Laplacian. Based on the local smoothness of R we remove these coarse areas from BW2 and finally, we iteratively convert each OFF pixel of BW1 into ON when it is a neighbor of an ON pixel in BW1 and of an ON pixel of BW2.

III. EVALUATION MEASURES

For the evaluation, the measures used comprise an ensemble of measures that have been widely used for evaluation purposes. These measures consist of (a) F-Measure; (b) pseudo F-Measure; (c) PSNR; (d) Distance Reciprocal Distortion Metric.

A. F-Measure

$$FM = \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \qquad (2)$$

where $\text{Recall} = \frac{TP}{TP + FN}$, $\text{Precision} = \frac{TP}{TP + FP}$

TP, FP, FN denote the True positive, False positive and False Negative values, respectively.

B. pseudo F-Measure

This measure has been introduced in [16]. It was motivated by the fact that each character has a unique silhouette which can be represented by its skeleton. In this respect, we assume that a perfect recall can be achieved in the case that each skeleton constituent of the ground truth has been detected. Compared with the typical F-Measure as presented in III.A, there exist a difference which concerns an alternate measure for recall, namely *pseudo-Recall (p-Recall)* which is based on the *skeletonized ground truth image*.

The skeletonized ground truth image is defined by the following equations:

$$SG(x, y) = \begin{cases} 0, & \text{background} \\ 1, & \text{text} \end{cases} \qquad (3)$$

Taking into account the skeletonized ground truth image, we are able to automatically measure the performance of any binarization algorithm in terms of recall.

p-Recall is defined as the percentage of the skeletonized ground truth image *SG* that is detected in the resulting MxN binary image *B*. p-Recall is given by the following equation:

$$\text{p-Recall} = \frac{\sum_{x=1, y=1}^{x=M, y=N} SG(x, y) \cdot B(x, y)}{\sum_{x=1, y=1}^{x=M, y=N} SG(x, y)} \; 100\ \% \qquad (4)$$

$$p-FM = \frac{2\times \text{p-Recall}\times \text{Precision}}{\text{p-Recall}+\text{Precision}} \quad (5)$$

C. PSNR

$$PSNR = 10\log(\frac{C^2}{MSE}) \quad (6)$$

where $MSE = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N}(I(x,y)-I'(x,y))^2}{MN}$

PSNR is a measure of how close is an image to another. Therefore, the higher the value of PSNR, the higher the similarity of the two MxN images is. We consider that the difference between foreground and background equals to C.

D. Distance Reciprocal Distortion Metric (DRD)

The Distance Reciprocal Distortion Metric (DRD) has been used before to measure the visual distortion in binary document images [17]. It properly correlates with the human visual perception and it measures the distortion for all the S flipped pixels as follows:

$$DRD = \frac{\sum_{k=1}^{S} DRD_k}{NUBN} \quad (3)$$

where DRD_k is the distortion of the k-th flipped pixel and it is calculated using a 5x5 normalized weight matrix W_{Nm} as defined in [17]. DRD_k equals to the weighted sum of the pixels in the 5x5 block of the Ground Truth GT that differ from the centered kth flipped pixel at (x,y) in the binarization result image B (Eq. 4).

$$DRD_k = \sum_{i=-2}^{2}\sum_{j=-2}^{2} |GT_k(i,j)-B_k(x,y)|\times W_{Nm}(i,j) \quad (4)$$

Finally, NUBN is the number of the non-uniform (not all black or white pixels) 8x8 blocks in the GT image.

IV. EXPERIMENTAL RESULTS

The H-DIBCO testing dataset consists of 14 handwritten document images for which the associated ground truth was built for the evaluation. Representative example of the dataset along with the associated ground truth image is shown in Fig. 1(a),(b). The document images of this dataset originate from the collections of the Library of Congress [18]. The selection of the images in the dataset was made so that should contain representative degradations which appear frequently (e.g. variable background intensity, shadows, smear, smudge, low contrast, bleed-through).

The evaluation was based upon the four distinct measures presented in Section III. At Table I, for each encountered measure, the detailed performance of each algorithm along with the widely used binarization techniques of Otsu [11] and Sauvola et. al. [19], is given. The final ranking as shown in Table I ('Score') was calculated by firstly, sorting the accumulated ranking value for all measures for each test image. The summation of all accumulated ranking values for all test images denote the final '*Score'* which is shown in Table I. Let $R^i(j,m)$ be the rank of the i^{th} method concerning the j^{th} image when using the m^{th} measure. Then, for each binarization method i, the final Score S_i is given by Eq. (9):

$$S_i = \sum_{j=1}^{14}\sum_{m=1}^{4} R^i(j,m) \quad (9)$$

To measure the processing time required by each method, we used only one image ("H06") from the testing dataset with an average size and a PC with dual processor at 2.2 GHz with 2GB of RAM memory.

TABLE I. DETAILED EVALUATION RESULTS FOR ALL METHODS SUBMITTED TO H-DIBCO 2012.

Rank	Method	Score	FM (%)	p-FM (%)	PSNR	DRD	Time (min)
1	**6**	**172**	**89.47**	**90.18**	**21.80**	**3.440**	**29.66**
2	11	340	92.85	93.34	20.57	2.660	0.26
3	4.a	412	91.54	93.30	20.14	3.048	0.09
4	13	435	90.38	95.09	19.30	3.348	0.92
5	7	494	91.85	92.19	19.65	3.056	0.66
6	8	501	89.98	92.07	19.44	3.761	19.44
7	18.a	531	91.16	93.36	19.32	3.163	7.15
8	18.c	570	90.20	92.96	19.07	3.819	8.50
9	5.a	575	89.69	92.47	19.00	3.961	0.42
10	3	601	89.48	91.99	18.74	4.547	4.93
11	19	620	89.62	93.05	18.78	3.741	0.50
12	18.b	664	89.48	92.23	18.65	4.407	8.00
13	5.b	676	89.21	91.94	18.79	4.162	0.63
14	15	708	87.73	91.67	18.50	3.929	296.79
15	12	725	89.29	92.50	18.59	3.859	0.43
16	4.b	806	87.01	91.44	18.26	4.418	10.49
17	10	893	86.53	88.38	17.83	4.777	5.66
18	2.b	896	85.91	91.36	17.58	5.476	0.16
19	14	902	86.37	90.10	17.90	4.627	0.02
20	9	978	84.04	88.42	16.82	7.116	0.01
21	2.a	1012	80.75	85.04	16.72	6.516	0.09
22	1	1049	82.06	79.41	16.72	7.526	9.68
23	17	1095	81.00	80.96	15.36	23.107	0.60
24	16	1145	75.23	80.15	15.94	7.792	0.73
-	Otsu	-	80.18	82.65	15.03	26.458	-
-	Sauvola	-	82.89	87.95	16.71	6.594	-

Overall, the best performance is achieved by ***Method 6*** which has been submitted by Nicholas Howe from Smith College, Department of Computer Science, Northampton

(MA), in USA. Example binarization result of the winning method is shown in Fig. 1(c). For the sake of clarity, the complete set of binarization results of each participating binarization method can be found in the following link: *http://utopia.duth.gr/~ipratika/HDIBCO2012/ hdibco2012results.htm*.

V. CONCLUSIONS

The H-DIBCO 2012 Handwritten Document Image Binarization Contest attracted 19 research groups that are currently active in document image analysis. The general objective of the contest is to identify current advances in handwritten document image binarization using meaningful evaluation performance measures. This objective is fulfilled by firstly, providing short descriptions of each submitted algorithm, thus, enabling the interested researchers to be aware of the highly performing algorithms and be able to push forward the state of the art by a new more advanced approach. Secondly, the public availability of the testing dataset and the evaluation software permits further benchmarking and comparison with H-DIBCO 2012 results.

REFERENCES

[1] B. Gatos. K. Ntirogiannis and I. Pratikakis (2009). "ICDAR 2009 Document Image Binarization Contest (DIBCO2009)". 10th International Conference on Document analysis and Recognition (ICDAR'09). July. 26-29. 2009. Barcelona. Spain. pp. 1375-1382.

[2] I. Pratikakis, B. Gatos and K. Ntirogiannis, "H-DIBCO 2010 – Handwritten Document Image Binarization Competition", ICFHR'10, Kolkata, India, pp. 727-732, 2010.

[3] Pratikakis I., Gatos B., and Ntirogiannis K. (2011), "ICDAR 2011 Document Image Binarization Contest (DIBCO2011)", 11th International Conference on Document analysis and Recognition (ICDAR'11), September 18-21, 2011, Beijing, China, pp. 1506-1510.

[4] B. Gatos, I. Pratikakis, S.J. Perantonis, Adaptive Degraded Document Image Binarization, Pattern Recognition, Volume 39, 2006, pp. 317-327.

[5] N. Howe, "A Laplacian Energy for Document Binarization", ICDAR 2011, pp. 6-11

[6] A. Hassaine, E. Decenciere, B. Besserer, Efficient restoration of variable area soundtracks, Image Analysis & Stereology, 28 (2), 2009, pp. 113-119.

[7] Ramírez-Ortegón, M. A.; Tapia, E.; Ramírez-Ramírez, L. L.; Rojas, R. & Cuevas, E., 'Transition pixel: A concept for binarization based on edge detection and gray-intensity histograms', Pattern Recognition, 2010, 43, 1233 – 1243.

[8] Ramírez-Ortegón, M. A.; Tapia, E.; Rojas, R. & Cuevas, E. 'Transition thresholds and transition operators for binarization and edge detection', Pattern Recognition, 2010, 43, 3243 - 3254.

[9] C. A. B. Mello, "Segmentation of Images of Stained Papers Based on Distance Perception," in IEEE International Conference on Systems Man and Cybernetics, Istanbul, Turkey, 2010, pp. 1636 - 1642.

[10] S. Lu et al. Document image binarization using background estimation and stroke edges, IJDAR 2010.

[11] N. Otsu, "A threshold selection method from graylevel histograms", IEEE Trans. Systems Man cybernet.Vol. 9 (1), 1979, pp. 62–66.

[12] J. Kittler, J. Illingworth, Minimum error thresholding, Pattern Recognition, Volume 19, Issue 1, 1986, pp. 41-47.

[13] Farrahi Moghaddam, Reza and Cheriet, Mohamed; AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization, Pattern Recognition, 2012, vol 45(6), pp 2419-2431.

[14] H. Ziaei Nafchi, H. Rashidy Kanan. A Phase Congruency Based Document Binarization. in proceedings of the IAPR International Conference on Image and Signal Processing, 2012.

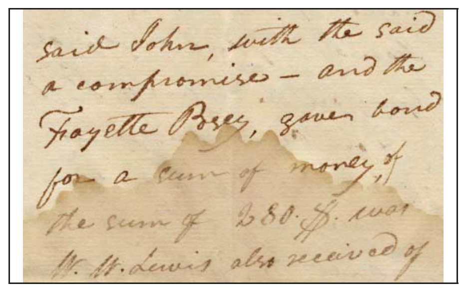
[15] P. Kovesi. Phase Preserving Denoising of Images. The Australian Pattern Recognition Society Conference: DICTA'99, pp. 212-217, 1999.

[16] K. Ntirogiannis. B. Gatos and I. Pratikakis. An Objective Evaluation Methodology for Document Image Binarization Techniques. The 8th IAPR International Workshop on Document Analysis Systems (DAS 2008). Nara Prefectural New Public Hall. Nara. Japan. September 17-19. 2008. pp. 217-224.

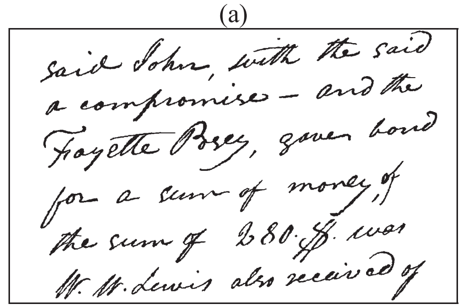
[17] H. Lu, A. C. Kot and Y.Q. Shi, "Distance-Reciprocal Distortion Measure for Binary Document Images", IEEE Signal Processing Letters, vol. 11, No. 2, pp. 228-231, 2004.

[18] Library of Congress (http://memory.loc.gov/ammem).

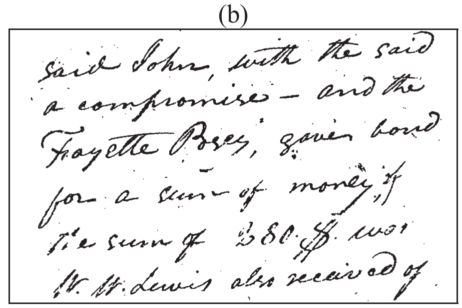
[19] J. Sauvola and M.Pietikainen. "Adaptive Document Image Binarization". Pattern Recognition. vol. 33. no. 2. pp. 225-236, 2000.



(a)



(b)



(c)

Figure 1. (a) Original image (b) Ground truth image; (c) Binarization result from the H-DIBCO 2012 winning ***Algorithm 6.***