A Binarization Method for Degraded Document Image using Artificial Neural Network and Interpolation Inpainting

Naouel Ouafek
Abdelhamid Mehri University Constantine2, and ENSC , Ali Mendjli, Constantine, Algeria.
Email:naouelouafek@yahoo.fr

Mohamed-Khireddine Kholladi
MISC laboratory Abdelhamid Mehri University Constantine2 . and El-Oued University.
Constantine, Algeria
Email:kholladi@yahoo.fr

***Abstract*—The binarization of degraded document image is an important phase in the document analysis to extract text from the original image. For this purpose, we present a new multi-stage approach started by, a preprocessing phase to enhance the contrast of the original image and to eliminate noises. Then an interpolation inpainting is performing to create the background estimation and the foreground (text) estimation, that will be used as feature to train and test a supervised artificial neural network. the final stage is the use of the ANN to classify the pixels of the original image into 0 for black and 1 for white to get the final binary image. The suggested method has been tested and evaluated over DIBCO'2009 database, the results show the successful of the proposed approach.**

***Index Terms*—Historical document binarization, Document Analysis, inpainting, ANN.**

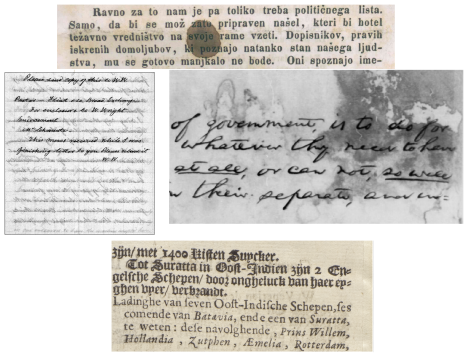
I. Introduction

The goal of the document image binarization is to create a perfect separation between text pixels and background pixels, The binarization is a type of classification problem that segment the document image into two classes: 0 or 1.
unfortunately, most of the historical document images suffer from different types of degradation, such as bleed-through, large stains, tears ink, blur. Figure (1) shows some examples of degradations. Those multiple degradations make the process of binarization a challenging task for researchers.

The good results of the binarization step are inescapable for improving the quality of text recognition system and document image analysis.
In the present work, a new multi-stage binarization method is suggest. The method exploit the benefit of the artificial network with an automatic extraction of features. This extraction is done with an inpainting phase in two way: one for erase the text pixels to get the background pattern , and the second one to erase the background pixels to get a text pattern.
This paper is organized as follows: an introduction in the first section, a second section describes all stages of the proposed method. The third section presents the experimental results and the last section suggests a conclusion for this work.

978-1-5386-4225-2/18/$31.00 © 2018 IEEE

Fig. 1. Degradation types



II. Literature Work

Many works were proposed in the literature to solve this problem. They can be classified as global, local and combined methods. In global methods like the Otsu's method [3] the author uses one value for thresholding the whole image, this method cannot adapt to the background noise and if the document suffer from more than one type of degradations. The Niblack's method [4] introduces the local thresholding by calculating for each pixel a new threshold according to local statics (standard deviation and mean of a window centred in this pixel). Those two methods are the most widely used so far as initial phase in multitude combined and novel methods. The Sauvola's method [5] is derived from Niblack's method, in their works they use a dynamic range of the standard deviation image to decrease the background noise. In Su et al [15], the authors use Otsu's method to calculate the high contrast pixels in "the image contrast" which is evaluated by the local maximum and minimum. Finally the image is thresholded by the statistics of a local neighborhood window. In 2013, Su et al [21] extend their previous work by combining the "contrast image" and the "gradient image". Moreover they use canny edge detector [18] and Otsu's method to detect the stroke width. Lastly, they improve their results by some post processing steps. Howe [16] proposed an algorithm to optimize the

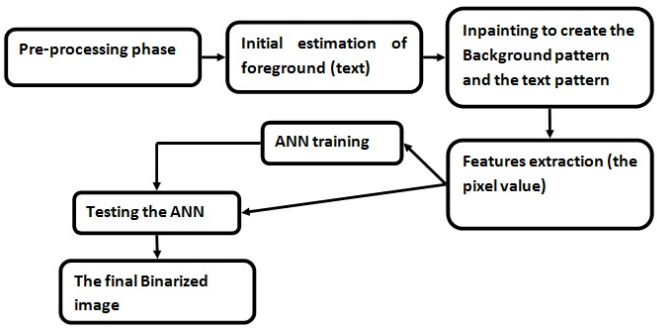


Fig. 2. Different phases of the proposed methods

global energy function based on Laplacian image and Canny edge detection [18]. Additional work of howe in 2012 [17] he proposed an automatic choice of parameters of his method for each image. The study of the background estimate is carried out by several authors in the last 10 years, Gatos 2006 [19] used Sauvola's method as initial binarized image to estimate background pixels. Then a local thresholding is applying to achieve final binarization. K. Ntirogiannis [22]2014 propose an inpainting algorithm to create the background image which is used to normalize the original image. After that they used a combined global and local adaptive binarization method to obtain the final binary image. The proposed method takes into account the stroke width and the contrast of the normalized image. Recent works of [23] in 2015, the authors use an adaptive median filtering to estimate the background pixels and this pixels are removing from the original image according to a threshold T. The final binary image is obtained by using a combination of two Gaussians clustering model. In [24] 2015, the proposed approach is a parametric configuration algorithm to improve outputs of other binarization algorithms.

There are several classification-based binarization methods have been proposed. In [26] a neural network is used for local image thresholding, the authors perform 8 statistical and textural image features to obtain a feature vector for each pixel from a window to train the network. In [27] the authors propose a selection of optimal parameters for the method [26], the main pillar of the method is that the user is the one who selects a pixel and sets it in text pixel or background pixel, than all the 8 features are calculated after this selection.

III. Proposed method

The main objective of this work, is to provide a final binarized image of the historical document, that maintains as much as possible text areas without any residual pixels of the background.
The proposed method is divided into four phases, which will be detailed in subsequent sections. Figure (2) represents the diagram of the proposed system.

A. Pre-processing

Let define $I(x, y)$ as the grayscale of the original historical document image. In the first step of pre-processing , we apply a two-dimensional Gaussian filter which is a low pass filter with $\sigma = 2$ where σ is the standard deviation. This filter eliminate the noise, and smooth the background texture to make it more uniform. Furthermore, it enhance the quality of text characters by increasing contrast between background and text areas. We have chosen $\sigma = 2$ because larger values of σ produce a wider peak (greater blurring).

B. Initial estimation of text area

To do the first estimation of text pixels , the enhanced image I is binarized with an effective binarization method. In our case, the approach proposed by Sauvola for adaptive thresholding [5], is used to create the initial binary image $S(x, y)$, which is shown in figure (3).
This image S will be used in the next stage as a mask for an

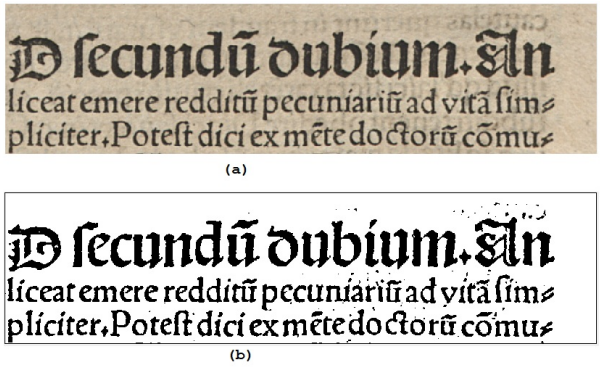


Fig. 3. (a): The original image , (b): The initial binary image S

inpainting process.

C. inpainting phase

The main purpose of this step is to bring out two images from the original image. The first image contains only text pixels and the second image contains only background pixels. Before performing the inpainting process two morphological operations are made on the initial binary image S:

- A dilation operation with a structuring element of radius $rad = 2$, is applied on the image S, in order to be sure that we take only pixels belonging to text class. Figure (4,(b)) presents the obtained image D_S.
- In parallel, we increase the thickness of the detected text for the purpose to cover more pixels from the background. To obtain this, an erosion operation with a structuring element of radius $rad = 2$, is made on the image S. Figure (4,(a)) presents the obtained eroded imageE_S.

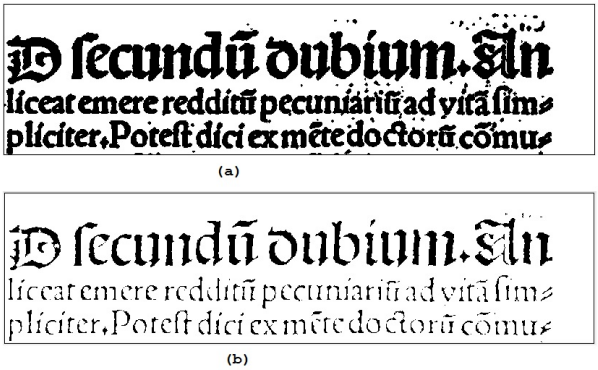


Fig. 4. (a): The eroded binary image E_S, (b): The dilated binary image D_S

The inpainting algorithm used in this work is a simple algorithm proposed by [1].
The algorithm propagate information from the outside of the area to be inpainted into its inside. The user provide a binary mask in which 0 represents the area that must be preserved and 1 the area to be painted.
We perform the process of inpainting in two ways:

- In the first way, we use the complementary image of the eroded E_S as a mask for the inpainting algorithm. The image obtained in this step is the images B_p in which the algorithm propagate background pixels into text areas, figure (5,(a)).
- In the second way, the algorithm propagate text pixels into background area by using the dilated image D_S as a mask of the inpainting process. The image obtained in this step is the image T_p shown in figure (5,(b)).

The images B_p, T_p represents the background pattern and text pattern respectively.
These patterns will be used in the next step as automatic selection of intervals of pixel values. The first interval pixels represents the text class and the second interval pixels is the background class.

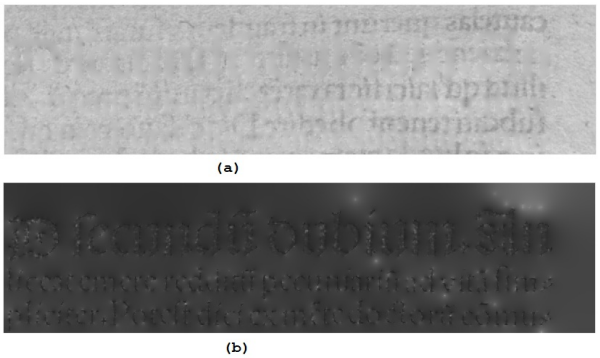


Fig. 5. (a): The background pattern B_p, (b): The text pattern T_p

D. Using ANN for binarization

First, the images I, B_p, T_p, are reshaped as vectors before using it by the ANN.
The ANN used in the proposed method perform back propagation algorithm with a sigmoid activation function which is defined by:

$$f(x) = \frac{1}{1 + e^{-ax}} \tag{1}$$

This ANN learns by a supervised learning because, we provide the Ground Truth image Gt as the target output. Gt is furnished in the dataset with its corresponding image.

Our training and test sets are composed of three input vectors, which are: the gray pixels values of the original image I, the background pattern B_p, and the text pattern T_p. The image Gt is also reshaped as a vectors than used as the predicted output vectors.

Every input vectors represent a feature for our ANN and every value in the vector (the pixel value) is a sample of the feature.

So we used 3 input neurons in the first layer and 1 output neuron in the last layer with one hidden layer.
Figure (6) presents the final binary image obtained by the suggested method.

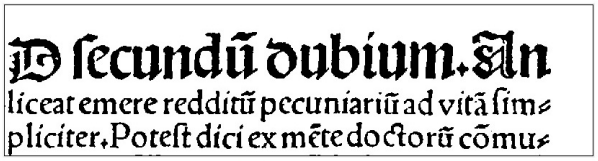


Fig. 6. The resulting binary image using the ANN

IV. Experiment results

In this section, we evaluate the performance of the proposed method by : Precision(3), Recall(2) and F-Measure metrics(4) [14].

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$F - measure = \frac{2 \times Recall \times Precision}{Recall + Precision} \tag{4}$$

Where TP, FP, FN are the true positive, false positive and false negative values respectively.
we have just used DIBCO'2009 [25] to test the idea of the proposed method. Future implementations on other datasets are the perspective of our work.
we randomly split the dataset into 20% for training the ANN and 80% for testing.

Pr2009	Precision	Recall	F-Measure
Proposed Method	0.9394	0.8927	0.9127
sauvola	0.89	0.9181	0.9009
Otsu	0.8943	0.9385	0.9127
Su et al	0.9800	0.8648	0.9168
Howe	0.9273	0.9578	0.9422
Hw2009	Precision	Recall	F-Measure
Proposed Method	0.8321	0.8426	0.8172
Sauvola	0.7897	0.8852	0.8184
Otsu	0.5805	0.9449	0.6594
Su et al	0.8357	0.8572	0.8251
Howe	0.9211	0.9271	0.9234

TABLE I
COMPARISON OF AVERAGES

The final output of the ANN is scaled to the interval [0, 1] (0 for black and 1 for white) to have a final binary image.

All stages of the proposed method are implemented in Matlab 2016.

For an objective evaluation we have reimplemented some known binarization methods in our PC ,so that all methods undergo for the same level of PC performance.
The methods used for comparison are : Sauvola's method ([5]) , Su et al ([6]) , Otsu ([3]) ,and Howe([17]).
We present in TABLE I , the experimental values for Machine printed images DIBCO'2009, Handwritten images DIBCO'2009. the values shown in the table are the averages of performance measures values of the proposed method and the comparison methods.

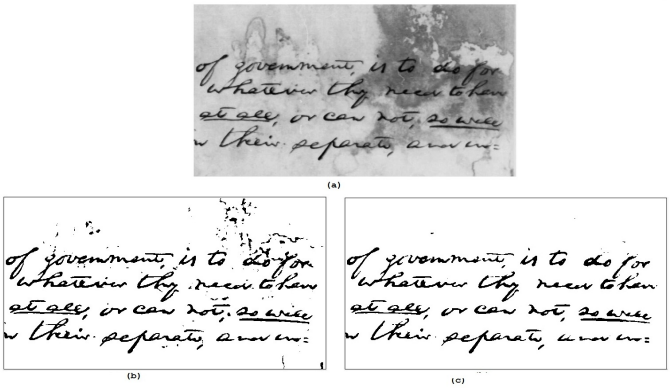


Fig. 7. (a):The original image , (b) The Sauvol's result,(c) The result of the proposed method.

The figure (7), (8), and (9) show a remarkable improvement obtained by the proposed method, there is a minimum of residual background pixels.

V. CONCLUSION

In this paper we have presented a new binarization method of historical degraded document. The method is composed of several stages and based mainly on the use of a supervised artificial neurone network ANN. We have presented the first version of the idea in which, the pixel value of the estimated background image, the estimated text image and,

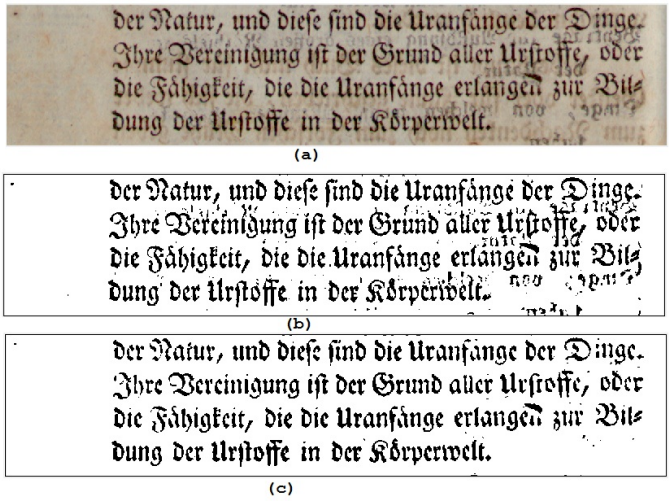


Fig. 8. (a):The original image , (b) The Sauvol's result,(c) The result of the proposed method.

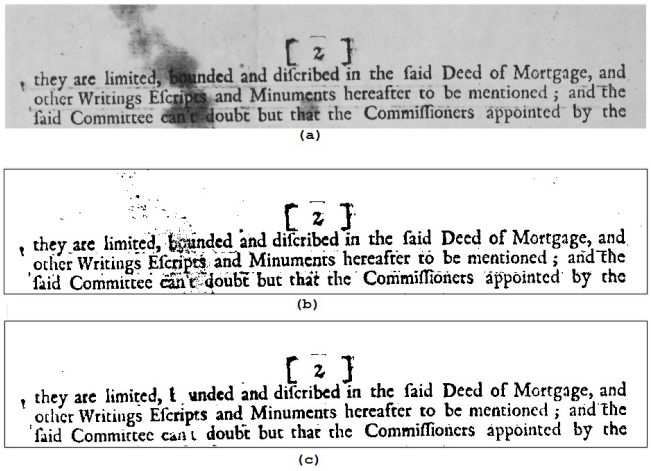


Fig. 9. (a):The original image , (b) The Sauvol's result,(c) The result of the proposed method.

the pixels value of the original image are used as feature to train and test the neural network. In a future version the use of measures of statistical textures which can be performed on a neighborhood group of pixels, for example : standard deviation ,the mean,etc.. instead the use of the pixel value to improve the process of classification.

REFERENCES

[1] M. Bertalmio and G. Sapiro and V.Caselles and C. Ballester, "Image Inpainting", Proceding of SIGGRAPH, 417–424, 2000
[2] K. Uhlr and V. Skala, "Reconstruction of damaged images using radial basis functions", In Proceedings of EUSIPCO 2005 (Bogazici University, 34342, Bebek, Istanbul, Turkey), p. 160,2005
[3] N. Otsu, "A threshold selection method from gray level histogram",IEEE Trans. Syst., Man, Cybern,vol. 19 no. 1,6266,Jan. 1979
[4] W. Niblack,An Introduction to Digital Image Processing,115-116,1986
[5] Sauvola and M. Pietikainen, "Adaptive document image binarization", Pattern Recognition, vol. 33, no. 2, 225236, 2000

[6] Su et al, " Robust document image binarization technique",In IEEE transactions on image processing,VOL. 22, NO. 4,225236, APRIL 2013.

[7] B. Gatos and I. Pratikakis, and S.J. Perantonis, " Adaptive degraded document image binarization ",In Pattern Recognition, vol. 39,317327,2006 .

[8] R.Hedjam and R.F Moghaddam and M. Cheriet, "A spatially adaptive statistical method for the binarization of historical manuscripts and degraded document images ",In Pattern Recognition , vol 44 (9),21842196,2011 .

[9] K. Ntirogiannis and B. Gatos b, I. Pratikakis," A combined approach for the binarization of handwritten document images ", In Pattern Recognition Letters, vol 35,(2014)

[10] R.Hedjam and M.Cheriet, "Historical document image restoration using multispectral imaging system", In Pattern Recognition, 2013 .

[11] N. Mitianoudis and N. Papamarkos, " Multi-spectral document image binarization using image fusion and background subtraction techniques",In Proceeding of IEEE International Conference on Image Processing,2014.

[12] N. Mitianoudis and N. Papamarkos, "Local cooccurrence and contrast mapping for document image binarization", In Proceeding of 14th International Conference on Frontiers in Handwritten Recognition (ICFHR14), Crete Island,Greece , 2014.

[13] A. Hyvarinen and J. Karhunen and E. Oja, "Independent Component Analysis",IN John Wiley, 481+xxii, 2001.

[14] R.Hedjam and H. Ziaei Nafchi and R.F.Moghaddam and M. Kalacska and M. Cheriet," ICDAR 2015 Contest on MultiSpectral Text Extraction (MS-TEx 2015) ", In Competitions @ 2015 13th International Conference on Document Analysis and Recognition (ICDAR), 2015 .

[15] B. Su and S. Lu and and C. L. Tan,Binarization of historical handwritten document images using local maximum and minimum filter, In Proceding International Workshop Document Anal. Syst,159166, Jun. 2010.

[16] N. Howe," A laplacian energy for document binarization",In International Conference on Document Analysis and Recognition, 610.6, 2011 .

[17] N. Howe, "Document binarization with automatic parameter tuning, In International Journal on Document Analysis and Recognition, 2012.

[18] J.Canny, "A computational approach to edge detection, In IEEE Trans. PatternAnal. Mach. Intell.8(6), 679698, 1986

[19] B. Gatos and I. Pratikakis and S. Perantonis, "Adaptive degraded document image binarization, In Pattern Recognit, 39, 3173278,2006.

[20] Z. Chi, K. W. Wong, "A two-stage binarization approach for document images", in: Proceedings of the International Symposium on Intelligent Multimedia, Video and Speech Processing, IEEE, 2001, pp. 275-278

[21] Su, B.and Lu, S.and Tan, C. L.r,"Robust document image binarization technique for degraded document images",In IEEE Transactions on Image Processing , 22, 14081417, 2013.

[22] K. Ntirogiannis and B. Gatos and I. Pratikakis,"A combined approach for the binarization of handwritten document images, In Pattern Recognition Letters, 35, 3-15, 2014.

[23] N. Mitianoudis and N. Papamarkos,"Document image binarization using local features and Gaussian mixture modeling,In Image and Vision Computing, 38, 3351, 29 April 2015.

[24] Rafael G. Mesquita and Ricardo M.A. Silva and Carlos A.B. Mello and Pricles B.C. Miranda,"Parameter tuning for document image binarization using a racing algorithm, In Expert Systems with Applications,42, 25932603,2015.

[25] Gatos, B., K. Ntirogiannis, and I. Pratikakis. 2009a. ICDAR 2009 Document Image Binarization Contest (DIBCO 2009) . In ICDAR09. p. 13751382.

[26] Y. Alginahi, Computer Analysis of Composite Documents with Non-uniform Background, PhD Thesis, Electrical and Computer Engineering, University of Windsor, ON, Canada, 2004.

[27] M. J. Islam, M. Ahmadi,M. A. Sid-Ahmed,Y. M. Alginahi "Optimal Parameter Selection Technique for a Neural Network Based Local Thresholding Method", Journal of Pattern Recognition Research 1 (2010) 69-94