Revue-IRS

Revue Internationale de la Recherche Scientifique
(Revue-IRS)

Vol. 1, No. 1, Décembre 2022

A New Hybrid Image Thresholding Technique Based On Local and Global Parameters Applied to Ancient Document Analysis

Saïd ECH-CHADI*, Hajar ELKIHEL and Youness ECHCHADI

Computer Science Department, Laboratory of Mathematic, Informatics and Communication Systems,
ENSA Safi Route Sidi Bouzid BP : 63, Safi, Morocco

https://orcid.org/0000-0002-2254-192X)

Abstract: Digital archiving has helped in the development of the exploitation of historical documents, but the number of old documents analysed in national archives and libraries that are specialized in old documents remains very little and dependent on advances in the field of image processing especially thresholding [1].

In this paper we propose a method for segmenting text in ancient documents based on a new hybrid adaptive binarization technique which integrates local and global parameters, in consideration of resolving unsolved problem in thresholding of degraded document images [2] due to the low quality of this type of image, the main idea of our method is to combine the local behaviour and the global behaviour of lighting and contrast variation within the image.

The direct application of our approach will improve the quality of processing of old documents, thus character recognition and in-depth analyzes based on contrasts and the types of inks used, will make it possible to better understand the cultural and historical riches existing in old documents.

Keywords: Hybrid Image Thresholding, Ancient Document Analysis, Local and global parameters, F_measure evaluation.

Digital Object Identifier (DOI): https://doi.org/10.5281/zenodo.7491859

1 Introduction

Today, we are taking more care of our cultural heritage, which constitutes a vivid and collective memory of our societies.
The evolution of our economies towards a model based on digital content has a profound impact on the interest of preserving this cultural heritage, especially for old documents and historical documents. Since the earlier attempts at digitizing documents, it has always contributed to the support of this impact. Enormous efforts have been

invested in cultural heritage digitization programs, mainly through maps, historical documents and manuscripts [3].
The major challenge in the evolution of digitization techniques have always been the quality of the document, which is degraded by the presence of certain phenomena such as discolored paper, ink expansion, uneven color shade, torn paper and other disturbances such as existence of small spots and others [4].
The current tendency to explore and exploit historical documents, call for the development of new techniques based on image processing. In fact, various processings of this kind use pre-processing chains whose main core is a binarization technique that extracts the true content, which will allow a correct interpretation of historical documents.
Image thresholding techniques - one family of binarization methods - are the most used in the pre-processing chains. Through their optical characteristics, ancient documents tend to require particular thresholding techniques. Indeed, among those methods, one developed by Sauvola (2000), another one called Integral Sum Image (2001) and one recently developed by Bataineh (2010) allow a binarization with quite satisfactory performance rate. Our research aims to develop a new method that happens to improve thresholding performance and quality of interpretation that will follow this preprocessing step (OCR) [5].
During the past decade, many adaptive binarization techniques have been proposed [6 - 8]. Sezgin and Sanker [9] have developed a comparative study of evaluation of these methods.
In current techniques, the binarization (threshold selection) is usually performed either globally or locally. Some hybrid methods have also been proposed. The global methods use one calculated threshold value to divide image pixels into object or background classes, whereas the local schemes can use many different adapted values selected according to the local area information. Hybrid methods use both global and local information to decide the pixel label [8][10].
The variability in the image influences the binarization technique to be developed. The development of a universal technique thus becomes complex and difficult, but attempts have been made in this direction [11][12].
In this paper, we propose a new formula to determine the threshold for a new adaptive binarization. It introduces and analyses this new method, which is able to process degraded document images altered by shadows, non-uniform illumination, low contrast, smears and noise [13]. Embedded and mobile development becomes a new challenge for these image processing methods [10].
In Section 2 we explore related works by developing the advantages and disadvantages of various methods. In Section 3 we present the theoretical basis of our approach while in Section 4 we put forward the experimental results and discussions. We drew conclusions at the end.

2 Related works

Thresholding is one of the most powerful tools for image segmentation. The segmented image obtained from thresholding has the advantages of smaller storage space, fast processing speed and ease in manipulation, compared with gray level image which usually contains 256 levels. Therefore, thresholding techniques have drawn a lot of attention during the past 20 years [14][15].

2.1 Adaptive binarization methods

2.1.1 Niblack's and Sauvola's Methods

We focus our study on adaptive binarization methods. Indeed, in these methods, the value of the threshold is determined by its local neighbours, thereof Niblack has developed the first adaptive thresholding method [9], its principle [15-16] is the determination of a threshold surface around the local mean m, Standard deviation s of the gray values calculated in a local matrix in the following form :

$$T = m + k \cdot s \qquad k = -0.2 \tag{1}$$

Sauvola et al [10] hypothesized the value of the pixels representing the text (transformed into a value 0) and those representing the non-text (transformed into a value 255) [15-17], proposing a formula for determining local thresholds:

$$T = m \cdot \left(1 - k \cdot \left(1 - \frac{s}{R}\right)\right) \qquad k = 0.5 \tag{2}$$

R represents the dynamic deviation of the gray values of the image; its value is set to 128.

However, outside from the hypothesis where it was issued, the Sauvola method is confronted with problems [16-17].

In another attempt, Wolf et al [16] [18] used the normalization of contrast and the mean gray value of the image to propose the following formula:

$$T = (1-k) \cdot m + k \cdot M + k \cdot \frac{s}{R} \cdot (m - M) \qquad k = 0.5 \tag{3}$$

M is the minimum value of the gray values of the image.

R expresses the maximum of the standard deviations of the gray values in various windows.

Thus, the formula proposed by Wolf et al. allows to improve performances, but these performances are degraded during the sudden changes in the gray values of the background since the influence of the minimum value of gray and Standard deviation is notable [15].

2.1.2 Integral Sum Image

The Integral Sum Image method [18] has been integrated in the field of image processing by Viola and Jones [20-21], for an input image I, its integral sum image g is calculated in each Pixel from the sum of the values of the grays above and to the left of the concerned pixel, thus the value of the pixel g (x, y) is:

$$g(x,y) = \sum_{i=1}^{x} \sum_{j=1}^{y} I(i,j) \tag{4}$$

A single pass allows to calculate the integral image in an efficient way.

$$g(1,y) = I(1,y) + g(1,y-1) \qquad y = 2 \dots . n \tag{5}$$

$$g(1,y) = I(x,1) + g(x-1,1) \qquad x = 2 \dots . m \tag{6}$$

$$g(x,y) = I(x,y) + g(x,y-1) + g(x-1,y) - g(x-1,y-1) \tag{7}$$

Equation (5-7) calculates the integral sum image [17] in a single pass.

$$g(x,y) = \sum_{i=x-c}^{x+c} \sum_{j=y-c}^{y+c} I(i,j) \tag{8}$$

Where c= (w-1)/2 , since w is an odd number.

Then the local sum s(x,y) will be calculated using two additions and a subtraction as in equation (9).

$$s(x,y) = [g(x+d-1,y+d-1) + g(x-d,y-d)] - [g(x-d,y+d-1) + g(x+d-1,y-d)] \tag{9}$$

Where d=round(w/2)

The local arithmetic mean m (x, y) in (x, y) is the average of the pixels in the window of size w×w of the image I. From equation (9)

$$m(x,y) = \frac{s(x,y)}{w^2} \tag{10}$$

All proposed thresholding methods in previous works, have been demonstrated to be effective in constrained pre-processing environments with predictable images. None has proven effective in all cases of general document image processing.

2.2 Hybrid adaptive Binarization Methods: Bataineh's method

Based on the work of Niblack, Sauvola and NICK. Bataineh et al [12] proposed a new method using the mean, the standard deviation and other factors, this hybridization has overcome the weaknesses of adaptive methods described in the previous chapter [12].

The strength of the method proposed by Bataineh is found in the resolution of the problems of thin texts (pen strokes) and low contrast between the foreground and the background [12].

Another feature of Bataineh's method is that it deals with binarization problems, such as thin pen-strokes and low-contrast images [13]

To develop the thresholding formula, Bataineh have used the following expression [12]:

$$\sigma_{Adaptive} = \frac{\sigma_w - \sigma_{min}}{\sigma_{max} - \sigma_{min}} \cdot \max_{level} \quad (11)$$

At the end, the threshold calculation proposed by Bataineh [12] is as follows:

$$T_w = m_w - \frac{m_w^2 * \sigma_w}{(m_g + \sigma_w)(\sigma_{Adaptive} + \sigma_w)} \quad (12)$$

3 Our approach

3.1 Proposed binarization algorithm

A new method for calculating the adaptive threshold is proposed. This method is likely to improve thresholding in the image processing field, especially for ancient documents (degraded state).

We derive our formula from the work of Niblack.

The idea of our method is to vary the threshold over the image, based on the local mean m_L, local standard deviation σ_L, global mean m_G, and global standard deviation σ_G, computed in a small neighbourhood of each pixel. This approach considers both global and local parameters, in order to resolve specific situations by enhancing some light textures that represent elements belonging to the binarized object.

We suppose that the pixel intensity depends on two functions f and g, where f comprises the dependency of local parameters m_L and σ_L, and g comprises the dependency of global parameters m_G and σ_G.

$$T(x, y) = f(m_L, \sigma_L) + g(m_G, \sigma_g) \quad (13)$$

We suppose that random variations of the intensity are governed by the global mean and the variations of the constituents f and g, hence:

$$T(x, y) = m_L(f'(m_L, \sigma_L) + g'(m_G, \sigma_g)) \quad (14)$$

By expliciting the variations:

$$T(x, y) = m_L\left[\left(\frac{\partial f}{\partial m_L} \cdot dm_L + \frac{\partial f}{\partial \sigma_L} \cdot d\sigma_L\right) + \left(\frac{\partial g}{\partial m_G} \cdot dm_G + \frac{\partial g}{\partial \sigma_G} \cdot d\sigma_G\right)\right] \quad (15)$$

$$T(x, y) = m_L\left[\left(\frac{\partial f}{\partial m_L} \cdot dm_L + \frac{\partial g}{\partial m_G} \cdot dm_G\right) + \left(\frac{\partial g}{\partial \sigma_G} \cdot d\sigma_G + \frac{\partial f}{\partial \sigma_L} \cdot d\sigma_L\right)\right] \quad (16)$$

We consider:

$$\frac{\partial f}{\partial m_L} \cdot dm_L + \frac{\partial g}{\partial m_G} \cdot dm_G = 1 - \alpha_2 \cdot \frac{m_L}{m_G} \quad (17)$$

And

$$\frac{\partial g}{\partial \sigma_G} \cdot d\sigma_G + \frac{\partial f}{\partial \sigma_L} \cdot d\sigma_L = \alpha_1 \cdot \left(1 - \frac{\sigma_L}{\sigma_G}\right) + \alpha_2 \cdot \frac{\sigma_L}{\sigma_G} \quad (18)$$

We obtain the following formula:

$$T(x, y) = m_L \cdot \left[1 - \alpha_2 \cdot \frac{m_L}{m_G} + \alpha_1 \cdot \left(1 - \frac{\sigma_L}{\sigma_G}\right) + \alpha_2 \cdot \frac{\sigma_L}{\sigma_G}\right] \quad (19)$$

Where:

$$T(x, y) = m_L \cdot \left[1 + \alpha_1 \cdot \left(1 - \frac{\sigma_L}{\sigma_G}\right) - \alpha_2 \cdot \left(\frac{m_L}{m_G} - \frac{\sigma_L}{\sigma_G}\right)\right] \quad (20)$$

Where α_1 , α_2 parameters ($\alpha_1 = 0.99$ and $\alpha_2 = 0.08$, obtained empirically), m_L and σ_L are the local mean and local standard deviation, m_G and σ_G are the global mean and global standard deviation calculated in w×w window centered at pixel (x,y).

3.2 Contribution of each term

We split the formula 20 to several terms, to show the contribution of each one in the behavior of the proposed thresholding formula, noting the following terms:

$$A = m_L \tag{21}$$

$$B = m_L \cdot \alpha_1 \cdot \left(1 - \frac{\sigma_L}{\sigma_G}\right) \tag{22}$$

$$C = m_L \cdot \alpha_2 \cdot \left(\frac{m_L}{m_G} + \frac{\sigma_L}{\sigma_G}\right) \tag{23}$$

The results are illustrated in the following graph which shows the lines 90 and 210 of an image with 250 pixels.

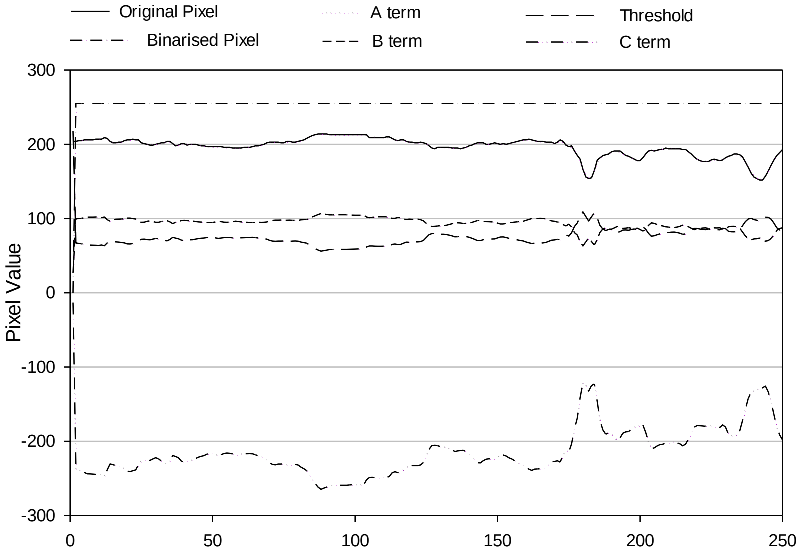

Figure 1. The contribution of each term in our binarization formula in line 90

In the graph illustrated in Figure 1, we see that the contribution of term A (mean value) has the same effect as the original pixel values, the impact of the term B is correlated with the evolution of the average value and the term C allows to balance the cumulative effect of the terms A and B.

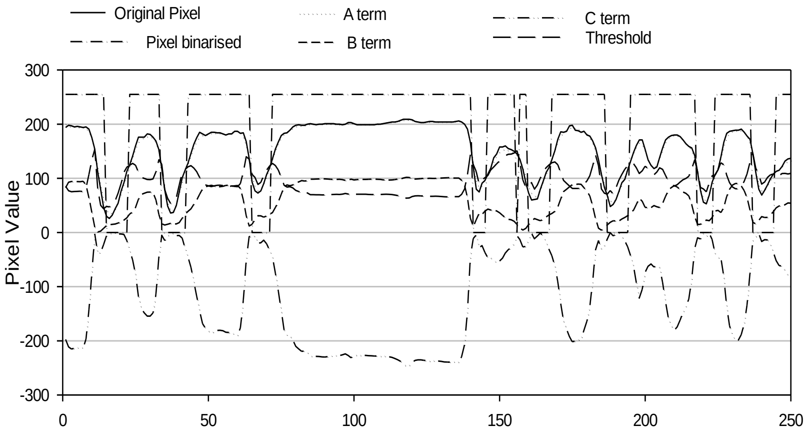

Figure 2. The contribution of each term in our binarization formula in line 210

In the graph illustrated in Figure 2, the term A (average value) coincides with the original pixel values. The term B

Evolves proportionally with the original values and at last the term C fulfills a compensatory role of the accumulation effect for the terms A and B.

These two graphs allow us to deduce that the term A is an average that recalls the existing global thresholding bases in our proposed formula, while the effect of B is moderator for integrating the impact of local standard deviation ratio to the global standard deviation, this ratio added to the ratio of global and local averages in the compensator term (C), limits the effect of the term A and/or the term B, thereby reducing the impact of pixel parasites which belongs neither to the object (text) nor to the background.

3.3 Quantitative evaluation by F-measure

In this experiment, we used F-measure to compare the performance of binarization algorithms. This criterion is utilized in the set of document images obtained from binarization contest [21]. It is defined as follows:

$$F\_measure = \frac{2 \cdot \mathrm{Re}\,call \cdot \mathrm{Pr}\,ecision}{\mathrm{Re}\,call + \mathrm{Pr}\,ecision} \qquad (24)$$

Where

$$\mathrm{Re}\,call = \frac{TP}{TP + FN} \qquad (25)$$

And

$$\mathrm{Pr}\,ecision = \frac{TP}{TP + FP} \qquad (26)$$

and TP, FP and FN denote true-positive, false-positive, and false-negative values, respectively.

4 Experimental results and discussion

The proposed binarization algorithm was tested with the benchmarking techniques and various scenarios against several known adaptive binarization techniques, such as Sauvola's [10] and Image Sum Integral [20], in addition to a hybrid method of global and local adaptive binarization [12]. Using a set of handwritten and printed images issued from DIBCO2009 [22].

The two categories of images (printed and handwritten) used to test our method had varying resolutions, sizes and contrasts to ensure a fair comparison of performances between our method and other methods.

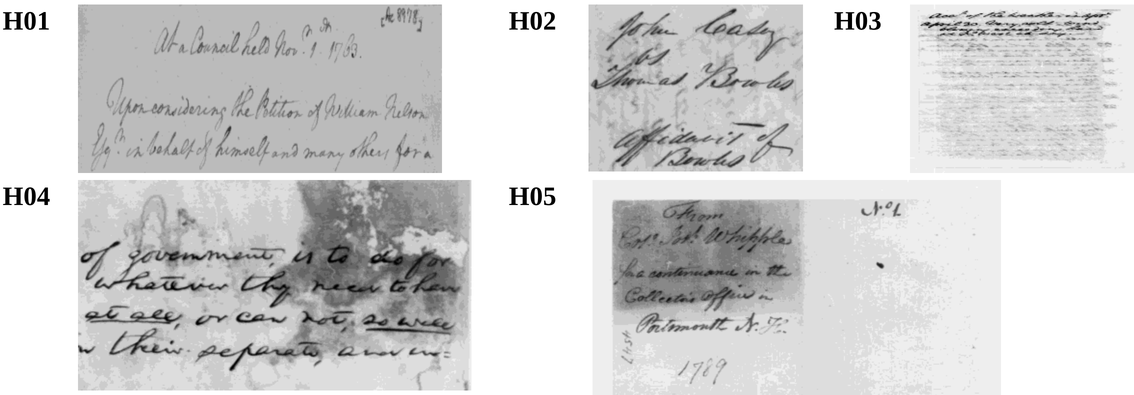

Figure 3. Sample of our set of handwritten documents (illuminated sample document + blotches and smudges)

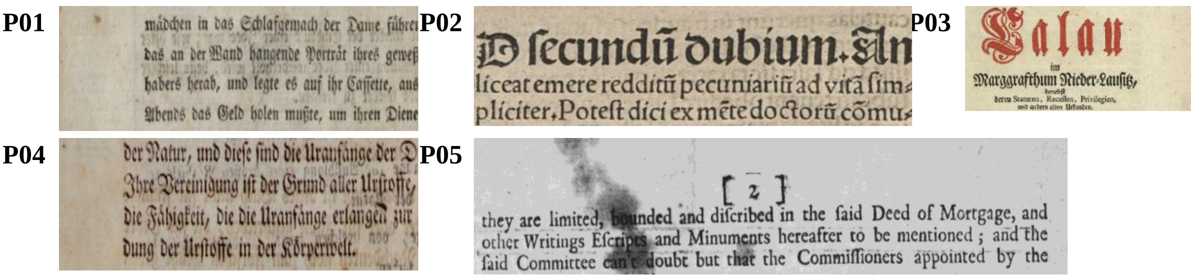

Figure 4. Sample of our set of printed documents (illuminated sample document + blotches and smudges).

Some results of our thresholding method are shown in Figure5 and in Figure 6.

Figure 5. Examples of results of our method compared to other threshold methods described in handwritten documents

Figure 6. Examples of results of our method compared to other threshold methods described in printed documents

Detailed performance of F_Measure for the predefined set of images are presented in Table 1, 2.

Table 1. Detailed performance of the F_Measure results for a sample of our set of handwritten images thresholded.

	H01	H02	H03	H04	H05	Mean
Image integral	90,325	52,939	85,609	67,730	69,484	**73,217**
Sauvola	84,742	57,237	84,199	55,788	61,967	**68,786**
Bataineh	89,243	57,477	88,038	87,874	78,104	**80.147**
Our Method	84,972	57,822	89,387	88,520	84,022	**80,945**

Table 2. Detailed performance of the F_Measure results for a sample of our set of printed images thresholded.

	P01	P02	P03	P04	P05	Mean
Image integral	90,215	92,315	95,245	83,216	87,222	**89,643**
Sauvola	85,654	93,077	95,157	77,704	83,834	**87,085**
Bataineh	87,645	92,462	95,967	86,087	89,404	**90,313**
Our Method	98,713	94,669	96,248	88,606	87,599	**93,167**

The proposed method is compared with other binarization techniques (Image Sum Integral, Sauvola's and Bataineh's methods) on many categories of images (handwritten and printed documents).

The proposed method has shown good performances in all type of images; it's particularly effective at removing blotches and smudges in the image yet still maintaining the handwritten or printed details.

The numerical test and results presented were gained using binarization metrics emphasizing the performance in textual image region binarization. Figures 5 and 6 present an example of benchmarking performed to a set of 10 textual document images having illumination and non-uniform contrast.

The background elimination technique generally achieved better precision and recall in all categories of our set of images. Our approach has given interesting precision values for both types of images, comparing them with the previously mentioned methods.

Our technique obviously outperforms the three comparison methods as illustrated in some images in table 1 and 2. Comparing the indicator F_measure, highlights this performance.

5 Conclusion

This paper demonstrated the effectiveness of combining two spatial components in adaptive binarization (local and global parameters of spatial distribution) for binaraizing in ancient document images.

Compared to methods like Integral Sum Image, Sauvola's and Bataineh's, the proposed method was even able to detect low contrast characters when other methods appeared to be inefficient. Furthermore, thanks to the consideration of local and global spatial context, we obtain lower noise within the background and within the text regions.

The efficiency of the proposed method was demonstrated by its implementation on a set of ancient documents that comprised historical document scans used in this paper as examples. A recognition rate is determined by the parameter F_Measure and in our case we have an average of 80.945% which was achieved using the first document set (handwritten images), and a 93.167% which was achieved using the second set (printed images).

A comparison was also drawn between our method and three other known and efficient thresholding methods, namely Integral Sum Image, Sauvola's and Bataineh. The comparison results show that our method is in most cases superior to other methods, thus it can be efficiently used for enhancing scanned documents as part of a document analysis scheme.

The improvement provided in terms of binarization can positively influence recognition in the following stages of ancient document analysis.

REFERENCES

[1] Likforman-Sulem L, Zahour A, and Taconet B. Text line segmentation of historical documents: a survey 2007; International Journal of Document Analysis and Recognition (IJDAR), 9(2-4): 123–138. doi : 10.1007/s10032-006-0023-z

[2] Shukle S, Sonawane A, Topale V, Tiwari P: Improving Degraded Document Images Using Binarization Technique 2014; International Journal Scientific & Technology Research, **3(**5): 333-338. doi : 10.15680/IJIRCCE.2016. 0408117

[3] Kavallieratou E, and Stathis. S. Adaptive Binarization of Historical Document Images. In : 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, China, August 2006, pp 742-745.

[4] Mat Som H, Mohamad Zain J, and Jihadi Ghazali A, Application of threshold techniques for readability improvement of Jawi historical manuscript images. 2011; Advanced Computing: An International Journal (ACIJ), **2(**2): 60-69. doi : 10.5121/acij.2011.2206

[5] Su B, Lu S, Tan C. L. Robust document image binarization technique for degraded document images 2013; IEEE Trans. Image Process. **22(**4): 1408–1417. doi :10.1109/TIP.2012.2231089

[6] Trier O.D. and Taxt T. Evaluation of Binarization Methods for Document Images 1995; IEEE Transaction on Pattern Analysis and Machine Intelligence, **17(**3): 31-35. Doi : 10.1109/34.368197

[7] Niblack W. An Introduction to Digital Image Processing. Englewood Cliffs, NJ: Prentice Hall 1986.

[8] Sauvola J. and Pietikainen M. Adaptive Document Image Binarization 2000, The Journal of the Pattern Recognition Society, **33(**2): 225-236. doi : 10.1016/S0031-3203(99)00055-2

[9] Sezgin M, Sankur B. Survey over image thresholding techniques and quantitative performance evaluation 2004, J. Electron. Imaging, **13(**1): 146–165. doi :10.1117/1.1631315

[10] Haritaoglu, I. Scene Text Extraction and Translation for Handheld Devices, In : IEEE Computer Vision and Pattern Recognition Conference, Kauai, HI, USA, December 2001, pp. 408-413.

[11] Bataineh B, Norul Huda S, Abdullah S, Omar K. An adaptive local binarization method for document images based on a novel thresholding method and dynamic windows 2011; Pattern Recognition Letters, **32(**14): 1805-1813. doi : 10.1016/j.patrec.2011.08.001

[12] Saini R. Document Image Binarization Techniques, Developments and Related Issues 2015; A Review International Journal of Computer Applications, **116(**7): 41 – 44. doi : 10.5120/20352-2541

[13] Halabi Y. S, Sa''Sa Z, Hamdan F, Haj Yousef K. Modeling Adaptive Degraded Document Image Binarization and Optical Character System 2009; European Journal of Scientific Research, **28(**1): 14 – 32. issn 1450-216X

[14] Huang, Q, Gao W, Cai W. Thresholding technique with adaptive window selection for uneven lighting image 2005; Pattern Recognition Letters, **26:** 801–808. doi :10.1016/j.patrec.2004.09.035

[15] Feng M. and Tan Y. Contrast adaptive binarization of low quality document images 2004; IEICE Electron. Express, **1(**16): 501-506. doi : 10.1587/elex.1.501

[16] Asirvatham J. and Agnes A. Survey on Restoration of Degraded Document using Different Thresholding Techniques 201;, International Journal of Engineering Sciences & Research Technology, **3(**4): 1868-1871. issn (Online): 1694-0814

[17] Feng M. and Tan Y. Adaptive Binarization Method for Document Image Analysis In : IEEE International Conference on Multimedia and Expo (ICME) Taipei, Taiwan 27-30 June 2004, pp. 339-342.

[18] Wolf C. and Jolion J.-M. Extraction and recognition of artificial text in multimedia documents 2004; Patt. Anal. Appl., 2004, **6(**4): 309-326. doi : 10.1007/s10044-003-0197-7

[19] Romen Singh T, Roy S, Imocha Singh O, Sinam T, Manglem Singh Kh. A New Local Adaptive Thresholding Technique in Binarization 2011; IJCSI International Journal of computer Sciences Issues, **8**, 6(2): 271-277. issn (Online): 1694-0814

[20] Viola P. and Jones M. J. Robust real-time face detection 2004; Int. Journal of Computer Vision, **57(**2): 137–154. doi : 10.1023/B:VISI.0000013087.49260.fb

[21] Crow, F. C. Summed-area tables for texture mapping - Computer Graphics – In : SIGGRAPH'84, Minneapolis, Minnesota USA, July 1984, **18(**3), pp. 207–212.

[22] Gatos, B, Ntirogiannis K, Pratikakis, I. ICDAR 2009 document image binarization contest (DIBCO 2009). In : 10th International Conference on Document Analysis and Recognition, Spain, 2009, pp. 1375–1382.