



HAL
open science

Local Thresholding Algorithm Based on Variable Window Size Statistics

Costin-Anton Boianiu, Alexandra Olteanu, Alexandru Stefanescu, Daniel Rosner, Nicolae Tapus, Mugurel Ionut Andreica

► **To cite this version:**

Costin-Anton Boianiu, Alexandra Olteanu, Alexandru Stefanescu, Daniel Rosner, Nicolae Tapus, et al.. Local Thresholding Algorithm Based on Variable Window Size Statistics. Proceedings of the 18th International Conference on Control Systems and Computer Science (CSCS) (ISSN: 2066-4451), May 2011, Bucharest, Romania. pp.647-652. hal-00767478

HAL Id: hal-00767478

<https://hal.science/hal-00767478>

Submitted on 20 Dec 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Local Thresholding Algorithm Based on Variable Window Size Statistics

Costin-Anton Boiangiu, Alexandra Olteanu,
Alexandru Stefanescu, Daniel Rosner, Nicolae Tapus,
Mugurel Andreica

* *Computer Science Department, Faculty of Automatic Control and Computers, University Politehnica of Bucharest, Splaiul Independentei 313, Sector 6, 060042 Bucharest, Romania (e-mails: costin.boiangiu@cs.pub.ro, alexandra.olteanu@cti.pub.ro, alexandru.stefanescu1708@cti.pub.ro, nicolae.tapus@cs.pub.ro, daniel.rosner@cs.pub.ro)*

Abstract: In an automatic document conversion system, which builds digital documents from scanned articles, there is a need to perform various adjustments before the scanned image is fed to the layout analysis system. This is because the layout detection system is sensitive to errors when the page elements are not properly identified, represented, denoised, etc. Such an adjustment is the detection of foreground and background of a document or simply called a document image binarization. This paper presents a new idea for treating the common problems which may occur during the binarization phase of the documents, that considers a parameter-free local binarization algorithm which dynamically computes the window size after it sets a threshold for the standard variation value of the window. This proved to offer more consistent results for a wide variety of scanned documents consisting of various old newspapers and old library documents in different languages, both handwritten and textual documents. Two methods for computing the local gray level threshold are proposed using the mean and standard deviation of the pixels in the current window:

Keywords: Binarization, Local Thresholding, Variable Window, Niblack, Otsu

1. INTRODUCTION

In recent years, the problem of converting scanned documents into electronic files, especially for large electronic libraries, became more and more studied since the number of degradation categories manifested in the documents to be processed is significant. An automatic content conversion systems, based on optical character recognition (OCR), enable operations such as editing, word searching, easy document storing and multiplication, and the application of a large set of text techniques including text-to-speech and text mining to be performed on the digitalized document. In addition, this ensures both a better preservation of original documents, due to minimizing the need for physical use and makes it suitable for automatic data processing or usage under a large spectrum of devices including mobile phones or other mobile devices.

Content conversion systems involve a number of basic steps to be performed to obtain an output in which we can classify the scanned document elements correctly. These steps are represented by: image binarization, which classifies it in foreground and background, skew detection

* The work presented in this paper was funded by CNCIS-UEFISCDI under research grant *PD.240/2010* (contract no. 33/28.07.2010) and by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labour, Family and Social Protection through the financial agreement POSDRU/89/1.5/S/62557.

and correction, noise removal, classification into textual and non-textual regions, text detection, segmentation into words and characters, character recognition, and a post processing step for correcting the resulting text file in terms of lexical and semantic perspective.

Document image binarization is the processing step at the base of every content conversion system and requires maximum quality for the output, since it affects all subsequent processing steps. A large number of methods have been proposed in the literature Mehmet and Bulent (2004); Yahia S. Halabi and Yousef (2009) but none delivers the best results for all types of input documents. Therefore, this remains an open and active area for research.

The main problems with existing binarization methods are:

- Failure to work on images with varying levels of illumination due to the use of a global gray level threshold for the entire image.
- Dependence on parameters which must be fine tuned manually to produce optimum results.

The proposed approach is at the same time *parameter-free and able to handle well many types of image degradation*, producing comparable results to classic and current top performing algorithms.

The next section presents an analysis of some related work in the field. The description of the algorithm can be found in section 3, while section 4 is composed of the experimental results and a short comparison between the proposed method and some traditional binarization methods. Finally, conclusions and future work perspectives are mentioned in section 5.

2. RELATED WORK

Existing binarization techniques are in most cases thresholding related. According to the most common classification, image thresholding methods can be performed either globally as in Otsu (Jan. 1979) or locally, Gatos et al. (2006); Boiangiu et al. (August 27-29, 2009); Boiangiu and Dvornic (July 2008). In the case of global binarization methods, a single calculated threshold value is used to classify the image pixels into object/foreground (black) or background (white), while for the local methods the local area information is used for determining a threshold value for each pixel. However, similar variations of the threshold value produces dramatic changes in the output in some cases, and only minor differences in others, depending on the number of objects in the image or local region, and on the difference of weight values of the pixels that represent these objects. Consequently, techniques of both categories mentioned above perform poorly under very complex backgrounds.

Global binarization methods, such as Zhang and Hu (2008), offer good results only when there is a clear separation between the foreground and the background; otherwise, an adaptive method is required. Otsu's algorithm introduced by Otsu (Jan. 1979) represents one of best global techniques and can achieve high performance in terms of correctness of segmentation boundaries and uniformity of the threshold zones. The algorithm finds the threshold that minimizes the weighted within-class variance, which is equivalent to maximizing the between-class variance, using the gray level histogram. In this approach, local statistic variations are actually suppressed or lost, even if they could bring an important contribution, due to the information contained by them, to pixel classification. To overcome the aforementioned problems local adaptive methods are used. Niblack's method, W.Niblack (1986), is a well known local threshold binarization method which involves computing for each pixel in the grayscale image the mean and standard deviation of the colors of the neighboring pixels in an area (window) of predefined size. The size of the neighborhood should be small enough to preserve local details and large enough to suppress noise. The threshold for determining if the pixel is converted to black or white is computed using the formula:

$$T = mean + k * stdev \quad (1)$$

where *mean* is the average value of the pixels in the local area, *stdev* the standard deviation of the same pixels and *k* is a constant, preselected coefficient. Drawbacks of this method include a considerable sensitivity to the window size and the persistence of background noise in the binary image. An improved version was proposed by Sauvola, presented in Gatos et al. (2004); Kasar et al. (2007); He et al. (2005), to reduce the amount of background noise in homogeneous regions larger than the window size. This

method introduces a hypothesis on the gray values of text and background pixels such that object pixel values are close to 0 and background pixels to 255, and is described by the formula:

$$T = mean * (1 + k(\frac{stdev}{R} - 1)) \quad (2)$$

where parameters *R*, the dynamic range of the standard deviation, and *k* are fixed to 128 and 0.5, respectively. In case of high contrast image regions, the standard deviation value is approximately equal with *R* and this suggests that the computed threshold could be approximated to the mean color value of the window, producing similar results as through Niblack's method. On the other hand, the differences are more obvious when the contrast in the local neighborhood is relatively low. In this situation the local threshold value is below the mean value and thus darker regions are successfully removed from the background.

The well known methods mentioned above are considerably old and since their first publication new methods have been proposed. We have confronted our approach with the methods claiming the top two places at ICDAR 2009 Document Image Binarization Contest, Gatos et al. (2009, 13 May 2010).

The winner algorithm was proposed by S. Lu and C.L. Tan, see Gatos et al. (13 May 2010), and includes four parts: document background extraction, stroke edge detection, local thresholding, and post-processing. They estimate the local threshold by averaging the detected edge pixels within a local neighborhood window.

J. Fabrizio and B. Marcotegui, Gatos et al. (13 May 2010); Fabrizio et al. (2009), proposed the second best rated algorithm, which is based on the so called toggle mapping operator. Corresponding morphological erosion and dilation masks are applied on the image and then each pixel is marked as background if the pixel value is closer to the erosion or as foreground otherwise. Certain pixels are removed to avoid noise and the remaining are reclassified into foreground, background and homogeneous regions, the latter being reassigned according to the class of the boundaries.

A lot of research areas currently cover the domain of image binarization: global thresholding, local thresholding, segmentation and classification, etc. The problem is that none of these areas have developed fully-grown algorithms that are able to deal successfully with all of the aspects involved in the plethora of scanned image documents. There have been developed mathematically bulletproof methods such as Otsu (Jan. 1979) for computing global thresholds, blur-based local thresholds Boiangiu and Dvornic (July 2008), divide-et-impera category approaches like region-thresholding Boiangiu et al. (August 27-29, 2009), but none of them can successfully tackle both noise and variable lighting in the same time. Well known algorithms like Sauvola and Niblack are very rigid in their approach about computing local statistical functions because they are operating on fixed (and image-dependant, not easy to compute) window size. The current approach tries to be a bridge between two worlds by taking advantage on both local statistics and variable window size for local threshold computation.

To deal with the aforementioned issues the algorithm described in the next section keeps the threshold equal with the mean, which is computed for different window sizes for each pixel, such that the standard deviation for the pixels in the current window respects a precomputed threshold. To obtain a parameter-free algorithm, we propose a modification to Niblack's method, in which instead of using predefined values, the window size is computed dynamically.

3. ALGORITHM

The starting point of the proposed binarization algorithm is represented by Niblack's method, the algorithm being based on sliding a rectangular window over the document image, calculating the mean and standard deviation of the gray pixel values in each window. However, instead of using predefined values, the window size is computed by gradually growing it until the value of the standard deviation of gray pixel levels within the window multiplied by the logarithm of the window size reaches the first local maximum or reaches the image height or width, as seen in Figure 1, obtaining in this way a parameter-free algorithm. In this way the most appropriate windows size is selected in order to preserve local proprieties and at the same time to suppress noise. In addition, this is also a good compromise between the quality of the output and the speed of conversion.

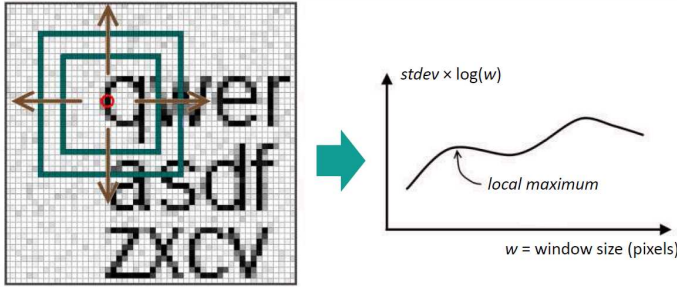


Fig. 1. Expanding the window until first local maximum is reached. The graph represents the dependency between the standard deviation values multiplied by the logarithmic of the window size and the window size

However, we started with a simple version of this approach in which the window size is computed dynamically after setting a threshold for the standard variation value. The window size is increased until the standard deviation of the gray pixel values contained in the window exceeds this threshold or reaches the image height or width. In this version, the threshold T for determining if the pixel is mapped to black or white is computed using the formula:

$$T = mean * (1 + t) \quad (3)$$

where t is a preselected coefficient. The *mean* value is dependent on the window's size, and consequently to the standard deviation threshold. This indicates that choosing an appropriate value for the standard deviation threshold will lead to better results. Equation (3) highlights the fact that because of a fixed threshold for standard deviation, the binarization threshold should be equal to the window's mean gray level. However, the preselected coefficient t

plays a significant role. It should be chosen negative when the window size imposed by the standard deviation threshold is small, in order to suppress noise. When the computed window size is larger, a positive coefficient t should be used for preserving local details. When selecting this threshold the standard deviation value for the entire image should be considered. In the simplest approach, the standard deviation threshold for the current window is computed as the standard deviation value for the entire image weighted by a preset coefficient w :

$$T_{stdev} = w * stdev_{image} \quad (4)$$

where T_{stdev} is the standard deviation threshold and $stdev_{image}$ is the standard deviation of the entire image.

Thus, the main algorithm steps would be:

- compute the threshold for standard deviation using (4)
- for each image pixel do
 - grow the window size until the standard deviation of the current window surpasses the threshold
 - determine the mean of the current window and then compute the local threshold using (3)
 - set the pixel to 0 (black) if lower than obtained threshold, or to 255 (white), otherwise.

Yet, we want to obtain a parameter-free algorithm with a clear advantage over the others, by not requiring parameter tuning. In a high volume content conversion workflow, this can be extremely important as it can lead to significant improvements of the total processing time, eliminating the need for adjustment at every batch of documents.

Taking this aim into account, we redesign the algorithm as follow:

- For each image pixel, compute the sum and the squared sum of the gray levels of pixels contained in a rectangular area defined by the current pixel and the pixel in the image top-left corner;
- For each image pixel do:
 - grow the window size until the standard deviation of the current window multiplied by the logarithm of the window size is smaller than the value computed for the previous window;
 - determine the mean of the current window and then compute the local threshold using (5) or (6);
 - set the pixel to 0 (black) if lower than the obtained threshold, or to 255 (white), otherwise.

We propose two methods for computing the local gray level threshold using the mean and standard deviation of the pixels in the current window:

$$T = mean \quad (5)$$

$$T = \frac{\min(stdev, stdAvg) \times (mean - T_{OTSU})}{stdAvg} + T_{OTSU} \quad (6)$$

where T_{OTSU} is the global Otsu threshold computed as in Otsu (Jan. 1979) and $stdAvg$ is a constant equal to 64 (half the value of the maximum standard deviation).

Basically, this approach considers that for a standard deviation value, determined for a window that is centered on each image pixel and limited to a fixed threshold, the neighborhood defined by the window offers adequate local

statistics so that the most appropriate threshold value for the center pixel can be determined based on the mean gray level values in the neighborhood.

Note that for images having large areas of identical color levels (e.g. synthetic images - computer generated drawings), when analyzing pixels belonging to these regions the window may grow to larger sizes than for document scans to meet the required threshold, since the standard deviation of the pixel colors in windows contained in such regions would be constant 0. As result, binarization of this type of images will require longer processing times. However, since content conversion systems generally work with document scans, such monochromatic areas rarely occur due to the inherent imperfections in nuance of paper and ink.

In the next section we will present some visual examples of the results obtained and we will discuss different test cases that these examples represent.

4. VISUAL EXAMPLES AND DISCUSSION

Binarization tests have been performed on scanned documents consisting of various old newspapers and old library documents in different languages, both handwritten and textual documents. The input datasets used for our evaluation was considered relevant for the measurement of quality of the proposed method because most of the images contained were suffering from problems like inconsistent lighting across the whole page, poor contrast and noisy aspect. The two variants of the proposed method have been compared against the methods claiming the top two places at ICDAR 2009 Document Image Binarization Contest and the well known methods of Niblack (15x15 window size and $k = -0.2$) and Otsu. The required parameters for each method have been kept fixed in all tests.

The Figures in this section present 9 test cases for images which are converted to black and white using a set of well-known algorithms from scientific literature (Niblack W. Niblack (1986), Otsu Otsu (Jan. 1979)) which have been previously mentioned, the best binarization algorithms (DIBCO Gatos et al. (2009, 13 May 2010)), as well as with the proposed technique using both methods for computing the local gray level threshold.

————— de aici trebuie readaptat textul —————

The first case is considered representative for historical books containing Fraktur-style fonts, with a lot of background noise, high density text and presenting difficulties due to uneven exposure during document acquisition, highlighted by the Figure 4.

The proposed approach managed to correctly classify the foreground pixels without misclassifying the noise pixels, offering a result very similar to the one corresponding to Sauvola's method, which delivered the best bitonal conversion for the first test case.

The second case, shown in Figure 3, is for historical newspapers which present large areas having uneven contrast and lighting, Antiqua-style fonts, medium to low text density and blurry textual regions.

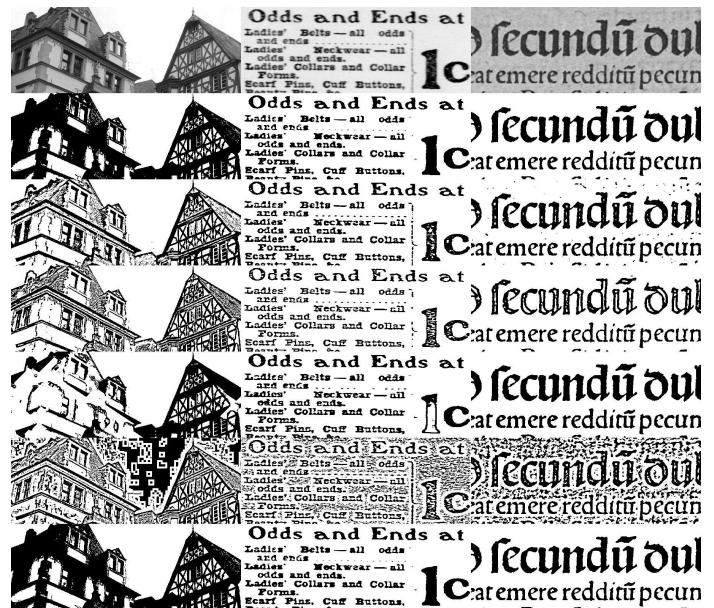


Fig. 2. Test set 1.: 3 test cases arranged in 3 columns - images and large text fonts. Each line represents in the same order the following: line 1. the original grayscale image; Then results obtained on the original image with: line 2. The proposed Adaptive Otsu approach; line 3. The proposed Adaptive approach; line 4. Lu and Tan's method; line 5. J. Fabrizio and B. Marcotegui's method; line 6. Niblack's method; line 7. Otsu's method

————— pana aici —————

The proposed approach was able to deal in a successful manner with the differences in ink color and with the variable lighting across the page, also preserving the original character edges which are extremely useful for a correct OCR reading. In addition, while Sauvola's method offered the worst result for the second case, the proposed approach succeeded to obtain one of the best results. This demonstrates that, even if some methods obtain slightly better results for some cases, as it was shown in the first set of images, the latter provided more consistent results offering a good classification of pixels for a larger number of types of degradation manifested in the documents.

One of the advantages of using an algorithm capable of dealing with more types of degradation or document problems is saving costs by requiring less human intervention in the binarization process. Although the proposed method may not produce the best bitonal conversion in all cases, it has a clear advantage over the others by not requiring parameter tuning. In a high volume content conversion workflow, this leads to significant improvements of the total processing time, eliminating the need for adjustment at every batch of documents.

5. CONCLUSIONS

A good document image binarization algorithm is vital to any document digitization system. This paper explores the possibilities of variable-window statistics in order to implement such a binarization algorithm. The proposed method is a very important tool for image color conversion, mainly

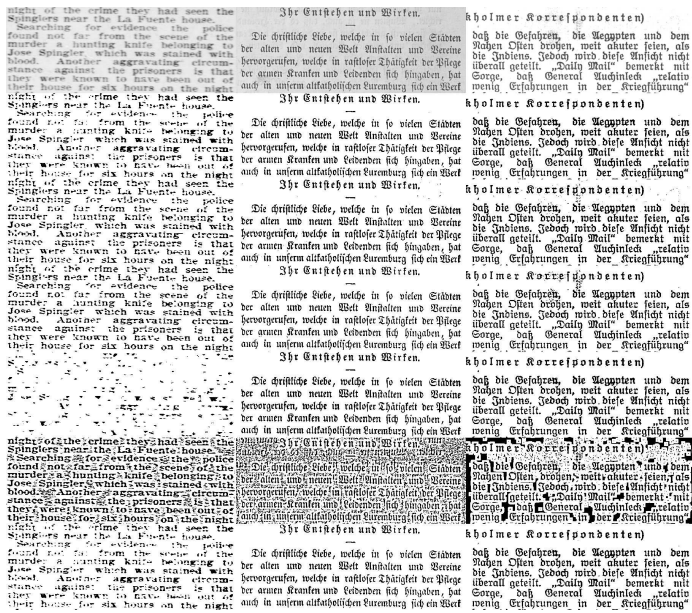


Fig. 3. Test set 2.: 3 test cases arranged in 3 columns - old news papers. Each line represents in the same order the following: line 1. the original grayscale image; Then results obtained on the original image with: line 2. The proposed Adaptive Otsu approach; line 3. The proposed Adaptive approach; line 4. Lu and Tan's method; line 5. J. Fabrizio and B. Marcotegui's method; line 6. Niblack's method; line 7. Otsu's method

because of its no need to parameter tuning, avoiding the necessitate for adjustment at every batch of documents, and its resistance to noise and imperfections, allowing the subsequent stages of a retroconversion system to take benefit of a clean foreground/background classification for image data.

The comparison between the results obtained using well known image binarization algorithms, both local and global, and the proposed algorithm on the selected test images has shown that the latter provides nearly best results in the majority of cases, proving its superior performance in a larger number of situations than any other algorithm that has been used for output comparison.

The main idea for developing the proposed algorithm was that in order to have a successful binarization approach, a local type of algorithm is needed because the image documents are not consistent in respect to contrast and illumination across the entire page. The next observation is that for a successful local approach, a variable neighborhood should be used because documents may contain a lower or greater density of information in different locations. The current approach tries to compute the most appropriate neighborhood size based on the local variance and proposes the neighborhood to have a square shape centered in the currently processed point. This will work in general but it is not perfect, since a close-to-perfect vicinity will need at least a size which varies according to the degree of the disorder in the image as two-dimensional vicinity and not as a set of one-dimensional measurements, as provided by the (histogram-based) variance. However, this works as an acceptable compromise between quality

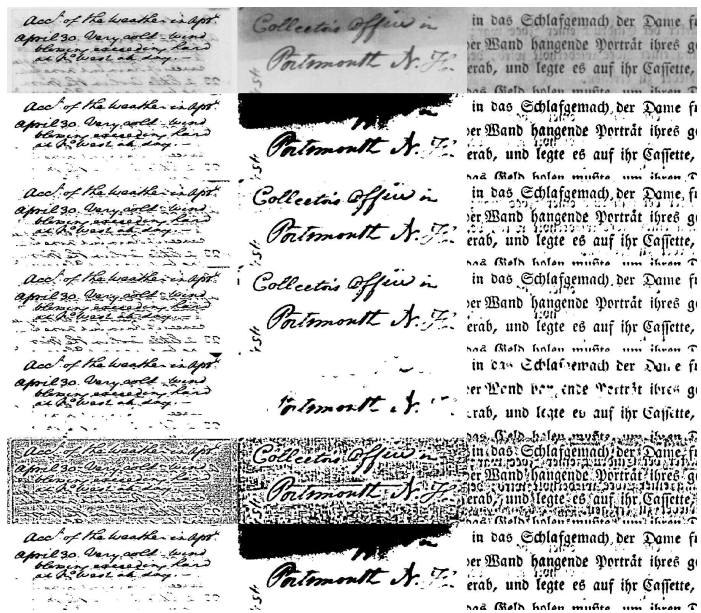


Fig. 4. Test set 3.: 3 test cases arranged in 3 columns - extremely noisy background. Each line represents in the same order the following: line 1. the original grayscale image; Then results obtained on the original image with: line 2. The proposed Adaptive Otsu approach; line 3. The proposed Adaptive approach; line 4. Lu and Tan's method; line 5. J. Fabrizio and B. Marcotegui's method; line 6. Niblack's method; line 7. Otsu's method

of the results and the speed of the conversion which is also an important factor. In the future there will be developed a more flexible method for determining the neighborhood which will vary locally in size in an intelligent manner, based on a more accurate measurement of the local two-dimensional information density in the image document.

ACKNOWLEDGEMENTS

The work presented in this paper was funded by CNCIS-UEFISCDI under research grant PD_240/2010 (contract no. 33/28.07.2010) and by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labour, Family and Social Protection through the financial agreement POS-DRU/89/1.5/S/62557.

REFERENCES

Boiangiu, C.A. and Dvornic, A.I. (July 2008). Methods of bitonal image conversion for modern and classic documents. In *WSEAS Transactions on Computers*, volume 7, 10811090. Cluj-Napoca, Romania.

Boiangiu, C.A., Dvornic, A.I., and Cananau, D.C. (August 27-29, 2009). Binarization for digitization projects using hybrid foreground-reconstruction. In *Proceedings of the 2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing*, 141-144. Cluj-Napoca, Romania.

Fabrizio, J., Marcotegui, B., and Cord, M. (2009). Text segmentation in natural scenes using toggle-mapping. In *Image Processing (ICIP), 2009 16th*

- IEEE International Conference on*, 2373–2376. doi: 10.1109/ICIP.2009.5413435.
- Gatos, B., Pratikakis, I., and J.Perantonis, S. (2006). Adaptive degraded document image binarization. volume 39, 317–327. Elsevier Science Inc., New York, NY, USA. doi: <http://dx.doi.org/10.1016/j.patcog.2005.09.010>.
- Gatos, B., Pratikakis, I., and Perantonis, S.J. (2004). An adaptive binarization technique for low quality historical documents. In *Lecture Notes in Computer Science*, volume 3163 of *Document Analysis Systems VI*, 102–113. Springer, Berlin.
- Gatos, B., Ntirogiannis, K., and Pratikakis, I. (13 May 2010). Dibco 2009: document image binarization contest. In *International Journal on Document Analysis and Recognition (IJDAR)*, ICDAR '09.
- Gatos, B., Ntirogiannis, K., and Pratikakis, I. (2009). Icdar 2009 document image binarization contest (dibco 2009). In *Proceedings of the 2009 10th International Conference on Document Analysis and Recognition*, ICDAR '09, 1375–1382. IEEE Computer Society, Washington, DC, USA. doi: <http://dx.doi.org/10.1109/ICDAR.2009.246>. URL <http://dx.doi.org/10.1109/ICDAR.2009.246>.
- He, J., Do, Q.D.M., Downton, A.C., and Kim, J.H. (2005). A comparison of binarization methods for historical archive documents. In *ICDAR '05: Proceedings of the Eighth International Conference on Document Analysis and Recognition*, 538–542. IEEE Computer Society, Washington, DC, USA. doi: <http://dx.doi.org/10.1109/ICDAR.2005.3>.
- Kasar, T., Kumar, J., and Ramakrishnan, A.G. (2007). Font and background color independent text binarization. In *Second International Workshop on Camera-Based Document Analysis and Recognition (CBDAR2007)*. Bangalore, INDIA.
- Mehmet, S. and Bulent, S. (2004). Survey over image thresholding techniques and quantitative performance evaluation. 13, 146–165. doi:doi:10.1117/1.1631315.
- Otsu, N. (Jan. 1979). A threshold selection method from gray-level histograms. volume 9, 62 – 66. Systems, Man and Cybernetics, IEEE Transactions on. doi: 10.1109/TSMC.1979.4310076.
- W.Niblack (1986). An introduction to digital image processing. 115–116. Prentice-Hall, Englewood Cliffs, NJ.
- Yahia S. Halabi, Zaid SA SA, F.H. and Yousef, K.H. (2009). Modeling adaptive degraded document image binarization and optical character system. volume 28, 14–32. European Journal of Scientific Research.
- Zhang, J. and Hu, J. (2008). Image segmentation based on 2d otsu method with histogram analysis. In *CSSE '08: Proceedings of the 2008 International Conference on Computer Science and Software Engineering*, 105–108. IEEE Computer Society, Washington, DC, USA. doi:<http://dx.doi.org/10.1109/CSSE.2008.206>.