Available online at www.sciencedirect.com
SCIENCE DIRECT®

Pattern Recognition 39 (2006) 317–327

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

# Adaptive degraded document image binarization

B. Gatos*, I. Pratikakis, S.J. Perantonis

*Computational Intelligence Laboratory, Institute of Informatics and Telecommunications, National Center for Scientific Research "Demokritos", 153 10 Athens, Greece*

Received 11 January 2005; accepted 9 September 2005

**Abstract**

This paper presents a new adaptive approach for the binarization and enhancement of degraded documents. The proposed method does not require any parameter tuning by the user and can deal with degradations which occur due to shadows, non-uniform illumination, low contrast, large signal-dependent noise, smear and strain. We follow several distinct steps: a pre-processing procedure using a low-pass Wiener filter, a rough estimation of foreground regions, a background surface calculation by interpolating neighboring background intensities, a thresholding by combining the calculated background surface with the original image while incorporating image up-sampling and finally a post-processing step in order to improve the quality of text regions and preserve stroke connectivity. After extensive experiments, our method demonstrated superior performance against four (4) well-known techniques on numerous degraded document images.


*Keywords:* Degraded document images; Local adaptive binarization

## 1. Introduction

Document image binarization (threshold selection) refers to the conversion of a gray-scale image into a binary image. It is the initial step of most document image analysis and understanding systems. Usually, it distinguishes text areas from background areas, so it is used as a text locating technique. Binarization plays a key role in document processing since its performance affects quite critically the degree of success in a subsequent character segmentation and recognition. When processing degraded document images, binarization is not an easy task. Degradations appear frequently and may occur due to several reasons which range from the acquisition source type to environmental conditions. Examples of degradation influence may include the appearance of variable background intensity caused by non-uniform intensity, shadows, smear, smudge and low contrast.

In general, approaches that deal with document image binarization are either global or local. In a global approach, threshold selection leads to a single threshold value for the entire image. Global thresholding [1–5] has a good performance in the case that there is a good separation between the foreground and the background. However, very often, document images are exposed to degradations that weaken any guaranty for such a separation. Unlike global approaches, local area information may guide the threshold value for each pixel in local (adaptive) thresholding techniques [6–14]. These techniques have been widely used in document image analysis because they have a better performance in extracting the character strokes from an image that contains spatially uneven gray levels due to degradations.

While it is essential to have a binarization technique that can correctly keep all essential textual information, it is of equal importance to apply this technique automatically without requiring from the user to adjust a set of parameters each time it is applied.

Taking all the above into account, in this paper, we propose a novel locally adaptive thresholding scheme which binarizes and enhances poor quality and degraded documents for the location of meaningful textual information

* Corresponding author. Tel.: +30 210 6503183; fax: +30 210 653 2175.
*E-mail addresses:* bgat@iit.demokritos.gr (B. Gatos), ipratika@iit.demokritos.gr (I. Pratikakis), sper@iit.demokritos.gr (S.J. Perantonis)
*URL:* http://www.iit.demokritos.gr/cil.

0031-3203/$30.00 
doi:10.1016/j.patcog.2005.09.010

without requiring any parameter tuning. The proposed scheme consists of five basic steps. The first step is dedicated to a denoising procedure using a low-pass Wiener filter. We use an adaptive Wiener method based on local statistics. In the second step, we use a first rough estimation of foreground regions. Next, as a third step, we compute the background surface of the image by interpolating neighboring background intensities into the foreground areas that result from the previous step. In the fourth step, we proceed to the final binarization by combining information from the calculated background surface and the original image. Text areas are located if the distance of the original image from the calculated background exceeds a threshold. This threshold adapts to the gray-scale value of the background surface in order to preserve textual information even in very dark background areas. In the last step, we proceed to a post-processing that eliminates noise, improves the quality of text regions and preserves stroke connectivity. The proposed method has been extensively tested with a variety of degraded image documents and has demonstrated superior performance against four (4) well-known techniques.

The paper is organized as follows. Section 2 briefly reviews the state-of-the-art with particular emphasis on local adaptive methods used during our experiments for comparison purposes. In Section 3, our methodology is described in detail while in Section 4 we discuss our experimental results. Finally, conclusions are drawn in Section 5.

## 2. Related work

In the literature, binarization is performed either globally or locally. For the global methods (global thresholding), a single calculated threshold value is used to classify image pixels into object or background classes [1–5], while for the local methods (adaptive thresholding), local area information guides the threshold value for each pixel [6,7]. Most of the image binarization algorithms rely on statistical methods, without taking into account the special nature of document images [8–10]. However, some document-directed binarization techniques have been developed [11–14]. For document image binarization, global thresholding methods are not sufficient since document images usually are degraded and have poor quality including shadows, non-uniform illumination, low contrast, large signal-dependent noise, smear and strains. Concerning local methods, a goal-directed performance evaluation of eleven popular local thresholding algorithms has been performed for map images [13]. According to this evaluation, for a slowly changing background, local algorithms work well. However, with a complex background, it appeared that none can be tuned up with a set of operating parameters good for all images. Furthermore, local algorithms were dependent on stroke width. A recent exhaustive survey of forty (40) image binarization methods, both global and local, is presented in Ref. [15]. Conducting a quantitative performance evaluation, local methods are shown to perform better. Nevertheless, this evaluation took into consideration only text document images that are degraded with noise and blur.

In the following, we review three (3) local binarization algorithms that are considered as the current state-of-the-art. These algorithms have been used for the comparison and evaluation of our approach.

Niblack [8] introduced an algorithm that calculates a pixelwise threshold by shifting a rectangular window across the image. The threshold $T$ for the center pixel of the window is computed using the mean $m$ and the variance $s$ of the gray values in the window

$$T = m + ks, \tag{1}$$

where $k$ is a constant set to $-0.2$. The value of $k$ is used to determine how much of the total print object boundary is taken as a part of the given object. This method can distinguish the object from the background effectively in the areas close to the objects. The results are not very sensitive to the window size as long as the window covers at least 1–2 characters. However, noise that is present in the background remains dominant in the final binary image. Consequently, if the objects are sparse in an image, a lot of background noise will be left.

Sauvola and Pietikainen [11] propose a method that solves this problem by adding a hypothesis on the gray values of text and background pixels (text pixels have gray values near 0 and background pixels have gray values near 255), which results in the following formula for the threshold:

$$T = m + (1 - k(1 - s/R)), \tag{2}$$

where $R$ is the dynamics of the standard deviation fixed to 128 and $k$ takes on positive values (usually set to 0.5). This method gives better results for document images.

Kim et al. [7] propose a local adaptive thresholding method where an image is regarded as a 3D terrain and its local property is characterized by a water flow model. The water flow model locally detects the valleys corresponding to regions that are lower than neighboring regions. The deep valleys are filled with dropped water whereas the smooth plain regions keep up dry. The final step in this method concerns the application of a global thresholding such as Otsu's method [2] on a difference image between the original terrain and the water-filled terrain. A shortcoming of this method is the selection of two critical parameters for the method, namely, the amount of rainfall, $w$, and the parameter of mask size, $s$, which is done on an experimental basis.

## 3. Methodology

The proposed methodology for degraded and poor quality document image binarization, text preservation and

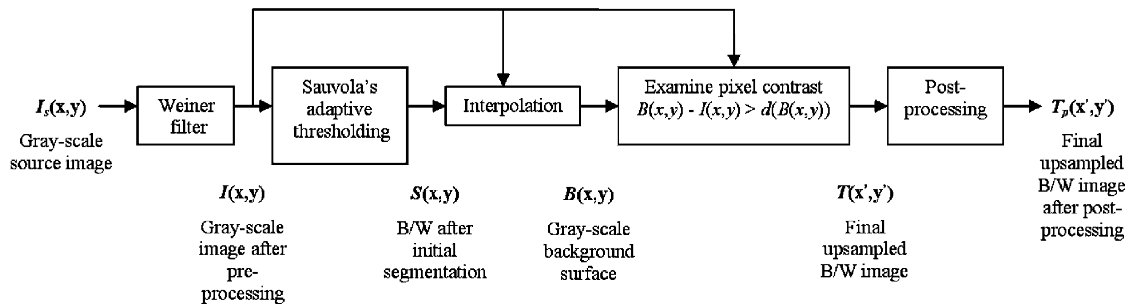



Fig. 1. Block diagram of the proposed methodology for low quality historical document text preservation.


Fig. 2. Image pre-processing: (a) original image; (b) 3 × 3 Wiener filtering.

enchancement is illustrated in Fig. 1 and fully described in this section.

### 3.1. Pre-processing

For degraded and poor quality documents, a pre-processing stage of the grayscale source image is essential for the elimination of noisy areas, smoothing of background texture as well as contrast enhancement between background and text areas. The use of a low-pass Wiener filter [16] has proved efficient for the above goals. The Wiener filter is commonly used in filtering theory for image restoration. Our pre-processing module involves an adaptive Wiener method based on statistics estimated from a local neighborhood around each pixel. The grayscale source image $I_s$ is transformed to the filtered grayscale image $I$ according to the following formula:

$$I(x, y) = \mu + \left(\sigma^2 - v^2\right)(I_s(x, y) - \mu) / \sigma^2, \tag{3}$$

where $\mu$ is the local mean, $\sigma^2$ the variance at a $3 \times 3$ neighborhood around each pixel and $v^2$ is the average of all estimated variances for each pixel in the neighborhood. Fig. 2 shows the results of applying a $3 \times 3$ Wiener filter to a document image.

### 3.2. Rough estimation of foreground regions

At this step, we obtain a rough estimation of foreground (text) regions. Our intention is to proceed to an initial segmentation of foreground and background regions that will provide us a superset of the correct set of foreground pixels. This is refined at a later step (Section 3.3). Sauvola's approach for adaptive thresholding [11] using $k = 0.2$, is suitable for this case (see Fig. 3). At this step, we process image $I(x, y)$ in order to extract the binary image $S(x, y)$, where 1's correspond to the rough estimated foreground regions.

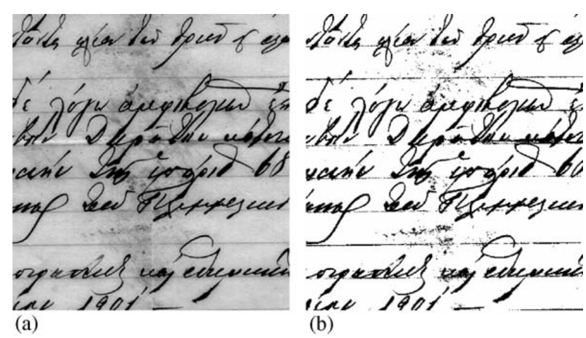



Fig. 3. Adaptive thresholding using Sauvola's approach: (a) original image; (b) rough estimation of foreground regions.

### 3.3. Background surface estimation

At this stage, we compute an approximate background surface $B(x, y)$ of the image $I(x, y)$. A similar approach has been proposed for the binarization of camera images [17]. Background surface estimation is guided by the valuation of $S(x, y)$ image. For pixels that correspond to 0's at image $S(x, y)$, the corresponding value at $B(x, y)$ equals to $I(x, y)$. For the remaining pixels, the valuation of $B(x, y)$ is computed by a neighboring pixel interpolation, as

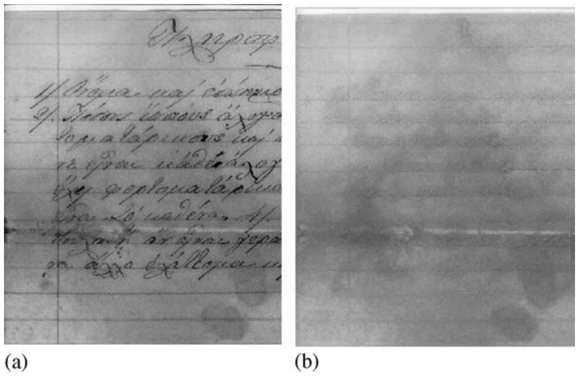


Fig. 4. Background surface estimation: (a) image $I$; (b) background surface $B$.

described in

$$B(x,y) = \begin{cases} I(x,y) & \text{if } S(x,y)=0, \\ \dfrac{\sum_{ix=x-\mathrm{d}x}^{x+\mathrm{d}x}\sum_{iy=y-\mathrm{d}y}^{y+\mathrm{d}y}(I(ix,iy)(1-S(ix,iy)))}{\sum_{ix=x-\mathrm{d}x}^{x+\mathrm{d}x}\sum_{iy=y-\mathrm{d}y}^{y+\mathrm{d}y}(1-S(ix,iy))} & \text{if } S(x,y)=1. \end{cases} \tag{4}$$

The interpolation window of size $\mathrm{d}x \times \mathrm{d}y$ is defined to cover at least two image characters. An example of the background surface estimation is shown in Fig. 4.

### 3.4. Final thresholding

In this step, we proceed to final thresholding by combining the calculated background surface $B(x, y)$ with the preprocessed image $I(x, y)$. Text areas are located if the distance of the preprocessed image $I(x, y)$ with the calculated background $B(x, y)$ exceeds a threshold $\boldsymbol{d}$. We suggest that the threshold $\boldsymbol{d}$ must change according to the gray-scale value of the background surface $B(x, y)$ in order to preserve textual information even in very dark background areas. For this reason, we propose a threshold $\boldsymbol{d}$ that has smaller values for darker regions. The final binary image $T(x, y)$ is given by the following formula:

$$T(x,y) = \begin{cases} 1 & \text{if } B(x,y) - I(x,y) > d(B(x,y)), \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

A typical histogram of a document image (see Fig. 5) has two peaks. One peak corresponds to text regions and the other peak corresponds to background regions. We may note that we consider gray value document images in the range of [0, 255] and textual information has a low gray level profile. The average distance $\delta$ between the foreground and background can be calculated by the following formula:

$$\delta = \frac{\sum_x \sum_y (B(x,y) - I(x,y))}{\sum_x \sum_y S(x,y)}. \tag{6}$$

In the case of document images with uniform illumination, the minimum threshold $d$ between text pixels and background pixels can be defined as $q \cdot \delta$, where $q$ is a weighting parameter. Fixing the value of $q$ at 0.8 we achieve total character body preservation that leads to successful OCR results [17]. At low contrast regions that appear in degraded and poor quality documents, we require a smaller value for the threshold $d$. To achieve this, we first compute the average background values $b$ of the background surface $B$ that correspond to the text areas of image $S$, denoted as

$$b = \frac{\sum_x \sum_y (B(x,y))(1 - S(x,y))}{\sum_x \sum_y (1 - S(x,y))}. \tag{7}$$

We require that the threshold be equal to the value $q \cdot \delta$ when the background value is high (greater than the average background value $b$ at Eq. (7)) and equal to $p_2 \cdot q \cdot \delta$ when the background value is low (less than $p_1 \cdot b$), with $p_1, p_2 \in [0, 1]$. To simulate this requirement, we use the following logistic sigmoid function that exhibits the desired saturation behavior for large and small values of the background as shown in Fig. 6:

$$d(B(x,y)) = q\delta\left(\frac{(1-p_2)}{1+\exp\left(\frac{-4B(x,y)}{b(1-p_1)} + \frac{2(1+p_1)}{(1-p_1)}\right)} + p_2\right). \tag{8}$$

After experimental work, for the case of degraded and poor quality document images, we suggest the following parameter values: $q = 0.6$, $p_1 = 0.5$, $p_2 = 0.8$.

### 3.5. Upsampling

In order to achieve a better quality binary image we incorporate in the previous step an efficient upsampling technique. Among available image upsampling techniques, bicubic interpolation is the most common technique that provides satisfactory results [18]. It estimates the value at a pixel in the destination image by an average of 16 pixels surrounding the closest corresponding pixel in the source image. Bicubic interpolation for image upsampling can be incorporated in our algorithm by substituting formula (5) with the following formula:

$$T(x',y') = \begin{cases} 1 & \text{if } B(x,y) - I_u(x',y') > d(B(x,y)), \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

where $T(x', y')$ is the binary image of size $M$ times the size of the original gray scale image, $x = (\text{int})x'/M$, $y = (\text{int})y'/M$ and $I_u$ is given by the following formula:

$$\begin{aligned} I_u(x',y') = & -b(1-b)^2 F(x', y-1) \\ & + \left(1 - 2b^2 + b^3\right) F(x', y) \\ & + b\left(1 + b - b^2\right) F(x', y+1) \\ & - b^2(1-b) F(x', y+2), \end{aligned} \tag{10}$$

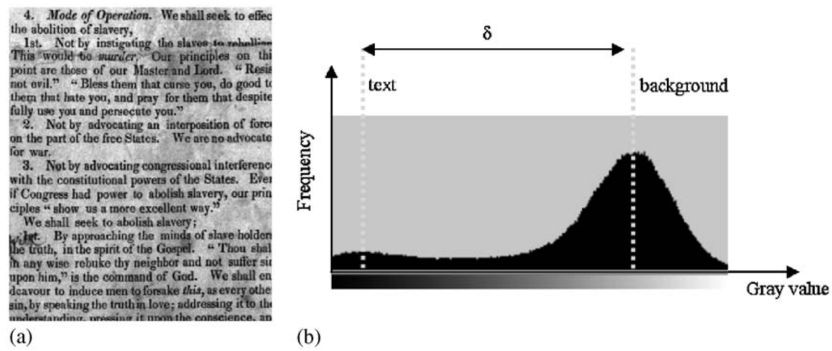



Fig. 5. Document image histogram: (a) original image; (b) gray level histogram.

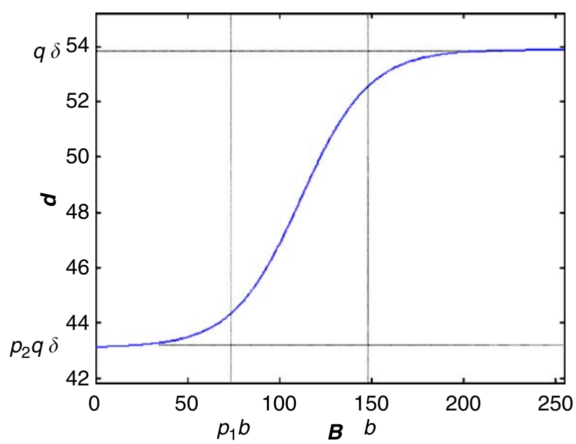



Fig. 6. Function $d(B(x, y))$.

where $F()$ is calculated as follows:

$$
\begin{aligned}
F(x', m) = & -a(1-a)^2 I(x-1, m) \\
& + \left(1 - 2a^2 + a^3\right) I(x, m) \\
& + a\left(1 + a - a^2\right) I(x+1, m) \\
& - a^2(1-a) I(x+2, m),
\end{aligned}
\tag{11}
$$

where $a = (x'/M) - x, b = (y'/M) - y$.

In most cases, a double sized resulting binary image ($M = 2$) is adequate for an improved quality result.

### 3.6. Post-processing

In the final step, we proceed to post-processing of the resulting binary image in order to eliminate noise, improve the quality of text regions and preserve stroke connectivity by isolated pixel removal and filling of possible breaks, gaps or holes. Below follows a detailed step-by-step description of the post-processing algorithm that consists of a successive application of shrink and swell filtering [19].

*Step* 1: A shrink filter is used to remove noise from the background. The entire binary image is scanned and each foreground pixel is examined. If $P_{sh}$ is the number of background pixels in a sliding $n \times n$ window, which has the foreground pixel as the central pixel, then this pixel is changed to background if $P_{sh} > k_{sh}$ where $k_{sh}$ can be defined experimentally.

*Step* 2: A swell filter is used to fill possible breaks, gaps or holes in the foreground. The entire binary image is scanned and each background pixel is examined. If $P_{sw}$ is the number of foreground pixels in a sliding $n \times n$ window, which has the background pixel $(x, y)$ as the central pixel, and $x_a$, $y_a$ the average values for all foreground pixels in the $n \times n$ window, then this pixel is changed to foreground if $P_{sw} > k_{sw}$ and $|x - x_a| < \mathrm{d}x$ and $|y - y_a| < \mathrm{d}y$. The latter two conditions are used in order to prevent an increase in the thickness of character strokes since we examine only background pixels among uniformly distributed foreground pixels.

*Step* 3: An extension of the above conditions, leads to a further application of a swell filter that is used to improve the quality of the character strokes. The entire binary image is scanned and each background pixel is examined. If $P_{sw1}$ is the number of foreground pixels in a sliding $n \times n$ window, which has the background pixel as the central pixel, then this pixel is changed to foreground if $P_{sw1} > k_{sw1}$.

All the parameters used at this step depend on the average character height $l_h$, which is estimated by using connected component analysis. A height histogram of the connected components is formed and the largest peak is selected as the average character height. After experimental work on a representative training set, we suggest the following parameter values for the post-processing phase: $n = 0.15 l_h$, $k_{sh} = 0.9n^2$, $k_{sw} = 0.05n^2$, $\mathrm{d}x = \mathrm{d}y = 0.25n$, $k_{sw1} = 0.35n^2$. An example of a resulting binary image after post-processing steps is given in Fig. 7.

## 4. Experimental results

The proposed algorithm was tested using degraded document images which belong to three distinct categories:

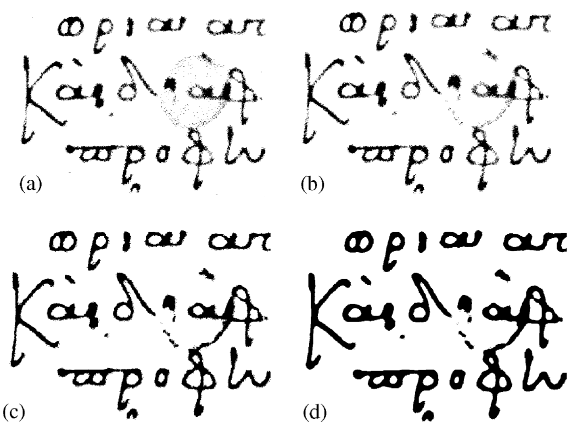
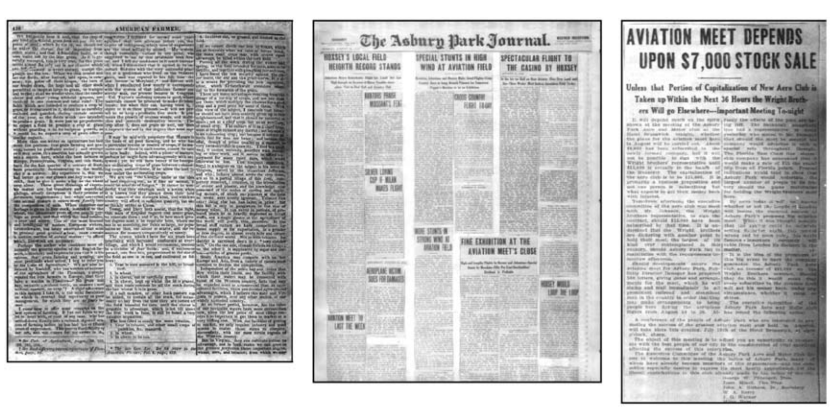



Fig. 7. Post-processing stage: (a) resulting binary image after final thresholding; (b) resulting image after post-processing Step 1; (c) resulting image after post-processing Step 2; (d) resulting image after post-processing Step 3.

historical handwritten documents, old newspapers and poor quality modern documents. All images have varying resolutions, stroke sizes, illumination contrast, background complexity and noise levels. The historical handwritten documents were selected from the Library of Congress on-line database [20], from the Mount Sinai Foundation Collection [21], from the Bodleian Library [22], as well as from private collections. All historical handwritten document images are of poor quality and have shadows, non-uniform illumination, smear and strain. In some of these documents, ink sipped through the pages and characters on the reverse side become visible and interfere with the characters on the front side. Example historical handwritten documents used for testing are shown in Fig. 8. Old newspaper images come from the Library of Congress on-line database [20] and suffer from problems similar to historical documents. Additionally, old newspaper images have extra noise due to the old printing matrix quality or ink diffusion. Example old newspaper documents used for testing are shown in Fig. 9. Poor quality modern documents originate from the MediaTeam Oulu Document Database [23] as well as from recent scannings of books and magazines. All modern documents are degraded with difficulties in distinguishing text and background regions.

We compared the performance of our algorithm with four (4) well-known binarization techniques. We evaluated the following: Otsu's global thresholding method [2], Niblack's adaptive thresholding method [8], Sauvola et al. adaptive

Fig. 8. Examples from the handwritten historical documents used for testing.


AMERICAN FARMER.

The Asbury Park Journal.

AVIATION MEET DEPENDS UPON $7,000 STOCK SALE

Fig. 9. Examples from the old newspapers used for testing.

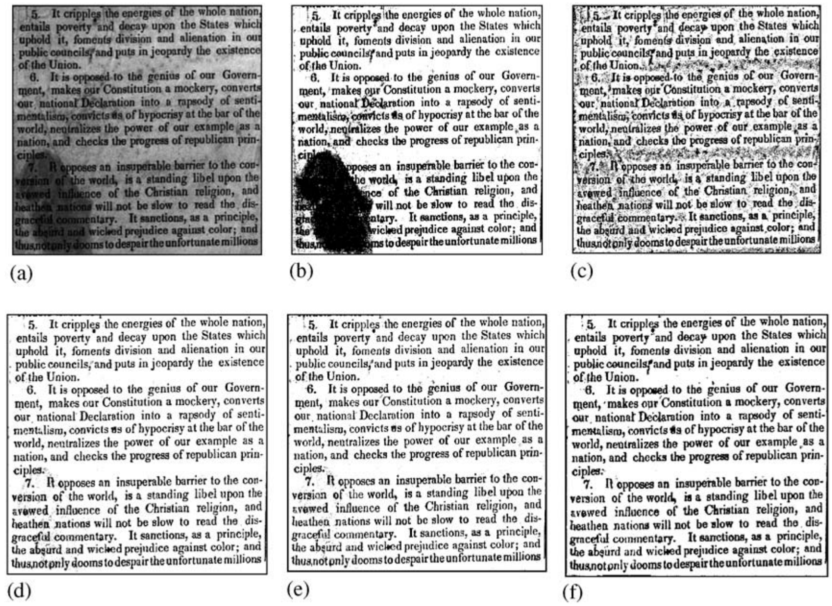



Fig. 10. Binarization of a typed document image: (a) original image; (b) Otsu's method [2]; (c) Niblack's method [8]; (d) Sauvola et al. method [11]; (e) Kim et al. method [7]; (f) the proposed method.


Fig. 11. Binarization of an historical handwritten document image: (a) original image; (b) Otsu's method [2]; (c) Niblack's method [8]; (d) Sauvola et al. method [11]; (e) Kim et al. method [7]; (f) the proposed method.

method [11], Kim et al. adaptive method [7] and our proposed adaptive method. Niblack's approach has been implemented with a $60 \times 60$ window in order to cover at least 1–2 characters in all testing samples. The value of $k$ (see Eq. (1)) is set to $-0.2$. Sauvola et al. approach has also been implemented with a $60 \times 60$ window. Concerning the value of $k$ (see Eq. (2)), it is set to 0.5. In Kim et al. approach the parameters chosen for the water flow model are:

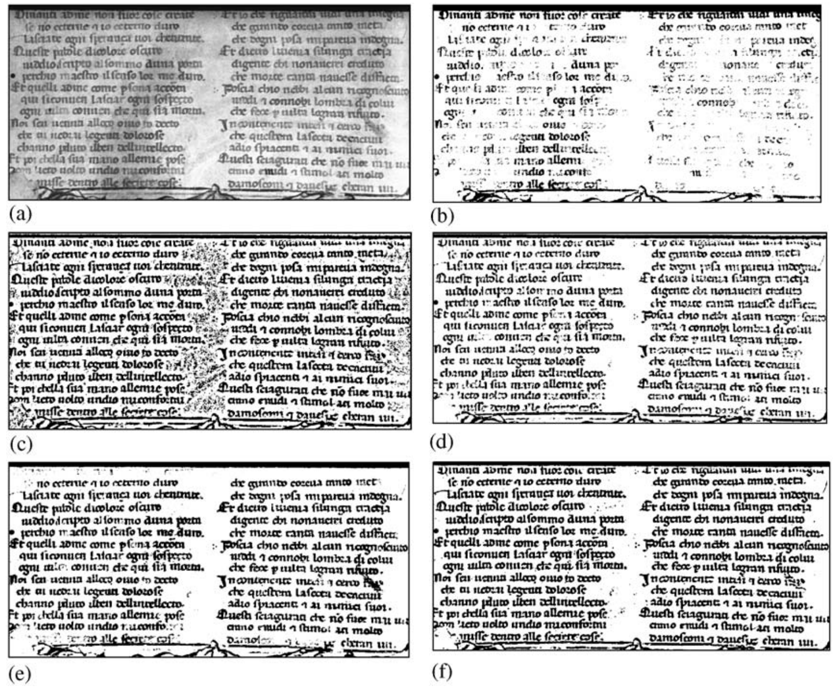


Fig. 12. Binarization of a historical handwritten document image: (a) original image; (b) Otsu's method [2]; (c) Niblack's method [8]; (d) Sauvola et al. method [11]; (e) Kim et al. method [7]; (f) the proposed method.

$w=20$ and $s=3$. All the parameters used for the above binarization algorithms have been proposed in the experiments by the authors of the corresponding algorithms as the best parameters for most document cases.

Based on visual criteria, the proposed algorithm outperforms all algorithms that it is tested against with respect to image quality and preservation of meaningful textual information. Example results are shown in Figs. 10–12. After a thorough examination of our experimental results, important observations can be summarized in the following:

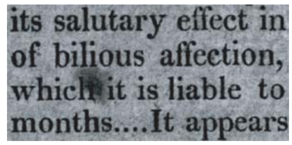
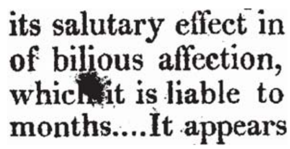
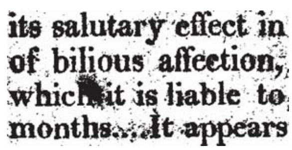
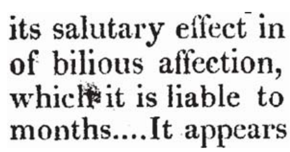
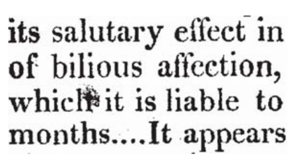
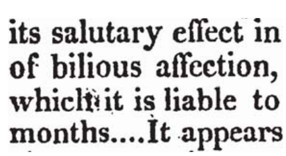
- The global thresholding technique of Otsu is not satisfactory for degraded documents that exhibit local variance problems.
- At Niblack's approach, the resulting binary image generally suffers from a great amount of background noise, especially in areas without text.
- The approach of Sauvola et al. solves the background noise problem that appears in Niblack's approach but in many cases characters become extremely thinned and broken.
- Although the approach of Kim et al. performs well in most of the testing cases, we have observed occasions that a great amount of noise remains in the resulting image as well as that characters become broken.
- The proposed method has superior performance compared with all other methods and performs well even when the documents are very noisy and highly degraded.

An additional experiment to quantify the efficiency of the proposed binarization method was also performed. We compared the results obtained by the well-known OCR engine ABBYY FineReader 6 [24], using as input image the binarization results of Otsu [2], Niblack [8], Sauvola et al. [11], Kim et al. [7] and the proposed method. A representative example of OCR results after the application of several binarization schemes example is shown in Table 1. To quantify the OCR results quality, we calculated the Levenshtein distance [25] between the correct text (ground truth) and the resulting text for five (5) representative degraded images. Table 2 shows that the application of the proposed binarization technique has demonstrated the best performance regarding the final OCR results. Overall results show that by using the proposed method, we achieve at least 37% improvement compared to the other approaches.

## 5. Conclusions

In this paper, we propose a novel locally adaptive approach for the binarization and enhancement of degraded documents. The proposed method does not require any parameter tuning by the user and can deal with degradations which occur due to shadows, non-uniform illumination, low contrast, large signal-dependent noise, smear and strain. We follow several distinct steps: a pre-processing procedure using a low-pass Wiener filter, a rough estimation of fore-

Table 1
Example document image portion and its corresponding OCR results. Marked words indicate the OCR errors

| Method | Document image portion | OCR result |
|---|---|---|
| Fine Reader [24] |  its salutary effect in of bilious affection, which it is liable to months....It appears | its salutary effect in of bilious affection, whicHKt is liable to months....It appears |
| Otsu [2] |  its salutary effect in of bilious affection, which it is liable to months....It appears | its salutary effect in of bilious affection, whicUlt is liable to months....It appears |
| Niblack [8] |  its salutary effect in of bilious affection, which it is liable to months....It appears | its salutary efex&m aof^Jious;affettio^il^ Wcw^^iiabte^flte^prive^0nths^^fc«pp^ar9 |
| Sauvola et al. [11] |  its salutary effect in of bilious affection, which it is liable to months....It appears | its salutary effect in of bilious affection, whictt*it is liable to months....It appears |
| Kim et al. [7] |  its salutary effect in of bilious affection, which it is liable to months....It appears | its salutary effect in of bilious affection, whicftit is liable to months....It appears |
| Proposed method |  its salutary effect in of bilious affection, which it is liable to months....It appears | its salutary effect in of bilious affection, which it is liable to months....It appears |

ground regions, a background surface calculation by interpolating neighboring background intensities, a thresholding by combining the calculated background surface with the original image while incorporating image up-sampling and finally a post-processing step in order to improve the quality of text regions and preserve stroke connectivity. After extensive experiments, our method demonstrated superior performance against four (4) well-known techniques on numerous degraded document images using visual criteria as well as quantitative measures.

Further research will focus on the challenges that emerge from the binarization of low resolution images and videos found on the Web.

**Acknowledgements**

This research is carried out within the framework of the Greek GSRT-funded R&D project, D-SCRIBE, which aims to develop an integrated system for digitization and processing of old Greek manuscripts.

The authors would like to thank Dr. Ergina Kavallieratou from Open University, Greece, for providing us with samples from her private historical document collection, Dr. Kim and Prof. Park from Sogang University, South Korea for the provision of their binarization algorithm as well as Prof. Pietikainnen from Oulu University, Finland, who kindly provided us the Oulu Document Database.

Table 2
OCR evaluation

| | Documents | | | | | |
|---|---|---|---|---|---|---|
| |  |  |  |  |  | |
| | Chars = 3600<br>Words = 646 | Chars = 883<br>Words = 149 | Chars = 2315<br>Words = 420 | Chars = 1797<br>Words = 288 | Chars = 2823<br>Words = 452 | Total<br>Chars = 11418<br>Words = 1955 |
| Methods | Levenshtein distance from the ground truth | | | | | |
| FineReader [24] | 159 | 81 | 124 | 98 | 85 | 547 |
| Otsu [2] | 417 | 262 | 162 | 324 | 184 | 1349 |
| Niblack [8] | 1198 | 532 | 1447 | 1249 | 2116 | 6542 |
| Sauvola et al. [11] | 132 | 65 | 157 | 696 | 82 | 1132 |
| Kim et al. [7] | 150 | 67 | 131 | 116 | 112 | 576 |
| Proposed method | **52** | **58** | **78** | **76** | **81** | **345** |

## References

[1] A. Rosenfeld, A.C. Kak, Digital Picture Processing, second ed., Academic Press, New York, 1982.
[2] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Systems Man Cybernet. 9 (1) (1979) 62–66.
[3] J. Kittler, J. Illingworth, On threshold selection using clustering criteria, IEEE Trans. Systems Man Cybernet. 15 (1985) 652–655.
[4] A.D. Brink, Thresholding of digital images using two-dimensional entropies, Pattern Recognition 25 (8) (1992) 803–808.
[5] H. Yan, Unified formulation of a class of image thresholding techniques, Pattern Recognition 29 (12) (1996) 2025–2032.
[6] P.K. Sahoo, S. Soltani, A.K.C. Wong, A survey of thresholding techniques, Comput. Vision, Graphics Image Processing 41 (2) (1988) 233–260.
[7] I.-K. Kim, D.-W. Jung, R.-H. Park, Document image binarization based on topographic analysis using a water flow model, Pattern Recognition 35 (2002) 265–277.
[8] W. Niblack, An Introduction to Digital Image Processing, Prentice-Hall, Englewood Cliffs, NJ, 1986 pp. 115–116.
[9] J. Yang, Y. Chen, W. Hsu, Adaptive thresholding algorithm and its hardware implementation, Pattern Recognition Lett. 15 (2) (1994) 141–150.
[10] J.R. Parker, C. Jennings, A.G. Salkauskas, Thresholding using an illumination model, ICDAR'93, 1993, pp. 270–273.
[11] J. Sauvola, M. Pietikainen, Adaptive document image binarization, Pattern Recognition 33 (2000) 225–236.
[12] M. Chang, S. Kang, W. Rho, H. Kim, D. Kim, Improved binarization algorithm for document image by histogram and edge detection, ICDAR'95, 1993, pp. 636–643.
[13] O.D. Trier, A.K. Jain, Goal-directed evaluation of binarization methods, IEEE Trans. Pattern Anal. Mach. Intell. 17 (12) (1995) 1191–1201.
[14] L. Eikvil, T. Taxt, K. Moen, A fast adaptive method for binarization of document images, International Conference on Document Analysis and Recognition, France, 1991, pp. 435–443.
[15] M. Sezgin, B. Sankur, Survey over image thresholding techniques and quantitative performance evaluation, J. Electron. Imaging 13 (1) (2004) 146–165.
[16] A. Jain, Fundamentals of Digital Image Processing, Prentice-Hall, Englewood Cliffs, NJ, 1989.
[17] M. Seeger, C. Dance, Binarising Camera Images for OCR, Sixth International Conference on Document Analysis and Recognition (ICDAR'01), Seattle, Washington, 2001, pp. 54–58.
[18] H. Hsieh, H. Andrews, Cubic splines for image interpolation and digital filtering, IEEE Trans. Acoustics Speech Signal Process. 26 (6) (1978) 508–517.
[19] R.J. Schilling, Fundamentals of Robotics Analysis and Control, Prentice-Hall, Englewood Cliffs, NJ, 1990.
[20] The Library of Congress (http://memory.loc.gov/).
[21] Holy Monastery of St. Catherine at Mount Sinai (http://www.sinaimonastery.com/).
[22] Bodleian Library—University of Oxford, (http://www.bodley.ox.ac.uk/).
[23] Media Team Oulu Document database (http://www.mediateam.oulu.fi/MTDB/).
[24] ABBYY (www.finereader.com).
[25] V.I. Levenshtein, Binary codes capable of correcting deletions, insertions and reversals, Sov. Phys. Dokl. 6 (1966) 707–710.

**About the Author**—BASILIOS GATOS received his Electrical Engineering Diploma in 1992 and his Ph.D. degree in 1998, both from the Electrical and Computer Engineering Department of Democritus University of Thrace, Xanthi, Greece. He is currently working as a Researcher at the Institute of Informatics and Telecommunications of the National Center for Scientific Research "Demokritos", Athens, Greece. His main research interests are in Image Processing, Pattern Recognition, Document Image Analysis, OCR, Processing and Recognition of Historical Documents.

**About the Author**—IOANNIS PRATIKAKIS received the Diploma degree in Electrical Engineering from the Demokritus University of Thrace, Xanthi, Greece in 1992, and the Ph.D. degree in Applied Sciences from Vrije Universiteit Brussel, Brussels, Belgium, in 1998. From March 1999 to March 2000, he was at IRISA/ViSTA group, Rennes, France as an INRIA postdoctoral fellow. He is currently working as a Research Scientist at the Institute of Informatics and Telecommunications of the NCSR "Demokritos". His research interests include 2D and 3D image analysis, image and volume sequence analysis as well as content-based image/3D models search and retrieval.

**About the Author**—STAVROS J. PERANTONIS is the holder of a BS degree in Physics from the Department of Physics, University of Athens, an M.Sc. degree in Computer Science from the Department of Computer Science, University of Liverpool and a D.Phil. Degree in Computational Physics from the Department of Physics, University of Oxford. Since 1992 he has been with the Institute of Informatics and Telecommunications, NCSR "Demokritos", where he currently holds the position of Senior Researcher and Head of the Computational Intelligence Laboratory. His main research interests are in Image Processing and Document Image Analysis, OCR and Pattern Recognition.