

A New Mixed Binarization Method Used in a Real Time Application of Automatic Business Document and Postal Mail Sorting

Djamel Gaceb, Véronique Eglin, and Frank Lebourgeois
LIRIS Laboratory, National Institut of Applied Science, France

Abstract: *The binarization is applied in the first stage of segmentation process and has a very strong impact on the performances of the system of the automatic sorting of company documents and mail. We present in the beginning of this paper a complete study of the different existing binarization mechanisms that are developed to meet the needs of specific applications. These conventional approaches, present weaknesses that it is crucial to overcome and unfortunately they remain unsuitable for our real time application. The separation between the thresholding and the text zones location stages considerably increase the computation time and lead to an over-segmentation of the noise and of the paper texture on empty zones of the image. Indeed, none of the traditional methods (whether global or local) efficiently meets all the required conditions. We have managed to optimize this stage by applying a local threshold only near the text zones that can be located by the cumulated gradients method with the multi-resolution and mathematical morphology. We demonstrate the consistent performance of the proposed method on several types of business documents and mail with wide-ranging content and image quality.*

Keywords: *Binarization, text zones location, real time processing, automatic sorting of company documents, mail.*

Received November 9, 2010; accepted May 24, 2011

1. Introduction

Automatic document and mail sorting machines of most recent systems process about 17 mail pieces per second that requires a fast and precise optical reading of interest zones using the OCR technology. This phase is mainly conditioned by a correct binarization process. In our work, we should take into account two constraints: a very large mail variety (size, quality, colour and different paper textures), real-time constraints and high spatial resolution of the images. Image acquisition of documents and delivered post mails is performed in greyscale because this permits the use of adaptive thresholding to generate the binarized text. The OCR analysis that constitutes a key step in the process of sorting post mails and companies documents requires a reduction in the amount of information passing through an inescapable preliminary step of binarization which, in itself, has a strong impact on the performance of all subsequent steps of automatic processing of the document.

The problem will be simpler if the grey level that is associated with the background was uniform, if the grey level associated to objects was also, and in the end if these grey levels were sufficiently different so that, compared with a threshold assumed to be known, we can label with white all the pixels with a grey level greater or equal to this threshold and with black all the pixel with grey level smaller to the same threshold. In practice, this ideal situation occurs very rarely and this

dichotomy is obviously not perfect because of the lighting defects or the noise that is introduced by the sensor itself. Therefore, a poor choice of binarization threshold can destroy a big amount of information contained in the image by degrading the quality of the characters to be recognized by the OCR, these characters may well be fragmented or merged shows that Figure 1.

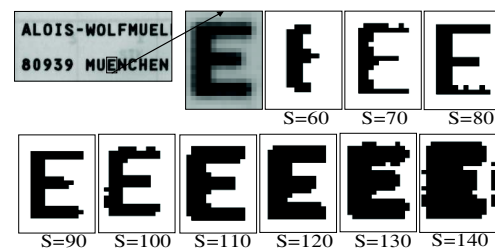


Figure 1. The effect of threshold on the character quality.

Thus, a binarization that is directly applied on degraded document images introduces many artefacts that lead to errors in the subsequent analysis modules. Therefore, it is necessary to apply pre-treatment of contrast enhancement, histogram equalization and noise reduction by filtering to be able to improve the quality of this binarization. In the context, we can cite for example the works of Ramponi, Shan and Fontanont [12, 26, 30], that have proposed quadratic filtering approaches to ameliorate the quality of mail images in the binarization phase shows that Figure 2.

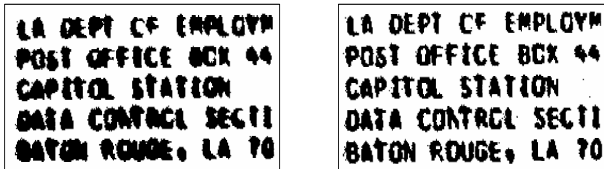


Figure 2. Example of a binary image without pre-treatment (left) and to the right with pre-treatment (quadratic filtering).

This figure illustrates the fact that a good binarization must be able to keep both the characters and objects without recovering too much noise. To solve this problem, we have identified a large number of works on the binarization of documents. We can distinguish three categories of methods depending on the nature of the used thresholding: global methods, local methods and hybrid methods that exploit the two previous approaches.

The separation between the thresholding and the text zones location phases considerably increase the computation time and lead to an over-segmentation of the noise and of the paper texture on empty zones of the image. Indeed, none of the traditional methods (whether global or local) efficiently meets all the required conditions.

This manuscript describes a new mixed method of fast binarization by applying a local threshold only near the text zones that can be located by the cumulated gradients method with the multi-resolution and mathematical morphology.

The paper is organized as follows: various existing binarization methods are quoted in section 2. We point out all previous works of the domain and their limits. Section 3 describes our mixed method of binarization. Experiments and results of our method are then commented and discussed. We also demonstrate in this last part the effectiveness of our proposition with a detailed description and evaluation of the performance.

2. The Study of Existing Methods

2.1. The Global Thresholding Methods Analysis

The Global methods determine a single threshold for any image starting from a point of view that objects must have a distribution of the grey levels that is relatively distinct from the background part. In this case, the search of the threshold is done by analyzing the histogram of grey levels and by determining a local minimum. Pixels with grey level below the threshold are set in black and others in white.

2.1.1. Methods Based on the Separation of Distribution

The localization of the binarization threshold is done by the separation of the distributions from the grey level of the entire image. This can be achieved by a separation of two Gaussians that model the histogram

or by using a minimization threshold of the amount of inertia, that is associated with the classes (objects, background) [7]. The famous method of Fisher [11] is to model the bimodal histogram of the grey levels of an image by a weighted sum of two Gaussian distributions and to locate a threshold that is regarded as a separator of the distributions. For this, we use a minimization criterion of the sum of inertia associated with the two classes of grey levels (C_1 and C_2). The grey level solution to this maximization corresponds to the threshold sought to distinguish the two classes. Other methods consist of finding a threshold by iteratively separating the histogram into two classes with an a priori knowledge of values associated with each class as is the case for the method ISODATA [7]. Another widely used method is that of the K-means that consists of assigning to each class a pixel that will constitute its initial centre of gravity. Each pixel of the image is then assigned to the class whose centre of gravity is closest. The centres of gravity are calculated again and the process continues iteratively until it convergence. The k-means can be either global applied directly to the entire image or local applied to each window of the image to which we select a size. The serialization of the k-means is to initialize the centres of gravity of each window with the centres of gravity of the final window before. The serialization gives very interesting results but is very consuming in terms of computation time. It requires more initialization of centres by the user [22].

2.1.2. Methods Based on the Discriminant Analysis

Otsu [24] equationtes the problem of binarization as discriminante analysis, for which he uses a particular criterion function as a measure of statistical separation. The statistics are calculated for the two intensity values classes separated by a threshold i . It calculates the statistics for each level of intensity k , i.e., for all possible thresholds. As part of the binarization by Otsu's method, the separation occurs from the mean and variance. We therefore calculate:

$$\mu(i) = \sum_{k=1}^i k \times h_{Normalized}(k) \quad et \quad \omega(i) = \sum_{k=1}^i h_{Normalized}(k) \quad (1)$$

In the end, we calculate for each value of k the value:

$$S_2(i) = \alpha(i) \times [1 - \alpha(i)] \times [\mu(255) \times \alpha(i) - \mu(i)]^2 \quad | \quad i=1...255 \quad (2)$$

The level that maximizes the criterion function is chosen as the threshold of binarization. Thus the threshold value is obtained for i such that $S_2(i) = \max(S_2)$ for any value of i varying from 1 to 255.

A variant of this approach was proposed by Tsai in [40]: it initially cuts the image recursively in quad-tree, and then applies a thresholding of type Otsu in each block. This method adapts well to the forms of the plots and solves the problem related to a non-uniform distribution of the light intensity. The disadvantage of

this approach is its computational cost and the risk of generating blocks entirely black.

2.1.3. Methods Based on the Entropy Principle

These methods are based on the optimization of a criteria function, the entropy in this case [9, 17]. Couto *et al.* [8] divides the image into sub-images and calculates the fuzzy entropy of each one. The best threshold is associated with the A-IFS of lowest entropy. Recent similar approach based on fuzzy entropy is presented in [3].

2.1.4. Methods Based on the Histogram Transformation

In these approaches, the threshold is not selected directly, but after the transformation of the grey levels histogram of the image. This transformation aims to raise the peaks and lower the valleys, by associating with each pixel a weight depending on its local properties, which can well discriminate the histogram modes [38].

2.1.5. Methods Based on the Cooccurrence Matrix

The co-occurrence matrix $M(d, \theta)$ is a matrix whose entries are the frequencies that are relative for the two neighbouring pixels Ng_i and Ng_j , separated by a distance d with an orientation θ . Köhler [18] gives a contrast measure by using the co-occurrence matrices, whose elements correspond to pairs of values of grey levels. The optimal threshold is determined for a maximum contrast.

2.1.6. Methods Based on the Neural Networks

Babaguchi [1] have proposed a binarization method based on a connectionist model (CMB). This technique is structured in two phases (learning and binarization). In the learning phase, the network uses the back propagation algorithm on the whole database of images, each represented by its own histogram and a desired threshold. In the binarization phase, the network receives the input histogram of an unknown image and returns as output the optimal threshold of binarization.

2.1.7. Discussion of the Global Thresholding Methods

The global thresholding methods have the advantage of being extremely fast, but their major drawback is that it ignores the spatial relationships between pixels of an object. Therefore, there can be no assurance that the selected pixels are properly adjacent and that the forms are indeed being isolated from the background. In this sense, the background pixels can be easily integrated into objects and vice versa the object pixels can be classified as background points.

This phenomenon occurs particularly around the contours and very noisy objects. Another frequently raised objection to thresholding method is a not necessarily constant illumination on the image: the change of light on the document lowers the quality of binarization by creating entirely black areas once the image have been binarized. In that case we must consider approaches that use a local thresholding: the threshold at any point of the image is then defined as a function of illumination in the vicinity of each point.

2.2. The Local Thresholding Methods

The local thresholding methods (adaptive) adapt to the context of each pixel by calculating a single threshold for each pixel of the image based on the information contained in its neighbourhood. This compensates for the brightness variations and the local degradations of an image. If the window covers an area of low image contrast, the sensitivity of the detection threshold is automatically increased. There are several methods of local thresholding. We will show in what will follow those that are most used in the field of automatic processing of documents.

2.2.1. Methods Based on the Concept of Niblack

The concept of Niblack [23] is based on the calculation of a thresholding value by dragging a window on the image. For each pixel a threshold T is calculated based on some statistics such as the local mean m and standard deviation s that are calculated on the grey levels of the neighbouring pixels in the window by the following equation: $T = m + k \times s$ with k is a negative constant.

A comparative study conducted by Trier and Jain [33] showed that the method of Niblack segmented well text characters and gives better performances on images of documents with respect to other methods of global and local binarization. This efficiency was also confirmed by [15] who compared the Niblack method with other newer methods. However, this algorithm produces noise on images whose background is degraded, resulting in the need of a post-processing that is very time consuming.

Sauvola [27] decided to ameliorate the equation ($T = m + k \times s$) by adding a hypothesis on the grey level values of text and background pixels (all the text pixels have grey levels close to 0 and all the background pixels have grey levels near 255), the local thresholding equation becomes:

$$T = m \cdot (1 - k \cdot (1 - \frac{s}{R})) \quad (3)$$

k is a parameter defined as $k = 0.5$ and R is the dynamic of the standard deviation, defined as $R = 128$. The results obtained with the method of Sauvola *et al.* show an actual reduction of noise due to the initial hypothesis on the data. From a quality point of view,

the problem has been resolved in the case of company document and post mail images by Sauvola by adding in the calculation that a dark pixel belongs more probably to the text than to the background. Today many segmentation algorithms methods use Niblack and Sauvola algorithms as basic approaches that they improve either by adjusting the equation 3 [10, 36] or by adding post-processing, [14, 35].

Valverde and Grigat [35] have attempted to overcome the disadvantages of the method of Niblack by the combination of two stages of post-processing to have more reliability on technical documents. The stages of post-processing use mainly morphological closing to reduce the noise and fill holes and a Sobel gradient to improve the shape of characters. In the case of video material that are characterised by different properties (low contrast, important deviations of greyscale, etc.,) the chosen Hypotheses by Sauvola are not always justified, sometimes causing “holes” in the characters. The method proposed by Wolf *et al.* [36] solves this problem for videos and improves the performance on the level of noise by a normalization of contrasts and the averages of grey levels as follows:

$$T = (1-k) \cdot m + k \cdot M + k \cdot \frac{s}{R} \cdot (m-M) \quad (4)$$

where M is the minimum of the grey levels of the whole image and the dynamic deviation R is set to the maximum of the standard deviations of all the windows. This algorithm concentrates on the maximum deviation of the entire image which may still cause damage on video images that are exposed to large changes in background luminance. Faced with this problem, Feng and Tan [10] offers a more reliable version with a new approach to the calculation of m , M and s estimated in a local primary window whose size is large enough to cover one or two characters of text. To compensate the effect of luminance variation, the dynamic of the deviation R is calculated this time on a local secondary window of a larger size instead of the entire image. The threshold T can then be given by the following equation:

$$T = (1-\alpha_1) \cdot m + \alpha_2 \cdot \frac{s}{R_s} \cdot (m-M) + \alpha_3 \cdot M \quad (5)$$

With $\alpha_2 = k_1 \cdot \left(\frac{s}{R}\right)^\gamma$ and $\alpha_3 = k_2 \cdot \left(\frac{s}{R_s}\right)^\gamma$

where α_1 , γ , k_1 and k_2 are positives constants.

2.2.2. Local Methods Based on Neural Networks

These methods use neural networks to give to the binarization system the learning capability to optimize the threshold values. In this context, Chigusa [6] use Hopfield network in a parallel algorithm. The threshold of each neuron is selected in an adaptive manner in function of the neighbouring grey levels to the central pixel represented by this neuron. The synaptic weights

between remote neurons are set to 0, and between neighbouring neurons are set to 1. This neighbourhood is measured by Euclidean distance and the equilibrium state of the network represents the final binary image.

Chi and Wong [5] propose to make a relation between the binarization phase and the phase of segmentation into block. They binarize the initial image by a global thresholding, and then they segment the binary image into blocks. They then apply a feedback to binarize each block by a neural network. This method cannot be effective on documents that have a non-uniform luminance as the global binarization may cause the fusion of the blocks and degrade the quality of the segmentation into block. To avoid this drawback, Hamza *et al.* [14] combined a Self-Organizing Map (SOM) with the K-means methods, Sauvola's and Niblack's methods. This combination gives better results on degraded document images but the performance is always conditioned by the choice of the number of neurons and the representativeness of the training set.

2.2.3. Methods Based on the Markov Random Fields

These methods use the Markov random field model after learning of calculating local threshold for each pixel according to the different configurations in its neighbourhood. We can cite as examples the method of Wolf *et al.* [36], for low quality text binarization (video). Kuk and Lelore *et al.* [19, 21] have improved this concept to binarize degraded old documents. These methods have the advantage of being robust to noise, non-uniform illumination and degradations, but the calculation times remain very important. Moreover, these methods give good performance results on very thin handwritten strokes.

2.3. Hybrid Thresholding Methods

These methods use a global and local analysis simultaneously. Trier and Taxt [34] proposed a method adapted to technical documents. For this, they calculate the Laplacian from the partial derivatives of the result issued from a median filter of the initial image, then they create a labelled image and separate the background of the form according to the algorithm described in [34]. In the end, they apply post-treatments that are presented in [39] to improve the quality of the binary image. This method works well on good quality documents, but has the drawback of being governed by several parameters that are difficult to resolve in practice. After separating the background from foreground, Chang [4] propose to equalize the histogram of the foreground so as to facilitate the distinction between characters and noise. Later on they use a Laplacian to reconstruct the shape of binary characters. Her in Vakis [28] proposed two thresholding algorithms for images suitable for

scanning of documents at high speed. The first algorithm uses an adaptive thresholding that works by commutation: they apply either a local thresholding on the pixels that possess a local gradient responding to a strong transition on the level of the path, either a global thresholding on the pixels with weak gradient belonging to homogeneous pixels of the background. The second algorithm is based on path tracing by using a regrouping based on a variant type algorithm K-means. Both approaches can be used independently or combined for a better result. In the same principle, Kamada and Fujimoto [16] propose a sequential technique for low resolution such as the images taken from camera phone. This binarization is done by two tasks: the first separates the foreground from the background by global thresholding, while the second applies an extraction of neighbourhood of characters and a linear interpolation of the image and then uses a local thresholding the foreground pixels, thereby improving the quality of the characters for OCR.

Wu and Amin [37] propose a hybrid thresholding method of postal envelopes images which is applied in two steps. The first step applies a global thresholding on the original image. The second step adjusts the threshold value based on spatial characteristics of the components formed in the first step. The method gives good results on most of the simple envelopes images, well contrasted and good quality. Badekas and Papamarkos propose in [2] a binarization system integrating the results of six independent local and global binarization techniques. Each of these techniques has a coefficient (weight) of contribution that must be assigned by the user so that the sum of all coefficients is equal to 100%. This method is primarily intended for documents with high degradation or low luminance. In the same context, Gosselin and Thillou proposed in [32] cooperation between the binarization phase and the segmentation phase of characters that are ameliorated mutually. And more recently, Tanaka [31] used a post treatment based on the analysis of the background of document images to correct the result of adaptive thresholding of the method of Niblack.

2.4. Assessments of the Binarization Methods

In fact, the simplest methods using a global thresholding has the advantage of being extremely fast but the change of lighting, the presence of various graphics printed on envelopes with different colour inks are rapidly decreasing the quality of binarization. The local methods exceed these limits and are more adapted to local changes in contrast. However they require more calculations, thus they are slower and unsuitable for real-time applications. Although they provide a good efficiency the documents that are concerned in our application, these local binarization approaches have mainly the following disadvantages: prohibitive time computation depending on the size of

the analysis window, over segmentation of the defects and textures of the background of the image, and difficult treatment of documents whose characters vary in size (the analysis window is fixed throughout the treatment).

Faced with these drawbacks and in regards of our real-time industrial application, it is necessary to devise a binarization method by respecting the very low computation times while restoring binary images of very good quality. The quality of binarized images remains a difficult issue to assess. Currently, few studies have been made to solve the problem of assessing the result of binarization. It is possible to group them into three families of approaches: The first family of approaches assess the quality of binarization by measuring the performance of text recognition [15, 33, 41]. The results of evaluation performed by Trier [33] and He [15] show that the methods of Niblack and Sauvola give better results on document images. The second family of approaches assesses this quality by measuring its similarity with a ground truth [29]. The third family of approaches is based on an unsupervised evaluation criterion for estimating the quality of a segmentation result and can be used to assess the outcome of a binarization [25].

We chose an evaluation approach of our binarization method of document that is based on the direct evaluation of the OCR results on site.

3. Our Binarization Approach

The localization (or detection) of text zones must be applied directly on the greyscale images acquired by a linear high speed and resolution CCD camera. The acquisition of images of documents or envelopes moving in a sorting process generates a slight blur on the level of characters. The ink used in printing, the texture and paper quality vary from one document to another. Moreover, the brightness is not always uniform because of the presence of some folds. All these parameters make the task of localization very difficult and costly in computing time. To better overcome these difficulties, we used the technique of accumulated gradients which poses no constraints on lighting or the taking of pictures. This regularity is calculated from a sequence of pixels with high gradients. The principle relies on the fact that the text characters form a regular texture, the higher gradient amplitudes then correspond to the strong luminance transitions on the contours level intra or inter characters. To avoid introducing new thresholds or lose the relevance of certain points in the application of this filter, we perform locally on the neighbourhood V_s of each point (x_0, y_0) a simple summation of the normalized gradients by the number N , number of the neighbour pixels $V_s(x_0, y_0)$.

$$Gr(x_0, y_0) = \frac{1}{N} \sum_{(x,y) \in V_s(x_0, y_0)} \frac{\partial f(x, y)}{\partial \vec{v}} \quad (6)$$

This “cumulated gradient” filter, originally developed for text localization in video images (small image) [20] was used to locate particular unconstrained titles in videos such as Television Archive [36] and segment of the printed color composite. This filter in its original version assumes that the text direction is a priori known; the derivatives are calculated in the supposed direction of the text (often horizontal) and summoned in the same direction. This strict convention inevitably raises issues on inclined documents. Therefore, we propose to improve and adapt the filter as follows, we calculate the horizontal and vertical derivatives, and then we summon them in the two directions (11) to make the filter insensitive to the rotation of images of documents.

To better adapt to the large size of our images and the real time constraint, we also implemented a quick but rough approximation for the calculation of the derivatives (12). The cost of calculating the summation at each point of the image in a neighbourhood V_s is too high for our application. We will therefore reduce the computational cost by performing the summation by blocks in multi-resolution. We divide the image into rectangular blocks of size $d_x \times d_y$, then we calculate in each block the sum of vertical and horizontal gradients as written in the following equations:

$$Gr(x_0, y_0) = \frac{1}{dx \, dy} \sum_{j=1}^{d_x} \sum_{i=1}^{d_y} \left| \frac{\partial I(x_0 + j, y_0 + i)}{\partial x} \right| + \left| \frac{\partial I(x_0 + j, y_0 + i)}{\partial y} \right| \quad (7)$$

$$\begin{aligned} \text{with } \frac{\partial I}{\partial x}(u, v) &= I(u-2, v) - I(u+2, v) \\ \text{and } \frac{\partial I}{\partial y}(u, v) &= I(u, v-2) - I(u, v+2) \end{aligned} \quad (8)$$

The accumulation of the gradients by blocks gives a quick low resolution picture Gr where the text zones are clearly the brightest areas see Figure 3. To be able to obtain a filter equivalent to the original algorithm and to move closer the result of the summation by block to the summation of the summation by pixel, we apply a morphological smoothing on the image Gr . The additional calculation of the pre-treatment in low resolution is negligible.

For this, we apply successively on the greyscale image Gr , d_1 dilations, e_1 erosions, d_2 dilations and e_2 erosions. These morphological operations can be given for any structural element ξ by the following equation 9:

$$M(x, y) = \left[\left[\left[(Gr(x, y) \oplus \xi_{d1}) \ominus \xi_{e1} \right] \oplus \xi_{d2} \right] \ominus \xi_{e2} \right] \quad (9)$$

With

$$\xi_{d1} = \xi \times d_1, \quad \xi_{e1} = \xi \times e_1, \quad \xi_{d2} = \xi \times d_2, \quad \xi_{e2} = \xi \times e_2.$$

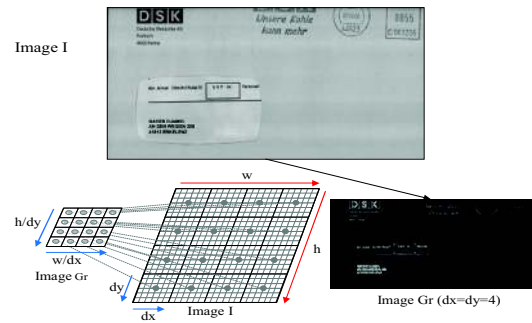


Figure 3. Accumulation of the neighborhood gradients in a low resolution.

The application of this morphological treatment allows, on one hand, to re-densify the text and therefore to agglomerate it into blocks and, and on the other hand, to take an adequate margin around the line to be able to include pertinent information that are carried by background (texture and colour) for better thresholding. This also improves the quality of binary characters which has the effect of increasing the rate of the OCR. The parameters d_1 , e_1 , d_2 and e_2 of the mask and the window size $d_x \times d_y$ have been set for the moment empirically. We can notice that an increase of d_x and d_y leads to a coarse and faster detection of the text, while the increase of d_1 and e_1 detect better text areas agglomerate together. The increase of d_2 and e_2 enhances the smoothing effect on noisy images. The experiments that we conducted on images of different types of documents show that the stability result can be achieved by respecting the following convention: $e_1=d_1=2$, to detect areas of texts and $e_1=d_1=1$, to detect words. The ordered succession of these elementary operations gives morphological openings and closings that filter all sorts of responses of gradient to the texture, background noise and even the defects of paper.

The morphological processing is continued by a fisher global thresholding section 2.1.1 that gives a binary mask F which contains the different blocks of the textual zones. This method quickly calculates a global threshold from the histogram of the greyscale image M . This binary mask is used to direct the local thresholding in full resolution to the zones of text and may be considered as a first segmentation into blocks of physical structure of the image of document.

This rapid emphasizing on blocks plays two important roles in terms of computation time: on one hand, it can effectively reduce the local thresholding time to make it almost similar to that of global thresholding on the other hand; it speeds up the extraction phase of the physical structure that we will see later on in details.

3.1. Final Binarization Issued from the Local Thresholding of Regions of Interest

To obtain a binary map B of the foreground in full resolution, we decided to use the method of Sauvola

for its rapidity with respect to the other local methods see Table 1 and for its performances (the Wolf method is specific to videos and is not suitable for our application). This local thresholding is applied only on textual zones located in the mask F shows that Figure 4. The saved time allowed us to use a large size for the window (21×21 for example) which allows obtaining very good results on printed documents and manuscripts with largely varying sizes of characters.

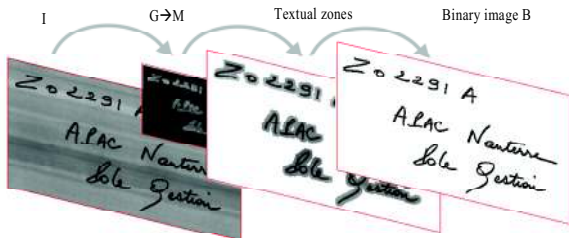


Figure 4. Steps of the hybrid binarization (local thresholding/localization of the textual zones).

4. Results and Evaluation of Our Method

The targeting of the text areas allowed us to binarize the image with a reduced time similar to those of the global methods, while using the performance of local methods.

4.1. Comparison of Computed Times

To compare the time of our binarization method with those of global methods (Otsu) and local methods (Sauvola, Niblack and Wolf), we used a database of 29225 images of documents and internal mails. The database is divided into 9 categories: circulating checks (CCH), IASP, Visa Card (CB), Listings A3 (LA3), Listings A4 (LA4), Forms (FMR), Planus (PLN), Handwritten Internal Mail (HIM), Printed Internal Mail (TIM).

The curves in the Figure 5 shows the average time elapsed to binarize the documents of each of the 9 categories by our mixed method then by the global method of OTSU and the local methods of Sauvola, Niblack and Wolf.

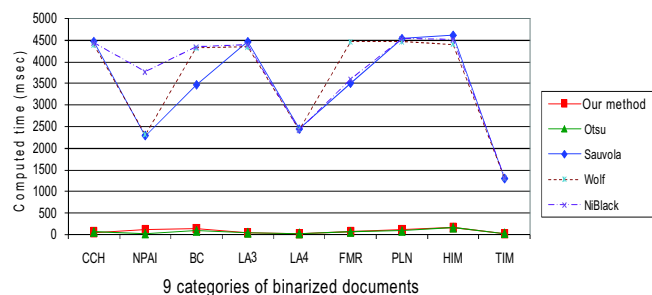


Figure 5. Comparison of the binarization time of our method in relation to different methods.

The comparison of times in the above figure shows that the average computation time of our hybrid method (thresholding / localization) of binarization are

similar to those of global methods and much less smaller than local methods.

4.2. Evaluation of the Binary Characters Quality

To assess the quality of the binary characters offered by our joint binarization method, we applied the OCR on the images of the same base binarized by our method and then binarized by the conventional method of Sauvola. The following table shows the increase of the rates of OCR by our mixed method with respect to the method of Sauvola. This means that our binarization method preserves perfectly the quality of the characters compared to other methods. The increase in OCR-ratio by our mixed binarization method is given by : (Categories, Increase in OCR-ratio)={ (CCH,+2%) (NPAL,+26%) (CB,+13%) (LA3,+11%) (LA4,+11%) (FMR,+23%) (PLN,+20%) (HIM,+76%) (TIM,+16%) }. Figure 6 shows that our mixed thresholding method offers a binarization of better quality compared to the global methods and local methods.

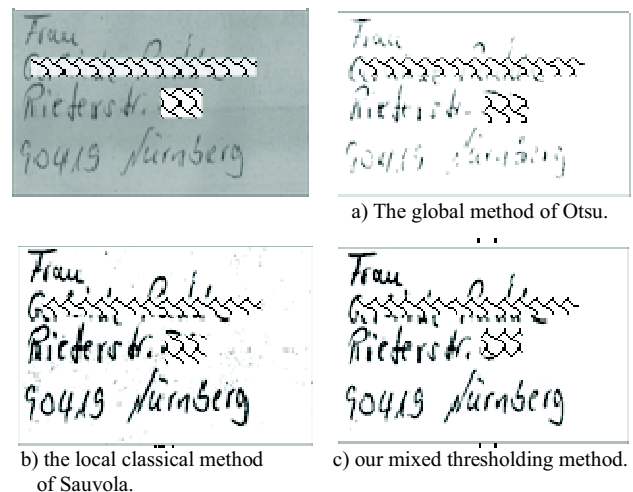


Figure 6. Result of binarization of an address block.

4.3. Interest of the Binarization Approach on the Detection Step of the Connected Components

In addition to the advantages that we just explained, our hybrid method of thresholding has also reduced the computation time of the connected components by the reduction of black pixels in all large black areas (that most often correspond to pictures or publicity indications) that are located as black edges with white centres. The Figure 7 shows an example on a concrete case. Our tests have shown that this binarization approach allowed dividing by three the time of detection of CCs in relation to that of detection of CCs obtained by the exclusive use of the algorithm of Sauvola.

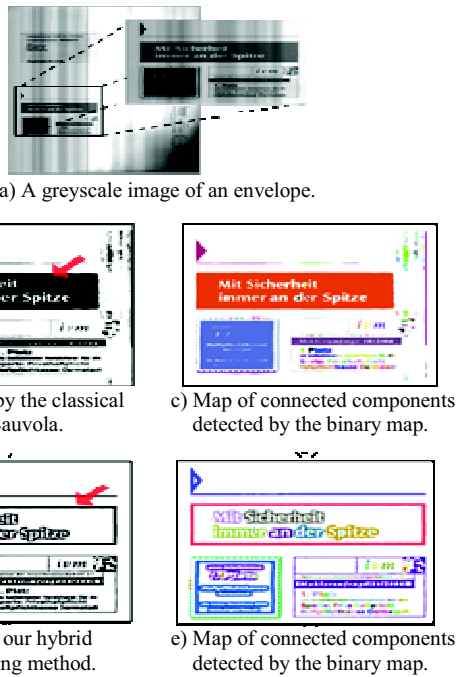


Figure 7. Example on concrete case.

5. Conclusions and Assessment

We have presented in this paper a new mixed method of binarization adapted to real time constraints and precision. We showed how our method significantly reduces the number of iterations typically used in a conventional binarization approach (global or adaptive). It begins by locating the zones of texts from the estimation of the cumulated gradient and morphological operations on low resolution.

This step is followed by local thresholding of type Sauvola focusing exclusively on those areas with high gradients in the version of the images at full resolution. The binarization process can be regarded as a first low level crude segmentation step, result of the binary maps issued from the gradient images that are thresholded according to the criterion of Fisher. A first cooperation between the detections of regions of interest step and directed thresholding step is proposed in this binarization. We showed how our method greatly reduces the computation time by exploiting the gradients thus saving the exhaustive analysis of fully filled regions (large black areas often visible on the logos that are present on the envelopes of letters). The detection step of the connected components is considerably accelerated. This work is adopted and used by CESA Company (www.cesa.fr). A robust approach of address block localization in business mail by graph coloring is published in the IAJIT journal [13].

References

- [1] Babaguchi N., "Connectionist Model Binarisation," in *Proceedings of 10th International Conference on Pattern Recognition*, NJ, vol. 2, pp. 51-56, 1990.
- [2] Badekas E. and Papamarkos N., "A System for Document Binarisation," in *Proceedings of 3rd International Symposium on Image and Signal Processing and Analysis*, vol. 2, pp. 909-914, 2003.
- [3] Bruzzese D. and Giani U., "Automatic Multilevel Thresholding Based on a Fuzzy Entropy Measure," *Studies in Classification, Data Analysis, and Knowledge Organization*, vol. 3, pp. 125-133, 2011.
- [4] Chang M., "Improved Binarization Algorithm for Document Image by Histogram and Edge Detection," in *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Montreal, Que, vol. 2, pp. 636-639, 1995.
- [5] Chi Z. and Wong K., "A Two-Stage Binarization Approach for Document Images," in *Proceedings of International Symposium on Intelligent Multimedia, Video and Speech Processing*, Hong Kong, pp. 275-278, 2001.
- [6] Chigusa Y., "An Image Binarization System for Composite Pictures," *International Symposium on Circuits and Systems*, CA, vol. 5, pp. 2292-2295, 1992.
- [7] Cocquerez J. and Philipp S., *Analyse D'Images: Filtrage et Segmentation*, Edition Masson, 1995.
- [8] Couto P., Bustince H., Pagola M., Jurio A., and Melo-Pinto P., "A-IFSs Entropy Based Image Multi-Thresholding," in *Proceedings of Trends in Applied Intelligent Systems, Lecture Notes in Computer Science*, vol. 6098, pp. 341-349, 2010.
- [9] Esquef I. and Albuquerque M., "Nonextensive Entropic Image Thresholding," in *Proceedings of XV Brazilian Symposium on Computer Graphics and Image Processing*, pp. 402, 2002.
- [10] Feng M. and Tan Y., "Adaptive Binarization Method for Document Image Analysis Multimedia and Expo," in *Proceedings of IEEE International Conference on Multimedia and Expo*, Taipei, vol. 1, pp. 339-342, 2004.
- [11] Fisher W., "On Grouping for Maximum Homogeneity," *Journal of the American Statistical Association*, vol. 53, no. 284, pp. 789-798, 1958.
- [12] Fontanot P. and Ramponi G., "A Polynomial Filter for the Preprocessing of Mail Address Images," in *Proceedings of IEEE Winter Workshop on Nonlinear Digital Signal Processing*, Finland, pp. 21-26, 1993.
- [13] Gaceb D., Eglin V., Lebourgeois F., and Emptoz H., "Robust Approach of Address Block Localization in Business Mail by Graph Coloring," *The International Arab Journal of Information Technology*, vol. 6, no. 3, pp. 221-229, 2009.

- [14] Hamza H., Smigiel E., and Belaid E., "Neural Based Binarization Techniques," in *Proceedings of 9th International Conference on Document Analysis and Recognition*, vol. 1, pp. 317-321, 2005.
- [15] He J., Do Q., Downton A., and Kim J., "A Comparison of Binarization Methods for Historical Archive Documents," in *Proceedings of 8th International Conference on Document Analysis and Recognition*, vol. 1, pp. 538-542, 2005.
- [16] Kamada H. and Fujimoto K., "High-Accuracy Binarization Method for Recognizing Text in Images of Low Spatial Resolutions," in *Proceedings of 5th International Conference on Document Analysis and Recognition*, Bangalore, pp.139-142, 1999.
- [17] Kapur J., Sahoo P., and Wong A., "A New Method for Gray-Level Picture Thresholding Using the Entropy of the Histogram," *Graphical Models Image Process*, vol. 29, no. 3, pp. 273-285, 1985.
- [18] Köhler R., "A Segmentation System Based on Thresholding," *Graphical Models and Image Processing*, vol. 15, no. 4, pp. 319-338, 1981.
- [19] Kuk J., Cho N., and Lee K., "Map-Mrf Approach for Binarization of Degraded Document Image," in *Proceedings of 15th IEEE International Conference on Image Processing*, San Diego, CA, pp. 2612-2615, 2008.
- [20] LeBourgeois F. and Emptoz H., "Document Analysis in Gray Level and Typography Extraction Using Character Pattern Redundancies," in *Proceedings of 5th International Conference on Document Analysis and Recognition*, India, pp. 177-180, 1999.
- [21] Lelore T. and Bouchara F., "Document Image Binarisation Using Markov Field Model," in *Proceedings of 10th International Conference on Document Analysis and Recognition*, Barcelona, pp. 551-555, 2009.
- [22] Leydier Y., LeBourgeois F., and Emptoz H., "Sérialisation Du K-Means Pour La Segmentation Des Images en Couleur," in *Proceedings of Semaine du Document Numérique Conférence Internationale Francophone sur l'Ecrit et le Document*, pp. 191-196, 2004.
- [23] Niblack W., "An Introduction to Digital Image Processing," *Prentice Hall*, pp.115-116, 1986.
- [24] Otsu N., "A Threshold Selection Method From Gray-Level Histograms," *IEEE Transaction on System, Man, and Cybernetics*, vol. 9, no. 1, pp. 62-66, 1979.
- [25] Philipp-Foliguet S., "Evaluation de la Segmentation," *Rapport Technique*, 2001.
- [26] Ramponi G. and Fontanot P., "Enhancing Document Images with a Quadratic Filter," *Signal Process*, vol. 33, no. 1, pp. 23-34, 1993.
- [27] Sauvola J., "Adaptive Document Binarisation," in *Proceedings of 4th International Conference on Document Analysis and Recognition*, Ulm, vol. 1, pp. 147-152, 1997.
- [28] Savakis A., "Adaptive Document Image Thresholding Using Foreground and Background Clustering Image Processing," in *Proceedings of International Conference on Image Processing*, Chicago, vol. 3, pp. 785-789, 1998.
- [29] Sezgin M. and Sankur B., "Selection of Thresholding Methods for Nondestructive Testing Applications," in *Proceedings of International Conference on Image Processing*, Thessaloniki, vol. 3, pp. 764-767, 2001.
- [30] Shan M. and Mathews J., "Adaptive Quadratic Preprocessing of Document Images for Binarisation Image Processing," *IEEE Transactions*, vol. 7, no. 7, pp. 992-999, 1998.
- [31] Tanaka H., "Threshold Correction of Document Image Binarization for Ruled-line Extraction," in *Proceedings of 10th International Conference on Document Analysis and Recognition*, Barcelona, pp. 541-545, 2009.
- [32] Thillou C. and Gosselin B., "Combination of Binarization and Character Segmentation using Color Information," in *Proceedings of the 4th IEEE International Symposium on Signal Processing and Information Technology*, pp.107-110, 2004.
- [33] Trier O. and Jain A., "Goal-Directed Evaluation of binarization methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI*, vol. 17, no. 12, pp. 1191-1201, 1995.
- [34] Trier O. and Taxt T., "Improvement of 'Integrated Function Algorithm for Binarization of Document Images," *Pattern Recognition Letters*, vol. 16, no. 3, pp. 277-283. 1995.
- [35] Valverde J. and Grigat R., "Optimum Binarization of Technical Document Images," in *Proceedings of International Conference on Image Processing*, Vancouver, vol. 3, pp. 985-988, 2000.
- [36] Wolf C., Jolion J., and Chassaing F., "Text Localization, Enhancement and Binarization in Multimedia Documents," in *Proceedings of 16th International Conference on Pattern Recognition*, Canada, vol. 4, pp. 1037-1040, 2002.
- [37] Wu S. and Amin A., "Automatic Thresholding of Gray-Level Using Multistage Approach," in *Proceedings of 7th International Conference on Document Analysis and Recognition*, USA, vol. 1, pp. 493-497, 2003.
- [38] Yang F., Ma Z. and Xie M., "A Novel Binarization Approach for License Plate," in

Proceedings of 1st IEEE Conference Industrial Electronics and Applications, Singapore, pp. 1-4, 2006.

- [39] Yanowitz S. and Bruickstein A., "A New Method for Image Segmentation," *Computer Vision, Graphics and Image Processing*, vol. 46, no. 1, pp. 82-95, 1989.
- [40] Yao-Hong T., "A New Approach for Image Thresholding under Uneven Lighting Conditions," in *Proceedings of 6th IEEE International Conference on Computer and Information Science*, Melbourne, pp. 123-127, 2007.
- [41] Zhang Y., "A Survey on Evaluation Methods for Image Segmentation," *Pattern Recognition*, vol. 29, no. 8, pp. 1335-1346, 1996.



Frank Lebourgeois is graduated in 1987 from University of Lyon I with a MS degree of science in mathematics then he gets a PhD in 1992 at INSA de Lyon in computer sciences. He is currently an assistant professor in the LIRIS laboratory and works on documents images restoration and analysis.



Djamel Gaceb received the engineer degree in electronics, signal and image processing Laboratory, from BLIDA University, Algeria, in 2002, and the MSc degree in computer science from Lyon I University, France, in 2005. From 2006 to 2009 he worked on automatic mail sorting system at CESA company. He received his PhD degree in 2009 from INSA of Lyon in computer sciences. He is working since September 2009 as temporary teaching and research assistant (ATER) at the Lyon I University and is attached since 2005 to the LIRIS UMR 5205 laboratory. He is currently working on the topic of mail sorting and document recognition, retrieval and analysis, iPhone vision and real time applications.



Véronique Eglin is graduated from the INSA of Lyon in 1995 and holder in 1998 of the PhD in computing science on "the document structure analysis", she is working since September 2000 as an associate professor in the INSA of Lyon and is attached since 2003 in the LIRIS UMR 5205 laboratory. She is contributing since 2003 to different digitization and valorisation projects of cultural inheritance. Her research domains are essentially centred on the characterization and the classification of handwritten and printed documents, the writer identification, the texture analysis for the typographies, and for documents layouts characterization.