A HEURISTIC BINARIZATION ALGORITHM FOR DOCUMENTS WITH COMPLEX BACKGROUND*

George D. C. Cavalcanti[1,2], *Eduardo F. A. Silva*[1], *Cleber Zanchettin*[1,2], *Byron L. D. Bezerra*[1,2], *Rodrigo C. Dória*[1] and *Juliano C. B. Rabelo*[1]

AiLeader Technologies - Recife - Pernambuco - Brazil[1]
Federal University of Pernambuco (UFPE) - Centro de Informática (CIn)[2]
Recife - Pernambuco - Brazil

{gdcc,efas,cz,bldb,rcd,jcbr}@aileader.com.br, {gdcc,cz,bldb}@cin.ufpe.br

ABSTRACT

This paper proposes a new method for binarization of digital documents. The proposed approach performs binarization by using a heuristic algorithm with two different thresholds and the combination of the thresholded images. The method is suitable for binarization of complex background document images. In experiments, it obtained better results than classical techniques in the binarization of real bank checks.

***Index Terms*—** Binarization, Documents with complex background, Document image processing, Automatic bank check processing

1. INTRODUCTION

A document image contains text, symbols and graphics. The starting step of most document image analysis systems refers to the conversion of the gray-scale image to a binary image. In many practical applications, the analyzed document image has poor quality, shadows, nonuniform illumination, low contrast, large signal-dependent noise, smear and strain. These characteristics determine a difficult separation between the image background and the interest area [1, 4, 2].

Binarization plays a key role in document processing since its performance critically affects the success of the image characters segmentation and recognition [3]. However, the binarization process often introduces noise, and many of the difficulties in reading documents are due to poor thresholding [4]. Another problem caused by binarization is the generation of broken or touched strokes [5].

There are two classes of binarization techniques: global and local thresholding. Global thresholding algorithms use a unique threshold per image. On the other hand, local approaches compute a separate threshold based on the pixel neighborhood. When there is a good separation between the background and the characters, global thresholding algorithms usually achieve a satisfactory efficacy. However, many document images have complex backgrounds, what makes the separation not so simple. Local or adaptive thresholding present a better performance when treating documents with complex backgrounds. By contrast, one weakness of the adaptive thresholding algorithms is related to the problem of preserving stroke connectivity [6].

Many binarization techniques have been developed over the years [7], but all of them aimed to be a generic approach to deal with many kinds of documents. In this work, a new heuristic binarization approach to deal with complex backgrounds is presented. While global thresholding algorithms are not good enough to treat complex backgrounds and local approaches do not preserve stroke connectivity (something essential for OCR and ICR applications), the proposed approach successfully removes the background, yet keeping stroke connectivity untouched. The proposed method was tested with complex background images of real Brazilian bank checks. The experimental and comparative results confirm the effectiveness of the method*.

The rest of the paper is organized as follows: section 2 describes the binarization approach. In section 3, the experimental study is described. Section 4 presents the final considerations.

2. BINARIZATION APPROACH

Global thresholding techniques aim to find a unique threshold to eliminate all pixels of the image background, while preserving all pixels of the image foreground. Unfortunately, many images present complex backgrounds or weak image foregrounds (some foreground pixels have gray values very close to those of some background pixels). In such cases, it is not possible to find a single threshold that completely separates the foreground image from the background. Thus, if

*This proposed approach is registered under the iBinarize® software intellectual property submitted to INPI-DEPE (Brazil) in 05/24/2006 through the protocol number 000822.

the algorithm decides to eliminate all the background, some of the foreground will also be eliminated, generating broken images and not preserving the stroke connectivity.

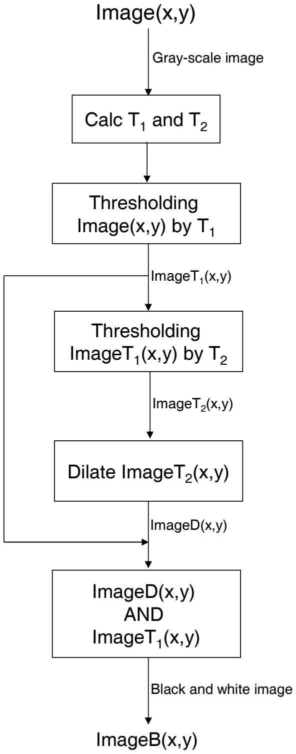

Fig. 1. Block diagram of the proposed approach to binarize images with complex backgrounds.

To try to solve that problem, a heuristic binarization algorithm was developed using two threshold values, T_1 and T_2, and the combination of the thresholded images. The key idea is that those pixels whose gray levels are greater than T_1 are directly eliminated. The pixels with gray levels below T_2 are preserved and the ones with a value between T_1 and T_2 are preserved if they are located near a pixel with a gray value below T_2, and deleted otherwise.

The proposed approach to binarize gray-level images with complex backgrounds is illustrated by the block diagram in Figure 1.

To perform the binarization, the first step is to calculate the two thresholds values T_1 and T_2. The thresholds are estimated based on the original image histogram - $Image(x, y)$. The threshold procedure aims to split the image in two parts: background and foreground. The main idea is to estimate the gray-level intensities of the background T_1 and to estimate the intensity of the foreground T_2.

In Figure 2, a gray-level image and the image histogram are presented. T_1 is the first gray-level intensity of the histogram. It is considered the first intensity with a frequency greater than $np = \frac{width \times height}{n}$. In the image histogram, ΔT represents the minimum distance of gray-level intensity between the background T_1 and the foreground T_2. After empirical tests, we adopted $n = 350$ and $\Delta T = 40$.

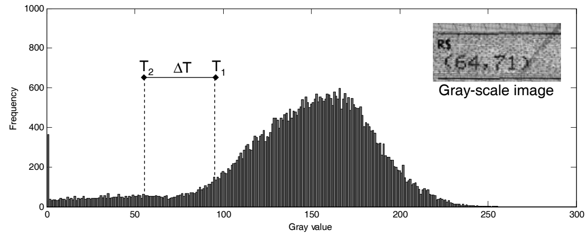

Fig. 2. Image histogram and the threshold values T_1 and T_2.

After the thresholds estimation, all pixels in the original image greater than T_1 (background intensity threshold) are set to white. This procedure eliminates most of the image background, generating $ImageT_1(x, y)$.

The resulting $ImageT_1(x, y)$ still has some noise, but all foreground information is preserved. To eliminate the remaining background noise, the pixels in $ImageT_1(x, y)$ greater than T_2 (foreground intensity threshold) are set to white, generating $ImageT_2(x, y)$.

The $ImageT_2(x, y)$ contains only the real area of interest, all the background is eliminated, but some of the foreground pixels were also eliminated in the process. To preserve the stroke connectivity and restore broken images, some pixels of $ImageT_1(x, y)$ are recovered. The recovered pixels are selected depending on their distance to the remaining pixels of $ImageT_2(x, y)$. The pixel recovering process is implemented by a morphological binary dilation operation (Equation 1) in $ImageT_2(x, y)$ followed by a binary AND operation (Equation 2) between the dilated image and $ImageT_1(x, y)$.

$$A \oplus B = \{x | [(\hat{B})_x \cap A] \subseteq A\}, \tag{1}$$

where: A and B are two sets in Z^2 (in particular, two bidimensional images). $(B)_x$ represents the translation of B by $x = (x_1, x_2)$, given by $(B)_x = \{c | c = b + x, \forall b \in B\}$ and $\hat{B}$ is the reflection of B, given by $\hat{B} = \{x | x = -b, \forall b \in B\}$. The set B is often called structuring element of the dilation and, in this work, a mask of 3 × 3 was used.

When dilation is applied to binary images it causes the original objects to grow larger. As it can be seen in Equation 1, that operation can be implemented in to parts: first, marking all white pixels which have at least one black neighbor, and then setting the marked pixels to black.

$$A \wedge B = \begin{cases} black, & \text{if } A(x,y) = B(x,y) = black \\ white, & \text{otherwise} \end{cases} \tag{2}$$

The dilation process is used to delimitate the neighborhood area of the remaining pixels in $ImageT_2(x, y)$. The binary AND is used to recover those pixels of $ImageT_1(x, y)$

that are in the neighborhood area of the pixels in $ImageT_2(x, y)$. This step improves the image quality and preserves the stroke connectivity.

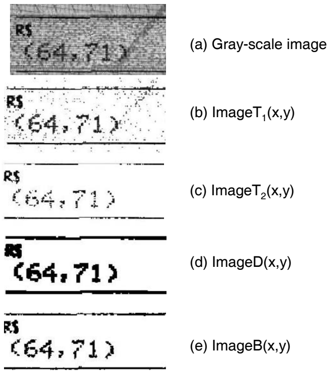

Fig. 3. Step-by-step images generated by the proposed method.

In Figure 3 the input image and the images generated in each step of the proposed algorithm are shown. (a) Gray-level image; (b) Thresholded T_1 image, most of the background has been eliminated, but some noise remains; (c) Thresholded T_2 image, only foreground pixels remain, but the stroke connectivity is not preserved; (d) Dilated $ImageT_2(x, y)$, open regions are closed and digits are connected; (e) Final image, all background eliminated and foreground preserved.

3. EXPERIMENTAL STUDY

The proposed binarization approach was tested on real bank check images. Bank checks are a class of document images difficult to be automatically recognized. Noise on the checks makes accurate segmentation and recognition very difficult. The relatively large size of the character set, the wide variety of fonts and sizes, handwritten styles, cultural characteristics and very cursively written scripts, the shape similarity of different characters, and the ambiguity of the grammar are some factors that contribute to make the manual character annotation and the automatic training and recognition processes very difficult. Moreover, a check possesses a complex background with considerable variations in color, pictures, stylistic characters or symbols, institutional logos, security graphics and overlapping images, which also compromises the success of the binarization process.

The bank check layout can be partitioned into interest regions, like the courtesy amount, legal amount, payee's name, handwritten signatures and MICR CMC-7 area. In experiments with the binarization techniques, a total of 1,350 gray-scale images of real Brazilian bank checks were used. The experiments consist in comparing the accuracy rate of the proposed approach against two well-known binarization algorithms: (1) the global thresholding algorithm Otsu [8], and (2) the local adaptive threshold algorithm Niblack [9]. Based on a previous study [3], the Niblack algorithm was implemented with the following parameters values: a 60×60 window, that aims to cover 1-2 characters, and $k = -0.2$.

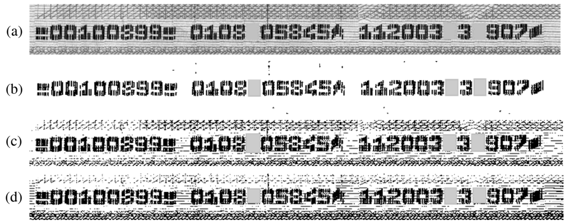

Fig. 4. CMC-7 Binarization. (a) Source image. (b) Proposed method. (c) Otsu. (d) Niblack.

After visual inspection, the proposed method outperforms all algorithms tested in terms of image quality and stroke connectivity. Figures 4 and 6 show some samples of CMC-7 and courtesy amount areas binarized by the three algorithms. It is possible to see that the Otsu and Niblack algorithms did not perform well on such images. These approaches could not eliminate the background correctly, leaving a great amount of background noise.

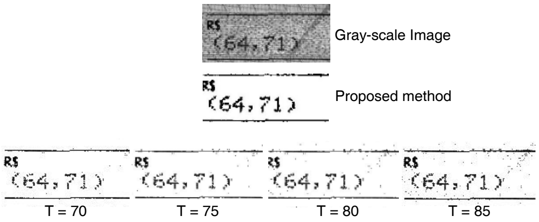

Fig. 5. The proposed approach versus different global thresholds (T).

Based on the images, Otsu (which is a global thresholding approach) achieved better results than an adaptive Niblack algorithm. Thus, another test was done aiming to select the best global threshold and validate if it was better than the proposed method. Figure 5 shows one courtesy amount image that was binarized by the proposed approach and four other images that were binarized by global threshold T. When $T = 70$ the digits are broken. Increasing the value of T, the amount of noise also increases. After several visual inspections, we could verify that the proposed approach is better than any global threshold.

Two interest areas of these checks were considered in the recognition study: courtesy amount and CMC-7. After the binarization procedure, the images were classified by an engine of automatic bank checks recognition produced by AiLeader Technologies©[1].

In Table 1, it is possible to observe the recognition rates achieved by the engine using the three binarization algorithms.

[1] www.aileader.com.br

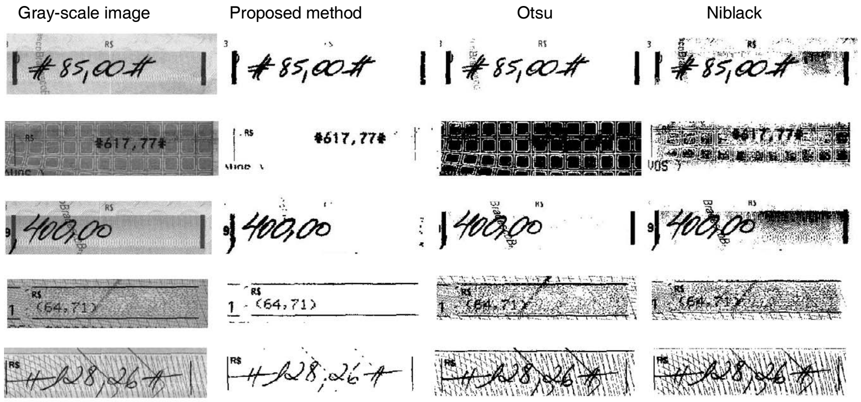

Fig. 6. Courtesy amount binarization by the proposed approach, Otsu and Niblack.

Check Field	Otsu	Niblack	Proposed
CMC-7	67.25%	40.75%	**89.58%**
Courtesy Amount	31.75%	10.37%	**57.36%**

Table 1. Recognition accuracy rates using three different thresholding approaches.

In all cases, the engine reached the best recognition rates when the proposed binarization approach was used. In the CMC-7 area, the recognition engine achieved a recognition rate of 89.58% using the proposed method, 67.25% using Otsu and 40.75% using Niblack. In the courtesy amount recognition, the engine achieved 57.36% of correct classification using the proposed method.

4. FINAL REMARKS

In this paper, a new heuristic binarization algorithm was proposed to treat documents with complex backgrounds. The database used in the tests was composed by real images of Brazilian bank checks, which usually contain very complex structures on the background. The proposed method was compared against two well-known binarization algorithms: Otsu (global thresholding) and Niblack (local thresholding). The results demonstrated that the proposed method outperforms the other approaches, mainly due to the incapability of the Otsu and Niblack algorithms to produce the desired separation between background and foreground.

Further research will focus on evaluating the proposed method against other binarization methods and validating it over other kinds of images, such as: historical handwritten documents, old newspapers, mail envelopes and maps.

5. REFERENCES

[1] C. Y. Suen, L. Lam, D. Guillevic, N. W. Strathy, M. Cheriet, J. N. Said, and R. Fan, “Bank check processing system,” *International Journal of Imaging Systems and Technology*, vol. 7, no. 4, pp. 392403, 1996.

[2] Shyang-Lih Chang, Li-Shien Chen, Yun-Chung Chung, and Sei-Wan Chen, “Automatic license plate recognition,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 1, pp. 42–53, 2004.

[3] B. Gatos, I. Pratikakis, and S.J. Perantonis, “Adaptive degraded document image binarization,” *Pattern Recognition*, vol. 39, pp. 317–327, 2006.

[4] H. P. Graf, C. J. C. Burges, E. Cosatto, and C. R. Nohl, “Analysis of complex and noisy check images,” in *International Conference on Image Processing*, 1995, pp. 316–319.

[5] Xiangyun Ye, Mohamed Cheriet, Ching Y. Suen, and Ke Liu, “Extraction of bankcheck items by mathematical morphology,” *International Journal on Document Analysis and Recognition*, vol. 2, no. 2-3, pp. 53–66, 1999.

[6] Il-Seok Oh, “Document image binarization preserving stroke connectivity,” *Pattern Recognition Letters*, vol. 16, no. 7, pp. 743–748, 1995.

[7] Oivind Due Trier and Anil K. Jain, “Goal-directed evaluation of binarization methods,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 12, pp. 1191–1201, 1995.

[8] N. Otsu, “A threshold selection method from gray-level histograms,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[9] Wayne Niblack, *An introduction to digital image processing*, Strandberg Publishing Company, Birkeroed, Denmark, Denmark, 1985.