A Recursive Thresholding Technique for Image Segmentation

M. Cheriet, J. N. Said, and C. Y. Suen

***Abstract*— In this correspondence, we present a general recursive approach for image segmentation by extending Otsu's method. The new approach has been implemented in the scope of document images, specifically real-life bank checks. This approach segments the brightest homogeneous object from a given image at each recursion, leaving only the darkest homogeneous object after the last recursion. The major steps of the new technique and the experimental results that illustrate the importance and the usefulness of the new approach for the specified class of document images of bank checks will be presented.**

***Index Terms*—Grey-scale images, image enhancement, image segmentation.**

I. Introduction

The problem of *image segmentation* has received considerable attention in the literature [12], [17], [20], [21]. Several methodologies have been proposed to tackle this problem, however, two approaches are widely used in this context: edge-based-like [4], [9], [22] and region-based-like [1], [2], [7], [8], [16]–[18]. In edge-based methods, the local discontinuities are detected first and then connected to form longer, hopefully complete, boundaries. In region-based methods, areas of an image with homogeneous properties are found, which in turn give the boundaries. The two methods are complementary, and one may be preferred to the other for some specific applications like document image analysis. In their remarkable work, Sahoo *et al.*, Pal *et al.*, and Trier *et al*. presented an excellent survey on image thresholding that tested the performance of many automatic global and thresholding techniques.

In fact, there is currently a substantial and growing interest in the field of document image understanding where several research groups are developing and designing systems to automatically process and extract relevant information from documents such as engineering drawings, maps, magazines, newspapers, forms, and mail envelopes [3], [5], [6], [13]–[15], [24], [25].

Our interest in this work is to develop a general recursive thresholding technique for grey-scale images, and demonstrate the need for such a new technique in the field of document processing for document analysis and understanding. Of course, the new technique also aims at preserving the topological properties of the extracted information that will be used for further processing.

Manuscript received November 20, 1995; revised July 28, 1997. This work was supported by the IRIS National Networks of Centres of Excellence of Canada, the Natural Sciences and Engineering Research Council of Canada, the Ministry of Education of Québec, and Bell Québec. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Patrick A. Kelly.

M. Cheriet is with the Laboratoire d'Imagerie, de Vision et d'Intelligence Artificielle, École de Technologie Supérieure de l'Université du Québec 1100, Montréal, P.Q. H3C 1K3 Canada (e-mail: cheriet@gpa.etsmtl.ca).

J. N. Said is with the Centre for Pattern Recognition and Machine Intelligence, Concordia University, Montréal, P.Q. H3G 1M8 Canada.

C. Y. Suen is with the Centre for Pattern Recognition and Machine Intelligence, Concordia University, Montréal, P.Q. H3G 1M8 Canada (e-mail: suen@cenparmi.concordia.ca).

Publisher Item Identifier S 1057-7149(98)04005-6.

II. General Scope of the Method

In segmenting several objects from a given image, the determination of multilevel threshold values should not be constrained by the number of objects in that particular image. For example, [16] is adequate when the image has exactly three or fewer objects. Moreover, [18] is based on statistical and spatial information requiring that the histogram consists of only Gaussian distributions. Despite the fact that grey-level distributions, small objects, and object overlapping are some of the most complicated issues that create several challenging difficulties for multilevel threshold selection in images, a thresholding technique must be able to segment a digitized image into different objects with similar properties. In this correspondence, we will present an extension of Otsu's approach [11] as a general technique for image segmentation and limit the scope of the problem to document images. The new approach goes beyond segmenting only one bright object from an image to an approach that recursively segments the brightest object at each recursion, leaving the darkest object in the given digitized image. The recursive method is developed without any constraints on the number of objects in the digitized image. Section III will briefly present Otsu's method, Section IV will present the general scope of the recursive thresholding technique, and Sections V and VI will present the experimental results and the conclusion of the new approach on document images as a particular application.

III. Otsu's Approach

This method, as proposed in [11], is based on discriminant analysis. The threshold operation is regarded as the partitioning of the pixels of an image into two classes C_0 and C_1 (e.g., objects and background) at grey-level t, i.e., $C_0 = \{0, 1, \cdots, t\}$ and $C_1 = \{t+1, t+2, \cdots, l-1\}$. As stated in [11], let σ_W^2, σ_B^2, and σ_T^2 be the within-class variance, between-class variance, and the total variance, respectively. An optimal threshold can be determined by minimizing one of the following (equivalent) criterion functions with respect to t:

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2}, \quad \eta = \frac{\sigma_B^2}{\sigma_T^2}, \quad k = \frac{\sigma_T^2}{\sigma_W^2}.$$

Of the above three criterion functions, η is the simplest. Thus, the optimal threshold t^* is defined as

$$t^* = \operatorname*{ArgMin}_{t \in G} \eta \tag{1}$$

where

$$\sigma_T^2 = \sum_{i=0}^{l-1} (i - \mu_T)^2 P_i, \quad \mu_T = \sum_{i=0}^{l-1} i P_i \tag{2}$$

$$\sigma_B^2 = w_0 w_1 (\mu_1 \mu_0)^2, \quad w_0 = \sum_{i=0}^{t} P_i, \quad w_1 = 1 - w_0 \tag{3}$$

$$\mu_1 = \frac{\mu_T - \mu_t}{1 - \mu_0}, \quad \mu_0 = \frac{\mu_t}{w_0}, \quad \mu_t = \sum_{i=0}^{t} i P_i, \quad P_i = \frac{n_i}{n} \tag{4}$$

where n_i is the number of pixels with grey-level i and n is the total number of pixels in a given image defined as $n = \sum_{i=0}^{l-1} n_i$. Moreover, P_i is the probability of occurrence of grey-level i defined as $P_i = \frac{n_i}{n}$.

Otsu's method as proposed affords further means to analyze further aspects other than the selection of the optimal threshold for a given image. For a selected threshold t^* of a given image, the class probabilities w_0 and w_1 indicate the portions of the areas occupied

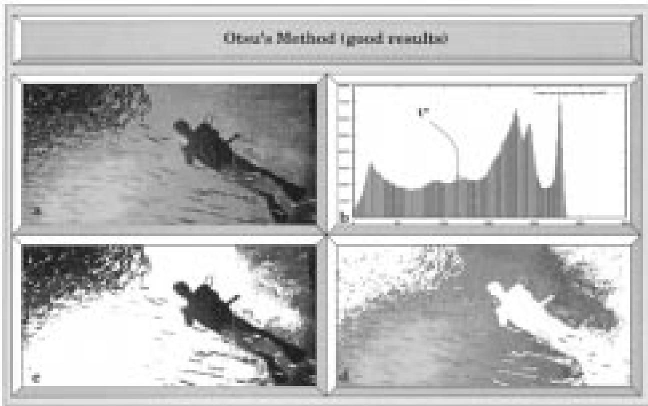

Fig. 1. Image segmentation using Otsu's method on an image with two objects. (a) Original grey-scale image. (b) Histogram where t^* is the optimal threshold. (c) Final result (class C_1). (d) Thresholded background (class C_0).

by the classes C_0 and C_1. The class means μ_0 and μ_1 serve as estimates of the mean levels of the classes in the original grey-level image. Moreover, the maximum value of η, denoted by η^*, can be used as a measure to evaluate the separability of classes C_0 and C_1 in the original image or the bimodality of the histogram. This is a very significant measure because it is invariant under affine transformations of the grey-level scale. It is uniquely determined within the range

$$0 \leq \eta \leq 1.$$

The lower bound (zero) is obtained when and only when a given image has a single constant grey level, and the upper bound (unity) is obtained when and only when two-valued images are given. This property is an important criteria that we will use in extending Otsu's approach.

In digital images, the uniformity of objects plays a significant role in separating these objects from the background. In fact, Otsu's method in thresholding grey-level images is efficient on the basis of uniformity measure between the two classes C_0 and C_1 that should be segmented. Fig. 1 illustrates the results after applying Otsu's method on the given image.

IV. Extension of Otsu's Approach (Dynamic Thresholding)

In the previous section we presented Otsu's method, which performs well on images with two classes. However, Fig. 2(c) illustrates the limitation of this method. In fact, Fig. 2 shows that if an image has more than two objects, then by applying [11] only one object is thresholded. Hence, we propose thresholding recursively and proceed in the following way: according to [11], a global threshold value is evaluated for a given image, Fig. 3(a), and a new image with one object (the brightest) being thresholded is produced, Fig. 3(c). Now, the regions, as in Fig. 3(c), are considered as the given image and the process of histogramming, peak selection, and thresholding as in [11] is recursively repeated until no new peaks can be found or regions become too small. In this sense, an effective recursive segmentation technique that segments the brightest object at each recursion can be achieved.

Formally, let N be the set of natural numbers, (i, j) be the spatial coordinates of a digitized image, and $G = \{0, 1, \cdots, l-1\}$ be a set of positive integers representing grey levels.[1] Then, an image function can be defined as the mapping

$$f : N \times N :\rightarrow G$$

[1] In our model, we assume $l = 256$.

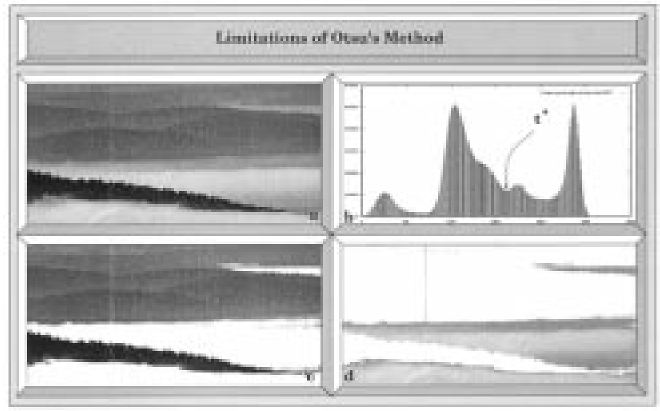

Fig. 2. Image segmentation using Otsu's method on an image with multiobjects.(a) Original grey-scale image. (b) Histogram where t^* is the optimal threshold. (c) Final image containing the darkest object. (d) Thresholded background (brightest object).

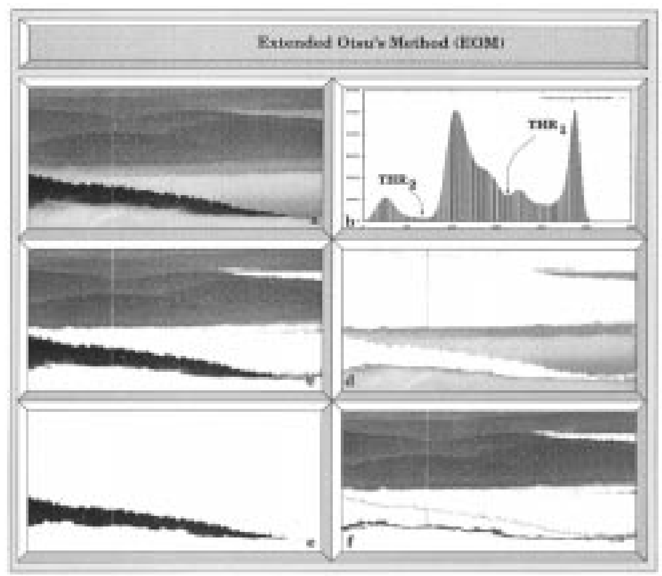

Fig. 3. Image segmentation using recursive method on images with multiobjects. (a) Original grey-scale image. (b) Histogram where THR_1 and THR_2 are the optimal thresholds during the first and the second recursions. (c) Image after thresholding the first object (C_0) using THR_1. (d) First thresholded background object C_0. (e) Image f_p that contains the darkest object. (f) Second thresholded background object C_1.

where $f(i, j)$ is the grey level of a pixel whose coordinates are (i, j), $0 \leq i \leq H$, and $0 \leq j \leq W$, and H and W are the height and width of a digitized image f, respectively. Also, by convention, the grey level zero is the darkest and the grey level $l - 1$ is the brightest.

As a first step, given an image f, a smoothed image f_{filt} is generated whose intensity at every pixel (i, j) is obtained by averaging the intensity values of the pixels of f contained in a 3 × 3 neighborhood S of (i, j). Formally

$$f_{\text{filt}}(i, j) = \frac{1}{K} \sum_{(n,m) \in S} f(n, m), \quad \forall (i, j) \in W \times H$$

where K is the total number of points in S and W and H are the width and height of the image f. This smoothing is required in order to produce better histograms to get more accurate thresholds.

Now, given f_{filt}, thresholding as in [11] is applied recursively. Initially, a grey-level histogram is evaluated for the entire image

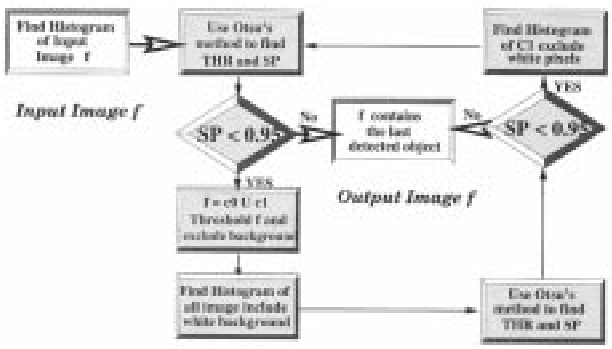

Fig. 4. Diagram of recursive thresholding technique.

$f_{\rm filt}$ as illustrated in Fig. 3(b). From this histogram, the method [11] automatically calculates the threshold THR_1 (t^* as in Section III) and the separability factor SP_1 (η^* as in Section III) for the entire image. SP_1 is a number ranging from zero to one, indicates the likelihood of separating the class that has the lowest intensity in the image from the other classes in the same image. If the separability factor is less than S ($S = 95$% which is determined in the training phase), which indicates the existence of two classes that are not 95% (or above) the same class, the threshold value THR_1 is used to separate the two classes or objects. This results in a new image f_1 [Fig. 3(c)] from which one object (denoted by C_0) has been segmented from the original image $f_{\rm filt}$.

At recursion t ($t \geq 1$), we will determine the grey-level histogram of the whole image, including the white pixels of the background and use [11] to calculate the separability factor SP_t and the threshold value THR_t. If SP_t exceeds 95%, the process will terminate, since the image has only one object remaining against a white background. However, if SP_t does not exceed 95%, the recursive process is repeated to threshold an object C_t. Recursion is repeated p times until no more objects can be segmented. As a result, we obtain

$$f_{\rm filt} = C_0 \bigcup C_1 \bigcup \cdots \bigcup f_p$$

where f_p [$f_p = f_{\rm thr}$, as in Fig. 3(e)] is the object that remains after segmenting the different objects. A diagram of the system is presented in Fig. 4.

V. Application to Check Images

As pointed out in Section II, the purpose of the new proposed thresholding method is to develop a general recursive thresholding technique for grey scale images and apply it to document processing for document analysis and understanding. In fact, the new technique will facilitate the process of information extraction from document images and preserve the topological properties of the extracted information that will be presented for further processing and recognition as in [23].

In pursuing this line, 220 real-life bank checks were used to train the system, and 505 different checks were used to test its performance. The training set is used to select an optimal value for S, which is set to 95% in this experimental application. For other applications, however, the value of S will differ depending on the intensity of the objects in the scene image. The optimal value of S should produce the best performance results even though performance evaluation of image processing techniques such as image segmentation is a nontrivial task to pursue. In tackling this problem, a set of human visual criteria were defined where each criterion is given a weight. Several performance evaluation techniques for image segmentation [21] and thinning [19] follow the same approach. In this application, the performance was analyzed by visual inspection considering criteria such as the thickness, intensity, and connectivity of strokes of the items to be extracted from the check images.

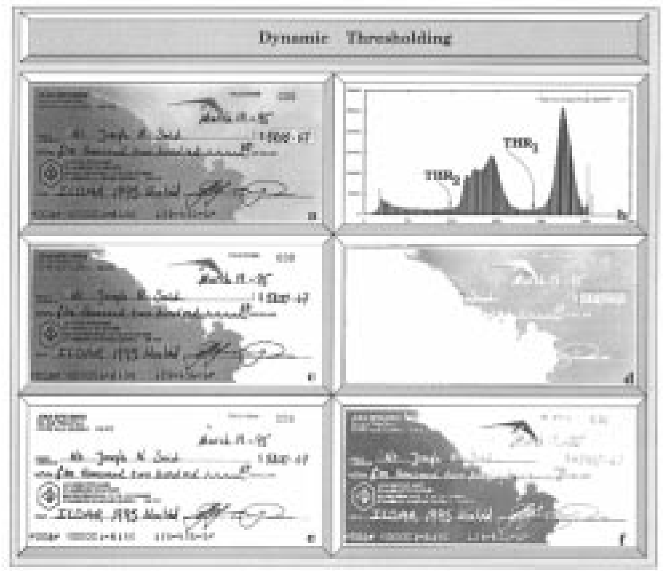

Fig. 5. Image segmentation using recursive method on a sample check image. (a) Original grey-scale check image. (b) Histogram where THR_1 and THR_2 are the optimal thresholds during the first and the second recursions. (c) Image after thresholding the first background brightest object (C_0) using THR_1. (d) First thresholded background object C_0. (e) Image f_p that contains the foreground darkest object which is the handwritten information. (f) Second thresholded background object C_1, which is the brightest object in the grey-scale image shown in (c).

In Table I we used the following convention.

- **D:** dark background or dark handwritten
- **D+:** very dark background or very dark handwritten information
- **S:** simple and homogeneous background with no figures
- **C:** complex background with homogeneous figures considered as one object (refer to Section IV)
- **C+:** complex background with nonhomogeneous figures considered as multiobjects (refer to Section IV)
- **t:** thin handwritten information
- **T:** thick handwritten information
- **L:** very light handwritten information
- **L+:** light but not dark handwritten information

As an example to illustrate this notation, the check image presented in Fig. 5 is considered as having the attributes of **D**, **C+** for its background [Fig. 5(d) and (f)] and $\{t, T\}$ and **D+** for its foreground [Fig. 5(e)]. Moreover, columns $A, B, C, D, E, F,$ and G represent the number of check images, average size of an image in megabytes, background and foreground of a given cheque image f, time in seconds in calculating $f_{\rm filt}$, $f_{\rm thr}$, and their corresponding total time, and the performance analysis of the system, respectively.

Table I indicates that the training and testing checks have good and satisfactory results when the target object required is the darkest in the given image. However, when the targeted object in a given image is not the darkest object, recursive segmentation will not properly segment the targeted object. The last two entries in Table I illustrate this point.

The system is developed under a Sun Sparc 330 (Sparc 1), with 24 M of memory and 25 MHz. An HP scanner is employed to capture the image of checks. The resolution of digitization of the experiment can vary between 200 and 300 dpi.

TABLE I
EXPERIMENTAL RESULTS

A	B	C	D	E	F	G
Training Set						
86	1.3	D S / t D+	14	2	16	99%-100%
61	1.4	D C / t D+	15	3	18	96%-100%
47	1.5	D C / {t, T} L+	16	5	21	85%-100%
26	1.5	D C+/ {t, T} D+	17	6	23	73%-84%
Testing Set (the training images were not used for testing)						
68	1.4	D S / {t, T} D+	15	2	17	99%-100%
74	1.5	D S / {t, T} L+	17	2	19	96%-100%
58	1.4	D S / {t, T} L	14	2	16	85%-100%
33	1.5	D+ S / {t, T} {D+, L+, L }	16	2	18	73%-84%
47	1.4	D C / {t, T} D+	15	5	20	98%-100%
56	1.4	D C / {t, T} L+	14	5	19	93%-100%
31	1.5	D C / {t, T} L	16	5	21	81%-100%
21	1.4	D+ C / {t, T} {D+, L+, L }	15	5	20	59%-81%
48	1.6	D C+ / {t, T} D+	18	5	23	98%-100%
35	1.6	D C+ / {t, T} L+	17	5	22	90%-100%
20	1.4	D C+ / {t, T} L	14	5	19	0
14	1.5	D+ C+ / {t, T} {D+, L+, L }	16	7	13	0

VI. CONCLUSION

In this work, we introduced a new and general method for image segmentation by extending Otsu's approach [11]. At each recursion, the new approach segments first the object with the lowest intensity from the given image. The recursive process continues until there is one object (the darkest) left in the image. As an application, the new method is trained on 220 real-life bank checks and tested on another 505 checks to eliminate their background to facilitate further processing [23] on these checks. Results of the performance analysis indicated that the new method is very efficient and effective in segmenting the darkest object in a given image. Currently, we are developing other segmentation methods [10] based on edge detection to locate targeted objects that are not the darkest in a given image.

ACKNOWLEDGMENT

The authors thank Davis and Henderson Intercheques, Inc. for samples provided, as well as their comments and interest in our research.

REFERENCES

[1] S. Baukharouba, J. M. Rebordao, and P. L. Wendel, "An amplitude segmentation method based on the distribution function of an image," *Comput. Vis. Graph. Image Process.*, vol. 29, pp. 47–59, 1985.

[2] M. J. Carlotto, "Histogram analysis using scale-space approach," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 121–129, 1987.

[3] M. Cheriet, J. N. Said, and C. Y. Suen, "A formal model for document processing of bank cheques," in *Proc. 3rd Int. Conf. Document Analysis and Recognition*, Montreal, P.Q., Canada, Aug. 24–26 1995, pp. 210–213.

[4] L. Hertz and R. W. Schafer, "Multilevel thresholding using edge matching," *Comput. Vis. Graph. Image Process.*, vol. 44, pp. 279–295, 1988.

[5] R. Kasturi and L. O'Gorman, Guest Eds., *Machine Vis. Applic. Spec. Issue: Document Image Anal. Tech.*, vol. 5, pp. 141–248, 1992.

[6] ——, *Machine Vis. Applic. Spec. Issue: Document Image Anal. Tech.*, vol. 6, pp. 67–180, 1993.

[7] J. N. Kapur, P. K. Sahoo, and A. K. Wong, "A new method for grey-level picture thresholding using the entropy of the histogram," *Comput. Vis. Graph. Image Process.*, vol. 29, pp. 273–285, 1985.

[8] J. Kittler and J. Illingworth, "Minimum error thresholding," *Pattern Recognit.*, vol. 19, pp. 41–47, 1986.

[9] R. Kohler, "A segmentation system based on thresholding," *Comput. Graphics Image Process.*, vol. 15, pp. 319–338, 1981.

[10] K. Liu *et al.*, "Automatic extraction of baselines and data from cheque images," *Int. J. Pattern Recognit. Artif. Intell.*, to be published.

[11] N. Otsu, "A threshold selection method from grey-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 62–66, 1978.

[12] N. R. Pal and S. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, pp. 1277–1294, 1993.

[13] *Proc. First Int. Conf. Document Analysis and Recognition*, Saint-Malo, France, 1991.

[14] *Proc. Second Int. Conf. Document Analysis and Recognition*, Tsukuba Science City, Japan, 1993.

[15] *Proc. Third Int. Conf. Document Analysis and Recognition*, Montreal, P.Q., Canada, 1995.

[16] S. S. Reddi, S. F. Rudin, and H. R. Keshavan, "An optimal multiple threshold scheme for image segmentation," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-14, pp. 661–665, 1984.

[17] P. K. Sahoo, S. Soltani, and A. K. C. Wong, "SURVEY: A survey of thresholding techniques," *Comput. Vis. Graph. Image Process.*, vol. 41, pp. 233–260, 1988.

[18] M. Spann and R. Wilson, "A quad-tree approach to image segmentation which combines statistical and spatial information," *Pattern Recognit.*, vol. 18, pp. 257–269, 1985.

[19] S. W. Lee, L. Lam, and C. Y. Suen, "Performance evaluation of skeletonizing algorithms for document image processing," in *Proc. First Int. Conf. Document Analysis and Recognition*, Saint-Malo, France, 1991, pp. 260–271.

[20] O. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 1191–1201, Dec. 1995.

[21] O. D. Trier and T. Taxt, "Evaluation of binarization methods for document images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 312–315, Mar. 1995.

[22] S. Wang and R. M. Haralick, "Automatic multithreshold selection," *Comput. Vis. Graph. Image Process.*, vol. 25, pp. 46–67, 1984.

[23] J. N. Said, K. Khorasani, and C. Y. Suen, "A new back-propagation learning algorithm with application to unconstrained handwritten character recognition," in *Proc. World Congr. Neural Networks*, Washington, DC, July 1995, vol. 2, pp. 217–221.

[24] J. N. Said, M. Cheriet, and C. Y. Suen, "An extended formal model for visual data processing of business forms and bank cheques," in *Proc. Interface'96 Conf.*, Toronto, Ont., Canada, pp. 84–92.

[25] ——, "Dynamic morphological processing: A fast method for base line extraction," in *Proc. 13th Int. Conf. Pattern Recognition*, Vienna, Austria, Aug. 25–30, 1996, vol. II, pp. 8–12.