Stroke-Model-Based Character Extraction from Gray-Level Document Images

Xiangyun Ye, Mohamed Cheriet, *Senior Member, IEEE*, and Ching Y. Suen, *Fellow, IEEE*

***Abstract*—Global gray-level thresholding techniques such as Otsu's method, and local gray-level thresholding techniques such as edge-based segmentation or adaptive thresholding method are powerful in extracting character objects from simple or slowly varying backgrounds. However, they are found to be insufficient when the backgrounds include sharply varying contours or fonts in different sizes. In this paper, a stroke model is proposed to depict the local features of character objects as double-edges in a predefined size. This model enables us to detect thin connected components selectively, while ignoring relatively large backgrounds that appear complex. Meanwhile, since the stroke width restriction is fully factored in, the proposed technique can be used to extract characters in predefined font sizes. To process large volumes of documents efficiently, a hybrid method is proposed for character extraction from various backgrounds. Using the measurement of class separability to differentiate images with simple backgrounds from those with complex backgrounds, the hybrid method can process documents with different backgrounds by applying the appropriate methods. Experiments on extracting handwritings from check image, as well as machine-printed characters from scene images demonstrate the effectiveness of the proposed model.**

***Index Terms*—Background removal, binarization, document image analysis, local thresholding.**

I. INTRODUCTION

EXTRACTION of pertinent data from document images with complex backgrounds remains a challenging problem in character recognition applications, such as automatic bank check processing [1]–[3], document search on the World Wide Web [4], vehicle license recognition [5], recognition of compact disk labels [6], indexing of video frames [7], database organization, image indexing and retrieval, etc. The problem of character extraction is critical to the performance of the aforementioned applications, in that the recognition results depend mostly on the quality of the characters extracted. The variability of the background and structure of document images, together with the intrinsic complexity of the character recognition problem, makes the development of general algorithms and strategies for automatic document processing systems difficult.

Modeling character objects is thus an attractive topic in character extraction methods. Global thresholding methods take account for only gray level information in the extraction of characters and classify a pixel in the original image as an object or background pixel simply according to its gray level in the global view. To extract character objects from complex or nonuniformly illuminated backgrounds, Cheriet *et al.* [3] proposed a *Recursive Ostu* algorithm, while Liu and Srihari [8] proposed a feature-based method. Problems remain unsolved when the histogram of the objects overlaps with that of the background. The solutions to overcome the disadvantages of global thresholding can be classified into two classes. The first one includes deliberate locally adaptive thresholding methods, such as Bernsen's *adaptive thresholding* [9] that estimates a threshold for each pixel according to its local neighbors. The drawback of this method is that when a background region with varying gray level is encountered, the contour pixels distinguishing dark backgrounds from brighter ones are classified as object pixels, known as "ghost" phenomenon. Ohya *et al.* [10] proposed a connected-component analysis method, assuming the gray level within character objects is homogeneous. Gu *et al.* [11] proposed a *differential top-hat* transform to extract characters from scene images, considering characters as particles of limited size. Since fixed parameters are used in these methods, problems arise when the input images manifest variable contrasts. The second solution is to extract features directly from gray scale images. Wang and Pavlidis [12] considered the gray scale image as a surface, with topological features corresponding to the shape features of the original image. Each pixel of the image is classified as *peak, pit, ridge, ravine, saddle, flat,* or *hillside*, according to rules based on the estimated first and second directional derivatives of the underlying image intensity surface, and the features are used to extract or recognize the characters. Cheriet [13] proposed a multiscale approach to extract the full shape of handwritten data from noisy gray level images, in order to overcome the intensity variation of strokes over a broad range of spatial scales. Parker [14] proposed a *local intensity gradient* technique based on the local contrast estimated by the gray level difference between a pixel and its neighbors. These methods are based on the assumption of binary document images, which limits the application of the methods to character extraction from backgrounds that are no more complex than only monotonically increasing gray tone transformations. More precise description of the character objects necessitates a model, not only for gray level information, but also for local geometric properties. Among the few papers endeavoring to model character objects and background, we find that Kamel

Manuscript received January 18, 2000; revised April 24, 2001. This work was supported by grants received from the Ministry of Education of Quebec, the IRIS Network of Centres of Excellence, and the Natural Sciences and Engineering Research Council of Canada. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Uday B. Desai.

X. Ye and C. Y. Suen are with the Centre for Pattern Recognition and Machine Intelligence, Concordia University, Montréal, PQ H3G 1M8, Canada (e-mail: xyye@cenparmi.concordia.ca; suen@cenparmi.concordia.ca).

M. Cheriet is with the Imagery, Vision, and Artificial Intelligence Laboratory, École de Technologie Supérieure, University of Québec, Montréal, PQ H3C 1K3, Canada (e-mail: cheriet@gpa.etsmtl.ca).

Publisher Item Identifier S 1057-7149(01)06042-0.

and Zhao's *logical level technique* [15] emphasizes the local linearity of character strokes, and Djeziri *et al.*'s *filiformity* [16] is the closest description of the local features of character objects. These methods are found to have the following weak points. First, loose restrictions are applied to the stroke width of characters to be extracted, leaving residues to some background regions. Second, only symmetric local windows are used, so that character strokes thinner than three pixels are lost. Third, fixed thresholds are applied to the local features, so that character strokes with varying contrasts can not be extracted. To contribute a unified solution to character extraction problems, we hereby define a descriptive model for character strokes, accounting for both gray level features and local geometric features.

As we observed, most binarization methods are based on intuitive observation of the gray level distinctions between characters and backgrounds, regardless of the complexity of the document image. To efficiently process the input images, we propose a hybrid method that applies different techniques to extract characters from document images with different complexities. A class separability measurement is used to distinguish documents printed with complex backgrounds from those printed with simple backgrounds. Since the problem of binarizing handwritings on simple background has already been solved by many global or local thresholding techniques, this paper will focus on the problems caused by uneven background of documents. A model of strokes, namely *double-edge*, is proposed as a unified solution to the problem of character extraction from complex backgrounds. The *double-edge* features can be extracted by a set of morphological operators, enabling either a fast algorithm in sequential processing or parallel processing by special-purpose hardware. Feature thresholding followed by conditional dilation for stroke extraction is discussed. Finally, we present the computational cost of the proposed method, and its goal-directed evaluation in a real-life application involving bank check processing.

II. Gray-Level Thresholding Techniques

Binarization techniques based on gray levels can be divided into two classes: global and local. Global binarization algorithms use a single threshold for the entire image. Pixels having gray levels lower than the threshold value are labeled as print (black), otherwise background (white). Locally adaptive binarization methods compute a separate threshold for each pixel based on its neighbors. Some of the methods calculate a threshold surface over the entire image. If a pixel $f(x,y)$ in the raw image has a higher gray level than the threshold surface value at (x,y), it is labeled as background, otherwise as print. For document images such as checks or maps, which may contain different backgrounds consisting of pictures and designs printed in various colors and intensities, the scanned images are often contaminated by variable backgrounds, low contrast, and stochastic noise. These difficulties exclude global algorithms and lead us to adopt a locally adaptive algorithm that is insensitive to the varying background.

A comparative study of binarization methods for document images has been conducted by Trier *et al.* [17], [18]. Eleven most promising locally adaptive algorithms have been studied: *Bernsen's* method, *Chow and Kaneko's* method, *Eikvil's* method, *Mardia and Hainsworth's* method, *Niblack's* method, *Taxt's* method, *Yanowitz and Bruckstein's* method, White and Rohrer's *dynamic threshold* algorithm, White and Rhorer's *integrated function* algorithm, *Parker's* method, *Trier and Taxt's* method, (corresponding references can be found in [17] and [18]) of which the former eight use explicit thresholds or threshold surfaces, while the latter three search for printed pixels after having located the edges. Trier and Jain concluded that Niblack's [19] and Bernsen's [9] methods along with post-processing step proposed by Yannowitz and Bruckstein [20] were the fastest and best ones based on a set of subjective and goal-directed criteria. The post-processing step here improves the binary image by removing 'ghost' objects. The average gradient value at the edge of each printed object is calculated and objects having an average gradient below a threshold are labeled as misclassified and therefore removed.

In a recent bank check item extraction method [21], we applied a local thresholding method based on Bernsen's method. In this method, for each pixel of the original image with gray level $f(x,y) \in [0, l-1]$, the local threshold $g(x,y)$ and local variance $c(x,y)$ are calculated, respectively

$$g(x,y) = (F_{\max}(x,y) + F_{\min}(x,y))/2 \qquad (1)$$

$$c(x,y) = F_{\max}(x,y) - F_{\min}(x,y) \qquad (2)$$

in which $F_{\max}(x,y)$ and $F_{\min}(x,y)$ are the highest and lowest values in a local neighborhood centered at pixel (x,y). Each pixel (x,y) is classified as an object pixel (indicated by value 1) or a background pixel (indicated by value 0) according to the following conditions:

$$b(x,y) = \begin{cases} 1, & \text{if } f(x,y) < g(x,y) \quad \text{AND} \quad c(x,y) > c^* \\ 0, & \text{otherwise} \end{cases} \qquad (3)$$

where the threshold c^* of local variance is decided by applying Otsu's method to the histogram of $c(x,y)$. The characters are extracted and stored as $B = \{b(x,y) \mid (x,y) \in F\}$.

This local thresholding method can binarize document images with slowly varying nonuniform backgrounds, yet when the background tends to vary at sharp edges; these edges can be taken as objects, causing great difficulty to the ensuing recognition stage. Fig. 1 shows some examples of character strokes on different types of backgrounds, each posing a certain degree of difficulty in extracting the strokes. Fig. 1(a) shows the ideal case for a simple background that can be removed by a global threshold; Fig. 1(b) illustrates a slow-varying background that can be removed by the above mentioned local thresholding method; Fig. 1(c) illustrates a background image with a sharp edge in the middle, which will produce an artifact if a local thresholding method is applied. Actually, Fig. 1(c) is a common case when legal amounts of bank checks are considered. As an example of difficult bank checks, Fig. 2 shows a bank check designed deliberately with stylistic, colorful and complex backgrounds, which cause difficulties in both the extraction and recognition of such items as legal amount, courtesy amount, date, signature, etc. The solution

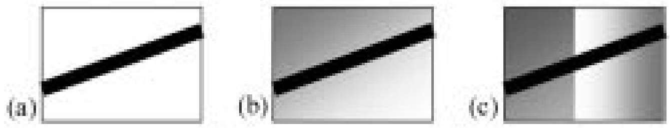

Fig. 1. Character stroke on different types of background. (a) Simple background; easy to binarize by global thresholding; (b) slowly varying background; can be removed by lical thresholding; (c) fast varying background; need to be removed by feature thresholding.

Fig. 2. Typical difficult bank check image.

to this problem requires a more precise and robust feature to describe the objects of our interest.

III. Character Extraction by Stroke Model

A. Topological Model for Character Strokes and Feature Extraction

Before we discuss the model of character objects, it is necessary to lay down some basic assumptions. Without losing generality, we consider the case of a dark stroke written on a lighter background, and regard a pixel with the following characteristics as an object pixel

1) it has a gray level lower than that of its local neighbors;
2) it belongs to a thin connected component with width less than a predefined value.

The above assumptions do not exclude the possibility of overlapping gray level distribution between objects and backgrounds, which is common in a variety of practical systems.

Edge-based image analysis is widely used in computer vision. Marr [22] describes any type of significant scene structures in an image as discontinuities in the intensity, generally called edge. From the viewpoint of extracting scene structures, the objective is not only finding significant intensity discontinuities, but also suppressing unnecessary details and noise while preserving positional accuracy. Bergholm [23] proposed that in any image, there are only four elementary structures: the *step edge*, the *double-edge*, the *corner edge* (L-junction), and the *edge box* (blob)

$$\text{step edge}: f(x,y)=\begin{cases}1, & \text{if } x>a\\ 0, & \text{elsewhere}\end{cases}$$
$$\text{double-edge}: f(x,y)=\begin{cases}1, & \text{if } a<x<b\\ 0, & \text{elsewhere}\end{cases}$$
$$\text{corner edge}: f(x,y)=\begin{cases}1, & \text{if } x>a \text{ and } y>b\\ 0, & \text{elsewhere}\end{cases}$$
$$\text{edge box}: f(x,y)=\begin{cases}1, & \text{if } a<x<b \text{ and } c<y<d\\ 0, & \text{elsewhere.}\end{cases}$$

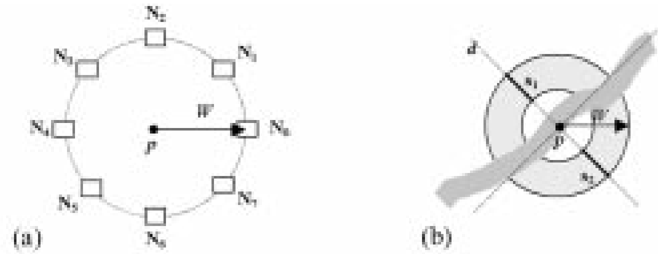

Fig. 3. Symmetric masks used in the literature. (a) Logical Level Technique [15]; $N_I(i=0,\ldots,7)$ are local windows that are W pixels away from p; (b) Filiformity [16]; s_1 and s_2 are two segments defined by an exterior and an interior rings aroung pixel p.

An *edge box* can be regarded as *double-edges* in two orthogonal directions, and the *double-edge* model, delimited by a positive edge and a negative one nearby, is a good description of a stroke in the local region. Based on this observation, Palumbo [24] proposed a second derivative method to extract the characters; White and Rohrer [25] proposed an integrated function technique to extract the sequence of $+,-$, and 0 in the differential image, and take the pixels between sequences "$+-$" and "$-+$" as objects; Liu *et al.* [26] extracted baselines from bankchecks by searching for a pair of opposite edges in a predefined distance. One common problem for these methods is the sensitivity to noise, because the depths of the strokes have not been taken into account. Better solutions can be found in Kamel and Zhao's *logical level technique* [15] and Djeziri *et al.*'s *filiformity* [16] measurement, expressed as

$$L_d(p)=\mathrm{Min}\{\mathrm{Avg}_{q\in N_i}(f(g)),\mathrm{Avg}_{q\in N_j}(f(q))\}-f(p), \quad \text{(Logic Level)} \quad (4)$$
$$\mu_d(p)=\mathrm{Min}\{\mathrm{Max}_{q\in s_1}(f(q)),\mathrm{Max}_{q\in s_2}(f(q))\}-f(p), \quad \text{(Filiformity).} \quad (5)$$

Here, $d = 0,1,2,3$ refer to four directions, $\{0,\pi/4,\pi/2,3\pi/4\}$, that approximate most possible edge directions around a pixel; s_1 and s_2 are two segments defined by an exterior ring and an interior one around pixel $p; Ni$ and N_j are two $(2W+1)^*(2W+1)$ squares that are W pixels away from p, in opposite directions (Fig. 3); W refers to the width constraints for the strokes to be extracted. The following estimation is used to evaluate whether pixel p belongs to the stroke or the background

$$L_W(p)=\mathrm{Max}_{d=0}^{3}\{L_d(p)\}$$
$$\mu_W(p)=\mathrm{Max}_{d=0}^{3}\{\mu_d(d)\}. \quad (6)$$

Having observed that the *logic level technique* (*LLT*) tends to break character strokes near darker neighbors, Oh [27] proposed a *bright-average scheme*, which replaces the average value in local windows by the average of the maximum in columns and rows. Although the above-mentioned approaches tend to describe character strokes in different ways, we found the common features behind are the thickness and darkness of a stroke. In this sense, the local threshold techniques based on stroke width restriction can be unified into searching for *double-edge* features occurring within a certain distance. In a one-dimensional (1-D)

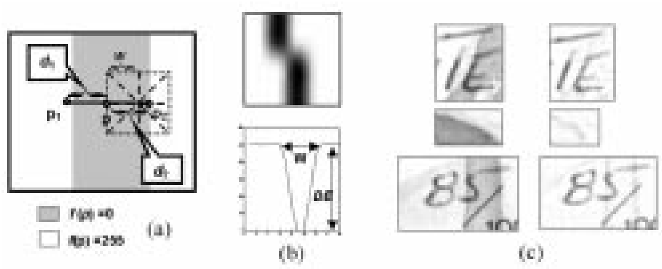

Fig. 4. (a). Close-up of an ideal stroke in a local neighborhood; (b) 1-D profile of an ideal stroke; (c) some examples of double-edge feature extracting (sharply varying backgrounds whose size is larger than $W = 5$ are eliminated).

profile, the darkness of a stroke with thickness W, can be estimated as a *double-edge*, whose intensity is

$$DE(x) = \mathrm{Max}_{i=1}^{W-1}\{\mathrm{Min}(f(x-i), f(x+W-i))\} - f(x). \quad (7)$$

The darkness of a two-dimensional (2-D) stroke can therefore be approximated as $DE_W(p) = \mathrm{Max}_{d=0}^{3}\{DE_{Wd}\}$, whose positive values indicate the existence of a dark stroke written on a lighter background. Although both *filiformity* and *LLT* try to extract narrow connected components, they are not applicable to strokes that are narrower than two pixels. Meanwhile the usage of symmetric masks limits the size of the extracted objects. This can be understood from a close-up of an ideal black stroke (represented as $f(p) = 0$) written on a white background (represented as $f(q) = 255$), shown in Fig. 4(a) and (b). For a given pixel p on the stroke, we assume that d_1 and d_2 are the distances between p and the nearest opposite boundaries p_1 and p_2. The relationships between (d_1, d_2) and the feature values at pixel p are

$$\text{Logic Level}: L_W(p) = \begin{cases} 1, & \text{if } \mathrm{Max}(d_1, d_2) < W - (2W+1)T/255 \\ 0, & \text{else} \end{cases}$$

$$\text{Filiformity}: \mu_W(p) = \begin{cases} 255, & \text{if } \mathrm{Max}(d_1, d_2) < W \\ 0, & \text{else} \end{cases}$$

$$\text{Double-Edge}: DE_W(p) = \begin{cases} 255, & \text{if } (d_1 + d_2) < W \\ 0, & \text{else.} \end{cases}$$

Fig. 4(c) illustrates some typical examples of *double-edge* features extracted from a document with sharply varying backgrounds. In the experiments we observed that strokes extracted by thresholding *filiformity* and *logic levels* remain artifacts, caused by the noise regions larger than the strokes, while *double-edge* feature eliminated all objects exceeding size W. Properly used, this *double-edge* model can not only extract characters from complex backgrounds [28], but also remove machine-printed textual noise that touches the handwriting [29].

According to our basic assumption of dark strokes/lines appearing on light backgrounds, it is easy to prove that only the existence of a dark stroke-like object gives rise to a positive value of $DE_W(x, y)$. Decomposing the *double-edge* image into positive and negative parts, the original image can be separated into information and noise. In [30] and [31], Cheriet *et al.* have given a detailed discussion about transferring

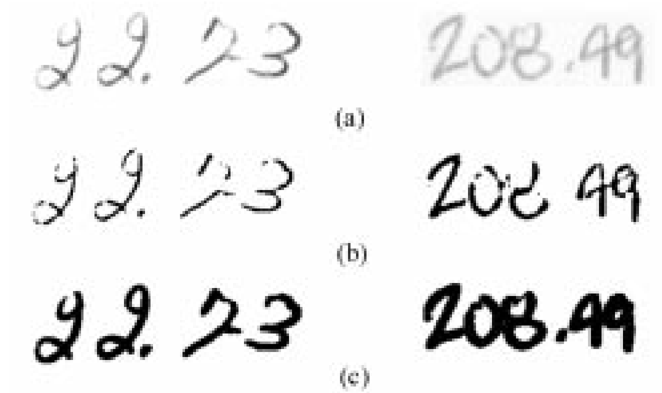

Fig. 5. Binarization of a courtesy amount. (a) Raw images of a courtesy amount with class separability $\mu = 0.856$ and $\mu = 0.803$, classified as "medium" difficult; (b) binarization by local gray level thresholding [Section III, formulae (4) and (5)], in which the global threshold of the local contrast is obtained by Ostu's method. The variation of intensity in strokes causes broken strokes in the binarized image; and (c) Binarization by feature thresholding and conditional dilation [Section III, formula (10)], in which the connectivity of strokes is preserved in the binarized image.

information among image representations at different scales. In their discussion, the character object is modeled as the concave part of an image surface, while the convex part is treated as background. Hence the zero-crossings are chosen as thresholds in the Laplacian scale-space, in order to differentiate objects from noise. We take his method of decomposing the original image into information and noise, as the primary extraction of character objects. Similarly, the feature image is decomposed into positive part DE_W^+ and negative part DE_W^-, with positive part corresponding to the candidate character objects

$$DE_W = DE_W^+ + DE_W^- \quad (8)$$

$$DE_W^+(x,y) = \begin{cases} DE_W(x,y), & \text{if } DE_W(x,y) > 0 \\ 0, & \text{otherwise} \end{cases}$$

$$DE_W^-(x,y) = \begin{cases} (-1)^* DE_W(x,y), & \text{if } DE_W(x,y) < 0 \\ 0, & \text{otherwise.} \end{cases}$$

The existence of dark strokes is then measured by the intensity of the positive part, denoted by

$$S_W(x,y) = DE_W^+(x,y). \quad (9)$$

B. Feature Thresholding for Stroke Extraction

The transformation from the original image into a double-edge feature image enables us to detect fine objects in spite of the complex backgrounds. Each pixel in the image can then be classified into object or background according to the intensity of its local feature.

As discussed in [30] and [31], images of unconstrained handwritten data contain strokes with intensities over a wide range of spatial scales [Fig. 5(a)]. Although the local features of either contrast, *Filiformity* or *double-edge* can help to eliminate the intensity variation in background regions, the variation within strokes themselves may cause severe information loss when a global thresholding technique is applied to the feature images [9]–[16], [19]. In [30], a top-down and a bottom-up multiscale approaches to extract handwritten data from gray level images were proposed. In the bottom-up approach, information is transferred among several scales, and full shape data are extracted. The multiscale approach assumed a simple background for ZIP code images, which does not hold in document images with complex backgrounds. Yet, the idea of information transferring among different scales supports the following post-processing based on conditional dilation.

Define $|F|_T$ as a thresholding operation with either a constant threshold or a threshold surface defined by an image

$$|F|_T = \{f_T(x,y) \,|\, (x,y) \in F\},$$
$$f_T(x,y) = \begin{cases} \begin{cases} 1, & \text{if } f(x,y) > T \\ 0, & \text{otherwise} \end{cases} & \text{if } T \text{ is a constant} \\ \begin{cases} 1, & \text{if } f(x,y) > T(x,y) \\ 0, & \text{otherwise.} \end{cases} & \text{if } T \text{ is an image} \end{cases} \tag{10}$$

We may use the following conditional dilation to extract the character strokes:

$$B_i = (B_{i-1} \oplus D) \bigcap |S_W|_{T_1}, \quad i = 1, 2, \ldots, B_0 = |S_W|_{T_0}. \tag{11}$$

Here, $0 < T_1 < T_0 < 255$ are two constant thresholds for the intensity of strokes, and D is a $3 * 3$ square structuring element. The iteration in (11) is repeated until there is no change from B_{i-1} to B_i. The characters are thus extracted and stored as B_n, where $B_n = B_{n-1}$ holds. B_n differs from $|S_W|_{T_1}$ in that only those stroke pixels with relative high intensities are taken as seeds, and the stroke pixels with lower intensities are extracted if they are connected to the seeds. By this method, the variation of gray scales within strokes is eliminated.

Let us recall (1) and (2), and rewrite (3) as $B = |C|_{\dot{c}} \bigcap |G|_F$, with the thresholding operation defined in (10), then the conditional dilation can be similarly applied as a post-processing procedure of the local gray level thresholding

$$B_i = (B_{i-1} \oplus D) \bigcap \left(|C|_{c_1} \bigcap |G|_F\right) \quad i = 1, 2, \ldots,$$
$$B_0 = |C|_{c_0} \bigcap |G|_F, \quad 0 < c_1 < c_0 < 255. \tag{12}$$

Here $G = \{g(x,y) \,|\, (x,y) \in F\}$ and $C = \{c(x,y) \,|\, (x,y) \in F\}$ are the local threshold and variance of the gray level image $F = \{f(x,y)\}; B_0 = |C|_{c_0} \bigcap |G|_F$ is the local thresholding result obtained from (3). In our application, we heuristically chose the thresholds T_0, T_1, c_0, c_1 as 1.2 and 0.8 times the Ostu optimal threshold for feature images, i.e., the *double-edge* and local contrast images. Therefore the conditional dilation is able to eliminate the "ghost" objects while connecting broken strokes. The thresholds T_0, T_1, c_0, c_1 are so selected, since lower values of T_0 and c_0 tend to expand the extracted strokes to background regions, while values higher than T_1 and c_1 will exclude isolated light handwritten strokes from the final results. Visual inspection in the results of local gray level thresholding [Fig. 5(b)] shows the broken strokes caused by the noticeable variation of stroke intensities. While the local gray-level thresholding method produces some broken strokes,

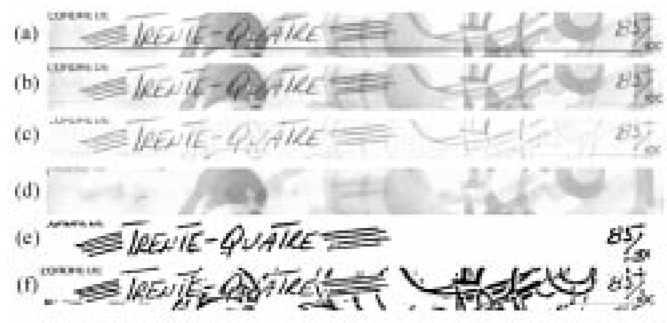

Fig. 6. An example of Legal Amount extraction. (a) Raw image A (1341×99 pixels); (b) new image A' after baseline removal; (c) double-edge feature detected in image A'; (d) background image estimate from image A'; (e) character extraction by feature thresholding; and (f) binarized image by applying Bernsen's algorithm to the raw image A'.

significant improvement can be obtained by the proposed conditional dilation method, without introducing noise into character extraction [Fig. 5(c)]. Encouraging results have been obtained even when complex backgrounds are encountered. An example of legal amount extraction shown in Fig. 6, a complex case posing difficulty to most binarization methods. The estimation and elimination of baselines shown in Fig. 6(b) is implemented by a set of morphological operations [21]. Within a unified scheme, the proposed method is able to extract character objects from complex backgrounds with satisfactory quality [Fig. 6(f)].

IV. Comparison of Computational Complexity

Similar to many local thresholding methods that can be implemented efficiently on parallel machines [32], the proposed method can be implemented in parallel by special-purpose hardware. This is a useful property when large-scale industrial applications are considered. In this paper, we assume software implementation without special-purpose image processing hardware. Therefore, the computational complexity related to the sequential implementation of the thresholding is measured by the required number of comparisons in carrying out the local operations.

A. Local Gray Level Thresholding (Enhanced Bernsen's [9], [21] Method)

The local gray level threshold at size W is calculated by (1) to (3). The major part of the computations involved in this method is the calculation of local maximum and local minimum described in (1). Since $N-1$ comparisons are required to extract the maximum or minimum value of N elements, $(W^2-1) \times 2$ comparisons are required for each pixel in the worst case. Douglas [33] has proven that regardless of the size of the structuring elements, erosion and dilation by linear structuring elements can be implemented with only three comparisons for each pixel, and with six comparisons by rectangular structuring elements. Therefore, the total number of comparisons required in calculating the local maximum and minimum is

$$C_1 = 6 \times 2 = 12. \tag{13}$$

B. Double-Edge Based Feature Extraction

The *double-edge* at thickness W is calculated by (7). Similarly, $[2+(W-1)]$ comparisons are required for each $DE_d(d=0,1,2,3)$, and the total number of comparisons required in calculating *double-edge* for each pixel is

$$C_2 = [2+(W-1)] \times 4 + 3 = 4W + 7. \tag{14}$$

When W becomes larger than three, we may use the fast algorithm described in [33] to calculate the local maxima. In this case, the total number of comparisons required for calculating *double-edge* at each pixel in the raw image becomes

$$C_3 = \begin{cases} 4W+7, & \text{if } W \le 3 \\ (3+1) \times 4 + 3 = 19, & \text{otherwise.} \end{cases} \tag{15}$$

V. Complexity of Document Image

Real-life documents are sometimes designed deliberately with stylistic, colorful, and complex backgrounds, causing difficulties in choosing appropriate character extraction methods. While global thresholding techniques can extract objects from simple or uniform backgrounds at high speed, local thresholding methods can eliminate varying backgrounds at a price of processing time. When batch-processing speed becomes a major concern in practical applications of an automatic document processing system, it is very important to have a balanced solution to process large quantities of documents at both high speed and high reliability. The first step of a balanced solution is to differentiate complex images from simple ones, so that various types of images can be processed with the appropriate methods. Meanwhile, the classification of document images should involve as little computation as possible, so that no extra computation is needed for processing simple images.

For this purpose, we start with a well-known global thresholding method in retrospect. Based on discriminant analysis, Ostu's method of threshold selection [34] is ranked as the best and fastest global thresholding method [17], [18]. Thresholding operation corresponds to partitioning the pixels of an image into two classes, C_0 and C_1, at threshold t. For an image with gray level ranges within $G = \{0, 1, \ldots, l-1\}$, the object and background can be represented by two classes, as $C_0 = \{0, 1, \ldots, t\}$ and $C_1 = \{t+1, t+2, \ldots, l-1\}$ or *vice versa*. Without losing generality, we define the objects as dark characters against lighter backgrounds. As elegant and simple criteria for class separability, the within and between class scatter matrices are widely used in discriminant analysis of statistics [35]. The within-class variance, between-class variance and total-variance are represented by σ_W^2, σ_B^2 and σ_T^2, respectively, and they reach the maximum at equivalent threshold t. Therefore, the optimal threshold t^* can be determined by maximizing one of the following criterion functions against the threshold

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2}, \quad \eta = \frac{\sigma_B^2}{\sigma_T^2}, \quad k = \frac{\sigma_T^2}{\sigma_W^2}. \tag{16}$$

Taking η for example, the optimal threshold t^* is determined as

$$t^* = \mathrm{ArgMax}_{t \in [0, l-1]}(\eta) \tag{17}$$

in which

$$\begin{aligned} &P_i = n_i/n, \quad w_0 = \sum_{i=0}^{t} P_i, \quad w_1 = 1 - w_0 \\ &\mu_T = \sum_{l=0}^{t-1} iP_i, \quad \mu_t = \sum_{i=0}^{t} iP_i, \quad \mu_0 = \frac{\mu_t}{w_0} \\ &\mu_1 = \frac{\mu_T - \mu_t}{1-\mu_0} \\ &\sigma_T^2 = \sum_{i=0}^{l-1} (i-\mu_T)^2 P_i, \quad \sigma_B^2 = w_0 w_1 (\mu_1 - \mu_0)^2. \end{aligned}$$

Here, n_i is the ith element of the histogram, i.e., the number of pixels at gray-level $i; n = \sum_{i=0}^{l-1} n_i$ is the total number of pixels in the image; $P_i = n_i/n$ is the probability of occurrence at gray-level i. For a selected threshold t^* of a given image, the class probabilities w_o and w_1 represent the portions of areas occupied by object and background classes respectively. The maximal value of η, denoted by η^*, can serve as a measurement of the separability between the two classes and the bimodality of the histogram. The class separability η lies in the range $[0, 1]$; the lower bound is reached when the image has a uniform gray level, and the upper bound is obtained when the image consists of only two gray levels. Since the larger η is, the simpler the background, we propose to classify the original images according to their class separabilities, and treat them with different techniques. For an ideal case when characters and background are composed of only two distinct gray levels, the class separability reaches the maximum, $\eta^* = 1$.

By choosing one or several proper thresholds for class separability, we are able to classify document images as relatively "simple" or "complex," and, thus, apply corresponding binarization methods to optimize the batch processing speed and meanwhile preserve recognition rate of the entire system. The thresholds for class separability are established from a training set of real-life document images, depending on the application. The training set should be composed of document images at different complexity levels, and the distribution of the complexity should conform to that of the real inputs to the system. Although we can not obtain an explicit function of image complexity versus class separability, we can choose a threshold for class separability, since the complexity of a document image is a nonincreasing function of the class separability (Fig. 7). For example, we may choose two thresholds η_0 and $\eta_1, 0 \le \eta_0 \le \eta_1 \le 1$, to classify the document images into three classes: "simple" images with $\eta_1 < \eta \le 1$, "medium" images with $\eta_0 < \eta \le \eta_1$, or "complex" images with $0 \le \eta \le \eta_0$. Hence the binarization procedure of a document image can follow the diagram shown in Fig. 8. The entire binarization procedure can be specified as global gray-level thresholding if $\eta_1 = 0$; as local gray-level thresholding when $\eta_0 = 0$ and $\eta_1 = 1$; or as feature thresholding which is the worst case in terms of processing time, when $\eta_0 = 1$. Compared with global gray-level thresholding techniques, local gray-level and feature thresholding methods take the local features of a document image into account, and can be applied to more complex images with one-tenth of the normal speed or even lower.

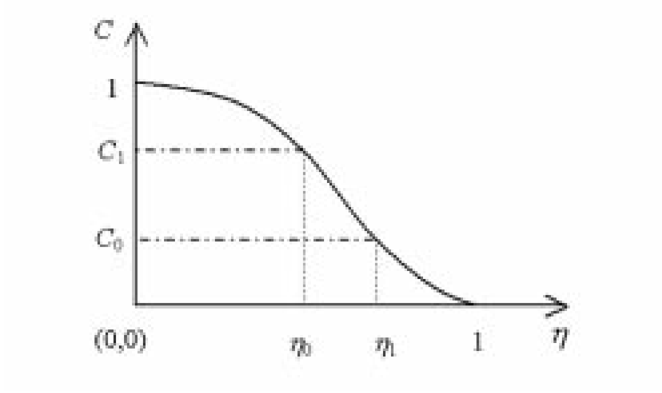

Fig. 7. Image complexity versus class separability.

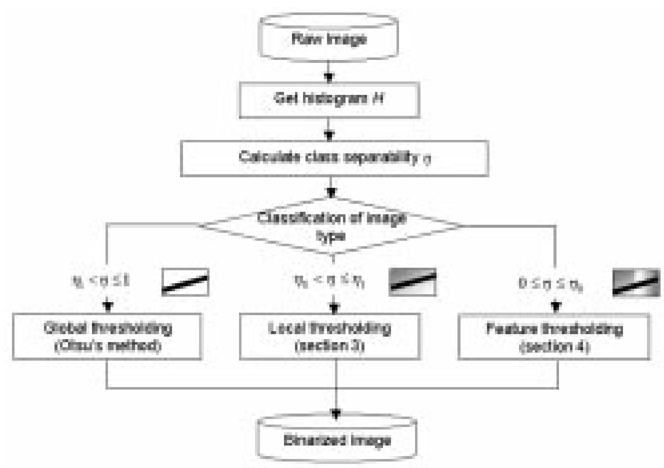

Fig. 8. Overview of binarization procedure for document images.

Suppose the average processing time of one image by global gray-level, local gray-level, and feature thresholding methods are t_0, t_1, and t_2, respectively; the average processing time can be expected as

$$\bar{t} = \sum_{i=0}^{2} P_i \cdot t_i \tag{18}$$

in which p_i represents the probability of different types of images in a testing set, estimated by the total number of images in the testing set and the number of images belonging to various classes. For the simplest case, when $p_i = 1/3(i = 0, 1, 2)$, the expectation of the processing time is the algebraic average of $t_i(i = 0, 1, 2)$, which is a substantial improvement when large quantities of images are processed.

VI. Experimental Results and Discussion

A. Subjective Evaluation

We evaluate the performance of the proposed *double-edge* model based method for character extraction in both subjective and objective manners. The subjective evaluation relies mostly on visual effects, including the continuity of strokes, and the suppression of the backgrounds. Two images taken from a French check, and a tourism magazine, which represent different types of documents are tested [Figs. 9(a) and 10(a)]. Local binarization methods are evaluated, including the enhanced Bernsen's method, Kamel-Zhao's method, Oh's method, Tophat Transform [36] followed by Ostu's method, Niblack's method, Djeziri's method followed by Ostu's method, and the proposed method. We observed the largest thickness to be detected as $W = 4$ and $W = 5$ for the two images, and set the threshold in Kamel-Zhao's and Oh's methods to 20. The constant k in Niblack's method is set to $k = -0.2$. Visual inspection shows the ability of the

Fig. 9. Local binarization results of "French Cheque." (a) Raw image; (b) enhanced Bernsen's method; (c) Kamel-Zhao's logic level technique; (d) Oh's LLT based on bright-average scheme; (e) Tophat Transform followed by Ostu's method; (f) Niblack's method; (g) Djeziri's filiformity followed by Ostu's method; and (h) proposed double-edge followed by conditional dilation.

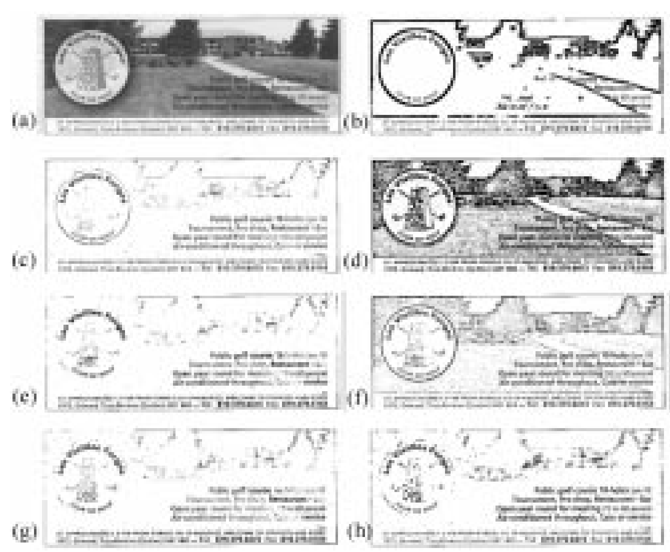

Fig. 10. Local binarization results of the "Tourism magazine." (a) Raw image; (b) enhanced Bernsen's method; (c) Kamel-Zhao's logic level technique; (d) Oh's LLT based on bright-average scheme; (e) Tophat Transform followed by Ostu's method; (f) Niblack's method; (g) Djeziri's filiformity followed by Ostu's method; and (h) proposed double-edge followed by conditional dilation.

TABLE I
COMPARISON OF COMPUTATIONAL TIME REQUIRED BY DIFFERENT LOCAL BINARIZATION METHODS

Method	Processing time (seconds) (measured on a Pentium® II 400 PC) French Check (Fig. 9a) (400x198 pixels) W=4	Tourism magazine (Fig. 10a) (673x264 pixels) W=5
Enhanced Bernsen's	0.27	0.60
Kamel-Zhao's	0.61	2.04
Oh's	0.67	2.72
Top-Hat + Ostu	1.17	3.64
Niblack's	2.09	7.47
Djeziri's + Ostu	1.43	4.01
Double-edge model	1.21	3.35

double-edge model based method in distinguishing the character strokes of desired thickness from the rest. In the French check (Fig. 9), most parts of the character string "RESERVED" and thick diagonal lines are removed since all objects larger than the predefined stroke width are considered as noise. For the tourism magazine (Fig. 10), most methods break the strokes at low contrast into pieces, while Niblack's and the *double-edge* model based method keep good continuity in characters. More experiments on images taken from various sources including Canadian checks and maps show that the *double-edge* model based method can extract characters in varying contrasts. Compared to other methods, it generally keeps good continuity in character strokes while suppressing the backgrounds. The sensitivity to stroke width enables the *double-edge* model based method to distinguish characters with the desirable stroke width from other characters or backgrounds. Meanwhile, it limits the method to extract characters with a known stroke width. When this information can not be obtained *a priori*, the characters can not be extracted properly.

B. Objective Evaluation

In terms of processing time, the proposed method is amid several local thresholding methods. We have the following results on the French check and the tourism magazine (Table I).

For quantitative evaluation, we follow the goal-directed rules used by Trier *et al.* [17], [18], in which an image understanding module based on the results of the low-level image processing routine in question is used. The experiments are carried out in a bank check processing system [1]. In multibank check processing, the main difficulties stem from variations in text locations and the complexity of their backgrounds. Since each bank has its specific check design and the location of items, the performance of low-level image processing, especially character extraction, actually plays an important role in the overall specification of multi-bank check processing systems. Instead of evaluating the quality of the extracted items by visual or machine-computed criteria [37], [38], an objective evaluation method should be conducted to investigate the effects of different extracting strategies on the subsequent image analysis steps. Since we have more reliable recognition techniques for courtesy amounts, we have chosen the courtesy amount to evaluate the proposed method of character extraction. The courtesy

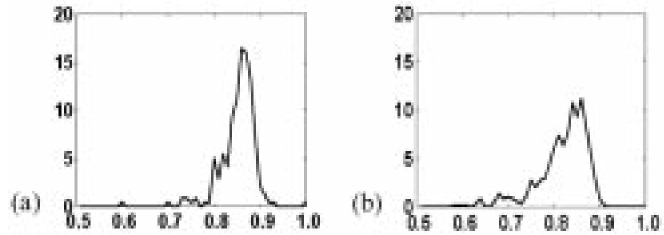

Fig. 11. Distribution of class separability. (a) Class separability of courtesy amount and (b) class separability of legal amount (250 real-life checks are analyzed).

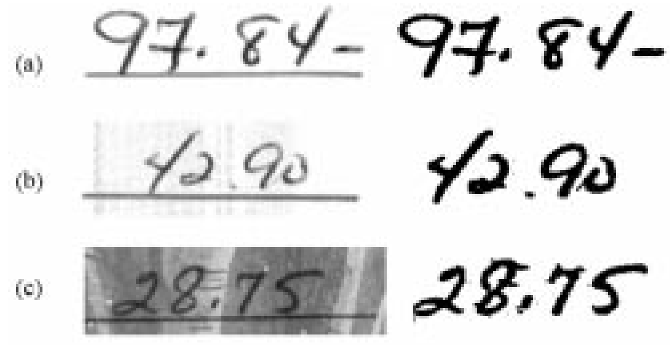

Fig. 12. Examples of different types of background. (a) "Simple" case, $\mu = 0.906$; (b) "medium" case, $\mu = 0.843$; and (c) "complex" case, $\mu = 0.674$.

amount recognizer developed by Strathy et al. [1], [39] is used. The digit recognizer was trained on about 340 000 digits from the NIST training database, plus about 15 000 digits collected by CENPARMI in experimental checks. The extraction and elimination of baselines from bank check images, as well as restoration of broken strokes, described in [21], are based on a set of morphological operations with structuring elements depicting the linear feature.

We conducted both qualitative and quantitative experiments on a set of 870 real-life gray-level check images scanned at Bell Quebec with a resolution of 300 DPI. Scanned at this resolution, most handwritings have a stroke width of 3–5 pixels. Therefore, we set $W = 6$ in the tests. The distributions of class separability in both courtesy and legal amounts estimated from a subset of 250 images are shown in Fig. 11. According to subjective observation, we classify the images into three classes, "simple," "medium," and "complex." Two thresholds $\eta_0 = 0.8$, and $\eta_1 = 0.88$ are chosen empirically by matching the subjective observation with the class separability [Fig. 12]. In fact, the local gray level thresholding and feature thresholding methods always outperform the global gray level thresholding, which means the higher the thresholds, the better the results. A comparison of computational time is shown in Table II; the processing time required for each image increases drastically when more precise and deliberate algorithms are applied. Among the several types of thresholding techniques we have analyzed, the *double-edge* model is ranked as the most computation-intensive method. However, the classification of original images according to their complexity helps to reduce the processing time. A balance between the computing speed and extracting quality should be taken into consideration when choosing the thresholding methods.

TABLE II
COMPUTATIONAL TIME REQUIRED BY DIFFERENT BINARIZATION STRATEGIES FOR COURTESY AMOUNTS (AVERAGE SIZE OF IMAGE IS $250*100$. RESULTS OBTAINED ON A PENTIUM® II 400 PC)

Binarization methods	Ostu's	Enhanced Bernsen's	*Double-edge* model	Hybrid method
Time (sec/image)	0.01	0.18	0.83	0.64

TABLE III
COMPARISON OF NONREJECTION ERROR RATES OBTAINED BY DIFFERENT BINARIZATION METHODS (ONLY THE RESULTS FOR COURTESY AMOUNT ARE PROVIDED)

Method	Non-rej. Err. Rate (%)
Ostu's	37.1
Enhanced Bernsen's	32.7
Kamel-Zhao's	94.3
Oh's	34.9
TopHat + Ostu	37.4
Niblack's	82.8
Djeziri's + Ostu	89.5
Double-edge model	21.1
Hybrid	25.4

In order to evaluate the performance of different character recognition methods quantitatively, we compared the proposed *double-edge* feature thresholding method and the hybrid method with several locally adaptive thresholding methods (Table III). We observed that due to the highly variable contrast in the handwritten strokes, some locally adaptive thresholding techniques that involve fixed parameters are found inadequate in this application.

The last column of Table II and the last row of Table III illustrate an important result, that is the combination of different character extraction methods can improve the processing speed of the system without affecting the recognition rate substantially. Although the experiments are carried out on courtesy amount recognition, we observed that more complex background images are usually found in the legal amount region (Fig. 2). The examples of extracting legal amount as well as other types of documents in Figs. 9 and 10 demonstrate the ability of the proposed method in dealing with complex backgrounds, and we plan to evaluate of the proposed method with a robust legal amount recognizer and a common database in due course.

VII. CONCLUSION AND FUTURE DIRECTION

The subject of character extraction from document images with complex backgrounds is still one of the most challenging topics in the field of document image analysis. In this paper, a model-based character extraction method using morphological operations is proposed. According to different background image complexities, we classify the raw images into three categories:

1) "simple," from which the binarized document image can be obtained by a global thresholding method;
2) "medium," which requires a locally adaptive thresholding followed by a post-processing step;
3) "complex," which involves sharply varying backgrounds, posing difficulties to most binarization methods.

The complexity of the document image is estimated by class separability, which is a weak yet useful measurement in classifying the document images. A hybrid character extraction approach can reduce the computational cost, which can be prohibitive in processing a large number of check images. The experimental result shows the natural adaptation of the method to complex and varying backgrounds.

For the cases involving complex backgrounds, we define character strokes as trenches with a limited size, described by a *double-edge* appearing in a local region. The experimental results on the goal-directed evaluation of courtesy amount extraction show the robustness of the method to complex and varying backgrounds. However, the *double-edge* feature does not distinguish noise in linear form from true strokes, and other features may be required to remove such interfering marks. Since the computational cost of *double-edge* feature thresholding is far more than that of global gray-level thresholding, we are looking for a robust and explicit measurement of image complexity, in order to improve the tradeoff between performance and processing speed. In other words, we suggest the application of simple algorithms to clean images and more sophisticated methods to complex images.

As observed in real-life checks, complex background images are often found in legal amount regions, causing errors in recognition. Although this method has not been evaluated by legal amount recognition, experimental results show promising improvement in visual aspects. Further evaluation on legal amount extraction would require a reliable recognizer and a common database, which are under investigation. It is noticeable that the proposed approach can be applied not only to check images, but also to other document images with complex backgrounds and poor illumination.

ACKNOWLEDGMENT

The authors would like to thank N. Strathy for introducing the digit recognizers, C. P. Nadal and A. N. Bloch for their help in the writing, and J. Zhou for his help in preparing the final draft.

REFERENCES

[1] C. Y. Suen, L. Lam, D. Guillevic, N. W. Strathy, M. Cheriet, J. N. Said, and R. Fan, "Bank check processing system," *Int. J. Imag. Syst. Technol.*, vol. 7, pp. 392–403, 1996.
[2] *Int. J. Pattern Recognit. Artif. Intell.*, vol. 11, no. 4, 1997.
[3] M. Cheriet, J. N. Said, and C. Y. Suen, "A recursive thresholding technique for image segmentation," *IEEE Trans. Image Processing*, vol. 7, pp. 918–921, June 1998.
[4] D. Lopresti and J. Zhou, "Document analysis and the world wide web," in *Proc. IAPR Workshop Document Analysis Systems*, 1996, pp. 651–669.
[5] H. A. Hegt, R. J. Haye, and N. A. Khan, "A high performance license plate recognition system," *Proc. IEEE Int. Conf. Systems, Man, Cybernetics*, pp. 4357–4362, 1998.
[6] Y. Zhong, K. Karu, and A. K. Jain, "Locating text in complex color images," *Pattern Recognit.*, vol. 28, no. 10, pp. 1523–1535, 1995.
[7] A. K. Jain and B. Yu, "Automatic text location in images and video frames," in *Proc. 14th Int. Conf. Pattern Recognition*, Brisbane, Qld., Australia, 1998, pp. 1497–1499.
[8] Y. Liu and S. N. Srihari, "Document image binarization based on texture features," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 540–544, May 1997.

[9] J. Bernsen, "Dynamic thresholding of gray-level images," in *Proc. Int. Conf. Pattern Recognition*, vol. 2, 1986, pp. 1251–1255.
[10] J. Ohya, A. Shio, and S. Akamatsu, "Recognizing characters in scene images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 16, pp. 214–220, Feb. 1994.
[11] L. Gu, T. Kaneko, N. Tanaka, and R. M. Haralick, "Robust extraction of characters from color scene image using mathematical morphology," in *Proc. 14th Int. Conf. Pattern Recognition*, Brisbane, Qld., Australia, 1998, pp. 1002–1004.
[12] L. Wang and T. Pavlidis, "Direct gray-scale extraction of features for character recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1053–1067, Oct. 1993.
[13] M. Cheriet, "Extraction of handwritten data from noisy gray-level images using a multi-scale approach," in *Proc. Vision Interface 98*, vol. 1, Vancouver, BC, Canada, 1998, pp. 389–396.
[14] J. R. Parker, "Gray level thresholding in badly illuminated images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 813–819, Aug. 1991.
[15] M. Kamel and A. Zhao, "Extraction of binary character/graphics images from grayscale document images," *CVGIP: Graph. Models Image Process.*, vol. 55, no. 3, pp. 203–217, 1993.
[16] S. Djeziri, F. Nouboud, and R. Plamondon, "Extraction of signatures from check background based on a filiformity criterion," *IEEE Trans. Image Processing*, vol. 7, pp. 1425–1438, Oct. 1998.
[17] Ø. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 1191–1201, 1995.
[18] Ø. D. Trier and T. Taxt, "Evaluation of binarization methods for document images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 312–314, 1995.
[19] W. Niblack, *An Introduction to Digital Image Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1986.
[20] S. D. Yanowitz and A. M. Bruckstein, "A new method for image segmentation," *Comput. Vis., Graph., Image Process.*, vol. 46, no. 1, pp. 82–95, 1989.
[21] X. Ye, M. Cheriet, C. Y. Suen, and K. Liu, "Extraction of bankcheck items by mathematical morphorlogy," *Int. J. Doc. Anal. Recognit.*, vol. 2, pp. 53–66, 1999.
[22] D. Marr and E. Hildreth, "Theory of edge detection," in *Proc. R. Soc. Lond.*, vol. B-207, 1980, pp. 187–217.
[23] F. Bergholm, "Edge focusing," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 9, pp. 726–741, June 1987.
[24] P. W. Palumbo, P. Swaminathan, and S. N. Srihari, "Document image binarization: Evaluation of algorithms," *Proc. SPIE*, vol. 697, pp. 278–285, 1986.
[25] J. M. White and G. D. Rohorer, "Image thresholding for optical character recognition and other applications requiring character image extraction," *IBM J. Res. Develop.*, vol. 27, no. 4, pp. 400–411, 1983.
[26] K. Liu, C. Y. Suen, M. Cheriet, J. N. Said, C. Nadal, and Y. Y. Tang, "Automatic extraction of baselines and data from check images," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 11, no. 4, pp. 675–697, 1997.
[27] I.-S. Oh, "Document image binarization preserving stroke connectivity," *Pattern Recognit. Lett.*, vol. 16, pp. 743–748, 1995.
[28] X. Ye, M. Cheriet, and C. Y. Suen, "Model-based character extraction from complex backgrounds," in *Proc. 5th Int. Conf. Document Analysis Recognition*, Bangalore, India, 1999, pp. 511–514.
[29] ——, "A generic system to extract and clean handwritten data from business forms," in *Proc. 7th Int. Workshop Frontiers Handwriting Recognition*, Amsterdam, The Netherlands, 2000, pp. 63–72.
[30] M. Cheriet, "Extraction of handwritten data from noisy gray-level images using a multiscale approach," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 13, no. 5, pp. 665–684, 1999.
[31] M. Cheriet, R. Thibault, and R. Sabourin, "A multi-resolution based approach for handwriting segmentation in gray level images," *Proc. IEEE Int. Conf. Image Processing*, pp. 159–168, 1994.
[32] N. B. Venkateswarlu, "Implementation of some image thresholding algorithms on a connection machine-200," *Pattern Recognit. Lett.*, vol. 16, pp. 759–768, 1995.
[33] S. C. Douglas, "Running max/min calculation using a pruned ordered list," *IEEE Trans. Signal Processing*, vol. 44, pp. 2872–2877, Nov. 1996.
[34] N. Ostu, "A threshold selection method from gray-scale histogram," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 62–66, 1978.
[35] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. New York: Academic, 1990.
[36] J. Serra, *Image Analysis and Mathematical Morphology*. London, U.K.: Academic, 1982.
[37] L. Heutte, P. Barbosa-Pereira, O. Bougeois, J. V. Moreau, B. Plesis, P. Courtellemont, and Y. Lecourtier, "Multi-bank check recognition system: Consideration on the numeral amount recognition module," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 11, no. 4, pp. 595–618, 1997.
[38] S. Knerr, V. Anisimov, O. Baret, N. Gorski, D. Price, and J. C. Simon, "The A_2iA intercheque system: Courtesy amount and legal amount recognition for French checks," *Int. J. Pattern Recognit. Artif. Intell*, vol. 11, no. 4, pp. 505–548, 1997.
[39] N. W. Strathy, C. Y. Suen, and A. Krzyzak, "Segmentation of handwritten digits using contour features," in *Proc. 5th Int. Conf. Document Analysis Recognition*, vol. 1, 1993, pp. 577–580.

Xiangyun Ye recieved the Ph.D. degree in communication and electronic systems from Shanghai Jiao Tong University, China, in 1997, and the B.E. and M.E. degrees in biomedical engineering from Southeast University, China, in 1991 and 1994, respectively.

Since October 1997, she has been with CENPARMI, Concordia University, Montréal, PQ, Canada, and LIVIA at École de Technologie Supérieure, University of Québec, Montréal, as a Postdoctoral Fellow and Research Associate. Her research interests include image processing, machine vision, pattern recognition, character recognition, and document analysis.

Mohamed Cheriet (M'95–SM'00) received the B. Eng. degree in computer science from the Université des Sciences et de Technologie d"Alger, Algiers, in 1984, and the M.Sc. and Ph.D. degrees in computer science from the University of Pierre et Marie Curie, Paris, France, in 1985 and 1988, respectively.

He was appointed Assistant Professor in 1992, Associate Professor in 1995, and Full Professor in 1998 in the Department of Automation Engineering, École de technologie superieure (ETS), University of Quebec, Montreal, QC, Canada. Currently, he is the Director of the Laboratory for Imagery, Vision and Artificial Intelligence (LIVIA), ETS, and an active member of the Centre for Pattern Recognition and Machine Intelligence (CENPARMI). His research focuses on image processing, pattern recognition, character recognition, text processing, documents analysis and recognition, and perception. He has published more than 90 technical papers in the field. He was a Guest Editor of the *International Journal of Pattern Recognition and Artificial Intelligence* and the *Machine, Perception, and Artificial Intelligence Series* books (Singapore: World Scientific).

Dr. Cheriet was the Co-Chair of the 11th and 13th Vision Interface Conferences in 1998 and 2000. He is currently the General Co-Chair of the 8th International Workshop on Frontiers on Handwriting Recognition, to be held in 2002.

Ching Y. Suen (M'66–SM'78–F'96) received the M.Sc.(Eng.) degree from the University of Hong Kong and the Ph.D. degree from the University of British Columbia, Vancouver, BC, Canada.

In 1972, he joined the Department of Computer Science of Concordia University, Montréal, PQ, Canada, where he became Professor in 1979 and served as Chairman from 1980 to 1984, and as Associate Dean for Research of the Faculty of Engineering and Computer Science from 1993 to 1997. Currently he holds the Distinguished Concordia Research Chair of Artificial Intelligence and Pattern Recognition, and is the Director of CENPARMI. He is the author/editor of 11 books and more than 300 papers on subjects ranging from computer vision and handwriting recognition, to expert systems and computational linguistics. He is the founder and Editor-in-Chief of a journal and an Associate Editor of several journals related to pattern recognition. He has served several professional societies as President, Vice-President, or Governor. He is also the founder and chair of several conference series including ICDAR, IWFHR, and VI. Currently he is the General Chair of the International Conference on Pattern Recognition to be held in Quebec City in August 2002.

Dr. Suen is a Fellow of the IAPR and the Academy of Sciences of the Royal Society of Canada. He is the recipient of several awards, including the ITAC/NSERC Award in 1992 and the Concordia "Research Fellow" Award in 1998.