An Extended Formal Module for Visual Data Processing of Business Forms and Bank Cheques

J. N. Said ‡ †, M. Cheriet †, C.Y. Suen ‡

‡ Centre for Pattern Recognition and Machine Intelligence, Concordia University
1455 De Maissonneuve Blvd. West, Montréal, Québec, H3G 1M8, Canada.
† Laboratoire d'Imagerie, de Vision et d'Intelligence Artificielle
École de Technologie Supérieure du l'Université du Québec
4750 Avenue Henri-Julien, Montréal, Québec., H2T 2C8, Canada.
email: said@cenparmi.concordia.ca, said@gpa.etsmtl.ca, cheriet@gpa.etsmtl.ca,
suen@cenparmi.concordia.ca

Abstract

Recently a formal model for document processing of business forms has been developed for visualizing and understanding various types of documents. In this paper, we extend this approach to allow the system to visualize and understand the finest granularity of data written or printed on these documents. An extended formal model is proposed to seek specific regions inside a document and perform appropriate local, instead of global, thresholding techniques that lead to extracting valuable information from the discovered regions with the challenge of preserving its topological properties. We will introduce the approach, illustrate how the extended formal model, which uses a new local thresholding technique to eliminate the background of a given document, and present the experimental results of both systems to compare their performances.

1 Introduction

Document processing of business forms has attracted considerable interests in recent years ([4, 5, 6, 7, 12, 16, 18, 19, 25, 28, 29, 30, 32]). [7] and [16] introduced a formal approach to extract information from bank cheques and present it to further processes that are able to recognize the written information on these scanned bank cheques [16, 23]. According to Figure 1, the process of visualizing, understanding, and extracting handwritten and hand-printed data from bank cheques is divided into the following steps: Image enhancement (average filtering), image segmentation (background elimination), object detection (detection and elimination of base lines), morphological processing (closing operation), image enhancement (binarization and median filtering) topological processing (restoration of lost information), and finally segmentation of required information.

In this paper, we present an important extension to [7] to enable the new system to preserve, visualize, and understand the finest granularity of data written or printed on these documents. [7] can be used in the modeling and the implementation of many areas of document processing where the preservation of the printed and the written information on documents is a very challanging task. In such systems, the problem of eliminating the background information from the foreground and at the same time preserving the topological properties of the foreground is a non-trivial task to model. This is true because many document images could contain different backgrounds with multi-level objects that form a complex and sometimes dark background that is very difficult to eliminate using global thresholding techniques as in [7, 16]. Preserving the topological properties of the foreground (written or printed) information will increase the productivity of the system and will have a direct impact on the recognition results which is a further step taken in [16].

Thresholding a gray-scale image by a value t means converting the image to a binary one. The two levels may represent two classes or, more specifically, objects and background in an image. The process of thresholding will classify pixels that have values greater than t to one category and the rest to another category. The threshold value t is global if the same critical value t is used across the whole image as in [7]. Many thresholding techniques

have been developed for the purpose of thresholding or segmenting foreground objects from backgound objects for a given image (see, for example [2, 3, 9, 11, 13, 14, 15, 20, 10, 21, 24, 27, 31]). Some developed techniques simply use the histogram of an image. That is the numbers of pixels at each gray-level, whereas other techniques use contextual information such as gray-level occurrences in adjacent pixels.

Despite the benefits that can accrue from using contextual information techniques and allowing the threshold t to vary over an image, histogram-based algorithms are the most commonly used. They are simple to understand and implement, and computationally fast once the histogram has been obtained. However, within this restricted class, a large number of thresholding techniques have been proposed in the last decade, which present a potential user with a bewildering choice. Lee, Chung, and Park[17] compared three histogram-based algorithms with two contextual ones.

In this paper, a new local approach is presented and compared to the global approach presented in [7] towards a more intelligent system that is able to model the human perception and understanding of document images. Section 2 will introduce the model in brief; Section 3 will present the novelty of the approach of Section 2 and show the limitations of the model; Section 4 will present the new technique to enhance the productivity of the system; Section 5 will present the results of the comparative study performed on both approaches; finally, Section 6 will present the conclusion and highlight the future direction of our work.

2 The Model in Brief

Let N be the set of natural numbers. (i, j) be the spatial coordinates of a digitized image, and $G = \{0, 1, \ldots, l-1\}$ be a set of positive integers representing gray-levels[1]. Then, an image function can be defined as the mapping $f : N \times N :\longrightarrow G$, where $f(i, j)$ is the gray-level of a pixel whose coordinates are (i, j). Also, by convention, the gray-level 0 is the darkest and the gray-level $l - 1$ is the lightest.

2.1 Image Enhancement

Given an image $f(x, y)$, the procedure is to generate a smoothed image $g(x, y)$ whose intensity at every pixel (i, j) is obtained by averaging the intensity values of the pixels of f contained in a predefined neighborhood S of (i, j). In other words, the smoothed image is formally defined as $g(i, j) = \frac{1}{N} \sum_{(n,m) \in S} f(n, m), \ \forall i, j \in f(x, y)$, where S is the set of coordinates of points in the neighborhood of (i, j), including (i, j) itself, and N is the total number of points in the neighborhood, i.e., N is equal to the cardinality (number of elements) in S. Formally $N = Card(S)$. In our work we used a 3×3 neighborhood.

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

2.2 Image Segmentation

Let $t \in G$ be a threshold and $B = \{b_0, b_1\}$ be a pair of binary gray-levels such that $b_0, b_1 \in G$. The result of thresholding an image $f(x, y)$ at gray-level t is a binary image function $f_B : N \times N :\longrightarrow B$, such that

$$f_B(i, j) = \begin{cases} b_0 & \text{if } f(i, j) < t. \\ b_1 & \text{if } f(i, j) \geq t. \end{cases}$$

In general a thresholding method is one that determines the value t^* of t based on a certain criterion.

2.2.1 Extended Otsu's approach

In the histogram method for peak selection, the contributions of a small region may be masked by other larger ones. One solution is to apply thresholding recursively in the following way: the thresholded regions are considered as new images and the process of histogramming, peak selection, and thresholding is repeated until no new peaks can be found or the remaining regions are too small.

In this sense, effective segmentation can be achieved in some classes of images by a recursive method extending Otsu's approach [7, 8, 20]. In the first stage of this new approach a gray-level histogram is evaluated for the whole image. According to [20], given a gray-level histogram of an image $f(x, y)$ the threshold THR, the mean μ, the variance σ, and the separability factor SP are calculated for the whole image. The separability factor SP, a number between 0 and 1, tells us the possibility of separation between the two classes C_0 and C_1. Now, given that the separability factor is less than 95% we will assume the calculated threshold value THR for segmenting the two classes or objects. The result of segmentation is a new image $f_{filt}(t + 1)$ with one class (or object), denoted by $C_1(t)$ is being segmented from the main image call it $f_{filt}(t)$ at time t.

At time $t+1$ we will calculate again the gray-level histogram of $f_{filt}(t+1)$ and test Otsu's approach for the separability factor. If the separability factor at time $t+1$ is greater than 95% we will terminate our process since the image $f_{filt}(t+1)$ has one and only

one object with the background eliminated. However, if the separability factor at time $t+1$ calculated by the Otsu's approach on $f_{filt}(t+1)$ is less than 95%, then a new histogram is calculated, excluding all previously found objects, to test the separability factor if it is greater than 95% to segment another object $C_2(t+1)$.

This process will continue until time $t+p$ where there are no more objects to be segmented. As a result we will obtain: $f_{filt}(t) = C_1(t) \bigwedge C_2(t+1) \bigwedge \ldots \bigwedge C_i(t+i-1) \bigwedge \ldots \bigwedge C_p(t+p-1)$ where $f_{thr} = C_p(t+p-1)$. Based on the experiments we did on 505 real life bank cheques[2] [8], the number of iterations was mostly 2 and rarely 3.

2.3 Guide Lines Elimination

Once the background has been eliminated from the gray-scale image, base lines will be located and abstracted from the gray-scale image f_{thr} as follows:

$$f_{lr}(i,j) = \begin{cases} 255 & \text{if } f_{thr}(i,j) \in Line \\ f_{thr}(i,j) & otherwise \end{cases}$$

where, $f_{lr}(i,j)$ represents the image after removal and *Line* represents a recorded base line. Unfortunately, when we cut the base lines from f_{thr} there is a loss in the handwritten information that intersects with the base lines (Figure 2 (b)). As a remedy to this problem, we will introduce a formal approach using morphological processing followed by topological processing to restore (fill the gaps) the lost information because of base lines elimination.

2.4 Morphological Processing

Given f_{filt}, we will use morphological processing to restore the lost handwritten information that intersects with the eliminated base lines. To restore this lost information we will perform a morphological closing operation as in [1, 26] on f_{filt}, threshold the base lines from the morphological image, and extract the information within the line's region to be "ANDed" ($\bigwedge$) with f_{lr}. In this section, we will present the morphological closing operation and its important property as in [1] before we deal with the other processes we mentioned in this paragraph.

The gray-scale closing is defined as $f \bullet V = f \oplus V \ominus V$, where $(f \ominus V)(i,j) = min\{f(i+k, j+l) \mid (k,l) \in V_\Delta\}$ is the erosion of image f by the structuring element V and $(f \oplus V)(i,j) = max\{f(i+k, j+l) \mid (k,l) \in V_\Delta\}$ is the dilation of the image f by the same structuring element V.

[2]These cheques are obtained in colaboration with Bell Québec.

Closing has an interesting property: it is an *idempotent* transform; once it has been applied, it is useless to apply it again. This is defined as $f \bullet V \bullet V = f \bullet V$.

According to [1], every nonlinear transformation results in some loss of information; but if a transformation is idempotent, the amount of information it would lose is self-controlled which is an important property to be considered in our case. The resulting image is named f_{clos} (Figure 2 (c)).

2.5 Restoring Lost Information

Knowing the base line's positions and lengths, we can process now the morphological image f_{clos} by thresholding base line pixel values according to their MEAN as follows

$$f_{clos}(i,j) = \begin{cases} f_{clos}(i,j) & \text{if } f_{clos}(i,j) \leq MEAN \\ 255 & otherwise \end{cases}$$

$$f_{AND}(i,j) = \begin{cases} f_{clos}(i,j) & \text{if } f_{clos}(i,j) \in Line \\ f_{lr}(i,j) & otherwise \end{cases}$$

Figures 2 (d), (e), and (f) illustrate this approach.

2.6 Topological Processing

The purpose of the morphological closing operation is to reduce the intensity of base lines and increase the intensity of the handwritten data in order to reduce the loss of useful data after thresholding base lines in f_{clos}. Unfortunately, this morphological approach, sometimes, does not completely preserve the handwritten data that intersect with the base lines to be eliminated. The reason behind this is that sometimes the intensity of the base lines is very close to the intensity of the handwritten data in the original gray-scale image f_{filt}. In such cases, thresholding the base lines as in f_{lr} will result in a loss in the topological property of the handwritten information.

Three possible configurations of gaps have been considered according to the information loss with diagonal handwriting or perfect superposition. The four extremities that delimit gaps are labeled: L+, L-, R+, and R- as illustrated by Figure 3 (b). Identification of gaps is achieved by performing an edge detection algorithm, as in Section 2.7 on the image f_{AND}, whose purpose is to find the edges of the handwritten data that were lost due to the elimination of guide lines in the image f_{lr} and not restored in f_{clos} using the morphological closing operation.

2.7 Edge Detection

For digital images, the gradient magnitude $D(i,j)$ and the gradient direction $\theta(i,j)$ of f at (i,j) are $D(i,j) = \sqrt{(\triangle_x f(i,j))^2 + (\triangle_y f(i,j))^2}, \theta(i,j) = tan^{-1}(\triangle_y f(i,j))/(\triangle_x f(i,j))$, where $\triangle_x f(i,j) = f(i+1,j+1) + 2f(i+1,j) + f(i+1,j-1) + f(i-1,j+1) - 2f(i-1,j) - f(i-1,j-1)$, and $\triangle_y f(i,j) = f(i-1,j+1) + 2f(i,j+1) + f(i+1,j+1) f(i-1,j-1) + 2f(i,j-1) + f(i+1,j-1)$.

2.8 Gaps Identification and Filling

As a second step after calculating the magnitude and the digital gradient of each pixel, the algorithm is able to go through the image searching for different extremities that delimits gaps (Figure 2 (g)); the algorithm uses different rules to identify the four possible extremities. Having identified all possible extremities, gaps are filled according to the following rules in the order: (L+, L-), (R+, R-), and (L+, R+) as in Figure 3 (a). (Figure 2 (h)) shows the results.

2.9 Information Extraction

Pursuing our formalism, if $\alpha_f = \{\alpha_{f_1}, \alpha_{f_2}, \ldots, \alpha_{f_m}\}$ is a finite set of filled data and $L_s = \{L_{s_1}, L_{s_2}, \ldots, L_{s_q}\}$ is the finite set of base lines, then we will define a finite set of relations denoted by $\Gamma = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_k\}$ between the two sets α_f and L_s represented by a matrix M:

$$M = \begin{cases} \Gamma_l & \text{if } (\alpha_{f_i}, L_{s_j}) \in \Gamma \\ 0 & \text{if } (\alpha_{f_i}, L_{s_j}) \notin \Gamma \end{cases}$$

which satisfies the following condition: $\forall l(\Gamma_l = (\alpha_{f_i} \mathcal{R} L_{s_j})), \mathcal{R} = \{R, L, A, B, I\}$, where R, L, A, B, and I represent Right, Left, Above, Below, and Intersect respectively.

It is clear that the filled data $\alpha_{f_i} \forall i \in \{1, \ldots, m\}$ could be easily located using the matrix M. We define the location description of α_{f_i} as a tuple whose components are: (a) the filled data α_{f_i} (b) any element of the set $\mathcal{R}$, and (c) a line L_{s_j}. Formally, we have: $(\alpha_{f_i}, r, L_{s_j})$, $\alpha_{f_i} \in \alpha_f$, $r \in \mathcal{R}$, $L_{s_j} \in L_s$. An item description may have close relations with several graphs, so several location description tuples can be associated with each α_{f_i}.

Now given the matrix M and the location of all base lines $L_{s_i} \in L_s$, the system is able to extract α_{f_i} $\forall i \in \{1, \ldots, m\}$. The extracted set α_f could further be presented to a recognition system. Referring to Figure 2, the matrix M is represented as follows:

$$M = \begin{array}{c} \\ \alpha_{f_1} \\ \alpha_{f_2} \\ \alpha_{f_3} \end{array} \begin{array}{c} \begin{array}{ccc} L_{s_1} & L_{s_2} & L_{s_3} \end{array} \\ \begin{pmatrix} A & A & A \\ B & R & A \\ B & B & I \end{pmatrix} \end{array}.$$

In fact, we can express the location description of each α_{f_i} as follows: For α_{f_1}: $(\alpha_{f_1}, I, L_{s_1})$; $(\alpha_{f_1}, A, L_{s_2})$; $(\alpha_{f_1}, A, L_{s_3})$; $(\alpha_{f_1}, A, L_{s_4})$. This is interpreted as α_{f_1} intersects with L_{s_1} and is located above L_{s_2}, L_{s_3}, and L_{s_4}. For α_{f_2}: $(\alpha_{f_2}, B, L_{s_1})$; $(\alpha_{f_2}, R, L_{s_2})$; $(\alpha_{f_2}, A, L_{s_3})$; $(\alpha_{f_2}, A, L_{s_4})$. This is interpreted as α_{f_2} is located above L_{s_1}, on the right of L_{s_2}, and above L_{s_3}, and L_{s_4}. For α_{f_3}: $(\alpha_{f_3}, B, L_{s_1})$; $(\alpha_{f_3}, B, L_{s_2})$; $(\alpha_{f_3}, I, L_{s_3})$; $(\alpha_{f_3}, A, L_{s_4})$. This is interpreted as α_{f_3} is below L_{s_1} and L_{s_2}, intersects L_{s_3}, and above L_{s_4}.

3 Discussion

Section 2 introduced a very efficient and novel approach for processing the gray-scale images of bank cheques towards the extraction of specific fields located inside these documents. Readers are advised to refer to [7, 8] and [16] for more details. In fact, according to [7] and [8] when the handwritten information is darker than the background, background elimination becomes less complicated to threshold and the resulting image will contain less noise as well as the topological properties of the handwritten information is preserved. Figure 2 illustrates this point. Unfortunately, having a dark handwriting and a light background is not always the case. When the handwritten information is lighter than or is about to be the same as that of the background, background elimination using a global thresholding technique on the whole gray-scale cheque image will cause a significant loss in the luminance and topological properties of the handwritten information as illustrated in Figure 4. To overcome this difficulty, a new approach will be presented in Section 4 to increase the productivity of [7] and [8]. The new approach will seek specific and dynamically located regions inside the gray-scale documents and perform a local thresholding technique to better preserve the finest granularity of extracted information. This will result in a more robust system that is able to visualize, understand, and extract handwritten information from different types of document images that have various intensities of handwriting against a simple or a complex background.

4 The Extended Approach

As stated earlier the intention in this new approach is to enable the system to preserve and visualize the

finest granularity of data written or printed on the cheque images. In pursuing this line, we will use a local segmentation technique as opposed to the global segmentation technique of Section 2.2. This will enhance the quality of the extracted information and further increase the visibility and understandability of the system. This is a very important contribution towards improving the recognition results of the extracted information. Figures 5 and 6 illustrate the new approach.

In what follows a detailed Algorithm of the new approach is presented taking as an application only the extraction of date courtesy amount and legal amount written on bank cheques:

STEP 1: Given a gray-scale cheque image f, an enhanced image f_{filt} is produced as follows: $f_{filt}(i,j) = \frac{1}{N}\sum_{(n,m)\in S} f(n,m),\ \forall i,j \in f(x,y)$

STEP 2: Use the thresholding technique of Section 2.2.1 on $f_{filt}(i,j)$ to produce the image f_{thr}: $f_{filt}(t) = C_1(t) \bigwedge \ldots \bigwedge C_p(t+p-1)$, where $f_{thr} = C_p(t+p-1)$. Figure 2 (a,b) illustrates.

STEP 3: Determine $L_s = \{L_{s_1}, L_{s_2}, L_{s_3}\}$ the finite set of base lines. Determine the matrix M that represents the understanding of the structure of the image as in Section 2.9. Figure 2 (a) illustrates. Extract $\alpha_f = \{f_D, f_C, f_L\}$ the finite set of filled data as shown in Figure 7. f_D represents the gray-scale image of the extracted date, f_C represents the gray-scale image of the extracted courtesy amount, and f_L represents the gray-scale image of the extracted legal amount.

STEP 4: Enhance f_k, $\forall k \in \{D,C,L\}$ as follows:

$$f_{kfilt}(i,j) = \frac{1}{N}\sum_{(n,m)\in S} f_k(n,m),\ \forall i,j \in f_k(x,y).$$

STEP 5: Locally use the extended Otsu's Approach for background elimination of the gray-scale images f_{kfilt} $\forall k \in \{D,C,L\}$ as follows: $f_{kfilt}(t) = k_1(t) \bigwedge \ldots \bigwedge k_p(t+p-1)$, where $f_{kthr} = k_p(t+p-1)$. Figure 8 (b,d,e) illustrate.

In f_{Dthr}, determine the location description of α_{f_D}, the filled in date, as the tuple $(\alpha_{f_D}, A, L_{s_D})$ and eliminate the lines as in Equation 1 to produce the gray-scale image f_{Dlr}. In f_{Cthr}, determine the location description of α_{f_C}, the filled in courtesy amount, as the tuple $(\alpha_{f_D}, A, L_{s_C})$ and eliminate the lines as in Equation 1 to produce the gray-scale image f_{Clr}. Similarly, in f_{Lthr}, determine the location description of α_{f_L}, the filled in legal amount, as the tuple $(\alpha_{f_L}, A, L_{s_L})$ and eliminate the lines as in Equation 1 to produce the gray-scale image f_{Llr}.

STEP 6: Perform the morphological closing operation on f_{kfilt} $\forall k \in \{D,C,L\}$ $f_{kclos} \bullet V = (f_{kfilt} \oplus V) \ominus V$.

STEP 7: Restore the lost information from the morphological images f_{kclos} $\forall k \in \{D,C,L\}$

$$f_{kres}(i,j) = \begin{cases} f_{kclos}(i,j) & \text{if } f_{kclos}(i,j) \leq MEAN \\ 255 & otherwise \end{cases}$$

$$f_{kAND}(i,j) = \begin{cases} f_{kres}(i,j) & \text{if } f_{kres}(i,j) \in L_{s_k} \\ f_{klr}(i,j) & otherwise \end{cases}$$

STEP 8: Within the lines region determine the edges of the images f_{km}, $\forall k \in \{D,C,L\}$, where f_{Dm}, f_{Cm}, and f_{Lm} are the images after performing a median filtering operation on the images f_{DAND}, f_{CAND}, and f_{LAND}.

$D_{f_k}(i,j) = \sqrt{(\triangle_x f_{km}(i,j))^2 + (\triangle_y f_{km}(i,j))^2}$

$\theta_{f_k}(i,j) = tan^{-1}(\triangle_y f_{km}(i,j))/(\triangle_x f_{km}(i,j))$

The result after detecting the edges and filling the gaps are the following images f_{Dgaps}, f_{Cgaps}, and f_{Lgaps}

STEP 9: Use the rules of figure 3 to fill in the gaps to restore all the lost information that were not completely restored by the morphological operation.

STEP 10: As a final step produce the binarized images after performing a median filtering on the following images f_{Dgaps}, f_{Cgaps}, and f_{Lgaps}.

5 Experimental Results

Four hundred and twenty two bank cheques were used for training and 203 cheques were used for testing the performance of both approaches (the approach that uses the global thresholding technique to eliminate the background and the approach that uses a local thresholding after segmenting the document into various regions of interest). In **TABLE I** we used the following convention to compare and summarize the experimental results performed using the testing set: **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 2.2.1)*, **C+:** *complex background with non-homogeneous figures considered as multi-objects (refer to Section 2.2.1)*, **t:** *thin handwritten information*, **T:** *thick handwritten information*, **L:** *very light handwritten information*, **L+:** *light but not dark handwritten information*, **E:** *completely eliminated background*, **E+:** *eliminated background with some noise left*, **l:** *luminance property of handwritten data is preserved*, **l-:** *luminance property of handwritten data is not preserved*, **O:** *topological property of handwritten data is preserved*, **O-:** *topological property of handwritten data is not preserved.*

TABLE I A/B: Accuracy of Previous/New Approach		
Background/Foreground (f)	*A*	*B*
DS/{t,T}D^+ 40	99-100(ElO)	99-100(ElO)
DS/{t,T}L^+ 23	96-100(El-O)	99-100(ElO)
DS/{t,T}L 17	88-100(El-O-)	95-100(ElO)
D^+S/{t,T}{D^+,L^+,L} 31	76-87(El-O-)	90-94 (El-O)
DC/{t,T}D^+ 25	99-100(ElO)	99-100(ElO)
DC/{t,T}L^+ 26	94-100(El-O)	99-100(ElO)
DC/{t,T}L 14	82-100(El-O-)	89-100(ElO)
D^+C/{t,T}{D^+,L^+,L} 8	61-82(E+l-O-)	75-88 (El-O)
DC^+/{t,T}D^+ 8	99-100(ElO)	99-100(ElO)
DC^+/{t,T}L^+ 3	91-100(ElO)	95-100(ElO)
DC^+/{t,T}L 3	0()	86-94(El-O)
D^+C^+/{t,T}{D^+,L^+,L} 5	0()	76-85 (El-O)

Line one in **TABLE I** inidates the accuracy of both systems on 40 BELL cheques whose background is dark and simple and whose forground is dark think or dark thick.

6 Conclusion

In this paper, we extended the approach presented by [7] to enable the system to preserve and visualize the finest granularity of data written or printed on the cheque images. We introduced the architectural design as presented in [7] and show its limitations in preserving the printed as well as the handwritten information. As an extension to [7], we presented a new approach that is able to better preserve the topological properties of the printed and handwritten information. The novelty of the new approach is to segment the document into various regions of interest and then proceed to threshold these regions using local thresholds as opposed to using one global threshold value on the whole image. The result of using the new approach is to increase the visibility of the system in extracting the required information. As a result, the recognition process as introduced in [16] will be improved because the quality of extracted data is enhanced.

Currently, we are testing the performance of the recognizers on the resulting images of the new approach. Moreover, we are increasing the training set to have a more robust and intelligent system that is able to mimic the human vision and intelligence in extracting information from different documents that have different layout strutures. Increasing the efficiency of the system is another very important area that we will be conducting intensive experiments and research to have a very efficient system that can process a batch of images in a very short period of time.

Acknowledgment: We would like to acknowledge the financial supports received from the IRIS National Networks of Centres of Excellence of Canada, the NSERC of Canada, the FCAR of Québec, and Bell Québec.

References

[1] A. Bleau, J. De Guise, and A. Robert Le Blanc, "A new set of fast algorithms for mathematical morphology," *CVGIP: Image Understanding,* vol. 56, No. 2, pp. 178-209, September 1992.

[2] S. Boukharouba, J. M. Rebordao, and P. L. Wendel, "An amplitude segmentation method based on the distribution function of an image," *Compt. Vision Graphics Image Process.,* vol. 29, pp. 47-59, 1985.

[3] M. J. Carlotto, "Histogram analysis using scale-space approach," *IEEE Trans. Pattern Anal. Mach. Intell.* vol. 9 (1), pp. 121-129, 1987.

[4] R. G. Casey, D. R. Ferguson, K. M. Mohiuddin, and E. Walach, "An intelligent forms processing system," *Machine Vision and Applications,* vol. 5, no. 3, pp. 143-155, 1992.

[5] D. Wang and S. N. Srihari, "Analysis of Form Images," *International Journal of Pattern Recognition and Artificial Intelligence,* Vol. 8 No. 5 1994, pp. 1031-1052.

[6] E. R. Doughert, AN Introduction to Morphological Image Processing, Vol. TT9, SPIE Press, Bellingham, Washington, U.S.A., 1992

[7] M. Cheriet, J. N. Said, and C. Y. Suen, "A formal Model for Document Processing of Bank Cheques," 3^{rd} *Inter. Conf. on Document Analysis and Recognition,* ICDAR'95, Montreal, Aug 24-26, 1995 pp. 210-213.

[8] M. Cheriet, J. N. Said, and C. Y. Suen, "A Recursive Approach for Image Segmentation," 3^{rd} *IEEE Trans. Image Processing,* (submitted 1995). pp. 210-213.

[9] L. Hertz and R.W. Schafer, "Multilevel thresholding using edge matching," *Compt. Vision Graphics Image Process,* vol 44, pp. 279-295, 1988.

[10] S. L. Horowitz and T. Pavlidis, "Picture segmentation by a tree traversal algorithm," *J. Assoc. Comput. Mach.,* vol. 23, pp. 368-388, Apr. 1976.

[11] J. N. Kapur, P. K. Sahoo, and A. K. Wong, "A new method for gray-level picture thresholding using the entropy of the histogram," *Compt.*

Vision Graphics Image Process, vol. 29, pp. 273-285, 1985.

[12] R. Kasturi et al, "A system for interpretation of line drawings," *IEEE Trans. on Pattern Analysis and Machine Intelligence,* Vol. PAMI-12, No. 10, Oct. 1990.

[13] I. Leplumey, J. Camillerapp, and C. Queguiner, "Kalman Filter Contributions Towards Document Segmentation," *Proc. of the 3rd International Conference on Document Analysis and Recognition,* Montrèal (Canada), Aug. 14-16, 1995, Vol. 3, pp. 765-769.

[14] J. Killer and J. Illingworth, "Minimum error thresholding," *Pattern Recognition,* vol. 19, pp. 41-47, 1986.

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

[16] L. Lam, C. Y. Suen, D. Guillevic, N. W. Strathy, M. Cheriet, K. Liu, and J. N. Said, "Automatic processing of information on cheques," *Proc. 1995 IEEE Int. Conf. on Systems, Man and Cybernetics,* Vancouver, Canada, Oct. 22-25, 1995, Vol. 3, pp. 2353-2358.

[17] S. U Lee, S. Y. Chung, and R. H. Park, "A comparative performance study of several global thresholding techniques for segmentation," *Comput. Vision Graphics Image Process.* bf 52 1990, pp. 171-190.

[18] *Machine Vision and Appl.,* Special Issue: Document Image Analysis Techniques, vol. 5, no., 3, 1992.

[19] Y. Nakano, H. Fujisawa, O. Kunisaki, K. Okada, and T. Hananoi, "A document understanding system incorporating with character recognition ," in *Proc. 8th Int. Conf. on Pattern Recognition,* 1986, pp. 801-803.

[20] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Systems Man Cybernet,* vol. SMC-8, pp. 62-66, 1978.

[21] S. S. Reddi, S. F. Rudin, and H. R. Keshavan, "An optimal multiple threshold scheme for image segmentation," *IEEE Trans. Systems Man Cybernet,* vol. 14 (4), pp. 661-665, 1984.

[22] L.G. Roberts, "Machine perception of three-dimensional solids," in *Optical and Electro optical Information Processing,* (J.T. Tippett, *et al., Eds.*), pp. 159-197, MIT Press, Cambridge, Mass., 1965.

[23] J. N. Said, K. Khorasani, and C. Y. Suen, "A New Back-Propagation Learning Algorithm with Application to Unconstrained Handwritten Character Recognition," *World Congress on Neural Networks WCNN'95,* Washington DC, July 1995, vol. 2, pp. 217-221.

[24] P. K. Sahoo, S. Soltani, and A. K. C. Wong, "SURVEY: A survey of thresholding techniques," *Compt. Vision Graphics Images Process,* vol. 41, pp. 233-260, 1988.

[25] J. C. Salome, M. Leroux and H. Oiry, "Retrieval of script information appearing in bank cheques for automatic purposes," *Proc. 5th Visual Communications and Image Processing'90,* (Switzerland), Oct 1990.

[26] J. Serra, "Image analysis and mathematical morphology," Academic Press, London, 1982.

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

[28] S. L. Taylor, R. Fritzson, and J. A. Pastor,"Extraction of data from preprinted forms," *Machine Vision and Applications,* vol. 5, no., 3, pp. 211-222, 1992.

[29] Y. Y. Tang, C. D. Yan, M. Cheriet, C. Y. Suen, "Automatic analysis and understanding of documents." C. H. Chen, L.F. Pau, and P.S.P. Wang, editors, *Handbook of Pattern Recognition and Computer Vision,* World Scientific, (Singapore), 1993.

[30] T. Y. Tang, C. Y. Suen, C. D. Yan, M. Cheriet, "Financial Document Processing Based on Staff Line and Description Language," *IEEE Trans. Systems, Man. and Cybernetics,* vol. 25, no. 5, pp. 738-754, May 1995.

[31] S. Wang and R. M. Haralick, Automatic multithreshold selection, *Compt. Vision Graphics Image Process,* vol. 25, pp. 46-67, 1984.

[32] C. D. Yan, Y. Y. Tang, and C. Y. Suen, "Form understanding system based on form description language," in *Proc. First Int. Conf. on Document Analysis and Recognition,* (Saint-Malo), France, Sept. 30-Oct2, 1991, pp. 283-293.

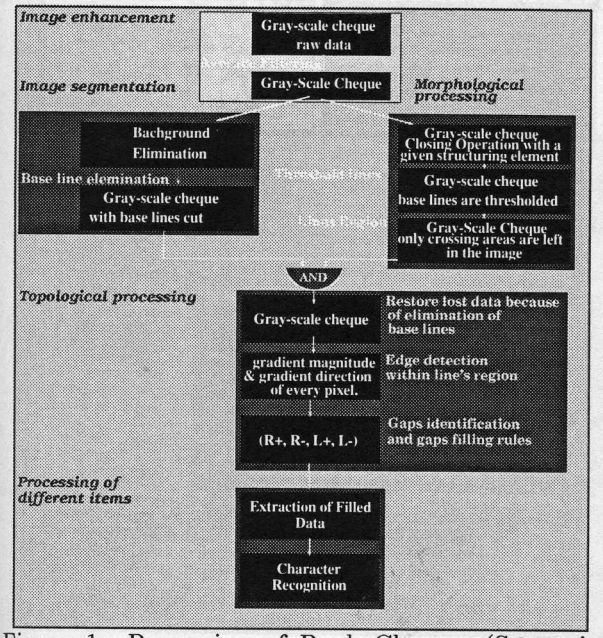

Figure 1: Processing of Bank Cheques (System's Design).

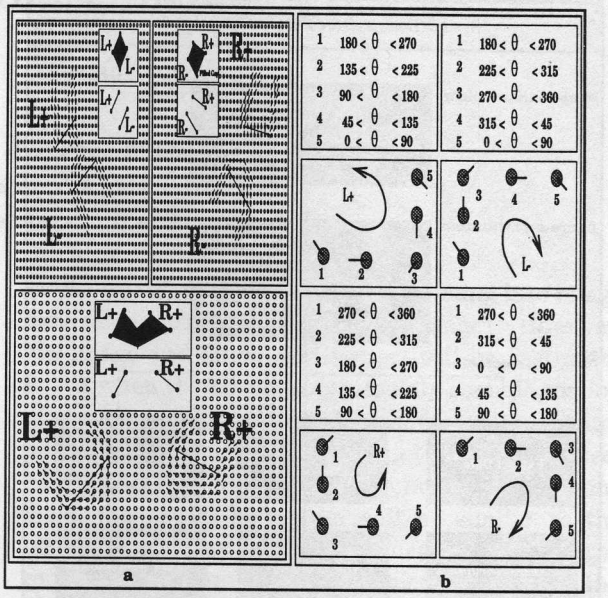

Figure 3: Rules for gaps identification and gaps filling.

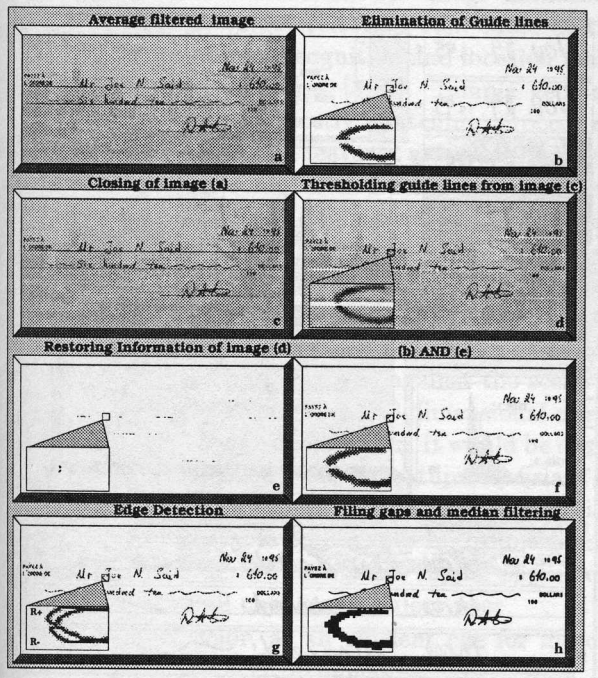

Figure 2: A sample output process.

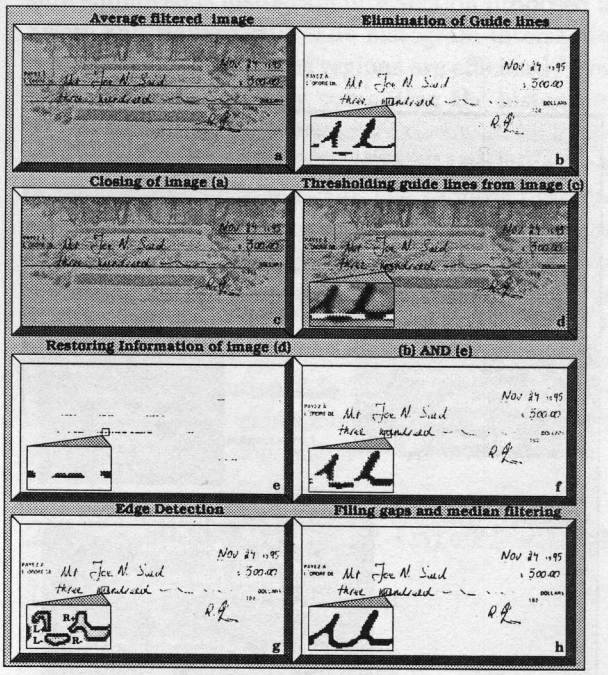

Figure 4: Luminance and topological properties are not preserved because the background is dark and complex and the handwritten information is about to have the same luminance as the background.

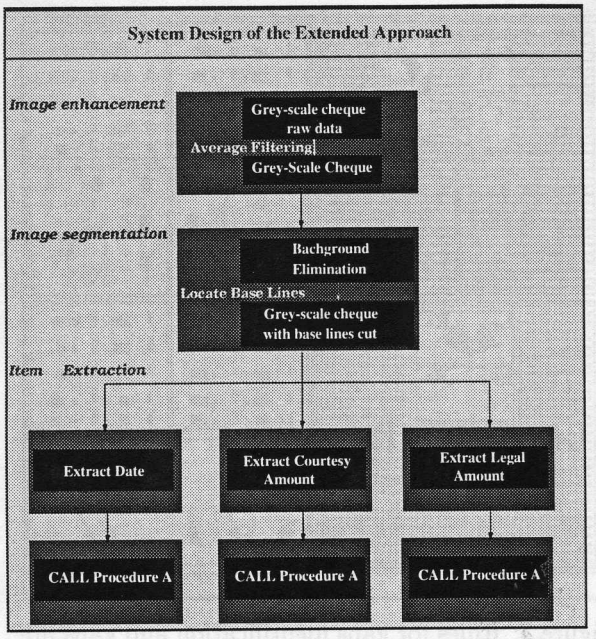

Figure 5: System Design of the Extended Approach.

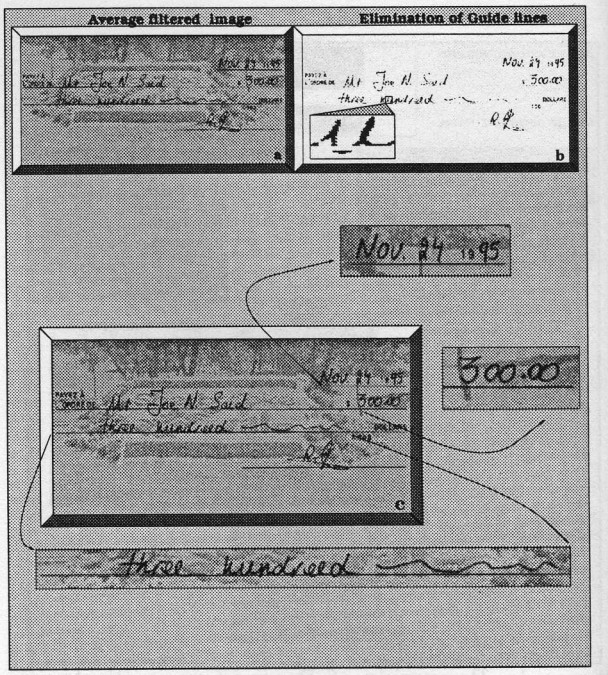

Figure 7: The new system and early extraction.

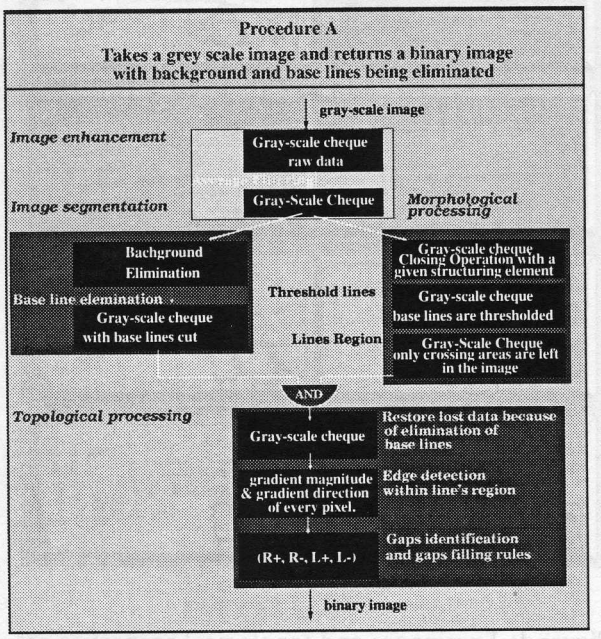

Figure 6: Procedure A.

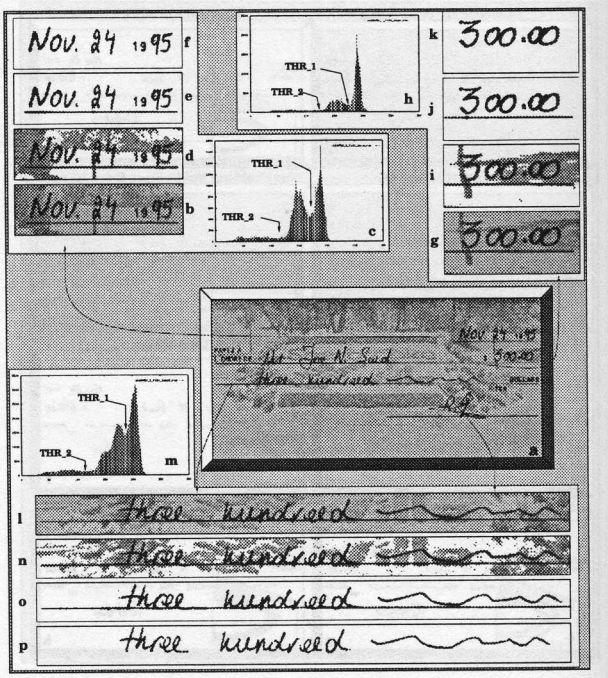

Figure 8: Local processing of extracted information.