Automatic Segmentation and Recognition of Bank Cheque Fields

Vamsi Krishna Madasu,
School of ITEE, University of Queensland
madasu@itee.uq.edu.au

Brian Charles Lovell
NICTA & School of ITEE, University of Qld.
lovell@itee.uq.edu.au

Abstract

This paper describes a novel method for automatically segmenting and recognizing the various information fields present on a bank cheque. The uniqueness of our approach lies in the fact that it doesn't necessitate any prior information and requires minimum human intervention. The extraction of segmented fields is accomplished by means of a connectivity based approach. For the recognition part, we have proposed four innovative features, namely; entropy, energy, aspect ratio and average fuzzy membership values. Though no particular feature is pertinent in itself but a combination of these is used for differentiating between the fields. Finally, a fuzzy neural network is trained to identify the desired fields. The system performance is quite promising on a large dataset of real and synthetic cheque images

1. Introduction

The widespread use of bank cheques in daily life makes the development of cheque processing systems of fundamental relevance to banks and other financial institutions. Bank transactions involving cheques are still increasing throughout the world in spite of the overall rapid emergence of electronic payments by credit cards [1]. However, fraud committed in cheques is also growing at an equally alarming rate with consequent losses [2]. According to the American Banker Association's (ABA) 1998 Cheque fraud survey, financial institutions alone incurred $512.3 millions in cheque fraud losses. Automatic bank cheque processing systems are hence needed not only to counter the growing cheque fraud menace but also to improve productivity and allow for advanced customer services.

The automatic processing of a bank cheque involves extraction and recognition of handwritten or user entered information from different data fields on the cheque such as courtesy amount, legal amount, date, payee and signature [3]. This is a formidable task and requires efficient image processing and pattern recognition techniques. The only two fields on a cheque that can be processed automatically with near-perfect accuracy by character recognition systems are the account number and the bank code as they are printed in magnetic ink. The other fields may be handwritten, typed, or printed; they contain the name of the recipient, the date, the amount to be paid (textual format), the courtesy amount (numerical format) and the signature of the person who wrote the cheque. The official value of the cheque is the amount written in words; this field of the cheque is called "legal amount". The amount written in numbers is supposed to be for courtesy purposes only and is therefore called "courtesy amount". Nevertheless, most non-cash payment methods use only the amounts written in numbers. The information contained in a cheque is frequently handwritten, especially considering that most of the cheques that were written by computer systems have been gradually replaced by newer methods of electronic payment. Handwritten text and numbers are difficult to read by automatic systems (and sometimes even for humans); so cheque processing normally involves manual reading of the cheques and keying in their respective values into the computer. Accordingly, the field of automatic cheque processing has witnessed sustained interest for a long time. This has led to complete systems with reading accuracy in the range of 20–60% and reading error in the range of 1–3% beginning to be installed in recent years [5].

The performance in handwriting recognition is greatly improved by constraining the writing, which addresses the problem of segmentation and makes the people write more carefully. Nevertheless, banks are not willing to change the format of the cheques to impose writing constraints such as guidelines or boxes to specify the location where each particular piece of information should be recorded. Instead they are interested in reducing the workload of the employees manually reading the paper cheque. Since employees also make mistakes reading or typing the amount of the cheques, a single manual read rarely drives the whole process. A system that is able to read cheque automatically would be very helpful, especially if it is fast and accurate. Even if misclassification occurs, the mistake could potentially be detected during the recognition process; however it is more desirable that the system rejects a cheque in case of doubt so that it can be directed to manual processing from the beginning.

In order to produce a successful cheque processing system, many sub-problems have to be solved such as background and noise removal, recognition of the immense styles of handwriting and signatures, touching and overlapping data in various fields of information and errors in the recognition techniques [3]. Other works include systems implemented to read courtesy amount, legal amount and date fields on cheques [4, 8, 9, 10, and 12]. However, till now, most of the handwritten cheques have to be processed with substantial human intervention due to high recognition errors, reflecting the fact that automatic recognition of handwritten data on bank cheques is a challenging task. Another main drawback of these systems is that, for each filled bank cheque, the recognition system has to maintain an unused bank cheque image sample requiring, therefore, a large memory size for each bank cheque. Apart from requiring additional storage these systems fail to process cheques that are not in their database.

2. Pre-processing of Bank Cheques

Before the segmentation phase, one of the most important steps is the pre-processing of bank cheques. This involves background elimination and baseline removal techniques to maintain the physical integrity of the rest of the cheque image information. Liu et al. [11] described a simple and robust solution for the extraction of baselines from bank cheques. On the other hand, thresholding techniques like Otsu's Method that gives good results for background elimination.

2.1. Background Elimination

One of the most important aspects of the threshold selection is the capability of identifying the peaks reliably in a given histogram. This capability is particularly important for automatic threshold selection in situations where image characteristics can change over a broad range of intensity distribution.

One approach for improving the shape of histograms is to consider only those pixels that lie on or near the boundary between objects and the background. But this information is clearly not available during segmentation, however, an indication of whether pixel is on an edge maybe obtained by computing its gradient. In addition, use of the Laplacian can yield information whether a given pixel lies on the background or object side of an edge.

The gradient and the Laplacian can be used to form a three level image as follows:

$$S(x,y) = \begin{cases} l; \nabla(pixel) < T \\ m; \nabla(pixel) \geq T \cap \nabla^2(pixel) \geq 0 \\ n; \nabla(pixel) \geq T \cap \nabla^2(pixel) < 0 \end{cases}$$

where,

T : the threshold value used for detecting the boundaries

l, m, n : are the assigned values of the gray levels for classifying images

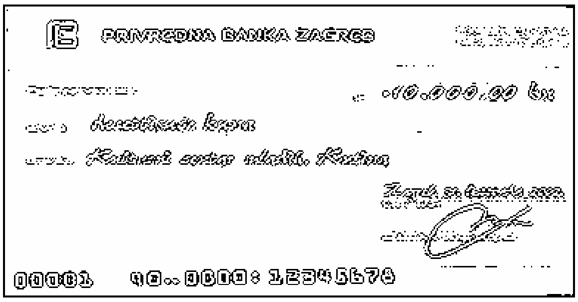

Figure 1. Background elimination using boundary characteristics

2.2. Noise elimination

We employ three morphological operators to get rid of stray marks and other isolated noisy blots which may cause problems in later stages of cheque processing.

Clean: removes isolated pixels such as individual ones surrounded by zeros.

H-break: removes H-connected pixels as illustrated below.

1	1	1
0	1	0
1	1	1

→

1	1	1
0	0	0
1	1	1

Figure 2. H-break morphological operator

Line Masking: removes horizontal and vertical lines. The structuring element is a string of ones of appropriate length, placed either horizontally or vertically.

2.2. Removal of lines

For line removal, we employ Radon Transform. The Radon transform computes projections of an image matrix along specified directions. A projection of a two dimensional function $f(x, y)$ is a line integral in a certain direction. For example, the line integral of $f(x, y)$ in the vertical direction is the projection of $f(x, y)$ onto the x-axis: the line integral in the horizontal direction is the projection of $f(x, y)$ onto the y-axis. Projections can be computed along any angle. In general, the Radon transform of $f(x, y)$ is the line integral of f parallel to the y' axis and is given as,

$$R_\theta(x') = \int_{-\infty}^{\infty} f\left(x'\cos\theta - y'\sin\theta, x'\sin\theta - y'\cos\theta\right)dy'$$

where,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

3. Segmentation and Extraction of Fields

In order to recognize the various information fields, we should segment the image into the target object regions and extract them one by one. In that case, coarse region segmentation that ignores small parts in each object region must be performed. The sliding window method in [8] can be used for coarse region segmentation after certain modifications.

ALGORITHM 1

1. The cheque image must be first converted to binary mode image.
2. The threshold image must be subjected to line masking techniques to eliminate all horizontal and vertical lines.
3. The width of sliding window is set to some initial value. The entire cheque image is then traversed by sliding this window and calculating the density at each step. The entropy is simply calculated using the formula:

   Entropy (E) = - (pixel density) × log (pixel density).

   The entropy is a better choice than density because it introduces a larger range of values so segmentation is easier and more accurate. A modified image of the cheque is then constructed by making use of the entropy calculated in the previous step.
4. The image obtained in step 3 is then subjected to optimal thresholding using Otsu's method as shown in figure 3.
5. Finally after thresholding we get an image where the coarse regions can be clearly identified.

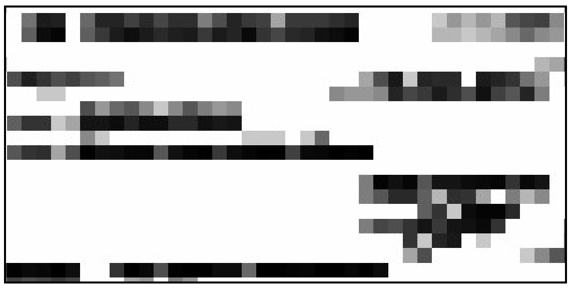

Figure 3. Entropy based image

3.2. Connected Component Labeling

Finding all equivalence classes of connected pixels in a binary image leads to what is called *connected component labeling*. The result of connected component labeling is another image in which everything in one connected region is labeled "1" (for example), everything in another connected region is labeled "2", etc.

We now describe the algorithm that we have employed for labeling the coarse segmented regions on the cheque and the extraction of the various information fields.

ALGORITHM 2

1. Scan through the image pixel by pixel across each row in order:
2. If the pixel has no connected neighbors that have already been labeled with the same value, create a new unique label and assign it to that pixel.

3. If the pixel has exactly one label among its connected neighbor that has already been labeled with the same value, give it that label.
4. If the pixel has two or more connected neighbors with the same value but different labels, choose one of the labels and remember that these labels are equivalent.
5. Resolve the equivalencies by making another pass through the image and labeling each pixel with a unique label for its equivalence class.

The above-mentioned algorithm was employed to locate and extract the regions of interest, or in other words the regions containing the useful information. Shown below are the images of labeled regions of interest in the cheque.

Figure 4. Labeled regions of interest

4. Recognition of Bank Cheque Fields

The final step in automatic bank cheque processing is to recognize the segmented and extracted fields of interest on any typical bank cheque. In this work, we have tried to classify the most important information fields which are the handwritten signature, bank logo, machine printed text and lastly the numerical amount. To achieve a fairly high recognition rate, it is essential that the features selected are able to discriminate the different information fields accurately.

Initially, we used the ‘Differential Box Counting’ (DBC) method proposed by Choudhari & Sarkar [7] to calculate the fractional dimensions of the data fields. This choice was motivated by the observation that fractal dimension is relatively insensitive to image scaling. Although DBC method is faster and more accurate than other box counting approaches, the results obtained were unconvincing, probably because fractals are based on the concept of self-similarity that was not evident in the fields which we had extracted from the cheque.

We then employed four features of our own which were then fed to a neural network for the purpose of identification. The features considered are:

Fuzzy Features: The information present on a bank cheque can be fuzzified and represented by a membership function which is defined by a Gaussian type function. The motivation behind taking fuzzy features is to capture the effect of neighboring pixels on the current pixel in a window. Here, the spatial arrangement of gray levels over a window is considered. A suitable window size is assumed. Since we do not know *a priori* how the gray levels are distributed in the extracted image, we considered each pixel with its relative response over all the neighbouring pixels in the specified window. A membership function to this effect is defined by the following equation:

$$u_j(i) = \exp\left[-\left\{ \frac{x(j) - x(i)}{b} \right\}^2 \right]$$

where,

$x(i)$ is the gray level of current pixel

$x(j)$ is the gray level of the neighboring pixel

b is the fuzzifier equal to size of the window

The cumulative response of the current pixel is given the weighted sum method which is defined by the expression:

$$y(i) = \sum_{j=1}^{n} \frac{\mu_j(i) \cdot x(i)}{\sum_{j=1}^{n} \mu_j(i)}$$

This constitutes the defuzzified response of the current pixel. This process is repeated for all pixels lying within a window.

Entropy: This feature is a measure of information contained in the extracted data field. The entropy of an information field is calculated using the following equation:

$$E_i = -P_i \log P_i$$

where, P_i is the pixel density in a window numbered i.

Energy: This feature gives the measure of the energy contained in the extracted data field and is related to the number of black pixels present in the image. The energy of the data fields is calculated using the equation,

$$E = \left(\sum_j \sum_i x_{ij}^2 \right) \Big/ \left(\sum_j \times \sum_i \right)$$

Aspect Ratio: is the ratio of length to breadth of the extracted data field, i.e.,

Aspect Ratio = (Width of the extracted field) / (Height of extracted field)

Table 1. Feature values of different information fields of a bank cheque

Feature	Bank Logo	HW Text	Sign	MP Text	Num- -erals
Fuzzy Values	0.6428	0.8468	0.8468	0.5742	0.7145
Entropy	0.1567	0.1247	0.1200	0.1566	0.1576
Energy	0.3572	0.1530	0.1400	0.4435	0.3089
Aspect Ratio	1.5833	10.545	2.7600	24.600	8.6667

5. Implementation and Results

The feature vector consisting of the four extracted features for various data fields were used to train a fuzzy integral based neural network [6]. The essence of feature vector approach is that when the features are clubbed together, they are more helpful in classification as compared to any of the features used alone.

The algorithm for implementing the neural network is briefly described below.

ALGORITHM 3

1. For each pixel, calculate its membership function over all the neighboring pixels in the specified window (3X3). If 'S' is the total number of pixels in the extracted field, compile the values thus obtained into an SX8 vector.
2. Each of the columns thus obtained will give rise to an input vector. Hence, there should be 8 input vectors.
3. Assign random values to fuzzy measure densities corresponding to each input value using a random number generator.
4. Assume that the aggregate information from all input source, **h (.),** is a linear function.
5. For each input vector, calculate the choquet fuzzy integral using the equation:

$$E_g(h) = \sum_{i=1}^{n} h(x_i)[g(A_i) - g(A_{i+1})]$$

6. The back propagation algorithm is then used for the learning process. Calculate the sum of the squared error using the following equation:

$$E = \sum_k E_k = \sum_k \left[\sum_i \left(f_i^k - Y_i^k \right)^2 \right]$$

7. Optimize the neural network by minimizing E with respect to the synaptic weights (fuzzy densities) of the network to obtain a new set of fuzzy measure densities.
8. Repeat step 7 till we obtain the desired error tolerance.

Owing to privacy and confidentiality laws, there are no publicly available standard or benchmark cheque databases to apply different techniques or to perform a comparative analysis. Hence, we implemented the neural network on a database of cheque images scanned by our research team and also cheques provided to us by an American cheque processing company under a non-disclosure agreement. The size of the database is fairly large with over 900 signatures obtained from different bank cheques in US, Europe and Asia thereby reflecting different cheque patterns and styles. We also performed a comparative analysis of our method with an earlier reported technique in literature to test the efficacy of the proposed algorithms.

The neural network was trained to classify the various extracted data fields based on the patterns present in the feature vector and the final decision was made on the basis of those values. The training set consisted of one sample of each type of bank cheque whereas the rest of the samples constituted the testing set. The system performed quite well when all the features were used in combination with each other (see table 2) as compared to just fuzzy or fractal features.

Table 2. Performance evaluation of recognition Systems

Method	Database size	Error rate
Fractals	300	45%
Fuzzy features	923	20%
All features	923	10%

We achieved a maximum classification error rate of 10% when all the information fields on each cheque were considered. While calculating the error in classification, we considered only those cheques whose the fields were segmented properly. An error in segmentation was determined on the basis of comparison with the sample cheque image used for training. On the other hand, errors in classification were calculated on the basis of thresholds which were set for different features of the cheque image. A cheque was classified as true when the mean of feature or features were inside the acceptable range. However, when the feature values were outside the range, they were taken as errors in recognition.

6. Conclusions

Segmentation of different fields such as text, signature, courtesy amount and logo is attempted in this work. We use the sliding window to estimate the pixel density through the entropy function. The traversed regions are then thresholded by Otsu's method and then labeled. Fuzzy based four texture features, viz., fuzzy membership values, energy, entropy, aspect ratio are extracted from these labeled regions. Treating these features as input sources, we constructed a neural work that evaluates the Choquet fuzzy integral by aggregating the information from the input sources giving rise to an output useful for identifying the fields. The results obtained substantiate the fact that different fields can be segmented and recognized without the need for any *a priori* information about the cheque data fields. This is quite significant for the creation of new age automatic bank cheque systems.

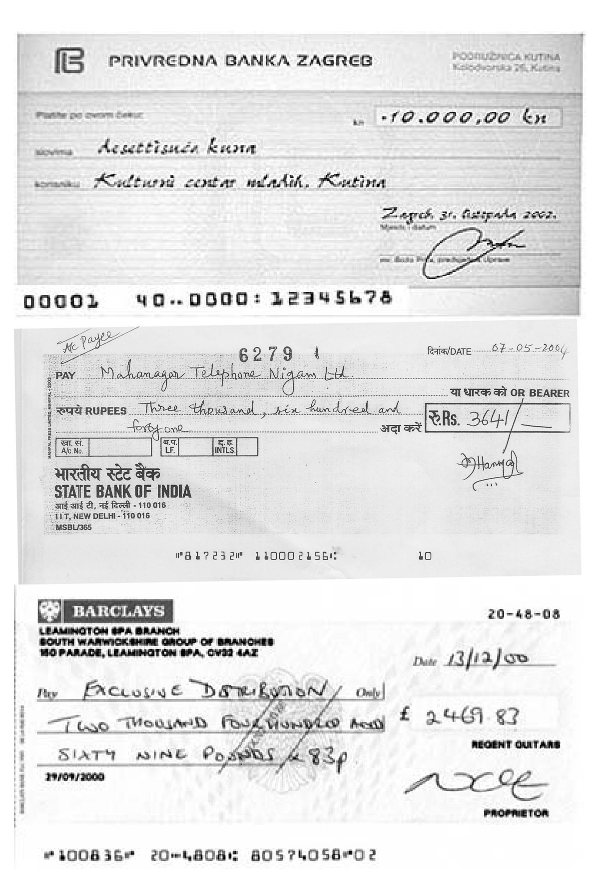

Figure 5. Samples from Cheque Database

7. References

[1] J. Blodgett, "Beyond cheque image statements: A new strategy for the 1990s", Advance Imaging, Vol. 9, pp. 73-75, 1994.

[2] T.G. Holland, "Cheques and Balances", Security Management, Vol. 43, pp. 76-82, 1999.

[3] C.Y. Suen, Q. Xu and L. Lam, "Automatic recognition of handwritten data on cheques – Fact or Fiction?" Pattern Recognition Letters, Vol.20, pp. 1287-1295, 1999.

[4] R. Fan, L. Lam, and C.Y. Suen, "Processing of date information on cheques", Progress in Handwriting Recognition, World Scientific, pp.473-479, 1997.

[5] S. Kelland and S. Wesolkowski, "A comparison of Research and Production Architectures for Cheque Reading Systems", Proceedings of the 5th International Conference on Document Analysis and Recognition, pp. 99-102, 1999.

[6] J-H Chiang, "Choquet fuzzy integral based hierarchical networks for decision analysis", IEEE Transactions on Fuzzy Systems, Vol.7, no.1, 1999 pp. 63-71.

[7] B.B. Choudhari and N. Sarkar, "Texture segmentation using fractal dimension", IEEE Transactions on PAMI, Vol.17, No.1, 1995, pp. 72-77.

[8] V.K. Madasu, M.H.M. Yusof, M. Hanmandlu and K. Kubik, "Automatic extraction of signatures from bank cheques and other documents", Proceedings of DICTA, vol. 2, pp. 591-600, 2003.

[9] K. Han and I. Sethi, "An off-line cursive handwritten word recognition system and its application to legal amount interpretation", Automatic Bankcheque Processing, World Scientific, pp. 295-308, 1997.

[10] K. Liu, C.Y. Suen, and C. Nadal, "Automatic extraction of items from cheque images for payment recognition," Proceedings of the Intl. Conference on Pattern Recognition, pp. 798–802, 1996.

[11] K. Liu, C.Y. Suen, M. Cheriet, J.N. Said, C. Nadal, and Y.Y. Tang, "Automatic Extraction of Baselines and Data from Cheque Images," International Journal of Pattern Recognition and Artificial Intelligence, Vol. 11, pp. 675-697, 1997.

[12] S. Impedovo, P. Wang and H. Bunke (eds), **Automatic Bankcheque Processing**, World Scientific Publishing Co, Singapore, 1997.