IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-4, Special Issue-1 E-ISSN: 2347-2693

Handwritten Character Recognition from Bank Cheque

Siddhartha Banerjee*
Department of Computer Science
Ramakrishna Mission Residential College,
Narendrapur, Kolkata-700103
sidd_01_02@yahoo.com

Bibek Ranjan Ghosh
Department of Computer Science
Ramakrishna Mission Residential College,
Narendrapur, Kolkata-700103
bibekghosh2003@yahoo.co.in

Arka Kundu
Department of Computer Science
Ramakrishna Mission Residential College,
Narendrapur, Kolkata-700103

***Abstract* –Handwritten Character Recognition (HCR) is the ability of a computer to receive and interpret handwritten input captured by as digital cameras or other devices.Recognition of handwritten characters by computer is a serious problem because there are a many variety of writing styles, character shapes written by different individuals. The main objective of our project is to recognize the handwritten characters present in a bank cheque. Paper cheques still play a big role in the non-cash transactions in the world even after the arrival of credit cards, debit cards and other electronic means of payment. Huge volumes of handwritten bank cheques are processed manually every day in developing countries. The present cheque processing procedure requires a bank employee to read and manually enter the information present on a cheque (or its image) and also verify the entries like signature and date. An attempt is made in this project to recognize the characters present payee name and in the cheque amount by using image processing techniques on handwritten cheque images. The system uses broad steps like thresholding, image segmentation, thinning and pattern matching for extraction of characters. The pattern matching is done using graph based method. Graph matching techniques areintroduced to compute the similarity of characters extracted from bank cheque with theinformation of characters store in the database.**

Keywords: *Segmentation, Thinning, Scaling, Pattern matching, COG, complete Bipartite Graph, Adjacency matrix*

I. INTRODUCTION

Handwritten Character Recognition (HCR) is the ability of a computer to receive and interpret handwritten input taken from digital cameras or other devices. Handwriting recognition is defined as the task of transforming a language represented in its spatial form of graphical marks into its symbolic representation. HCR system consists of a number of stages which are preprocessing, feature extraction, classification and followed by the actual recognition.

Recognition of handwritten characters by computer is a serious problem because there are a many variety of writing styles, character shapes written by different individuals. It makes reading difficult for humans. Most of the researchers tried to solve problem based on the image processing and pattern recognition technique.

Paper cheques still play a big role in the non-cash transactions in the world even after the arrival of credit cards, debit cards and other electronic means of payment. Huge volumes of handwritten bank cheques are required to be processed manually every day in developing countries. The present cheque processing procedure requires a bank employee to read and manually enter the information present on a cheque (or its image) and also verify the entries like signature and date.

An attempt is made in this project to develop an automated system to recognize the characters present in the payee name and in the cheque amount from a handwritten bank cheque image using image processing techniques. After this short introduction, the proposed method is described in section II and result is discussed in section III. The concluding remarks are given in section IV.

II. PROPOSED METHODOLOGY

The steps that are followed to recognize the handwritten characters from the bank cheque image is describe by the flowchart in Fig. 1. It consist two broad steps:

- Preprocessing and
- Character recognition.

A. Preprocessing:

The image of the bank cheque captured by camera may contain some noise. To remove these noises present in the image and make it compatible for further processing we apply the following steps.

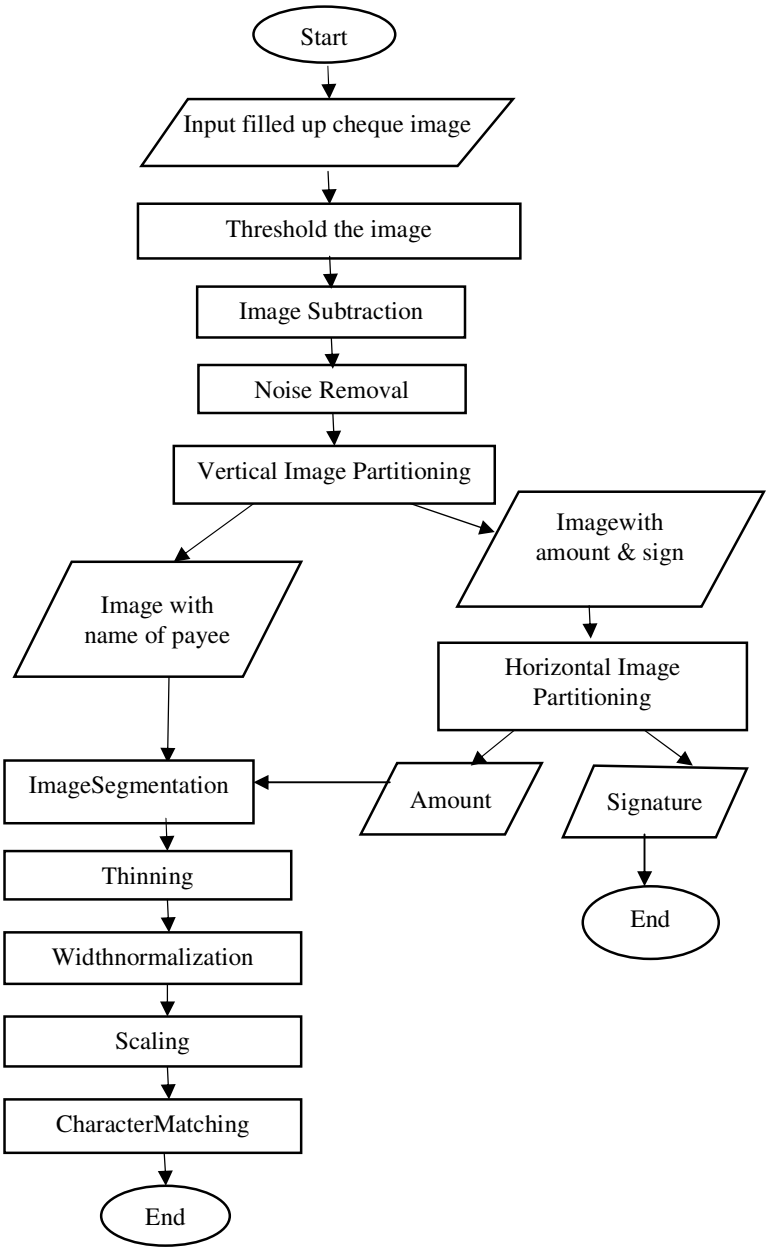


Fig. 1: Steps of Proposed Method

Thresholding – Two gray scale images of blank bank checque and filled up cheque are taken as input. Otsu [6] thresholding is used for converting both the image from grayscale to binary image. In Otsu method the threshold is selected by nonparametric and unsupervised manner. This is an automatic threshold selection technique from gray level histogram. It is an iterative process through all the possible threshold values and finds a measure of spread for the pixel levels each

side (background and foreground) of the threshold. The final threshold is the minimum value of the sum of background and foreground spreads. In this method the weight, mean variance and within class variance are calculated for all possible thresholds and then find the threshold that minimizes the weighted within-class variance. The result is shown in Fig. 2.

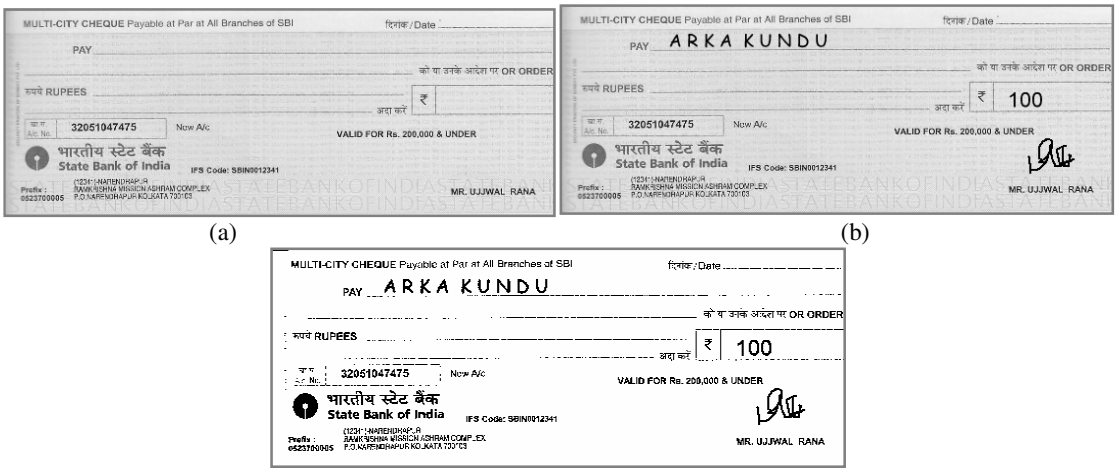


Fig 2: (a) Gray level image of Blank Cheque (b) Gray level image of Filled Up cheque (c) Binary image of Blank cheque and (d) Binary image of Filled up cheque.

Image Subtraction – The binary image of blank cheque is subtracted from the binary image of filled up cheque. The method simply subtracts the two images such that any pixel which is black in filled up cheque but not in the blank one is kept black in the resultant image, everything else is kept white. This step will remove the field names of a standard cheque and will leave only what the user wrote on the filled up cheque, nothing else. The result of image subtraction is shown in Fig. 3.

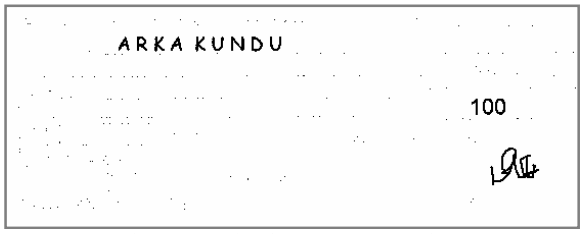


Fig 3: Result of Image subtraction.

Noise Removal – It is clear from the Fig. 3 that the subtracted image contains the fields filled up by hand written character with some noise. The median filter is used to remove the grainy noises from the subtracted image. The result is shown in Fig 4.

Fig 4: Result of Image subtraction

Partitioning – Now the need is to partition the image into sub-images containing the field data like name, account number, signature separately. Knowing the standard structure of the bank cheque, partition can be doe easily by cutting the image vertically and horizontally into corresponding parts. One thing should be kept in mind that for different banks, partitioning

delimiters will be different. In this experiment the structure of the default SBI chequeis considered. Fig 5 shows the result of partitioning.

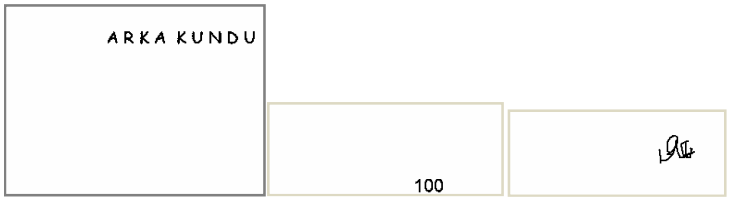


Fig. 5: Result of partitioning.

Segmentation - After partitioning we will need to segment each sub image to obtain characters within it. For Image segmentation region growing algorithm [7] given in ALGORITHM 1 is followed to level each black region with a level. Then separate image is created for each level. The application of this step on Figure 5 is shown in Figure 6.

ALGORITHM 1:

Step 1: An arbitrary black pixel (r, c) known as s*eed* pixel is chosen from the image.
Step 2: This pixel is given a unique label *L.*
Step 3: 8-neighbouring pixel of the seed pixel is examined. The neighbor(s) which are black are also assigned the unique number *L.* Once a new pixel is accepted as a member the neighbors of this new pixel are examined. This process goes on recursively until no more pixels is accepted. All the pixels of the current region are marked with *L.*
Step 4: Repeat Step 1 to Step 4 until every pixel is assigned to some region.

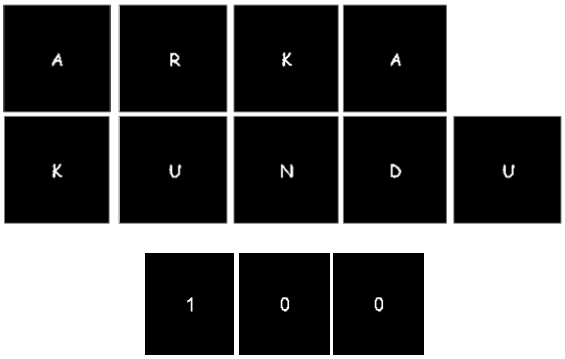


Fig. 6: Segmentation of characters (image negative)

Thinning - The segmented image is then thinnedto get single pixel skeleton of the character. The standard thinning algorithm [1] by using a set of eight structuring element is used for this purpose.

Width normalization - This step is performedto crop the rectangular area containing the charcter and eliminating the black space. For these purpose we apply the following algorithm.

ALGORITHM 2

Step 1: Scan the image from left to right to findthe column C_1 that contains the first black pixel.
Step 2: Scan the image from right to left to findthe column C_2 that contains the first black pixel.
Step 3: Scan the image from top to bottom to findthe row R_1 that contains the first black pixel.
Step 4: Scan the image from bottom to top to findthe row R_2 that contains the first black pixel.
Step 5: Extract the image from pixel (R_1, C_1) to(R_2, C_2).

Figure 8: After width normalization of Figure 7.

Scaling - After width normalization we getonly the rectangular area containing the character. But the size of the character may vary from number plate of different vehicles. We apply standard scaling [1] operation to make all the character of same size.

B. Character Recognition

After Scaling the characters are recognized by the following steps [5].

Feature Extraction - After preprocessingstep, we select a set of representative points, known as feature points, against which the properties of the given character are compared with the features of characters stored in database. The set of feature points are extracted depending on Centre of Gravity (COG) [7] of the character using the following algorithm. Fig. 9 shows the calculated feature points.

ALGORITHM 3
Step 1: Calculate the COG of the character image.
Step 2: Calculate set of feature points of theimage(s) using following steps:
2.1: Draw a horizontal line through the eachCOG to break each image into two sub-images.
2.2: Calculate COG of the signature belongingto each of the sub-image.
2.3: Draw a vertical line through each of theCOG to break each sub-image into two sub parts.
2.4: Calculate COG of the signature belongingto each of the sub-image of Step 2.3.
Step 3: Repeat Step 2 for N number of times.
Step 4: Consider the COGs obtained in Step 2.2and Step 2.4, in each iteration as set of feature points.

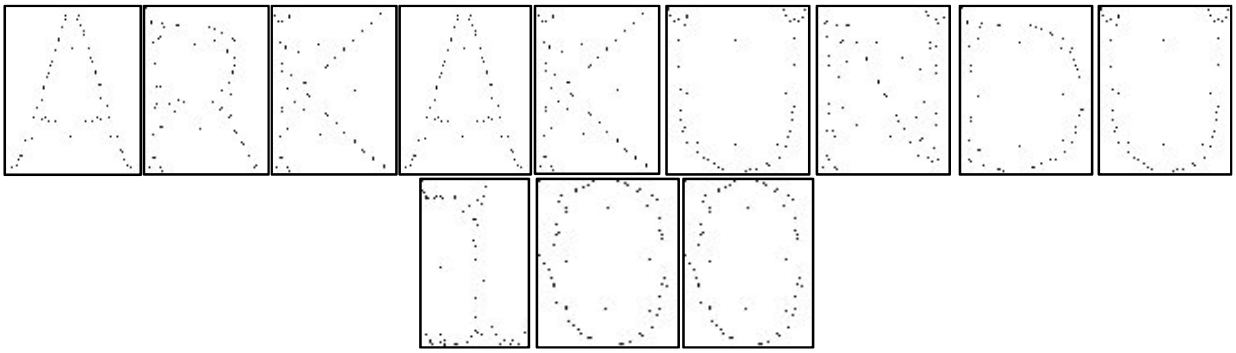
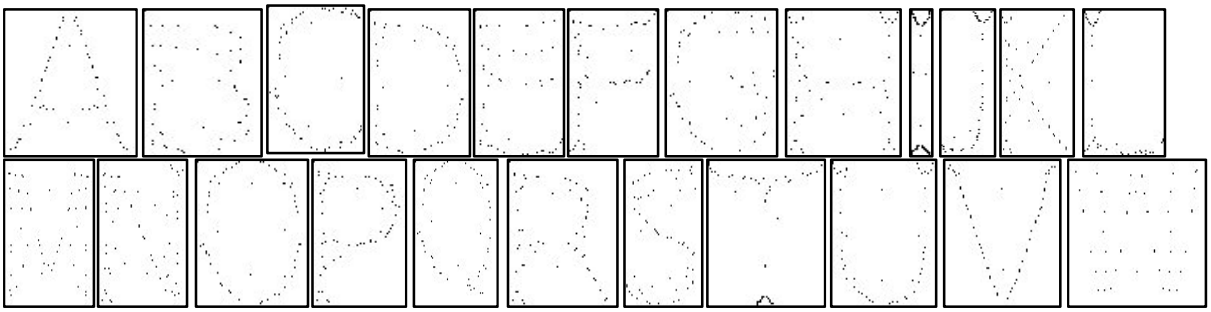


Fig. 9: Calculated feature points.

Matching – Thelast and final technique of our project is character matching. To detect the handwritten characters we stored the digitized form of all the 26 alphabets and 10 different numbers in a specific database. These characters are then thinned, normalized and scaled to the same size as of the handwritten characters. Then the feature points of each and every character is found out and stored in the database. These feature points are then matched with the feature points of each handwritten character. If more than 80% of the feature points of a handwritten character matches with any of the feature points of a character present in the database, then the handwritten character is recognized as that character of the database.The database used had these character feature point set is shown in Fig. 10.



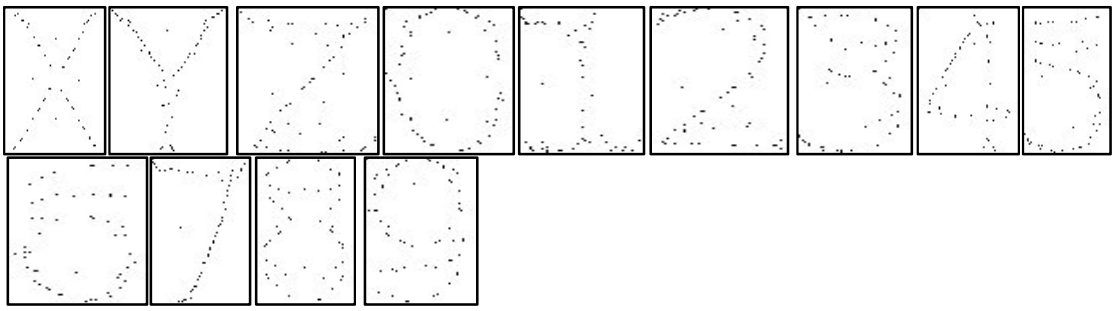


Fig. 10: Feature points of alphabets and digits stored in the database.

III. RESULT

In this system, Recognition of handwritten characters is still limited to certain handwriting styles. Due to the presence of large variation of shape, scale and format in hand written characters, this system sometimes may not give desired results. This system will be able to recognize only those handwritten characters if the style of writing is similar to the digitized characters stored in the database. We collect 150 filled up cheques from 100 different people and applied our proposed method. The method is able to recognize the characters with 87.2% accuracy.

IV. CONCLUSION

The proposed method is able to recognize hand written characters with limited accuracy. If the database is increased to store the feature points of characters and digits written in different style the accuracy can be increased. Proposed method can recognize handwritten characters if they are written separately. But if the characters are written such a way that the characters have connection between them, this method fails to recognize. Our next project is to remove these limitations.

REFERENCES

[1]. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd Edition, Prentice Hall, 2009.
[2]. Youssef EsSaady, Ali Rachidi, Mostafa El Yassa, DrissMammass, "AmazighHandwritten Character Recognition based on Horizontal and Vertical Centerline of Character", International Journal of Advanced Science and Technology Vol. 33, August, 2011
[3]. R. Palacios, A. Gupta, P. S. Wang, "Handwritten Bank Cheqe Recognition",International Journal of Image and Graphics, Vol. 4, No. 2 (2004)
[4]. R. Jayadevan, S. R. Kolhe, P. M. Patil, U. Pal, "Automatic processing of handwritten bank cheque images: a survey"
[5]. B. R. Ghosh, S. Banerjee, S. Dey, S. Ganguli, S. Sarkar, "Off-Line Signature Verification System Using Weighted Complete Bipartite Graph", 2nd International Conference on Business and Information Management (ICBIM), ISBN : 478-1-4799-3264-1/14/$31.00 ©2014 IEEE pp.109-113.
[6]. Otsu, N.: A threshold selection method from gray-level histograms, IEEE Trans. Sys., Man., pp.62–66.
[7]. B. Chanda and D. D. Majumder, Digital Image Processing and Analysis. Prentice Hall India, 2009.