Handwritten Character Recognition - A Comprehensive Review

Mahesh Goyani, Harsh Dani, Chahna Dixit
Gujarat Technological University
Ahmedabad, India
mgoyani@gmail.com, daniharsh28@gmail.com, chahdixit@gmail.com

Abstract— Character recognition comes into picture when various patterns of handwritten or optical characters are to be recognized digitally. Many researchers have proposed different approaches for character recognition in different languages. In this paper, we have reviewed several techniques of character recognition. The main important phases of character recognition include pre-processing, segmentation, feature extraction and classification. Various feature extraction techniques and classification techniques have been surveyed in this paper and an attempt is made here to draw a conclusion regarding HCR techniques from the literature survey.

Index Terms— HCR, Segmentation, Neural Network, Feature Extraction, Classification

I. INTRODUCTION

Since writing styles of people vary to a great extent, it becomes practically difficult for the computer to recognize the handwritten characters. To fulfill this task, handwritten character recognition comes into picture. The extensive application of character recognition is in recognizing the characters in bank cheques and transaction forms, 3D object recognition, car plates, automatic text entry for desktop publication, business card reading, library cataloguing, ledgering, automatic sorting of postal mail, ZIP code, automatic invoice processing [1, 2, 3]. The automation of entire process requires a high recognition rate, as well as maximum reliability [4].

Character recognition is classified into two categories: Online character recognition and Offline character recognition [5, 6]. Offline recognition means the text written on a paper is captured optically which is then available as an image. Online recognition means the text written on any digital device is represented as a function of time and the order of strokes [5, 6]. Online character recognition is more robust compared to offline character recognition since additional information about various styles of writing, pressure, order of strokes etc is available in case of online character recognition.

Based on the type of input image for the recognition system there are two types of character recognition: Handwritten Character Recognition (HCR) and Optical Character Recognition (OCR). In handwritten character recognition system, input image is handwritten text; the text may be written by a person with hands or by any digital devices like stylus, digital pen etc. Whereas in optical character recognition system, the characters in the input image are not written by hand; rather the characters may be printed, entered through any input device like keyboard. In Handwritten Character Recognition, the characters in any single input image may or may not have any fixed font size or style, since it depends on the handwriting of a person. While in Optical Character Recognition, the characters have a definite pattern and fixed font size and styles.

Character recognition systems consist of four major phases: Pre-processing, Segmentation, Feature Extraction, and Classification [6, 7, 8, 9, 10, 11]. There are many classification techniques of character recognition such as template matching [6, 10, 12], artificial neural networks (ANN) [5, 10, 13, 14, 15], syntactical analysis, wavelet theory, hidden Markov models (HMM) [9, 16, 17, 18, 19, 20], Bayesian theory [21, 22] and minimum distance classifiers etc [1]. A lot of methods have been developed for character recognition, among them the neural network is mainly used due its classification efficiency and ease of use [21]. Research has been made in the field of HCR considering varieties of languages like English, Devnagari, Arabic, Thai, Chinese, Tamil and many more.

The structure of the paper is as follows: The working principle of handwritten character recognition followed by literature survey of the same. Conclusions are made in the later part.

II. THE WORKING PRINCIPLE

The process for HCR involves the following steps as shown in figure 1.

Figure 1. Process of HCR
2.1. Image Acquisition: The input image for the recognition system is obtained through a scanner in the form of a scanned image [23]. Hence, the handwritten characters format is given as input to the system and a
digital form of the input characters is achieved in the form of image as output after image acquisition.

2.2. Pre-Processing: Pre-processing is done on the image to make it suitable for the recognition process. Basically, it includes removal of unwanted noise and inconsistent data. The sub-processes involved in pre-processing are binarization, noise reduction, normalization and skew correction, thinning and slant removal [16, 18].

Binarization refers to the conversion of the scanned image into gray scale image and then into binary image. It is also known as thresholding [6] or digitization. Extraction of foreground (ink) from the background (paper) is called as thresholding [19].

Noise may be introduced due to poor scanning. Spatial domain filters and frequency domain filters may be used for noise removal. Median filtering, Wiener filtering, Butterworth low pass filter, Gaussian low pass filtering and morphological operations can be performed to reduce noise [14, 24, 25].

Normalization is the process of converting a random sized image into an image of standard size [16].

Thinning is the process by which a one-pixel width representation of an image is obtained. The correctness of the image and its endpoints are preserved in thinning [6]. Thinning process simplifies the further analysis. The slant is the deviation of the strokes from the vertical direction, depending on the writing styles. Estimation and correction of the slant angle is therefore necessary.

2.3. Segmentation: Segmentation means the separation of clear character area from non-character area in the given image. Segmentation includes thresholding of the image, skeletonization of the image and the final step of segmentation which involves pruning [13].

Skeletonization is done to reduce the character area to just one pixel line. Due to skeletonization some unwanted branches may develop on the output skeletons. So in the final step of segmentation, these spurs are removed with the help of pruning [13].

2.4. Feature Extraction: Feature extraction is done to extract unique features of different characters for its recognition process. In feature extraction stage every character is assigned a feature vector to identify it. This vector is used to distinguish the character from other characters [14].

Feature extraction techniques fall among three categories. 1. Statistical methods 2. Structural methods [16, 17, 21, 26] 3. Global Transformation and Series Expansion Features. Statistical method is based on estimating how data are collected and selected. It is based on probability theory and hypothesis. Structural method involves extraction of such values that contains some information about the structure of the character [21]. Global Transformation and Series Expansion are invariant to global deformations like translation and rotations [27].

Various examples of Statistical methods include profiles, distances and crossings, zoning [12, 17, 26], belt shape pixel number [22], projection histogram [6, 10, 12, 17, 18, 23], celled projections , Fourier Transform [10, 11, 12] whereas Aspect Ratio [17], Cross points, Loops [26], Branch points, Strokes and their directions, inflection between two points, Horizontal curves at top and bottom, chain codes, Extremum points, Zero crossings, isolated dots, Profile on the left or right, Symmetry of character, width of a stroke, spline approximation of strokes etc are the Structural methods [6, 21, 27]. Various Examples of Global Transformation and Series Expansion feature involves Walsh-Hadamard Transform, Hough Transform [16], Gabor Transform [10, 12, 28], Wavelets, Moments, Karhunen-Loee Expansion [27].

2.5. Classification: Classification is the main decision making stage of character recognition system. It uses the features extracted in the previous stage to identify the text segment according to preset rules. Various classifiers are Artificial immune system, Kohonen Network, Associative Memory, Support Vector Machine [12, 15, 17, 19, 23], Nearest Neighbour, K-nearest neighbours [3, 16, 29], Bayesian classification, Projection Distance, Subspace method, Linear Discriminant Function, Modified quadratic discriminant function, Mirror Image Learning, Euclidean Distance, Modified Projection distance, Compound projection distance, Compound modified quadratic discriminant function, Regularized Discriminant Analysis [21].

2.6. Post processing: The output obtained from classification may contain some errors. Such errors can be removed or reduced by post processing methods.

Some group of characters like O and Q, M and N, V and Y, R and P have to be provided with special care due to their similarity. For example, the letters O and Q can be identified using signature features [18]. Another method for post-processing is the fusion method which is based on the Borda count [12].

III. LITERATURE SURVEY

Lot of work has been done in the field of handwritten character recognition considering many different languages using various techniques and algorithms.

Gaganjot Kaur et al. [13] have selected 7 features, namely number of boundaries, extent, filled area, major axis length, minor axis length, orientation and solidity, and have used Levenberg-Marquardt back propagation training algorithm to train the neural network. They achieved maximum accuracy of 93% for 130 samples by using 50 hidden layer neurons. Om Prakash Sharma et al. [6] considered Euler number computation to categorize the alphabets and further proposed an improved zone based hybrid feature extraction model that could give higher accuracy with reduced time for training and classification as compared to the diagonal based feature extraction model which gave 98.5% accuracy. In [30], weight matrix was formed by using the learning rule for neural network and recognition score was calculated. The system had high accuracy rate but mismatch occurred for similar type of characters. So, Euclidean
distance metric was further used to recognize such class of characters.

In [21], Rajbala Tokas et al. have experimented on various feature extraction methods and concluded that among those methods, cross-corner, diagonal and direction methods are the most accurate methods and that other methods can be combined to obtain higher accuracy rate. Perceptron learning method has been used by Rakesh Kumar Mandal et al. [29] for character recognition. Row-wise segmentation technique was used for feature extraction and an overall accuracy of 80% was obtained. Euclidean distance metric was used in [1] for character recognition and an average accuracy of 90% was achieved by using wavelet transform multiresolution technique for feature extraction.

In [4], histogram band information was extracted from the character string and multilayer perception neural network was used for recognition. Some patterns correctly classified with 100 % accuracy whereas remaining number of patterns were classified with absolute relative error. Muhammad Naem Ayyaz et al. [17] have used hybrid feature extraction (combination of correlation based features and some statistical/structural based features like end points and junction points, invariant moments, projection histogram, profiles) and multiclass SVM classification and obtained 96.5% accuracy for digits and 96% for alphabets. In [7], handwritten characters have been recognized through multiset of puzzle pieces; the pattern primitives are segmented, identified, and labeled for further recognition using fuzzy function and four basic features – horizontal stroke, vertical stroke, left slant stroke, right slant stroke have been used for labelling purpose and the classification is done in two ways: Exact matching and Probability matching. The authors have concluded that probability matching is better than exact matching.

J. Pradeep et al. [10] have used three different ways of feature extraction for character recognition (horizontal based, vertical based and diagonal based) and feed forward back propagation neural network is used for classification. The diagonal based feature extraction has been found to be the most accurate. A system to recognize handwritten characters has been developed by Anita Pal et al. [11] by using the boundary tracing technique for feature extraction and the experimental results show that Fourier descriptors with back propagation network yields a recognition accuracy of 94%. In [18], multiple level hidden markov model has been used for classification to recognize the characters by finding local and global features (gradient features, projection features, curvature features) and an average accuracy of 98.26% has been obtained. For a significant number of letters, the accuracy rate was close to 100%. Stuti Asthana et al. [14] have used combination of five different scripts and two different approaches are used for recognition in feed forward neural network – one with single hidden layer and other with two hidden layers. They achieved 96.53% average recognition rate using double hidden layer and 94% using single hidden layer.

In [12], Anshul Gupta et al. have discussed character recognition technique which uses four Fourier features and achieved 98.75% accuracy using SVM classifiers. Two neural classifiers – Back Propagation (BP) and Radial Basis Function (RBF) network had been used by M. Blumenstein et al. [31] and two feature extraction techniques namely transition features and direction features were performed. The experimental results showed that in most of the cases the direction features yielded more accuracy compared to the transition features. In [32], an efficient fuzzy method is used for handwritten character recognition which shows recognition rate of above 95% if the character is written by the same person who presented the characters to the system during training and recognition rate of around 70% for any other person.

Nguang Sing Ping et al. [33] have extracted data from the character by 13-point zone division and used two types of back propagation training algorithm - Batch Gradient Descent and Levenberg Marquardt and concluded that Batch Gradient Descent algorithm shows better recognition accuracy; although the algorithm is slow, it doesn't require a large memory for network training. In [34], character recognition is done using fuzzy membership function and the method has been applied on different unknown characters. It has been found that the pattern having highest membership grade is the recognized character in most of the cases; however few mismatch occurs if there is completely different style of handwritten character. Velappa Ganapathy et al. [35] have performed neural network training on different characters with and without using selective thresholding minimum distance technique and concluded that networks that use selective thresholding minimum distance technique generally produce higher percentage accuracies compared to the networks that do not use it.

In [36], the authors have proposed character recognition system which uses 280 dimension feature vector which uses histogram (radial) and profiles (in-out and out-in) and achieved varying accuracy of 72.8% to 98.8% depending on the database difficulty and character category. Alex Graves et al. [37] have introduced a novel approach for both online and offline handwritten character recognition using Recurrent Neural Network and compared the result using HMM based system. The experiments showed that the new approach outperformed with the HMM based system and was more robust to the changes in sample size.

S. Arora et al. [38] have extracted chain code histogram features and moment based features for Devanagari characters in their proposed multiple classifier system. The recognition accuracy obtained from chain code histogram based classifier was 88.19% and moment based classifier was 65.67%, whereas by combining both the classifiers using weighted majority scheme gave 98.03% accuracy. Twelve different classifiers and four sets of features (two sets from binary and two sets from
gray-scale images) that are based on gradient and curvature are used for Devnagari character recognition by the authors in [39] and the results show that Mirror Image Learning (MIL) classifier provides the best results among the 12 classifiers with an average accuracy of 94.94%. They also concluded that curvature features provided higher results compared to the gradient features for almost all the classifiers. In [26], Devnagari characters were first identified by regular expression matching, if they did not match with any pattern they were passed to minimum edit distance filter and overall accuracy achieved was 82%.

SVM and ANN classifiers have been used for Devnagari character recognition on Shadow based, Chain code Histogram, Longest Run, and View based features in [15] and a conclusion is made that reliable classification is possible using SVM classification. Gurumukhi Numerals have been recognized by Anita Rani et al. [23] by using various combinations of feature sets, and the highest accuracy of 99.4% is obtained for two different combinations of feature sets – 1. Diagonal features and background directional distribution 2. Projection histograms, Diagonal features and background directional distribution. P. Phokharatkul et al. [40] have used Ant-Miner Algorithm for the recognition of Thai characters and 97% of the training set had been recognized.

In [28], a comparative study of Gabor feature against Gradient feature for handwritten Chinese character recognition was conducted and it was observed that Gradient feature performed much superior to Gabor feature. Z. Shaaban [41] has described a new approach for the recognition of multi-font Arabic texts which depends on multiple parallel neural networks; and the classifier achieved a high recognition rate of 98%. Ismael Ahmad Jannoud [9] has proposed Arabic Handwritten Text Recognition system using Discrete Wavelet Transform method for feature extraction and the best recognition rate was for the isolated characters (99%), while the middle characters had the worst recognition rate (91%). In [42], Mahesh Goyani et al have discussed top down hierarchical histogram based approach for printed devnagari script character isolation. Extracted characters may be used as binary feature or some advanced features can be computed to train classifier. In [43], Mahesh Goyani et al have discussed euler number, symmetry and histogram based approach which is classifying set of characters at each level. This approach is effective for fixed font family. In [44], authors have proposed correlation coefficient based HCR system. Though it gives good result, it suffers from extensive computation of correlation. Chain code based approach is proposed in literature [45], in which the accuracy of chain code feature is tested against feed forward back propagation neural network and support vector machine. Results show that SVM outperforms ANN. Performance of chain code features is compared with binary features and holistic features like PCA and FDA in literature [46]. L2 norm, neural network and SVM are as classifier.

Discussion shows, FDA outperforms all other features with SVM.

IV. CONCLUSION

Several conclusions have been drawn regarding the feature extraction techniques and the classification techniques from the literature survey that we have made. We have come across many feature extraction techniques and found that hybrid feature extraction technique (combining correlation based features and some statistical/structural features) produces results with more accuracy than a single technique. Also, if the characters are written in box sheets, not deviating from the actual position, then Row-wise segmentation feature extraction technique is robust yielding better accuracy with fast convergence and comparatively less number of epochs for classification using neural network.

Varieties of classifiers have also been explored in the survey and among the many classifiers, Mirror Image Learning classifier has been found to give better results. Generally, Artificial Neural Network or Multi-class SVM classification is used for classification. While using neural network for classification, if the number of hidden layers is increased, the recognition rate accuracy increases but there will be an increase in number of epochs and the processing speed will be slow. However, SVM classification is more reliable because it gives high efficiency with respect to speed, memory and classification accuracy.

REFERENCES


