A Survey on Handwritten Character Recognition

Manish Kumar Sahu[1], Dr. Naveen Kumar Dewangan[2]

M.Tech Scholar, BIT, Durg[1]

Associate Professor, Department of Electronics & Telecommunication, BIT, Durg[2]

Abstract: Handwritten character recognition has been one of the active and challenging areas of research in the field of image processing and pattern recognition. It has number of applications which include, reading aid for blind, and conversion of any hand written document into structural text form. In this paper we present a brief survey of various methods which recognizes of English alphabet in a given scanned text document. The first step is image acquisition which acquires the scanned image followed by noise filtering, smoothing and normalization of scanned image, rendering image suitable for segmentation where image is decomposed into sub images. Feature Extraction improves recognition rate and misclassification. We have to surveyed and compared various methods in this paper.

Keywords: Handwritten Character Recognition, Pattern Recognition, Image Processing, Feature Extraction.

I. INTRODUCTION

Handwriting recognition has been one of the most interesting and challenging research areas in field of image processing and pattern recognition in the recent years. It contributes immensely to the advancement of automation process and improves the interface between man and machine in numerous applications. In general, handwriting recognition is classified into two types as off-line and on-line handwriting recognition methods. The on-line methods have been shown to be superior to their off-line counter parts in recognizing handwritten characters due to the temporal information available with the former. However, in the off-line systems shown in figure (1) we get comparably high recognition accuracy levels.
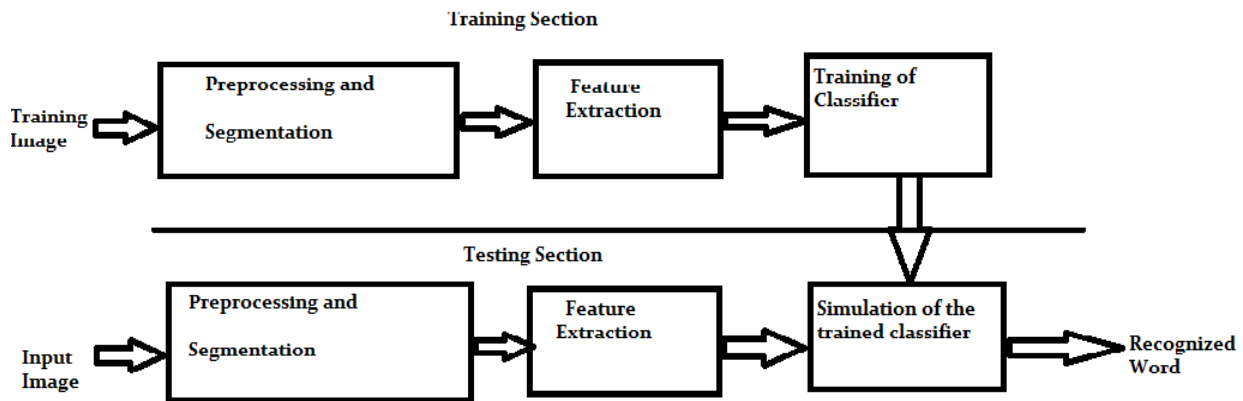

Figure (1):- Generic character recognition model

The recognition of characters from scanned images of documents has been a problem that has received much attention in the fields of image processing, pattern recognition and artificial intelligence. Classical methods in pattern recognition do not as such suffice for the recognition of visual characters due to the following reasons:

(1.) The 'same' characters differ in sizes, shapes and styles from person to person and even from time to time with the same person.
(2.) Like any image, visual characters are subject to decaying due to noise.
(3.) There are no hard-and-fast rules that define the appearance of a visual character. Hence rules need to be heuristically deduced from samples. As such, the human system of vision is excellent in the sense of the following qualities:

(a.) The human brain is adaptive to minor changes and errors in visual patterns. Thus we are able to read the handwritings of many people of different styles of writing.
(b.) The human vision system learns from experience, Hence we are able to hold newer styles and scripts with amazingly high speed.
(c.) The human vision system is immune to variations of size, aspect ratio, colour, location and orientation of visual characters.

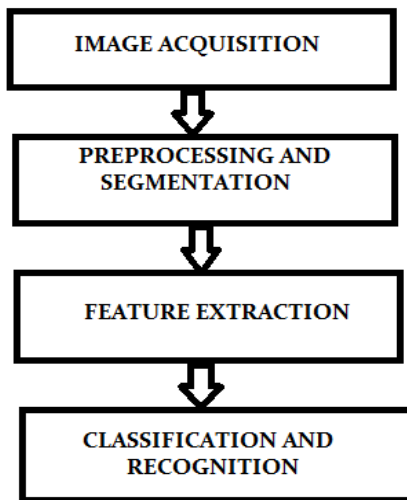## II. STEPS IN GENERAL OFFLINE CHARACTER RECOGNITION SYSTEM

Figure (2):- General offline character recognition system

(1.) Image acquisition:-

In Image acquisition, the recognition system acquires a scanned image as an input image. The image should have a specific format such as JPEG, BMT, etc. This image is acquired through a scanner, digital camera or any other digital input device.

(2.) Pre-processing:-

The pre-processing is essentially enhances the image rendering it suitable for segmentation. The role of pre-processing is to segment the interesting pattern from the background. Generally, noise filtering, smoothing and normalization should be done in this step. The pre-processing also defines a compact representation of the pattern. Binarization process converts a gray scale image into a binary image.

(3.) Segmentation:-

In the segmentation stage, an image of sequence of characters is decomposed into sub-images of individual character. In this system, the pre-processed input image is segmented into isolated characters by assigning a number to each character using a labeling process. This labelling provides information about number of characters in the image. Each individual character is uniformly resized into pixels.

(4.) Feature Extraction:-

In this stage, the features of the characters that are crucial for classifying them at recognition stage are extracted. This is an important stage as it improves the recognition rate and reduces the misclassification. Diagonal feature extraction scheme for recognizing off-line handwritten characters is proposed in this process. Every character image is divided into equal zones, each of size 10x10 pixels. The features are extracted from each zone pixels by moving along the diagonals of its respective 10x10 pixels.

(5.) Classification and Recognition:-

The classification stage is the decision making part of the recognition system. A feed forward back propagation neural network is used in this work for classifying and recognizing the handwritten characters. The pixels derived from the resized character in the segmentation stage form the input to the classifier. The neural consists of two hidden layers besides an input layer and an output layer.

III. THE DIFFERENT METHODS OF CHARACTER RECOGNITION

Prasad et al., (2013) Proposed ANN based system perform character recognition due to ANNs high noise tolerance. The feature extraction step of optical character recognition is the most important. A poorly chosen set of features will yield poor classification rates by any neural network.

At the current stage of development, the software does perform well either in terms of speed and accuracy but not better. It was likely to replace existing OCR methods, especially for English text.

Sharma et al., (2013) Proposed a useful method for the recognition of handwritten character to a great extent. The proposed method applied on different unknown characters. This neural network based method gives the accuracy 85%. Developed for proposed algorithms cannot be applied to recognize cursive handwriting recognition.

Dass et al., (2013) Stated that offline character recognition is a difficult problem, not only because of the great quantity of variations in human handwriting style, but also, because of overlapping of the neighbor characters. Recognition strategy heavily depend on the nature of data to be recognized. There are different network in neural network for classification purpose and pattern network was used for classification and observe almost 96 recognition rate.

Kaur et al., (2013) Proposed system which recognizes English numeral using gradient descent based back propagation training algorithm. The main aim of using this algorithm was to reduce the error which is difference between the computed value of neural network and desired value. In this system by using artificial neural network methodology of English numeral recognition, 99% accuracy was achieved.

Verma et al., (2014) Proposed a system in which neural network and Surf Feature has demonstrated its capability for solving complex character recognition problems. Noise reduction plays an important role in character recognition.

The proposed research work used neural network algorithm and surf feature extraction with their implementation details for solving character recognition problems. This algorithms has been performed based on Noise in input image provide promising results in terms of PSNR and MSE.

Verma et al., (2014) Proposed SURF feature and developed neural network has the capability of strong robustness performance and good distinction between feature points. It has greatly improved computing speed. The average success rates of recognition of all characters are 98.7753% by using neural network and SURF feature. These algorithm has been performed based on noise in input image and provides promising results in terms of PSNR and MSE.

Aggrawal et al., (2014) Proposed the reason for choosing artificial neural networks as a part of the research to perform character recognition due to their high noise tolerance. The designed systems have the ability to produce accurate results, provided the correct dataset is available at the time of training the network. The current stage of research depicts that the software does perform well both in terms of speed and accuracy. But the character location is not efficient since the size of every block varies. This can be taken care of by initializing the weights during training of dataset.

Prasad et al., (2015) Proposed that the feature extraction step of character recognition is the very important. A poorly chosen set of features will yield poor classification rates by any artificial neural network. At current stage of development, the software does perform well either in terms of speed or accuracy but not better. A simple approach for recognition of characters using artificial neural networks has been described.

IV. CONCLUSION

Different method of character recognition was reported that the Verma et al., (2014) gives the accuracy of approximately 99% by using neural network and SURF feature, exhibit outcomes of noise reduction and image quality improvements, with different noise levels, which qualify it to be suitable for image processing and Pattern matching.

REFERENCES

[1] Kauleshwar Prasad, Devvrat C. Nigam, Ashmika Lakhotiya: Character Recognition Using Matlab's Neural Network Toolbox, International Journal of u- and e- Service, Science and Technology Vol.6, No. 1, February, 2013, 13-20.

[2] Ankit Sharma, Dipti R Chaudhary: Character Recognition Using Neural Network, International Journal of Engineering Trends and Technology (IJETT) – Volume 4 Issue 4 – April 2013, 662-667.

[3] Vipin, Rajeshwar Dass, Rajni: Character Recognition Using Neural Network, International Journal of Advanced Trends in Computer Science and Engineering Volume 2, No.3, May-June 2013, 62-67.

[4] Amritpal kaur, Madhavi Arora: Neural network based Numerical digits Recognition using NNT in March, 2015/90-92.