Handwritten Digits Recognition Special point

P. P. Chaudhari  
Computer Engg., Govt. Polytechnic  
India

K.R. Sarode  
Dept. of Computer Engg Govt. Engineering  
India

Abstract: One of the most challenging topics is the recognition of handwritten numeral characters. In this paper is concerned with recognition of an offline handwritten numeral character using artificial neural network. In this technique we extracted the multiple features. This multiple features such as end point, junction point, straight lines are extracted from image. Artificial neural network is used as classifier of the character in classification stage. The recognition system of experimental results shows that this technique is effective and reliable. The overall procedure results in recognition rate are 94.1%.

Keywords: Handwritten numeral character recognition, multiple feature, feature extraction, neural network, classification.

1. INTRODUCTION

The recognition of optical characters is known to be one of the earliest applications of Artificial Neural Networks, which partially emulate human thinking in the domain of artificial intelligence. The numeral character recognition is most challenging and tantalizing field, because the big research and development effort that has gone into it, has not solved all commercially urgent and intellectually interesting problems. Handwritten numeral character recognition not only is very challenging problem, but also because it provides a solution for processing large volume of data automatically and large variety of scientific applications. Digital document processing is gaining popularity for application to office and library automation, bank and postal services, publishing houses and communication technology. The complexity of the problem is greatly increased by noise in data and infinite variability of handwriting as a result of mood of writer and nature of writing. To recognize the handwritten characters is very difficult task. Character recognition has been the subject of intensive research for more than three decades [2]. Many different methods have been explored by large number of scientists to recognized characters. Many statistical and syntactic approaches have been developed for handwritten character recognition by researchers [3,4]. In character recognition a review by V. K. Govindan and Shivprasad explained two main to approaches pattern recognition. They are statistical / decision / theoretic / syntactic / linguistic / grammatical and structural approaches. Each has certain merits and demerits [11]. In syntactic approaches for handwritten character recognition still have problems [7]. The division of two approaches is sometime not clear cut particularly in terms of practical applications.

Artificial neural network has been successfully applied to handwritten digit recognition numerous times, with vary small error margins [5, 9]. This paper proposes a new technique for the recognition of hand-printed numeral characters using neural networks. There are four main advantages of this technique. First, the proposed technique combines rule-based (structural) and classification tests. Second, it is more efficient for large and complex sets, such as numeral characters. Third, feature extraction is inexpensive. Fourth, the execution time is independent of both the character font and size..This paper describes the neural network method applied in the classification stage. Thus, this system takes a hybrid approach to character recognition. The paper is organized as follows: In section II, the basic preprocessing digitization and thinning. In section III, contains feature extraction steps are described. Section IV describes classification technique using neural network. In section V, we show results of character recognition and draw conclusion in section VI. The block diagram handwritten numeral character recognition is shown in figure 1.
II. CHARACTER ACQUISITION

The first step in process is to acquire handwritten numeral characters. We used scanner to do this job. Other sensors can be camera, video camera, paintbrush etc. The handwritten numeral samples were specifically collected from different peoples for this study. Each participant was asked to enter a random sequence of digits. All digits used in this study are disconnected. The first step in this process is to acquire handwritten numeral characters by scanner at 300 dpi. that yielded a binary image which subsequently stored in compressed format in memory.

III. PREPROCESSING

The goal of preprocessing is to increase the quality of recognition, that means more precisely that character is transformed to such that they are more similar to the class they belongs. In this stage no any recognition process is performing. There is important interaction between preprocessing of character images and feature extraction process. In preprocessing the preliminary steps include normalization and digitization.

A. Normalization

Normalization is considered as the most important pre-processing factor for character recognition. Normally, the character image is linearly mapped onto a standard plane by interpolation / extrapolation. The size and position of character is controlled such that the x / y dimensions of normalized plane are filled. The implementation of interpolation/ extrapolations is influential to recognition performance [6]. The normalization procedures simplify the classification and enhance the performance.

B. Digitization

In digitization object is converted into binary form by binarization method. Object is separated from background (1-represent as a region, 0 represent as a no region). This binarized image is put though preprocessing routines that smooth the image and eliminate noise, artificial holes and other artifacts produced by the digitizing process.

C. Thinning

The thinning of elongated objects is a fundamental preprocessing operation in image analysis. Thinning is a morphological operation that is used to remove selected foreground pixels from binary images, somewhat like erosion or opening. It can be used for several applications, but is particularly useful for skeletonization[12]. In this mode it is commonly used to tidy up the output of edge detectors by reducing all lines to single pixel thickness. Thinning is normally only applied to binary images, and produces another binary image as output. It gives thinned image. This thinned image is used for the locating special points.

IV. FEATURE EXTRACTION

Features extraction is the crucial stage in numeral identification as each numeral is unique in its own way, thus distinguishing itself from other numerals. Hence, it is very important to extract features in such a way that the recognition of different numerals becomes easier on the basis of the individual features of each numeral. Feature extracted special points of digitized image.

A. Special points

Special points such as endpoints and junction points are detected while scanning through and storing thinned image in matrix form. They are defined as points vertically, horizontally, or diagonally connected to respectively one or two pixels of the thinned image.

The End point is defined as only one of the eight neighbors as a “1s” and which “E” denotes. The junction point is defined as pixel which has more than two “1s” among the eight neighbors. But this is getting many undesired junction points, so we applying condition more than two “1s” as its two neighbors, the number 0 to 1 and 1 to 0 transitions in the eight neighborhoods of pixel should be greater than or equal to six. “J” denotes the junction point. Both features such as end points and junction point are...
used in classification stage. Starting from a J points, one of its neighbors is selected. The pattern of numeral character three image denoted E point and J point of numeral character.

V. CLASSIFICATION

Classification is the important stage for numeral character recognition. The extracted features special points are used to identify numeral. Having extracted the features, it is required to store them in some form. Each pattern should uniquely identify a character and each character may represent by several distinct pattern. A multilayer feed forward network is used for classification stage. Multiple layer networks has layers of processing element and each element makes independent computations on data that it receives and passes the result to another layer The next layer will in turn make independent computation and pass on results to yet another layer. The neural network trained with to identify each character. After this, if the pattern of unknown character were presented neural network, it would be able to classify the character based on the identical pattern to which it has been trained. Multilayer feed forward network can be successfully applied to a variety of classification and recognition problems. For 100 handwritten numerals character 0 through 9. Fifty vectors have been used to train the network, fifty remaining vectors were needed for test purposes. The architecture was trained with 10 inputs and 10 output neurons. The best result is obtained by multilayer feed forward network.

VII. RESULT

The techniques adopted in this paper for recognizing hand-printed numeral characters using Neural Networks combines with location of special points and a classification test. This approach is efficient for feature extraction and recognition. As indicated by the experiments performed, the algorithm results in a 94.1% recognition rate.

The various types of large variability of handwritten numeral character, test samples are taken. There recognition rate for each digit is shown in given in table 1. Performance of result is satisfactory

Table 1. Recognition rates of handwritten numeral characters

<table>
<thead>
<tr>
<th>NUMERAL</th>
<th>RECOGNITION RATE (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>96</td>
</tr>
<tr>
<td>1</td>
<td>93</td>
</tr>
<tr>
<td>2</td>
<td>92</td>
</tr>
<tr>
<td>3</td>
<td>94</td>
</tr>
<tr>
<td>4</td>
<td>93</td>
</tr>
<tr>
<td>5</td>
<td>95</td>
</tr>
<tr>
<td>6</td>
<td>92</td>
</tr>
<tr>
<td>7</td>
<td>96</td>
</tr>
<tr>
<td>8</td>
<td>94</td>
</tr>
<tr>
<td>9</td>
<td>96</td>
</tr>
</tbody>
</table>

VII. CONCLUSIONS

In this paper, hybrid approach is used for recognition of handwritten numeral characters. Multiple features extracted from image. These features were analyzed using neural network. The best recognition rate achieved by our proposed system was 94.1%. This result is obtained by extracted the multiple features and neural network is used as classifier.

REFERENCES


