ISSN No. 0976-5697

Volume 3, No. 5, Sept-Oct 2012

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Handwritten Numerals Recognition using Hough Transformation Technique

Deepika Ahuja
Lecturer, Department of Computer Science
Engineering Bhai Maha Singh College
of Engineering & Technology,
Muktsar, Punjab
deepika.ahuja551@gmail.com

Munish Kumar*
Assistant Professor, Computer Science
Department Punjab University
Constituent College,
Muktsar, Punjab
munishcse@gmail.com

Abstract: Handwriting recognition has always been a exigent task in image processing and pattern recognition. In this paper, we have presented Hough transformation technique for extracting the features of handwritten numerals. SVM based approach has been used to classify the numerals based on the extracted features. In this study, we have taken the samples of handwritten numerals from 120 different writers. Each writer has written each numeral five times. We have collected total 6000 samples of handwritten numerals for training and testing. We have taken 75% data as training data and 25% data as testing data. The proposed system achieves a maximum recognition accuracy of about 91.75%.

Keywords - SVM classifier, Hough transform, OCR.

I. INTRODUCTION

Handwritten numeral recognition has received wide attention in academic and production fields. India is a multi-lingual and multi-script country, where eighteen official scripts are accepted and have over hundred regional languages [1]. Handwritten numeral recognition has a variety of applications like telephone numbers, number plates, pin codes, passports, identity numbers *etc*. In United State, the offline handwritten numeral recognition system is widely used. This system was proposed for the development of zip code or postal code recognition that could be engaged in mail sorting, which reduces human efforts to sort mails with postal codes or zip codes that were difficult to identify. For recognition of numerals of many different scripts, there are various feature extraction methods identified in the literature review. These include chain code, projection histograms, Hough transform, Fourier descriptor, Contour profile, Laplace transform. Selection of good feature extraction technique is the most important factor for attain high recognition accuracy in handwritten numerals.

There are various feature extraction methods have been proposed for handwritten numerals. Now, we briefly review them below: Bagde *et al.* [2] have proposed a method that deal with MLP (Multilayer perception) neural network that is often used in handwritten numerals recognition. Majumdar and Chaudhuri [3] presents structural features like the position of holes, intersection points, terminal points *etc.* were used to identify a numeral. They have used MLP classifier for recognition purpose. Rajput and Mali [4] have presented chain code and Fourier descriptors that extract the information as feature of the numeral and they obtain the accuracy of about 98.12%. Romeo *et al.* [5] have reported a method based on directional two dimensional continuous wavelet transform. The transform has four parameters by using this they acquire the wavelet descriptor that is in the form feature vector for each numeral image. Pradeep *et al.* [6] have proposed a novel technique, namely, diagonal based feature extraction for numeral recognition. For further classification multilayer feed forward neural network is used. Roy *et al.* [7] have used directional features extracted from the contour of each numeral. Impedovo *et al.* [8] focus on membership functions for zoning based classification. Jayadevan *et al.* [9] have proposed a scheme named lexicon driven segmentation-recognition used for recognition the legal amount of words of the banks. Alaei *et al.* [10] have proposed a method based on the modified contour chain code for the feature extraction and SVM classifier for the classification. They have achieved recognition accuracy of about 93.37%. Asthana *et al.* [11] have used the multilayer feed forward back-propagation algorithm using two hidden layer for recognition of multiscript numerals. Benne *et al.* [12] have used structural features such as water reservoirs maximum profile system distances, and fill holes density. Yamaguchi *et al.* [13] have implemented a technique based on Hough transform technique for digit recognition. In this paper, we have used SVM classifier for recognition.

II. DATA SET

For this work, we have collected 6000 sample of handwritten numerals. Data is collected from different schools and colleges. We have taken the samples of handwritten numerals from 120 writers and each writer has writes each numeral 5 times for my research.

III. THE PROPOSED RECOGNITION SYSTEM

The block diagram of the proposed system is shown in figure 1. For extracting the features, we came across the various stages during numeral recognition such as: Digitization, Pre-processing, Feature extraction and classification.

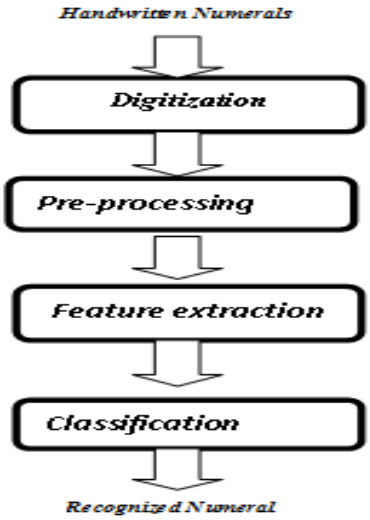

Figure 1: Block diagram of Handwritten Numeral Recognition System.

A. Digitization:

The conversion of handwritten numeral into digital form with the help of optical devices such as scanner which produces digital image of handwritten numeral image is known as digitization.

B. Pre-processing:

In this phase the numeral image is normalized into size of 100 × 100. After normalization, we perform the thinning step and produce bitmap image in the form of (0, 1).

C. Feature extraction:

In this phase, the features of input numerals are extracted. The selection of appropriate feature extraction method is probably the single most important factor in achieving high recognition performance. So, the extracted features should be able to classify each numeral uniquely. We have used Hough transformation technique for feature extraction of handwritten numerals. Singh and tyagi [14] have presented radial based function handwritten numeral recognition of Devanagari script. Birajdar and Subhedar [15] have proposed JPEG algorithm for recognition of handwritten devanagri numerals. Smereka and Duleba [16] have proposed a modified Hough transform technique that improves the recognition of low contrast circular objects.

IV. CLASSIFICATION

Classification phase is the decision making phase of the handwritten numeral recognition system. This phase uses the features, which have been extracted in previous stage for making the class membership of numerals. In this work, we have used SVM classifier for recognition. SVM is a very useful technique for data classification in the field of pattern recognition. SVM is a supervised learning, in which a machine is trained with training data instead of programmed.

V. HOUGH TRANSFORMATION TECHNIQUE

The Hough transformation is a feature extraction technique which is used for image analysis and digital image processing. The objective of this technique is to find inadequate instances of object with the help of voting procedure, which is carried out in parameter space. In image space the straight line can be described as $y = mx + b$ can be graphically plotted for each pair of image points (x, y). In Hough transform characteristics of straight line not as image points $(x1, y1), (x2, y2)$ but instead in terms of its parameters *i.e* slope parameter *m* and intercept parameter *b*. The straight line $y = mx + b$ can be represented as a point (b, m) in parameter space. We have used a different pair of parameters θ (*theta*) and r for lines in Hough transform, these are polar coordinates.

The parameter *r* represents the distance between the line & origin while θ is the angle from origin to the closest points. The equation of line can be written as

$$r = xcos(\theta) + ysin(\theta)$$

VI. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the results of recognition system for handwritten numerals based on Hough transformations are presented. We have used 6000 samples of handwritten numerals for training and testing. Out of 6000 samples, we have taken 4500 samples for training set and 1500 samples for testing data set. Using this approach, we have achieved an accuracy of about 91.75% when we use SVM with polynomial kernel classifier as shown in Table 1. This accuracy can probably be increased by considering a larger data set while training the classifier. This work can also be extended for handwritten character recognition of Indian scripts.

VII. CONCLUSION

The work presented in this paper proposes a handwritten numeral recognition system based on Hough transformation technique. The classifier that has been employed in this work is SVM. The maximum recognition accuracy of 91.75% is achieved in this work for the case when we input the Hough transformation technique to SVM classifier with polynomial kernel. This accuracy can probably be increased by considering a larger data set while training the classifier. The results of recognition accuracy with this features and SVM with three kernels are depicted in Table 1.

Table 1: Handwritten numeral recognition accuracy.

Classifier	Accuracy (%)
Linear Kernel	89.45 %
Polynomial Kernel	91.75 %
RBF Kernel	86.70 %

VIII.REFERENCES

[1] S. V. Rajashekaradhya, P. V. Ranjan, "Numeral/mixed numeral recognition of South-Indian scripts: the zone based feature extraction method," Journal of Theoretical and Applied information Technology, Vol. 7, pp. 63-79, 2009.

[2] P. R. Bagde, S. A. Ladhake, K Chinchkhede, "Handwritten Marathi numeral recognition: classification system with a neural network approach," International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2, pp. 258-265, March 2012.

[3] Majumdar and B .B. Chaudhuri, "A MLP classifier for both printed and handwritten bangle numeral recognition," Proceedings of the 5th Indian conference on Computer Vision, Graphics and Image Processing, 2006, p. 796 – 804.

[4] G. G. Rajput, S. Mali, "Marathi handwritten numeral recognition using Fourier descriptors and normalized chain code," Recent Trends in Image Processing and Pattern Recognition RTIPP, 2010, p. 141- 145.

[5] D. Romero, M. S Leticia, A. M. Ruedin, "Directional continuous wavelet transform applied to handwritten numerals recognition using neural networks," Journal of Computer Science & Technology, Vol. 7, pp. 66 -71, 1 April 2007.

[6] J. Pradeep, E. Srinivasan and S. Himavathi, "Diagonal based feature extraction for handwritten alphabets recognition system using neural network," International Journal of Computer Science & Information Technology (IJCSIT), Vol 3, pp. 27- 38, Feb 2011.

[7] K. Roy, C. Chaudhuri, U. Pal and M. Kundu, "A Study on the Effect of Varying Training set Sizes on Recognition Performance with Handwritten Bangla Numerals," IEEE Indicon Conference, 2005, p. 570-574.

[8] S. Impedovo, R. Modugno, G. Pirlo, "Membership functions for zoning-based recognition of handwritten digits," International Conference on Pattern Recognition, 2010, p. 1876- 1879.

[9] R. Jayadevan, U. Pal, F. Kimura, "Recognition of words from legal amounts of Indian bank cheques," 12th International Conference on Frontiers in Handwriting Recognition, 2010, p. 166 -171.

[10] Alaei, U. Pal and P. Nagabhushan, "Using modified contour features and SVM based classifier for the recognition of Persian/Arabic handwritten numerals," International Conference on Advances in Pattern Recognition 2009, p. 391- 394.

[11] S. Asthana, F. Haneef, R. K. Bhujade. "Handwritten multiscript numeral recognition using artificial neural networks," International Journal of Soft Computing and Engineering, Vol. 1, pp. 1-5, 2011.

[12] R. G Benne, B.V. Dhandra and M. Hangarge, "Tri-scripts handwritten numeral recognition: a novel approach," Advances in Computational Research, Vol. 1, pp. 47-51, 2009.

[13] T. Yamaguchi, Y. Nakano, M. Maruyama, H. Miyao and T. Hananoi, "digit classification on signboards for telephone number recognition," proceedings of Seventh International Conference on Document Analysis and Recognition, Vol. 1, pp. 359-363, 2003.

[14] P. Singh, N. Tyagi, "Radial basis function for handwritten devanagari numeral recognition," International Journal of Advanced Computer Science and Applications, Vol. 2, pp. 126-129, 2011.

[15] G. Birajdar and M. Subhedar, "Use of jpeg algorithm in handwritten devnagri numeral recognition," International Journal of Distributed and Parallel Systems, Vol. 2, pp.152-160, July 2011.

[16] M. Smereka, I. Duleba, "Circular object detection using a modified Hough transform," International Journal of Applied Math Computer Science, Vol. 18, pp. 85–91, 2008.

