Recognition of Handwritten Devanagari Numerals

[1]**Navneet Kaur,** [2]**Monica Goyal**

[1,2]Dept. of CSE, Guru Kashi University, Talwandi Sabo, India

Abstract

Recognition of Devanagari numerals is a difficult task. Extensive research has been done on character recognition in the last few decades. In optical character a numeral, character or symbol to be recognized can be machined printed or handwritten character or numeral. There are several approaches that deal with the study of handwritten Devanagari numerals depending on the type of feature extracted and different way of extracting them. In this paper, a new is used for recognition of handwritten Devanagari characters. We explain the segmentation, feature extraction, water reservoir method, and neural network techniques. Feature extraction provides 92% accuracy of numeral recognition; water reservoir provides 94.34% accuracy. We also discuss various problems exist in numeral recognition.

Keywords

Segmentation, Handwritten Character, Numeral Recognition, Water Reservoir

I. Introduction

Recognition of handwritten numerals is important because of its applicability to a number of problems, like postal code recognition and information extraction from fields of different forms. Handwritten character Recognition was begun to come into existence in research during 1980s. Many documents are computer aided in the present era. Handwritten recognition has its own significant and it is adoptable in many variousfields such as onlinehandwritten recognition on computer tablets, recognize zip code on mail for postal mail sorting, processing bank check amounts, numeric entries in forms filled up by hand. There are different challenges faced while attempting to solve this problem. The handwritten digits are not always of same size, thickness, or orientation and position relative to the margins. Numerous amount of work has been done on English and other sub continental languages and even on Indian script, but the result didn't come out of the constraint of laboratories.

Recognition of Handwritten Devanagari Numerals or Characters is complicated task due to the unconstrained shape variations, different writing style and different kind of noise. Also, handwriting depends on the writer and because we do not always write the same digit in exactly the same way, building a general recognition system that would recognize any digit with good reliability in every application is not possible.

A. Typesof Numeral Recognition

According to the mode of data acquisition, numeral recognition methodologies are categorized into two systems as:

• Online Numeral Recognition System
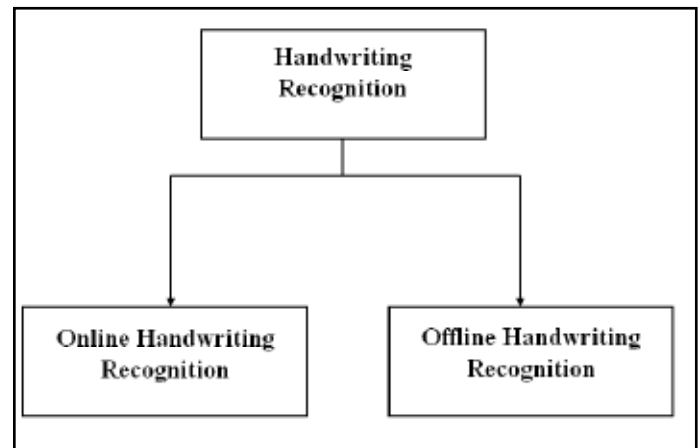• Offline Numeral Recognition System

Fig. 1: Areas of Handwritten Recognition

1. Online Numeral Recognition System

Online numeral recognition is the process of recognizing handwriting, recorded with a digitizer, as a time sequence of pen coordinates. It captures the temporal and dynamic information of the pen trajectory.

2. Offline Numeral Recognition System

Offline numeral recognition is the process of converting the image of writing into bit pattern by an optically digitizing device such as optical scanner or camera. The recognition is done on this bit pattern data for machine-printed or handwritten text.

II. Devanagari Script

Devanagari also called Nagari, is an abugida alphabet of Indian and Nepal. It is written from left to right, does not have distinct letter cases, and is recognizable by a horizontal line that runs along the top of full letters.

Fig. 2: Sample of Devanagari Numerals

Devanagari is the main script used to write standard Hindi, Marathi, and Nepali. Since the 19th century, it has been the most commonly used script for Sanskrit. Devanagari is also employed

for Bhojpuri, Gujarati, Magahi, Maithili, Bhili, Newari, Santhali, Tharu, and sometimes Sindhi, Dogri, Sherpa and by Kashmiri-speaking Hindus. It was formerly used to write Gujarati.

III. Preprocessing

The goal of pre-processing is to increase the quality of hand printed data. That means more precisely that numeral is transformed to such that they are more similar to mean the class they belong to. In pre-processing the preliminary steps include normalization, digitization and thinning. Pre-processing aims to produce accurately. After pre-processing phase, a cleaned image is available that goes to the segmentation phase.

IV. Digitization And Thinning

In digitization, object is converted into binary form by binarization method. Object is separated from background. This binarized image is put through pre-processing routines that smooth the image and eliminate noise, artificial holes and other artefacts produced by the digitizing process. Edge detection technique is used to convert original image into skeleton image. Skeleton of image means reducing the width of the line of the character to one pixel only. This helps in better extraction of feature from the text images.
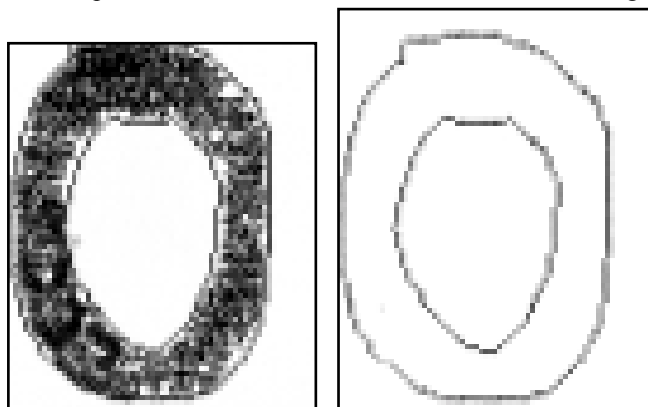

Fig. 3: The Original Image and Skeleton Image

V. Feature Extraction

The selection of good feature set is the most important aspect of handwritten character recognition. This method provides the ease of implementation and good recognition. Feature extraction extracts the well-defined numerals, which classify the numeral in classification stage. The features are extracted by projection methods from original and skeleton image. We also extracted special points from thinned image.

VI. Conclusion

This paper is concerned with the problem of recognition of unconstrained, isolated, handwritten Devanagari numerals and various techniques used to solve these problems. We concluded that it is possible to enhance recognition rate if a numeral is divided in systematic order. Most of the errors came from those components where touching numerals have some common portion and different writing style of each person. An accuracy of numeral can be increased by using the combination of more than one classifier.

References

[1] U. Bhattacharya, B. B. Chaudhuri, R. Ghosh, M. Ghosh, "On Recognition of Handwritten Devanagari Numerals", In proc. of the Workshop on learning Algorithms for pattern Recognition, Sydney, pp. 1-7, 2005.

[2] C. Vasantha, Ritu Jain, C. Patvardhan,"Optical Recognition of Handwritten Devanagari Numerals", Proceedings of NSC-2007,System Society of India, Manipal, pp. 14-15,dec 2007.

[3] Gaurav Jain, Jason Ko,"Handwritten Digits Recognition", Multimedia Systems, Project Report, University of Toronto, pp. 1-3,November 21, 2009, Vol. 8.

[4]. Migual Po-Hsien Wu,"Handwritten Character Recognition", A thesis report, University of Quinsland, pp-26, October 29, 2010, Vol. 6.

[5] V. K. Govindan, A. P. Shivprased,"Character Recognition a Review", Pattern Recognition,Vol. 23, pp. 671-683, 2007.

[6] Kartar Singh Siddharth Renu Dhir Rajneeh Rani "Handwritten Gurumukhi Numeral Recognition using different Feature Sets", International Journal of Computer Application, pp. 85, August 2011, Vol. 28.

[7] Bikash Shaw et al. [22],"Offline Handwritten Devanagari Word Recognition: a segmentation based approach", 2008 IEEE.

Navneet Kaur received his B-Tech degree in computer science Engg. from College of Engg. & Management, Rampura Phul, under Punjabi University Patiala, Punjab in 2011, and Pursuing the M-Tech. degree in Computer Science Engineering at Guru Kashi University, TalwandiSabo, Punjab.

Monica Goyal received his B-Tech degree in computer science & Engineering from Guru TegBhadhurKhalsa Institute of Engineering & Technology College, of Chappianwali, Malout, under PTU, Punjab in 2007, and M-Tech. degree in Computer Science Engineering fromYadwindra college of Engineering andTechnology from Punjabi University, Patiala, Punjab. Her research area is in digital image processing. At present, she is engaged in Guru Kashi University, Talwandi Sabo, Punjab as an Assistant Professor in Computer Science Engineering & Information Technology department.