

## A Hybrid Approach for Recognition of Hand Written Devnagri Compound Characters

Prashant Yawalkar<sup>1</sup> and M. U. Kharat<sup>2</sup>

<sup>1</sup>Research Scholar, <sup>2</sup>Professor and Head

<sup>1&2</sup>Department of Computer Engineering, MET's Institute of Engineering, Bhujbal Knowledge City, Maharashtra, India  
E-Mail: prashant25yawalkar@gmail.com, mukharat@rediffmail.com

**Abstract** - Being an effective tool in the world of communication, numerous techniques have been developed for documenting the handwritten text. Few of the exceptional techniques describe the environment of handwritten scripts and further convert it into electronic data by implementing various algorithms. Devnagri is one of the widely used scripts for most popular and commonly used languages like Marathi and Hindi. Recent development in the field of handwritten character recognition based on different methodologies like neural network, fuzzy logic, and deep neural networks has shown remarkable improvement in character recognition accuracy from 75% to 96%. We propose a fuzzy-Neural hybrid approach for recognition of hand written Devnagri compound character that uses a rotation invariant rule-based thinning algorithm as one of the major pre-processing activity. Thinning the characters to their central line, preserving the shape of the character are the distinctive features of thinning algorithm. Concurrent application of different rules to each pixel of the character image results into symmetrical thinning as well as improves the overall speed of the system. The system is trained using Neural Network where the weights are optimized using fuzzy rules improving the accuracy of the system. Results obtained for the fuzzy-neural based system with thinning helps in preserving the topology of the characters written in Devnagri and prove that accuracy of the system has stabilized in the band of 92-97% which was fluctuating in the band of 89-94% for the previously implemented systems. The system also shows a substantial improvement in accuracy for recognition of compound characters in comparison with our previously implemented system.

**Keywords:** Character Recognition, Neural Network, Rotation Invariant, Thinning, Fuzzy Logic, Fuzzy-Neural, Image Acquisition, Segmentation, Feature Extraction

### I. INTRODUCTION

Recognition of handwritten characters is a difficult task for machines while the same task is carried out accurately when given to the humans. Devnagri is one of the widely used scripts for most popular and commonly used languages like Marathi and Hindi. Optical Character recognition (OCR) for Devnagri script becomes specifically complex due to complicated curves and various shapes present in these languages [1]. Optical character recognition is nothing but electronic translation of scan images of handwritten, typewritten or printed wording interested in system-encoded content [2]. This process can be used for converting books and documents into electronics form, managing a record-keeping system in an office, or publishing the text on a website. Using OCR we can also modify the text, search for

a phrase, store it efficiently, display a replica free of scanning artifacts, as well as use processes like machine translation, text-to-speech and text mining to it.

**A. Fundamental Stages in Character Recognition:** Many times differentiating the hand written characters is almost impossible even to the human eye, and that they can only be distinguished by context for some hand-writing. It becomes critical to identify the minute differences between them in order to distinguish such characters. An important issue of different relative proportions while writing the character by different writers needs to be considered, while dealing with handwritten characters. Even the same person may not always write the same letter with the same proportions [3]. The fundamental stages in hand written character recognition system like acquiring the image, applying the pre-processing activities, decomposing in to segments, extracting the features, and finally classification and recognition. The process of scanning and translating a paper document in to digital form is called image acquisition. Almazan *et al.*, addressed the issues related to word spotting and word identification of images [4].

While capturing the image depending on the quality of the scanner noise may be introduced. Preprocessing focuses on deduction of such noise. Few more activities like thinning, normalization and segmentation of image needs to be carried out for improving the process of recognition [5]. Normalization involves resizing of characters for stroke width, slant, slope, height of the characters. Further trimming down each character image to vertical letters of uniform height made up of one pixel-wide stroke is carried out. A Normalization free, flexible and size invariant system that recognizes handwritten Devnagri characters was presented by Mukerji and Rege. Separation of characters in its constituent strokes by modification in thinning technique and direction codes worked efficiently and improved the overall accuracy of the system [6]. A 8-step Rotation invariant thinning algorithm constructing 20 rules that are concurrently applied to each pixel iteratively, in order to remove every point lying on the exterior boundaries of the symbol was developed that helped in preserving the shape of the character as well as improving the recognition accuracy [7].

Segmentation which is a decisive activity in the process of character recognition decomposes the input image into sub