## Character Recognition: An Optimal Feature Based Approach

Kaiser J. Giri, Rumaan Bashir, Javaid Iqbal Bhat

Islamic University of Science and Technology, Awantipora, J&K kaiser.giri@islamicuniversity.edu.in, rumaan.bashir@islamicuniversity.edu.in, javaidonnet@gmail.com

## Abstract

This paper presents an optimal feature based recognition technique for English characters. In order to improve the efficiency of the recognition system, the entire process is divided into two phases. In the first phase, the features necessary for recognition are extracted from the characters under consideration in an optimal manner and a feature matrix is generated so as to create a knowledge base. In second phase, a time efficient searching algorithm is used to recognize the characters using the already created feature matrix. The technique has been successfully tested on various fronts and satisfactory results have been obtained.

**Keywords:** Character, Recognition, Feature Extraction, Node Matrix, Decision Tree, Active Nodes, Passive Nodes.

## **1. INTRODUCTION**

Character recognition is the process of extracting characters from a document image and converting them to their equivalent text [1]. Character recognition plays an important role in converting an electronic document stored as document image to its textual form which can later be edited and processed by the computers. Character recognition comprises of a series of steps such as digitization, preprocessing, segmentation, extraction of features, their classification and knowledgebase creation. Character recognition being an essential activity in document image processing & analysis, has undergone through research over past few decades [2-7]. A number of character recognition techniques related to various languages have been developed so far, however the development of methods to read the text with same capability as those of humans is still an open area for researchers.

Among many other languages, English has comparatively a wide range of audience and is therefore being most widely used across the world. Recognition of characters belonging to English language has therefore resulted into an active area of research and many techniques have been developed accordingly [3, 8]. The traditional character recognition systems suffer from low recognition accuracy due

