Abstract

This paper presents a novel methodology for recognizing machine printed Devanagari script document. Shirorekha Chopping based preprocessing is chosen to enable the segmentation of printed text into various characters. Fuzzy Directional Features have shown improvement over commonly used Directional features. A set of 8 directional Fuzzy Directional Features (FDF) for each character is extracted and classified to the appropriate character class. Radial Basis function (RBF) kernel based Support Vector Machines (SVM) model is used for training the various multi font characters and testing the Devanagari document to be recognized. Experiments are conducted for the multi font Devanagari document recognition. The recognition rate of the proposed OCR system with the image document of Devnagari Script has been found to be 97.9% for two fonts Mangal and Krutidev.

References

Robust Printed Devanagari Document Recognition using Hybrid Approach of Shirorekha Chopping, Fuzzy Directional Features and Support Vector Machine


**Index Terms**

Computer Science

Pattern Recognition

**Keywords**

Devanagari OCR Shirorekha Chopping Fuzzy Directional Features Support Vector Machine