Abstract

In this paper, we describe the line, word, character and top character segmentation for printed Hindi text in Devanagari script. And also describe the line and word segmentation for printed text in Gurmukhi script. A performance of 100% at line level, approximately 100% at word level, 99% at character level, and 97% at top character level for Devanagari script and performance of 100% at line level and 99% at word level for Gurmukhi script is obtained. Here we have observed the performance of segmentation with the help of five documents in devanagari script and five documents in gurmukhi script.

Reference

- S. Antani and L. Agnihotri, Gujrati Character Recognition, in Proceedings of the international conference on Document Analysis and Recognition (ICDAR-99), Bangalore, India,

**Index Terms**

Computer Science

Pattern Recognition

**Key words**

Optical character recognition

Character Segmentation

Middle Zone

Upper Zone
Lower Zone

line

word

character

top character