

Segmentation of Lines and Words of Handwritten Devanagari Text using Connected Components with Statistics Method

Vijay More

Faculty, MET's Institute
of Engineering, Bhujbal
Knowledge City, Nashik,
vbmore2005@rediffmail.
com

Madan Kharat

Head, Dept. of Computer
Engineering, MET's
Institute of Engineering,
Bhujbal Knowledge City,
Nashik,
mukharat@rediffmail.co
m

Shyamrao Gumaste

Head, Dept. of
Information Technology,
MET's Institute of
Engineering, Bhujbal
Knowledge City, Nashik,
svgumaste@gmail.com

Abstract: In preprocessing step of handwritten Devanagari text recognition, segmentation is the required step. Segmentation accuracy of Devanagari text characters is completely depends on accurately segmented lines and words from handwritten documents. There are many issues and challenges for segmenting lines and words correctly from these documents. Segmentation of lines and words from Devanagari text documents is not seen addressed in more detail as of now, whereas it has been addressed in more sense for other script documents other than Devanagari. Here we worked on segmentation of lines and words using Connected Components with Statistics Method on PHDIndic_11 dataset and received line segmentation accuracy of 91.91% and word segmentation accuracy of 72.89%.