Abstract

Text line localization and segmentation is an important preprocessing stage in the context of document image analysis. Text lines must be localized first and then segmented in the logical order. The final recognition results are highly dependent on the results of text line segmentation. Historical documents have free form handwritten text and pose a great challenge for text line segmentation. The presence of connected and overlapping characters make the segmentation task more challenging. Smearing techniques have been conventionally used for the purpose of text line localization. In this paper, performance analysis of various smearing techniques is carried out on the text line localization of Kannada historical scripts.

References

- Vassilis Katsouros and Vassilis Papavassiliou, "Segmentation of handwritten Document images into text lines", Institute for Language and Speech Processing / R. C. Athena; Greece.
- Partha Pratim Roy, Umapada Pal, Joseph Llados; Morphology based Handwritten Line Segmentation Using Foreground and Background Information; International Conference on Frontiers in Handwriting Recognition, Montreal, Canada, pp. 241-246, 2008
on Document Image Analysis for Libraries (DIAL), 2006
  - Abderrazak Zahour, Brunco Taconet, Laurence Likforman-Sulem, Wafa Boussellaa, "Overlapping and Multi - touching text line segmentation by Block Covering analysis &quot; Pattern Analysis and Applications, Volume 12, Issue 4, pp335-351, 2009

Index Terms

Computer Science  
Pattern Recognition

Keywords

Smearing Techniques  
Run Length Smoothing Algorithm  
Binary Transition Count Map  
Adaptive Local Connectivity Map

Touching / Overlapping Components