Impact of Smearing Techniques on Text Line Localization of Kannada Historical Scripts

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 Historical documents have free form handwritten text and pose a great challenge for text line segmentation.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.
- Douglas J. Kennard, William A. Barrett; Separating Lines of Text in Free-Form Handwritten Historical Documents; Proceedings of the Second International Conference
on Document Image Analysis for Libraries (DIAL), 2006
- Z. Shi, S. Setlur and V. Govindaraju, Text extraction from gray Scale Historical document images using adaptive local connectivity map, In Proceedings of 8th International Conference on Document Image Analysis and Recognition, 2005
- Dr. M. G. Manjunath and G. K. Devarajswamy, Kannada Lipi Vikasa, Yuvasadhane, Bengaluru.

Index Terms

Computer Science
Pattern Recognition

Keywords
Smearing Techniques Run Length Smoothing Algorithm Binary Transition Count Map Adaptive Local Connectivity Map Touching / Overlapping Components