Abstract

Optical Character Recognition (OCR) is one of the challenging areas in the domain of image processing, where the handwritten or printed characters are digitized by using an optical scanner. The image is then analyzed broadly by two methods – (i) matrix space analysis method and (ii) feature space analysis method. Matrix space analysis method takes more memory space and time, compared to feature space analysis. However, it works fine for the scripts in which the strokes are prominent, e.g. English numeric scripts. On the other hand, the feature analysis method is useful where the scripts are complex and having more similarity between the letters in its writing style. Hence, the feature analysis approach is more useful to
many of the regional languages. In this paper, we have used the Ant-miner algorithm (AMA) for offline OCR of hand written Oriya scripts, popularly known as Utkal lipi. The AMA is a rule-based approach. The rules are incrementally tuned during the training. The Oriya language contains more than 50 distinct characters i.e. 12 Swara-varnas (i.e., vowels) and 38 Byanjan-varnas (i.e., consonants) and their composite characters. In this work, for the analysis, we define three types of block as per the writing styles of the scripts. AMA is then tested with four characters from each block: finally, a character recognition tool has been developed using Matlab for observation and validation.

References

- Dash T., Nayak T., Chattopadhyay S. "Offline Verification of Hand Written Signature Using Adaptive Resonance Theory Net (Type-1)"; In the proceedings of the 4th International Conference on Electronic Computer Technology (ICECT-2012 Vol-2) Kanyakumari, India (6-8 April); ISBN: 978-1-4673-1849-5; DOI: 978-1-4673-1/12; IEEE catalog number: CFP1295F-PRT, IEEE Xplore.
Optical Character Recognition using Ant Miner Algorithm: A Case Study on Oriya Character Recognition

- Mishra S, Nanda D., Mohanty S "Oriya Character Recognition using Neural Networks"; ICCT Vol. 2 Spl. Issue 2, 2010 for International Conference [ICCT-2010],


Index Terms

Computer Science  
Pattern Recognition

Keywords

Off-line character recognition  
Optical character recognition  
Image enhancement  
Oriya scripts

Ant-minor algorithm

Feature space analysis