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 handwritten 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 blocks 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

the Human Pattern Matching System


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


- Chattopadhyay S. "A Prototype Depression Screening Tool for Rural Healthcare: A
Step towards e-Health Informatics,” Journal of Medical Imaging and Health Informatics vol. 2, issue 3, pp. 244-249, 2012.


Index Terms

Computer Science

Pattern Recognition

Keywords

Off-line character recognition  Optcal character recognition  image enhancement

Oriya scripts

Ant-minor algorithm

Feature space analysis