Abstract

The main difficulty in developing a successful Optical Character Recognition (OCR) system lies in the confusion between the characters. In the case of Amazigh writing (Tifinagh alphabets), some characters have similarities based on rotation or scale. Most of the researchers attempted to solve this problem by combining multiple descriptors and/or classifiers which increased the recognition rate, but at the expense of processing time that becomes more prohibitive. Thus, reducing the confusion of characters and their recognition times is the major challenge of OCR systems. In this paper, we present an off-line OCR system for Tifinagh characters. Using a proposed key point extraction algorithm, character skeleton is divided into several segments. The length and orientation of every segment are stored into a feature vector, then the relationship between these segments are represented by a graph in the form of an incidence matrix. Classification is done by searching for similarity between test images and their counterpart in the reference database by comparing their incidence matrix. When search provides multiple classes, feature vectors are compared and input image is assigned to the closest class. Based on experiments done on 3300 images, an accuracy of 99% is achieved.
References


Index Terms

Computer Science  Pattern Recognition

Keywords

Feature points Extraction, Graph Theory, Incidence Matrix