Combination of Different Feature Sets and SVM Classifier for Handwritten Gurumukhi Numeral Recognition

Abstract

A lot of research has been done in recognizing handwritten characters in many languages like Chinese, Arabic, Devnagari, Urdu and English. This paper focuses on the problem of recognition of isolated handwritten numerals in Gurumukhi script. We have used different feature extraction techniques such as projection histograms, background directional distribution (BDD) and zone based diagonal features. Projection Histograms count the number of foreground pixels in different directions such as horizontal, vertical, left diagonal and right diagonal creating 190 features. In Background Directional Distribution (BDD) features background distribution of neighbouring background pixels to foreground pixels in 8-different directions is considered forming a total of 128 features. In the computation of diagonal features, image is divided into 64 equal zones each of size 4×4 pixels then features are extracted from the pixels of each zone by moving along its diagonal, thus consisting of total 64 features. Different combinations of these features are used for forming different feature vectors. These feature vectors are classified using SVM classifier as 5-fold cross validation with RBF (radial basis function) kernel. The highest accuracy achieved is 99.4% of whole database using combination of background directional distribution and diagonal features with SVM classifier.
References

- G. S. Lehal and Chandan Singh, "A Complete Machine printed Gurmukhi OCR System.,"
Combination of Different Feature Sets and SVM Classifier for Handwritten Gurumukhi Numeral Recognition

Index Terms

- Computer Science
- Pattern Recognition

Keywords

- Handwritten Gurumukhi Numeral Recognition
- Feature Extraction
- Projection
- Histograms
- Background Directional Distribution (bdd) Features
- Diagonal Features
- Svm Classifier
- Rbf Kernel