K-WAY TREE CLASSIFICATION BASED ON SEMI-GREEDY STRUCTURE APPLIED TO MULTISOURCE REMOTE SENSING IMAGES

Yang-Lang Chang[1*], *Zhi-Ming Chen*[2], *Hsuan Ren*[3], *Jyh-Perng Fang*[1], *Wen-Yew Liang*[4]
Tung-Ju Hsieh[4], and *Kun-Shan Chen*[3]

[1]Department of Electrical Engineering
National Taipei University of Technology, Taiwan

[2] Department of Computer Science and Information Engineering
National Central University, Taiwan

[3]Center for Space and Remote Sensing Research
National Central University, Taiwan

[4]Department of Computer Science and Information Engineering
National Taipei University of Technology, Taiwan

*Correspondence author: ylchang@ntut.edu.tw

Abstract

In traditional supervised pattern classification, the *Bayes* classifier models the classification as a distribution estimation problem. *Bayes* theory can be used to solve classification problems, but it requires complete knowledge of probability distributions. It is a difficult task to determine the right parametric form of the distribution. It is also necessary to have good parameter estimation methods and sufficient training samples that provide accurate information. For remotely sensed images, obtaining such complete knowledge is generally difficult. This led to the development of the *minimum classification error* (MCE) to overcome these problems [1]. MCE learning attempts to minimize either the empirical error rate or the expected error rate, given an arbitrary choice of the discriminant function. The goal of MCE training is to be able to correctly discriminate the observations for best classification results rather than to fit the distributions to the data.

Recently, a generalized positive Boolean function (GPBF) classifier scheme is proposed by the authors based on MCE criteria to the multisource data fusion for landslide classification [2]. GPBF implements the MCE as a criterion to improve classification performance. It makes use of MCE criteria to apply both positive and negative samples as training parameters for GPBF classifier. GPBF can also take all competing classes into consideration, according to individual significance, when searching for the GPBF classifier parameters. This is very important because we lack complete knowledge of the data

distribution particularly in dealing with multisource remotely sensed image data sets. It also makes multiclass classification possible for our proposed method.

In this paper we present a new supervised classification method, referred to as the *k-way tree semi-greedy* (KTSG) classifier, for supervised classification of multisource remote sensing images. The *k-way tree,* also known as the *k-ary tree*, is a tree with no more than k children for each node. It is a *k-dimensional* space partitioning structure for constructing the labeled sample points. The proposed KTSG is organized by a *k-way tree* in which every node (*semi-independence part in semi-greedy* structure) is composed of a set of *k-dimensional* positive and negative labeled samples as represented as a percentage, i.e. the corresponding ratio of number of samples between a specific class and other classes. It iteratively divides the *k-dimensional* hyperplane into 2^k subspaces according to the centroids of the labeled (training) samples of all classes. KTSG uses an unbalanced *k-way tree* in which each leaf node with *semi-greedy* structure has different region sizes regarding the distributions of all classes. Each *semi-greedy* leaf node are constructed of training samples of different classes in which the instances belonging to a specific class are labeled as positive samples and the rest are labeled as negative samples. By constructing the *k-way tree* based on each specific class, the statistical ratios between different classes are then compared as a basis for stopping the new subspace separating and identifying which class belongs to which subspace. By delivering both positive and negative samples of different classes to KTSG learning modules, KTSG outperforms traditional classifiers in terms of classification accuracies. The effectiveness of the proposed KTSG is evaluated by fusing *MODIS/ASTER airborne simulator* (MASTER) images and *airborne synthetic aperture radar* (AIRSAR) images for land cover classification during the Pacrim II campaign. According to experimental results, the KTSG performs as we expect.

Keywords: generalized positive Boolean function (GPBF), minimum classification error (MCE), k-way tree semi-greedy (KTSG) classifier

REFERENCES

[1] B. H. Juang, W. Chou, and C. H. Lee, "Minimum classification error rate methods for speech recognition," *IEEE Trans. Speech Audio Process.* Vol. 5, Issue 3, pp. 257-265, 1997.

[2] Yang-Lang Chang, Long-Shin Liang, Chin-Chuan Han, Jyh-Perng Fang, Wen-Yew Liang, and Kun Shan Chen, "Multisource Data Fusion for Landslide Classification using Generalized Positive Boolean Functions," *IEEE Trans. Geosci. Remote Sensing*, Vol. 45, Issue 6, Part 1, pp. 1697-1708, 2007.