Acoustic Scene Classification by using Combination of MODWPT and Spectral Features

Mie Mie Oo
mimioo.mdy@gmail.com
University of Computer Studies, Mandalay, UCSM

Lwin Lwin Oo
lwinlwinoo@gmail.com
University of Computer Studies, Mandalay, UCSM

Acoustic Scene Classification (ASC) is classified audio signals to imply about the context of the recorded environment. Audio scene includes a mixture of background sound and a variety of sound events. In this paper, we present the combination of maximal overlap wavelet packet transform (MODWPT) level 5 and six sets of time domain and frequency domain features are energy entropy, short time energy, spectral roll off, spectral centroid, spectral flux and zero crossing rate over statistic values average and standard deviation. We used DCASE Challenge 2016 dataset to show the properties of machine learning classifiers. There are several classifiers to address the ASC task. We compare the properties of different classifiers: K-nearest neighbors (KNN), Support Vector Machine (SVM), and Ensembles Bagged Trees by using combining wavelet and spectral features. The best of classification methodology and feature extraction are essential for ASC task. In this system, we extract at level 5, MODWPT energy 32, relative energy 32 and statistic values 6 from the audio signal and then extracted feature is applied in different classifiers.