

Music Genre Classification Using Modulation Spectral Features and Multiple Prototype Vectors Representation

李建興, 周智勳, 連振昌, 方仁政

Computer Science & Information Engineering

Computer Science and Informatics

chlee@chu.edu.tw

Abstract

In this paper, we will propose an automatic music genre classification approach based on long-term modulation spectral analysis of spectral (OSC and MPEG-7 NASE) as well as cepstral (MFCC) features. A modulation spectrogram corresponding to the collection of modulation spectra of MFCC/OSC/NASE will be constructed. The modulation spectrum is then decomposed into several logarithmically spaced modulation subbands. For each modulation subband, a new set of modulation spectral features, including modulation spectral contrast (MSC), modulation spectral valley (MSV), modulation spectral energy (MSE), modulation spectral centroid (MSCEN) and modulation spectral flatness (MSF) are then computed from each modulation subband. To cope with the problem that the feature vectors extracted from the music tracks of identical music genre might differ significantly, each music genre is modeled with a number of representative prototype vectors generated by c-means clustering algorithm. An information fusion approach which integrates both feature level fusion method and decision level combination method is then employed to improve the classification accuracy. Experiments conducted on ISMIR 2004 music dataset have shown that our proposed approach can achieve higher classification accuracy than other approaches with the same experimental setup.

Keyword : Mel-frequency cepstral coefficients, modulation spectral analysis, music genre classification, normalized audio spectrum envelope, octave-based spectral contrast