World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:12, No:10, 2018

Music Genre Classification Based on Non-Negative Matrix Factorization Features

Authors : Soyon Kim, Edward Kim

Abstract : In order to retrieve information from the massive stream of songs in the music industry, music search by title, lyrics, artist, mood, and genre has become more important. Despite the subjectivity and controversy over the definition of music genres across different nations and cultures, automatic genre classification systems that facilitate the process of music categorization have been developed. Manual genre selection by music producers is being provided as statistical data for designing automatic genre classification systems. In this paper, an automatic music genre classification system utilizing non-negative matrix factorization (NMF) is proposed. Short-term characteristics of the music signal can be captured based on the timbre features such as mel-frequency cepstral coefficient (MFCC), decorrelated filter bank (DFB), octave-based spectral contrast (OSC), and octave band sum (OBS). Long-term time-varying characteristics of the music signal can be summarized with (1) the statistical features such as mean, variance, minimum, and maximum of the timbre features and (2) the modulation spectrum features such as spectral flatness measure, spectral crest measure, spectral peak, spectral valley, and spectral contrast of the timbre features. Not only these conventional basic long-term feature vectors, but also NMF based feature vectors are proposed to be used together for genre classification. In the training stage, NMF basis vectors were extracted for each genre class. The NMF features were calculated in the log spectral magnitude domain (NMF-LSM) as well as in the basic feature vector domain (NMF-BFV). For NMF-LSM, an entire full band spectrum was used. However, for NMF-BFV, only low band spectrum was used since high frequency modulation spectrum of the basic feature vectors did not contain important information for genre classification. In the test stage, using the set of pre-trained NMF basis vectors, the genre classification system extracted the NMF weighting values of each genre as the NMF feature vectors. A support vector machine (SVM) was used as a classifier. The GTZAN multi-genre music database was used for training and testing. It is composed of 10 genres and 100 songs for each genre. To increase the reliability of the experiments, 10-fold cross validation was used. For a given input song, an extracted NMF-LSM feature vector was composed of 10 weighting values that corresponded to the classification probabilities for 10 genres. An NMF-BFV feature vector also had a dimensionality of 10. Combined with the basic long-term features such as statistical features and modulation spectrum features, the NMF features provided the increased accuracy with a slight increase in feature dimensionality. The conventional basic features by themselves yielded 84.0% accuracy, but the basic features with NMF-LSM and NMF-BFV provided 85.1% and 84.2% accuracy, respectively. The basic features required dimensionality of 460, but NMF-LSM and NMF-BFV required dimensionalities of 10 and 10, respectively. Combining the basic features, NMF-LSM and NMF-BFV together with the SVM with a radial basis function (RBF) kernel produced the significantly higher classification accuracy of 88.3% with a feature dimensionality of 480.

Keywords : mel-frequency cepstral coefficient (MFCC), music genre classification, non-negative matrix factorization (NMF), support vector machine (SVM)
Conference Title : ICMIR 2018 : International Conference on Music Information Retrieval
Conference Location : London, United Kingdom
Conference Dates : October 15-16, 2018