MUSIC GENRE/MOOD CLASSIFICATION: MIREX 2012

Shin-Cheol Lim¹, Karam Byun¹, Jong-Soel Lee², Sei-Jin Jang², and Moo Young Kim¹

¹Dept. of Information and Communication

Engineering, Sejong University, Seoul, Korea {en.shincheol,ge2ode2}@gmail.com, mooyoung@sejong.ac.kr

ABSTRACT

In this submission system, the spectro-temporal features are extracted based on the timbre features such as melfrequency cepstral coefficients. The feature-selection algorithm such as a support vector machine (SVM) ranker is applied for the spectro-temporal features to reduce the computational complexity. An SVM is used as a classifier.

1. INTRODUCTION

We proposed the novel music genre/mood classification system. Firstly, timbral features are extracted from the audio signals. Then, the spectro-temporal features are extracted using the timbral features. Secondly, the featureselection algorithm is applied for the spectro-temporal features to reduce the computational complexity. Lastly, genre/mood modeling and classification are performed based on a support vector machine (SVM).

2. FEATURE EXTRACTION

2.1 Timbral features

The following timbral features are extracted: melfrequency cepstral coefficients (MFCC), decorrelated filter banks (DFB), and octave-based spectral contrast (OSC). DFB are extracted using a high-pass filter instead of discrete cosine transform in MFCC. The timbral feature vectors are extracted using a hamming window of around 92ms with 50% overlap.

2.2 Spectro-temporal features

Instead of directly using the timbral features, we extract statistical features and modulation features. The statistical features such as mean, variance, min, and max are extracted from the timbral feature vectors for each texture window [1]. The modulation features include modulation spectral flatness/crest measures (MSFM/MSCM) [2], modulation spectral contrast/valley (MSC/MSV) [3], feature-based MSV/MSC [4], and feature-based MSFM/MSCM [5]. For texture window of around 3s with 50% overlap, the spectro-temporal features are extracted.

2.3 Feature selection

In our system, a feature-selection algorithm is applied to increase the recognition performance and decrease the

²Digital Media Research Center, Korea Electronics Technology Institute, Seoul, Korea {leejs,sjjang}@keti.re.kr

computational complexity. As a feature-selection algorithm, an SVM ranker is used. The selected feature sets are different for genre and mood classification.

3. CLASSIFICATION

In this system, we use LIBSVM [6] as a classifier. The feature is normalized between 0 and 1 before using an SVM classifier. A Gaussian radial basis function (RBF) kernel is used. RBF kernel parameters such as *C* and γ are determined by a grid-search method. For the GTZAN database, the proposed system yields 87.4% accuracy with 10 fold-cross validation.

4. ACKNOWLEDGEMENTS

This work was supported by the Ministry of Knowledge Economy grant funded by the Korea government (No. 10037244).

5. REFERENCES

- [1] G. Tzanetakis, "MARSYAS submissions to MIREX 2009," in *Music Information Retrieval Evaluation eXchange (MIREX)*, 2009.
- [2] D. Jang and C. D. Yoo, "Music information retrieval using novel features and a weighted voting method," in *Proc. IEEE Int. Symposium Industrial Electronics*, pp. 1341-1346, 2009.
- [3] C.-H. Lee, J.-L. Shih, K.-M. Yu, and J.-M Su, "Automatic music genre classification using modulation spectral contrast feature," in *Proc. IEEE Int. Conf. Multimedia and Expo*, pp. 204-207, 2007.
- [4] C.-H. Lee, J.-L. Shih, K.-M. Yu, and H.-S. Lin, "Automatic music genre classification based on modulation spectral analysis of spectral and cepstral features," *IEEE Trans. Multimedia*, vol. 11, no. 4, pp. 670-682, 2009.
- [5] S.-C. Lim, S.-J. Jang, S.-P. Lee, and M. Y. Kim, "Music genre/mood classification using a featurebased modulation spectrum," in *Proc. IEEE Int. Conf. Mobile IT Convergence*, pp. 133-136, 2011.
- [6] C.-C. Chang and C.-J. Lin, "LIBSVM A Library for Support Vector Machines," *ACM Trans. on Intelligent Systems and Technology*, 2011.