MUSIC GENRE/MOOD/COMPOSER CLASSIFICATION: MIREX 2014 SUBMISSIONS

Seung-Ryoel Baek
Dept. Information and Communication Engineering, Sejong University, Seoul, Korea
daey7979@nate.com

Moo Young Kim
Dept. Information and Communication Engineering, Sejong University, Seoul, Korea
mooyoung@sejong.ac.kr

ABSTRACT

In this submission system, novel features are extracted based on the timbre and visual features. The feature-selection algorithm such as a support vector machine (SVM) ranker is applied to reduce the computational complexity and to increase the recognition accuracy. An SVM is used as a classifier.

1. INTRODUCTION

We propose the novel music genre/mood/composer classification system. Firstly, timbre and visual features are extracted from the audio signal. Then, the temporal features are extracted using the timbre and visual features. Secondly, the feature-selection algorithm is applied for the temporal and visual features to reduce the computational complexity and to increase the recognition performance. Lastly, genre/mood/composer 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: mel-frequency 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 Visual features

Visual features [1] are extracted from the music clip spectrograms. The spectrograms are divided into the 7 octave based subband sub-images. Lastly, the visual features are extracted via Gabor filter with 6 orientations and 7 scales.

2.3 Temporal features and visual features

Instead of directly using the timbral features, we extract statistical features, modulation features and visual features. The statistical features such as mean, variance, min, and max are extracted from the timbral feature vectors for each texture window [2]. The modulation features include modulation spectral flatness/crest measures (MSFM/MSCM) [3], modulation spectral contrast/valley (MSC/MSV) [4], feature-based MSV/MSC [5], feature-based MSFM/MSCM [6], and feature-based autoregressive model features. For texture window of around 3s with 50% overlap, the temporal features are extracted. Timbral-based-visual features are extracted with timbral feature spectrogram. The spectrograms are divided into 3 subbands. And the Gabor filter is applied.

2.4 Feature selection

In our system, a feature-selection algorithm is applied to increase the recognition performance and decrease the computational complexity. As a feature-selection algorithm, an SVM ranker is used.

3. CLASSIFICATION

In this system, we use LIBSVM [7] 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.

4. ACKNOWLEDGEMENTS

This research was supported by the MSIP (Ministry of Science, ICT and Future Planning), Korea, under the ITRC (Information Technology Research Center) support program (NIPA-2014-H0301-14-1021) supervised by the NIPA (National IT Industry Promotion Agency).

5. REFERENCES

[1] M. Wu, Z. Chen, J. R. Jang, J. Ren, "Combining visual and acoustic features for music genre classification," in *10th International Conference on Machine Learning and Applications*, 2011, pp. 124–129.

[2] G. Tzanetakis, "MARSYAS submissions to MIREX 2009," in Music Information Retrieval Evaluation eXchange (MIREX), 2009.

[3] 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.

[4] 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.

[5] 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.

[6] S.-C. Lim, S.-J. Jang, S.-P. Lee, and M. Y. Kim, "Music genre/mood classification using a feature-based modulation spectrum," in Proc. IEEE Int. Conf. Mobile IT Convergence, pp. 133-136, 2011.

[7] C.-C. Chang and C.-J. Lin, "LIBSVM - A Library for Support Vector Machines," ACM Trans. on Intelligent Systems and Technology, 2011.