MUSIC GENRE CLASSIFICATION MIREX 2014 SUBMISSIONS

Shumin Xu
Dept. Electrical and Computer Engineering,
University of Rochester, NY
Shumin.xu@rochester.edu

Yating Gu
Dept. Electrical and Computer Engineering,
University of Rochester
yatinggu@rochester.edu

ABSTRACT

The submission intends to classify the music pop genre. The features are extracted from spectral pattern, delta spectral pattern, spectral contrast, correlation pattern, beat spectrogram, logarithmic fluctuation pattern, audio texture and beat-level texture. A one versus one support vector machine (SVM) is used as classifier and the principal component analysis (PCA) is applied to reduce the computational complexity.

1. INTRODUCTION

We propose several novel features for genre classification, instead of traditional acoustic features, the task extracts image texture features and adds rhythm pattern features. A new dataset Pandora is built for experiment, which reach an accuracy rate of 67%.

2. FEATURE EXTRACTION

The submission uses several timber features: spectral pattern, delta spectral pattern, spectral contrast pattern; rhythm pattern features: logarithmic fluctuation pattern, correlation pattern; beat-level features: texture, beat-level texture and beat spectrogram.

The task first take STFT with the sampling rate 22050Hz, window size of 2048, hop size of 512. Later the spectrogram is transferred into cent-spectrogram, as a method for human perception. Block-level features is adopted in the submission. The block size is obtained by optimization. A final step of summarization is applied to all the block-level features into a single song-level feature for each song. All features are concatenated in a single feature vector for each song.

3. CLASSIFICATION

The task use libsvm toolbox built in Matlab, a one versus one SVM is used to classify the genre. All feature vectors are concatenated in a matrix, before classification, a normalization and rescale step is taken. For accuracy concern, the feature is normalized by subtract local-mean value in advance in the feature extraction step. We also use PCA as the method to reduce computational complexity and time cost. In the submission, PCA is used in two way, one is used for each feature dimension reduction and the other way is using PCA after all feature extraction.

4. REFERENCES

[1] Klaus Seyerlehner, Gerhard Widmer, Markus Schedl, PeterKnees, “Automatic music tag classification based on block-level features”, ISMIR- 2010.

[2] Jonathan Foote, Shingo Uchihashi, “The Beat Spectrum: A New Approach to Rhythm Analysis”, IEEE International Conference on Multimedia and Expo, Tokyo, Japan, pp.881-884, 2001.

[3] Hugo Fastl, “Fluctuation strength and temporal masking patterns of amplitude-modulation broadband noise”, Hearing Research, 8 (1982) 59-69, Elsevier Biomedical Press.

[4] D.P.W. Ellis, “Beat tracking by dynamic programming,” Journal of New Music Research, vol. 36, no. 1, pp. 51-60, July 2007.

[5] 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.

[6] B. S. Manjunath and W.Y. Ma, “Texture Features for Browsing and Retrieval of Image Data,” IEEE Trans. Pattern Analysis and Mach. Intel., pp. 837-842, 1996.