MIREX 2015 Submission for Audio Classification Tasks

Ming-Ju Wu
Department of Computer Science
National Tsing Hua University,
Hsinchu, Taiwan
brian.wu@mirlab.org

Zhe-Cheng Fan
Department of Computer Science
and Information Engineering
National Taiwan University,
Taipei, Taiwan
lambert.fan@mirlab.org

Jyh-Shing Roger Jang
Department of Computer Science
and Information Engineering
National Taiwan University,
Taipei, Taiwan
jang@mirlab.org

ABSTRACT

Using one type of feature for learning may be inadequate to achieve optimal results for music classification. In this submission, we applied the confidence-based late fusion approach [1] to combine acoustic and visual features for audio classification tasks, where the confidence-based late fusion approach is based on SVMs using RBF kernel. The confidence score of an SVM prediction can be defined via two key factors. The first key factor is based on the distance between the test instance and the hyperplane in the Hilbert space of the SVM. The second key factor is based on the distance between the test instance and its nearest neighbor in the Hilbert space of the SVM. Because different types of feature may exhibit different discriminative powers for a given music clip, confidence-based late fusion selects a presumably more accurate prediction. Please refer [1] for a detailed description of our submission.

1. REFERENCES

[1] M. J. Wu and J. S. Jang, "Combining Acoustic and Multilevel Visual Features for Music Genre Classification," *ACM Transactions on Multimedia Computing, Communications, and Applications*, Vol. 12, No. 1, September 2015. Available: http://dl.acm.org/citation.cfm?id=2801127