

MusCat: A Music Browser Featuring Abstract Picture and Zooming User Interface

Kaori Kusama

Ochanomizu University

Itoh Takayuki

Ochanomizu University

1. INTRODUCTION

We present MusCat, a music browser featuring abstract pictures and zooming user interface. It visualizes collections of tunes by abstract pictures, based on features, not based on metadata. Our technique firstly calculates features of tunes, and hierarchically clusters the tunes according to the features. It then automatically generates abstract pictures for each tune and cluster, so that users can recognize characteristics of tunes more instantly and intuitively. It finally displays the tunes and their clusters by using abstract pictures.

2. PRESENTED TECHNIQUE

2.1 Music Feature Extraction

Current our implementation uses features calculated by MIR-toolbox [1]. We had a feasibility study of features applying many sample tunes, and subjectively found that the following 11 features (RMS energy, Low energy, Tempo, Zero crossing, Roll off, Brightness, Roughness, Spectral irregularity, Inharmonicity, Key, and Mode) were effective for our purpose.

2.2 Clustering

Next, the technique hierarchically clusters music files based on features. We experimentally applied various clustering techniques for our own collection of tunes, and compared the results by carefully looking at the dendrogram. As a result we selected Ward method as a clustering algorithm, because it successfully divides a set of tunes into evenly sized clusters.

We also apply a metric learning method for constrained clustering, so that we can generate more preferable clusters. We defined multiple constrains based on statistics of subjective tune grouping tests by examinees.

2.3 Generation of Abstract Picture

Our technique generates abstract pictures to express tunes, and displays so that users can intuitively select the tunes.

The technique selects colors of abstract pictures based on color image scale [2]. It is a color system that distributes combination of three colors in a two dimensional space, so called “sensibility space”, which has warm-cool and soft-hard axes. Based on our subjective, we assign Mode is to warm-cool axis and Roll off to soft-hard axis. We think listeners often use substitute adjective words such as “light” or “soft” for the impression of music, and often these impression is related to frequency-based tone balances. Calculating Mode and Roll off of a tune, our implementation places the tune onto the sensitivity space, and selects the color closest to the tune.

Our design first generates the following three layers, 1) gradation layer, 2) a set of circles, and 3) a set of stars, as shown in Figure 1.

2.4 Visualization

The technique displays a set of abstract picture by applying the image browser CAT [1], which displays representative pictures

of all clusters while zooming out, and pictures in the specific clusters while zooming in. We extend CAT so that we can use CAT as a music browser, where we call the extended CAT as “MusCat,” as an abbreviation of “Music CAT”. Figure 2 is an example snapshot of MusCat.

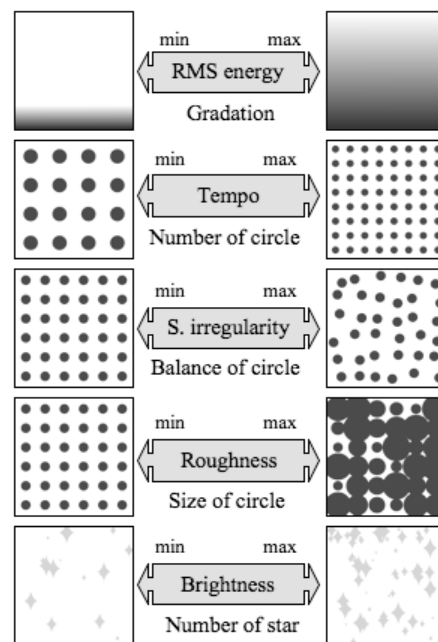


Figure 1. Automatic generation of three layers of images based on music features.

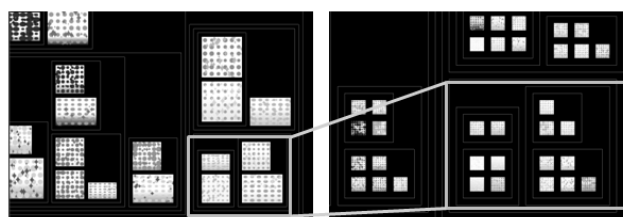


Figure 2. Example of visualization by MusCat. (Left) MusCat displays representative pictures while zooming out. (Right) MusCat displays thumbnail pictures while zooming in.

3. REFERENCES

- [1] O. Lartillot: “MIRtoolbox,” <http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>
- [2] S. Kobayashi: *Color System*, Kodansha, 2001.
- [3] A. Gomi, R. Miyazaki, T. Itoh, J. Li: “CAT: A Hierarchical Image Browser Using a Rectangle Packing Technique,” *12th International Conference on Information Visualization*, pp.82-87, 2008.