

MusCat: 楽曲の印象表現に基づいた一覧表示の一手法

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

草間 かおり 伊藤 貴之*

Summary. Today many people store music media files in personal computers or portable audio players, thanks to recent evolution of multimedia technologies. The more music media files these devices store, the more difficult it is to search for tunes that users want to listen to. We propose MusCat, a music browser to interactively search for the tunes according to features, not according to metadata (e.g. title, artist name). The technique firstly calculates features of tunes, and then hierarchically clusters the tunes according to the features. It then automatically generates abstract pictures, so that users can recognize characteristics of tunes more instantly and intuitively. It finally visualizes the tunes by using abstract pictures. The technique enables intuitive music selection with the zooming user interface.

1 はじめに

音楽の再生機材として PC やポータブル音楽プレイヤーが主流になり、その内蔵記憶装置の記憶容量の増大に伴い、個人が保有する楽曲数が膨大化している。これにより、ユーザが聴きたい楽曲を見つけるのが困難になることが多くなると考えられる。多くの場合においてユーザは、曲名やアーティスト名などのメタデータをもとに楽曲を検索する。しかし、これらのメタデータに頼らずに、旋律や調性などの楽曲特徴や印象に基づいて楽曲を検索する技術には、まだ検討の余地があると考えられる。

本報告では、膨大な楽曲をメタデータに依存することなく、楽曲そのものの特徴に基づいてクラスタリングし、各クラスタの印象表現のために生成した画像を一覧表示するブラウザを提案する。

2 処理手順

2.1 特徴量抽出

近年さまざまな楽曲特徴検出手法が提案されている。我々は本研究において、MIR toolbox[1] を用いて、楽曲から 11 個の特徴量を抽出している (表 1 参照)。

本手法ではこれら 11 個の特徴量を同等に扱うために、特徴量 f を正規化した f' を用いる。ここで $f' = (f - f_{min}) / (f_{max} - f_{min})$ であり、 f_{min} と f_{max} は特徴量の最大値と最小値である。

2.2 クラスタリング

続いて、特徴量に基づいて階層的にクラスタリングする。楽曲は様々な特徴すべてが複合的に作用していることから、各楽曲に対しすべての特徴を加味して分析する多変量クラスタ分析法を用いる。多変量解析に

Copyright is held by the author(s).

* Kaori Kusama and Takayuki Itoh, お茶の水女子大学大学院 人間文化創成科学研究科 理学専攻

表 1. 楽曲特徴.

特徴量	説明
RMS energy	音量
Low energy	弱音の割合
Tempo	テンポ
Zero crossing	波形が 0 値をとる回数
Roll off	85 % を占める低音域の割合
Brightness	1500Hz 以上の音域の割合
Roughness	不協和音の多さを示す値
Spectral irregularity	音質の変化の大きさ
Inharmonicity	根音に従っていない音の量
Key	主に使われている音
Mode	major と minor の音量の差

もいくつかの方法があるが、我々は分類感度が高く、クラスタの要素数が比較的均一なウォード法を採用している。

また、印象におけるユーザの満足度を高めるようクラスタリングに制約条件を付与させる。

2.3 画像生成

本手法では、ユーザが特定の印象を有する楽曲を直感的に発見するために、抽象的な印象画像を提示する。

2.3.1 色の選択

まず、印象画像を生成する色を決定する。我々の実装では、カラーイメージスケール [2] (図 1 参照) をもとに印象画像の色を選択する。カラーイメージスケールは、warm-cool 軸と soft-hard 軸の 2 次元の感性空間に色を分布させたものである。本手法では楽曲を感性空間に配置する。

そこで、warm-cool 軸、soft-hard 軸に対応する楽曲特徴量を考える。一般にメジャーコードに明るい・暖かいといったポジティブな印象を表わすコードであり、マイナーコードは冷たい・悲しいといったネガティブな印象を表わすコードである。このことから本手法では、マイナーコードとメジャーコードの音量差を示す

特徴量 mode を warm-cool 軸に割り当てる．また一般的に，低音が少なければ軽い・やわらかい音，低音が多ければ重い・重い音と表現することが多い．このことから本手法では，低音域の音量を示す特徴量 roll off を soft-hard 軸に割り当てる．以上により，楽曲を色の感性空間に配置し，楽曲の座標値と最も近い座標の色をその楽曲の色とする．

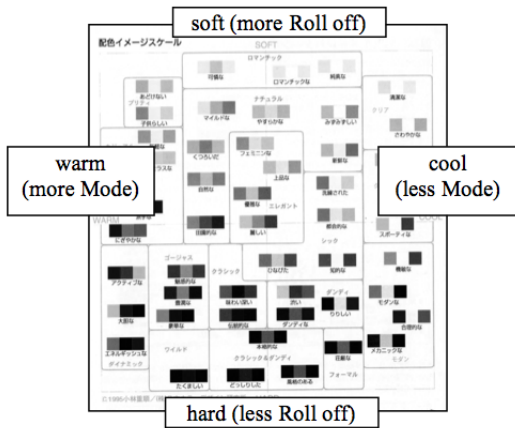


図 1. カラーイメージスケールの感性空間

2.3.2 デザインの選択

続いて，楽曲の特徴量に対応したデザインを生成する．本報告では，デザインの文法 [3] に基づいたデザインの一例を示す（図 2 参照）．

これらのデザインをそれぞれ生成した後，3 枚の画像を重ね合わせることによって，印象画像となる一枚の画像を生成する．

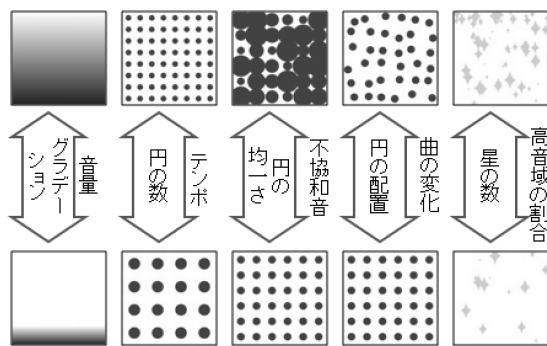


図 2. 楽曲特徴量に基づくデザイン生成

2.4 一覧表示

本手法では，画像ブラウザ CAT[4] に楽曲再生機能を設けたブラウザ MusCat を用いて，印象画像を一覧表示する．

CAT では，前処理として大量画像を階層型にクラスタリングし，それを互いに重ならずに等しいサイズで

一覧表示する．CAT では各画像をサムネイル表示し，サムネイルを長方形の枠で囲うことでクラスタを表現する．さらに CAT では，ズーム率に合わせた詳細度制御を設けている．ズームイン時は，低階層クラスタの各々の画像サムネイルを表示する．そしてズームアウト操作に伴って，低階層クラスタを示す長方形領域を，各低階層クラスタの代表画像で置き換えて表示する．さらにズームアウト操作を続けると，高階層クラスタを示す長方形領域を各高階層クラスタの代表画像で置き換えて表示する．このように，CAT は階層化された画像群に対するズーム操作によって，直感的に画像を絞り込みながら閲覧できる．MusCat は，以上のような CAT の諸機能を継承し，かつ印象画像に対するマウス操作での楽曲再生機能を有する．

MusCat による印象画像の表示結果を，図 3 に示す．

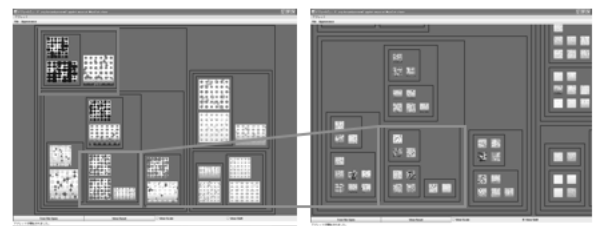


図 3. 図面の例

3 まとめと今後の課題

本報告では，特徴に基づいて分類した楽曲を一覧表示するとともに，その再生機能を有するブラウザ MusCat を提案した．なお，提案手法の詳細については文献 [5] をご参照いただきたい．今後の課題として，再生機能をさらに充実させるとともに，分類のユーザ満足度を向上させたい．

参考文献

- [1] O. Lartillot. MIRtoolbox. <http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>.
- [2] 小林 重順. カラーシステム. 日本カラーデザイン研究所 (編). 1999.
- [3] C. Leborg. デザインの文法. ピーエヌエヌ新社. 2007.
- [4] A. Gomi. R. Miyazaki. T. Itoh. J. Li. A Hierarchical Image Browser Using a Rectangle Packing Technique. In *12th International Conference on Information Visualization*, pp. 82–88, 2008.
- [5] 草間 かわり. 伊藤 貴之. MusCat: 楽曲データの印象表現に基づいた一覧表示の一手法. 第 81 回音楽情報科学研究会, MUS81-19. 2009.