

# MusCat: 楽曲データの印象表現に基づいた一覧表示の一手法

草間かおり（指導教員：伊藤貴之）

## 1 概要

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。本論文では、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 提案内容

### 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。本論文において、MIR toolbox[1]を用いて、楽曲から11個の特徴量を検出している。(表1参照) また、Marsyas[2]を用いて、MFCCという9次元の特徴量を検出している。

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

### 2.2 クラスタリング

音楽的特徴量に基づくクラスタリング 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を複合してひとつの楽曲であるので、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解析法を用いる。ここで、多変量解析にもいくつかの方法があるが、本手法は分類感度が高く、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。

音響的特徴量に基づくクラスタリング 一方、特徴量 MFCC に基づいたクラスタリングにおいては、群平均法を採用している。

### 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 音楽的特徴量による画像生成

色選択 まず、印象画像を生成する色を決定する。本論文の実装では、カラーイメージスケール [3] (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し、感性空間上で最も距離の近い色を楽曲の色とする。続いて、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、mode, roll off の2特徴量に基づいて、楽曲を色の感性空間に配置する。その後、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。

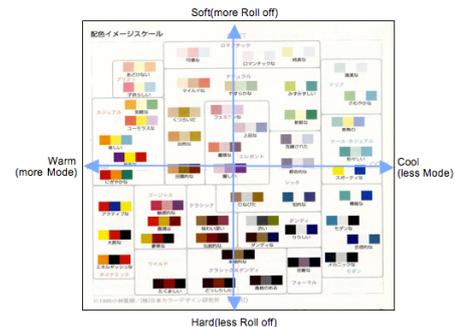


図 1: カラーイメージスケールの感性空間

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。本論文では、デザインの文法 [4] に基づいたデザインの一例を示す。このデザインでは、RMS energy を背景のグラデーション、Tempo を円の個数、Spectral irregularity を円の配置バランス、Roughness を各円のサイズのばらき、Brightness を星の個数に割り当てる(図2参照)。

### 2.3.2 音響的特徴量による画像生成

Marsyas によって抽出される特徴量 MFCC は、音量を表す値として、9つの正方形の大きさで表わすことで画像を生成する。

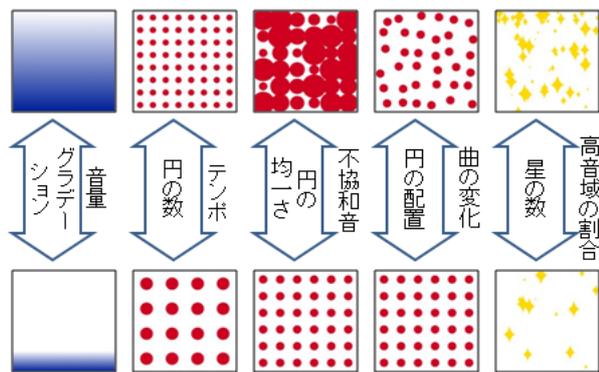


図 2: 楽曲特徴量に基づくデザイン生成

## 2.4 一覧表示

本手法では、画像ブラウザ CAT[5] 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。

## 3 結果

**音楽的特徴量に基づいた可視化結果** MIRtoolbox から得られた 11 個の特徴量を基にしてクラスタリング及び画像生成した表示結果を、図 3 に示す。図 3 (左) はズームアウト時 (右) は左の画像をズームイン操作することによって得られた画像である。ズームアウトにより、高階層クラスタの構成には色との相関性が大きいことを観察できる。各クラスタの個々の画像を見比べると、左上のクラスタは比較的円のばらつきの多い画像で構成されていることがわかる。また、ズームイン画面を見ると、色相が青に近い画像で構成されるクラスタにおいて、低階層では暖色系と寒色系に分類していること、また円のサイズが均一であること、などを観察できる。このことから、短調が多く協和音の多い楽曲を、さらに細かく比較的長調が多い曲か否かで細かく分類していることがわかる。

**音響特徴量に基づいた可視化結果** MFCC によってクラスタリングし、画像生成した結果を、図 4 に示す。各音域

の音量を視覚的に認識することが可能となり、緑色の正方形が大きい画像の楽曲は男性ボーカルによる伴奏楽器の少ない楽曲であることがわかる。このように、音響特徴に基づく分類においては、ジャンルに捕われずに、楽曲編成が類似している楽曲毎に分類された。

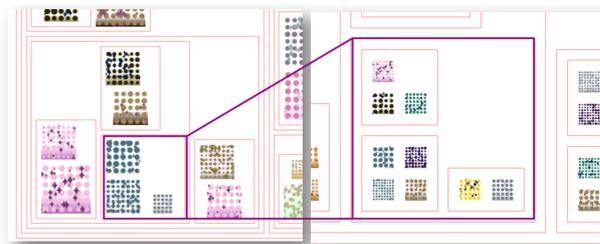


図 3: 音楽的特徴量による可視化結果例



図 4: 音響的特徴量による可視化結果例

また、音響的特徴に基づいたクラスタリング結果を、2 種の画像を切り替え表示することにより、ジャンルにとらわれずに楽器構成による探索するだけでなく、分類した各楽曲の音楽的特徴量を認識することが可能となった。これらのように MusCat を使うことで、楽曲を聴く前に楽曲の特徴を理解することが可能となる。

## 4 まとめ

本論文では、楽曲を印象に基づいて直感的に認識するために、楽曲の特徴量から印象画像を自動生成して、クラスタリングした楽曲に対して一覧表示する手法 MusCat を提案した。

## 謝辞

ユーザテストに協力頂いた諸氏に、謹んで感謝の意を表す。本論文の実験には、RWC 研究用音楽データベースに収録された楽曲を使用させていただいた。

## 参考文献

- [1] O. Lartillot. MIRtoolbox. <http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>.
- [2] Marsyas. <http://marsyas.info/>
- [3] 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.
- [4] C. Leborg. デザインの文法. ピーエヌエヌ新社. 2007.
- [5] 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.