

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。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ったと考えられる。ここで我々は、曲名やアーティスト名など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、例えばカフェやバーの BGM の選曲のように、楽曲のタイトルやアーティストにこだわらず、場所、時間、状況によるムードに基づいて楽曲を選曲したい場面があると考える。あるいはユーザが、アーティストや作曲者に関わらず、ある楽曲について似た曲調の楽曲を見つけたい場面があると考える。このような場面では、楽曲を印象に基づいて分類し、特定の印象を有する楽曲群を簡単に提示できるツールがあると便利であると考え、

以上の背景に基づいて本研究では、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から、印象に基づいて整理、分類し、わかりやすく一覧表示するツールを提案する。

2. 関連研究

音と感性について、コレスポンデンス分析による楽曲の特徴認識[1]では、楽曲に対する印象を表す形容詞を評価することによって、音楽の特徴認識には【重い—軽やか】、【スピードのある—ゆったりしている】、【powerのある—powerのない】という3軸で認識すると述べている。

色と感性について、色空間と感性の反映方法[2]では、単色刺激から受ける感性情報の因子として Evaluation, Potency, Activity があること、また2色配色の刺激から受ける感性情報の因子として Harmony が存在することを述べている。このことから、色を用いて感性を表現することが可能であると考えられる。

印象の合う画像で楽曲を表現する手法に MIST[3]がある。MIST ではサンプルとなる楽曲や画像の特徴量から相関性を学習させ、任意の楽曲や画像の感性語適合度を自動算出し、楽曲の印象に適した画像をアイコンとして表示する。Music icons[4]では、ニューラルネットワークを用いて適した幾何学模様の画像を選択している。これらの手法が予め用意した画像を選択提示するのに対して、本研究では印象画像の自動生成手法を提案している。

楽曲分類において、D. Liu[5]は、楽曲を音量、リズム、音質に基づいて階層的にクラスタリング手法を提案している。この手法ではまず、音量によって高階層クラスタを生成した後、リズム、音質で低階層クラスタを生成する。しかし本手法は楽曲分類に専念するものであり、表示方法は提案していない。

3. 処理手順

3.1 特徴量検出

本手法では最初に、楽曲データから表 1 に示す特徴量を検出する。楽曲には、次第に曲の印象が変わるものや、一つの楽曲に複数の印象を合わせ持った曲があり、楽曲全体の特徴量の一つに定めるのは困難である。そこで、一つの画像を生成するには一定区間ごとに特徴量を検出する。次第に楽曲が変化する場合は各区間の特徴量の平均値をその楽曲全体の特徴量とし、特定の時刻から明確

に印象が変わる楽曲に対しては、印象が変化する前と後に分けて複数の特徴量を1曲の特徴量とする。現段階では、楽曲の中間部のある5秒間から得た特徴量を楽曲全体の特徴量として暫定的に定めている。

表 1: 楽曲特徴

	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
Spectral	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

3.2 印象画像の生成

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。人間の感覚には、音楽や音を聞いて色を感じる「色聴」という知覚があることから、音と色の印象は密接な関係にある。よって、視覚的に表現され得ない楽曲の印象を可視化するには、色印象に重きを置いた印象画像を媒体として表示することにより、楽曲の背景知識や歌詞の意味よりも楽曲特徴に基づいた印象を重視して、視覚的表現ができると考えられる。

3.2.1 色の割当

まず、印象画像を生成する色を決定する。本手法ではカラーイメージスケール[6]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。楽曲もこの感性空間に分布させ、空間上で最も距離の近い色をその楽曲の色とする。

そこで、warm-cool 軸、soft-hard 軸に適応する楽曲特徴量を考える。一般にメジャーコードはポジティブな印象を表わすコードであり、マイナーコードはネガティブな印象を表わすコードであることから、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 軸に割り当てることにした。また低音が少なければ軽い音、低音が多ければ重い音と表わすことに基づいて、低音域の音量を示す特徴量 Roll off を色の感性空間の soft-hard 軸と同等に扱い、楽曲を色の感性空間に分布させる。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。

3.2.2 デザインの生成

続いて、デザインの文法[7]に基づいて、また楽曲の特徴量に応じてデザインを生成する手法を提案する。本報告では一例として、RMS energy を背景のグラデーション、Tempo を円の個数、Spectral irregularity を円の配値バラ

ス, Inharmonicity を円の半径のばらつき, Brightness を星の個数に割り当てるデザイン生成を示す(図1参照). 続いて, このデザインに従って生成された3枚の画像を重ね合わせることによって, 一枚の画像を生成する(図2参照).

以上のデザイン生成手法はあくまでも, 文献を参考に我々の主観的に基づいてデザインであり, まだ検討の余地が十分にある. デザイン生成結果の妥当性について今後検証し, 被験者からの意見をもとに, 楽曲の印象に合うデザインを考案する予定である.

絶対評価ではなく, 一覧表示する楽曲に対して相対的な印象画像を生成するために, 楽曲特徴量 f において楽曲特徴量のしきい値 f_i を以下のように定める.

$$f_i = \frac{i(\max(f) - \min(f))}{n} + \min(f) \quad (1 \leq i \leq n)$$

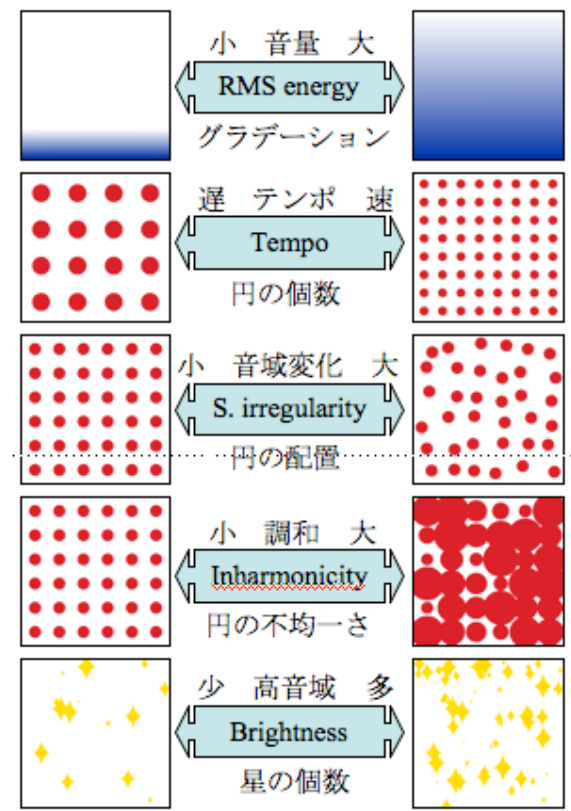


図1: 楽曲特徴量とデザインの対応関係

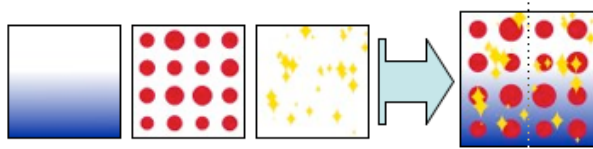


図2: 印象画像の生成

3.3 クラスタリング

続いて, 楽曲を特徴量に基づいて階層的にクラスタリングする. 本研究では, 多次元変数クラスタ分析およびウォード法を用いている.

3.4 一覧表示

続いて, 生成された印象画像およびクラスタリング結果を, 階層構造を有する画像群の可視化手法 CAT[8]を用

いて一覧表示する. 図3(上)に示すように CAT では, 各楽曲を表す印象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 は, ズーム率に合わせた詳細度制御機能として, ズームイン時は各々の画像サムネイルを表示し, ズームアウト時にはクラスタを示す長方形領域を代表画像に置き換えて表示する(図3(下)参照). このように CAT を用いることで, 階層化された画像群に対するズーム操作により, 直感的な楽曲群の絞り込み操作を実現する.

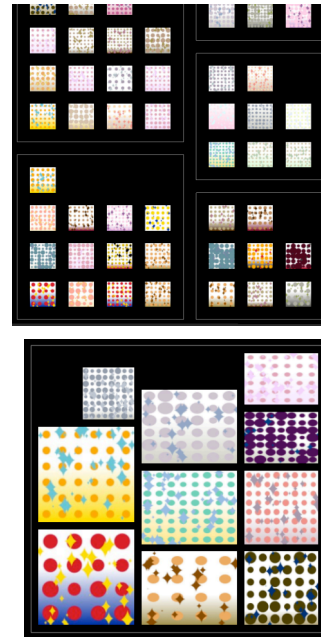


図3: CATでの表示結果例

4. まとめと今後の課題

本研究では, 楽曲を印象に基づいて直感的に一覧表示する手法として, 楽曲の特徴量から印象画像を自動生成して, 楽曲のクラスタリング結果と合わせて表示する手法を提案した. 以下に今後の課題を示す.

- ・ 楽曲特徴量の再検討
- ・ 印象画像デザインの改良
- ・ 音楽プレイヤーとしてのユーザインタフェース開発
- ・ 被験者実験による満足度評価

参考文献

- [1] 山脇他, コレスポンデンス分析による楽曲の特徴認識, 感性工学研究論文集, Vol. 7, No. 4, pp. 659-663, 2008.
- [2] 富岡他, 色空間と感性の反映方法, ISDL Report, 20040621002, 2004.
- [3] 小田他, MIST: 音楽アイコンの自動選択の一手法, WISS2007, pp. 115-116, 2007.
- [4] P. Kolhoff et al., "Music Icons: Procedural Glyphs for Audio Files," *IEEE SIBGRAP*, pp. 289-296, 2006.
- [5] D. Liu et al., "Automatic Mood Detection from Acoustic Music Data," *ISMIR-03*, 2003.
- [6] 小林, カラーイメージスケール, 日本カラーデザイン研究所(編), pp. 94-123, 講談社, 2001.
- [7] C. Leborg, Visual Grammar: デザインの文法, ビー・エヌ・エヌ新社, 2007.
- [8] A. Gomi et al., "CAT: A Hierarchical Image Browser Using a Rectangle Packing Technique," *IV08: 12th International Conference on Information Visualization*, pp.82-87, 2008.