

GRAPE: グラデーション画像によるプレイリスト単位の音楽の可視化

魚田 知美† 伊藤 貴之†

†お茶の水女子大学 理学部 情報科学科 〒112-8610 東京都文京区大塚 2-1-1

1. 概要

携帯型音楽プレイヤーの登場や音楽配信技術の発展で、音楽と触れ合う手段が増え、音楽は主要なコンテンツに成長したといえる。しかし「音楽」というメディアは、実際に聞いてみるまで中身を把握できないため、内容把握に時間がかかる。音楽の内容を短時間で直感的に把握する一手段として、画像で情報の全体像を表現する「可視化」が非常に有用であるといえる。

本報告では、「1画像で1プレイリスト」を表すプレイリスト単位の音楽の可視化、GRAPE (GRadation Arranged Playlist Environment)を提案する。我々は音楽鑑賞の際、1曲聞くごとに次に聞く曲を逐一選択する方法よりも、プレイリストやアルバムといった音楽のまとまりを選択する方法のほうが一般的だと考える。よって、プレイリストやアルバムの全体像を可視化することで、多くのユーザの選曲操作スタイルを変えることなく選曲操作を支援できると考える。

GRAPEは音楽再生機器上での選曲を容易にする他に、次のような効果も考えられる。まず、プレイリストやCDの中身に関する傾向や差異を知りたい時に、本手法によってまとまったオーディオ情報を概観できる。また、プレイリストを単純に画像として扱い、視覚的に楽しむ、ユーザ間で共有する、といったことが容易になる。可視化結果である画像が様々な機器やソフトウェアで扱いやすいメディアであることから、この他にも様々な活用ができると期待できる。

なお、本報告で言う「プレイリスト」とは、ユーザが作成したプレイリストや音楽プレイヤーによって自動作成されたプレイリストだけでなく、各アルバムや各アーティストまたは音楽コレクション全体などといった音楽のまとまりを総称するものとする。

本手法では SOM (Self Organization Map)を適用することで、グラデーション風の画像を自動生成する。グラデーションには、リズム感・メリハリを持つという表現効果がある。また、人の心を動かす心理効果がある。単色に比べてグラデーションの物には「触りたい」、「欲しい」といった要求が大きく作用する。この心理効果は GUI のボタン、その他の各種製品にも広く利用されている。このような効果は、可視化や情報を専門とする人だけではなく、音楽を聴く全ての人を対象とする音楽プレイヤーのデザインに有益と考えた。

2. 関連研究

楽曲特徴量から楽曲の可視化画像を生成する手法の例として、MusCat[1]やMusicThumbnailer[2]がある。

MusCatは、MIRtoolbox[3]を用いて楽曲特徴量を抽出し、その特徴量に基づいて図1(左)のような印象画像を生成する。MusicThumbnailerは、MARSYAS[4]を用い

て楽曲特徴量を抽出し、その特徴量に基づいて図1(右)のようなサムネイル画像を生成する。

以上にあげた手法はいずれも楽曲単位の視覚的表現を目指している。表示方法によって1画面にプレイリストを表示することはできるが、プレイリスト単位での音楽表現に特化しているとはいえない。一方GRAPEでは「1画像で1プレイリスト」を目標とし、プレイリスト単位での音楽表現に特化していることから、ユーザの選曲操作スタイルに沿った可視化であると考えられる。しかし代わりに、以上の手法と比べると各楽曲の細かい内容の描写に欠ける面はある。

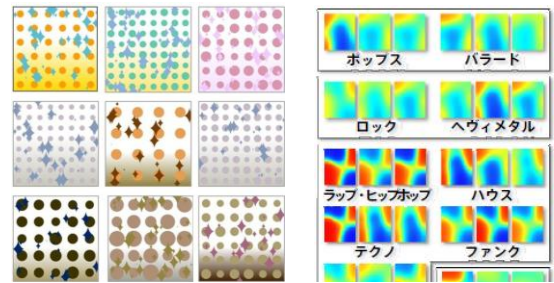


図1. (左) MusCat[1]による可視化例
(右) MusicThumbnailer[2]による可視化例

3. 提案手法

3.1 楽曲特徴量抽出

まず各楽曲から楽曲特徴量を抽出する。楽曲特徴量の抽出には、数値解析ソフトウェア MATLAB の上に実装された、楽曲特徴分析パッケージ MIRtoolbox を用いる。GRAPE では MIR toolbox で抽出可能な特徴量のうち、RMSenergy (音量平均値), Tempo, Brightness (高音域の割合) の3つの特徴量を、以下の式によって正規化して使用する。

$$f' = \frac{f - f_{\min}}{f_{\max} - f_{\min}}$$

$f_{\max}$ 楽曲特徴量の最大値
 $f_{\min}$ 楽曲特徴量の最小値

現在は、各楽曲の任意秒数を指定して特徴量を抽出しているが、多くの楽曲は曲の全体の中で特徴量に変化する。我々が実験に用いた楽曲の数ヶ所から試験的に特徴量を抽出してみた結果、楽曲中で値が比較的安定している上記3つの特徴量を代表として使用することとした。

3.2 SOM による楽曲配置

続いて、抽出した特徴量に基づいて各楽曲を配置する。各楽曲を配置する手法としてSOMを用いる。SOMとはKohonenが提案した人工ニューラルネットワークの1種であり、以下のステップで構成される。

1. 初期値としてランダムな値を持った2次元空間の中から、入力要素と1番近い点(勝者ユニット)を探し出す。

GRAPE: A music playlist visualization technique
featuring gradation images

†Tomomi Uota and Takayuki Itoh, Ochanomizu University

- 2.勝者ユニットとその近傍が入力要素に近づくように学習を行う。
- 3.全ての入力要素について1.と2.を適用し、これを反復する。

SOMを用いることで、入力データ群の距離を保ったまま、任意の次元へと写像することができる。1楽曲を1要素としてSOMを適用することにより、似ている楽曲は近い場所に集まり、あまり似ていない楽曲は遠い場所に配置される。これにより自動的にグラデーション風の画像が生成される。また今後、使用する特徴量を増やした際にも、多次元情報を持つ要素を2次元空間上に写像すること強みとするSOMであれば柔軟に対応できると考えられる。GRAPEでは各楽曲を正方形のタイルのように見立てプレイリストごとに配列して1枚の長方形を形成する。SOMを用いる以外の手段として、ボロノイ図などの領域分割手法を適用して各領域に楽曲を割り当てることも考えたが、各楽曲を表現する領域の大きさと形状が均一である方が、GUI上での選曲操作を容易にできると考え、本研究ではSOMを採用した。

3.3 YCbCr 色空間を利用した楽曲色

続いて、抽出した楽曲特徴量に基づいて各楽曲に色を割り当てる。色を割り当てる際に我々は、YCbCr色空間を採用した。我々が実験に用いた楽曲を、試験的にRGB、HSBなどの多様な色空間に配置した結果、YCbCr空間に配置したときに最も色の分布がばらつく傾向にあり、GRAPEで採用したときの視覚的効果が高いと判断した。現時点では図3のように、3つの特徴量をYCbCr色空間へ対応させて色を決定している。

Brightnessは単純に各楽曲の輝かしさと表現することができる特徴量であるから、Y軸である明度と対応させた。RMSenergyは楽曲の音量平均値を示し、アコースティックな曲では小さな値、エレクトリックな曲では大きな値になる傾向がある。これは赤色が持つイメージ[7]である「活動的」や「エネルギーッシュ」というキーワードに意味が近いと考えて対応させた。Tempoは値が大きいほど速い曲である。よって青色が持つイメージ[7]である「開放感」や「爽やか」というキーワードに意味が近いと考えて対応させた。また、YCbCr色空間ではCr軸とCb軸の値が小さくなると緑に近づく。つまりRMSenergyとTempoの値が小さい楽曲には緑が割り当てられるということである。これは緑色が持つイメージ[7]である「穏やか」や「落ち着いた」というキーワードに意味が近いと考えた。以上の理由からYCbCr色空間へ3つの特徴量を図3のように対応させた。

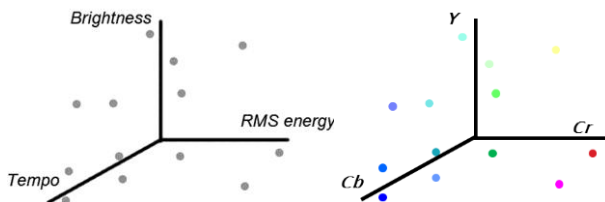


図3. 3つの特徴量と YCbCr 色空間の対応

3.4 GUI 上での画像一覧表示

以上により取得した楽曲の位置情報と色情報をプレイリスト毎にまとめて表示し、GUI上でプレイリスト群を一覧表示させる。我々が開発した GRAPE の GUI には、画像一覧表示の他に、各タイルをシングルクリックする

ことで曲名を表示し、ダブルクリックすることで特定の楽曲を再生する機能も備えている。

4. 実行結果と考察

我々は、GRAPE にプレイリスト A,B について可視化を実行した。プレイリスト A は 36 曲、プレイリスト B は 28 曲の楽曲が入っている。可視化結果を図 4 に示す。

2つのプレイリスト画像の色の明るさという観点から直感的に以下の点が推測できる。

- プレイリスト A は明るい色が多いことから、明るい楽曲が多い
 - プレイリスト B は色の明るさの分散が大きいことから、明るい楽曲も暗い楽曲も含む
- また、色相の分散という観点から直感的に以下の点が推測できる。
- プレイリスト A は赤から桃色が多いことから、Cr 軸に割り当てた RMSenergy の値が大きい楽曲が多い
 - プレイリスト B は色の分散が大きいことから、特出した傾向がない

以上のように、GRAPE によって各プレイリストの特徴を視覚的に示すことができた。



図4. (左) プレイリスト A の可視化結果
(右) プレイリスト B の可視化結果

5. まとめと今後の課題

本報告では、まとまった楽曲情報を表現する可視化手法 GRAPE を提案した。実行結果から、各プレイリストの特徴が読み取れることがわかった。

今後の課題として、ユーザテストを実施することと共に、以下の点に取り組みたい。

- 楽曲の色、位置の精度向上
- 生成画像の穴あき、楽曲重なり回避
- よりよい可視化結果を得るために必要な入力情報に関する検討

参考文献

- [1] K. Kusama, T. Itoh, MusCat: A Music Browser Featuring Abstract Pictures and Zooming User Interface, ACM Symposium on Applied Computing, Multimedia Visualization Track, pp. 1227-1233, 2011.
- [2] K. Yoshii, M. Goto, Visualizing Musical Pieces in Thumbnail Images Based on Acoustic Features, 9th International Conference on Music Information Retrieval (ISMIR), pp. 211-216, 2007.
- [3] O. Lartillot, MIRtoolbox, <http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>
- [4] G. Tzanetakis et al, MARSYAS, <http://marsyas.info/>
- [5] 芳原, 色彩の教科書, 洋泉社, 2011.