

音響特微量と歌詞情報に基づく楽曲嗜好解析支援ツールの製作 A music preference analyzer based on acoustic features and lyrics

南部理奈[†] 斎藤博昭[†]
RinaNanbu Hiroaki Saito

概要

近年の楽曲配信サービスにおいて、ユーザが楽曲を効率的に選ぶには自身の楽曲嗜好を理解する必要がある。よって従来行われていない音響特微量と歌詞情報の両者を反映した楽曲嗜好解析支援ツールを作成した。まず大量の楽曲群からそれぞれ音響特微量を抽出し、各歌詞の名詞の出現数に基づいてクラスタリングした結果を歌詞情報とした。さらにこれらの特微量を二次元平面上に可視化し、ユーザはその結果を俯瞰することで自己の楽曲嗜好を解析できる仕組みとなっている。実験の結果、本ツールの使用により多くのユーザが自己の楽曲嗜好が音響特微量と歌詞情報のどちらか、または両者の影響を受けており、楽曲嗜好に対する理解が深まることがわかった。

1. はじめに

近年、インターネットサービスの普及と多様化によって音楽と人との関わり方が多様化している。Apple 社の iTunes 等に代表されるようなオンライン楽曲購入サービスを通して、人々は Web 上で手軽に膨大な量の音楽を入手することが可能となった。また、携帯音楽プレイヤーの大容量小型化によって大量の音源データを持ち運ぶことが可能になった。これに伴い、ユーザは過剰な楽曲の中から自分の好みに合う楽曲を選ぶことが困難になってきている。そこでユーザの嗜好を反映させた楽曲推薦の仕組みが求められている。

音楽と人間の感情には深いつながりがあることが数多くの研究によって明らかにされている。音楽を選ぶ上で、個々人が重視する音響特微量や歌詞情報は異なってくる。よってユーザの嗜好がどのような音響特微量や歌詞情報に起因するかを解析することで、より精度の高い楽曲推薦が行えるといえる。しかしながら、既存の楽曲嗜好の解析に関する研究では音響特微量のみを考慮し、歌詞情報を加味して行われていない。したがって、音響特微量と歌詞情報の両者を反映し、個々人の嗜好に適した特微量を選出することが精密な楽曲推薦においては必要である。

そこで、本研究ではユーザが自分の嗜好に沿って楽曲を選出する際に重視する音響特微量や歌詞情報を明らかにするための仕組みを提案する。

2. 楽曲推薦

楽曲推薦システムは一般的に、協調型フィルタリングと内容に基づくフィルタリングに二分される。協調型フィルタリングにおいては、ユーザのプロフィールの類似度を比較し楽曲推薦を行うのに対し、内容に基づくフィルタリングは楽曲データを解析し、それぞれの楽曲の特徴を用いて

推薦を行うという手法である。

吉井ら[3]は協調型フィルタリングと内容に基づくフィルタリングを融合したハイブリッド型楽曲推薦システムを提案した。具体的には pLSI(確率的潜在意味インデクシング)を取り入れ、さらに改良手法として結び混合分布モデルやアーティスト情報の利用を提案したが、自身の嗜好が影響を受ける楽曲の特微量を特定することはできない。

また、MusiCube[6]は楽曲ファイルから抽出した音響特微量に基づき、対話型進化計算を用いてユーザの嗜好を学習し、その学習結果を GUI で提示するインタフェースである。これにより、ユーザは満足度の高い推薦結果を得るだけでなく、自身の嗜好が起因する音響特微量の提示を可能にした。我々は、音響特微量だけでなく歌詞情報も考慮することで個々の楽曲嗜好をより明確化でき、より精度の高い楽曲推薦が行えると考えている。

3. 提案手法

3.1 システム概要

本システムは、ある楽曲群の各楽曲の音響特微量と歌詞情報をまとめたファイルを入力とし、二次元平面上にマッピングしたものを出力する。システムの概要を図 1 に示した。各プロセスの詳細を以下に論じる。

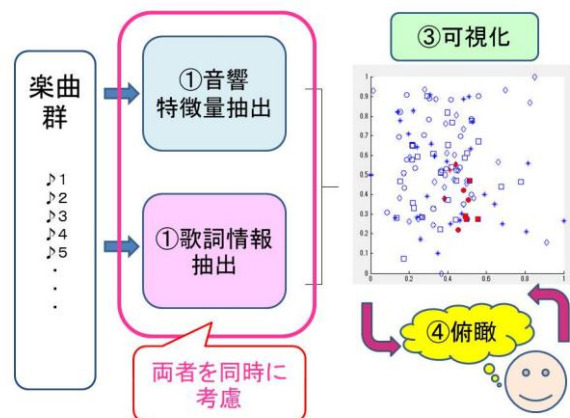


図 1 システム概要

3.1.1 音響特微量の抽出

本ツールでは、MIRtoolbox¹を用い表 1 に示す 10 個の特微量を抽出した。²

¹<http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>

²本研究では MIRtoolbox の表 1 に示す関数と 'Frame' オプションを利用し抽出した

[†] 慶應義塾大学大学院理工学研究科
Graduate School of Science and Technology, Keio University



図2 インタフェース概要

本手法では、これらの音響特徴量を同等に扱うため、音響特徴量を正規化した f を用いた。ここで、

$$f' = (f - f_{\min}) / (f_{\max} - f_{\min}) \quad (1)$$

とし、 $f_{\max}$ と $f_{\min}$ は音響特徴量の最大値と最小値である。¹

表1 音響特徴量一覧[3]

音量 (mirrms)
弱音の割合(mirlowenergy)
テンポ(mirtempo)
波形が0の値をとる回数(mirzerocross)
85%を占める低音域の割合(mirrolloff)
高音域の割合(mirbrightness)
不協和音の多さを示す値(mirroughness)
音質の変化の大きさ(mirregularity)
根音に従っていない音の量(mirinharmonicity)
長和音と短和音の音量の差(mirmode)

3.1.2 歌詞情報の抽出

以下の手順で歌詞情報の抽出を行った。

1. 各楽曲の歌詞のテキストデータを用意する。
2. 形態素解析ツール MeCab²を用いて各楽曲について名詞のみを抽出する。
3. 各名詞の出現頻度を計算する。

¹音響特徴量の抽出については MusiCube[3]で有効とされていた手法を踏襲した

²<http://mecab.sourceforge.net/>

4. k-平均法により楽曲をグループ 0,1,2,3 の4つにクラスタリングする。本研究ではユーザが大量の楽曲群の中から好みに合う楽曲を10曲選択し、それらの楽曲について解析することを想定している。したがって、選択した10曲の楽曲に対しクラスタリング結果を反映するためkの値を大きくするほどユーザの認識が困難となる。よって本研究ではk=4とした。

5. どのクラスタに所属するかを各楽曲の歌詞情報とする。

3.1.3 可視化

上記の2つの特徴量のデータをまとめたテキストファイルを入力とし、与えられたデータを二次元平面上へプロットした。表示の詳細については次節で述べる。

3.1.4 俯瞰

本システムを操作し、分布の変化を見ながら特徴量を視覚的にとらえ、自身の好みがどの特徴量に起因するかを解析する。本システムの詳しい操作については次節で述べる。

3.2 インタフェース概要

図2は本システムの全体図である。本システムは以下の5つの部分から構成される。

A. 楽曲選択部（図2-A）

楽曲群の中で、ユーザの好みの楽曲を10曲選択する。ポップアップメニューに表示されている曲番号をマウスでクリックすることで選択することができる。

B. 特徴量選択部（図2-B）

プロットの際に割り当てられるx軸、y軸それぞれに対し特徴量を選択する。いろいろな特徴量の組合せに基づく分布を見ることができる。

C. 実行部（図2-C）

「実行」ボタンをクリックすることで、上記で選択された音響特徴量に基づいてプロット結果を表示する。選択された楽曲については赤字で、それ以外の楽曲については青

字で表示される．さらに「歌詞を考慮して実行」ボタンをクリックすることで，歌詞情報に基づいてプロットの記号が○，□，◇，*の4種類に変化する．

D. プロット部 (図 2-D)

上記で選択された音響特徴量に従って入力ファイルのプロット結果を表示させる領域である．表示された分布の偏りや記号を見ることで，ユーザが選択した好みの楽曲がどのような特徴量に起因するかを視覚的に読み取ることができる．

E. 楽曲番号の表示部 (図 2-E)

「曲番号表示」ボタンをクリックし，プロット上に分布している各楽曲をクリックすると，各プロットが示す曲番号が表示される．これにより，各プロットがどの曲に対応しているかを知ることができる．

F. 最適な特徴量抽出部 (図 2-F)

特徴量選択部によりすべての特徴量の組合せを選択し分布を見ることはユーザにとってかなりの負担となる．よってこの部分ではエントロピーを用いて楽曲の密度を計算し，ユーザが選択した赤字の楽曲の分布が最も集まるような特徴量の組合せと分布を表示することができる．「No.1」ボタンをクリックすると，エントロピーが最小な特徴量の組み合わせと分布を表示する．また，「No.2」ボタンをクリックすると，特徴量「音量」を除外した組み合わせのうち，エントロピーが最小のものを表示することができる．これは，入力したデータ全体が特徴量「音量」のもので偏りやすいため，ユーザの選曲による偏りを分析するには，「音量」を除いた方が効果的な場合があるためである．

エントロピーの導出は以下の通りである．まず，特定の2つの音響特徴量を xy 軸に割り当て，その表示領域内を $N \times N$ で分割する．次に，各分割領域内の赤色のアイコンと赤色以外のアイコンを数え，以下の式よりエントロピーの合計を算出する．

$$E_{sum} = - \sum_{i=1}^{N^2} (p_{ri} \log p_{ri} + p_{qi} \log p_{qi}) \quad (2)$$

$$p_{ri} = r_i / (r_i + q_i) \quad (3)$$

$$p_{qi} = q_i / (r_i + q_i) \quad (4)$$

ここで i 番目の分割領域における r_i と q_i は赤色のアイコンの数と赤色以外のアイコンの数， p_{ri} と p_{qi} は赤色のアイコンと赤色以外のアイコンの割合とする．¹また本研究では $N=5$ とした．

図 3 は「No.1」による最適な特徴量抽出の実行結果の一例である．他の特徴量の組み合わせによる図 4 の実行結果と比較すると，赤色のアイコンが集中していることがわかる．

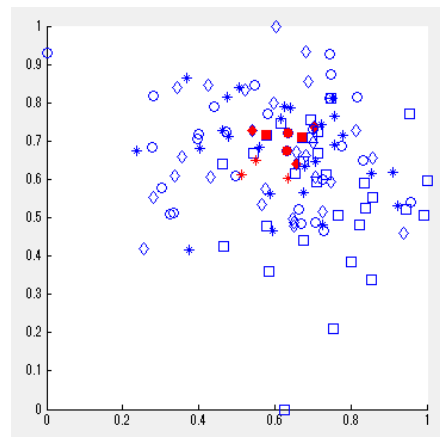


図 3 最適な特徴量抽出による結果

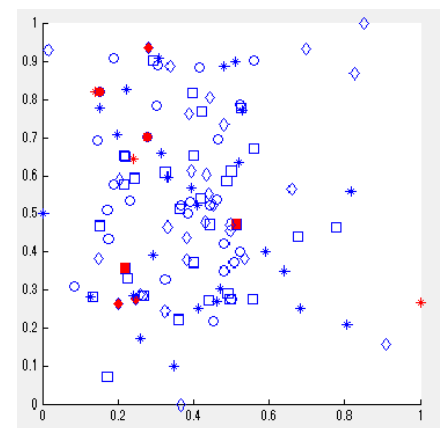


図 4 任意の特徴量選択による結果

以上のように，本システムでは音響特徴量と歌詞情報を同時に可視化することによって，ユーザの楽曲嗜好が両者に対しどのような影響を受けているかを視覚的に読み取ることができる．これにより，システム使用前には認識していなかった新たな嗜好に気づくことができると考えられる．さらに，好みの楽曲と似た特徴量を持つ他の楽曲も発見することができ，両者を反映した楽曲推薦に応用できると考えられる．

4. 実験

4.1 実験方法

本実験では，J-POP における 2000 年から 2004 年までの各年間シングル CD 売上ランキング TOP20²から 117 曲を楽曲データとして用い，10 人の被験者に以下のような実験を行った．

1. 被験者は，117 曲の楽曲リストの中から好みに合う楽曲を 10 曲選択する．

¹エントロピーの導出については MusiCube[3]で有効とされていた手法を踏襲した

²各年間のシングル CD 売上ランキングは”PRIVATeLiFE エンタメデータ&ランキング”<http://entamedata.web.fc2.com/index.html>より参照

2. 選択した楽曲について本ツールを使用し解析する.

3. 以下のアンケートに答える.

Q1:選んだ曲についてわかったことを具体的に教えてください.

Q2:ツール使用前と後で, 自分の楽曲嗜好に対する認識にどのような変化がありましたか.

Q3:このツールの改善点を教えてください.

4.2 結果

4.2.1 音響特徴量について

図 5 は被験者 A の特徴量選択部において x 軸に弱音の割合, y 軸に不協和音の多さを示す値を選択し実行した結果である.

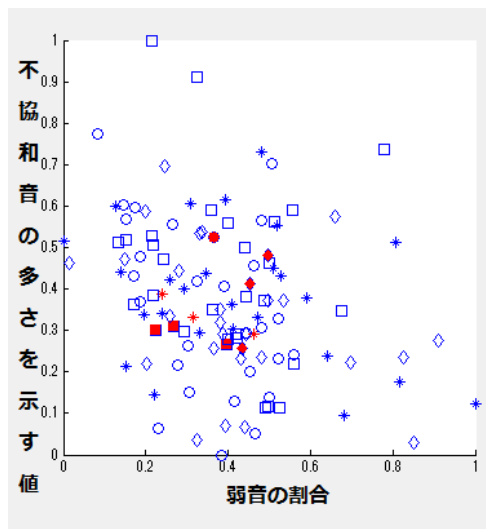


図 5 被験者 A の実行結果

また, 以下が被験者 A のアンケート結果である.

Q1:不協和音を選ぶと分布が集まっていたので不協和音が好きだとわかった.

Q2:テンポを重視していると思ったけれど不協和音の方を重視していると感じた.

これより, 本システムにおける音響特徴量の考慮により被験者 A は自身の楽曲嗜好に対して新たな理解を得たといえる.

4.2.2 歌詞情報について

図 6 は被験者 B の最適な特徴量抽出部「No.2」の実行結果である.

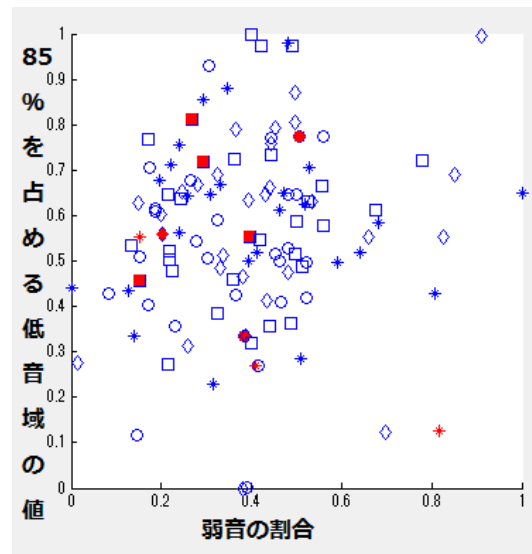


図 6 被験者 B の実行結果

また, 以下が被験者 B のアンケート結果である.

Q1:自分の音楽の趣味はバラバラなのだということがわかった.

Q2:ダイヤの歌詞の曲はあまり好きではないのかもしれないということがわかった. 確かにミスチルや EXILE はあまり好きではない.

これより, 本システムにおける歌詞情報の考慮により被験者 B は自身の楽曲嗜好に対して新たな理解を得たといえる.

4.2.3 音響特徴量と歌詞情報の両者について

図 7 は被験者 C の最適な特徴量抽出部「No.2」の実行結果であり, 図 8 は被験者 C の特徴量選択部において x 軸に 85%を占める低音域の値, y 軸に高音域の割合を選択し実行した結果である.

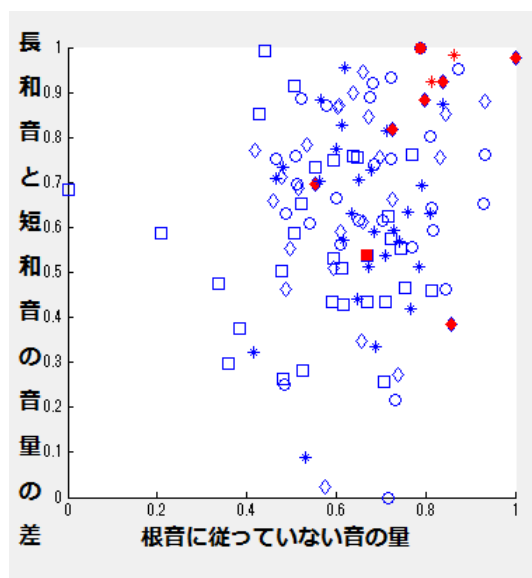


図 7 被験者 C の最適な特徴量抽出による結果

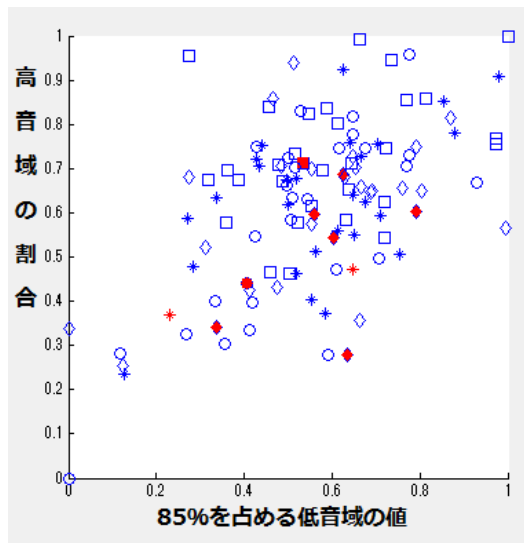


図 8 被験者 C の任意の特徴量選択による結果

また、以下が被験者 C のアンケート結果である。

- Q1:高音と低音がバランス良く入っている方が好みに近い。根音に従っていない音の量が 1.0 に近いほど好む傾向がある。ダイヤの割合が多い。
- Q2:もっと高音や低音で好みが変わると思っていたが、意外にも中間が好みだった。曲 2 と曲 69 は春の曲と思っていたが、共通した特徴が多いのかと思ったが、弱音の割合と同じダイヤであることぐらいしか共通点がなく、歌詞が好みに影響を与えていると再認識できた。

これより、本ツールにおける音響特徴量と歌詞情報の両者の考慮により被験者 C は自身の楽曲嗜好に対して新たな理解を得たといえる。

5. 考察

5.1 本システムの有用性

表 2 実験結果

被験者	音響特徴量	歌詞情報
A	○	×
B	×	○
C	○	○
D	○	○
E	○	○
F	○	×
G	×	○
H	○	×
I	○	×
J	×	×

表 2 は、全被験者のアンケート結果をもとに、音響特徴量と歌詞情報のそれぞれに対する自己の楽曲嗜好に関する発見の有無をまとめたものである。被験者 10 人のうち、音響特徴量に対し発見があったと述べた被験者は 7 人、歌詞情報に対しては 5 人、また両者に対しては 3 人という結果となった。すなわち、本システムの使用により、被験者 1 人を除き、音響特徴量と歌詞情報の少なくともどちらか一方に対して発見があり、楽曲嗜好に対する認識が深まったことがわかった。したがって、楽曲嗜好の解析においては音響特徴量と歌詞情報の両者を考慮することが必要であり、本システムが効果的であるといえる。

5.2 改善点

5.2.1 データ

本実験を行うにあたり、始めに被験者に好みに合う楽曲を選曲してもらうことから、被験者を問わず容易に選曲できるデータを扱う必要があった。そのため年間の CD 売上ランキングの結果を参照したが、被験者によっては知っている曲に限られており、その場合はあまり良い結果を得られなかった。したがって、あらゆるユーザに対応できるように、異なる世代の楽曲を用意するほか、データを固定せずに各ユーザのよく知る楽曲群を扱うことができるようにする必要があるだろう。

5.2.2 インタフェース

本システムでは、ユーザが実行結果で表示された各プロットがどの楽曲を示しているかを知るために、曲番号の表示機能を加えた。しかし、ユーザには与えられた曲番号と楽曲を照らし合わせる作業を強いることになり、使いづらい印象を与えていることが分かった。これより、インタフェース面では楽曲情報の表示方法について改善する必要があるといえる。

5.2.3 機能

本システムでは、ユーザの任意の選択に合わせて分布を表示するだけにとどまっていたため、実際に各特徴量が楽曲の中でどのような意味を成しているかの確認までユーザに促すことができなかった。このため、被験者によっては楽曲嗜好の深い理解につながらなかったと考えられる。これより、機能面では実験結果を視覚的に表示するだけでなく、楽曲の聴取や歌詞の閲覧により表示された結果を体感できるよう改善する必要があることがわかった。

6. 結論と今後の展望

6.1 結論

本論文では、音響特徴量と歌詞情報の両者を反映した可視化ツールを作成し、ツール使用を通して俯瞰することで、楽曲嗜好に対する新たな発見を見出し理解を深める仕組み作りを行った。実験の結果、多くのユーザが自己の楽曲嗜好について音響特徴量と歌詞情報のどちらか、または両者について発見を見出すことができた。よって、楽曲嗜好の解析には両者を反映する必要がある、個人の楽曲嗜好に対する理解を深める仕組みとしては有効であると言える。

6.2 今後の展望

6.2.1 本システムの向上について

今回のツール使用を含む実験において被験者は 10 人であった。本ツールの普遍的有用性を証明するために、多くの被験者に同様の過程を経てツール使用を試みる必要がある。また、その場合対象とする楽曲群は異なる年代のものを集めるほか、全被験者に対し固定とするのではなく被験者ごとに好みの楽曲を持ち寄るなど、多くの被験者に対応できるような仕組みが必要であると考えられる。

また、本ツールにおいては楽曲情報の表示に曲番号を利用した。しかし、今後あらゆる被験者、楽曲群に対し本ツールの利用を促す際、曲番号の表示だけでは対応に限界を有すると考えられる。被験者、楽曲群を問わず、一目で結果を認識できるインタフェース作りが求められる。

また、本ツールは二次元平面上へプロットするものであったため、得られた結果と実際の楽曲の内容を照らし合わせることができなかった。したがって楽曲の再生や歌詞の閲覧機能を加えることにより、得られた特徴量と楽曲との関連性をユーザがその場で体感できる仕組みは、楽曲嗜好への理解をより有効に促進すると考える。さらに、本研究では歌詞情報を考慮する際に k-平均法によるクラスタリングを用いたが、その他のクラスタリング手法により本システムが向上する可能性がある。よって歌詞情報を反映する際は、その他のクラスタリング手法についても検討する必要がある。

6.2.2 本システムの応用について

本システムの使用によって、楽曲推薦システムにおける「個人の嗜好」特定に役立つと考えられる。本システムにより得た楽曲嗜好に対する特徴と似た特徴を持つ楽曲を提示することにより、ユーザに適切な推薦がなされることが考えられる。また、継続的に本ツールを使用することで、常に変化する楽曲嗜好の変動も容易に認識でき、推薦に反映することができる。

参考文献

- [1] C.M.ビショップ, “パターン認識と機械学習”, シュブリンガー・ジャパン, 2008.
- [2] Lee, J.S. and Lee, J.C., “A Music Recommendation System Based on Context reasoning”, Lecture Notes in Computer Science, Vol.427, pp.190-203, 2006
- [3] Yoshii, K. and Goto, M., “Continuous PLSI and Smoothing Techniques for Hybrid Music Recommendation”, International Society for Music Information Retrieval Conference (ISMIR), pp.339-344, 2009
- [4] 梶克彦, 平田圭二, 長尾確, “状況と嗜好に関するアノテーションに基づくオンライン楽曲推薦システム”, 情報処理学会研究報告, Vol.2004, No.127, pp.33-38, 2004
- [5] 松原正樹, 李軒, “音楽への感性を開拓するメタ認知支援ツール”, 日本認知科学会第 27 回発表論文集, pp.842-845, 2010
- [6] 斎藤優理, 伊藤貴之, “MusiCube: 特徴量空間における対話型進化計算を用いた楽曲提示インタフェース”, 第 25 回人工知能学会全国大会論文集, 2011
- [7] 町田和嘉子, 伊藤貴之, “Lyricon-複数アイコンの自動選択による楽曲構成の可視化”, Vol.2009-MUS-81, No.18, 2009