

## 印象の個人性と順序関係を考慮した楽曲の印象推定

岩月 靖典<sup>†</sup>, 酒向 慎司<sup>†</sup>, 北村 正<sup>†</sup>

<sup>†</sup> 名古屋工業大学大学院工学研究科

### 1 はじめに

近年のインターネットを介した楽曲配信サービスの拡大に伴い、利用者が膨大な楽曲の中から目的の楽曲を効率良く見つけ出す方法が必要が高まっている。その中で、音楽的な特徴を用いた方法や、人間が楽曲を聴いたときに抱く印象を利用した方法のような楽曲の音響情報そのものの利用が検討されている。本研究では、人間にとって直感的な情報を利用した検索が可能になり、更には検索以外にも場面に合わせたBGMの推薦にも活用できると考え、音響的な特徴から印象を推定する方法について検討する。

楽曲の印象推定は、人間の印象があいまいで複雑であることから難しい問題であるとされている。複雑にしている大きな要因として印象の個人性がある。複数人が同一の楽曲を聴いた場合においても印象は個人毎に様々であり、個人性は印象推定においても考慮すべき重要な要素である。先行研究[1]では、楽曲に対する印象を利用者に評価させることで個人の印象の傾向を取得しているが、楽曲を評価する作業は時間がかかるため利用者の負担が大きい。そこで、本研究では比較的取得が容易な年齢や性別のようなプロフィール情報を利用して個人性を考慮する方法を提案する。

一方で、印象の傾向が類似していても評価の幅が個人によって異なるという問題がある。このような問題に対して、被験者に二つの楽曲の印象を比較させることで評価値を得る方法が提案され、間隔尺度で楽曲の印象を評価するよりも別の被験者と類似した印象評価をしているように感じるという調査結果が得られている[2]。本研究では、これを拡張して間隔尺度で得られた評価に対して順序関係を取り入れ、順序尺度として扱うことにより印象評価の正規化を行う。

### 2 個人性と順序関係を考慮した提案手法

#### 2.1 提案法の概要

提案法の概略図を図1に示す。提案法では、入力情報として利用者のプロフィール情報と欲しい楽曲の印象情報を用いる。プロフィールにより、いくつかの印象推定器から最適な推定器を選択し、選択された印象推定器によりデータベース内の楽曲の印象を推定する。ここで、印象推定器は人間の印象評価データを用いて音響信号と印象の関係が学習された重回帰モデルを指す。そして、入力された印象と推定された各楽曲の印象をケンドールの順位相関係数により類似関係の評価し、最も相関がある楽曲を出力する。

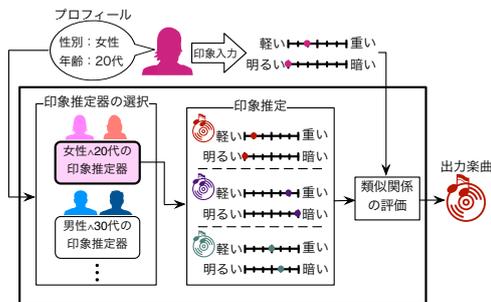


図1: 提案法の概略図

### 2.2 印象評価による人物のクラスタリングと順序関係の考慮

人間の印象には個人性が存在するが、その一方で印象は個人毎に完全に独立ではなく、類似した印象の傾向を持つ場合も存在する。そこで、類似した印象を持つ人物でクラスタ化し、クラスタ内の人物の印象評価を利用して印象推定器を学習させる。本研究では、楽曲の印象を表1に示す14対の感性語で表現し、1~7の7段階のSD (Semantic Differential) 法で取得した。SD法で取得した値に対して順序関係を考慮した正規化を施し、その後クラスタリングを行う。

正規化の手順として、最初に被験者毎に各々の印象について楽曲間で評価値が高い順に順位を付けを行う。これにより1~7で表現されていた値が1~楽曲数  $N$  の値に変換される。その後、楽曲間で順位付けられた値に対して、14対の感性語の間で順位付けを行い、最終的に1~14の値が付けられる。それぞれの順位付けには平均順位法を用いる。平均順位法は(8,5,5,4,2)のデータにおける同順位の値5に対する処理として、(2位+3位)/2 = 2.5位として扱い、(1位, 2.5位, 2.5位, 4位, 5位)と順位付けを行う方法である。

このようにして正規化された値に対して、被験者別に変数を音響特徴量とした回帰係数を算出する。そして、算出された回帰係数を用いてクラスタリングを行う。クラスタリングの手法としては、初期値に依存せず分類感度が高いクラスタリングができる方法で知られているウォード法を用いる。

表1: 楽曲の印象表現に用いた感性語対

軽い	-	重い	明るい	-	暗い
しみりした	-	うきうきした	迫力のある	-	静かな
穏やかな	-	激しい	陰気な	-	陽気な
のびやかな	-	抑えたような	速い	-	遅い
優雅な	-	荒々しい	寂しい	-	賑やかな
慌しい	-	のんびりとした	重厚な	-	軽快な
安らぐ	-	緊張した	華やかな	-	素朴な

#### 2.3 プロフィールによるクラスタの選択

楽曲を評価した人物であれば、評価値を用いてクラスタを選択することができるが、この方法では評価していない人物に対してはクラスタの選択ができない。新たに楽曲に対する印象評価を取得する方法も考えられるが、楽曲の印象評価は時間を要する作業であるため、利用者の負担となる。そこで、本研究では個人の趣味嗜好が楽曲に印象に影響を与えると仮定し、比較的に入力が容易である年齢や性別のようなプロフィール情報を用いてクラスタの選択を行う。

本研究で用いるプロフィールは全て yes か no の2択であるとする。クラスタ  $c$  において、プロフィール  $p_i$  に対する回答が yes であった場合の重み  $w_{p_i=yes}^{(c)}$  を式(1)のように計算する。

$$w_{p_i=yes}^{(c)} = \frac{n_{p_i=yes}^{(c)}}{n_{p_i=yes}^{(all)}} - \frac{n^{(c)}}{n^{(all)}} \quad (1)$$

これを各クラスタに対し計算し、また no の場合についても同様に計算する。ただし、上式で  $n_{p_i=yes}^{(c)}$  はクラスタ  $c$  内でプロフィール  $p_i$  に yes と回答した人数、 $n_{p_i=yes}^{(all)}$  は全体で  $p_i$  に yes と回答した人数を指し、 $n^{(c)}$  と  $n^{(all)}$  はそれぞれクラスタ内の人数と全体の人数を指す。

$w^{(c)}$  と入力されたプロフィールからクラス  $c$  に対する適合度  $f^{(c)}$  を式 (2) のように算出し、この値が最も高いクラスを選択する。

$$f^{(c)} = \sum_{i=1}^M w_{p_i=ans}^{(c)} \quad (2)$$

ただし、 $M$  はプロフィールの総数であり、 $ans$  は人物のプロフィールの回答である。

### 3 評価実験

#### 3.1 実験条件

対象楽曲として、RWC 研究用音楽データベース (DB) のクラシック音楽及びジャズ音楽から 15sec の区間を 200 サンプル抽出したものを使用した。これらの楽曲の音響特徴としては、MIRToolbox[3] を用いて抽出した強弱、リズム、音色、調性の 60 次元の特徴と音楽ゆらぎ特徴 [4] の 66 次元の特徴を合わせた 126 次元から、変数増減法で選択された特徴を使用した。

印象特徴は前述したように SD 法に基づいた聴取実験から得られた値を使用した。聴取実験は 118 名で行い、その中で個人別に学習した重回帰モデルによる推定精度が一定以上出会った被験者 93 名を評価実験に用いた。また、プロフィール情報についても各被験者からアンケート形式で取得した。実験で使用したプロフィールは表 3 の 15 種類である。

未知楽曲に対する印象推定精度を評価するために 10-fold のクロスバリデーションを行った。したがって、学習用として 180 サンプル、評価用として 20 サンプルとなる。印象推定精度の計算方法として、ある印象入力に対して 1 曲を選択する方法では、類似した印象の楽曲が存在する場合に適していないと考え、聴取実験での印象評価と推定された印象の間の類似度をケンドールの順位相関係数が高い順に順位付けを行い、未知の 20 サンプルの中で上位 5 以内に入ったサンプルを正解とした。したがって、ランダムに推定した場合には 25% の推定精度となる。

実験の手順として、最初に被験者にクローズの条件で、印象評価について順序関係を考慮した方法と考慮しない方法の比較を行う。続いてクローズの条件で精度が高かった方法を用いて、被験者にオープンでプロフィールによってクラスを選択する手法の評価を行う。

表 2: プロフィール (括弧内の数値は質問数や段階数)

楽曲を聴く場面の広さ (5)	楽曲のジャンルの好み (12)
クラシック音楽への親しみ (4)	ジャズ音楽への親しみ (4)
楽器経験 (2)	歌唱経験 (2)
CD の所有枚数 (7)	オーディオ機器への拘り (1)
アウトドア派・インドア派 (1)	ネガティブ・ポジティブ (1)
大雑把・几帳面 (1)	消極的・積極的 (1)
嫌なことを先にやるか後にやるか (1)	性別 (1)
年代 (3)	

#### 3.2 実験結果

以上の条件で、被験者にクローズとして自身が所属するクラスを選択する方法のクラス数 1~10 の実験結果を図 2 に示す。図中において、実線は順序関係を考慮した提案手法であり、点線は順序関係を考慮していない方法の結果である。この結果から、順序関係を考慮した方法が考慮していない方法に比べて高精度な推定が可能であることが分かり、順序関係を考慮することの有効性が示された。また、被験者全体で学習する方法であるクラス数 1 の条件よりもクラス数を増加させた方が精度が高くなることから、クラスリングが印象傾向の分類に有効であることが示された。

続いて、被験者にオープンとしてプロフィールによってクラスを選択した結果を図 3 の実線に示す。ただし、クラスの選択にあたっては、一定の人数がいない場合には、プロフィールの重みの算出に過学習が発生すると考え、15 名以上のクラスの中から選択するという条件を加えた。この結果で、個人性を考慮していないクラス数 1 の精度と比べて

プロフィールを用いた方法は低下傾向であり、期待した精度は得られなかった。クラス数が 10 の条件では精度の向上を確認できたが、92 名を 10 個のクラスに分割しており、15 名以上のクラス数が限られ、必ずしも有効なクラスが選択されたとはいえない。このような結果となった原因として、選択するクラスに 15 名以上という制約を持たせたがプロフィールの重みを計算するには人数が不十分であったため、必ずしも有効なクラスを選択できなかったためと考えられる。しかしながら、乱数でクラスを選択した 10 回の平均精度 (点線) と比較すると、プロフィールを用いた方法は全体的に高めの精度となっており、プロフィールを用いたクラスを選択方法に一定の有効性が示された。

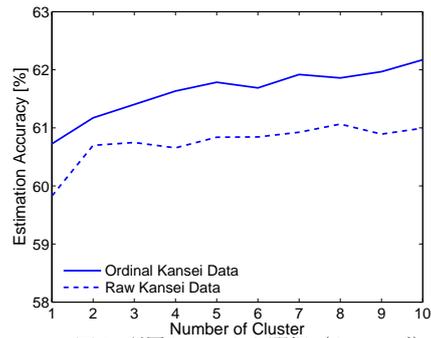


図 2: 所属クラスを選択 (クローズ)

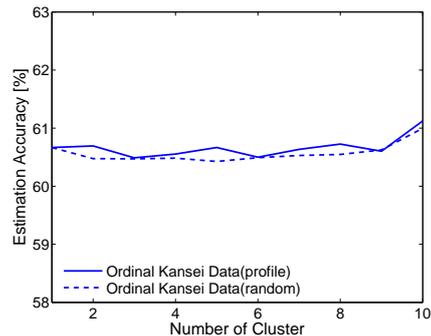


図 3: プロフィールによるクラスを選択 (オープン)

### 4 むすび

本研究では、印象の個人性と順序関係、それぞれを考慮した印象推定を行った。実験の結果から、人間の印象評価を順序関係で正規化することが印象推定の精度の向上に繋がることが明らかになった。これは人間の印象評価が、個人毎に異なる基準との比較によって行われており、この基準が提案手法の正規化によって除去できたためと考えられる。一方で、プロフィールを用いたクラスを選択による個人性の考慮は、ランダムに選択するよりも高い精度となったものの、大きな精度の向上には繋がらなかった。今後の課題として、印象の傾向に関与するプロフィールの選択、主観評価実験がある。

#### 参考文献

- [1] 杉原 太郎, 森本 一成, 黒川 隆夫, "m - R I K : 個人の感性特性に対応可能な音楽検索システム," 情報処理学会誌 46 巻 7 号, pp.1560-1570, 2009.
- [2] Y.-H. Yang, H.-H. Chen, "Music Emotion Ranking," In Proc. ICASSP, pp.1657-1660, 2009.
- [3] Lartillot O., Taivaiaine P., "MIR in Matlab(II): A Toolbox for Musical Feature Extraction form Audio," In Proc. IS-MIR, pp.237-244, 2007.
- [4] 伊藤 雄哉, 山西 良典, 加藤 昇平, "音楽ゆらぎ特徴を用いた楽曲印象の推定," 日本音響学会誌 68 巻 1 号, pp.11-18, 2011.