

柔軟な楽曲検索のための時間変化を考慮した楽曲の印象推定

寺島 大樹 酒向 慎司 北村 正

名古屋工業大学大学院工学研究科

1 はじめに

近年、キーワードを用いた楽曲検索の他にもさまざまな検索手法が研究されている。そのひとつの手法として印象に基づく検索があり、具体的なタイトルなどがわからない時に有効である。

従来の印象に基づく検索に関する研究は、楽曲の印象が時間によって変化することを考慮していないという問題があった。しかし、音楽は時間とともに印象が変化していくと考えられ、また過去の印象の影響を受けると考えられる。そのためこれらを考慮した楽曲検索の手法を検討する。本研究では印象の時間変化を印象値時系列として表す。楽曲を細かく分けたフレームごとに印象を推定することで印象時系列を生成し、DP マッチングを用いることで印象の時間変化と過去の印象による影響を考慮した楽曲検索を行う。

2 印象空間を用いた印象推定

本研究では印象を表現する方法として2次元の空間を設定し、その座標を印象値として扱う方法を用いる。以下に印象空間の定義をし、楽曲の音響信号から定義した空間の座標へ写像する重回帰式の推定法について述べる。

2.1 印象空間の設定

本研究では先行研究 [1] の主観評価データベースと印象推定のためのデータベース [2] の2つを用いる。[2] は2次元の印象空間を利用しているが、[1] は14対の感情語をSD法の7段階で印象を評価させたものである。[1] の評価結果を用いて14対の感情語を因子分析し、2次元空間へ変換した。その結果第1軸に明るさ、第2軸に激しさを表す2次元の印象空間が得られたが、[2] で利用されている印象空間であるVA空間 [3] と近いものであった。

2.2 楽曲の印象空間への写像

本研究では楽曲の推定に音響信号を用いるが、音響信号から直接印象を推定することは難しい。そのため、印象に関係していると考えられる音量、和音や調性などの音響特徴量を抽出し、推定に用いる。評価実験データベースで使われた楽曲からMIRtoolbox [4] を用いて13種類74次元の特徴量を抽出した。これを説明変数とし、評価実験結果の各軸の項目の評価値を目的変数として重回帰分析を行った。特徴量の選択には変数増減法を使用した。

3 印象値時系列の生成と検索

印象値時系列を生成するために、楽曲を4秒間の長さのオーバーラップ0.5秒のフレームごとに区切り、各フレームにおいて2節の方法で印象推定を行った。検索時には求めた印象値時系列を連続DPマッチングを用いて比較し、累計距離の少ない区間を同じ印象変化として出力した。

4 主観評価実験

4.1 実験条件

提案手法の有効性を検証するため主観評価実験を行った。表1に実験条件を示す。「印象の変化の有無」「学習データが過去の印象を考慮したかどうか」の組み合わせ4パターンで提案手法と既存手法の比較を行った。データベース [2] は被験者が楽曲を聴きながらリアルタイムで印象値時系列を記録した、過去の印象を考慮したといえる印象データが用意されている。逆に [1] で用意されているデータは、被験者が楽曲を一通り聴いてから1つ

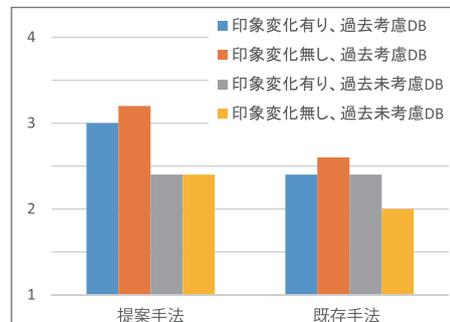


図1 実験結果

の印象値を設定したものである。この2つのデータベースを用いることで、印象推定の学習に過去の印象の考慮が必要かどうかを検討する。また、印象の有無については事前に検索対象の楽曲から印象時系列を求め、その変化が大きいものと小さいもの各30曲からランダムに5曲ずつ選択した。既存手法は楽曲全体の音響特徴量の平均を用いる方法である。入力楽曲1つに対して提案手法と既存手法それぞれで検索した2曲を提示し、各手法による検索結果が入力楽曲の印象とどれだけ合っているかを4段階で評価させた。なお、楽曲の印象は個人差が大きいため、被験者は1人とした。

4.2 実験結果

図2に主観評価実験の結果を示す。全体的に提案手法がよりよい結果であるが、過去の印象を考慮していないデータベースで学習したものは考慮したものに比べて評価が落ちている。提案手法による印象の時間変化を考慮した検索は一定の効果が出るが、印象の推定には過去の影響を考慮したデータベースを用いる必要があると考えられる。

5 まとめ

本研究では、楽曲は印象が時間経過で変化するものとし、印象の時間変化を考慮するために連続DPマッチングを用いた検索を行う方法を提案した。主観評価実験により提案手法の有効性が示された。

参考文献

- [1] 岩月ら：HCGシンポジウム2012論文集，pp.140-146，2012。
- [2] Mohammad Soleymani, et al. : Proceedings of the 2nd ACM international workshop on Crowdsourcing for multimedia, pp.1-6, 2013.
- [3] Robert E. Thayer : Oxford University Press, Oxford, 1989.
- [4] Olivier Lartillot, et al. : International Conference on Music Information Retrieval Vienna, pp.237-244, 2007.

表1 実験条件

楽曲数	[2] より 744 曲, [1] より 61 曲
質問数	20 (各パターン 5 つずつ)
質問内容	提案手法の検索結果は入力と印象が似ているか 既存手法の検索結果は入力と印象が似ているか
人数	1人