

メディア検索

歌唱

能動的音楽鑑賞

体験デモ

似た歌声の曲を探す

藤原 弘将 後藤 真孝

(後藤グループ：産業技術総合研究所)

概要 ボーカルの声質の類似度に基づく楽曲検索システムを開発した。本システムは、クエリとして与えられた楽曲と類似した声質を持つ楽曲を予め登録したデータベース中から検索する。本システムを実現するために、伴奏を含む音響信号中から伴奏音の影響を低減させ歌声の特性を表現する特徴ベクトルを抽出する手法と、相互情報量を用いて2つの特徴ベクトル列間の類似度を計算する手法を開発した。本システムを実装し、約2000曲をデータベースに登録し運用することで、システムが正しく動作することを確認した。

キーワード：歌声，声質，音楽情報検索，音楽類似検索

1. はじめに

本研究では、ユーザによって提示されたクエリの楽曲と、ボーカルの声質が類似した曲をデータベース中から検索出来る音楽情報検索システム(VocalFinder)[1,2]について述べる。既知の曲しか検索出来ない従来の書誌情報による検索に加えて、本システムを使用することで、楽曲の内容を用いて未知の楽曲を検索することが可能になる。

近年、このような内容に基づく音楽検索が重要性を増してきている。携帯型音楽プレーヤや Web 上でのオンライン楽曲販売サービスの普及により、ユーザ達は膨大な量の楽曲から好みの楽曲を検索し、どんな時でもどんな場所でも聴きたい音楽を鑑賞出来るようになった。これにより、自分の好みの楽曲を手がかりに、まだ聴いたことのない楽曲を発見したいという要求が生じてきた。しかし、探したい対象の楽曲が未知で、「好み」のような曖昧な情報のみが手がかりの場合、従来の書誌情報に基づく検索システムは無効だった。

そこで我々は、ボーカルの声質を手がかりに楽曲を検索可能な音楽情報検索システムを開発した。本研究では、ユーザはボーカルの声質というより具体的な内容に基づいて、好みの楽曲を検索することができる。従来の内容に基づく音楽検索の技術では、特徴量として MFCC などを使用していたため、楽曲全体のスペクトル形状など、大まかな情報に基づいて楽曲を検索することしかできなかった。

2. 実現方法

本システムは、データベースに予め登録された楽曲の中から、ユーザが提示したクエリの楽曲と類似した声質を持つ楽曲を検索し、ユーザに提示する。本システムの概要を図 1 に示す。本システムは、データベース構築段階(楽曲の音響信号の分析)と実行段階(楽曲の検索)の2段階からなる。データベース構築段階では、検索対象の楽曲はダウンロードされた CD からリッピングされた後データベースに貯蓄される。そして、システムは各



図1 VocalFinderの概要

楽曲を分析し、ボーカルの声質を表現する特徴量を抽出する。実行段階では、ユーザが好みの楽曲をクエリとしてシステムに入力すると、システムはクエリの楽曲を分析し、楽曲のボーカルの声質を表現する特徴量を抽出する。そして、システムはクエリの楽曲と、データベース中のそれぞれの楽曲で、ボーカルの声質の類似度を計算し、高い類似度を持つ楽曲を検索結果として出力する。

このような音楽情報検索システムを実現するためには、ボーカルの声質を表す特徴量の抽出手法と、その特徴量同士の類似度の計算手法を開発する必要がある。以下に各手法の概要について述べる。詳細は、文献[1,2]に記載されている。

2.1 特徴抽出

ボーカルの声質を表現する特徴ベクトルを計算するために、我々が以前提案した歌手名同定手法に

おける特徴抽出法を使用する。この手法は、楽曲中に混在する伴奏音の影響を低減することが出来る。そのため、伴奏音と歌声が混在した状態をそのまま表現してしまう MFCC などの特徴量と比較して、この特徴ベクトルは歌声の声質をよりよく表現することが出来る。本特徴抽出手法は、伴奏音抑制、特徴抽出、高信頼度フレーム選択の3つの処理からなる。伴奏音抑制とは、伴奏音の影響を低減させるために、歌声の高調波構造に基づき歌声を分離・再合成する手法である。特徴抽出では、分離された歌声から歌手性を表現するスペクトル包絡に関する特徴量を計算する。高信頼度フレーム選択では、抽出された特徴ベクトル列中から、歌声と信頼できる区間(フレーム)を選択し、歌声を含まない区間や伴奏音の影響を大きく受けている区間を除去する。

2.2 類似度計算

本研究では、2つの楽曲間の類似度尺度として相互情報量を使用する。まず、楽曲の特徴ベクトルの分布を GMM によりモデル化し、楽曲から抽出された特徴ベクトル列から、EM アルゴリズムを用いて各楽曲に対する GMM のパラメータを推定する。その後、それぞれの楽曲を表現する GMM の分布同士の距離を、相互情報量を用いて計算する。

3. システムの運用

図2は、システムのスクリーンショットである。システムを実装し、システムの検索対象データベース中に、「RWC 研究用音楽データベース：ポピュラー音楽」(RWC-MDB-P-2001)[3]から選んだ75曲を登録した場合と2000年4月から2007年7月までのオリコン CD シングル週間ランキングの上位10曲(合計約2700曲)を登録した場合の2通りでシステムを運用した。図2の運用例では、RWC-MDB-P-2001中の女性歌手「緒方智美」が歌う「PROLOGUE」(No.7)がクエリとして入力されている。ユーザがクエリの楽曲が提示した後、検索ボタンを押すと、検索結果としてデータベース中の楽曲が類似度の高い順に並べられて出力される。図2のように、検索結果として順位、楽曲名、歌手名、類似度の順に表示されている。ユーザが検索結果リスト中の楽曲をダブルクリックすると、楽曲再生ウィンドウにその楽曲が読み込まれ、再生が始まる。

実際の検索結果では、ほとんどの場合で上位10曲に選ばれた楽曲はクエリの楽曲とボーカルの声質が似ていると感じられた。また、大人数によって歌われる楽曲をクエリとした場合、検索結果の上位の曲は同様に大人数によって歌われる楽曲となることが多かった。一方、検索結果の下位にランクされた楽曲については、順位は楽曲の類似度を正しく反映していないように思われた。しかし、通常の検索用途では上位の10曲程度を提示すれば十分であり、問題にはならない。また、歌声に深くエフェクトがかけられている曲などでは、検索結果が直感的な類似度を反映していない場合もあった。

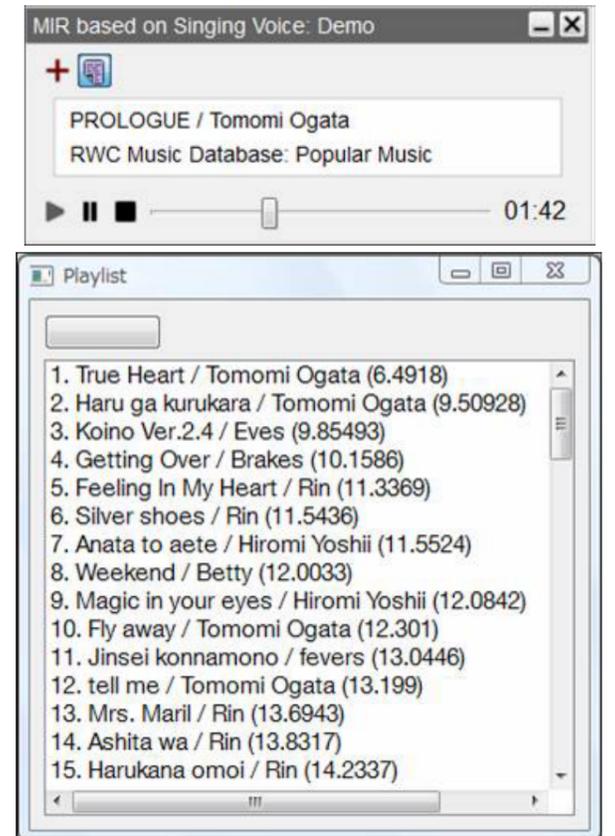


図1 スクリーンショット

4. まとめ

本稿では、ボーカルの声質に基づく類似楽曲検索システムについて述べた。さらに、本システムを実現するために必要な、多重奏の音響信号中から声質に関する特徴量を抽出する手法と、2つの楽曲間の声質の類似度を計算する手法について述べた。本システムを約2700曲を用いて実際に運用し、声質の類似度に基づいて楽曲を検索するのに有効であることを確かめた。

参考文献

- [1] Hiromasa Fujihara, Masataka Goto: A Music Information Retrieval System based on Singing Voice Timbre, In Proc. ISMIR2007, pp. 467-470, (2007).
- [2] 藤原弘将, 後藤真孝: "VocalFinder: 声質の類似度に基づく楽曲検索システム", 音楽情報科学研究会, 2007-MUS-71, Vol.2007, No.81, pp.119-126, 情報処理学会,(2007).
- [3] 後藤真孝, 橋口博樹, 西村拓一, 岡隆一: "RWC 研究用音楽データベース", 情報処理学会論文誌, Vol.45, No.3, pp.728-738, (2004).