

楽器音イコライザによる楽曲音響特徴変動と類似楽曲検索への応用

糸山 克寿^{†*}後藤 真孝[‡]駒谷 和範[†]尾形 哲也[†]奥乃 博[‡][†]京都大学 大学院情報学専攻

* 日本学術振興会 特別研究員 (DC1)

[‡]産業技術総合研究所

1. はじめに

例示に基づく類似楽曲検索 [1] とは、ユーザが指定した楽曲をクエリとして与え、楽曲を相互の類似性に基づいてランキングする検索手法である。類似楽曲検索ではユーザは事前にクエリとなる楽曲を準備する必要がある。また、検索結果に不満がある場合、よりよい検索結果を得るためにはユーザはクエリとなる他の楽曲を探す必要がある。たとえば、検索された楽曲の歌声やドラムスの音量が大きすぎるとユーザが感じた場合、クエリとした楽曲に雰囲気や音色などの特徴が類似しておりかつ歌声やドラムスの音量がより小さい楽曲を探す必要がある。このような条件を満たす楽曲を見つけ出すのは堂々巡りであり、直接的な検索手法ではない。

この堂々巡りを解消するため、楽曲のリミックス (楽器パートの音量操作) で類似楽曲検索におけるクエリを作成する手法 [2] を用いる。楽曲をリミックスすると、楽曲から抽出される音響特徴の分布が変化し、その結果として検索結果が変化する。ユーザはより好みに近い検索結果を得るため、オリジナルの楽曲とは異なるミックスバランスのもとで合成された新たなクエリを生成し、検索を行う。たとえば、歌声やドラムスの音量を下げたクエリを生成することで、前述の問題は解決される。

ミックスダウン前の音源に対してある楽器パートの音量を操作した場合と分離音に対して音量を操作した場合とに起こる音響特徴変動が同じ傾向を示していれば、分離音に対してその楽器パートの音量を操作することは類似楽曲検索において好みの楽曲を検索する良い手助けになる。一方、音響特徴変動の傾向が異なるとその楽器パートの音量操作は楽曲検索の手助けとして使うことは難しい。分離された楽器パートのうち、楽曲検索の良い手助けとなる楽器パートについて述べる。

2. 楽器音イコライザによる楽曲音響特徴変動

楽器音イコライザは、ユーザが多重奏音楽音響信号をリミックスすることを可能にするオーディオプレーヤーである。多重奏音楽音響信号をリミックスするためには音響信号を各楽器パートごとに分離する必要がある。本稿では、音源分離問題を入力混合音のパワースペクトルを楽曲中の各単音に対応するパワースペクトルへ分解することと定義する。このような分解を行うために、各単音のパワースペクトルを近似的に表現するモデルとして、調波・非調波統合モデル [3] を用いる。このモデルは、調波的な音のパワースペクトルを表現する調波構造モデルと非調波的な音のパワースペクトルを表現する非調波構造モデルとの和で定義される。分離処理の詳細については、紙面の制約上省略する。

Acoustic Feature Variation and Application to Similarity-based Music Retrieval using Instrument Equalizer: Katsutoshi Itoyama (Kyoto Univ.), Masataka Goto (AIST), Kazunori Komatani, Tetsuya Ogata, and Hiroshi G. Okuno (Kyoto Univ.)

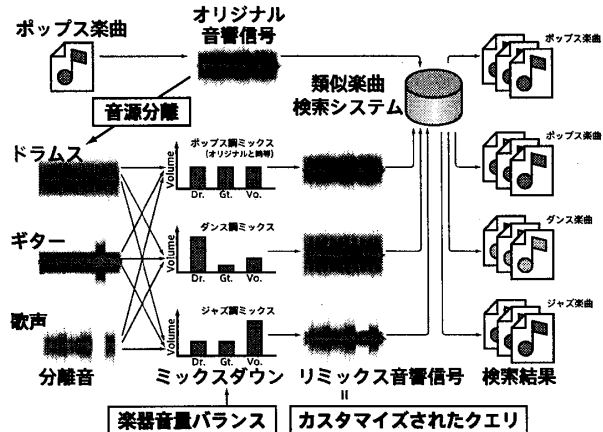
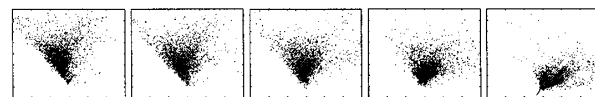


図1: 音響特徴変動を用いた類似楽曲検索の概要。楽器音イコライザで楽器をリミックスすることで各楽器音が存在する区間から抽出される音響特徴が変化し、楽曲全体の音響特徴の分布が変化する。単一の楽曲から様々な音響特徴分布を生成できるため、多様な検索結果を得ることができる。



(a) $-\infty$ dB (b) -5 dB (c) ± 0 dB (d) $+5$ dB (e) $+\infty$ dB

図2: RWC 音楽データベース：ポピュラー音楽、No. 1 を楽器パート毎に分離し、ドラムスパートの音量を変化させた音響信号から音響特徴量を抽出し、その特徴量の第1・第2主成分をそれぞれ横・縦軸としてプロットしたもの。音量変化に伴って、特徴量の分布が変化している。

楽器音イコライザで楽器音量バランスを操作することで、リミックス後の混合音から抽出した音響特徴は元の楽曲から抽出した音響特徴とは異なるものとなる。リミックスで音響特徴の分布が変化する様子を図2に示す。

楽曲の雰囲気は楽曲を構成する楽器の編成やその音量バランスに影響される。楽曲のリミックスによって音響特徴の変動とともに楽曲の雰囲気も変化する。例えば、ポップス楽曲の歌声を増幅させ、ギターやドラムスを減衰させると楽曲の雰囲気および音響特徴の分布はジャズやバラードに近づく。このようにリミックスされた楽曲を類似楽曲検索のクエリとすることで、図1に示すように、元の楽曲の音響特徴を継承しつつ、かつ雰囲気の異なる (上記の例ではジャズやバラードに近い) 楽曲を検索することが可能になる。

3. 類似楽曲検索システムの実装

音楽における雰囲気を表現するため、音量、スペクトル重心などからなる合計 33 次元の特徴量 [2] を設計した。特徴量はパワースペクトルからフレーム毎に抽出する。

楽曲ごとの全体的な雰囲気を表現するために、楽曲毎に特徴量を混合正規分布でモデル化する。本稿の実験で

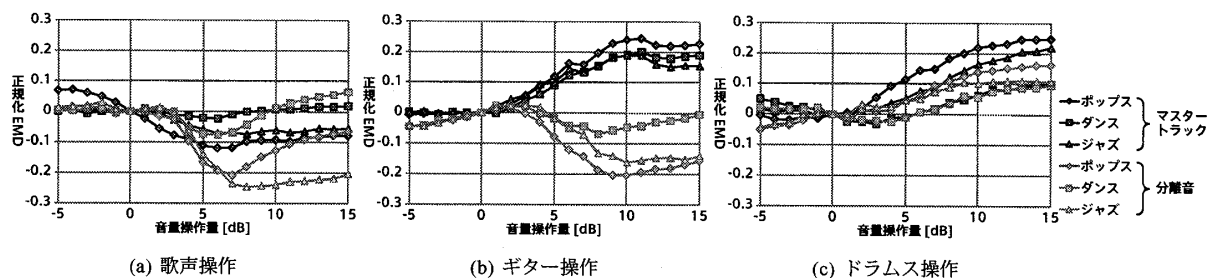


図 3: 楽器パートの音量を操作してジャンル毎に平均 EMD を計算し、音量操作量が 0 dB の際の EMD が 0 になるように正規化したもの。左から順に、歌声、ギター、ドラムスパートの音量を操作した場合。EMD が小さい (グラフの下方にプロットされている) ほどそのジャンルとクエリ楽曲との平均類似度は大きいことを表す。

は、混合数を 8 とした。各楽曲内での雰囲気的时间的な変動は扱わない。こうして得られた混合正規分布間の距離の逆数を、楽曲間の類似度として用いる。距離には Earth Movers Distance (EMD) [4] を用いる。

4. 実験

楽曲のリミックスによる音響特徴の変化について、第 3 章で述べた類似楽曲検索手法を基に調査した。

4.1 実験条件

クエリとなる楽曲には RWC 音楽データベース: ポピュラー音楽 [5], No. 1-10 の 10 楽曲を用いた。これらの楽曲と、RWC 音楽データベースへの AIST アノテーション [6] に含まれる標準 MIDI ファイルを用いて前述の音源分離手法で楽器パートごとの音源分離を行った。また、AIST アノテーションに含まれる各楽器パートをミックスダウンする前の音源 (マスタートラック) もリミックスの対象とし、音源分離性能が楽曲の音響特徴の変化に与える影響を比較する。検索対象となるデータベースに幅広い音響特徴を持つ楽曲を含めるため、同データベース: 音楽ジャンルからジャンルの大分類がポップス、ダンス、ジャズである合計 30 楽曲を選択した。

本実験では、歌声、ギター、ドラムスの 3 楽器パートの音量を操作した。楽器パートの音量操作で楽曲のジャンルを変化させるためには、操作対象の楽器パートは楽曲中での演奏時間が十分に長くなければならない。そこで、クエリとなる 10 楽曲の全てで演奏されており、各楽曲中の 60% 以上の区間で演奏されている、上記の 3 パートを選択した。歌声、ギター、ドラムスの平均音源分離性能 (SNR) は、それぞれ 2.7, -1.77, 0.53 dB であった。

これらの楽器パートの音量を -5 から +15 dB の範囲で操作し、リミックスされた音響信号をクエリとしてデータベース中の楽曲の特徴量分布との EMD を計算した。各楽器パート、各ジャンルの平均 EMD を求め、音量操作量が 0 dB の際の EMD が 0 になるように正規化したものを図 3 に示す。クエリ楽曲は全てポピュラー音楽であるため、クエリ楽曲とポップス以外のジャンルの楽曲との EMD はポップスに比べて平均的に大きくなる。本実験ではクエリ楽曲の音響特徴の変化に着目するため、各ジャンルとの距離そのものよりも音量操作に応じた EMD の変化が重要であり、このような正規化を行った。

4.2 考察

ドラムスの音量を操作した場合の EMD の変化量のグラフ (図 3(c)) は、いずれのジャンルにおいてもマスタートラックと分離音のそれぞれの EMD 比率がある程度類似した曲線を描いており、音量操作による音響特徴の変化も類似していると考えられる。一方、ギターの音量を操作した場合 (図 3(b)) はいずれのジャンルにおいてもマスタートラックと分離音とで大きく異なる曲線を描いている。これは、マスタートラックのギターの音量を操作した場合と分離音のギターの音量を操作した場合とで異なる音響特徴の変化が起きている。ギターの平均音源分離性能は -1.77 dB で 3 つの楽器パート中では最も分離性能が低く、不完全な分離音を楽器パート音量操作に用いたことが原因と考えられる。歌声の音量を操作した場合 (図 3(a)) は、ポップスおよびダンスに関してはマスタートラックと分離音とである程度類似した曲線を描いているが、ジャズに関しては曲線が大きく異なっており、検索対象の楽曲の音響特徴によって同じ分離性能であっても影響の大きさが異なる場合が確認された。

5. おわりに

本稿では、楽曲中の楽器パートの音量操作と楽曲の音響特徴変動の関係を楽曲の雰囲気に基づく類似楽曲検索の観点から議論し、楽器パートの分離性能の違いが雰囲気の変化に影響することを確認した。今後の課題として、被験者実験などを通じた本類似楽曲検索システムの評価などが挙げられる。

謝辞 本研究の一部は、科研費、グローバル COE, Crest-Muse の支援を受けた。

参考文献

- [1] A. Rauber *et al.*: Using Psycho-acoustic Models and Self-organizing Maps to Create a Hierarchical Structuring of Music by Sound Similarity, ISMIR, 2002.
- [2] K. Itoyama *et al.*: Query-by-Example Music Retrieval Approach Based on Musical Genre Shift by Changing Instrument Volume, DAFx, 2009.
- [3] 糸山他: 楽譜情報を援用した多重奏音楽音響信号の音源分離と調波・非調波統合モデルの制約付パラメータ推定の同時実現, 情報処理, Vol. 49, No. 3, 2008.
- [4] Y. Rubner *et al.*: A Metric for Distributions with Applications to Image Databases, ICCV, 1998.
- [5] 後藤他: RWC 研究用音楽データベース: 研究目的で利用可能な著作権処理済み楽曲・楽器音データベース, 情報処理, Vol. 45, No. 3, 2004.
- [6] M. Goto: AIST Annotation for the RWC Music Database, ISMIR, 2006.