

楽曲の曖昧型内容検索における 利用者の関連性フィードバックを用いた特徴量選択に関する研究

坂井 俊亮[†]亀山 啓輔[‡][†]筑波大学第三学群情報学類[‡]筑波大学システム情報工学研究科

1. はじめに

近年, ある楽曲データに類似した楽曲コンテンツを検索したいという要求を実現するため, 楽曲の内容検索に関する研究が進められている[1, 2, 3, 4].

本稿では, 内容に基づく (content-based) 楽曲検索を対象とし, さらに関連性フィードバックに基づく特徴空間の非線形写像を行うことで, よりユーザに適応的な楽曲検索システムの実現のために有効な手法を提案する.

2. 従来手法とその問題点

2.1 内容に基づく音楽情報検索技術

近年研究が盛んに行われている音楽情報内容検索では, 楽曲の音楽音響信号関連の特徴を利用することで, ユーザが検索したい楽曲に関するメタ情報を用いずに, 楽曲の検索が可能である.

2.2 従来手法

これまでに, メル周波数ケプストラム係数 (MFCC), 局所スペクトル特徴[1], 和音進行[2]などの特徴を利用した類似楽曲検索手法が報告されてきた. さらに, ユーザの嗜好に合った楽曲を推奨するシステムとしては, 内容に基づくフィルタリングと協調フィルタリングを連動させる手法[3]や, 感性語評価値を自動付与する手法[4]などの研究が報告されている.

2.3 従来手法の問題点

従来の手法では, 類似楽曲として提供される検索結果がユーザの要求に真に適合した楽曲であることは保証されていない. 感性情報や自然言語などによる類似度評価が可能なシステムの場合でも, ユーザの嗜好が正確に表現されとは限らず, それゆえ, 初期の検索段階においてはユーザの嗜好に合致しない楽曲が推奨される可能性がある.

3. 提案手法

本研究では, 上記のような問題点を解決し, ユーザ嗜好に適応可能な楽曲推奨システムを構築することを目的とする. そのために, ニューラルネットワークによる特徴空間の適応的な写像により, 検索結果に対するユーザの評価情報を有効活用する. ここでニューラルネットワークは, 当初の楽曲特徴空間から, 多様な楽曲ユーザ嗜好空間への非線形写像を可能にしている. また検索処理の高速化を図るための前処理として, データベース中の楽曲データをあらかじめクラスタリングを用いた大分類を行った上で保存しておく. これにより, 初期検索時だけでなく, 再検索時も検索時間の短縮を望むことができる. 提案手法の具体的な手順を図1に示す.

まず大分類システム部では, データベースに登録する楽曲に対して特徴量を抽出後, Bisecting K-means アルゴリズムによるクラスタリングを行い, データベースに保存しておく. ここで抽出した特徴量は表1のとおりである. 次に検索システム部では, 問い合わせ楽曲から同様の特徴量を抽出し, データベース中から最近傍クラスタ内の各データとのユークリッド距離を評価し, 小さいものから順に検索結果として出力する.

提案手法ではさらに, 順位付けされた初期検索結果に対してユーザが嗜好に基づいて順位の並べ替えを行う. そしてその評価を模倣するようニューラルネットワークに学習させ, 同クラスタ内の楽曲を対象として, 非線形写像を行うことで特徴量選択をし, 再度検索を実行する.

表 1. 抽出した楽曲特徴量

特徴量	内容
スペクトル変化度	フレーム間の平均振幅変化量
平均音数	フレーム間で 10[dB]以上変化のあったサンプリング点の時間平均
平均周波数強度	周波数に振幅値を重みとして掛けた値の平均
振幅最大周波数	フレームにおける振幅最大周波数

Feature Set Selection with User Relevance Feedback on Content-based Retrieval of Music

[†] Shunsuke Sakai [‡] Keisuke Kameyama

[†] College of Information Science, University of Tsukuba

[‡] Graduate School of SIE, University of Tsukuba

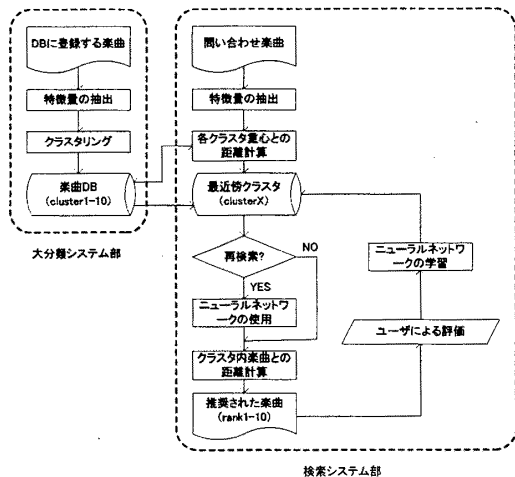


図 1. 提案手法のフローチャート

4. 実験

4.1 実験方法

本実験では以下の 100 楽曲データを用いる。

- ・ 44.1kHz, 16bit の wave フォーマットデータ
- ・ ピアノ独奏曲
- ・ CD から抽出した 10 楽曲各々に関して、ランダムな箇所から約 10sec のクリップを 10 個抽出

検索結果 10 曲に対してここでは”tempo”あるいは”pitch”類似性という仮想的なユーザ重視項目を選択し, ”tempo”の類似性評価は全特徴量中のスペクトル変化度・平均音数(”pitch”であれば平均周波数強度・振幅最大周波数)特徴空間におけるユークリッド距離を用いて行い, 値が小さい順にその順位を学習セットとして利用する。

4.2 実験結果

問い合わせ楽曲に対する初期検索結果(表 2, 4)と, 上記 2 つの重視項目を用いた再検索結果(表 3, 5, 6)を以下に示す。表中 error は検索順位(1~10)に対する平均絶対値誤差, precision は tempo (あるいは pitch) rank (1~10) を正解検索とした場合の検索精度を表し, tempo rank, pitch rank は, 問い合わせ楽曲と同クラスタ内の楽曲における各重視項目に基づく近さ順位を示しており, 括弧内は学習データとして用いる順位である。

表 2. 初期検索結果の評価指標

	tempo	pitch
error	3.20	2.10
precision	0.60	0.80

表 3. 再検索結果の評価指標 (10 試行平均値)

	tempo	pitch
error	0.44	0.60
precision	0.96	1.00

表 4. 初期検索出力

result rank	clip	tempo rank	pitch rank
1	ww.08	2 (2)	2 (2)
2	ww.06	1 (1)	3 (3)
3	ww.09	3 (3)	7 (6)
4	ww.04	15(10)	1 (1)
5	hr.07	5 (4)	11 (9)
6	m3.03	12 (7)	5 (4)
7	cm.03	14 (9)	6 (5)
8	ww.01	6 (5)	12(10)
9	cm.01	10 (6)	9 (7)
10	m3.09	13 (8)	10 (8)

表 5. 再検索出力 (tempo)

result rank	clip	tempo rank
1	ww.06	1
2	ww.08	2
3	ww.09	3
4	ww.03	4
5	hr.07	5
6	ww.07	8
7	ww.01	6
8	hr.10	9
9	hr.03	7
10	cm.01	10

表 6. 再検索出力 (pitch)

result rank	clip	pitch rank
1	ww.04	1
2	ww.10	4
3	ww.08	2
4	ww.06	3
5	m3.03	5
6	cm.03	6
7	ww.09	7
8	cm.01	9
9	ww.02	8
10	m3.09	10

表 5, 6 の再検索の結果から, tempo を重視した場合には, 問い合わせ楽曲にテンポがより近い曲順に出力され, 特に初期結果 10 位以内に存在しなかった類似曲が再検索時に得られているのがわかる。また同様のことが pitch 重視の場合にも当てはまる。

5. まとめ

本稿では, 従来の音楽情報検索システムに, 関連性フィードバックに基づく特徴量の非線形変換を取り入れることで, ユーザ嗜好に適応可能なシステムを構築し, その有効性を検証した。

参考文献

- [1] Li, et al., ”Integrating Features from Different Sources for Music Information Retrieval”, IEEE ICDM’06, 2006.
- [2] 莪山, ”楽音信号からの和音進行抽出手法と類似楽曲検索への応用”, PIONEER R and D, Vol.14, No.2, 2003.
- [3] 黒瀬ら, ”感性情報を用いた楽曲推薦システム”, DEWS2003, 8-P-06, 2003.
- [4] 大塚ら, ”PCM データに対応した感性語による音楽データベース検索システムに関する研究”, DEWS2002, 8-P-05, 2002.