

卒業論文要旨
時系列ウェーブレット変換を用いた楽曲推薦の有効性の検証

(知能情報システム学) 徳山竜也

1. 序論

近年、携帯音楽プレーヤや Web 上でのオンライン楽曲販売サービスの普及により、音楽に関する関心が高まってきている。しかし、膨大な量の楽曲から好みの楽曲を探し出すのは困難であるため、その中からユーザにふさわしい楽曲を推薦する仕組みが求められている。

そこで本稿では、既報^[1]で提案された時系列ウェーブレット変換を用いた楽曲特徴量と、音響特徴量として従来から利用されてきた MFCC(Mel-Frequency Cepstrum Coefficient)との比較を楽曲推薦に対する有効性という観点で行なった。また、ユーザによる楽曲評価から得た嗜好ベクトルと時系列ウェーブレット変換を利用した特徴量、MFCC、楽曲のリズム特性を表す特徴量を利用した楽曲推薦システムを構築し、その性能を検証した。

2. 特徴量抽出

2.1. ウェーブレット変換による特徴量

音響信号をウェーブレット変換によって多重解像度解析した各係数の標準偏差を用いた。まず、窓幅に区切った音声信号 $f(n)$ を、レベル 0 のスケーリング係数 $s_k^{(0)}$ とみなす。次に、レベル j のスケーリング係数 $s_k^{(j)}$ 、およびウェーブレット係数 $w_k^{(j)}$ を $s_k^{(0)}$ から、逐次レベル J まで求める。ここで得られた各係数の標準偏差の時系列データを楽曲の特徴量とした^[1]。

2.2. MFCC による特徴量

フレーム毎に音響信号の高速フーリエ変換(FFT)で得られるパワースペクトルに対してメルスケールのフィルタバンクを施し、周波数軸変換されたパワースペクトルに対して離散コサイン変換(DCT)を実行することにより MFCC 特徴量が抽出される。12 次元 MFCC、1 次元対数パワー、12 次元 Δ MFCC、1 次元 Δ 対数パワーの各平均と標準偏差の時系列データを楽曲の特徴量とした。

2.3. リズム特徴量

まず、ウェーブレット変換の多重解像度解析を利用して分解された周波数バンドに対して前処理を施す。次に、各バンドを逆ウェーブレット変換し、自己相関を計算する。自己相関におけるピークをビートヒストグラムとする。ビートヒストグラムでの次の特徴を、リズム特性を表す特徴量とした^[1]。

- A0, A1 : 1, 2 番目のピークの相対振幅
- P1, P2 : 1, 2 番目のピークの時間
- RA : 振幅の比率 $A1 / A0$
- SUM1, SUM2, SUM3 : BPM (Beats Per Minute) が 40~90, 90~140, 140~250 であるヒストグラムの合計

ここで、予備実験で高い推薦正答率を得ることができた $a:[SUM1]$, $b:[A0, A1, SUM1, SUM2, SUM3]$, $c:[A0, A1, P1, P2, RA, SUM1, SUM2, SUM3]$ を特徴量として使用した。

3. 楽曲推薦システムと評価実験

今回利用した楽曲は RWC 研究用音楽データベース (音楽ジャンル)^[2]の 100 曲で、ユーザ 10 名によって 5 段階の主観評価 (好き : 5 ~ 嫌い : 1) が付与された。

以下の A, B の条件で評価を行なった。

条件 A: 評価既知曲 99 曲に対して評価既知曲 1 曲を未知曲として取り扱い、判定する。これを 100 曲に対して行なう。

条件 B: 評価既知曲 98 曲に対して評価既知曲 1 曲を未知曲として取り扱い、判定する。これを 99 曲に対して行ない、各最良手法での残り 1 曲の推薦精度を判定する。

また、(1), (2)の手法で推薦を行なった。

(1)一種類の特徴量を用いた手法

全評価楽曲と対象未評価楽曲に対する特徴量を主成分分析し、その特徴ベクトルに対して、ユークリッド距離を用いて、類似度が最大となる評価済楽曲を決定する。そのスコアが 4 または 5 であるとき対象未評価楽曲を推薦する。

(2)複数の特徴量を組み合わせた手法

A) 各特徴量で求めたスコアが、一種類が 4 以上で他方が 3 以上であれば推薦する。

B) 一種類の特徴量でスコアが 4 以上であり、かつ他方で類似度最大となった楽曲が、他方において全評価楽曲の中での類似度の順位が上位 50%以内であった場合推薦する。

C) 一種類の特徴量でスコアが 4 以上であり、かつ他方の特徴量で類似度最大となった楽曲が、他方において求めた類似度が全評価楽曲の平均より大きければ推薦する。

*特徴量が三種類するとき、他方とは他の二種類を表す。その場合、上から手法 D, E, F とする。

4. 実験結果と結論

表 1 条件 A によるウェーブレット変換と MFCC の推薦平均正答率(%)

ウェーブレット変換		MFCC		
推薦	ランダム	推薦	ランダム	
正答率	との差	正答率	との差	
平均	46.19	14.19	42.95	10.95

表 3 条件 A, B におけるシステムの推薦平均正答率(%)

条件 A		条件 B		
推薦	ランダム	推薦	ランダム	
正答率	との差	正答率	との差	
平均	67.6	35.6	55.8	23.8

表 2 条件 A でシステムが選択した手法と特徴量

user	手法	特徴量
1	F	MFCC(main)&wavelet&Rhythm a
2	一種類	wavelet
3	E	MFCC(main)&wavelet&Rhythm a
4	E	MFCC(main)&wavelet&Rhythm a
5	C	wavelet(main)&Rhythm a
6	E	MFCC(main)&wavelet&Rhythm b
7	F	Rhythm b(main)&wavelet&MFCC
8	E	wavelet(main)&MFCC&Rhythm c
9	E	wavelet(main)&MFCC&Rhythm c
10	E	wavelet(main)&MFCC&Rhythm a

表 1 よりウェーブレット変換と MFCC を比較すると、ウェーブレット変換を用いた方が推薦精度が高くなった。これにより時系列ウェーブレット変換を用いて得た特徴量は楽曲推薦に有効であることが確認できた。表 2, 3 より三種類の特徴量を組み合わせることで、より精度の高い楽曲推薦システムを構築することができた。また表 3 より条件 B においては平均で約 56%と条件 A に比べて低い正答率となったが、類似度が 2 番目に高い評価済楽曲のスコアも 4 または 5 のときに推薦するという方式をとると、約 68%と高い推薦精度を得ることができた。

参考文献

- [1] 田仲 雅和, “ウェーブレット変換を用いたユーザの嗜好に基づく楽曲推薦システム”, 京都府立大学大学院人間環境科学研究科環境情報学専攻 修士論文, 2008.
- [2] 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: 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”, 情報処理学会論文誌, vol.45, no.3, pp.728-738, 2004.