

ユーザ評価と楽曲ゆらぎ特徴を用いた 楽曲嗜好のハイブリッド推定

ー共有楽曲数の差異が協調フィルタリングに与える影響ー

伊東 孝浩^{1,a)} 加藤 昇平^{1,b)}

概要：近年、音楽配信サービスの発展により個人が多くの楽曲を聴取することが可能となった。しかし、膨大な楽曲の中から嗜好に合った楽曲を選択する事はユーザにとって大きな負担となっている。この問題を解決するため、楽曲特徴を用いたコンテンツベースフィルタリングによる楽曲推薦手法が提案されている。しかし、コンテンツベースフィルタリングのみを用いた手法では、類似した楽曲ばかりが推薦されユーザの嗜好を特定する上での問題点が指摘されている。そこで本稿では、他ユーザの嗜好を反映する事が可能な協調フィルタリングの技術を取り入れた嗜好推定手法を提案する。また、ユーザ間の嗜好や聴取曲の差異による協調フィルタリングの影響についての考察を行う。

1. はじめに

近年、ミュージックプレイヤーの大容量化やインターネットを通じた音楽配信サービスの発展により、個々のユーザが多くの楽曲を聴取することが可能となった。所有する楽曲データベースの中から楽曲を聴取する場合、予め楽曲に付与されているテキスト情報(曲名、アーティスト名、ジャンル名、アルバム名等)を利用することが一般的である。しかし、膨大な楽曲データベースの中から嗜好に合った楽曲を選択することはユーザにとって負担となる。この負担を軽減するため、楽曲をランダムに再生する方法が挙げられるが、ユーザの嗜好に合った楽曲が必ずしも再生されない可能性がある。これらの問題点を解決するため、ユーザの嗜好を捉えた楽曲推薦システムが提案されている[1][2][3]。宇野ら[1]は楽曲がいつ、どこで聴取されたかを記録し、ユーザのライフログと楽曲の関係性について可視化するシステムを提案している。濱崎ら[2]はweb上に点在する音楽コンテンツの中から、ユーザにとって興味深い情報を推薦するための手法を提案している。コンテンツに付与されているカテゴリタグを利用し、コンテンツの関係性を可視化したインタフェースを提案している。清水ら[3]はユーザの知らない情報を特定し、情報の発見性を考慮した推薦手法を提案している。ユーザが知らないであ

る情報を推測し、従来の情報推薦手法に組み合わせることでユーザの満足度を向上させる手法を提案している。

また、多田ら[4]はユーザの嗜好を学習し、嗜好に合った楽曲を推薦する手法を提案している。ユーザの過去の楽曲評価と音響特徴の関係性をクラス判別手法である単純ベイズ分類器により定式化し、ユーザの嗜好を推定している。しかし、音響特徴による学習だけではユーザの嗜好を的確に捉えることが困難である。我々は既存手法の問題点を解決するため、情報フィルタリング手法の一つである協調フィルタリングを導入した楽曲推薦手法を提案し、よりユーザの嗜好を捉えた楽曲推薦の可能性を確認した[5]。本稿ではユーザ間の共通楽曲数の際による協調フィルタリングの影響について検証を行う。

2. 嗜好推定

本稿では、クラス判別手法である単純ベイズ分類器と他ユーザの評価傾向を考慮する協調フィルタリングにより、楽曲に対するユーザの評価値を推定する。

2.1 単純ベイズ分類器

情報の持つ特徴量 $f_1, f_2, f_3, \dots, f_n$ を基に、情報が属するクラス c_j を次の事後確率最大化仮説にてより計算する。

$$C_{MAP} = \arg \max_{c_j \in C} P(c_j) \prod_{i=1}^n P(f_i | c_j) \quad (1)$$

ここで C_{MAP} は分類されるクラス、 C はクラスの集合、 $P(f_i | c_j)$ は特徴量 f_i がクラス c_j に属する条件付き確率を表す。本稿では、ユーザの楽曲に対する評価値をクラスと

¹ 名古屋工業大学

Nagoya Institute of Technology, Dept. of Computer Science and Engineering, Graduate School of Engineering, Gokiso-cho, Showa-ku, Nagoya-si 466-8555, Japan

^{a)} tkitoh@katolab.nitech.ac.jp

^{b)} shohey@katolab.nitech.ac.jp

し, 3章で説明する楽曲特徴を情報の持つ特徴量として推定を行う. 単純ベイズ分類器は特徴量が独立である条件下でクラス分類が可能であるが, 特徴量は離散値でなければならない. 本稿ではガウス分布による確率近似により, 楽曲特徴量 f_i がクラス c_j に属する条件付き確率分布 $P(f_i|c_j)$ を決定している. これらを基にユーザの嗜好に合うか否かを確率値で算出する.

2.2 協調フィルタリング

協調フィルタリングとは, ユーザに情報を推薦する際にそのユーザの興味や嗜好が類似する他ユーザの利用履歴や嗜好傾向を用いて推薦情報を提示する手法である. 現在多くのオンラインショッピングサイト(例えば Amazon[6])でも使用されている方式であり, ユーザがある商品を購入した際に同じ商品を購入したことがある他ユーザの購入履歴から商品を推薦する機能が実装されている. GroupLens[7]の研究では, あるユーザと他ユーザとの類似度を Pearson の相関係数を用いて算出し, その結果を用いてユーザがまだ未評価の情報に対する推定評価値の計算が可能である.

本稿では, 楽曲 i' に対するユーザ A の推定評価値 $\hat{s}_{A,i'}$ を, 2.1 章で算出した推定値 $\tilde{s}_{A,i'}$ と協調フィルタリング手法を用いて次式により算出する.

$$\hat{s}_{A,i'} = \tilde{s}_{A,i'} + \frac{\sum_{B \in B_{i'}} r_{AB} \times (s_{B,i'} - \bar{s}_B)}{\sum_{B \in B_{i'}} |r_{AB}|} \quad (2)$$

ここで $B_{i'}$ は楽曲 i' を評価している他ユーザ集合, r_{AB} はユーザ A とユーザ B の嗜好類似度, $s_{B,i'}$ は楽曲 i' に対するユーザ B の評価値, \bar{s}_B はユーザ B の評価値の平均を表す. r_{AB} は次式により算出する.

$$r_{AB} = \frac{\sum_{i \in I_{AB}} (s_{A,i} - \bar{s}_A)(s_{B,i} - \bar{s}_B)}{\sqrt{\sum_{i \in I_{AB}} (s_{A,i} - \bar{s}_A)^2} \sqrt{\sum_{i \in I_{AB}} (s_{B,i} - \bar{s}_B)^2}} \quad (3)$$

ここで I_{AB} はユーザ A, B が共に評価している楽曲集合を表す.

3. 楽曲特徴

本稿では, 楽曲の持つ楽曲特徴量として音楽ゆらぎ特徴 [8] を用いることとする. 音楽ゆらぎ特徴は, 認知学においてもその重要性が述べられている音楽の時間的概念を考慮した音響特徴 [9] であり, 人が楽曲から受ける印象と大きく関わっていることが確認されている. ユーザの感性に適合する楽曲を推測し推薦する提案システムにおいても有効であると考えられる. 表 1 に提案システムで用いる 66 種の楽曲特徴を示す. 次節からはこれらの楽曲特徴の詳細について述べる.

3.1 音楽ゆらぎ特徴

音楽を構成する 3 大要素としてメロディ, ハーモニー, リズムが存在しており, これらの時間的変化は人の感性評価に影響を与えると考えられる. メロディー及びハーモニーを構成する音量, 音高の時間的変化, またリズムの時間的変化を音楽ゆらぎ特徴として抽出する. 本稿では,

表 1 音楽ゆらぎ特徴

特徴番号	ゆらぎスペクトル特徴量
1 - 6	音量 (all) の特徴量 (a) ~ (f)
7 - 12	音量 (low) の特徴量 (a) ~ (f)
13 - 18	音量 (middle) の特徴量 (a) ~ (f)
19 - 24	音量 (high) の特徴量 (a) ~ (f)
25 - 30	音量 (ultrahigh) の特徴量 (a) ~ (f)
31 - 36	音高 (周波数重心) の特徴量 (a) ~ (f)
37 - 42	音高 (比率) の特徴量 (a) ~ (f)
43 - 48	音高 (回帰直線の傾き) の特徴量 (a) ~ (f)
49 - 54	音高 (回帰直線の切片) の特徴量 (a) ~ (f)
55 - 60	音高 (ピーク周波数) の特徴量 (a) ~ (f)
61 - 66	リズム (ビートスペクトラム) の特徴量 (a) ~ (f)

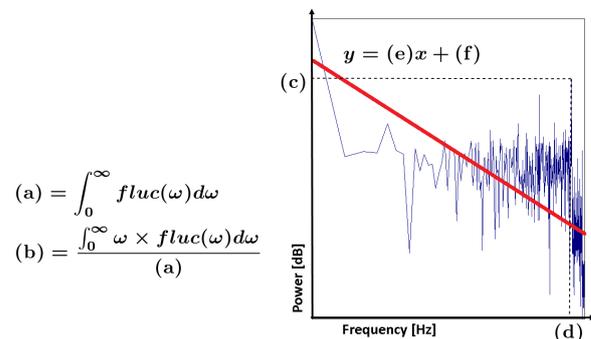


図 1 ゆらぎスペクトルの一例

WAV 形式で保存された楽曲データを扱う. 全ての楽曲データは音質を揃えるために市販されている CD を用いてサンプリング周波数 44100[Hz], 16 ビットリニア量子化, ステレオのデータとして保存した. 音楽ゆらぎ特徴を抽出するため, まず時間分解能 (25[ms]) 毎にシフトしながら窓幅 2048 点の Hanning 窓 (およそ 46[ms]) に対して楽曲データにフーリエ変換を施し, 各時刻における周波数スペクトル $fluc(\omega, t)$ (ω :周波数, t :時刻) を算出する. そして, 算出された $fluc(\omega, t)$ から音量に関する帯域ごとの特徴量 5 種, 音高に関する特徴量 5 種, リズムに関する特徴量 1 種を算出する.

- 音量 all (全周波数帯域) におけるパワースペクトルの積分値
- 音量 low (周波数帯域 200[Hz] 未満) におけるパワースペクトルの積分値
- 音量 middle (周波数帯域 200 以上 800[Hz] 未満) におけるパワースペクトルの積分値
- 音量 high (周波数帯域 800 以上 2000[Hz] 未満) におけるパワースペクトルの積分値
- 音量 ultrahigh (周波数帯域 2000[Hz] 以上) におけるパワースペクトルの積分値
- 音高 周波数重心
- 音高 低周波数成分 (low) の割合
- 音高 回帰直線の傾き
- 音高 回帰直線の切片

表2 第12主成分までの分散と累積寄与率

主成分番号	分散	累積寄与率
第1主成分	6.08	0.34
第2主成分	5.21	0.63
第3主成分	2.34	0.77
第4主成分	1.07	0.83
第5主成分	0.81	0.87
第6主成分	0.34	0.89
第7主成分	0.28	0.91
第8主成分	0.26	0.92
第9主成分	0.21	0.93
第10主成分	0.14	0.94
第11主成分	0.11	0.94
第12主成分	0.09	0.95

表3 実験楽曲60曲への評価割合

	被験者の評価値				
	-2	-1	0	1	2
被験者1	11	20	12	17	0
被験者2	1	18	14	22	5
被験者3	7	21	13	17	2
被験者4	4	19	18	17	2
被験者5	0	6	19	25	10
被験者6	2	8	12	25	13
被験者7	6	12	23	16	3
被験者8	0	16	19	17	8
被験者9	3	16	19	14	8
被験者10	6	19	20	11	4
被験者11	3	9	6	13	29
被験者12	1	5	28	20	6

- 音高 ピーク周波数 (パワースペクトル最大の周波数)
- リズム ビートスペクトラム

また、リズム特徴としては Foote らが提案した各時刻間における周波数特徴のコサイン類似度によってリズムを表現するビートスペクトラム [10] を用いる。

得られたそれぞれの特徴量を時系列化することで、ゆらぎ情報が得られる。これらのゆらぎ情報に対して再度高速フーリエ変換を施すことでゆらぎスペクトルが生成される。生成された各ゆらぎスペクトルにおける以下の6つの特徴量が音楽ゆらぎ特徴となる。

- スペクトル積分値 S
- 周波数重心 f_c
- スペクトル最大値 f_{luc}^{max}
- スペクトル最大の周波数 f_{max}
- スペクトル回帰直線の傾き α
- スペクトル回帰直線の切片 β

楽曲特徴毎に、全楽曲において平均値が0、標準偏差が1となるように正規化を行った。図1にゆらぎスペクトルの一例を示す。

4. 評価実験

提案手法の有効性を確認するために以下の実験、考察を行った。被験者は心身共に健康な20代前半の男女12名とした。

4.1 実験に用いた楽曲

楽曲推薦を行う上で、協調フィルタリングを用いるためには個別に所有する楽曲データベースを事前に準備する必要がある。しかし、協調フィルタリングの推定はユーザ間のデータベース内容に強く依存する。本稿では公平性を考慮してあらかじめ用意した楽曲60曲を用いて実験を行った。実験に用いる楽曲は聴取による被験者への負担を考慮し、一般的にサビと呼ばれる一部を切り取り使用した。一楽曲の長さは30秒に固定した。

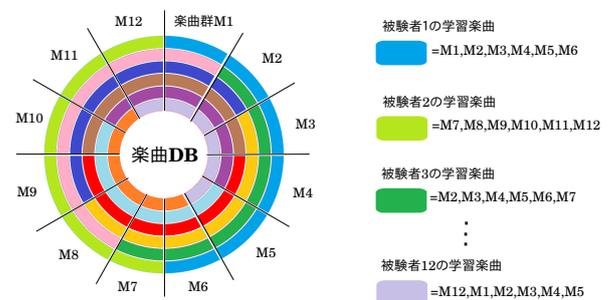


図2 学習楽曲の割当方法 (被験者12名の場合)



図3 共通楽曲数の異なる環境

4.2 実験に用いる特徴量

楽曲の持つ特徴量として3章で述べた楽曲ゆらぎ特徴を用いた。ただし、単純ベイズ分類器で評価値を推定する場合には特徴量の独立性が必要となるため、算出した計66種の特徴量を主成分分析を用いて変換した。ここでは、累積寄与率が95%を超えた第12主成分までを用いた。表2に第12主成分変量までの分散と累積寄与率を示す。

4.3 実験内容

被験者には楽曲60曲を聴取させ楽曲の好みを評価させた。評価は-2~2の5段階(最低評価を-2, 最大評価を2)とした。表3に評価実験の結果を示す。この結果を用い

表4 被験者評価値の推定一致数 (60 曲中の正答数)

	単純ベイズのみの推定		協調フィルタリング補正あり		
	最大確率	期待値	環境 1	環境 2	環境 3
被験者 1	15	8	20	22	22
被験者 2	6	18	18	15	15
被験者 3	18	18	23	23	32
被験者 4	14	21	27	24	24
被験者 5	7	16	16	20	28
被験者 6	16	16	29	37	29
被験者 7	11	14	19	17	16
被験者 8	28	28	26	37	33
被験者 9	8	19	27	28	21
被験者 10	19	26	22	20	22
被験者 11	22	12	24	26	22
被験者 12	15	22	29	31	22
平均	14.9	18.2	23.3	25.0	23.8
分散	41.4	31.8	19.3	51.1	32.3

て、被験者が評価づけを行った 60 曲のうち半数の 30 曲を学習曲とし、残りの 30 曲を未知曲としてシステムの推定性能を検証する。本稿では協調フィルタリングの影響を考察するため、被験者毎に異なる学習曲を設定し楽曲に対する評価値を推定する。図 2 に被験者毎の学習曲の割当を示す。楽曲データベース 60 曲を楽曲群 M1 ~ M12 に均等分割し、各被験者へ図 2 のように学習楽曲を割り当て、環境 1 とした。また、被験者間の学習曲の共通数が環境 1 と異なるような環境 2 と環境 3 を用意した (図 3)。環境 2 に関しては色毎に被験者 2 人、環境 3 に関しては色毎に被験者 3 を割り当てる。本稿では被験者の評価値推定を以下の 5 種類の方法で行い、推定結果に関して考察する。

- (1) 単純ベイズで確率が最大となる評価値 (最大確率)
- (2) 単純ベイズで計算した確率に基づく評価値 (期待値)
- (3) (2) に協調フィルタリングを加えた評価値 (環境 1)
- (4) (2) に協調フィルタリングを加えた評価値 (環境 2)
- (5) (2) に協調フィルタリングを加えた評価値 (環境 3)

表5 単純ベイズの結果の例

楽曲 i'	評価値の推定確率				
	-2	-1	0	1	2
楽曲 i'	0.11	0.07	0.23	0.34	0.25

学習曲 30 曲に付けられた被験者の評価値と楽曲の持つ特徴量を用いて、単純ベイズにより楽曲 i' の推定評価値を計算した場合、表 5 のような評価値毎の適合確率が算出される。この場合、評価値が 1 となる確率 0.34 が最大であるため (1) の推定では評価値 1 を取る。また、確率分布により期待値を計算すると推定評価値が 0.55 となるため (2) の推定では小数第 1 位で四捨五入を行い推定評価値 1 となる。(3)(4)(5) では (2) で算出した値に、それぞれ、環境 1、環境 2、環境 3 における協調フィルタリングの補正をかけた値を推定評価値とした。また、被験者毎に設定する学習

曲の差異によっても結果が異なるため、学習曲を変更して (1) ~ (5) の評価値推定を行った (12 試行)。

4.4 実験結果

表 4 に被験者評価値の推定結果を示す。協調フィルタリング補正を加えた推定法が単純ベイズのみを用いた推定法と比較して正答数が高いことがわかる。これより、楽曲特徴だけでは推定が難しかった嗜好推定が協調フィルタリングの効果により向上したことがわかる。中でも環境 2 に関しては被験者 12 人中 6 人の評価値推定で最大の正答数を示している。協調フィルタリングの環境の差異に関しては環境 2 が最も推定性能が高い結果となった。しかし、被験者全体の推定成功数の分散は環境 2 が最も高く推定性能は安定しない。環境 1 は推定一致数の分散が最も低く、他の環境と比較しての推定性能を安定的に向上させていることがわかる。同環境は単純ベイズのみの推定に比べて 8.4 ポイント (56%) の性能向上を示しており協調フィルタリングの効果を強く表している。

次に事前実験で楽曲評価を分散させて付けていた被験者 9 に着目し推定結果の詳細を調べる。表 6 に 5 種類の推定方法それぞれの推定正誤結果を示す。今回の実験では、楽曲に対する被験者評価を -2 ~ 2 の 5 段階で行っているため、評価値の極性を誤推定する場合、すなわち、正の評価値が与えられた楽曲が負に判別される場合や負の評価値が与えられた楽曲が正に判別される場合は嗜好推定の面で大きな問題となる。単純ベイズのみを用いた推定では、極性が反転して推定された楽曲が確率最大では 10 曲、期待値では 9 曲存在した。協調フィルタリングを用いた推定では、極性が反転して推定された楽曲が環境 1 では 3 曲、環境 2 では 2 曲、環境 3 では 4 曲に、それぞれ抑制された。他の被験者との協調フィルタリングの影響により推定結果を補正する効果が確認された。

表 6 被験者 9 の評価値推定の正誤結果

単純ベイズ (最大確率)

		推定結果				
		-2	-1	0	1	2
被 験 者 評 価	-2	2	0	1	0	0
	-1	4	3	5	2	2
	0	7	2	2	6	2
	1	3	2	4	0	5
	2	2	1	2	2	1

単純ベイズ (期待値)

		推定結果				
		-2	-1	0	1	2
被 験 者 評 価	-2	1	1	0	1	0
	-1	0	1	11	4	0
	0	0	3	14	2	0
	1	0	4	7	3	0
	2	0	0	4	4	0

協調フィルタリング補正あり (環境 1)

		推定結果				
		-2	-1	0	1	2
被 験 者 評 価	-2	2	1	0	0	0
	-1	1	7	5	3	0
	0	0	4	11	4	0
	1	0	0	6	5	3
	2	0	0	0	6	2

協調フィルタリング補正あり (環境 2)

		推定結果				
		-2	-1	0	1	2
被 験 者 評 価	-2	2	1	0	0	0
	-1	1	7	6	1	1
	0	0	4	10	5	0
	1	0	0	6	5	3
	2	0	0	0	4	4

協調フィルタリング補正あり (環境 3)

		推定結果				
		-2	-1	0	1	2
被 験 者 評 価	-2	2	1	0	0	0
	-1	3	4	5	3	1
	0	0	4	5	9	1
	1	0	0	3	7	4
	2	0	0	1	4	3

との共通曲数が嗜好推定に及ぼす影響の差異を確認した。今後は楽曲に含まれる歌詞情報を考慮した類似度計算を取り入れることで嗜好推定性能の向上を目指したい。

参考文献

- [1] 宇野愛, 伊藤貴之: “MALL ライフログを用いた楽曲推薦結果の可視化”, 情報処理学会 第 74 回 全国大会 (2012).
- [2] 濱崎雅弘, 後藤真孝: “Songrium: 多様な関係性に基づく音楽視聴支援サービス”, 情報処理学会 研究報告 (2012).
- [3] 清水拓也, 土方嘉徳, 西田正吾: “発見性を考慮した協調フィルタリングアルゴリズムに関する複数方式の検討”, DEWS2007.
- [4] 多田圭吾, 山西良典, 加藤昇平: “ユーザ感性へのインタラクティブ適応に基づく楽曲推薦システム”, 第 11 回情報科学技術フォーラム, Vol.11, pp.23-29 (2012).
- [5] 伊東孝浩, 加藤昇平: “ユーザ評価と楽曲ゆらぎ特徴を用いた楽曲嗜好のハイブリッド推定”, 第 12 回情報科学技術フォーラム, Vol.12, pp.239-242 (2013).
- [6] LINDEN Greg: “Amazon.com Recommendations: Item-to-Item Collaborative Filtering”, *IEEE Internet Computing* Vol.7 pp.76-80(2003).
- [7] P. Resnick, N. Iacovou, M. suchak, P. Bergstorm, and J. Riedl. GroupLens: “An open architecture for collaborative filtering of Netnews.” *In Proc. Of The Conf. on Computer Supported Cooperative Work* pp.175-186(1994).
- [8] 伊藤雄哉, 山西良典, 加藤昇平, 伊藤英則: “楽曲に対する感性評価と音響ゆらぎ特徴の対応付け”, 感性工学会論文誌, Vol.3, pp.341-348 (2011).
- [9] L. B. Meyer: “Emotion and Meaning in music”, University of chicago Press (1956).
- [10] J. Foote and S. Uchihashi: “The beat spectrum: A new approach to rhythm analysis”, *In Proc. IEEE Int. Conf. on Multimedia and Expo*, pp. 881-884 (2001).

5. まとめ

本稿ではユーザ評価と楽曲ゆらぎ特徴を用いた楽曲嗜好のハイブリッド推定を提案し, 共有楽曲数の差異が協調フィルタリングに与える影響を検証した。提案手法を用いて楽曲に対するユーザの嗜好推定実験を行い, 単純ベイズ分類器によるクラス判別と協調フィルタリングの影響を調べた。実験の結果, 協調フィルタリングの補正により音響特徴だけでは困難だった楽曲の嗜好推定の向上を確認した。また, 協調フィルタリングが働く環境において, 他ユーザ