

## 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

## Music Recommendation System Based on Interactive Adaptation for User's Personality

多田 圭吾† 山西 良典‡ 加藤 昇平†  
Keigo Tada Ryosuke Yamanishi Shohei Kato

## 1 はじめに

今日、生活のあらゆる場面に多種多様なエンターテインメントが存在している。その中でも、特に音楽は古くから世界中の人々に愛されきたエンターテインメントの一つであり、現代においても人々の生活をより豊かにするエンターテインメントコンテンツとして注目されている。

近年ではメディア情報処理技術の発展により大容量メモリ搭載型の携帯音楽再生機器が登場したことで、ユーザは膨大な楽曲データベースを保持・携帯することが可能となった。このような個人が所有する楽曲データベースからの選曲方法としては、あらかじめ楽曲に付与されたテキスト情報(曲名、アーティスト名等)を用いる選曲方法が一般的である。しかし、楽曲聴取時に一曲聴き終える度に膨大な楽曲データベースの中からテキスト情報を用いて選曲を行なうことはユーザにとって操作上の負担となり音楽がもつエンターテインメント性を損ないかねない。また、インターネットを利用した楽曲配信サービスの普及により、ユーザは自身が好む楽曲を1楽曲単位で購入することが可能となった。そのため、ユーザは時間や場所を問わずに自身の感性に合う多数の楽曲を享受可能となったが、曲単位で購入した楽曲によって構成された楽曲データベースには楽曲のジャンルや印象などに統一性が無くなってしまい、テキスト情報を用いた選曲はさらに難しい。

膨大な楽曲データベースを所有するユーザが煩雑な操作を必要とせず音楽を楽しむ代表的な手法としては、「ランダム再生」と「プレイリスト作成」の二つが挙げられる。ランダム再生は楽曲データベース内の楽曲をランダムに再生する手法であり、ユーザは一度ランダム再生を開始すれば一切の操作を必要とせず、音楽聴取時にユーザに課される操作上の負担は軽い。しかしながら、膨大で多種多様な楽曲を含む今日の楽曲データベースでは、ユーザ自身が収集した楽曲データベースでさえ、全ての楽曲がシステム使用時のユーザ感性に適した楽曲であるとは言えない。そのため、ユーザ感性に適さない楽曲が再生された場合、ユーザが聴きたいと感じる楽曲が再生されるまでスキップ操作を続けるか、再度テキスト情報を用いて選曲を行う必要が生じる。プレイリスト作成では、ユーザが聴きたいと感じる楽曲および再生する順序をあらかじめ指定するため、ユーザ感性に適した楽曲のみが再生される。しかし、あらかじめ大量の楽曲データベースからテキスト情報などを用いて選曲しリストを作成しなければならず、また、再生される楽曲と順序はプレイリスト毎に固定であるためユーザに「飽き」が生じやすい。

以上の問題から、煩わしい操作を必要とせずユーザ

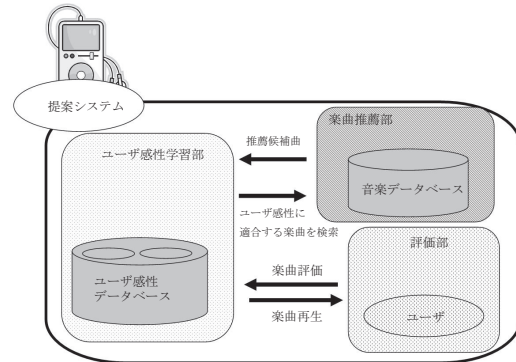


図1 提案システム概要

個々の感性に応じた楽曲聴取システムが求められていると考える。本稿ではユーザ個人が所有する多種多様な楽曲データの中からユーザが聴きたいと感じる楽曲、つまりユーザ感性に適合する楽曲のみを推薦・再生するシステムを提案する。図1に提案システムの概要を示す。ユーザ感性に適合する楽曲を推薦するためには、当然ながら提案システムはどのような楽曲がユーザ感性に適合するかを認識しなければならないが、あらかじめユーザ感性を学習する必要があるシステムではプレイリスト作成と同様にユーザに操作上の負担を掛けてしまう。そこで、提案システムではユーザが最後まで聴取した楽曲をユーザの感性に適合する楽曲、ユーザが楽曲の再生中にスキップした楽曲をユーザの感性に適合しない楽曲と仮定することでユーザの感性に適合する楽曲を認識し、システムの使用に伴いユーザの無意識下で自動的にユーザ感性を学習する。そのため、ユーザが提案システムの使用時に行う操作は「スキップ」のみとなり、ユーザにかかる操作上の負担を軽減することが可能と考える。提案システムはユーザにより適合と評価された楽曲群および不適合と評価された楽曲群の情報を記録し、それぞれの群の楽曲特徴からユーザ感性を学習し、推薦楽曲を決定する。また、提案システムで用いる楽曲特徴として音響特徴とメタ情報の2種類を用意し、音響特徴がユーザに与える印象を重要視するユーザには音響特徴を、アーティスト名やジャンルといったテキスト情報を重要視するユーザにはテキスト情報をそれぞれ用いることで多様なユーザ感性への適応を図る。

## 2 関連研究

楽曲推薦に関する研究はその目的により、ユーザが未所有の楽曲データベースからの楽曲推薦とユーザが所有する楽曲データベースからの楽曲推薦の2種類に分けられる。前者では、協調フィルタリング [1] を用いて感性

† 名古屋工業大学, Nagoya Institute of Technology

‡ 立命館大学, Ritsumeikan University

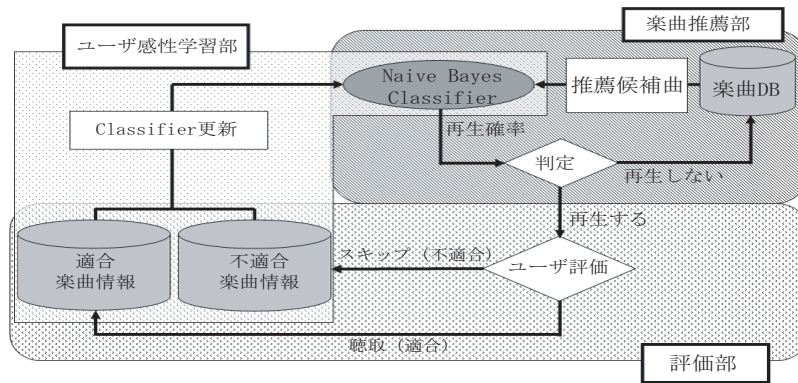


図 2 提案システムのアーキテクチャ

が似ているユーザ同士の楽曲情報を利用する楽曲推薦が多く、iTunes Genius [2] などの楽曲配信サービスで新たな楽曲を入手する場合において利用され有効性が確認されている [3]。本研究では、後者のユーザが所有する楽曲データベースからの楽曲推薦を目的としている。

関連研究として、彦坂ら [4] はユーザに煩わしい操作を要求することなく、ユーザの好みに合わせた自動選曲システムを提案している。このシステムでは、ユーザの主観的な好みと楽曲の音響的特徴との関連性を基にユーザの気分にも適合する楽曲をオンライン学習により選曲するシステムを提案している。このシステムは「ユーザが聴きたくない」と評価した楽曲から音響特徴的にもっとも掛け離れた楽曲を「ユーザが聴きたい楽曲」とであると仮定して推薦する。しかし、音響的に離れた楽曲を推薦するのみの仮定ではユーザの好みを捉えるためには十分とはいえない。音響特徴を用いた楽曲推薦としては他に Flexer [5] や Yoon ら [6] の研究も存在するが、選曲時にアーティスト名やジャンルなどのメタ情報を考慮する場合も考えられ、音響特徴のみを用いたこれらの手法では多様なユーザ感性に適応した楽曲推薦の実現は難しい。一方で、井原ら [7] はアーティストの性別やジャンルといったメタ情報から変化するユーザの好み（動的好み）を推定する「好み推定法」を提案しており、楽曲推薦において高いユーザ満足度を確認している。Bogdanov らの研究 [8] で報告されているように、メタ情報のみを楽曲特徴として用いた場合でもユーザが満足する楽曲推薦は可能であるが、メタ情報のみでは音響特徴から受ける楽曲の印象を重視した選曲には対応することができず、感性に適応する楽曲推薦システムとしては不十分と考える。以上から、幅広く多様なユーザ感性を捉えるためには、音響特徴とメタ情報をハイブリッドに用いるべきと考える。

Yoshii ら [9] は音響特徴とメタ情報をハイブリッドに用いた楽曲推薦を提案し、両特徴を扱った楽曲推薦システムの有用性を確認している。しかし、Yoshii らのシステムでは楽曲推薦にインターネットを用いて獲得した他者の楽曲評価を必要としており、また、ユーザは楽曲を 5 段階で評価しなければならず楽曲評価のためにインターネット環境および煩雑な操作を必要とする。寺田ら [10] や斉藤ら [11] は、それぞれ特別なグラフィックユーザインターフェース (GUI) を用意してユーザの感

性を捉え、ユーザの個性に応じた楽曲推薦を実現している。これらのシステムは、データベース内の楽曲を感性に基づき分類するためには有用であると考えられる。しかし、用意された GUI はランダム再生やプレイリスト選択などの従来手法に比べ複雑な操作を必要としており、聴取時に逐次的にユーザ感性を捉え、感性に応じた楽曲推薦を実現するためには、より単純な操作のみで実現される楽曲推薦が期待される。

また、我々は先行研究として印象言語を用いた感性的選曲システムを提案している [14]。この感性的選曲システムでは、ユーザが能動的に聴取したい楽曲を検索する場合を想定しており、GUI を用いてユーザ感性を印象言語の種類とその度合いにより、感性的な要求に応じた選曲を可能にしている。一方で、本稿で提案する楽曲推薦システムは、音楽を作業のバックグラウンドミュージック (BGM) として聴くなど半受動的な音楽聴取を想定しており、GUI や複雑な操作を必要とせずユーザ感性に適応した楽曲を推薦・再生することを目的としている。

### 3 提案システム

図 2 に本稿で提案するインタラクティブ楽曲推薦システムのアーキテクチャを示す。提案システムでは、ユーザが所有する楽曲データベースからユーザ感性に適合する楽曲の推薦を目的とし、「推薦楽曲を全て聴取」、「スキップ」の単純な操作のみからユーザ感性をインタラクティブに学習する。そして、sec:musicfeature 節に詳細を示すユーザ感性への適応を図って用意した楽曲特徴を用いて、単純な操作のみでユーザ感性に逐次適応可能な楽曲推薦を実現する。

まず、楽曲推薦部で楽曲データベースからユーザ感性に適合すると推測される推薦候補曲が Naive Bayes Classifier に入力される。そして、これまでのユーザ感性の学習に基づき推薦候補曲のユーザ感性への適合確率 (再生確率) が算出され、推薦候補曲を再生するか否かが確率的に決定される。ただし、提案システムを初めて利用する場合にはユーザ感性の学習データが存在しないため、再生確率は 1 となり推薦候補曲は必ず再生される。確率的に楽曲推薦を行うことでセレンディピティを創発しユーザの飽きを防止するとともに、変化する多様なユーザ感性を捉える。



次に評価部で、ユーザは再生された推薦楽曲が自身の感性に適合する場合は楽曲の最後まで聴取し、適合しない場合はスキップを行って楽曲を評価する。システムはユーザ感性データベースとしてユーザ感性に適合した楽曲、不適合であった楽曲のそれぞれの楽曲特徴を記録する。ここで、ユーザ感性データベースには適合と不適合についてそれぞれ最新  $N$  曲分のみを記録することで、変化するユーザ感性への対応を図る。例えば、提案システム使用中にユーザ感性が大幅に変化した場合でもユーザ感性に適合、不適合な楽曲特徴をそれぞれ新たに学習し、ユーザ感性データベースを更新すれば、システムは変化したユーザ感性に応じた楽曲推薦が可能になる。本稿では予備実験を通して、ユーザ感性をうまく捉えられ、かつ、感性の変化に適応可能な  $N$  の数値をヒューリスティックに決定し、 $N = 30$  とした。

ユーザ感性学習部では学習したユーザ感性データベースを基に Naive Bayes Classifier が更新され、次楽曲の推薦では更新された Naive Bayes Classifier を用いて楽曲を再生するか否かが決定される。つまり、推薦楽曲が再生され、ユーザの評価を得るに従って提案システムはユーザ感性をより捉えたものへとインタラクティブに適応していく。

#### 4 楽曲特徴

本稿では、楽曲特徴として音楽ゆらぎ特徴とメタ情報の 2 種類の特徴を用いる。音楽における時間的変化は認知学においても重要性が述べられており [12]、先行研究では音楽の時間的変化を捉える音楽ゆらぎ特徴と楽曲印象に関係性があることを確認し [13]、これらの関係性を基にした感性的選曲システムを提案している [14]。そこで本稿では、楽曲特徴の一種として音楽ゆらぎ特徴を用いることで音響特徴から受ける印象に基づいて選曲するユーザ感性への対応を図る。また、楽曲の音響特徴から受ける印象を重要視して選曲するユーザがいる一方で、特定のアーティストやジャンルの楽曲を好むユーザもいると考えられる。そこで、これらのユーザ感性を捉えるためにアーティスト名やアルバム名などのメタ特徴も楽曲特徴の一種として扱う。

表 1 に提案システムで用いる 70 個の楽曲特徴を示す。ここで、楽曲特徴 1 から 66 は音楽ゆらぎ特徴であり、楽曲特徴 67 から 70 がメタ特徴となる。次節からはこれらの楽曲特徴について詳細を述べる。

##### 4.1 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、リズムが知られており、これらの時間的変化が楽曲印象に影響を与えられられる。音楽ゆらぎ特徴では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、リズムの時間的変化を捉える。

本稿では、一般的に市販されている CD から収集した 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サンプリング周波数 44100Hz、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。まず事前処理として、各楽曲データに対して時間分解能 (25ms) 毎にシフトしながら窓幅 2048 点のハニング窓 (およそ 46ms) でフーリエ変換を施し、各時刻における周波数スペクトル  $fluc(\omega, t)$  ( $\omega$ : 周波数,  $t$ : 時刻) を算出する。

表 1 楽曲特徴

特徴番号	ゆらぎスペクトル特徴量
1 - 6	音量 (all) の特徴量 (1) ~ (6)
7 - 12	音量 (low) の特徴量 (1) ~ (6)
13 - 18	音量 (middle) の特徴量 (1) ~ (6)
19 - 24	音量 (high) の特徴量 (1) ~ (6)
25 - 30	音量 (ultrahigh) の特徴量 (1) ~ (6)
31 - 36	音高 (周波数重心) の特徴量 (1) ~ (6)
37 - 42	音高 (比率) の特徴量 (1) ~ (6)
43 - 48	音高 (回帰直線の傾き) の特徴量 (1) ~ (6)
49 - 54	音高 (回帰直線の切片) の特徴量 (1) ~ (6)
55 - 60	音高 (ピーク周波数) の特徴量 (1) ~ (6)
61 - 66	リズム (ビートスペクトラム) の特徴量 (1) ~ (6)
67 - 70	アーティスト名, アルバム名, ジャンル, 発売年代

そして、算出された  $fluc(\omega, t)$  から以下の音量に関する 5 特徴、音高に関する 5 特徴、リズムに関する 1 特徴をそれぞれ算出する。

- 音量 all (全周波数帯域) における音量
- 音量 low (周波数帯域 200Hz 未満) における音量
- 音量 middle (周波数帯域 200Hz 以上 800Hz 未満) における音量
- 音量 high (周波数帯域 800Hz 以上 2000Hz 未満) における音量
- 音量 ultrahigh (周波数帯域 2000Hz 以上) における音量
- 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
- 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

ここで、各周波数帯域における音量は各周波数帯域におけるパワースペクトラムの積分値とし、下式によって算出される。

$$S_{\omega}(t) = \int^{\omega} freq(\omega, t) d\omega. \quad (1)$$

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 [15] を用いる。

ここで、得られた特徴量を時系列化することで音楽ゆらぎ情報が得られる。それぞれの音楽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、各ゆらぎスペクトルにおける以下の 6 つの特徴量を音楽ゆらぎ特徴として抽出する。

- (1) スペクトル積分値  $S$
- (2) 周波数重心  $f_c$
- (3) スペクトル最大値  $fluc^{max}$
- (4) スペクトル最大の周波数  $f_{max}$
- (5) スペクトル回帰直線の傾き  $\alpha$

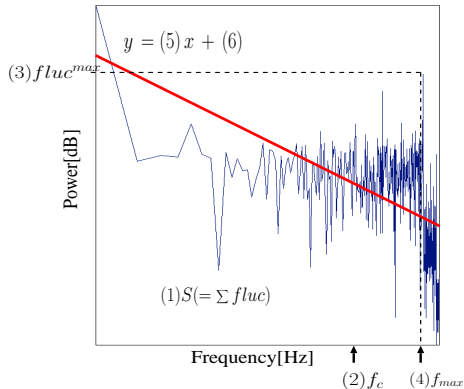


図 3 ゆらぎスペクトルの一例 (周波数重心)

### (6) スペクトル回帰直線の切片 $\beta$

このとき、全楽曲で平均値が 0、標準偏差が 1 となるように音楽ゆらぎ特徴毎に正規化を行った。図 3 に、周波数重心のゆらぎスペクトルの一例と音楽ゆらぎ特徴の抽出例を示す。ここで、図中の (1) ~ (6) の特徴量がそれぞれ、表 1 中の [31] ~ [36] の特徴量に該当する。音楽ゆらぎ特徴のより詳細な抽出方法については、文献 [13] を参照されたい。

#### 4.2 メタ情報

楽曲を表す特徴として、音響的な特徴だけではなくアーティスト名やジャンルなどのメタ情報も存在する。楽曲特徴としてこれらのメタ情報も音響特徴と併せて用いることで、音響特徴とメタ情報が組み合わさった複雑なユーザ感性にも対応した楽曲推薦システムが構築できると考えた。本研究では Apple 社の iTunes で音楽情報として用いられているアーティスト名、アルバム名、ジャンル、及び、発売年代の 4 種類のメタ情報を収集し、楽曲特徴として扱う。

## 5 ユーザ感性に基づいた楽曲推薦

提案システムではユーザが評価済みの楽曲特徴を学習し、楽曲データベース内の未呈示楽曲がユーザの感性に適合するかを推定する。本研究では、ユーザ感性推定器として確率的分類手法の一つである Naive Bayes Classifier を応用する。Naive Bayes Classifier は非常に単純な式でクラス分類が可能であることから迷惑メール分類などで多く利用されている [16]。そのため、提案システムの携帯音楽再生機器などへの応用にも適した有用性の高いクラス分類器であると考えられる。

Naive Bayes Classifier は以下の式で表される。

$$\begin{aligned} & \text{classify}(f_1, \dots, f_n) \\ & = \operatorname{argmax}_c p(C = c) \prod_{i=1}^n p(F_i = f_i | C = c). \end{aligned} \quad (2)$$

ただし、決定規則は最大事後確率決定規則を用いた式である。

提案システムでは、クラス変数  $C$  は推薦候補となる楽曲がユーザ感性に適合か不適合かを表すクラスであり、特徴変数  $F_i$  は楽曲特徴 (4 節参照) を表し、 $n = 70$  となる。推薦候補曲の特徴量  $f_1, \dots, f_n$  が与えられたとき、推薦候補曲が  $C = \text{適合}$  である確率  $p(C = \text{適合})$



図 4 提案システム実験用インターフェース

と、 $C = \text{適合}$  であるときの特徴量が  $F_i = f_i$  である確率  $p(F_i = f_i | C = \text{適合})$  を学習データから算出し、それらの積が推薦候補曲が適合となる確率を示す。同様に推薦候補曲が不適合となる場合についても確率を算出する。提案システムでは適合、不適合についてそれぞれ算出した  $p(C = c) \prod_{i=1}^n p(F_i = f_i | C = c)$  を正規化した値に基づいて推薦候補曲の再生率を確率的に決定し、ユーザ感性への柔軟な適応および意外性を持つ選曲による飽きの防止を図る。このとき、事前分布  $p(C = c)$  は一様分布とした。

Naive Bayes Classifier を用いる場合、ゼロ頻度問題と連続値への対応問題について考えなければならない。ゼロ頻度問題とは推薦候補曲の特徴が学習データに含まれなかった場合にその特徴のクラスらしさを表す確率  $p(F_i = f_i | C = c)$  が 0 になるという問題であり、本稿ではラプラススムージング [17] を用いて、仮想的なカウント加算を行うことで対応した。また、本来 Naive Bayes Classifier で扱うことのできる特徴量は離散値であり、連続値である音楽ゆらぎ特徴をそのまま扱うことは出来ない。この問題を解決し得る手法として、Geiger ら [18] が一つの特徴量を一つの確率密度分布で近似することで連続値を扱うことが可能であることを示している。本稿では、ガウス分布による確率近似を行うことで連続値の特徴量についても  $p(F_i = f_i | C = c)$  を算出可能となるよう対応した。

## 6 提案システム評価実験

提案システムの有用性を確認するため、3 種類の評価実験を行った。本来の提案システムではユーザインターフェースにはスキップボタンのみが存在すればよく、楽曲聴取の継続を適合、楽曲のスキップを不適合とみなしてユーザ感性を学習する。実験では、提案システムの有用性を数量的に確認するために図 4 のような実験用のインターフェースを用意した。実験用インターフェースには推薦楽曲の評価用に 4 つの評価ボタンが用意されており、ユーザは推薦楽曲に対して自身の感性に推薦楽曲がどの程度適合するかを評価する。推薦楽曲を聴取後、楽曲が自身の感性にとっても適合する場合は「とても合っている」、適合する場合は「合っている」の評価ボタンをそれぞれ押し、ユーザ感性に適合しない場合は「合っていない」、全く適合しない場合は「全く合っていない」の評価ボタンをそれぞれ押し楽曲をスキップする。ただし、これらのボタンはあくまでシステムの有用性を検証する為だけに存在しており、システム内ではユーザ感性を適合 (最後まで聴取)、不適合 (楽曲をスキップ) の 2 値のみで学習している。このとき、「とても合っている」と「合っている」は適合、「合っていない」と「全く合っていない」は不適合とした。実験では、これらの 4 段階

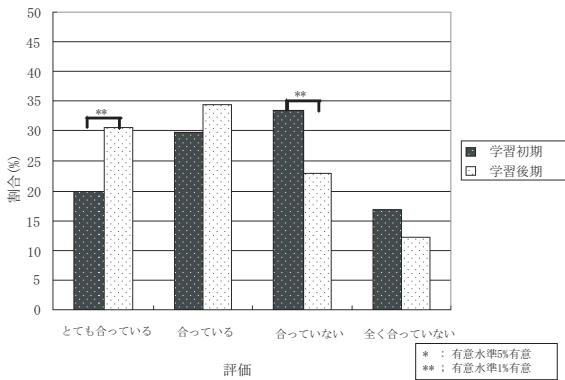


図 5 実験結果 1: 学習初期と学習後期における全被験者の各評価の割合 (%)

の評価の割合がどのように変化するか考察する。

評価実験では、全ての被験者で同一の計算機およびスピーカを用いた。提案システムで使用する楽曲データベースには本来であれば個人がそれぞれ所有するものを用いるべきだが、便宜上全被験者で共通の実験用楽曲データベースを用意した。実験用楽曲データベースは被験者に所有している音楽 CD を持参してもらい、楽曲の印象やジャンルなどが偏らないよう考慮したうえで全 909 曲の楽曲を収集した。また、実験用楽曲データベース内の楽曲は被験者の実験参加への負担を考慮し、一般的にサビと呼ばれる楽曲の一部を切り取り再生時間が 30 秒前後となるように編集している。被験者は、心身ともに健康な 20 代前半の男性 16 名とした。

### 6.1 実験 1: 提案システムの有用性確認

本実験では、提案システムがユーザ感性を未学習の状態を初期状態とし、被験者の 100 曲分の楽曲評価を終了条件とする。推薦楽曲の評価には実験用インターフェースを用いて、被験者自身の感性に適合する度合いを評価してもらった。ここで、本実験ではユーザ感性の学習が不十分であると考えられる 1 曲目から 20 曲目までを「学習初期」、ユーザ感性の学習が進んだと考えられる 81 曲目から 100 曲目までを「学習後期」と呼ぶ。そして学習初期と学習後期における全被験者の評価の割合から、提案システムがユーザ感性を捉えることが可能であるか考察する。

図 5 に学習初期の各評価の割合と、学習後期の各評価の割合を示す。同図から、学習初期に比べ学習後期ではポジティブな評価の割合が高くなっており、ネガティブな評価の割合が下がっていることがわかる。比率の差の検定 [19] を用いて学習初期と学習後期の各評価の割合について有意差検定を行ったところ、「とても合っている」と「合っていない」では学習初期と学習後期の評価の割合に有意差が確認された。つまり、学習後期ではユーザ感性に適合した楽曲が推薦されるようになり、また、ユーザ感性に不適な楽曲は推薦されにくくなったと考えられる。このことから、提案システムはシステム使用に伴って次第にユーザ感性を捉えることが可能であり、ユーザ感性に応じた楽曲推薦システムとして十分な有用性をもつことが示唆された。

### 6.2 実験 2: ユーザ感性の変化への適応性

感性はユーザごとに固有であるとともに、一人のユーザの感性も時間の経過と共に変化すると考えられる。本実験では、提案システムがユーザ感性の変化を捉え、その変化に順応して楽曲を推薦可能であるかを確認する。

被験者には推薦楽曲への評価にテーマを課し、推薦楽曲の 1 曲目から 100 曲目までは「激しい楽曲」を適合とし、101 曲目から 200 曲目までは「穏やかな楽曲」を適合として全 200 曲の推薦楽曲を評価させた。実験 1 の結果から、100 曲目までの推薦楽曲を評価し終えた時点で提案システムは被験者が激しいと感じる楽曲を推薦するシステムであると考えられる。そして、101 曲目からは被験者は穏やかな楽曲を適合とするため、101 曲目からしばらくの間、被験者の評価として「合っていない」「全く合っていない」の割合が高くなると考えられる。提案システムがユーザ感性の変化に順応可能であれば、システムの使用に伴い次第に変化したユーザ感性を捉え 200 曲目までの推薦楽曲を評価された時点では「とても合っている」「合っている」の割合が高くなると考えられる。

まず、1 曲目から 100 曲目までの被験者の推薦楽曲に対する評価について考察する。1 曲目から 20 曲目までを「激しい楽曲学習初期」、81 曲目から 100 曲目までを「激しい楽曲学習後期」とし、これら 2 つの期間の被験者の評価の割合を図 6 に示す。全体の傾向として、学習初期に比べ学習後期は被験者のポジティブな評価が高くなっていることが見てとれる。また、比率の差の検定を用いた有意差検定では、学習初期と学習後期の間で「とても合っている」と「全く合っていない」の評価に有意差が確認された。このことから、推薦楽曲を 100 曲評価した時点で提案システムは被験者の「激しい楽曲」についてのユーザ感性を学習できたことが示唆された。

次に、101 曲目から 200 曲目までの被験者の推薦楽曲に対する評価について考察する。ここで、101 曲目から 120 曲目までを「穏やかな楽曲学習初期」、181 曲目から 200 曲目までを「穏やかな楽曲学習後期」とし、これら 2 つの期間における被験者の評価の割合を図 7 に示す。比率の差の検定を用いた有意差検定の結果、全ての評価において学習初期と学習後期とで有意差が確認された。学習初期では、提案システムは学習した被験者の激しい楽曲についてのユーザ感性に適合する楽曲を推薦するが被験者は穏やかな楽曲を適合として評価している。そのため、「とても合っている」「合っている」の評価が非常に低く、逆に「合っていない」「全く合っていない」の評価が高い。しかし学習後期では評価の割合は逆転しており、「とても合っている」「合っている」の評価が非常に高く、逆に「合っていない」「全く合っていない」の評価が低い結果となっている。このことから、提案システムがユーザ感性の変化に適応可能であり、変化したユーザ感性に応じた楽曲を推薦可能であることが示唆された。

### 6.3 実験 3: 音楽ゆらぎ情報とメタ情報の重要性

選曲時に楽曲のどの要素に着目するのか、つまりユーザ感性に関連する楽曲要素はユーザ毎に異なると考えられる。そのため、多様なユーザ感性を捉えた楽曲推薦を実現するためには、楽曲の音響特徴から受ける印象を重要視するユーザ感性に適応するためには楽曲の音響特徴を、一方「誰が歌っているのか」「ジャンルは何か」といったメタ情報を重要視するユーザ感性にはメタ情報を



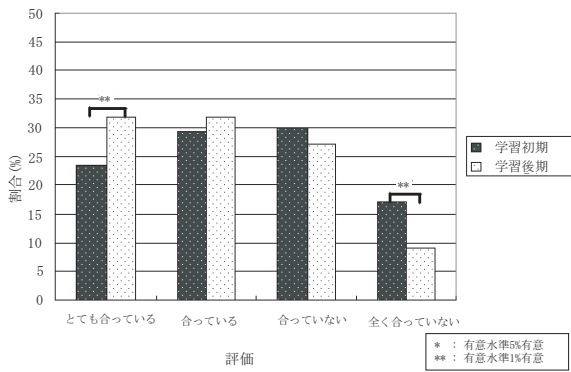


図 6 実験結果 2 - 1:「激しい楽曲」学習初期と後期における全被験者の各評価の割合 (%)

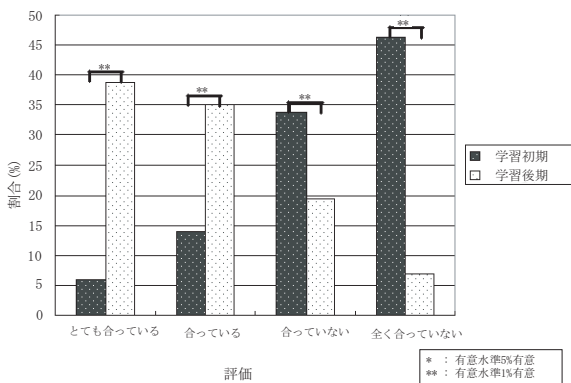


図 7 実験結果 2 - 2:「穏やかな楽曲」学習初期と後期における全被験者の各評価の割合 (%)

それぞれ考慮する必要がある．そのため提案システムでは、音楽ゆらぎ情報とメタ情報の両方を楽曲特徴として扱っている．本実験では、音楽ゆらぎ特徴のみを扱う A system、メタ情報のみを扱う M system、音楽ゆらぎ情報およびメタ情報を扱うシステム AM system (提案システム) の 3 つの実験用楽曲推薦システムを用意し、各システム使用時における被験者の評価について考察する．本実験ではまず事前に、被験者に楽曲データベース内から自身の感性に適合する楽曲を 30 曲、不適合な楽曲を 30 曲ずつそれぞれ選出させ、各実験用楽曲推薦システムの学習データとして用いた．被験者は実験用インターフェースを用いて各システムによって推薦された楽曲をそれぞれ 30 曲評価した．

各システムによって推薦されたそれぞれ 30 曲の推薦楽曲のうち、ユーザ感性の学習が進んだと考えられる 21 曲目から 30 曲目までの楽曲に対する各価別の割合を図 8 に示す．比率の差の検定により各システムの評価について有意差検定を行ったところ、「合っている」の評価について AM system が A system および M system に対して有意に高く、「全く合っていない」の評価について M system が A system および AM system に対して有意に高い結果を示した．つまり、音楽ゆらぎ特徴およびメタ情報を楽曲特徴として用いた AM system (提案システム) は高いユーザビリティ評価を示す一方で、メタ

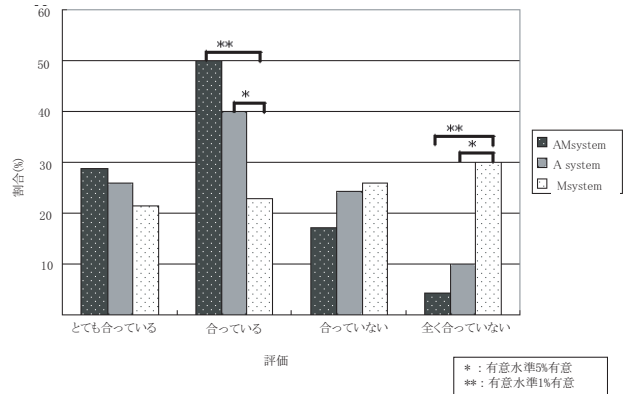


図 8 実験結果 3: 各実験用楽曲推薦システムによる推薦楽曲についての各評価の割合 (%)

情報のみをユーザ感性の学習に用いた M system は低いユーザビリティ評価となった．また、有意差は確認されなかったが AM system は A system に比べて比較的高いポジティブな評価を得ていることが見てとれる．これらのことから、楽曲推薦システムにおいて音楽ゆらぎ特徴およびメタ情報の併用が多様なユーザ感性へ適応するために有用であると示唆された．

## 7 まとめと今後の展望

本研究では、楽曲特徴とユーザそれぞれの感性の関係をインタラクティブに学習し、ユーザ感性に適合する楽曲推薦システムを提案した．このとき、楽曲特徴として音楽ゆらぎ特徴およびメタ情報を扱うことで多様なユーザ感性への適応を図った．提案システムでユーザ感性を学習するために必要となる操作は、推薦楽曲がユーザ感性に不適合である場合に行う楽曲のスキップのみであり、直感性の高い楽曲推薦システムと考える．そのため、自動車の運転中など、GUI の操作が不自由な状況において、特に高い有用性が期待される．

評価実験を行ったところ、提案システムはシステム使用に伴ってユーザ感性を捉えることが可能であり、ユーザ感性がシステム使用中に変化した場合でもユーザ感性の変化に順応し、ユーザ感性に適合した楽曲を推薦可能であることを確認した．また、比較実験により楽曲特徴として音楽ゆらぎ特徴とメタ情報の両方をを用いた提案システムは、それぞれを単独で楽曲特徴として用いたシステムに比べてユーザ感性をよりうまく捉えた楽曲推薦が可能であることを確認した．

提案システムの今後の展望として、以下の 4 点が挙げられる．

- (1) Naive Bayes Classifier における連続値対応  
本稿で楽曲特徴として扱った音楽ゆらぎ特徴は連続値であるためガウス分布を用いて離散値に近似し、Naive Bayes Classifier の特徴として用いた．今後は、複数のガウス分布から算出される混合分布 [20] を用いた連続値の離散化を行い、よりの確、柔軟にユーザ感性を捉える手法を検討する．

## (2) 学習データ数の動的変化

提案システムは、学習データとして  $N$  曲分の楽曲特徴を保持することでユーザ感性を捉え、感性の変化にも対応した楽曲推薦を実現している。しかしながら、ユーザ感性の変化が生じたことを自律的に認識し動的に学習データの保持数を変更することで、より素早くユーザ感性の変化へ適応可能と考える。

## (3) 歌詞情報の取り扱い

本稿では楽曲の性質を示すメタ情報としてアーティスト名、アルバム名、ジャンル、発売年代の4種類を用いた。この他に、歌詞情報もまた同様に楽曲の特徴を示すメタ情報の一つとして考えられる。現在、様々な歌詞情報提供サービス(例えば、SONY社の「歌詞ピタ」[21])が存在しており、これらのサービスを利用することで容易に歌詞情報の取得が可能である。取得した歌詞情報について、形態素解析などの言語処理技術を応用して抽出した「歌詞の印象」等をメタ情報として扱うことで楽曲の歌詞も考慮し、より多様なユーザ感性に適応した楽曲推薦システムの構築が期待される。

## (4) 他メディアへの応用

また、本稿で提案したシステムアーキテクチャは、楽曲推薦以外にも、動画像における動きの特徴などを用いることで様々なメディア推薦システムとして発展、応用させることが可能であると考えられる。

今後はこれらの課題に取り組み、直感性および有用性の高い楽曲推薦システムを実装していくとともに、感性的でインタラクティブなメディア情報推薦システムとしての発展をねらう。

## 謝辞

本研究は、一部、文部科学省科学研究費補助金(課題番号 20700199)の助成のもと行なわれた。

## 参考文献

- [1] L. Terveen and W. Hill: "Beyond recommender systems: Helping people help each other", HCI in the New Millennium, Addison-Wesley, pp. 487-509 (2001).
- [2] Apple Inc.: "iTunes". <http://www.apple.com/itunes/>.
- [3] L. Barrington, R. Oda and G. Lanckriet: "Smarter than genius? human evaluation of music recommender systems", Proceedings of 2009 International Society for Music Information Retrieval, pp. 357-362 (2009).
- [4] 彦坂, 谷口, 菅田, 白井: "ユーザの嗜好に適應させた選曲システムの提案", 情報処理学会研究報告. [音楽情報科学], **2006**, 133, pp. 19-24 (2006).
- [5] A. Flexer, M. Gasser and D. Schnitzer: "Limitations of interactive music recommendation based on audio content", Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound, pp. 96-102 (2010).
- [6] T. Yoon, S. Lee, K. ho Yoon, D. Kim and J.-H. Lee: "A personalized music recommendation system with a time-weighted clustering", Proceedings of Intelligent Systems, 2008. IS '08. 4th International IEEE Conference, pp. 10-48-10-52 (2008).
- [7] 井原, 金田, 上野, 金山: "情報をもつ属性を用いたユーザ動的好み推定法の提案およびその評価", 電子情報通信学会論文誌. A, 基礎・境界, **80**, 4, pp. 700-707 (1997).
- [8] D. Bogdanov and P. Herrera: "How much metadata do we need in music recommendation? a subjective evaluation using preference sets", Proceedings of 2011 International Society for Music Information Retrieval Conference, pp. 97-102 (2011).
- [9] K. Yoshii, M. Goto, K. Komatani, T. Ogata and H. G. Okuno: "An efficient hybrid music recommender system using an incrementally trainable probabilistic generative model", Proceedings of Audio, Speech, and Language Processing, IEEE Transactions on, pp. 435-447 (2008).
- [10] 寺田, 玉垣, 小原: "neut: 選曲の可視化による調整可能なランダム再生システム", デザイン学研究作品集, **11**, pp. 76-79 (2006).
- [11] 斉藤, 伊藤: "Musicube: 特徴量空間における対話形進化計算を用いた楽曲提示インターフェース", 人工知能学会全国大会(第25回) (2011).
- [12] L. B. Meyer: "Emotion and Meaning in music", University of Chicago Press (1956).
- [13] 伊藤, 山西, 加藤: "音楽ゆらぎ特徴を用いた楽曲印象の推定", 日本音響学会誌, **68**, 1, pp. 11-18 (2012).
- [14] 豊田, 山西, 加藤: "楽曲の音響ゆらぎに基づく感性的選曲システムの提案と選曲アルゴリズムの検討", 日本感性工学会論文誌, **11**, 2, pp. 223-231 (2012).
- [15] 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).
- [16] 山内, 上村, 畑中: "判定に利用するトークンの限定によるページアンフィルタの判定精度向上手法", 情報処理学会論文誌, **52**, 9, pp. 2686-2696 (2011).
- [17] C. D. Manning, P. Raghavan and H. Schütze: "Introduction to Information Retrieval", Cambridge University Press (2008).
- [18] D. Geiger and D. Heckerman: "Learning gaussian networks", Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence, pp. 235-243 (1994).
- [19] R. C. Sprinthal: "Basic Statistical Analysis (9th Edition)", Prentice Hall (2011).
- [20] A. Pérez, P. Larrañaga and I. Inza: "Bayesian classifiers based on kernel density estimation: Flexible classifiers", International Journal of Approximate Reasoning, **50**, 2, pp. 341-362 (2009).
- [21] SONY Inc.: "歌詞ピタ". <http://ss.kashi-ism.jp/kashipita/>.