

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

Hybrid Estimation of Music Preference Using User Ratings and Music Fluctuation Features

伊東 孝浩 † 加藤 昇平 †
Takahiro Ito Shohei Kato

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。ユーザが所有する楽曲データベースの中から楽曲を選択する場合、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、ユーザの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になると考える。

この負担を軽減するため、楽曲データベース内の楽曲をランダムに再生する方法が挙げられる。しかし、ユーザの負担が軽減できる一方でユーザの感性に合った楽曲が再生されない可能性がある。また、ユーザ自身があらかじめ楽曲データベースからプレイリストを作成する方法も挙げられる。しかし、あらかじめプレイリストを作成することで、選択される楽曲群が固定されてしまう問題点や、ユーザの感性が変化した場合に再びプレイリストを作成し直さなければならない課題も存在する。

これらの課題を解決するため、多田ら [1] 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。この手法では、システムが推薦した楽曲をユーザが最後まで聴取したか途中でスキップしたかをシステムが学習することでユーザの嗜好を特定し、学習結果を用いてユーザの感性に適合した楽曲を決定する。過去の聴取履歴から、ユーザが未評価である楽曲がユーザの嗜好に合うか否かを推定するために、確率モデルを用いてクラス判別を行う単純ベイズ分類器を用いている。単純ベイズ分類器によりユーザの嗜好を学習し、楽曲特徴の類似性を基に推定評価値を算出する。この結果から、楽曲とユーザの嗜好との適合確率を決定し、確率的な楽曲推薦を行う。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り、また分類器の学習が進むにつれて多様性のある楽曲がし難くなる問題点が存在する。

そこで本稿では、この既存手法の問題点を解決するために協調フィルタリングを導入することで多様性のある楽曲推薦手法を提案する。提案手法の概要図を図 1 に示す。協調フィルタリングとは情報に対するユーザの嗜好と他ユーザの嗜好の類似度を用いて、ユーザにとって有効な情報を推薦する手法である。他ユーザの嗜好を取り入れることで、音響特徴だけでは難しかった潜在的な嗜好を特定し、推薦性能の向上を目指す。提案手法の有効性を確認するため、既存手法と提案手法によって推薦される楽曲について比較を行った。実験の結果、推薦楽曲へのユーザ評価の半数以上が向上し、ユーザ評価に大きな変化が見られなかったユーザに関しては推薦楽曲の多様性の向上を確認した。

2 関連研究

宇野ら [2] はユーザが持つ楽曲データベース内の楽曲がいつ、どこで聴取されたかを記録し、ライフログと楽曲の関係性について研究している。ツリー構造を持つデータを可視化する平安京ビュー [3] を用いることで、楽曲の持つ特徴量とユーザ自身の嗜好の変化を一覧可視化する手法を提案してい

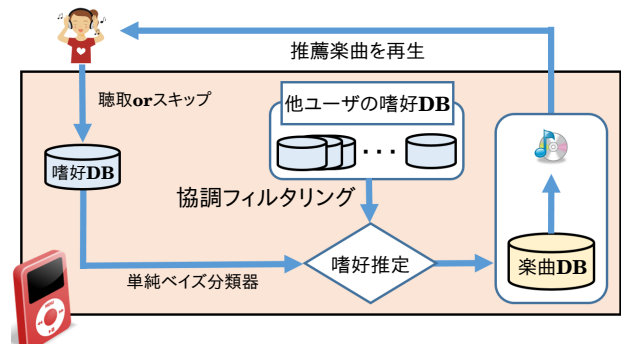


図 1 提案手法の概要

る。独自に開発した Android 端末用のアプリケーションにより、楽曲再生中の日付や曜日などのライフログ情報を蓄積し、楽曲とユーザの状況、環境との関連性を確認している。

また、濱崎ら [4] 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。動画コミュニティサービスである「ニコニコ動画」[5] 上に存在するコンテンツには、そのコンテンツのジャンルを示すカテゴリタグが付与されているが、濱崎らはそのカテゴリタグによるコンテンツ間の関係性には方向性があると述べている。ユーザが興味のあるコンテンツから他のコンテンツへの関係性を矢印タグとして定義することにより、新たな音楽コンテンツと出会うための方向リンクとして表現している。これらのリンクとコンテンツの持つ特徴量を用いることで、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。

また、清水ら [6] はユーザの知らない情報を特定し、その中からユーザの嗜好に合った情報を推薦するため、情報の発見性を考慮した推薦手法を提案している。従来の手法では、ユーザの嗜好に適した情報を推薦することに重点を置いているため、ユーザが既知の情報が多く推薦されるという問題があった。この問題点を解決するため、ユーザにとってどのような情報が既知であるかという観点でのユーザ間および情報間の類似度を計算することでユーザが知らないであろう情報を推測し、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。

3 先行研究

多田らによるこれまでの先行研究 [1] では、ユーザが所有する楽曲データベースからユーザの嗜好に合う楽曲の推薦を目的としている。「推薦楽曲を最後まで聴取」あるいは「途中でスキップ」の単純な操作のみからユーザの嗜好をインタラクティブに学習する。そして、5 章に示す楽曲特徴を用いて単純な操作のみでユーザの嗜好に適応可能な楽曲推薦を実現している。ユーザ嗜好の学習にはクラス分類を確率的に行う単純ベイズ分類器 (Naive Bays Classifier) を用いている。

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

3.1 推定評価値の算出

楽曲データベース中の推薦候補の楽曲に対して、過去のユーザ評価と楽曲特徴量を基に評価値推定を行う。単純ベイズ分類器では、楽曲の持つ特徴量 $f_1, f_2, f_3, \dots, f_n$ を基に、楽曲情報が属するクラス c_j を次のような式で分類する。

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

ここで C_{MAP} は分類されるクラス、 C はクラスの集合を表す。特徴量が独立である条件下で事後確率最大仮説によりクラス分類が可能であるが、単純ベイズ分類器で扱う特徴量は離散値でなければならない。そこで多田ら [1] 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 $P(f_i | c_j)$ を推定している。これらを基にユーザの嗜好に合うか否かを確率値で決定する。

3.2 推薦システム

推薦システムの流れについて説明する。まずシステムはユーザの感性に適合すると推定される推薦候補楽曲を選択する。3.1 節で述べたように推薦候補曲には嗜好に適合する確率値が付与されている。この確率に基づき、推薦候補曲を再生するかしないかを確率的に決定する。ユーザ感性は常に一定のものではなく、学習上は嗜好に合わないと推定されていても実際に推薦した場合に嗜好に合う可能性を考慮するためである。また、確率的に推薦曲を再生することで意外性を持たせユーザに飽きを生じにくくさせるも目的としている。楽曲が再生されると、ユーザの嗜好に合っている場合は最後まで聴取し、適合しない場合はスキップを行い評価する。システムは評価を受け取り、再度学習して次の推薦候補曲を決定する。

しかしながら、この手法では確率的ではあるものの過去に高い評価を与えた楽曲と類似した曲が推薦されやすい問題点がある。確率的に意外性のある楽曲を推薦するだけでなく、協調フィルタリングを用いてユーザの嗜好を捉えつつ多様な楽曲を次章から説明する。

4 提案手法

先行研究では確率的に推薦する楽曲を決定し多様な楽曲推薦を可能としているが、楽曲の持つ音響特徴量とユーザの過去の聴取履歴を基に決定しているため、過去に聴取した楽曲と類似したものが推薦されやすい問題点がある。この問題点を解決するため、情報推薦手法の一つである協調フィルタリングを導入し、ユーザの嗜好を捉えつつより多様性のある楽曲推薦を行う。

4.1 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[7]）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[8] 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。

本稿ではユーザ A が未評価の楽曲 i' に対して、先行研究によるユーザ A の嗜好推定値 $\hat{s}_{A,i'}$ に協調フィルタリングの要素を加えた推定評価値 $\hat{s}_{A,i'}$ の計算式を提案する。

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

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

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)

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} は下式により算出する。

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

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。

5 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 [9] 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 [10] 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。

5.1 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られられる。メロディ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、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

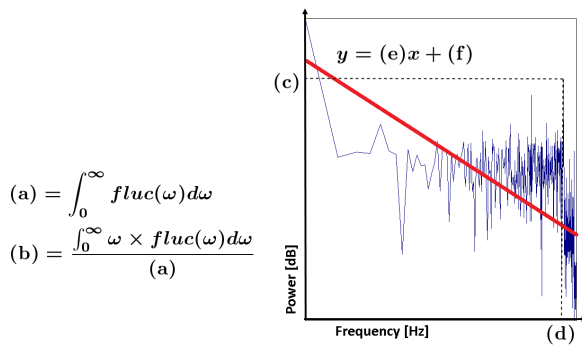


図 2 ゆらぎスペクトルの一例

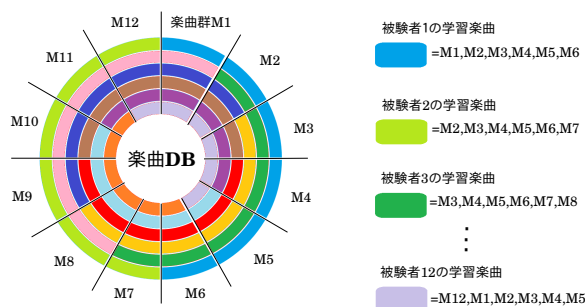


図 3 学習楽曲の割当方法 (被験者 12 名の場合)

- 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
- 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 [11] 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 $fluc^{max}$
- 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 2 にゆらぎスペクトルの一例を示す。

6 提案手法の評価実験

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 代前半の男性 12 名とした。

6.1 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データベース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。

6.2 実験に用いる特徴量

楽曲の持つ特徴量として第 5 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。

6.3 実験内容

事前準備としてシステムは被験者の嗜好学習を行う。被験者はデータベース内の楽曲 30 曲に好きか嫌いかの評価づけを行い、この結果を基に文献 [1] の手法を用いて被験者の嗜好を学習する。図 3 に被験者毎の嗜好学習楽曲の割当を示す。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 3 のように学習楽曲を割り当てることで、協調時における被験者間の公平性を考慮している。

実験では、事前準備で聴取していない 30 曲の推定評価値を算出し、既存手法と提案手法それぞれで推定評価値の上位 10 曲を選曲し、それぞれの手法で推薦された楽曲として被験者が評価実験する。評価は 1~5 の 5 段階 (1 を最低評価, 5 を最高評価) で行った。

6.4 実験結果

表 2 に各手法の推薦楽曲への評価値合計及び楽曲特徴量の分散を示す。結果から、12 人中 7 人の被験者で推薦楽曲への評価及び楽曲の持つ特徴量の分散の向上が確認された。被験者全体の平均についても評価値合計は 31.92 から 34.25 に、特徴量の分散は 1.31 から 1.34 に上昇した。評価値の合計と特徴量の分散の両方に向上が見られた被験者に注目すると、被験者 4、被験者 9 の合計評価値が大きく上昇していることが分かる。この結果からこれらのユーザに関して既存手法では困難であった多様な楽曲推薦を提案手法による協調フィルタリングが実現し、ユーザの嗜好をより捉えた推薦システムであるためだと考えられる。また、特徴量の分散のみ向上した被験者に注目すると、評価値の合計が大きく減少していないことが分かる。評価値の合計が最も減少した被験者 12 に関して、特徴量の分散が大きく上昇している。この結果から、被験者の嗜好を捉える点では僅かながらの問題が見られたが、トレードオフとして推薦楽曲の多様性が大きく上昇したと言える。さらに、評価値の合計のみ上昇した被験者に注目すると特徴量の分散が大きく減少している。この結果から、多様性のある楽曲推薦だけがユーザの満足度を向上させるのではなく、ユーザによっては特定の要素 (アーティストや曲調など) に重点を置いた楽曲推薦が有効であると思われる。同様に、被験者 8 に関しては特徴量の分散は大きく減少しているにもかかわらず、楽曲に対する評価値の合計は変化おらず、楽曲推薦において多様性だけでなくユーザの感性を正しく捉えた学習の必要性が示唆された。本稿の提案手法では、単純ベイズ分類器と協調フィルタリングのどちらを重視するかをユーザ毎に調整していなかった。この調整をユーザ毎に行うことで、よりユーザの満足度の高い推薦システムが実現できると思われる。

7 まとめと今後の課題

本稿では、ユーザの単純ベイズ分類器を用いた楽曲推薦手法に協調フィルタリングを導入することで、ユーザの嗜好を捉え

表2 各手法の上位10曲に対する被験者の評価値合計と楽曲特徴量の分散

	推薦楽曲への評価値の合計		楽曲特徴量の分散	
	提案手法	既存手法	提案手法	既存手法
被験者 1	32	31	1.72	1.43
被験者 2	31	33	2.69	1.63
被験者 3	34	33	1.95	2.29
被験者 4	29	24	0.68	0.95
被験者 5	45	35	1.16	0.95
被験者 6	31	32	1.40	0.59
被験者 7	40	31	2.70	3.51
被験者 8	36	36	2.38	4.41
被験者 9	37	30	0.36	0.30
被験者 10	32	34	0.17	0.15
被験者 11	38	34	0.29	1.20
被験者 12	26	30	0.53	0.31
平均	34.25 ± 5.22	31.92 ± 3.15	1.34	1.31

つつ、より多様性のある楽曲推薦手法を提案した。提案手法の有効性を確かめるため被験者実験を行った。結果として、提案手法においてユーザへの推薦楽曲の多様性の向上、単純ベイズ分類器ではなし得なかった楽曲推薦の可能性を確認した。しかし、被験者実験の結果から以下の問題点があげられる。

7.1 個人の嗜好傾向に合わせた重み付け

本稿では、単純ベイズ分類器による楽曲評価の推定と、他ユーザとの協調フィルタリングによる推定を単純に加算している。そのため、ユーザによって有効な楽曲推薦が必ずしも成功しないことがある。この問題点を解決するため、ユーザの嗜好傾向を捉え2種類の評価推定の重みをユーザ毎に調整することが考えられる。

7.2 ライフログの導入

ユーザが音楽を聴取する時には、時間帯や場所などの状況・環境が感性に大きく影響すると考えられる。宇野ら [2] が述べているように、楽曲とライフログの関係をシステムに反映することでユーザに有効な推薦が可能と考えられる。今後は過去の聴取履歴に加えて、聴取した際のライフログ情報を記録することにより、ユーザの状況に合った推薦精度の向上を目指すたい。

7.3 ユーザが未知である楽曲の発見性

本稿では、ユーザが持つ楽曲データベースからユーザの嗜好に合った楽曲推薦手法を提案した。今後のシステムの発展として、ユーザが音楽配信サービスなどを利用する状況下では、ユーザが未知の楽曲の中から嗜好に合うであろう楽曲の推薦も考えられる。本稿で用いた単純ベイズ分類器や協調フィルタリングの手法に加えて、楽曲の持つメタ情報（曲名、アーティスト、ジャンルなど）やタイアップ情報（ドラマ、アニメ、映画、CM など）を反映させることで、より多様な楽曲推薦が可能と考えられる。濱崎ら [4] のように、ウェブマイニング技術を応用することで、ウェブ上から楽曲推薦に活用できる情報を選別することでより多様な楽曲推薦を目指したい。

今後はこれらの課題に取り組み、よりユーザの嗜好を捉えた楽曲推薦システムの実装を行う予定である。

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 25280100、および、25540146）の助成により行われた。

参考文献

- [1] 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”, FIT2012 第11回情報科学技術フォーラム, Vol.11, pp.23-29 (2012).
- [2] 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).
- [3] 伊藤貴之, 山口裕美, 小山田耕二: “長方形の入れ子構造による階層型データ視覚化手法の計算時間および画面占有面積の改善”, 可視化情報学会論文集, Vol.26, No.6, pp.51-61, (2006).
- [4] 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).
- [5] 戀塚昭彦: “ニコニコ動画の創造性: 動画コミュニティサービス「ニコニコ動画」の5年間”, 情報処理学会 情報処理, Vol.53, No.5, pp.438—488(2012).
- [6] 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
- [7] LINDEN Greg: “Amazon.com Recommendations: Item-to-Item Collaborative Filtering”, *IEEE Internet Computing* Vol.7 pp.76-80(2003).
- [8] 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).
- [9] 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).
- [10] L. B. Meyer: “Emotion and Meaning in music”, University of Chicago Press (1956).
- [11] 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).