

楽曲レコメンドシステム

The System of Recommendation for Music Contents

児玉泰輝† 鈴木康悟† 松下文雄† 小田川智† 莪山真一† 塩田岳彦†
Yasuteru Kodama Yasunori Suzuki Fumio Matsushita Satoshi Odagawa Shinichi Gayama Takehiko Shioda

1. まえがき

2001年からHDDを利用したカーナビゲーションが市場に登場し、現在では高級機のカーナビゲーションはHDD搭載タイプが主流となっている。また、HDDを有効に活用する機能の一つとして、音楽録音機能を搭載した製品も多い。HDDに膨大な音楽を保存し、従来にはなかった自由な音楽の再生を可能とするものである。この機能は、市場で評価され、新たなカーAVシステムとして注目されている。

一方で、多くの楽曲の中から所望の曲を探し出すことが困難になるという問題が生じており、解決するための技術開発が行われている。例えば、インターネット経由でダウンロードした既成情報(楽曲イメージなど)や固定情報(アーティスト、タイトル、ジャンルなど)を利用して楽曲を検索するものや、あらかじめ規定した楽曲の印象表現を用いて楽曲を検索するものがある。しかし、全ての楽曲の情報をデータベース化することは困難であり、個々のユーザの嗜好に適合したレコメンドを行うことは難しい。また、自動車におけるインターネット接続は、通信の安定性、通信料金の面から普及しているとは言いがたく、カーナビゲーションにおいては不便なものであった。

そこで我々は、ユーザの嗜好を学習する機能をもつ楽曲レコメンドシステムを開発した。このシステムは、これまでのカーナビゲーションをより快適に、使いやすくするものとして期待されている。

本システムは、CDから自動的に音楽的な特徴量を抽出するため、ネットワークなどから楽曲に関する情報を取得する必要はない。また、音楽の印象を示す検索語(明るい、ノリがいい、静かな、かなしい、癒される)を選択するだけで、HDD内の多くの楽曲から、イメージにあった楽曲を再生する。さらに、ユーザの嗜好を学習し、各個人の好みを音楽のレコメンドに反映することができる。

本稿では、本システムの概要と実験結果について報告する。

2. レコメンド方法

CDから自動的に抽出される音楽的な特徴量を楽曲特徴量、検索語を規定する特徴量を検索特徴量と定義する。本システムは、これらの情報を用いて検索語と音楽の合致度を算出し、この結果に基づいてレコメンドする。

2.1 楽曲特徴量の算出

本システムは、楽曲を録音する際、楽音信号を分析することにより楽曲特徴量を求める。

具体的には、楽曲の調性、和音出現頻度の分散(HVL)、一分あたりのリズム量(BPM)、最大ビートレベル(MBL)、最大信号レベル(MSL)、および平均信号レベル(ASL)の6つの数値である。

最初に、楽音信号から和音と振幅情報を時系列データとして数値化する。次に、楽曲全体において最も優位な調性と和音出現頻度の分散を求める[1]。さらに、和音変化数と振幅極大値の個数から一分あたりのリズム量を算出する。なお、最大ビートレベル、最大信号レベル、平均信号レベルは各々振幅情報の極大値、最大値および平均値から求められる。

2.2 検索語の定義

本システムは指定された検索語に対して、楽曲をレコメンドするものである。検索語は音楽を表現する言葉で、車を運転中に利用すると考えられる検索語(「明るい」「ノリがいい」「静かな」「かなしい」「癒される」)を採用した。

2.3 検索特徴量の算出

検索特徴量とは、検索語と音楽のイメージを関連付ける情報である。この情報は、検索語に合致すると判断された曲の楽曲特徴量の記録(FIT記録)と、非合致すると判断された曲の楽曲特徴量の記録(UNFIT記録)から求められる、平均、分散(FIT平均、FIT分散、UNFIT平均)である。FIT、UNFIT記録には、あらかじめ初期値が与えられているが、3章に述べる学習処理によって適切に更新される。

2.4 合致度算出

これらの楽曲特徴量と検索特徴量を利用して音楽と検索語の合致度を算出し、レコメンドする楽曲を決定する。

選択された検索語にFITする曲をレコメンドするために、ここでは非合致度を算出し、この値が小さい曲を相対的に合致度の高い曲であるとしてレコメンドを行う。楽曲特徴量ごとの非合致度を特徴量非合致度(Score[i], $i=0, \dots, N-1$, ScoreKey)とし、2.4.1、2.4.2に示す方法で算出する。そして、すべての特徴量非合致度の合計を検索語とその曲の非合致度ScoreTotalとする。

$$\text{ScoreTotal} = \sum_{i=0}^4 \text{Score}[i] + \text{ScoreKey} \quad \dots \text{式1}$$

こうして求められた非合致度により、レコメンドを行う。

2.4.1 調性以外の特徴量非合致度算出方法

調性以外の楽曲特徴量の数をNとし、対象とする楽曲の楽曲特徴量と、所定の検索語に対するFIT平均をそれぞれElement[i], AgreeAve[i], $i=0, \dots, N-1$ として、式2でFIT平均との距離AgreeDistance[i]を求める。

$$\text{AgreeDistance}[i] = |\text{AgreeAve}[i] - \text{Element}[i]| \quad \dots \text{式2}$$

式2の結果とFIT分散AgreeDeviation[i]からElement[i]についての特徴量非合致度Score[i]を式3のように算出する。

$Score[i] = AgreeDistance[i] * 1/AgreeDeviation[i]$...式 3
 その他の特徴量非合致度も同様に算出する。

2.4.2 調性の特徴量非合致度算出方法

調性とは、音楽のもつキー（ハ長調、イ短調）であり、0 から 23 の値をとる。

調性 j に対する FIT、UNFIT 平均の差を $KeyAvrGap[j]$, $j=0, \dots, 23$ とすると、特徴量非合致度 $ScoreKey$ は式 4 のように調性以外の特徴量非合致度とのバランスを取るための定数 K を乗じたものとする。

$ScoreKey = K * KeyAvrGap[Key]$ 式 4

3. 学習方法

ユーザはレコメンドされた曲を鑑賞し、その曲が選択した検索語に FIT しているか判断する。システムはこの判断により FIT 記録または UNFIT 記録に楽曲特徴量を登録し、統計的解析により検索特徴量を更新する。この値は学習が適切に行われると、ある値（目標値）に収束することが期待される。

まず、FIT/UNFIT 記録の生成方法について説明する。

調性以外の特徴量は、非合致度算出において同様に扱うため、既知の平均と標準偏差で標準化したものを登録する。

例えば、ある楽曲特徴量(調性 = 0)を合致記録に追加する場合、調性の合致記録 $AgreeKey$ は、 $AgreeKey[0] = 1$ とし、 $AgreeKey[1 \sim 23] = 0$ とする。

次に、それぞれの記録の平均、分散を求め検索特徴量を更新する。

このようにして求めた検索特徴量を次回のレコメンドに利用する。

4 結果

ある被験者による「ノリがいい」について 200 回学習させ、FIT 曲の推移と検索特徴量の変化の様子を観察した。詳細は以下に説明する。ここで、レコメンド対象とする曲は 216 曲とする。また、未学習状態から開始させるため、FIT/UNFIT 記録の初期値は一樣な乱数を設定した。

216 曲の合致度を算出することにより、合致度の大きい順にソートされたレコメンドリストを作成する。図 1 は、作成されたレコメンドリストにおける、ユーザが FIT と判断する曲の分布が、学習によって変化していく様子を示している。学習する度に合致度は変化するので、レコメンドリストにおける曲順が変化する。図 1 の縦方向にレコメンドされた 216 曲のリストを表示している。このリストされた曲の表示色で、ユーザが判断する FIT/UNFIT 曲を示している。図 1 では、ユーザが判断する FIT 曲を黒、UNFIT 曲を白で示した。

学習回数 65 付近までは上位から下位までユーザが判断する FIT 曲が一樣に分布しているが、それ以降の学習では、上位はほとんど FIT 曲で、UNFIT 曲は下位を中心に分布するようになり、この状態で安定した。このように、学習を繰り返すことにより合致度の上位にユーザが判断する FIT 曲が移動し、適切にレコメンドを行える状態が持続することが確認できた。

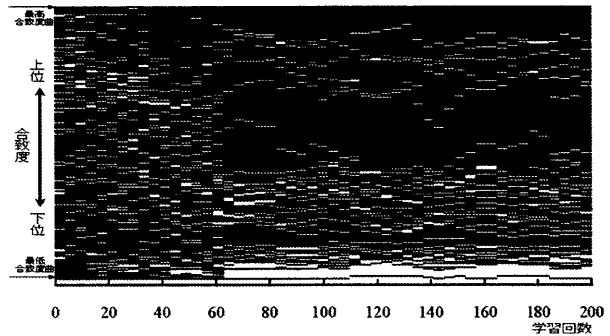


図 1 学習による FIT 曲分布の推移

図 2 に、検索特徴量である FIT 記録の平均と分散を示す。学習回数 65 付近で各平均は、ほぼ目標値になったことがわかる。

また、FIT 記録の初期値は一樣乱数のため、分散は非常に大きいですが、FIT 記録が形成されるに従い小さい値になり、学習回数 65 付近で、FIT 記録の値が平均値に集中してくるのことがわかる。このことは、学習がほぼ終了し、ユーザに対して最適なレコメンドができるように検索特徴量が正しく設定されたことを示している。これは、前述の FIT 曲分布の推移と整合が取れている。

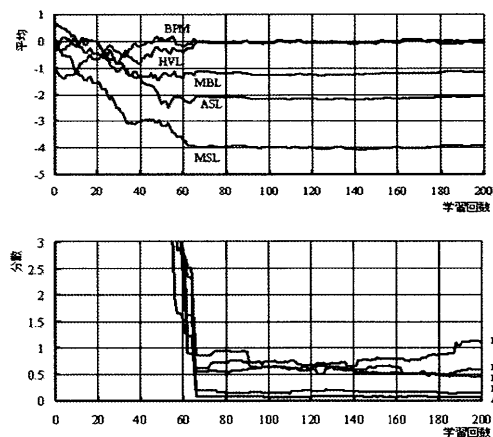


図 2 FIT 記録の平均、分散

5 まとめ

任意の曲をユーザの判断で学習することにより、検索特徴量が正しく設定され、ユーザの感性に合った曲のレコメンドを行う楽曲レコメンドシステムを実現した。

本システムは、音楽録音機能を有効に利用するためのアプリケーションとしてカーナビゲーションに実装されている。

今後は、学習方法の改善や、より最適な楽曲特徴量の抽出などを検討し、合致度の精度を更に向上させたい。また、楽曲レコメンドシステムを利用した、新たなエンターテイメントの創造を目指したい。

参考文献

[1] 萩山真一, “類似楽曲検索を目的とした楽音からの和音進行抽出手法,” FIT2003, 一般講演論文集, Vol.2, p245-246, 2003