

脳波解析に基づく曲中の感情変化を予測する楽曲推薦システム

Music Recommendation System that Predicts Affective Changes Based on Brain Wave Analysis

*1 杉本知仁 *2 福井健一 *2 森山甲一 *2 栗原聡 *2 沼尾正行
Toshihito Sugimoto Kenichi Fukui Koichi Moriyama Satoshi Kurihara Masayuki Numao

*1 大阪大学大学院情報科学研究科情報数理学専攻

Department of Information and Physical Sciences, Graduate School of Informartion Science and Technology, Osaka University

*2 大阪大学産業科学研究所

The Institute of Scientific and Industrial Research, Osaka University

Recently, the number of music on the Internet has increased rapidly. It becomes difficult to find music suited to person's feelings. In this paper, we propose the automatic music recommendation system that predicts affective changes. By using the chord progression and brain wave analysis, partial structures of the music that affects person's affect is extracted. This structures are combined with frame structures optimally and used for recommendation. Performances of proposed system are discussed based on a set of experiments to confirm whether the recommendation system can actually predict the target affective.

1. はじめに

一般に楽曲推薦システムが広く研究されている背景として、情報の肥大化が挙げられる。近年インターネットを通じてさまざまな情報が配信されており、流通する情報量は増加の一途をたどっている。音楽に関しても同様に流通する量が増え、自分の感性に合う音楽を探すことが困難になりつつある。この問題を解決するために、様々な楽曲推薦システムが研究されている。本研究ではこのような背景を踏まえ、さらに感性を操作するという目的を持って楽曲推薦システムの構築を行っている。現代ストレス社会において、自身の気分を主体的にコントロールする能力が求められている。音楽聴取によってユーザを悲しい状態から楽しい状態へ、さらにはストレスを強く感じている状態からリラックスした状態へとといった方向に導くことができれば非常に有意義であると考えられる。

従来研究では楽曲特徴量としてテンポや拍子、演奏楽器といった曲全体として抽出される枠組構造が使用されていた。特徴量として和音構造を取り入れている研究もあるが、メジャーコードやマイナーコードの割合などというように、平均化して枠組構造の一部として用いられている。このため推薦される楽曲は「全体として特定の感性を喚起する楽曲」のみであった。

そこで本研究では楽曲内での感情の変化を予測し、それを基に「曲中で感情が変化する楽曲」の推薦を行うシステムを提案する。枠組構造に加えて新しい楽曲構造として和音進行を取り入れ、さらに従来用いられていたアンケート方式の評価法 (SD法) ではなく、脳波解析法によって個人の感性を評価する。これにより、楽曲の部分構造と感性との関係性を学習することができ、曲中で感情が変化する楽曲が推薦可能となる。

本稿ではまず推薦の精度を向上させるため、枠組構造と和音進行が個人の感性に与える影響度について検証する。これらはどちらも感性に影響を与えるものであると考えられるが、その影響の強さは個人によって、また感性によっても異なると考えられる。そこで、学習の結果得られる枠組構造モデルと和音構

造モデルにユーザの感性に合わせて最適な重み付けを行い、それをもとに楽曲推薦を行う。システムの概要図を図1に示す。

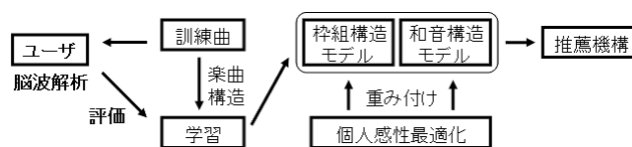


図 1: システム概要図

2. 関連研究

楽曲を個人に合わせて推薦するシステムは数多く提案されており、様々な特徴量が用いられている。特徴量として拍子、テンポといった基本的なもの他に、平均音高差や平均音長差、メジャーコードやマイナーコードの割合などを取り入れているもの [4] や、音符数等を考慮にいれているもの [5] もある。また、歌詞、楽曲情景、視聴状況を特徴量としてプレイリストを推薦する研究 [6] や、MIDI ファイルから特徴量を抽出する手法が多い中、楽音信号から特徴量を抽出して推薦する手法 [7] も研究されている。

3. 感性の抽出と学習

本研究では、ユーザからあらかじめ用意した訓練曲に 5 段階評価をつけてもらい、それを基に個人の感性に影響を与える楽曲構造を学習する。

3.1 楽曲特徴量

ユーザに提示する楽曲は一階述語表現で表される。楽曲から抽出される特徴量は全体構造から 10 種類、和音構造から 5 種類である。本研究では和音の表記方法の一つであるコードネームから和音の特徴量を抜き出して使用している。図 2 に本研究で用いた特徴量を示し、これらの特徴量を用いて実際に楽曲を一階述語表現したものを図 3 示す。

連絡先: 沼尾正行, 大阪大学産業科学研究所

〒 567-0047 大阪府茨木市美穂ヶ丘 8-1,

Tel:06-6879-8425 Fax:06-6879-8428

E-mail: numao @ ai.sanken.osaka-u.ac.jp

全体構造	
Genre	曲のジャンルを示す。Pops, Rock, Jazz等がある。
Key	調の中心であり、音階の起点となる音を表す。
Scale	曲の音階(1オクターブ12音内、その曲の中心となる7音)を表す。
Tonality	長調か短調かを表す。
Tempo	曲の速さを示す。
Time Signature	曲の持つ基本的なリズムの定義。4分の4拍子が一般的である。
Melody Instrument	曲の旋律(メロディー・パート)を演奏する楽器。
Melody Category	旋律楽器の分類。本研究では Piano系, Sax系等を用意した。
Chord Instrument	曲の和音(コード・パート)を演奏する楽器。
Chord Category	和音楽器の分類。本研究では Piano系, Guitar系等を用意した。
和音構造	
Key	その和音が属する楽曲の主音
Root degree	その調における度数を表す。
Root name	和音の根音を表す。
Type	和音の構成する音の数および構成音ごとの関連を表す。
Tension	和音に付加的に用いられる音を表す。

図 2: 使用した楽曲特徴量

```
music (4,
song_frame ( rock, a, c, minor, four_four, allegretto,
             sawtooth_lead, lead, distortion_guitar, guitar),
[
  chord( a_minor, vii , g, major, null, null ),
  chord( a_minor, iv , d, minor, null, null ),
  chord( a_minor, iv , d, minor, null, null ),
  chord( a_minor, vi , f, major, null, null ),
  :
  chord( a_minor, vi, f, major, null, null )
]).
```

図 3: 楽曲の一階述語表現

3.2 脳波解析の基づく楽曲評価

個人の感性に影響を与える部分構造を学習するため、楽曲提示時の被験者の脳波を感性スペクトル解析法 [2] を用いて解析する。

従来使用されていた SD 法は楽曲を全体で評価することしかできなかったが、脳波は時系列ごとの連続数値データとして得られるため、これを解析することにより楽曲提示中の感性の変化を知ることができる [1]。しかし、被験者によって個人差があるため、全被験者の感性に対して一律に同じ閾値を設けて評価付けすることができない。そこで感情の振幅や定常状態といった個人差を吸収するため、被験者が提示された楽曲に対して図 4 のように上位 10% を評価 5、次の 20% は評価 4 というように評価付けを行った。これによりどの被験者も 1~5 の評価が [1 : 2 : 4 : 2 : 1] の割合で付与される。この評価割合は SD 法を用いた従来研究 [3] を参考にして定義した。

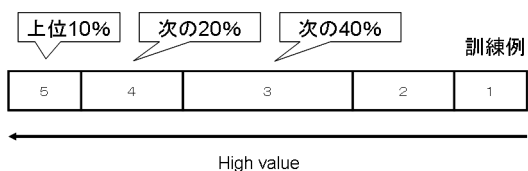


図 4: 5 段階評価への切り分け

今回は感性に影響を与える部分構造を学習するため、和音進行をフレーズごとに 5 段階評価する。フレーズとは、楽曲構造で「小楽節」つまり 4 小節の単位のことを指すものであり、このように音楽的に意味のある単位で楽曲を区切ることで楽曲の流れを分断することなく楽曲構造を切り分けることが可能となる。枠組構造の評価は各フレーズに対する評価の平均値を与える。図 5 に評価値付けイメージを示す。

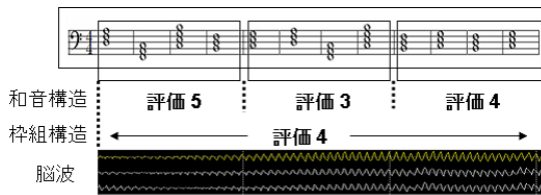


図 5: 評価付けイメージ

3.3 帰納論理プログラミングによる学習

本システムでは、帰納論理プログラミングの一つである FOIL を利用してレベル別学習を行う。感性の抽出を高い精度で行えるかどうかは、先で述べた音楽知識表現と FOIL が学習結果として獲得する目標述語の定義に依存する。獲得する目標述語は全体構造に関する「frame」と和音進行に関する「chords」であり、これらの述語は別々に学習される。

目標述語に対して、被験者の評価値が 5 の訓練例を正例として学習した結果得られる述語集合をレベル 5 とし、評価が 4 以上のものを正例として得られる述語集合をレベル 4 とする。レベル 1, 2 の述語集合は評価が 1 以下、2 以下のものを正例として学習した結果得ることとする。枠組構造モデルと和音構造モデルはこれらの述語集合を合わせたものとして得られる。図 6 に得られる両構造の感性「Stress」に関するモデルを示す。それぞれの述語の第二引数はレベルを表す。

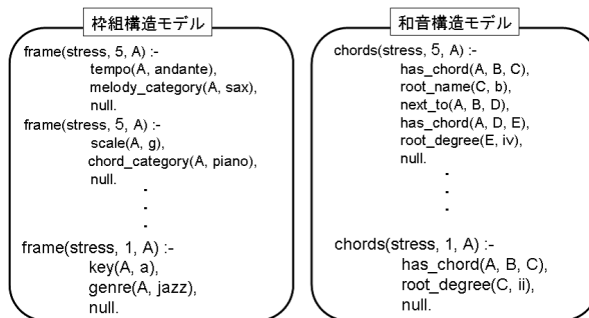


図 6: 述語モデル例

4. 個人感性最適化

枠組構造と和音構造のそれぞれが個人の感性に与える影響の強さを調べ、それによって最適な重み付けを行う手法について述べる。この手法は交差検定を用いて行う。まずユーザから評価値を得ている訓練例を 5 セットに分け、1 セットをテスト例とし、残りの 4 セットで学習を行う。この結果得られる枠組構造モデルと和音構造モデルのそれぞれからテスト例に対する予測評価値を得る。

4.1 評価値予測手法

モデル内の述語がテスト例に存在した場合、評価値として当てはまった述語のレベルを与える。複数の述語が当てはまった場合はそれらのレベルの平均値を、当てはまる述語がなかった場合には評価 3 を与える。この結果、楽曲 X に対して枠組構造モデルと和音構造モデルのそれぞれから得られた予測評価値を $F(X)$, $C(X)$ とする。このとき、和音進行に関してはテスト例をフレーズごとに評価するため、テスト例 1 曲に対して $F(X)$ と $C_1(X)$, $C_2(X)$... $C_n(X)$ (n: フレーズ数) が得られる。

4.2 個人感性最適重みの獲得手法

以下の式から楽曲に対する最終的な予測評価値 $P(X)$ を得る。 w_f と w_c は個人の各感性に与えられる両モデルに対する重みである。

$$P(X) = F(X) * w_f + \left(\sum_{k=1}^n C_k(X)/n \right) * w_c \quad (1)$$

上式で付与する重みの比率を段階的に変化させたときの予測値 $P(X)$ と実評価値が一致した場合の精度を検証し、最も精度のよかった際の重みを「個人感性最適重み」として得る。この重みは個人の感性ごとに得られる。今回は両者の重みを [5:0] ~ [5:5] と [1:5] ~ [4:5] まで変化させて精度を検証した。図7、図8に被験者A、Bの予測精度の推移を示す。

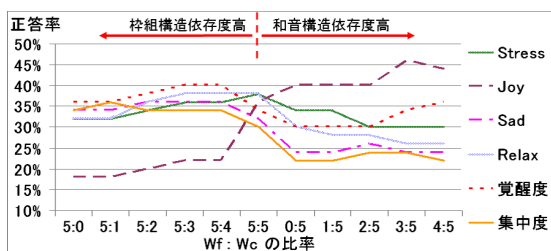


図7: 被験者Aの予測精度

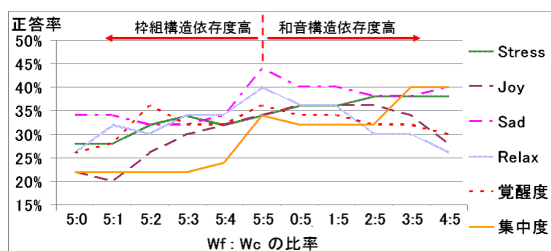


図8: 被験者Bの予測精度

被験者Aは“Joy”に関しては和音構造に強く依存し、他の感性については枠組構造に依存していることがわかる。また、被験者Bについては全体的に和音構造に依存していることがわかる。この結果、個人によって影響を強く受ける構造に違いがあることがわかり、また同じユーザでも感性によってどちらの構造に強く影響されるか異なることがわかった(個人最適化手法をSD法による評価に対して適用した場合についても同様の結果が得られることが示されている[3])。

5. 楽曲推薦手法

個人感性最適重みを用いて楽曲推薦を行う。推薦される楽曲は「(I) 特定の感性を強く喚起する楽曲」と「(II) 曲中での感性が大きく変化するような楽曲」の2種類である。まず、全推薦用楽曲に対する予測評価値 $F(X)$ 、 $C(X)$ を算出しておく。

(I) 特定の感性を強く喚起する楽曲の推薦

個人感性最適重みを上述の式(1)の重み w_f と w_c に適用し、評価値 $P(X)$ を算出する。こうして全曲に対して一意に付けられた評価値 $P(X)$ が被験者に対する推薦度となる。例えば感性“楽しさ”であれば $P(X)$ の値が最も大きい曲は正方向の形容詞“楽しい”を、 $P(X)$ の値が最も小さい曲は負方向の形容詞“つまらない”を最も強く喚起するであろう曲となる。

(II) 曲中での感性が大きく変化する楽曲の推薦

以下の条件式を全て満たす楽曲が推薦される。

$$\begin{cases} F(X) = 3 \\ C_i(X) > 3 \\ C_m(X) < 3 \end{cases} \quad (\exists l, m)$$

この式を満たす楽曲は感性“悲しみ”を例に考えると、あるフレーズでは正方向の“悲しい”が喚起されるが、あるフレーズでは負方向の“嬉しい”が強く喚起される曲といえる。条件をより強くすることによって、曲中の感情の変化がより激しいものを推薦することができる。

図9に推薦される楽曲例について示す。

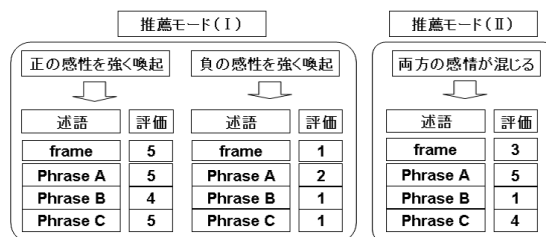


図9: 推薦される楽曲

6. 被験者実験

被験者13人に訓練曲50曲を提示し、脳波解析法で4感性(Stress, Joy, Sad, Relax)と2指標(覚醒度, 集中度)対して5段階評価を付けを行った。推薦用楽曲としてあらかじめ124曲用意してある。

評価値予測精度の検証

次に個人感性最適化のために訓練例を10曲1セットとして5セットに分割して交差検定を行い、予測評価値 $F(X)$ と $C(X)$ を算出した。図10にそれらが実評価値と完全に一致した場合の予測精度を全ユーザで平均したものを感性ごとに示す。

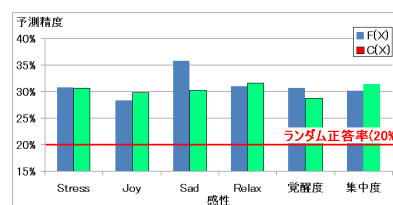


図10: 感性別予測評価精度

この結果、 $F(X)$ 、 $C(X)$ それぞれを単独で用いた場合、ランダム予測精度に比べ全体的に10%程度高い精度で予測できることがわかる。また、感性“Sad”は $F(X)$ を用いた場合の予測精度が他の感性と比較して5%程度高いことがわかった。

個人感性最適重みの検証

得られた $F(X)$ と $C(X)$ を用いて個人感性最適化を行い、各個人の各感性に対して最適な重みを算出した。このとき、全被験者の全感性に対する個人感性最適重みの度数分布を図11に示す。この結果から、予測が枠組構造のみに依存する割合は10%以下であり、90%以上の割合で和音構造を加味した方が予測精度は向上することがわかる。

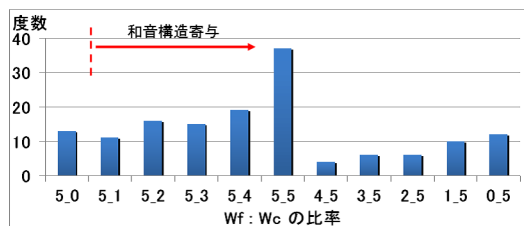


図 11: 個人感性最適の選出率

次に、枠組構造 $F(X)$ のみで予測した場合と和音構造を取り入れて最適重みを用いた場合の予測精度の比較を行い、図 12 に示す。この結果から最適化手法を用いることにより、全感性で 5% 以上予測精度が向上することがわかり、本重み付け手法が有効に働いていることが示された。

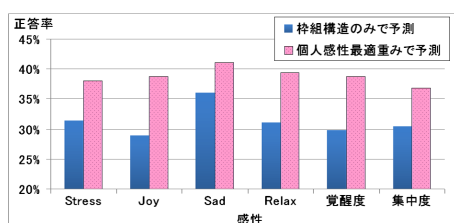


図 12: 個人感性最適化の有用性

推薦 (I) の検証

推薦 (I) のため、各感性について推薦度の上位 3 曲と下位 3 曲をそれぞれ選んだ。推薦の精度を確かめるため「正と負の感性をそれぞれ喚起する曲に対する評価の平均値に差がない」という仮説のもと t 検定を行った。感性別に検定したところ、有意水準 5% で “Joy” と “覚醒度” に関して有意差が認められた。“集中度” を除く他の全ての感性性についても統計的に有意とは言えないが評価に正の差が見られた。次に、被験者別に検定を行ったところ、4 人の被験者に対して有意水準 5% 有意差が認められ、さらにそのうち 3 人の被験者に関しては有意水準 1% でも有意差が認められた。

推薦 (II) の検証

推薦 (II) のために条件を満たす楽曲を被験者に提示して、曲中の感性の変化を調べた。感性の強さの変化（正方向、不変、負方向）を予測したものと実際の感情の変化が一致しているか検証した結果を図 13 に示す。この結果、特に “Joy”、“Relax” に関しては高い精度で曲中の感情変化を予測できていることがわかった。

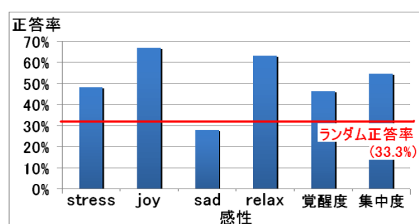


図 13: 感情変化予測率

7. 考察

本研究では脳波解析を基に抽出した感性情報を用いて楽曲を推薦するシステムを構築した。これは従来の SD 法を基にした推薦システムとはまったく異なるものである。なぜなら従来の推薦システムは音楽が持っている性格を受け止め、それを評価するというものであったが、本研究では例えば “悲しい” という性格を持っている曲が提示されても、それに対して実際に悲しいという感情が喚起されなければ “悲しい” と評価されない。つまり、従来研究はある感性を喚起させるであろう性格を持つ楽曲を推薦するシステムであり、本システムは実際にある感性を喚起させる楽曲を推薦するシステムであるといえる。このため推薦を行うことは従来研究に比べて非常に困難であると考えられる。

このような中で、推薦 (I) を行い、感性別、被験者別ともに部分的ではあるが統計的に有意な結果を得ることができた。有意差の得られた 2 感性は枠組構造と和音構造の両構造から等しく影響を受ける場合が多く、そのため推薦に用いる情報量が多くなり、推薦精度が高くなったと考えられる。

また、推薦 (II) に関しては “Sad” を除いた全ての感性では曲中の感情変化を予測することが可能であった。“Sad” は図 10 からわかるように枠組構造に大きく依存するため楽曲内で感情の変化が少なく、予測することが困難であったと考えられる。

8. おわりに

本研究では楽曲特徴量として新たに和音進行を取り入れ、脳波解析によって個人感性を抽出することにより、曲中での感情変化を予測し、それをもとに楽曲推薦を行うシステムを構築した。これにより、感性という曖昧で捉えがたいものが楽曲提示中にどのように変化するかを予測することが可能であることを示した。

参考文献

- [1] Toshihito Sugimoto, Roberto Legaspi, Akihiro Ota, Koichi Moriyama, Satoshi Kurihara and Masayuki Numao, Modelling Affective-based Music Compositional Intelligence with the Aid of ANS Analyses, *Knowledge-Based Systems*, 21 (3), pp. 200-208 (2007)
- [2] Toshimitsu Musha and Yuniko Terasaki, Feature extraction from EEGs associated with emotions, *Artificial Life and Robotics*, Vol.1, pp. 15-19, 1997.
- [3] 杉本 知仁, 西川 敬之, 森山 甲一, 栗原 聡, 沼尾 正行, 自動推薦システムの構築における感性と楽曲特徴量の関係性に関する考察, 第 83 回 知識ベースシステム研究会, 2008.
- [4] 竹川和毅, 土方嘉徳, 西田正吾, 内容に基づく音楽データの探索・推薦システム, *人工知能学会論文誌*, Vol. 23 No. 5, pp. 330-343 (2008)
- [5] 池添剛, 梶川嘉延, 野村康雄, 音楽感性空間を用いた感性語による音楽データベース検索システム, *情報処理学会論文誌*, Vol. 42, No. 12, pp. 3201-3212, (2001).
- [6] 梶克彦, 平田圭二, 長尾確, 状況と嗜好に関するアノテーションに基づくオンライン楽曲推薦システム, *情報処理学会研究報告 [音楽情報科学]*, Vol. 2004, No. 127(20041212), pp. 33-38, (2004).
- [7] 小田川智, 児玉泰輝, 我山真一, 鈴木康悟, 松下文雄, 塩田岳彦, 楽曲レコメンドシステム, *電子情報通信学会技術研究報告*, Vol.06, No. 160(20060707), pp. 49-53, (2006).