

# 脳波の解析に基づく音楽に関する感性推定器

Estimating Affective Based on Brain Wave Analysis on Music

\*1山野悠  
Yu Yamano

\*1Paul Inventado

\*1Rafael Cabredo

\*2Roberto Legaspi

\*2福井健一  
Kenichi Fukui

\*2森山甲一

Koichi Moriyama

\*2栗原聡

Satoshi Kurihara

\*2沼尾正行

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, to gain a better understanding of humankind, many researches about sensitivity are made. In the area of music, they conduct extensive researches on the influence of sensitivity from music. They have researched getting information about sensitivity by analysing electroencephalogram. However, greater part of sensitivity which is got by analysing electroencephalogram is not known. In this paper, to come out that part, we analysis electroencephalogram corresponding to sensitive when subjects listen music. Then, we formulate estimating instrument of sensitivity by analysing electroencephalogram with machine learning. This lead to make automatic music composition system which reflects individual sensitivity more accurately.

## 1. はじめに

近年、人間に対する理解を深めるため、感性を扱う学際的な研究が盛んに行われている。音楽に関する分野でも、音楽が人間の感性に与える影響についての研究や、人間の感性を用いた音楽検索等、感性に関連する幅広い研究が行われ始めた。本研究でも、作曲において人間の感性を扱う手法を提案し、開発を行ってきた。[2]

楽曲には、「楽しい」や「悲しい」といったような感情を喚起する要素が含まれているが、ある人にとって「悲しい」楽曲が別の人にとってもそうであるとは限らない。これは、人によって楽曲に対する感じ方、つまり感性が異なる為である。このように感性とは各人に固有のものであり、単純に定義できるものではない。その為、個人の感性を扱うシステムを構築することは困難である。しかしながら、作曲という創造的な活動においては、人間の感性が重要な役割を担う為、個人の感性を考慮に入れた作曲を行うシステムの構築は非常に有用であると考えられる。

感性を学習するうえで、脳波解析による手法は有力なものであり、現在まで脳波解析によって感性情報を取得する様々な研究が行われてきた。脳波と感性の対応関係は、武者氏らの研究[1]により、「ストレス」「喜び」「悲しみ」「リラックス」の4つの感性に関して既知となっているが、その他未知である部分が多い。

そこで本研究では、未知である部分を解明すべく、楽曲を聴いている間の脳波とその間に抱いている感性を対応させ、機械学習により脳波データからの感性の推定器の構築を行った。これにより、個人の感性をより適切に反映させた自動作曲が期待できる。

## 2. 関連研究

音楽と感性に関する研究として、音楽の旋律が心理的・生理的反応に及ぼす影響をSD法と脳波解析を用いて調べた研究、音楽を聴いたときに起きる感情的反応について調べた研究[4]などがある。

これらの研究により、音楽が感性に影響を与えることが明らかとなった。

## 3. 感性情報の取得

感性は非常にあいまいなものであり、情報を取得するのが非常に困難である。感性を取得する方法として、感性をアンケート形式で評価するSD法 (semantic differential method) や、RussellのAV空間を用いる方法、生理信号から感性を抽出する方法などが提案されている。以下でSD法とRussellのAV空間について紹介する。

### 3.1 SD法

感性評価の方法の一つとして用いられているSD法は、形容詞や形容動詞の感性語対をアンケート形式で回答することにより感性を数値化するものである。SD法の例を図1に示す。評価の尺度は5ランク、もしくは7ランクが一般的である。

SD法を用いての感性評価は、日常生活での評価とシチュエーションが非常に類似しているため、被験者にとって理解しやすいというメリットがある。しかし、1つの楽曲に対し、1つのアンケートで評価するため、時間を追った感性評価が不可能であるというデメリットがある。

### 3.2 RussellのAV空間

本研究で使用した240曲のデータセット[3]には、あらかじめ行われていた被験者実験により1秒ごとのAV空間の座標値が与えられている。RussellのAV空間は、図2のようにenergetic-silent(Arousal)、positive-negative(Valence)の二軸

連絡先: 山野悠, 大阪大学 産業科学研究所沼尾研究室,  
〒567-0047 大阪府茨木市美穂ヶ丘8-1,  
Tel:06-6879-8426 Fax:06-6879-8428  
E-mail:yamano@ai.sanken.osaka-u.ac.jp

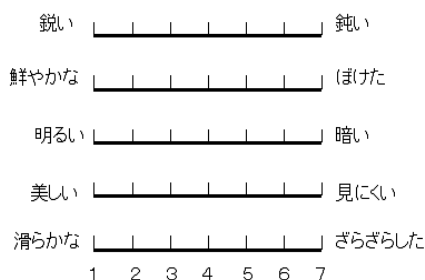


図 1: SD 法の例

(AV 軸) から成る二次元平面で人間の感性を表現するものである。AV 値を計算することで、楽曲を AV 空間上の点として表現することが出来るため、楽曲と感性とを対応付ける際によく用いられている。データセットで与えられていた AV 値を図 2 に示す。

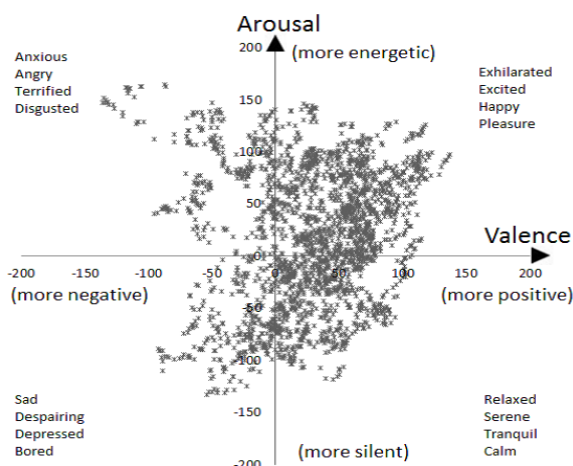


図 2: 楽曲の AV 値分布

Russell の AV 空間を用いることで、feeltrace や moodswing[3] といった感性情報取得のためのソフトウェアによってインターネット上で世界中から感性データを集めることも可能となっている。Russell の AV 空間を用いての感性評価は時間を追った感性の推移が見て取れるため、1 つの楽曲から得られる情報量が非常に多い。また、提示する楽曲が少なくても良いため、SD 法を用いる場合に比べて被験者にかかる負担が少ない。本研究で使用した楽曲の AV 値は、Moodswing を用いてデータを取得している。

### 3.3 脳波の解析に基づく感性の取得

脳波とは、脳の活動電位差を記録したものである。活動電位は数  $\mu V$  の微弱な電位であり、その変動により様々な周波数成分を含む合成波の形を示す。脳波は、図??に示すように大きく 4 つの周波数帯域に分類される。[6]

波は、安静・閉眼・覚醒時に後頭部及び頭頂部から顕著に出現し、脳の活動状態を示す指標として用いられる。開眼や緊張時には波は減衰し、外部からの刺激に対して脳の活動状態が活発になると波が多く表れるようになる。一方、覚醒の度合いが小さく脳の活動状態が低下すると波が多く表れるよう

周波数	波形	脳の活動状態
$\delta$ 波(0.5~4Hz)		睡眠状態
$\theta$ 波(4~8Hz)		脳の機能低下
$\alpha$ 波(8~13Hz)		安静、閉眼、覚醒した状態
$\beta$ 波(13Hz~)		開眼、緊張、興奮した状態

図 3: 脳波の活動状態

になる。このように脳波の周波数解析からは、脳の活動状態や覚醒の程度を調べることができる。また、波・波・波の増減と、喜び・悲しみ・怒りといった感性に対応関係があることが武者・小川ら [?, 7] の研究で証明されており、脳波データから感性の推定が可能であると言える。

脳波解析による感性獲得は、Russell の AV 空間を用いた時と同じく、時間を追った感性の推移が見て取れるため、提示する楽曲が少なくても良く、被験者にかかる負担が少ない。しかし、脳波は外界からの影響を受けやすく、楽曲以外から受ける影響を反映してしまうことや、被験者の体調・精神状態によっても左右されてしまうことがデメリットとなる。そのため被験者実験には細心の注意を払う必要がある。

## 4. ニューラルネットによる推定器の構築

### 4.1 ニューラルネットワークの概要

ニューラルネットワークは、脳神経系における情報処理の方式を人工的・工学的に模倣したもので、並列分散的で自己学習的な能力をもつシステムである。その原理は、非線形特性をもつニューロン阻止を多数結合した人工ニューラルネットワークを用いて並列分散的に計算を行うことである。ニューラルコンピューティングでは、ニューラルネットワークの結合状態や構造およびダイナミクスを、学習と呼ばれる機能に基づき、問題の解を与える情報構造のシステムに適応変化させることが重要な手法となっている。このようにニューラルネットとは、脳神経系回路網を人工的・工学的に模した非線形特性をもつ並列分散処理的なシステムであり、パターン認識や制御システムなどへ応用されている。

### 4.2 階層型ニューラルネットワーク

ニューラルネットはその構造から、図 4 のような階層的なネットワークと図 5 のような相互結合のある非階層的なネットワークに分類して考えることができる。階層型ネットワークは、図 4 のように、ユニットが複数の階層をなすようになり、入力層から出力層へ向かう一方方向の結合のみが許されるネットワークである。一方、相互結合ネットワークは、図 5 のように、任意のふたつのユニット間に双方向の結合を許すようなネットワークである。

### 4.3 多層パーセプトロン

本研究で使用した多層パーセプトロンの概要について記す。多層パーセプトロンは単純パーセプトロン [8] を層状に繋ぎ合わせたネットワークである。図 6 に単純パーセプトロンの例を、図 7 に多層パーセプトロンの例を示す。

例えば、中間層が 1 層のネットワークでは、入力信号  $x$  に対する出力信号  $z = (z_1, \dots, z_K)^T$  は、

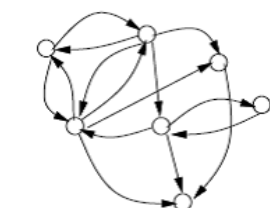
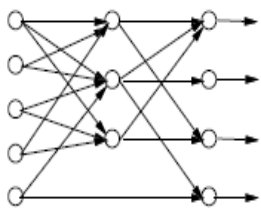


図 4: 階層型ネットワーク

図 5: 相互結合ネットワーク

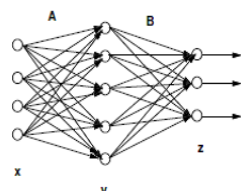
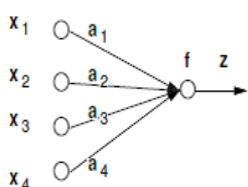


図 6: 単純パーセプトロンの例

図 7: 多層パーセプトロンの例

$$\zeta_j = \sum_{i=1}^I (a_{ij}x_i) + a_{0j} \quad (1)$$

$$y_j = f_{hidden}(\zeta_j) \quad (2)$$

$$\eta_k = \sum_{j=1}^J (b_{jk}y_j) + b_{0k} \quad (3)$$

$$z_k = f_{out}(\eta_k) \quad (4)$$

のように計算される。ただし、 $a_{ij}$  は、 $i$  番目の入力から中間層の  $j$  番目のユニットへの結合荷重であり、 $b_{jk}$  は、中間層の  $j$  番目のユニットから出力層の  $k$  番目のユニットへの結合荷重である。 $a_{0j}$  および  $b_{0k}$  は、それぞれ、中間層の  $j$  番目のユニットおよび出力層の  $k$  番目のユニットのバイアスである。また、 $f_{hidden}$  および  $f_{out}$  は、それぞれ中間層のユニットの入出力関数および出力層のユニットの入出力関数である。中間層のユニットの入出力関数としては、普通ロジスティック関数が使われる。出力層のユニットの入出力関数は、利用目的に応じて決められる。例えば、関数近似のためにネットワークを利用する場合には線形関数が使われ、パターン認識に利用する場合にはロジスティック関数とすることが多い。

多層パーセプトロンを用いることで表現可能な関数を考えた時、非常に多層パーセプトロンの汎用性が高いことがわかる。例えば、中間層のユニットの入出力関数  $f_{hidden}$  が

$$f_{hidden}(t) = \begin{cases} 1 & (t \rightarrow +\infty) \\ 0 & (t \rightarrow -\infty) \end{cases} \quad (5)$$

のような性質を持つ非線形の連続なシグモイド関数であり、出力層の入出力関数が線形関数のとき、中間層が 1 層の多層パーセプトロンによって、任意の連続関数が近似可能である。もちろん、任意の連続関数を近似するためには中間層のユニットの数を非常に多くする必要があるが、この結果は、多層パーセブ

トロンを入出力関係を学習するために使うには、理論的には中間層が 1 層のみのネットワークで十分であることを示している。

## 5. 実験

### 5.1 実験内容

Jacquelin A. Speck らの研究 [3] で使用されていたデータセットに含まれている、1 秒ごとの Arousal-Valence 値が付与されている 240 曲の楽曲に対応する脳波を得るために実験を行った。実験では 240 曲、各楽曲につきデータセットと同じ 15 秒間を被験者に提示し、その間の脳波を計測した。脳波計は、取り付けが非常に容易である、Emotiv EPOC という 14 個の電極からなる脳波計を使用した。Emotiv EPOC を図 8 に、その電極配置を図 9 に示す。



図 8: Emotiv EPOC

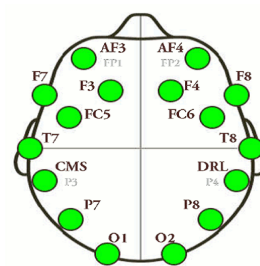


図 9: 電極配置

その後、得られた脳波データとデータセットの AV 値を交差検定により、多層パーセプトロン (MLP) の推定精度を検証した。MLP の入力値を実験で得られた 14 極の脳波電位の値、出力値を Arousal 値もしくは Valence 値とし、Arousal・Valence それぞれに関して 1 秒ごとの推定精度を 10-fold cross-validation により求めた。入力値・出力値共に、各秒ごとの被験者全員の平均値を用いている。

### 5.2 Cross-Validation による検証

得られた脳波データから MLP の推定精度を検証するために、統計的検証手法の一つである cross-validation (交差検証法) を用いて検証を行う。交差検証法とは、まずデータセット  $S$  を均等に  $n$  個の部分集合 ( $S_1, S_2, \dots, S_n$ ) に分割し、次に部分集合  $S_i$  を取り除いた残りの集合  $S' (= S - S_i)$  を訓練集合として学習を行い、そこで得られたモデルに対して  $S_i$  を用いて予測を行うとうものである。これを  $n$  回繰り返して、学習・予測率の平均を得ることで、データセット全体にわたる推定の性能について検証を行うことができる。本研究では  $n = 10$  (10-fold Cross-Validation) を用いて検証を行った。

### 5.3 被験者手順

測定は以下の流れにしたがって行なった。

1. 30 秒の閉眼状態
2. 15 秒の閉眼状態
3. 曲 (15 秒) の提示

以下、2・3 を全曲聞き終わるまで繰り返した。

実験は 19~25 歳の身体的・精神的に健康な男女 8 名ずつ、計 16 名の被験者に対して行った。脳波は様々な外界の影響を受けやすいので、実験環境には細心の注意を払い、静かな部屋で、閉眼状態で実験を行った。また脳波計は振動にも反応するため、脳波計測中に被験者が体を動かさないよう注意した。

## 5.4 実験結果

実験で得られた脳波データの一部を図 10 に示す。得られた脳電位を入力値、出力値を Arousal 値もしくは Valence 値とし、weka[5] を用いて 10-fold cross-validation により推定精度を検証した結果、推定精度は表 1 のようになった。表 1 からわかるように、Arousal 値の推定精度はおよそ 70%、Valence 値の推定精度は 77% という値が得られた。

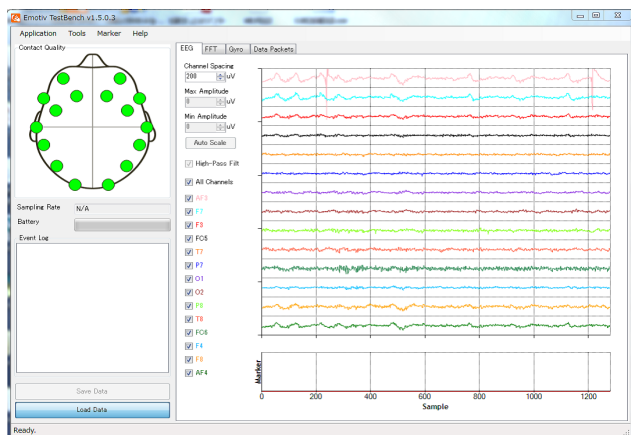


図 10: 脳波データ

表 1: Arousal・Valence 値の推定精度

中間層のニューロン数	Arousal 値	Valence 値
1	68.6508%	74.5238%
2	68.4524%	75.7143%
3	69.0079%	76.5079%
4	68.2937%	76.1508%
5	69.2063%	75.873%
6	69.0079%	75.7937%
7	68.6508%	75.8333%
8	68.7302%	76.1905%
9	69.0079%	76.1111%
10	69.3254%	76.5079%
11	69.7619%	76.5873%
12	69.0873%	76.1905%
13	69.8016%	76.5873%
14	69.9206%	77.1032%
15	69.2063%	77.6984%
16	70.3571%	77.2222%
17	69.2063%	77.0635%
18	69.2857%	77.0635%
19	69.8413%	76.8254%
20	69.5238%	77.4603%

## 6. 考察

実験結果で記したとおり、推定精度が 50% を超えているため、推定器としての役割を果たしているといえる。しかしながら、推定精度向上の余地があるため、今後さらなる研究が必要

である。今回の実験で推定精度が低かった原因として、静電遮蔽を行っておらず交流障害（ハム）を含んでいる可能性がゼロではないこと、楽曲情報以外の外部からの影響を完全に除去できなかったことなどが挙げられる。また、脳電位の値をフーリエ変換等の作業をすることなく入力値として用いているが、フーリエ変換等の変換を行った後、入力値として用いることで推定精度が向上する可能性も否めない。今後の課題として、以上のような事柄を改善する予定である。

## 7. おわりに

本研究では脳波データと Arousal・Valence 値との対応関係を 10-fold cross validation により検証することにより、脳波と感性の対応関係を明らかにする推定器の構築を行った。推定精度は Arousal 値がおよそ 70%、Valence 値がおよそ 77% という結果が得られた。両感性共に 50% を上まわる結果となっているため、本研究の目的は達成できたと言える。脳波に含まれるノイズを除去できていないせいか、現段階では推定精度はあまり良いものとは言えないため、今後改良を加え、推定精度を向上させる必要がある。また、MLP 以外の方法のクラス分類の推定精度の検証も行う予定である。

## 参考文献

- [1] T. Musha, H. Terasaki, H. A. Haque and G. A. Ivanitsk, " Feature extraction from EEGs associated with emotions ", *Artificial Life and Robotics*, Vol.1, pp.15-19, 1997.
- [2] 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 ", 人工知能学会 第 83 回知識ベースシステム研究会資料 (SIG-KBS-A802), pp.33-36, 2008.
- [3] Jacquelin A. Speck, Erik M. Schmidt, Brandon G. Morton and Youngmoo E. Kim, " A COMPARATIVE STUDY OF COLLABORATIVE VS. TRADITIONAL MUSICAL MOOD ANNOTATION ", Proc. 12th International Society for Music Information Retrieval Conference, pp.549-554, 2011.
- [4] Suzanne Filipic and Emmanuel Bigand. " Emotion and cognition in music: Which comes first? ", Proceeding of the 5th Triennial ESCOM Conference, p.159, 2003.
- [5] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten, " The WEKA Data Mining Software: An Update ", *SIGKDD Explorations*, pp.10-18, 2009.
- [6] 日本生理人類学会計測研究部会, 人間科学計測ハンドブック, pp.331-376, 技報堂出版, 1996.
- [7] 小川嗣夫, " 過去の情動経験想起に伴う脳波の測定 " 人間文化研究, Vol.3, pp.101-109, 2000.
- [8] Rumelhart, D.E., Hinton, G.E. and Williams, R.J., " Learning representations by backpropagating errors ", *Nature*, Vol.323-9, pp.533-536, 1986.