

人工知能とユーザ、聴衆のコラボによるコンテンツの生成



技術解説

沼尾 正行*

Content Creation in Collaboration among AI, Users and Audience

Key Words : Artificial Intelligence, Machine Learning, Music, Affective Computing, Genetic Algorithm

1 はじめに

人工知能技術が長足の進歩を遂げ、チェス、将棋、囲碁において人を凌駕し、自動運転の実用化も近づいたことで、技術的特異点まで議論されるようになってきた。その中で人工知能のフロンティアと考えられるのは、人の感情の扱いと小説や音楽などの創造行為である。人工知能による小説の著作や Google による自動作曲が話題になっているが、創作の種は人が仕込んできた。本稿では、人の与える創作の種を減らして、柔軟性と自律性を上げることを考える。

創作の目的は受容者（音楽の場合、聴き手）に感動を引き起こすことにある。この考え方に基づき、聴き手の反応を見ながら創作を行う。聴き手の反応、すなわち、感情を測定するのである。さらに、いちいち聴き手にお伺いをたてずに済むように、今作りつつある新たな曲に対する聴き手の反応を予測する。

これらを行うため、音楽を対象とした感情の測定技術と予測技術、それらをベースにした自動作曲について論じる。

既存の楽曲を学習することにより、新たな曲を作る手法として、たとえば、モーツァルトの曲からモーツァルト風の曲を自動作曲することが考えられる。たとえば、人工知能における機械学習の分野では、楽曲の構造を学習する手法がある [10]。筆者も既存

の曲の和音進行を学習することにより、別の曲を編曲するシステムを開発したことがある。これらのシステムは、楽曲から機械学習を行うが、その曲の聴き手が何を感じたかは全く考慮していない。聴き手の反応を無視しているのである。

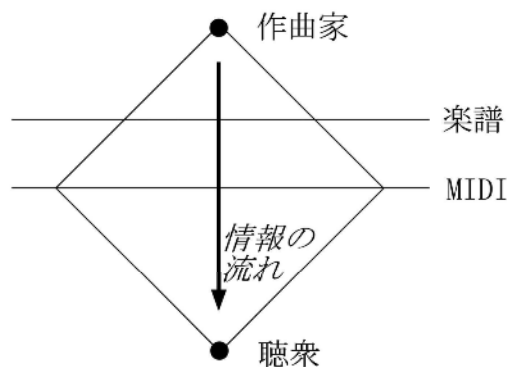


図1: 情報の流れとオーサリング

音楽には、作曲家が楽譜を書き、演奏家が演奏して、聴き手が聞くという、作曲家から聴き手への図1に示すような「情報の流れ」[2]がある。楽譜や演奏情報（コンピュータ演奏なら MIDI データ）は、その流れの一断面を表示するものである。したがって、楽譜の断面や MIDI データの断面を通過する情報を調べれば、楽譜や演奏の各部分が喚起する感性を知ることができる [3]。人間の感性は、非常に複雑なものであり、各個人に依存するし、時代背景を含む状況にも影響を受けやすい。したがって、「感性は何か」ということを追究するのではなく、「感性に影響を与えるような音楽構造は何か」を調べることで、個人の聴き手、もしくは聴衆の感性に合わせた作曲を行うことを試みるわけである。以上のような考え方にに基づき構築したシステムについて述べる。

このシステムでは、図2に示すようにまず個々の



* Masayuki NUMAO

1959年2月生
東京工業大学 工学部 電気電子工学科
1982年卒、大学院理工学研究科 情報工
学専攻博士課程 1987年修了
現在、大阪大学 産業科学研究所 教授
工学博士 人工知能、機械学習
TEL : 06-6879-8425
FAX : 06-6879-8428
E-mail : numao@sanken.osaka-u.ac.jp

聴き手に幾つかの楽曲を聞いてもらい、各々の楽曲への感じ方を調べる。それにより、各聴き手に印象を引き起こす音楽構造を作曲用ルールとして抽出する（モデル化）。作曲は、抽出した作曲用ルールにより、音楽構造を組み合わせて音楽を生成することにより、指定した印象の曲を生成する（創作）。以下、音楽構造の抽出、作曲の手法、および得られた曲を被験者実験により評価した結果について述べる。

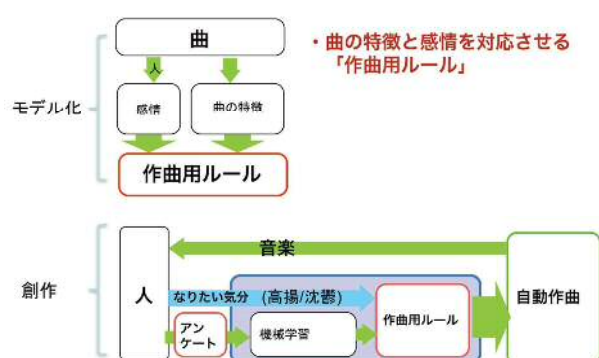


図2: 聴き手の印象に基づく創作

2 アンケートを用いた音楽構造の抽出

この節では、楽曲への感じ方を意味差分（SD）法によるアンケートを用いて調べる手法について述べる。感性に影響を与える音楽構造を抽出するためには、楽譜全体を一塊で扱うだけでなく、それを分析して楽譜を構成する情報を抽出せねばならない。そのために、西洋調性音楽の音楽理論を制約条件として使い、調性、終止形、借用和音といった音楽構造を扱えるようにした。

音楽構造の抽出は、以下のような手順で行われる。

1. 被験者に既存の曲をいくつか提示する。心理実験では、図3に示すように75曲を提示していた。2015年度に作成したデモでは50曲を用意し、10曲程度を聴くようにしている。
2. 数種類¹の形容詞対を提示し、提示した曲の印象を形容詞対ごとに5段階評価してもらう。デモでは、簡単のため高揚度：高揚（正）－沈鬱（負）を答えるだけで、以下の手順を省

¹嗜好度（好き、嫌い）、明るさ（明るい、暗い）、安心度（安心、不安）、美しさ（美しい、汚い）、嬉しさ（嬉しい、悲しい）、せつなさ（せつない、せつなくない）



図3: アンケートによる印象収集

略している。

3. 被験者が5の評価を与えた曲に含まれている音楽構造を正例、その他の曲に含まれている構造を負例とし、機械学習システムにかける。これにより、5以上の評価を得る音楽構造が得られる。
4. 被験者が4以上の評価を与えた曲に含まれている音楽構造を正例、その他の曲に含まれている構造を負例とし、前段階で抽出結果に修正を加える。これにより、4以上の評価を得る音楽構造が得られる。

また、音楽構造としては、和音進行、リズム [4]、メロディを扱えるようにしている。たとえば、和音進行の場合だと、次のような属性（述語）を用いた：

1. 調性、拍子、使用楽器といった楽曲の枠組を表す frame/1
2. 連続する2和音の連結形態を表す pair/2
3. 連続する3和音の連結形態を表す triplet/3
4. (より一般的に) 和音の系列を扱う述語（複数）

例えば、調性がハ長調もしくはハ短調、テンポが Allegretto、伴奏楽器がピアノ、拍子は4分の4であり、ある連結形態の2和音が含まれる楽曲がある



図4: 制約条件の獲得例

被験者が「好き」と感じる、といった制約条件を獲得できる。

制約条件の獲得には、正例と負例に基づく機械学習の手法として、帰納論理プログラミング (Inductive Logic Programming: ILP) [12] を用いている。正例のみを用いて、相関ルールマイニング [12] により系列を取り出すこともできる。楽譜は記号で表現されており、記号間の関係も重要であるので、記号や関係表現が直接扱える ILP との親和性はよい。図4に獲得された制約の例と、それと合う楽譜を示す。制約は prolog で表現され、図中に示すように各変数 C1,C2,C3 は楽譜の各小節に対応している。これらの制約は被験者 A が暗いと感じる曲は、楽曲の枠組みが短調で、テンポがラルゲットであり、次の和音進行を含んでいることを示す。最初の和音が短調で V 度、二番目が VI 度で 5 の和音、三番目が 7 の和音の基本形である。

3 作曲

作曲は遺伝的アルゴリズムを用いて行われる。遺伝的アルゴリズムは、染色体表現、遺伝的操作及び適合度関数の実装が非常に重要である。そこで、染色体表現として図6を用い、音楽理論に基づいた最低限度の禁則事項を表現する評価関数と、学習により抽出された個人の感性情報を基に作成した評価関数とを組み合わせることにより、適合度関数を作成した。これにより、楽曲として成り立つ最低限度の曲の生成を実現すると共に、個人の感性に合わせた

- ・ 被験者 A 暗い
 - frame(S) :-
tonality_moll(S),
tempo_larghetto(S).
 - triplet(C1, C2, C3) :-
moll(C1),
form_V(C2),
chord_VI(C2),
chord_V(C1),
inversion_Zero(C3),
form_VII(C3).

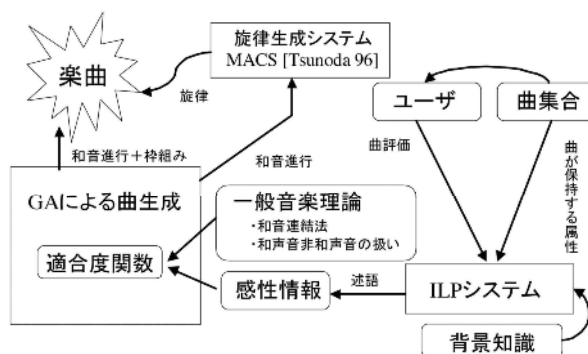


図5: 作曲システムの構成



図6: 染色体

楽曲の生成を実現している。

また、この遺伝的アルゴリズムで生成した楽曲は和音進行だけで構成されているため、旋律生成システムにより旋律を付加している。たとえば、初期のシステムで用いた MACS [11] は、和音進行や音長規則等を与えることにより確率的に旋律を生成する。

以上の手法によれば、たとえば、明るさに関する

知識を用意していなくても学習により明るさに影響を与える音楽構造を獲得し、さらにそれを基に作曲を行える。

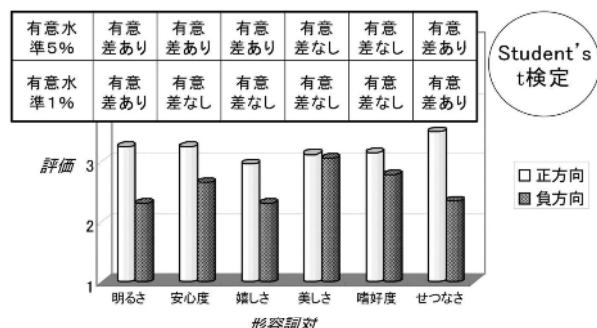


図 7: 作曲に対する被験者の評価

図 7 に示すように、被験者実験では、6 種類の形容詞対のうち、明るさ、安心度、嬉しさ、せつなさ、の 4 形容詞対についての作曲結果は、t 検定により有意水準 5% で有意差が確認されている。さらに、2 種類の形容詞対に対しては有意水準 1% で有意差が確認されている。

被験者実験の結果において、本システムは 6 種類の形容詞対のうち 4 種類について被験者の感性情報に基づいた作曲を行えることを実証した。しかし、対立概念のうち正方向への作曲は曲全体の統一感、構成感といった作曲技法に起因する問題から十分とまでいえる結果は得られていない。これを改善するため、リズムやメロディを学習したり [4]、曲全体の構成を工夫することを試みている [5]。

本システムで提案した手法は、機械学習により被験者から抽出した感性情報を用い作曲を行うという点において従来の作曲アルゴリズムと大きく異なる。また、対話的遺伝的アルゴリズムを用いた作曲手法に比べて、聴き手（被験者）へのアンケート負担を劇的に減らすことができる。

4 生体信号を用いた音楽構造の抽出

アンケートによる印象収集では、訓練曲を評価してアンケートに答える必要があり、多大な手間がかかっていた。そこで、心電、皮膚抵抗、脳波などの生体信号を解析し、楽曲の評価付けを行なう手法を開発した [6]。その中では、1 曲全体の評価付けだけではなく、一つの曲の中での評価の変化を取り出すようになっている。どの程度の長さの部分につい

て評価を行えるのかについても、検討を行った [1]。

生体信号を用いた作曲 → ブレインメロディ

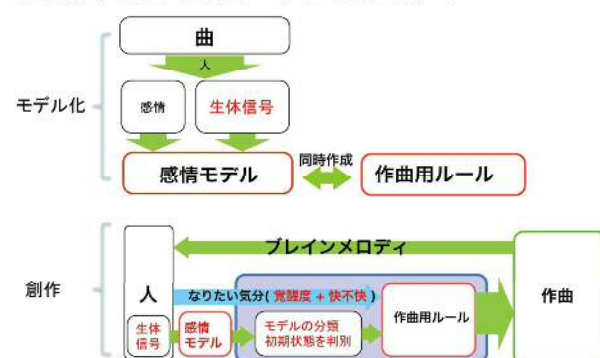


図 8: 生体信号に基づく創作

この手法の流れを図 8 に示す。図 2 に生体信号が加わっている。まず、生体信号と感情の関係を機械学習によりモデル化する。アンケートの代わりに、そのモデルに基づいて、聴き手の音楽に対する反応を得る。それを元に作曲用ルールを作り出す。MIDI/ 音声処理開発を行っているクリムゾンテクノロジー社の協力を得て、大阪大学 COI プロジェクトの一環として、これを実際に行うデモシステムを実現した。また、オランダの imec/NL の協力を得て、図 9 のようなヘッドホン付の脳波センサも試作した。ジェル不要の柔らかなポリマー電極により、快適な装着感が得られている。



図 9: 脳波センサ

生体信号と感情の関係を機械学習によりモデル化する手順を図 10 に示す。縦軸を覚醒度、横軸を快-不快として、聴き手の感情状態を表現することで、

マウスやジョイスティックで入力してもらう（アノテーション）。それと生体信号の関係を機械学習によって学習し、生体信号から感情状態を得られるようにする。

解析結果 [8, 9] を図 11 に示す。聴き手に入力してもらったアノテーションがばらつくことも考えると、かなりの精度で感情が抽出されている。心電計や皮膚抵抗計の結果も総合することで、さらに安定した結果が得られるようになってきている [7]。

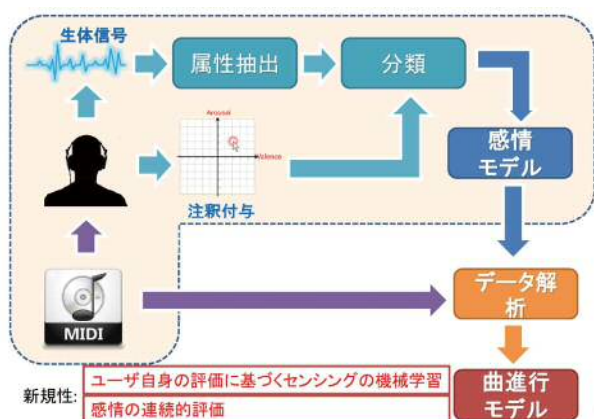


図 10: 脳波データの解析

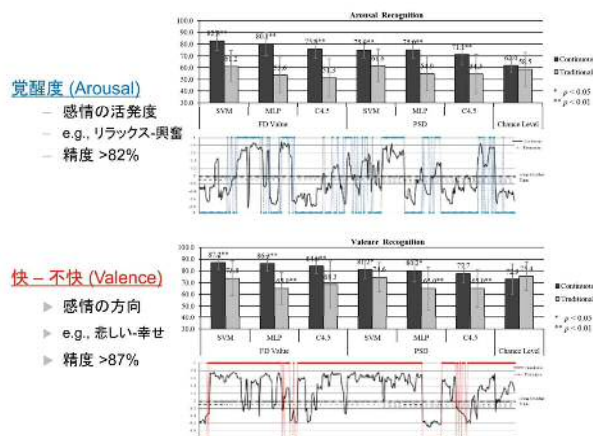


図 11: 解析結果

5 おわりに

音楽を対象とした感情の測定技術と予測技術、それらをベースにした自動作曲について述べた。この技術をベースに、東京都市大学の太谷教授が、フォークデュオ・ワライナキとのコラボにも取り組んでいる。

参考文献

- [1] Rafael Cabredo, Roberto Legaspi, Paul Salvador Inventado, and Masayuki Numao. Discovering emotion-inducing music features using EEG signals. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, Vol. 17, No. 3, pp. 362–370, 2013.
- [2] K. Devlin. *Logic and information*. Cambridge University Press, Cambridge, 1991.
- [3] Masayuki Numao, Shoichi Takagi, and Keisuke Nakamura. Constructive adaptive user interfaces — composing music based on human feelings. In *Proc. Eighteenth National Conference on Artificial Intelligence (AAAI-02)*, pp. 193–198. The AAAI Press / The MIT Press, 2002.
- [4] Noriko Otani, Ryoko Kamimura, Yu Yamano, and Masayuki Numao. Generation of rhythm for melody in a constructive adaptive user interface. In *4th International Workshop on Empathic Computing, a workshop in IJCAI-13*, 2013.
- [5] Noriko Otani, Satoshi Kurihara, and Masayuki Numao. Generation of chord progression using harmony search algorithm for a constructive adaptive user interface. In *Proc. 12th Pacific Rim International Conference on Artificial Intelligence (PRICAI 2012), Lecture Notes in Artificial Intelligence*, Vol. 7458, pp. 400–410, 2012.
- [6] 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*, Vol. 21, No. 3, pp. 200–208, 2008.
- [7] Nattapong Thammasan, Juan Lorenzo Hagad, Kenichi Fukui, and Masayuki Numao. Multimodal stability-sensitive emotion recognition based on brainwave and physiological signals. In *Context Based Affect Recognition (CBAR), a workshop in 7th Affective Computing and Intelligent Interaction (to appear in IEEExplore)*, 2017.

- [8] Nattapong Thammasan, Koichi Moriyama, Kenichi Fukui, and Masayuki Numao. Continuous music-emotion recognition based on electroen-cephalogram. *IEICE Transactions*, Vol. E99-D, No. 4, pp. 1234–1241, 2016.
- [9] Nattapong Thammasan, Koichi Moriyama, Kenichi Fukui, and Masayuki Numao. Familiarity effects in EEG-based emotion recognition. *Brain Informatics*, Vol. 4, pp. 39–50, 2017.
- [10] Gerhard Widmer. Using plausible explanations to bias empirical generalization in weak theory domains. *Machine Learning (EWSL-91)*, pp. 33–43, 1991.
- [11] 角田和也. 計算機における作曲支援システムの研究. 三重大学大学院博士前期課程論文, 1996.
- [12] 元田浩, 津本周作, 山口高平, 沼尾正行. データマイニングの基礎. オーム社, 2006.

