

ユーザの感性に合わせた自動編曲及び作曲

沼尾正行, 高木将一, 中村啓佑¹

東京工業大学 大学院情報理工学研究科 計算工学専攻

Email: numao@cs.titech.ac.jp

概要

ユーザインタフェースにおいて、ユーザに適応し、適切な挙動を示すシステムを構築する手法について述べる。この手法では、情報の移動度を測定し、それに基づいて情報の流れを制御することにより、情報フィルタリングや検索プロセスをユーザやその環境に適応させる。さらに、情報の内部構造を分析し、その構成要素の移動度を測定することにより、情報を選択するだけでなく、その構成過程をも制御できる。この原理に基づくオーサリングシステムの実例として、音楽の編曲および作曲システムについて述べる。このシステムではユーザがキーワード(形容詞対)を指定することにより、好みに応じた自動編曲や、自動作曲が行える。

Constructive Adaptive User Interfaces — Composing Music Based on Human Feelings

Masayuki NUMAO, Shoichi TAKAGI, and Keisuke NAKAMURA

Department of Computer Science, Tokyo Institute of Technology

Email: numao@cs.titech.ac.jp

Abstract

We propose a method to locate relations and constraints between a music score and its impressions, by which we show that machine learning techniques may provide a powerful tool for composing music and analyzing human feelings. We examine its generality by modifying some arrangements to provide the subjects with a specified impression. This talk introduces some user interfaces, which are capable of predicting feelings and creating new objects based on seed structures, such as spectrums and their transition for sounds that have been extracted and are perceived as favorable by the test subject.

¹現在、東京大学知能機械情報専攻

1 はじめに

音楽には、作曲家が楽譜を書き、演奏家が演奏して、聴衆が聞くという、作曲家から聴衆への図1に示すような情報の流れがある [8]。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。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」を調べることにより、個人の感性に合わせた作曲・編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最新のシステムについて述べる。

このシステムでは、まず個々のユーザに幾つかの楽曲を聞いてもらい、各々の楽曲への感じ方のアンケートに答えさせる。それにより、ユーザに印象を引き起こす音楽構造を抽出する。編曲の場合は、既存の曲を抽出した音楽構造になるべく合うように変形することにより、ユーザが指定した印象の曲を得る。作曲の場合は、抽出した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。以下、音楽構造の抽出、編曲、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。

2 音楽構造の抽出

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はなく、それを分

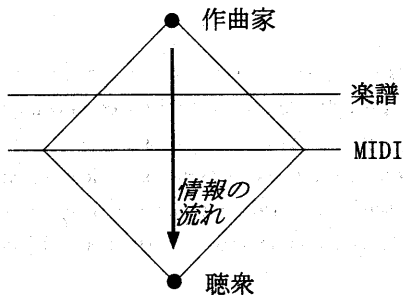


図 1: 情報の流れとオーサリング

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用い、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提示する。今回は 75 曲を提示した。
2. 6 種類 12 語²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。
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、音楽構造に直接対応する属性(述語)として、

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`

の 3 種類を用意した。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、帰納論理プログラミング (Inductive Logic Programming: ILP)[7] を用いている。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 2 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

²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)

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),
 choird_V(C1),
 inversion_Zero(C3),
 form_VII(C3).

図 2: 制約条件の獲得例

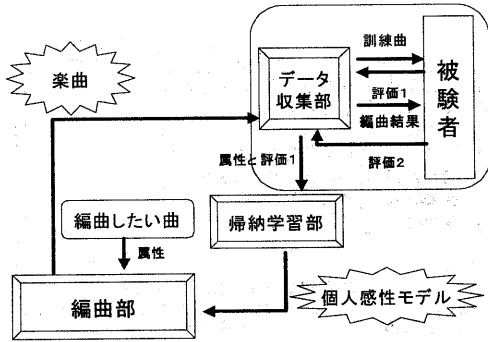


図 3: 編曲システムの構成

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 編曲

編曲の場合は、個人の感性を反映しつつ原曲の雰囲気がある程度は残す必要がある。そのために、以下のような手法により編曲を行っている。

1. 原曲の持つ和音構造を解析し、その和音が持つ機能 (Tonic, Dominant, Subdominant) を特定する。
2. 学習された制約をより多く満たすように各和音を変更する。ただし、原曲の各和音が持つ機能は変更しないように、変更候補の和音を限定する。
3. 原曲が持っていた旋律を、2番目の手順で確定された和音列を基に変更する。すなわち、編曲後の和音構成音と、編曲後の旋律の整合性を取るようになる。

具体的には次のような手法によった。

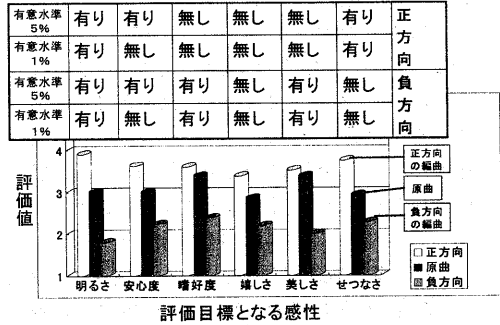


図 4: 編曲に対する被験者の評価

1. 編曲対象先頭の三和音を対象とする。
2. 三和音が取りうる和音を和音機能は変えずに述語 pair, triplet を満たすように列挙する。
3. それまでに対象としていた三和音中の最後の和音と次の二和音とを組み合わせた三和音に対象を移す。2と同様に列挙する。
4. 3を繰り返して合計して最終的に一番訓練例を満たしている和音列を選ぶ。同点の場合はなるべく原曲に近い和音を選ぶ。
5. これを全ての調 (12個) で行い、一番訓練例を満たしている調を選ぶ。
6. 5で選んだ調で最も良い frame(曲の構造) を選ぶ。具体的には最も多く訓練例を満たす frame を全探索する。同点の場合はより原曲に近い方を選ぶ。

実験では実際の曲を用いて、被験者の曲に対する感性の学習がどの程度できるかを調べる。実験においては、被験者にとってポピュラーであると思われる J-POP から 39 曲、クラシック、和声学の教科書から 36 曲の本格的転調のない 75 曲を用意した。実験では被験者の負担を考慮し、曲全体を対象とせず、適当な 8 ~ 16 小節程度を抜き出し、コード進行をデータとして被験者に質問した。又編曲用の曲はこれとは別に、既存の曲、3 曲を用意した。

実験方法 (図 3) を以下に示す。

1. 各曲に対し、明るさ、安定度、嗜好度、美麗度、うれしさ、せつなさの 6 つの感性の評定尺度について、各々 5 段階で質問した。
2. 被験者のデータをもとに、明るさについては“明るい”、“暗い”、安定度については“安心”、“不安”、嗜好度については“好き”、“嫌

い”、美麗度については“美しさ”、“汚さ”、うれしさについては“嬉しい”、“悲しい”、せつなさについては“せつない”、“せつなくない”の計6つの感性についてそれぞれ個人編曲モデルを生成した。

せつなさについては“せつない”の反対語が見つからなかった為、取って否定語である“せつなくない”という形容詞を用いた。

- 今回の実験は範囲評価を用いた。範囲評価とは1曲の内、全体の評価はXだが一部のここの評価はYであるという場合に用いるもので、曲全体を評価するよりも精度の高い感性モデルが生成可能となる。範囲評価を行うために曲を楽譜に直し、JavaAppletに取り込んでhtml上で視覚的に感性語を指定し、評価してもらった。なお、評定尺度には全体評価と同様の6つを用いた。
- 編曲対象の曲について、1曲ごとにそれぞれの感性に関する編曲を行ない、それをすべての曲について行なう。
- 編曲後の曲と元の曲、計36曲について同様の方法で被験者に質問した。この際、被験者にどのような編曲を行なった結果であるかは伏せて、質問している。
- 元の曲と特定の感性に対する編曲後の曲について、各形容詞についての改善度を比較・考察し、システムの性能の評価を行った。

明るさ・安定度・嗜好度・美麗度・うれしさ・せつなさの6つの評定尺度についての評価の平均値を図4に表す。

全ての形容詞において、正方向の編曲(好き、明るい、安心、美しい、嬉しい、せつない)は原曲に対して高い評価値に、負方向の編曲(嫌い、暗い、不安、汚さ、悲しさ、せつなくない)は原曲に対して低い評価値を示した。データ数が少ないということもあるが、平均値で見ると全て成功している。

全形容詞において、多少の個人差が見受けられるが全体的に良い値が得られた。これは、範囲評価を設けたこと、triplet(3和音の述語)、frame(曲の枠組)を新たに獲得する述語として設けたこと、編曲で全探索という形で和音進行を探索したことによると考えられる。

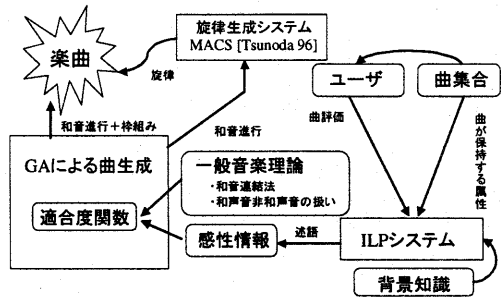


図5: 作曲システムの構成

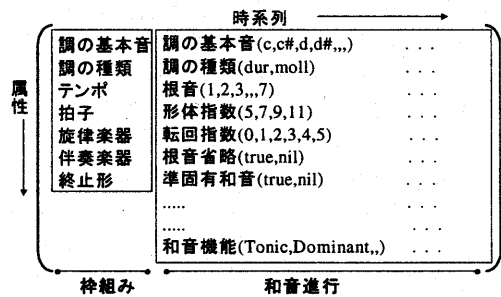


図6: 染色体

4 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で、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 MACS[4]により旋律を付加している。MACS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システムである。

以上の手法を基に、被験者実験を行った。図9は被験者Aが明るいと感じるような作曲例を、図10は被験者Aが暗いと感じるような作曲例を示している。

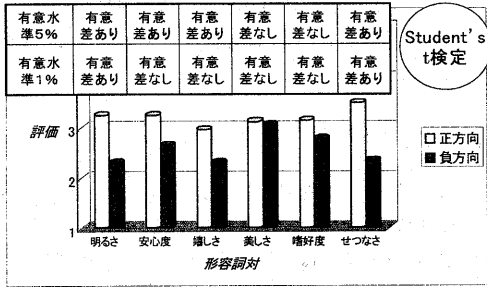


図 7: 作曲に対する被験者の評価

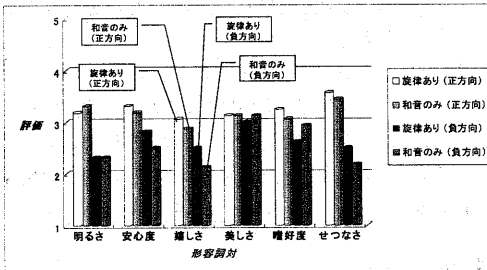


図 8: 旋律の効果

これらの作曲例は、背景知識に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たことを示している。

図 7 に示すように、今回の被験者実験では、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。図 8 に旋律の効果を示す。形容詞対によっては、大きな差が出ているのが分かる。

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。この問題は今後作曲技法及び獲得述語の種類を増やすことによって対応したい。本研究では一般的に定量的評価を与えにくい計算機による作曲手法という研究分野において、被験者実験を行うことによりその作曲結果に対



図 9: 被験者 A が明るいと感じる曲



図 10: 被験者 A が暗いと感じる曲

してある程度の定量的評価を行った。これについてはさらに考察を深めたい。

本システムで提案した手法は、帰納論理プログラミング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。本研究は、人間の評価を用いるという意味に於いて本研究に一番近いと思われる対話的遺伝的アルゴリズムを用いた作曲手法に比べて、被験者への負担が少ない上に抽出した情報の再利用性という点で大きな利点を持つ。また述語で感性情報を定義するために、被験者自身が述語形式で自分の感性情報を定義することができる。そのため帰納論理プログラミングシステムによって学習できないにもかかわらず本人が自覚している感性情報についてシステムに与えることによって、様々な作曲を試すことができる。これらのことは、このシステムが作曲能力のある人、無い人の両方にとって有用であることを示す。本システムでは作曲手法に関する背

景知識は最低限のものしか用意していないにも関わらず、人間の作曲家にはとてもおよばないものの聴くに耐え得る楽曲を生成できた。本研究の手法を改善していくことによって、さらに質の高い作曲を行えると考えられる。

5 関連研究

CONCERT[1]はニューラルネットワークを用いた作曲システムである。バックプロパゲーションを用い、入力音と文脈から次の音の候補を選択するように学習を行う。ピッチ(音程)表現を5度円(Circle of Fifth)等の音楽心理学の成果の一部を取り入れることにより表している。音楽心理学の成果を採用入れることにより生成された曲の質を高めているということに特徴がある。しかし、曲冒頭はある程度の質を持ったフレーズが並んでいるが、曲の終止部付近では旋律としてのまとまりを失ってしまうと言う欠点がある。

Wiggins[3]によるGAを用いた作曲、和声付けシステムでは、利用者が旋律を入力し、その旋律に対する和声付けを行う例と、利用者が和声を入力し、それに対する旋律を生成する例の2種類が提示されている。遺伝子表現、遺伝的操作、適合度関数は音楽知識を十分に考慮された上で形成されている。遺伝的操作は大きく交叉、突然変異の2つに分かれるが、特に突然変異を重要視し、意図的に音楽的意味がある操作にしている。この研究で述べられたGAによる作曲手法に於ける突然変異の重要性は、本研究に大きな影響を与えている。

平田らによって開発されたハービー君[5]、パービーブン[9]、パービーブン[10]は、演繹オブジェクト指向(DOOD)に基づく知識表現を使用したJazzにおけるリハーモナイズシステム、編曲システムである。ハービー君はジャズのリハーモナイズシステムで、各和音単位で演繹を行いリハーモナイズを行う。パービーブンはハービー君の改良システムで、複数和音の連結をケーダンス木で表し、和声的文脈を考慮してリハーモナイズを行う。Time-Span Reduction(TSR)の考え方を導入し、グルーピングを行っている。グルーピングにより、ケーダンスの和音機能を考慮することが可能である。パービーブンは事例ベースの編曲システムで、和音レベルではなく音符レベルで楽曲を扱う。音符単位で楽曲を扱うために、意外性のある編曲を行うことが可能

になっている反面、意外過ぎて、本研究で意図したような人間の感性への適合には向かない可能性もある。

6 おわりに

形容詞対に関する背景知識を用意しなくても、個人の感性情報を学習により抽出することにより個人に合わせた作曲・編曲を可能とする手法を提案し、実際に実験を行って作曲および編曲結果が被験者の意図に合っていることを確かめた。筆者らが以前に発表した実験[2]では編曲のみが対象であったが、それを作曲に拡張し、得られた曲の質をかなり改善することに成功した。

本手法はWWWのコンテンツ、ゲーム、ユーザインタフェースなどに活用することが可能である。また、ここで提案した手法は音楽以外のコンテンツにも応用可能であり、急増するコンテンツへの要求に応える手法になると考えられる。

参考文献

- [1] M.C. Mozer. Neural network music composition by prediction: Exploring the benefits of psychoacoustic constraints and multiscale processing. *Connection Science*, 1994.
- [2] M. Numao, M. Kobayashi, and K. Sakaniwa. Acquisition of human feelings in music arrangement. In *Proc. IJCAI 97*, pp. 268-273. Morgan Kaufmann, 1997.
- [3] G. Wiggins, et al. Evolutionary methods for musical composition. *International Journal of Computing Anticipatory Systems*, 1999.
- [4] 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.
- [5] 後藤, 平田. ハービー君: 演繹オブジェクト指向に基づいてジャズらしいコードにリハーモナイズするシステム. 情報処理学会研究報告, Vol. 96-MUS-16, pp. 33-38, 1996.
- [6] 高木将一, 沼尾正行. 機械学習の手法を用いた感性の抽出と作曲・編曲への応用. 人工知能学会全国大会(第15回)論文集, 2001.
- [7] 沼尾正行. 帰納論理プログラミングの応用—近年の国際会議の動向を踏まえて—. 人工知能学会誌, Vol. 12, No. 5, pp. 683-688, 1997.
- [8] 沼尾正行. Global intelligenceによる知識流通. 情報処理学会知能と複雑系研究会, Vol. 01-ICS-124, pp. 1-8, 2001.
- [9] 平田, 青柳. パービーブン: 誰でもどこでもインタラクティブに使える知的ジャズ和音生成システム. 情報処理学会研究報告, Vol. 99-MUS-31, pp. 7-12, 1999.
- [10] 平田, 青柳. パービーブン: 音符レベルでユーザ意図を把握して編曲を行う事例ベースシステム. 情報処理学会研究報告, Vol. 2000-MUS-37, pp. 17-23, 2000.