

楽曲の部分構造と全体構造を考慮した自動作曲システム

Automatic composition system considering both partial and overall music structure

西川 敬之*¹
Takayuki Nishikawa

大谷 紀子*²
Noriko Otani

福井 健一*³
Ken-ichi Fukui

森山 甲一*³
Koichi Moriyama

栗原 聡*³
Satoshi Kurihara

沼尾 正行*³
Masayuki Numao

*¹大阪大学大学院情報科学研究科

Graduate School of Information Science and Technology, Osaka University

*²東京都市大学環境情報学部

Faculty of Environmental and Information Studies, Tokyo City University

*³大阪大学産業科学研究所

The Institute of Scientific and Industrial Research, Osaka University

We have focused on the system that automatically composes music adapting to the user's sensibility. The system learns sensibility model of a user from his/her evaluations of tunes. The system composes music using Genetic Algorithm (GA) whose fitness function is based upon the user's learned sensibility model. In this paper, we refine on GA in composing stage to make it generate tunes considering both partial and overall music structure, taking advantage of motif in music. We empirically validated our approach improves the quality of composed music.

1. はじめに

本研究室では、従来行われてきた計算機による自動作曲の研究に、近年研究が盛んに行われている感性工学の観点を取り入れ、個人の感性を反映させた楽曲を自動生成するシステムの研究開発を行ってきた。この自動作曲システムは、既存楽曲に対するユーザの評価を基に帰納論理プログラミング (Inductive Logic Programming, ILP) によってユーザの音楽に対する感性と楽曲の構造の関係を学習し、ユーザの感性を反映していると思われる楽曲の部分構造を感性モデルとして獲得する。そして、そのユーザの感性モデルを適応度関数に組み込むことで、遺伝的アルゴリズム (Genetic Algorithm, GA) による個人感性を反映した楽曲の生成を実現している。

上記の作曲手法によって、一部の感性において被験者の感性に基づいた作曲が可能であることが被験者実験によって実証されている [Legaspi 07]。しかし、生成される楽曲の「音楽としての良さ」については特に言及されてこなかった。

この自動作曲システムが獲得する個人の感性モデルは「和音のルートがⅢ度、Ⅲ度、Ⅵ度と続く『優しい』と感じる」といった楽曲の部分構造のみを示すものであり、これらを遺伝的アルゴリズムの適応度関数として利用する従来の楽曲の生成手法は、楽曲の局所的な構造から全体の構造を生成するボトムアップ方式の作曲手法であると言える。すなわち、生成される楽曲の音楽としての展開や構成といった全体構造を考慮せずに楽曲を生成している。従来は生成楽曲の長さを8小節に限定して生成していたため、それでも楽曲として成立していたが、それ以上の長さの楽曲を生成しようとする、曲としてのまとまりがなくなる、非常に単調な楽曲が生成されるといった問題が生じる。したがって、生成楽曲の音楽としての完成度を高めるためには、楽曲の全体構造も考慮しながら作曲を行う必要がある。

本研究では、楽曲の部分構造と全体構造の両者を考慮した、音楽の生成に適した遺伝的アルゴリズムを設計する。これにより、個人の感性を反映しつつ、音楽的にも優れた楽曲を生成する自動作曲システムの構築を目指す。

以下、2.節で本研究の基礎となる自動作曲システムについて述べ、3.節で提案手法について述べる。次いで4.節で本研究で提案した手法を実装した自動作曲システムを用いた実験とその結果について記し、5.節で本研究についてまとめる。

2. 個人の感性モデルの獲得と感性モデルに基づく作曲

本節では、本研究のベースとなる自動作曲システムの詳細について述べる。システムはユーザの音楽についての感性情報を学習し感性モデルを獲得する部分と、得られた感性モデルを基に作曲を行う部分の二つに大別できる。

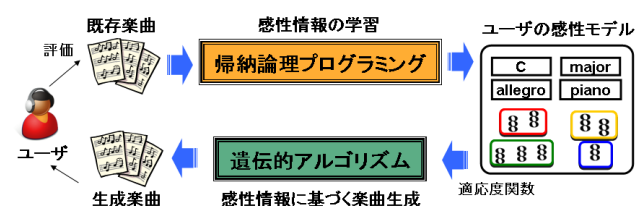


図 1: 自動作曲システム概要図

2.1 節で感性情報を学習し感性モデルを獲得する方法について、2.2 節で獲得した感性モデルを基に作曲を行う方法について述べる。

2.1 感性モデルの獲得

システムにユーザの音楽に対する感性と楽曲の構造の関係を学習させるためには、システムに入力する楽曲データを構造を持つデータとして表現する必要がある。そのため、楽曲を以下の述語を用いて述語論理で表現する。

連絡先: 沼尾研究室, 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:06-6879-8426, Fax:06-6879-8428,
E-mail:nishikawa@ai.sanken.osaka-u.ac.jp

表 1: 楽曲の枠組構造の特徴量

Genre	楽曲のジャンルを表す
Key	調の中心であり、音階の起点となる音を表す
Scale	楽曲の音階を表す
Tonality	長調か短調かを表す
Tempo	楽曲の速さを表す
Time Signature	楽曲の持つ基本的なリズムを表す
Melody Instrument	楽曲の旋律の音色を表す
Melody Category	Melody Instrument が示す音色の分類を表す
Chord Instrument	楽曲の和音の音色を表す
Chord Category	Chord Instrument が示す音色の分類を表す

表 2: 楽曲の和音構造の特徴量

Root	和音の基礎となる音が音階の何度の音に相当するかを表す
Type	和音を構成する音の数及び構成音ごとの関連を表す
Tension	和音に付加的に用いられている音を表す

- `song_frame/10`
楽曲の枠組を表す述語であり、その引数は表 1 に示す Key, Tempo 等である。
- `chord/3`
和音構造を表す述語であり、その引数は表 2 で示す Root, Type, Tension である。この述語をリストの要素として並べることにより、和音進行が表現される。
- `music/3`
楽曲全体を表す述語であり、第一引数にその楽曲の番号、第二引数に `song_frame/10` を、第三引数に `chord/3` のリストをとる。

`music/3` の例を以下に示す。

```
music(
  46,
  song_frame(pops,g,g,major,four_four,allegro,
             flute,flute,steel_guitar,guitar),
  [chord(i,major,null),    chord(i,major,null),
   chord(iv,major,null),  chord(iv,major,null),
   chord(ii,minor_7,null), chord(ii,minor_7,null),
   chord(i,major,null),   chord(i,major,null),
   chord(i,major,null),   chord(i,major,null),
   chord(iv,major,null),  chord(iv,major,null),
   chord(ii,minor_7,null), chord(ii,minor_7,null),
   chord(i,major,null),   chord(i,major,null)]
).
```

述語論理で表現した楽曲構造から個人の感性に影響する楽曲の部分構造を感性モデルとして獲得するために、帰納論理プログラミングの一つである FOIL [Quinlan 97] を利用する。ユーザには事前に既存楽曲のもつ感情的性格についての 5 段階評価を行ってもらい、評価により一定以上の評価値を得た楽曲を正例、それ未満の評価を得た楽曲を負例として FOIL に入力し、出力として個人の感性情報を反映した楽曲の部分構造を得る。このシステムでは楽曲の枠組構造と和音進行についての感性モデルを、それぞれ述語 `frame/2`, `chords/2` の形で獲得している。以下の述語例は、あるユーザが「明るい」と感じる楽曲が、メロディの音色がシンセ・リード系でコードの音色がギター系、調性が major であるという曲の枠組や、1 つめの和音が I 度 major で 2 つめの和音のタイプが minor 7th, 3 つめの和音のルートが VI 度となる連続した 3 和音を持つということを表している。

```
frame(bright, A) :- melody_category(A, lead),
                   chord_category(A, guitar).
tonality(A, major).

chords(bright, A) :- has_chord(A, B, C), root(C, i),
                    type(C, major),
                    next_to(A, B, D),
                    has_chord(A, D, E), type(E, minor_7),
                    next_to(A, D, F),
                    has_chord(A, F, G), root(G, vi).
```

2.2 感性モデルに基づく作曲

2.1 節で得られた感性モデルを適応度関数として用いた GA の最適解として、ユーザの感性情報を反映した楽曲の枠組構造と和音進行を獲得する。GA による枠組構造の獲得と和音進行の獲得は互いに独立したプロセスである。なお、両者共に世代交代モデルとして MGG[佐藤 97] を採用している。

2.2.1 楽曲の枠組構造の獲得

図 2 のように定義した染色体に対して遺伝的操作を行う。

Genre	Key	Scale	Tonality	Time Signature	Tempo	Melody Instrument	Melody Category	Chord Instrument	Chord Category
-------	-----	-------	----------	----------------	-------	-------------------	-----------------	------------------	----------------

図 2: 楽曲の枠組構造の染色体表現

二点交叉によって特徴量の組み合わせを変更し、突然変異によって特徴量の値を変化させる。

適応度関数 $Fitness_Function(F)$ は以下のように定義する。

$$Fitness_Function(F) = Fitness_Frame(F) \quad (1)$$

$Fitness_Frame(F)$ は楽曲の枠組構造に関するユーザの感性モデルに基づく関数であり、2.1 節で得られた楽曲の枠組構造に関するユーザの感性モデルを基に作成する。

2.2.2 楽曲の和音進行の獲得

1 つの遺伝子を 1 小節 1 和音に対応させ、染色体を定義する。図 3 は 8 小節の長さを持つ楽曲を表している。このように定義した染色体に対して遺伝的操作を行う。

chord 1	chord 2	chord 3	chord 4	chord 5	chord 6	chord 7	chord 8
---------	---------	---------	---------	---------	---------	---------	---------

図 3: 楽曲の和音進行の染色体表現

二点交叉によって和音の進行を組み換え、突然変異によって和音を構成する三つの特徴量 (Root, Type, Tension) を変化させる。

適応度関数 $Fitness_Function(C)$ は以下のように定義する。

$$Fitness_Function(C) = Fitness_Chords(C) + Fitness_Builtin(C) \quad (2)$$

$Fitness_Chords(C)$ は楽曲の和音進行に関するユーザの感性モデルに基づく関数であり、2.1 節で得られた楽曲の和音進行に関するユーザの感性モデルを基に作成する。 $Fitness_Builtin(C)$ は音楽理論における和音進行の禁則に基づく関数であり、不自然な和音進行の発生を抑制する。

3. 提案手法

1. 節で述べたように、和音進行獲得時における従来の GA の各個体の評価方法は、図 4 のように楽曲の部分構造についての感性モデルを適応度関数とした評価であったため、生成される楽曲は全体としてのまとまりに欠ける、非常に単調なものになるといった問題が生じていた。

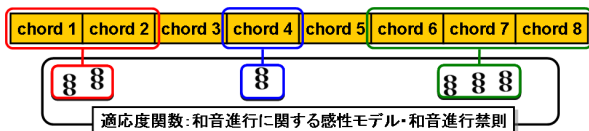


図 4: 従来の適応度関数の適用方法

本研究ではこれらの問題点を解決するために、楽曲を構成する最小単位であるモチーフ (2 小節) を基に遺伝的アルゴリズムを設計、利用する手法を提案する。この手法により、楽曲としてのまとまりを持ちかつ曲の中で展開を持つ楽曲の生成が可能になることが期待される。

3.1 モチーフについての感性モデルの獲得

本研究では従来の楽曲の枠組構造と和音進行についての感性モデルに加え、新たに楽曲のモチーフについての感性モデルを獲得する。そのために、従来の述語 music/3 による楽曲の述語表現に加え、新たに以下の述語を定義して楽曲を述語表現する。

- bar/4
楽曲の 1 小節を時系列に表現する。前から順に 1 つの属性値を四分音符 1 つ分の長さに対応させ、属性値によってその時刻にどの音が鳴るかを表現する。
- motif/2
第一引数と第二引数に bar/4 を持ち、全体で 2 小節分の時系列を表す。
- music_structure/2
楽曲全体を表す述語であり、第一引数にその楽曲の番号、第二引数に motif/2 のリストをとる。

music_structure/2 の例を以下に示す。

```
music_structure(
26,
[motif(bar((i,major),-, (i,major),-), bar((i,major),-, (i,major),-)),
motif(bar((vi,minor),-, (vi,minor),-), bar((vi,minor),-, (vi,minor),-)),
motif(bar((iv,major),-, (iv,major),-), bar((v,major),-, (v,major),-)),
motif(bar((i,major),-, (i,major),-), bar((i,major),-, (i,major),-)),
motif(bar((i,major),-, (i,major),-), bar((i,major),-, (i,major),-)),
motif(bar((vi,minor),-, (vi,minor),-), bar((vi,minor),-, (vi,minor),-)),
motif(bar((iv,major),-, (iv,major),-), bar((v,major),-, (v,major),-)),
motif(bar((i,major),-, (i,major),-), bar((i,major),-, (i,major),-))]
).
```

上記の述語論理で表現した楽曲データから、FOIL を用いて楽曲のモチーフについての感性モデルを述語 motif/2 の形で獲得する。

```
motif(bright,A) :- motif(A,bar(-,-,-,-),bar((v,7),-,-,,-)).
```

上記の述語例は、このユーザが「明るい」と感じる楽曲が、1 小節目が四分音符 4 拍分の任意の和音で 2 小節目が四分音符 4 拍分の V 度 7th の和音となるモチーフを持つということを表している。

3.2 モチーフに基づく楽曲生成

本研究における GA による楽曲生成は、楽曲の枠組構造の獲得、モチーフの獲得、和音進行の獲得の三つのプロセスに分けられる。枠組構造の獲得方法は 2.2.1 節で説明した手法と同様である。以下でモチーフの獲得手法と、獲得したモチーフを基に楽曲全体の和音進行を獲得する手法について述べる。

3.2.1 楽曲のモチーフの獲得

図 5 のように、2 小節分の和音を時系列に染色体として表現する。

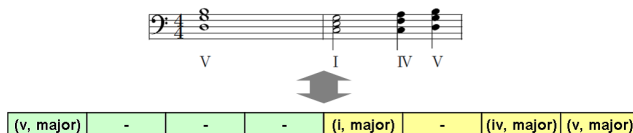


図 5: 楽曲のモチーフの染色体表現

二点交叉によって和音の進行を組み換え、突然変異によって特徴量 (Root, Type) の組み合わせを変化させる。

適応度関数 $Fitness_Function(M)$ は以下のように定義する。

$$Fitness_Function(M) = Fitness_Motif(M) + Fitness_Builtin(M) \quad (3)$$

$Fitness_Motif(M)$ は楽曲の和音進行に関するユーザの感性モデルに基づく関数であり、3.1 節で得られた楽曲のモチーフに関するユーザの感性モデルを基に作成する。 $Fitness_Builtin(M)$ は音楽理論における和音進行の禁則に基づく関数であり、不自然な和音進行の発生を抑制する。

3.2.2 楽曲の和音進行の獲得

3.2.1 節で獲得した楽曲のモチーフを基に、図 6 のように和音進行を染色体表現し、初期個体群を生成する。

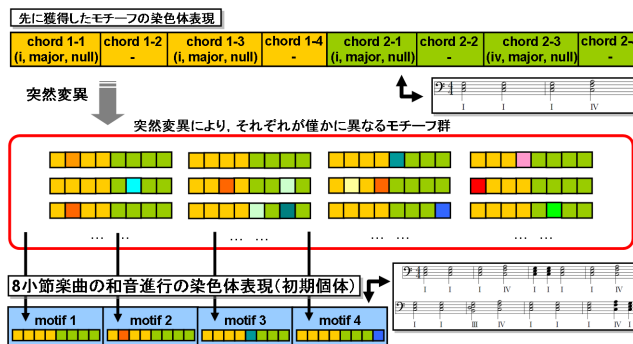


図 6: 楽曲の和音進行の染色体表現及び初期個体生成

このようにして生成された初期個体は、共通のモチーフから構成されているので、楽曲としての統一感を備えていると考えられる。

この初期個体にモチーフを単位とした二点交叉を行うことで、モチーフを保持したまま楽曲全体としての和音の進行を組み換え、突然変異を起こすことによってモチーフ内の和音の特徴量を変化させる。

適応度関数 $Fitness_Function(C)$ は以下のように定義する。

$$Fitness_Function(C) = Fitness_Chords(C) + Fitness_Builtin(C) + Fitness_Forms(C) \quad (4)$$

2.2.2 節で定義した適応度関数に新たに楽曲全体の展開に関するルールについての関数 $Fitness_Forms(C)$ を加える。本研究では、生成楽曲の展開が楽曲の展開としてわかりやすい「起承転結」の展開となるように関数を設定した。初期個体をこの適応度関数を用いて評価し、世代交代を繰り返すことで、楽曲としての統一感を保持しつつ展開のある楽曲の生成を誘導する。適応度関数を染色体に適用する際の概念図を図 7 に示す。

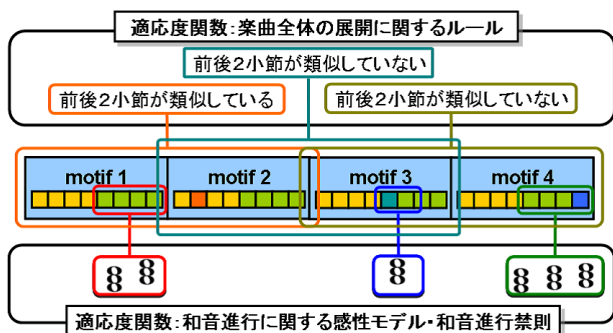


図 7: 楽曲の和音進行を表す染色体への適応度関数の適用

4. 実験

提案手法を実装した自動作曲システムの性能評価を行うために、二種類の被験者実験を行った。

実験 1 では、既存楽曲 53 曲に対する評価を基に被験者ごとの感性に合わせて生成した楽曲を各被験者に提示し、それらの楽曲についての評価値を見ることで実際に個人の感性に合わせた楽曲を生成できたかどうかを確かめる。評価のために五つの形容語対を用意し、各形容語対の間で 5 段階尺度の評価を行ってもらった。評価実験の結果を表 3 に示す。表 3 の各形容語ごとの平均値は、形容語ごとに生成した 5 曲に対する被験者の評価平均値から、さらに被験者 12 人分の平均値をとったものである。

表 3: 実験 1 の被験者の評価平均値及び t 検定結果

評価形容語	平均	分散	t 値	自由度	有意水準 5%
好き	2.45	0.56	-0.11	11	有意差なし
嫌い	2.47	0.77			
明るい	3.07	0.87	5.51	11	有意差あり
暗い	1.47	0.28			
嬉しい	2.93	0.60	6.16	11	有意差あり
悲しい	1.58	0.17			
優しい	3.42	0.59	5.82	11	有意差あり
優しくない	1.92	0.51			
おだやかな	3.60	1.05	7.25	11	有意差あり
おだやかでない	1.42	0.17			

実験 2 では、従来手法と提案手法で生成された楽曲に対して被験者に音楽的評価をつけてもらい、その評価値を比較することで実際に音楽的に優れた楽曲が生成できたかどうかを確かめる。評価のために七つの形容語を用意し、各形容語について 5 段階尺度の評価を行ってもらった。評価実験の結果を表 4 及び表 5 に示す。表 4 は生成楽曲の長さを 8 小節、表 5 は生成楽曲の長さを 16 小節としたときの結果であり、各評価項目の平均値は、それぞれの手法で生成した 10 曲に対する各形容語ごとの被験者の評価平均値から、さらに被験者 15 人分の平均値をとったものである。

これらの結果から以下のことが言える。

表 4: 実験 2 の被験者の評価平均値及び t 検定結果 (8 小節)

評価項目	平均	分散	t 値	自由度	有意水準 5%
楽曲としてのまとまりがある	3.10	0.18	1.35	14	有意差なし
曲に統一感がある	3.19	0.25	-0.23	14	有意差なし
曲に展開がある	3.17	0.24			
曲が単調である	2.47	0.27	4.43	14	有意差あり
楽曲としておもしろい	3.11	0.30			
楽曲としてつまらない	3.53	0.36	-3.71	14	有意差あり
楽曲として成立している	2.88	0.29			
	2.56	0.33	2.68	14	有意差あり
	2.95	0.41			
	3.27	0.32	-2.32	14	有意差あり
	2.98	0.43			
	2.94	0.43	1.34	14	有意差なし
	3.09	0.78			

上段...従来手法 下段...提案手法

表 5: 実験 2 の被験者の評価平均値及び t 検定結果 (16 小節)

評価項目	平均	分散	t 値	自由度	有意水準 5%
楽曲としてのまとまりがある	3.08	0.36	-0.69	14	有意差なし
曲に統一感がある	2.99	0.30	-0.51	14	有意差なし
曲に展開がある	3.29	0.35			
曲が単調である	3.24	0.38	3.66	14	有意差あり
楽曲としておもしろい	2.47	0.48			
楽曲としてつまらない	3.05	0.43	-4.30	14	有意差あり
楽曲として成立している	3.67	0.48			
	3.09	0.29	2.97	14	有意差あり
	2.34	0.57			
	2.79	0.64	-2.37	14	有意差あり
	3.44	0.67			
	3.18	0.68			
	2.89	0.80	1.05	14	有意差なし
	3.01	0.67			

上段...従来手法 下段...提案手法

- 明るさ、嬉しさ、優しさ、おだやかさについては、被験者の感性に応じた作曲を行うことができた。
- 従来手法と比較して、楽曲中に展開が感じられ、面白みのある楽曲を生成することができた。

5. おわりに

本研究では、楽曲の部分構造と全体構造を考慮した自動作曲の手法として、楽曲のモチーフを基に遺伝的アルゴリズムを用いて楽曲を生成する手法を提案した。この手法により、個人の感性を反映しつつ、従来の自動作曲システムよりも音楽的に優れた楽曲を生成することができることを示した。

参考文献

[Legaspi 07] Roberto Legaspi, Yuya Hashimoto, Koichi Moriyama, Satoshi Kurihara and Masayuki Numao. Music Compositional Intelligence with an Affective Flavor. *Proc. ACM International Conference on Intelligent User Interfaces*, pp. 216-224, 2007.

[Quinlan 97] J.R. Quinlan, Learning logical definitions from relations, *Machine Learning*, Vol.5, pp. 239-266 (1990)

[佐藤 97] 佐藤浩, 小野功, 小林重信. 遺伝的アルゴリズムにおける世代交代モデルの提案と評価. *人工知能学会誌*, Vol. 12, pp. 1-11, 1997.