

# ハーモニーサーチに基づくユーザの感性を反映した楽曲生成

Music Composition Adapting to User's Sensibility Based on Harmony Search

大谷 紀子\*<sup>1</sup>

Noriko Otani

田所 克敏\*<sup>1</sup>

Katsutoshi Tadokoro

栗原 聡\*<sup>2</sup>

Satoshi Kurihara

沼尾 正行\*<sup>2</sup>

Masayuki Numao

\*<sup>1</sup> 東京都市大学環境情報学部

Faculty of Environmental and Information Studies, Tokyo City University

\*<sup>2</sup> 大阪大学産業科学研究所

The Institute of Scientific and Industrial Research, Osaka University

Automatic composition systems that compose music adapting to personal sensibility have been proposed. The systems induces a personal sensibility model by using a listener's emotional impressions of music and composes music on the basis of that model. Though the experimental results show that it is possible to compose a musical piece that partially adapts to the listener's sensibility, the quality of the composed piece has not been considered thus far. In order to generate high-quality music, it is necessary to consider both the partial and the complete music structure. In this paper, we propose a method for generating chord progressions using a harmony search algorithm. The harmony search algorithm is a music-based metaheuristic optimization algorithm that imitates the musical improvisational process. We compared the evaluations of music generated using the proposed method and the other method that was proposed in the previous work.

## 1. はじめに

個人の感性を反映した楽曲を生成するシステムの研究が進められている。Legaspiらは、帰納論理プログラミング(Inductive Logic Programming; ILP)を用いて個人の感性モデルを獲得し、遺伝的アルゴリズム(Genetic Algorithm; GA)を用いて感性モデルに即する楽曲を生成する手法を提案している[Legaspi 07]。本手法により、個人の感性を反映した楽曲生成が一部の感性について可能であることが報告されている。しかし、楽曲の評価基準とされている感性モデルは楽曲の部分構造のみを示すものであり、音楽としての展開や構成といった全体構造が考慮されていない。

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ことを目的として、楽曲の部分構造と全体構造の両者を考慮した自動作曲システムが提案されている[西川 09, 大谷 09]。統一感をもちつつも、起承転結の展開がみられる楽曲が優れていると考えられるため、全体構造の評価では楽曲全体の統一感と楽曲中の展開という2つの指標を設定する必要がある。トレードオフの関係とも考えられる両指標を高めるために、モチーフを和音進行の生成に活用している。モチーフとは楽曲を構成する最小単位であり、基本的に2小節からなるとされる。西川ら[西川 09]は単純GAにより和音進行を生成する手法を提案しているが、感性モデルに即していないモチーフが用いられる可能性があることに加え、複数のモチーフを組み合わせた和音進行を生成することができない。以上の問題を解決する手法として、大谷らはGAの1手法である共生進化[Moriarty 96]を用いて和音進行を生成する手法を提案した。感性の反映度については単純GAよりも高評価が得られたが、統一感について劣るという結果が報告されている[大谷 09]。

本研究では、音楽家の即興過程を模倣した最適化アルゴリズムであるハーモニーサーチ(Harmony Search; HS)により

和音進行を生成する手法を提案する。聴者の楽曲に対する印象は、和音進行のみならずメロディに大きく左右される。前述の研究では和音進行の生成手法に主眼を置いており、メロディは和音進行に合わせてランダムに生成されるため、目的の感性に即したメロディが付与されるとは限らず、楽曲の評価が和音進行手法の評価に一致しない可能性がある。本研究における評価実験では、和音進行にベースパートを加えたメロディなしの楽曲を被験者による評価対象とし、HSと共生進化により生成された和音進行を比較する。

## 2. 部分構造と全体構造を考慮した楽曲生成

本節では、本研究における楽曲の表現形式および生成手法について概説する。なお、ここで扱う楽曲は4/4拍子とし、四分音符の和音の並びで楽曲が構成されるものとする。

### 2.1 表現形式

楽曲は、枠組構造と和音進行から構成される。枠組構造は、楽曲のジャンル、キー、音階、調、速さ、基本リズム、旋律の音色、旋律の音色のカテゴリ、和音の音色、和音の音色のカテゴリという10要素からなる。和音進行は和音の並びであり、和音はRoot, Type, Tensionの3要素からなる。Rootは和音のキー音、Typeは和音を構成する数と構成音ごとの関連性、Tensionは付加的に用いられている音を表す。前の和音を継続して演奏する場合は、(Root, Type, Tension)の代わりに-で和音を表現する。

小節は和音4つの並びで表現され、モチーフは小節2つの並びで表現される。いずれも表現形式は述語論理である。楽曲を表す述語musicの例を図1に示す。ここで、song\_frameは枠組構造を表す述語、barは小節を表す述語である。

感性モデルは、特定の感性に影響する楽曲の部分構造である。モチーフと和音進行の感性モデルの例を図2に示す。述語motifの定義は、あるユーザが「優しい」と感じる楽曲が1小節目がIV度majorの和音2拍分と任意の和音2拍分、2小節目が任意の和音4拍分となるモチーフを持つことを意味

連絡先: 大谷紀子, 東京都市大学環境情報学部

〒224-8551 横浜市都筑区牛久保西 3-3-1, 045-910-2938

E-mail: otani@tcu.ac.jp

```
music(1,
  song_frame(pops,a,a,major,four_four,andante,
    vibraphone,vibraphone,steel_guitar,
    guitar),
  [bar((i,dim7,null),-, (vi,7,null),-),
  bar((iiib,dim7,null),-, (i,major,ninth),-),
  bar((iii,7,null),-, (vi,seventh_aug,null),
    (vib,major,null)),
  bar((is,minor_7b5,null),-, (i,major,ninth),-),
  bar((ivs,minor_7b5,null),-, (iiib,add9,null),
    (vii,add9,null)),
  bar((ivs,minor_7b5,null),-, (vi,7,null),-),
  bar((iiib,add9,null),-, (i,add9,null),-),
  bar(-,-,-)]
).
```

図 1: music の例

```
motif(affectionate,A) :-
  motif(A,bar((iv,major),-,_,-),bar(-,-,-,-)).
chords(affectionate,A) :-
  has_chord(A,B,C),root(C,vi),type(C,minor),
  next_to(A,B,D,_),
  has_chord(A,D,E),root(E,vi).
```

図 2: モチーフと和音進行の感性モデルの例

する。述語 chords の定義は、あるユーザが「優しい」と感じる楽曲が 1 つめの和音が VI 度 minor, 2 つめの和音が VI 度となる連続した 2 和音を持つことを意味する。

## 2.2 生成手順

楽曲の生成手順は以下の通りである。

1. 感性モデルを学習するための訓練例を生成
2. ILP システム FOIL を用いて 1. から枠組構造, モチーフ, 和音進行の感性モデルを獲得
3. 枠組構造の感性モデルを元に枠組構造を生成
4. モチーフと和音進行の感性モデルを元に和音進行を生成
5. 4. を元にベースパートを生成
6. 3., 4., 5. を組み合わせて楽曲を生成

訓練例は、聴者により 5 段階尺度でつけられた既存楽曲の評価を元にして生成する。評価項目は嗜好度, 明るさ, 嬉しさ, 優しさ, 穏やかさであり, それぞれに関して「好き」「明るい」「暗い」など感性の評価軸上で正方向に相当する感性には大きい値, 「嫌い」「暗い」など負方向に相当する感性には小さい値がつけられる。すなわち, 嗜好度に関して「好き」な度合が最も高いときには 5, 「好き」な度合が最も低い, あるいは「嫌い」な度合が最も高いときには 1 の評価がつけられる。「好き」「嫌い」「明るい」「暗い」「嬉しい」「悲しい」「優しい」「優しくない」「穏やかな」「穏やかでない」という 5 対 10 語の形容語に関して, 聴者ごとにそれぞれ 2 種類の訓練例 a, b を生成する。正方向の感性の形容語に関しては, 訓練例 a では評価が 5 のデータを正例, 4 以下のデータを負例とし, 訓練例 b では評価が 4 以上のデータを正例, 3 以下のデータを負例とする。負方向の感性の形容語に関しては, 訓練例 a では評価が 1 のデータを正例, 2 以上のデータを負例とし, 訓練例 b では評価が 2 以下のデータを正例, 3 以上のデータを負例とする。

各訓練例を用いて FOIL により学習し, 10 語の形容語に関する枠組構造, モチーフ, 和音進行の感性モデルを獲得する。以降, 訓練例 a を用いた場合の感性モデルと訓練例 b を用いた場合の感性モデルを下付き文字 a, b で区別することとする。

3. では枠組構造の各要素を遺伝子として, 2 点交叉と突然変異により進化させる。枠組構造  $F$  の適応度  $ffit(F)$  は, 式 (1) で求める。

$$ffit(F) = cov(F, FM_a) \times 3 + cov(F, FM_b) - cov(F, FM'_a) \times 3 - cov(F, FM'_b) \quad (1)$$

$FM_a$  と  $FM_b$  は対象形容語に対する枠組構造の感性モデル,  $FM'_a$  と  $FM'_b$  は対象形容語の反意語に対する枠組構造の感性モデルである。 $cov(F, FM)$  は, 感性モデル  $FM$  の節のうち, 枠組構造  $F$  が満たしている節が感性モデル生成時に被覆していた正例の数である。

和音進行は HS により生成する。詳細は次節で述べる。各小節のベースパートとして, 和音進行のルート音を付加し, 枠組構造と合わせて 1 つの楽曲とする。

## 3. HS に基づく和音進行の生成

本節では, HS に基づく和音進行の生成手法について説明する。モチーフと和音進行の感性モデルを評価に用いることで, 感性モデルに即したモチーフが和音進行に含まれるようにする。また, 複数のモチーフを組み合わせた和音進行も生成され得るため, 単純 GA に基づく手法のような問題は発生しない。

### 3.1 HS の概要

HS は音楽家の即興過程を模倣した最適解探索アルゴリズムである [Geem 01]。音楽家が即興演奏する際には, 次のいずれかの方法を採る。

1. 既知のフレーズをそのまま奏でる
2. 既知のフレーズの一部を変更して奏でる
3. 新しいフレーズを作成して奏でる

HS では, 解候補をハーモニー, 解候補集合をハーモニーメモリ (Harmony Memory; HM) と呼ぶ。上記の演奏方法に対応する以下の 3 つの手法により新しいハーモニーを生成する。

1. HM から 1 つのハーモニーを選択する
2. HM 内の 1 つのハーモニーを調整する
3. 新しいハーモニーをランダムに生成する

最適化アルゴリズムには多様性と収束性が不可欠であるが, HS では, ハーモニーの調整とランダム生成により多様性を維持し, HM により収束性を保証する。HS の疑似コードを図 3 に示す。ここで  $G$  は解候補生成の繰返し回数,  $R_c$  は HM 内のハーモニーを選択する確率,  $R_a$  はハーモニーを調整する確率,  $f(x)$  は目的関数である。

### 3.2 和音進行生成における HS

1 つのハーモニーでは, 図 4 に示すように, root.type 音と tension 音が交互に並んでいる。root.type 音は Root と Type の組合せに振られた ID を表す。既存楽曲に含まれる 75 通りの (Root, Type) と - にそれぞれ 0~75 の整数を ID として割り当てている。tension 音は Tension に振られた 0~7 の ID を

```

HMの初期化;
worst := HM内の最悪ハーモニー;
worstfit := f(worst);
for i := 1 to G {
  r1 := 0.0~1.0の乱数;
  if(r1 < Rc) {
    new := HMからランダムに選択したハーモニー;
    r2 := 0.0~1.0の乱数;
    if(r2 < Ra) {
      new := 制限内で調整されたnew;
    }
  } else {
    new := 新たに生成したハーモニー;
  }
  newfit := f(new);
  if(newfit > worstfit) {
    worstをnewで置き換える;
    worst := HM内の最悪ハーモニー;
    worstfit := f(worst);
  }
}

```

図 3: HS の疑似コード

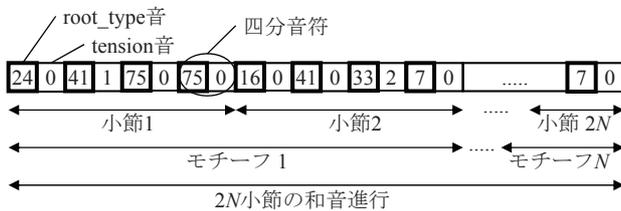


図 4: ハーモニーの構造

表す。root\_type 音と tension 音の組が 1 つの四分音符に変換され、4 つの組が 1 つの小節、8 つの組が 1 つのモチーフとなる。2N 小節の楽曲は、N 個のモチーフの列として表現される。和音進行を表すハーモニー C は目的関数  $cfit(C)$  の値により評価する。複数のモチーフを用いた場合の統一感を目的関数に盛り込むため、1 つの和音進行に含まれるモチーフ数 N を和音進行に関する項に乘じ、モチーフと和音進行を同程度に考慮するようにした。

$$\begin{aligned}
 cfit(C) = & \sum_{M \in C} \{ motif(M) + builtin(M) \} \\
 & + \{ chords(C) + builtin(C) \} \\
 & + forms(C) \times N
 \end{aligned} \tag{2}$$

$M \in C$  はモチーフ M が和音進行 C に含まれることを示す。また、音楽理論における和音進行の禁則に関するペナルティ関数を  $builtin$ 、楽曲全体の展開度合を示す関数を  $forms$  と表す。 $builtin$  により不自然な和音進行の発生を抑制し、 $forms$  により起承転結の展開をもつ楽曲の生成を促進する。 $motif$  と  $chords$  はそれぞれモチーフ、和音進行に関する感性モデルへの適合度であり、 $ffit$  と同様に  $cov$  を用いて算出される。

$$motif(M) = cov(M, MM_a) \times 3 + cov(M, MM_b) - cov(M, MM'_a) \times 3 - cov(M, MM'_b) \tag{3}$$

$$chords(C) = cov(C, CM_a) \times 3 + cov(C, CM_b) - cov(C, CM'_a) \times 3 - cov(C, CM'_b) \tag{4}$$

$MM_a$  と  $MM_b$  は対象形容語に対するモチーフの感性モデル、 $MM'_a$  と  $MM'_b$  は対象形容語の反意語に対するモチーフの感性

表 1: パラメータ

	パラメータ	値
共生進化	突然変異確率	0.25
	部分解集団の個体数	96
	全体解集団の個体数	100
	世代交代回数	1000
H S	繰返し回数 G	100000
	HM 内のハーモニーの選択確率 $R_c$	0.85
	HM 内のハーモニーの調整確率 $R_a$	0.30
	HM 内のハーモニー数	500

モデル、 $CM_a$  と  $CM_b$  は対象形容語に対する和音進行の感性モデル、 $CM'_a$  と  $CM'_b$  は対象形容語の反意語に対する和音進行の感性モデルである。

HM から選択したハーモニーを調整するときは、次のいずれかの変更を施す。

- (Root,Type) の組を別の組または - に変更
- Root を変更
- Type を変更
- Tension を変更

#### 4. 評価実験

提案手法の効果を確認するため、20 代前半の学生 19 名を被験者として、共生進化に基づく手法で生成された楽曲、および提案手法で生成された楽曲を比較する評価実験を行なった。以下の結果では、前者の手法を共生進化、後者の手法を HS と呼ぶ。共生進化では、全体解の適応度を式 (2) により算出することで、解候補を HS と同様に評価できるようにした。

被験者が既存楽曲 53 曲に付与した評価値を元に、10 個の形容語に関する感性モデルを獲得した。各感性モデルを用いて、各形容語に対応する 8 小節の楽曲を生成した。

和音進行の生成以外の部分は、両手法で同じ手順に従っている。両手法で用いたパラメータの値を表 1 に示す。HS では 1 度に 1 つの解候補しか生成しないのに対し、共生進化では各世代で全体解が進化するため、共生進化における全体解集団の個体数と世代交代回数の積が、HS における繰返し回数と等しくなるように設定した。

##### 4.1 感性の反映度に関する評価

生成された楽曲が被験者の感性を反映している度合を比較するため、既存楽曲の評価と同様に、生成された楽曲の評価を 5 つの評価項目について 5 段階尺度で調査した。

HS により生成された楽曲の評価結果を図 5 に示す。ここでの評価値平均とは、ある形容語の感性モデルに基づいて生成された曲の対象形容語に対応する評価項目の評価値の全被験者平均を意味する。例えば、「明るさ」の「正方向」の評価値平均は、ある被験者が「明るい」と感じるように生成された楽曲に対して同じ被験者により付与された「明るさ」の評価値の平均である。正方向の形容語でより高く、負方向の形容語でより低くなる方が感性を反映しているといえる。

全体的に評価値が低い傾向がみられ、正方向の形容語で評価値平均が標準の 3 を超えたのは「穏やかな」のみであったが、すべての評価項目に関して、正方向と負方向に正しく差異があらわれた。t 検定により、有意水準 5% で明るさと穏やか

さ、有意水準 1% で穏やかさに関して有意差があるとの結果が得られた。

HS と共生進化の評価値平均を図 6 で比較する。共生進化も HS と同様、全体的に評価値が低い結果となった。いずれの感性についても有意差は見られなかったが、「暗い」「悲しい」については HS で比較的良好な評価が得られている。また、共生進化では「嬉しい」と「悲しい」の評価が逆転している。

メロディの有無による評価値の変動を調査する追加実験を 10 人の被験者に対して行ったところ、対象形容語の正方向 / 負方向に関わらず、メロディがない楽曲の方が評価値が低くなった。嗜好度の正方向の形容語と穏やかさの負方向の形容語に関して特に変動が大きい。メロディの音列は和音進行を元にランダムに決定しているが、メロディを奏でる音色は枠組構造に含まれているため、感性モデルの一部が楽曲に対する印象に影響を与えた結果とも考えられる。

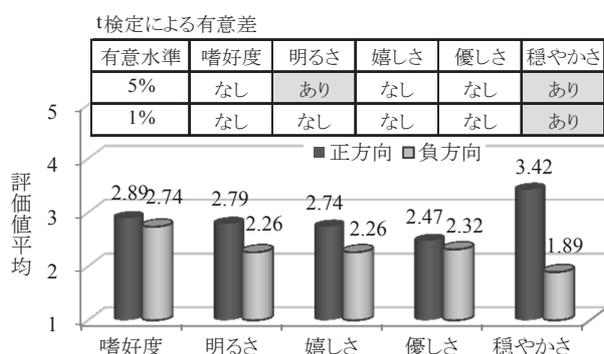


図 5: 感性の反映度に関する評価

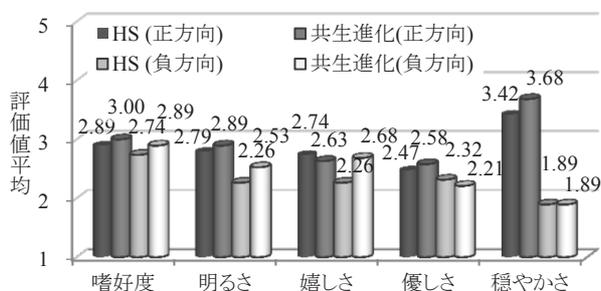


図 6: HS と共生進化の評価の比較

#### 4.2 楽曲の完成度に関する評価

統一感、展開性、面白さ、成立性という 4 つの評価項目について、生成された楽曲の完成度を調査した。各被験者は、自分の感性に合わせて生成された楽曲に対し、「統一感があるか」「展開があるか」「楽曲として面白いのか」「楽曲として成立しているか」に関してそれぞれ 5 段階尺度で評価値を付与する。

HS と共生進化の評価値平均を図 7 に示す。両手法の結果に有意差は見られないが、先行研究で課題となっていた統一感に関して、他の項目よりも評価値が高く、特に HS では 3.4 という高評価が得られた。いずれの評価項目に関しても、評価値平均が標準の 3 以上の値となり、評価項目ごとのばらつきも少ない。トレードオフの関係といえる統一感と展開性に加え、面白さに関して高評価が得られたことから、提案手法では単

純 GA に基づく手法における問題点を解消しつつ、完成度を高められたといえる。

メロディの有無による評価値の変動を調査する追加実験では、すべての評価項目に関して、メロディがない楽曲の方が評価値が高くなった。メロディは和音進行に合わせて生成しているものの、音列決定に音楽理論等は加味されていないため、完成度を低くする要因となっている。

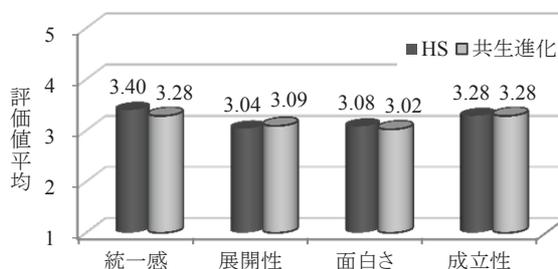


図 7: 楽曲の完成度に関する評価

## 5. おわりに

本研究では、個人の感性に即した完成度の高い自動楽曲生成を目的として、HS により和音進行を生成する手法を提案した。評価実験の結果、統一感や展開性のような相反する性質についても、バランスよく完成度の高い楽曲が生成されることが示された。また、メロディがない状態でも、正方向と負方向の感性を正しく反映できることがわかった。今後、追加実験で明らかになった傾向を考慮して、メロディの生成方法について検討するとともに、楽曲に感性をより反映する手法を考案する。

## 謝辞

本論文は、平成 23 年度物質・デバイス領域共同研究拠点における共同研究課題 (2011233) の研究成果の一部を取りまとめたものである。ここに記して謝意を表したい。

## 参考文献

- [Geem 01] Geem, Z., Kim, J., and Loganathan, G.: A New Heuristic Optimization Algorithm: Harmony Search, *Simulation*, Vol. 76, pp. 60–68 (2001)
- [Legaspi 07] Legaspi, R., Hashimoto, Y., Moriyama, K., Kurihara, S., and Numao, M.: Music Compositional Intelligence with an Affective Flavor, in *Proc. of ACM International Conference on Intelligent User Interfaces*, pp. 216–224 (2007)
- [Moriarty 96] Moriarty, D. and Miikkulainen, R.: Efficient Reinforcement Learning through Symbiotic Evolution, *Machine Learning*, Vol. 22, pp. 11–32 (1996)
- [西川 09] 西川 敬之, 大谷 紀子, 福井 健一, 森山 甲一, 栗原 聡, 沼尾 正行: 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, 第 23 回人工知能学会全国大会予稿集, 1F1-1 (2009)
- [大谷 09] 大谷 紀子, 西川 敬之, 栗原 聡, 沼尾 正行: 楽曲生成への共生進化の適用に関する検討, 第 23 回人工知能学会全国大会予稿集, 3H3-2 (2009)