

リズムを含めた音楽構造による個人感性獲得機構

Acquiring listeners' sensibility information using a music structure containing rhythm

上田 明頌*¹ 西川 敬之*¹ 福井 健一*² 森山 甲一*² 栗原 聡*² 沼尾 正行*²
 Akinobu Ueda Takayuki Nishikawa Kenichi Fukui Koichi Moriyama Satoshi Kurihara 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 Information 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

Sensibility is a significant factor in music. We've focused on the system that automatically composes music adapted to one's sensibility. The system learns a sensibility model represented by first order logic from subjects' evaluations of tunes on the assumption. The system composes music using GA whose fitness function is based upon the relevant elements obtained in learning. In this paper, we propose the introduction of rhythm. As a result, the performance of the composing music system is expected to improve in quality.

1. はじめに

1.1 研究背景

我々はこれまで個人の感性を反映させる自動作曲システムの開発・改良が行われてきた。[1] 個人の感性を反映させる作曲とは、あるひとのある感情（例えば明るい、暗い、楽しい、悲しいなど）を喚起させる曲を生成することである。

楽曲には「切ない」や「哀しい」といった感情を喚起する要素が含まれているが、ある人にとって「切ない」楽曲が、別の人にとってもそうであるとは限らない。また「明るい」「暗い」のようにある程度万人に共通する感情もあれば、「好き」「嫌い」のように人によって全く違うものもある。このように、人の感性とは非常に複雑なものであり、それを反映させることは難しい。さらに作曲だけをとっていても、人が普段自然に「音楽」として聞いているものは非常に有機的なものであり、ただ単に音楽理論をあてはめたシステムではまともな曲は生まれにくい。

このようなことを考えると、音楽として成立した曲であり、かつ個人の感性を反映させた曲をつくることは非常に難しい。しかしながら、作曲に対するニーズが多岐多彩にわたっている現代の音楽産業の中で、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。

1.2 研究の目的

本研究は、人間の感性を学習し、それをを用いて個人に合わせた「切ない曲」や「哀しい曲」等を自動生成することを目的としている。

これまでの研究では感性の学習で用いる楽曲構造の特徴量として、調（キー）、拍子、テンポ、メロディ楽器、伴奏楽器が扱われてきた。しかし、音楽の重要な要素であるリズムに関する特徴量は扱われておらず、この自動作曲システムによって多種多様なジャンルの音楽を生成することは困難であった。そこで、本研究ではリズムに関する新たな特徴量としてグルーブという概念を導入した。さらに、様々なリズムの音楽を表現するのに最低限必要と考えられるリズム楽器（ベース、ドラム）

を新たにシステムに追加した。これによって、より個人の感性に訴える作曲が可能となる。

2. システムの概要

まず、感性を学習する為に、ユーザに既存の楽曲を聴かせて評価値を得る。次に、楽曲の情報とユーザの評価値から、ユーザの感性と関連のある楽曲構造を、帰納論理プログラミング (ILP) を用いて一階述語論理形式で学習する。そして、得られた述語をもとにした適応度関数を組み込んだ遺伝的アルゴリズムにより作曲を行う。この一連のプロセスは従来のシステムと同じであるが、本研究ではこのシステムを基にリズムに関わる音楽的要素を新たにシステムに加えることによってより質の高い自動作曲を行う。

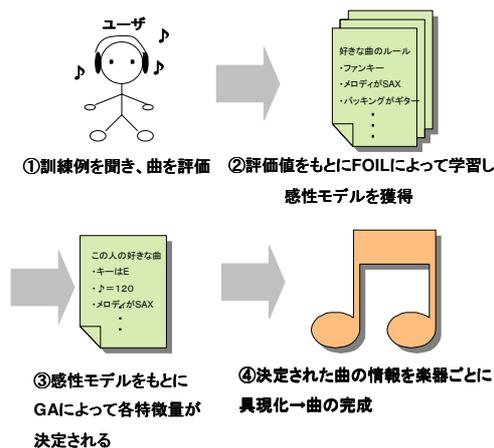


図 1: システム概要図

連絡先: 上田明頌, 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:ueda@ai.sanken.osaka-u.ac.jp

3. 帰納論理プログラミングによる感性情報の学習

感性を高い精度で学習できるかどうかは以下で述べる音楽知識表現と目標述語の定義に依存する。

3.1 楽曲の述語表現

本研究では、楽曲の枠組構造と和音構造の二種類を扱う。楽曲を以下の述語で表現する。

- `song_frame/10` 楽曲の枠組を表す述語であり、その引数はジャンル、キー、主音、調性、テンポ、拍子、メロディ楽器、メロディ楽器の種類、伴奏楽器、伴奏楽器の種類である。
- `chord/3` 和音構造を表す述語であり、その引数は調、ルート、コードのタイプである。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。

3.2 目標述語

以上に述べた知識表現を用い、被験者の感性を刺激する楽曲構造についての述語獲得を行う。述語は、`frame/2`、`chords/2` の 2 種類があり、以下の様な意味を持つ。

- `frame/2` 楽曲の枠組構造に関する述語である。被験者のある感性に影響を与える楽曲の枠組構造について説明する働きを持つ。
- `chords/2` 楽曲中に含まれる和音進行に関する述語である。被験者のある感性に影響を与える和音連結の形態について説明する働きを持つ。

これらの目標述語を学習するために本研究では ILP のひとつである FOIL を使用する。

3.3 被験者による楽曲評価と訓練例生成

被験者は楽曲を 5 つの形容詞対について 5 段階評価する。評価の仕方については、例えば「好き」な曲ほど 5 段階評価の高い数値を与えるように、また「嫌い」な曲ほど 5 段階評価の低い数値を与えるように被験者に指示した。他の形容詞対についても同様に「正方向」「負方向」をあらかじめ定義し、5 段階評価と対応させている。

得られた評価値と述語で表現された楽曲構造から、2 種類の目標述語にそれぞれ専用の訓練例を生成する。`music/3` の第二引数 `song_frame/10` から `frame/2` の学習のための訓練例、`music/3` の第三引数 `chord/3` のリストから `chords/2` の学習のための訓練例をそれぞれ生成する。

4. 遺伝的アルゴリズムを用いた作曲

ユーザの感性を反映させた楽曲の枠組構造と和音進行が遺伝的アルゴリズムによって決定される。枠組構造と和音進行が決められる過程は互いに独立しており、二つの過程の出力結果を合わせて一つの楽曲の構造データとする。以下で、それぞれの過程での遺伝的アルゴリズムの利用法について示す。

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

図 2: 枠組構造の染色体

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

4.1 感性を反映させた枠組構造の獲得

- 染色体表現
各遺伝子座が 2.1 節の楽曲の枠組構造に対応する様に染色体を定義する。(図 2)
- 適応度関数
3.1 節で得られたユーザの枠組構造についての感性モデルを適応度関数として用いる。その際、用いる感性モデルは次の 4 つである。

1. ユーザの評価で 5 を獲得した訓練例楽曲を正例、4 以下を獲得した訓練例楽曲を負例として学習した結果得られた感性モデル
2. ユーザの評価で 4 以上を獲得した訓練例楽曲を正例、3 以下を獲得した訓練例楽曲を負例として学習した結果得られた感性モデル
3. ユーザの評価で 2 以下を獲得した訓練例楽曲を正例、3 以上を獲得した訓練例楽曲を負例として学習した結果得られた感性モデル
4. ユーザの評価で 1 を獲得した訓練例楽曲を正例、2 以上を獲得した訓練例楽曲を負例として学習した結果得られた感性モデル

感性の評価軸上で正方向に相当する感性(「好き」「明るい」など)を反映させる際には、1. と 2. を正の適応度関数、3. と 4. を負の適応度関数として用いる。感性の評価軸上で負方向に相当する感性(「嫌い」「暗い」など)を反映させる際には、3. と 4. を正の適応度関数、1. と 2. を負の適応度関数として用いる。各感性モデルは、2 章で示したように、複数の規則から構成されている。これは、感性モデルを獲得する手法である FOIL がカバーリングアルゴリズムを採用しているためである。感性モデルを適応度関数として扱う時には、各規則に対して点数を割り当てる必要がある。本システムでは、FOIL が規則を出力した際にその規則がカバーしていた正例の数をその規則の得点としている。さらに、1. と 4. に含まれる規則の得点については、2. と 3. に含まれる規則の得点の 3 倍の重み付けをしている。

4.2 感性を反映させた和音進行の獲得

- 染色体表現
1 つの遺伝子を 1 小節 1 和音 (`chord/3`) に対応させ、染色体を定義する。図 3 で表された染色体は、8 小節の長さを持つ楽曲を表している。

● 適応度関数

適応度関数を以下のように定義する。

適応度関数の定義

$$\text{Fitness_Function}(C) = \text{Fitness_Chord}(C) + \text{Fitness_Builtin}(C)$$

C: 楽曲の和音進行を表す染色体
 Fitness_Function(C): C に対する適応度関数
 Fitness_Chords(C): 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
 Fitness_Builtin(C): 音楽理論における和音進行の禁則に基づく関数

Fitness_Chords(C) は、2.2 節で得られたユーザの和音進行についての感性モデルを適応度関数として用いる。感性モデルの適応度関数としての利用方法は枠組構造の獲得のときと同様である。

5. 提案手法：リズムの導入

本システムはリズムの導入に当たって主に 2 つのものを新たにシステムに追加することを提案する。1 つは楽曲の枠組構造に新たにグルーブ*1 という概念を加える。そして、もう 1 つはそのグルーブを表現するためにリズム楽器としてドラムとベースを追加する。

5.1 枠組構造へのグルーブの追加

2.1 節で扱った楽曲の枠組構造はジャンル、キー、主音、拍子、調性、テンポ、メロディ楽器、メロディ楽器の種類、伴奏楽器、伴奏楽器の種類 10 種類であったが、ここに新たにグルーブを追加する。

我々は、リズムを導入したシステムとそうでないシステムでの作曲の質を比較するために被験者実験を行ったが、その中で用いた訓練例 50 曲を 13 種類のグルーブに分類した。つまり、音楽的にみてノリが似ているものごとに分類した。以下に 13 種類のグルーブの名称とその特徴を示す。

- funky: 大きな特徴は、バックビート (裏拍) を意識した 16 ビートのリズム。
- j-pop: シンプルな 8 ビートとルートを中心に 8 分音符で刻むベースラインが特徴。
- rock 'n roll: 主に 2 拍目 4 拍目を強調した 8 ビート。
- rock: 小節の変わり目を特に強調したような 8 ビート。
- ballad: バラードの曲の雰囲気を生み出すようなリズムパターン。他のリズムパターンよりもドラム、ベースは音数も音量も控えめ。
- acoustic ballad: ドラムはスタンダードドラムセットではなくブラシセット、ベースはエレクトリックベースではなくウッドベースによるバラードに対するリズムパターン。
- acoustic ballad: ドラムセットではなくパーカッションを用いているバラードに対するリズムパターン。

*1 一般的に用いられる意味とは少し違うが、本文では曲が主にリズムに関する部分でどういうノリのものかという意味で用いている。音楽のジャンルと深く関係する言葉であり、例えばファンクにはファンクのノリがありジャズにはジャズのノリがある。ここでは、音楽の様々なノリを総称してグルーブと呼ぶことにする。

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

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

図 4: 枠組構造の染色体表現

- pop: 明るくかわいい雰囲気を出すベースパターンと軽快なドラム。(8 ビート)
- pop: ドラムは軽快に 16 分を刻み、ベースは主に 4 分を刻む。全体としてはゆったりとした印象を生む。
- pop: シャッフルよりのはねたビート。
- pop: rock に分類したもののよりはポップであるが、pop ~ pop よりも rock に近い 8 ビート。
- rock ballad: バラードに対するアプローチのリズムパターンであるが他の ballad のパターンよりも、音数も多く音量のやや大きい。
- dance&club: earth wind&fire を代表とするダンスミュージックでよく使われるような 16 ビート。

グルーブの種類とジャンルの種類は深く関連しているので本論文では個々のグルーブの名前に一般的なジャンルの名前を代用して表現している。

ここで、2.1 節の楽曲の枠組構造にグルーブが加わることで生じるシステムの変更点について説明する。まず、楽曲の枠組構造が 1 つ増えたので 3.1.3 節の song_frame/10 が song_frame/11 となる。また、枠組構造の染色体表現にも図 4 のようにグルーブが追加される。

5.2 ドラム・ベースの出力

決定されたグルーブを表現するためにリズム楽器としてドラムとベースを追加する。出力のさせ方としては予め個々のグルーブに対応するドラムパターン、ベースパターンをセットにしてリズムパターンとして用意しておき、ドラムは用意したパターンをそのまま出力し、ベースは楽曲のコード進行に対応するように変換させてから出力させる。本システムでは 13 種類 22 パターンのリズムパターンを用意した。13 種類のグルーブに対し、22 個のリズムパターンがあるのは、1 つのグルーブに対し 2 種類のパターンを用意したものと 1 種類しか用意していないものがあるためである。1 つのグルーブに対して 2 種類のパターンがあるものは GA によって枠組構造が決定された後、2 種のパターンのうちのどちらかのパターンがランダムに選択される。例えば、枠組構造でグルーブが funky と決定されると、出力されるリズムパターンは funky1, funky2 のどちらかがランダムに選択される。

ベースパターンの変換

ベースはドラムとともにその曲のグルーブを決める中心となり、更にはコードを強調する役割を担っている。したがって、ベースを出力させるには決定されたグルーブとコード進行の両方を反映させなくてはならない。本システムでは、各グルーブに対し予めベースのリズムパターンを用意しておき、そのパターンを決められたコード進行に添う音を用いて出力させるこ

とにより、グルーブとコード進行の両方を反映させるようにしている。

例えば、図5はグルーブが funky の場合に予め用意された実際のベースパターンの一つである。枠組構造が決定されグルーブが funky となればその時出力されるベースパターンは図5のベースパターンを元に作られる。そして、実際にあるコード進行に対応させたのが図6のベースパターンである。図6のベースパターンは図5のパターンと全く同じ位置に音符(休符)があるが、その音符の一つ一つは全く違う音となっている。このように音符(休符)の位置はそのままにして弾く音のあるルール(22パターンそれぞれに決められている)に従うように変化させることでグルーブとコード進行の両方に対応させたベースパターンを出力できるようにしている。このパターンの場合のルールは1小節分全てを書くとき1音目はルート、2音目は3度、3音目は5度、4音目は1オクターブ上のルート、5音目は2度、6音目は3度、7音目は5度、8音目は1オクターブ上のルートというルールに従ってベースパターンが生成されている。



図5: 予め用意したベースパターンの例



図6: 予め用意したベースパターンを元に別のコード進行に対応させたパターンの例

6. 実験

リズムを考慮していない旧システムとリズムを導入させた本システムでどちらがより感性を反映させた精度の高い作曲ができるのかを実験した。

実験方法

1. ポップスを中心とした既存の楽曲50曲を10代~30代の男女10人に提示し、嗜好度(嫌い・好き)、明るさ(暗い・明るい)、安心さ(不安・安心)、嬉しさ(悲しい・嬉しい)、切なさ(切なくない・切ない)の5対の感性語について5段階評価してもらう。
2. 被験者の評価データをもとに10感性語についての感性情報を旧システム、本システムの両方でそれぞれ学習する。
3. 学習した感性情報を基に10感性語についてそれぞれ3曲ずつ両システムで作曲し(計60曲)、それらの曲に対してと同様に5段階評価してもらう。被験者は提示された曲がどの感性語について生成されたものなのか、また提示された曲が2種類の作曲システムによって作られていることを知らされない。
4. さらに、提示された曲の表現の豊かさ、曲の完成度についてもそれぞれ5段階評価してもらう。

実験結果

新旧それぞれの作曲システムによって生成された曲の評価値

		平均値 (旧/新)	t 値 (旧/新)	有意差 (旧/新)
嗜好度	嫌い	2.20/2.40	1.28/2.57	なし/あり
	好き	2.60/2.80		
明るさ	暗い	1.63/2.27	4.58/2.49	あり/あり
	明るい	3.27/3.20		
安心さ	不安	1.80/2.10	3.98/6.05	あり/あり
	安心	2.57/3.23		
嬉しさ	悲しい	2.17/2.57	1.63/2.24	なし/あり
	嬉しい	2.53/3.23		
切なさ	切なくない	2.77/2.37	2.97/2.44	あり/あり
	切ない	3.47/3.13		

表1: 感性の反映度合いに関する評価結果

	旧システム	本システム
平均値	2.83	3.14
標準偏差	0.05	0.21
t 値	2.57	
有意水準 5% での有意差	有意差あり	
有意水準 1% での有意差	有意差なし	

表2: 楽曲の質的評価結果(表現の豊かさ)

	旧システム	本システム
平均値	2.55	3.10
標準偏差	0.15	0.31
t 値	3.41	
有意水準 5% での有意差	有意差あり	
有意水準 1% での有意差	有意差あり	

表3: 楽曲の質的評価結果(曲の完成度)

の t 検定結果を下の表に示す。旧システムによる結果では嗜好度、嬉しさにおいて有意差はないが本システムでは有意差が認められている。(有意水準 5%)

また、表現の豊かさと曲の完成度についても t 検定を行ったところ、いずれも本システムによる値の方が高いといえることが分かった。(表現の豊かさ:有意水準 5%, 完成度:有意水準 1% での有意差あり)

7. まとめ

音楽において重要でありながら、我々の自動作曲システムでは今まで扱われることのなかったリズムをシステムに導入することを提案した。実験結果からリズムを導入した本システムのほうがより個人の感性を反映させる作曲が可能であるということが示された。

参考文献

- [1] 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)
- [2] 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)