

## 強化学習に基づく MIDI 楽曲選曲システムの提案 MIDI Music Classifier System Based on Reinforcement Learning

森田 央<sup>†</sup>      六井 淳<sup>†</sup>  
Morita Hiroshi      Rokui Jun

### 1. はじめに

現在、我々が暮らす日常には数多くの音楽情報が存在している。個人でも数万曲の楽曲を所有し、容易に数万曲の音楽データを音楽機器や音楽再生インターフェイスに保存し持ち歩く事が可能となっている。しかし、音楽データが増えるとともに、それらの中からユーザの好みの楽曲を見つけ出すのは困難になっている。

多くの音楽情報を扱う音楽情報処理の分野では、音楽構造解析、音源分離、楽曲分類、ランダム再生などの様々な研究がなされている[1]。音楽構造解析の分野においては音楽認知モデルが注目を集めており、音楽理論 GTTM(Generative Theory of Tonal Music)[1]を用いた音楽構造解析の研究が盛んに行われている。GTTMを用いることにより、音楽理論に基づいた音楽構造解析が可能になると考えられている。本来、この理論に基づいた音楽構造解析を行うことが期待されているが、GTTMはコンピュータ上への実装が困難であるという問題点がある。

現在の音楽情報処理の分野では、ユーザの嗜好に合う音楽推薦・音楽検索システムの研究が行われており、ユーザの好みを判断する手法としてフィルタリングが用いられている。吉井らによる音楽推薦システム[2]では、履歴や楽曲の類似性によりユーザの好みに合う楽曲推薦の研究が行われている。黒瀬らによる感性情報を用いた楽曲推薦システム[3]では、感性情報を用いたフィルタリングによる楽曲推薦の研究がなされている。しかし、これらの研究ではユーザ毎の感性の個人差を反映した検索はなされていない。ユーザの好みを判断するためには、個人差のある感性を機械に学習させる必要がある。

そこで本研究では、ユーザ毎の感性に従って楽曲を分類し選曲するシステムを提案する。本システムは単なる好みの選曲システムではなく、個人の感性を関数として表現することを目的としている。楽曲を個体とみなし、ユーザの好みという探索空間内で生存競争を行わせていくことで、ユーザにとって最適な個体を発見していく。本システムは楽曲の音楽情報を解析し、個体の遺伝子として扱う。分類子を用いて楽曲をユーザの好みという探索空間へ写像し、楽曲=個体は探索空間内で選択、淘汰される。そこで決定した楽曲をユーザが視聴し、結果に対する評価をシステムへフィードバックすることで分類子を更新し、世代交代を行う。分類子がユーザに適さないと判断された場合、分類子自体を変更し、突然変異を行う。世代交代を繰り返すという操作により、分類子をユーザの好みへ収束させていく。本システムは分類子を用いることでユーザの感性に対応させる創発システムである。

本研究では、楽曲データとして MIDI 楽曲を用いた。本システムによって MIDI 楽曲の分類、評価、学習を繰り返す事により、個人に合った分類空間への収束を行う。構築したシステムを用い、被験者に対し分類空間の収束実験を行った。

### 2. MIDI 楽曲選曲システムの概要

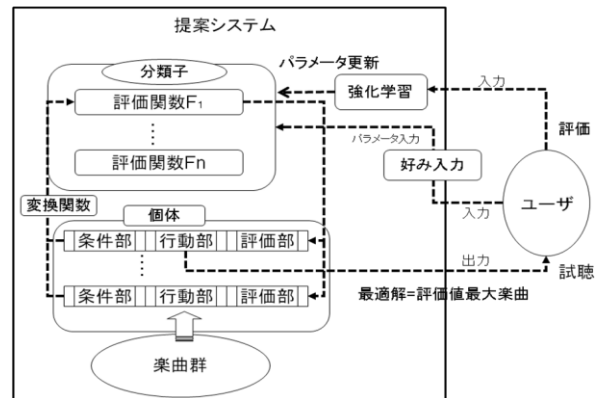


図1 概要図

本論文で提案する MIDI 楽曲選曲システムでは、MIDI 楽曲を利用する。本システムが選曲した MIDI 楽曲をユーザが視聴し、評価を繰り返す事により、ユーザ毎の分類空間を成長させるシステムである。図1に MIDI 楽曲選曲システムの概要図を示す。

本システムでは、楽曲を個体とみなし、個体を評価するために分類空間として分類子を使用する。分類子は複数評価関数からなり、個体は評価関数に基づき選択される。システムは選択によって決定した楽曲をユーザへ提示する。これに対し、ユーザは楽曲を試聴し出力結果への評価をシステムへ返す。システム側はユーザの評価を元に評価関数のパラメータを更新する。この操作を繰り返すことで、分類子を変化させ、収束させていく。つまり、本システムはユーザ毎の感性に対応した分類空間である分類子を生成するシステムである。膨大なデータベース上にある楽曲に対し、ユーザ毎の分類空間を使用、交換することでユーザ毎の好みの楽曲を検索、発見することが可能になる。

### 3. 分類子システム

本研究では遺伝的アルゴリズム(Genetic Algorithms : GA)[4]を利用した分類子システム(Classifier System:CS)[5]を用いる。分類子システムは、プロダクションシステムに基づく実行機能、信頼度割り当てに基づく強化学習機能、遺伝的アルゴリズムに基づくルール生成機能の三つの機能から構成される。

<sup>†</sup> 島根大学大学院総合理工学研究科数理・情報システム学専攻  
Interdisciplinary Graduate School of Science and Engineering  
Department of Mathematics and Computer Science/Shimane  
University

### 3.1 概要

本研究で用いる分類子システムでは、MIDI 楽曲データを個体とみなし、楽曲データの集合を個体群として扱う。図 2 のように、MIDI から生成された音データを行動部とし、音データのバイナリデータから抽出した値をコードデータとする。条件部、行動部を遺伝子情報とし、評価部を付加することで各個体として扱う。評価部は各個体の評価値である。遺伝子情報をもとに評価関数により各個体の評価値が与えられる。個体群中では評価値の高い順に選択が行われ、ユーザが選択結果を評価する。その評価に基づき、強化学習により評価関数のパラメータに報酬が与えられ、評価関数が更新される。これらを繰り返す事により、評価関数を収束させていく。

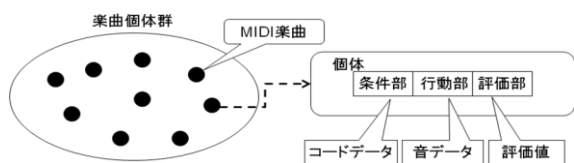


図2 楽曲の扱い

#### 3.1.1 条件部

本研究では、楽曲として MIDI 楽曲を利用する。MIDI は演奏情報を含んだ音楽データのため、詳細な音楽情報が得られる。本システムでは MIDI 楽曲の MIDI バイナリデータから演奏情報を解析し、そのデータを利用し扱う。MIDI はチャンクと呼ばれるブロックで構成される。トラック数などの基本情報が収められたヘッダーチャンクのあとに、複数個のトラックチャンクが並ぶ。最初のトラックチャンクはコンダクターチャンクといい、調、拍子テンポなどの演奏情報が格納されている。その後、続くトラックに実際の演奏情報が格納されている。図 3 に MIDI 楽曲の実際のバイナリデータについて示す。図 3 のように、出音や消音、テンポ等の情報が直接格納されている。この MIDI バイナリデータから音楽情報を抽出し、システム上で利用可能な数値へ変換する。

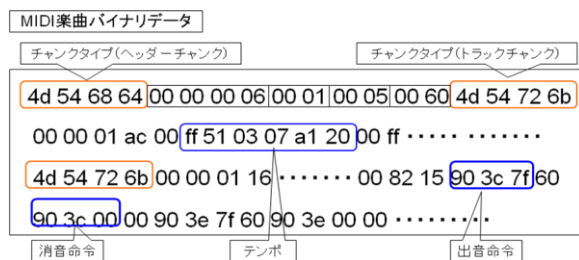


図3 MIDI バイナリデータ例

図 4 は条件部の概念図である。図 4 のように、バイナリデータを評価関数で扱うため、まず変換関数を用いてバイナリデータを評価関数用の値へ変換をする。評価関数は、変換関数により変換された値とパラメータの値を含む関数である。変換関数により変換されたコードデータが評価関数へ入力され、遺伝子毎の評価値が決定される。

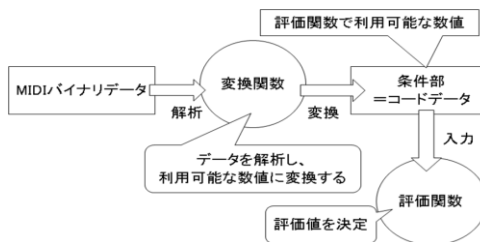


図4 条件部概念図

#### 3.1.2 評価部

評価部は、条件部より得られた値より計算される評価関数の値を扱う。図 5 は評価部の概念図である。図 5 のように、全ての個体の評価部に評価関数により得られた評価値が与えられる。この評価値に基づき、評価値の高い順に選択が行われ、順位付けされる。

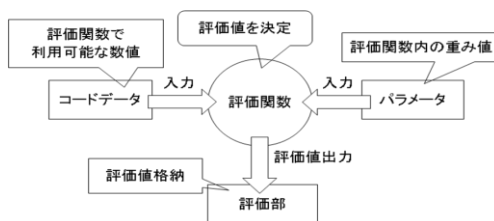


図5 評価部概念図

#### 3.1.3 行動部

行動部は MIDI より生成された wav データを扱う。図 6 は行動部の概念図である。図 6 のように、選択により最も評価値の高い個体が決定されると、評価値が最大の楽曲の行動部、つまり音データが出力される。これをユーザが試聴し、アンケートを行う事により、システムが評価される。このユーザのシステム評価判定に基づき、強化学習を行う。強化学習により、使用されたパラメータへの報酬を決定する。この報酬に基づき、評価関数のパラメータが更新される。

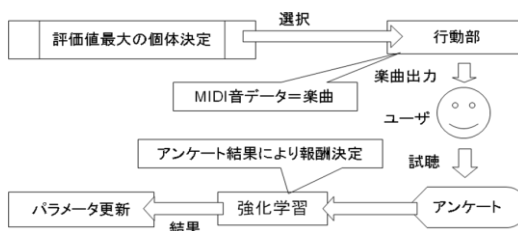


図6 行動部概念図

## 4. 分類空間の収束実験

### 4.1 実験目的

本研究で提案する MIDI 楽曲選曲システムを構築し、強化学習による分類空間の収束実験を行った。強化学習によるパラメータの変化をシステムへフィードバックすることにより、分類空間が変化し特定の楽曲へ収束するか検証を行った。本システムでは創発性に基づくため、アンケート内容とパラメータの間の相関関係は考慮しないものとする。

### 4.2 実験方法

まず、被験者にシステムが出力する楽曲を試聴してもらおう。それに対し、毎回システムが出力楽曲に対するアンケートを要求する。出力楽曲に対しユーザが感じた結果を、アンケート内容に添ってシステムに入力し強化学習を行う。強化学習によりパラメータ値が更新され、再評価が行われる。システムの収束判定により収束するまで繰り返し試聴とアンケートを行う。

次式が本実験で使用したシステム内の評価関数である。

$$F = \sum_{i=1}^n x_n p_n \quad (1.1)$$

$x_n$  はコードデータ値、 $p_n$  はパラメータ値である。本実験では、コードデータとして、特定範囲中の各音符頻度値とテンポ情報を MIDI 楽曲から取得し使用した。ここで、 $n=14$ であり、コードデータ  $x_1$  はテンポ情報である。本実験では、テンポへのバイアスを 0.0001 とし、

$$x_1 = \text{テンポ値} \times 0.0001 \quad (1.2)$$

とした。

楽曲のソート結果に対する収束判定条件は以下の二つとする。

収束条件 1: 今回選ばれた楽曲と、前回選ばれた楽曲の評価値の差の絶対値が 0.01% 未満

収束条件 2: 同じ楽曲が 2 回連続で選ばれる

このどちらかの収束条件を満たした場合、分類空間が収束したと判断する。

パラメータ値  $p_n$  は表 2 の  $n$  番のアンケート番号と対応している。アンケート時の入力に対するパラメータの変化を報酬変化幅  $p_n \pm 7$  とした。アンケートに対する入力が yes の時:  $p_n + 0.7$

no の時:  $p_n - 0.7$

となり、更新が行われる。初期値として入力するパラメータを各  $p_n = 0.5$  とし、実験を行った。楽曲は 1 分前後の MIDI 楽曲 50 曲を作成し使用した。

以上の条件で被験者 4 名に対し実験を行った。表 1 に使用楽曲情報を示す。表 2 に本システムが行ったアンケート内容を示す。

表 1: 使用楽曲情報

楽曲	MIDI(SMF)
フォーマット	Format1
曲数	50

表 2: アンケート内容

番号	内容
1	この曲は好きですか？
2	この曲はアップテンポですか？
3	この曲は明るいですか？
4	この曲は悲しいですか？
5	この曲はシンプルですか？
6	この曲はバラードですか？
7	この曲は面白いですか？
8	この曲は落ち着きますか？
9	この曲は暗いですか？
10	この曲はきれいですか？
11	この曲は楽しいですか？
12	この曲はリズムカルですか？

13	この曲は感動的ですか？
14	この曲は盛り上がりますか？

### 5. 実験結果

分類空間の収束実験により、各被験者において特定楽曲への収束が確認できた。以下に実験結果を示す。表 3 は各被験者における収束楽曲、収束回数、収束条件の実験結果である。

表 3: 実験結果

	収束楽曲	収束回数	収束条件
被験者 1	36 番	6 回	2
被験者 2	5 番	9 回	2
被験者 3	48 番	4 回	2
被験者 4	28 番	4 回	2

#### 5.1 評価値の変化

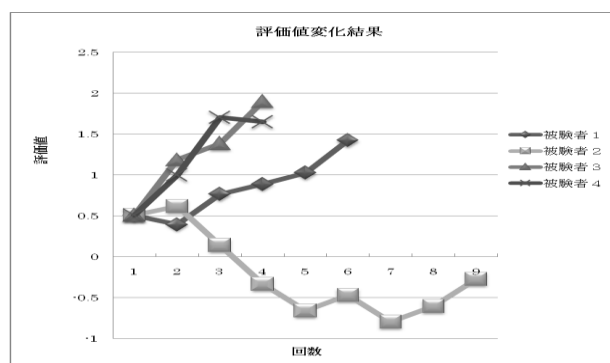


図 9 収束時 1 位楽曲の評価値変化結果

図 9 は被験者 4 人に対して行った実験結果より、各被験者における収束時順位 1 位楽曲の評価値変化のグラフである。このグラフから、被験者 1,3,4 では学習回数により評価値が単調増加傾向にあることが分かる。しかし被験者 2 においては安定しておらず、途中まで評価値の単調減少傾向が見られる。

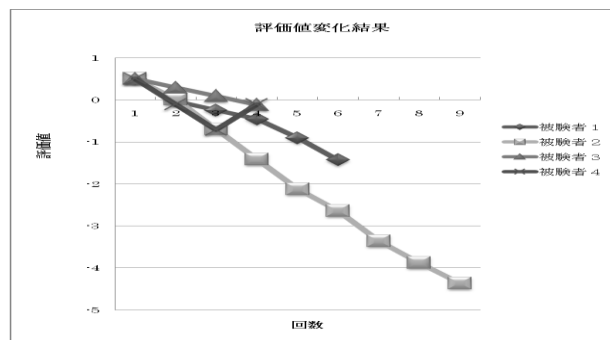


図 10 収束時 50 位楽曲の評価値変化結果

図 10 は被験者 4 人に対して行った実験結果より、各被験者における収束時 50 位の楽曲の評価値変化のグラフである。このグラフから、収束時下位の楽曲では評価値が単調減少傾向にあることが分かる。被験者 2 では学習回数の増加に伴い最下位楽曲の評価値も下がり続けており、他の被験者よりも極端に低い数値となっている。

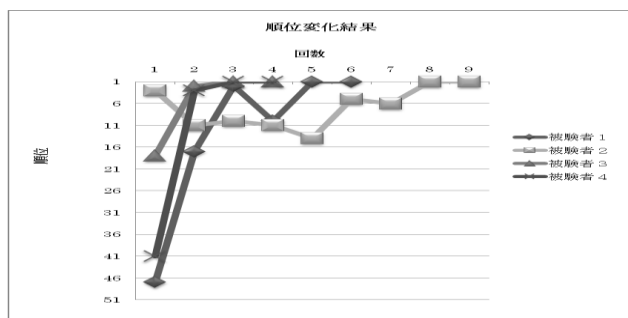


図 11 収束時 1 位楽曲の順位変化結果

図 11 は被験者 4 人に対して行った実験結果より、各被験者における収束時 1 位の楽曲の順位変化のグラフである。初回の順位に違いはあるが、楽曲の評価値の初期値はほぼ同一となっている。このグラフから、被験者 1,3,4 において図 9 の評価値の増加に伴い順位も上昇しているのが分かる。

## 5.2 パラメータ値の変化

5.1 の実験結果より、被験者 2 において他被験者と異なる結果が得られた。被験者 1,3,4 では類似したパラメータ変化が見られたが、被験者 2 では異なるものだった。ここでは被験者 2 のパラメータ値変化結果を示す。

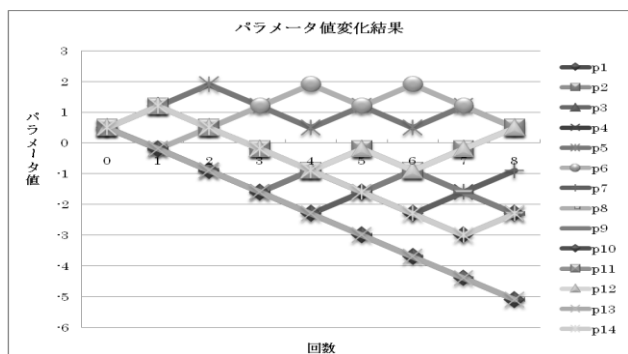


図 12 被験者 2 パラメータ値変化結果

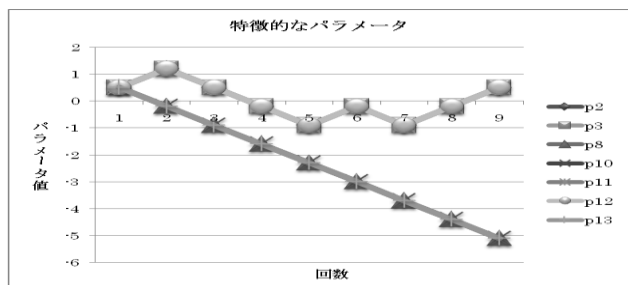


図 13 被験者 2 パラメータ値変化結果

図 12 は被験者 2 に対して行った実験結果より、全てのパラメータ値の変化のグラフである。図 13 はパラメータ値中で特徴的なパラメータを抜き出したものである。図 12 から、p8,p10,p13 が単調減少していることが分かる。そ

のため、p8,p10,p13 が評価値の減少に大きく影響していると考えられる。これに対し、主に増加しているパラメータは見られない。図 11 から、順位が上昇したのが 5 回目以降である。そのため、5 回目以降値が増加した p2,p3,p11,p12 が評価値に影響したと考えられる。図 13 は、図 9 と図 10 の評価値変化に類似しているのが分かる。このことから、p2,p3,p10,p11,p12,p13 のパラメータが最も評価関数へ影響し収束楽曲が決定したのだと考えられる。

## 5.3 考察と問題点

実験結果より、強化学習により分類空間が収束することが確認できた。分類空間の変化により、楽曲の評価値が変化し特定の楽曲への収束が確認できた。被験者 1,3,4 では収束楽曲の評価値が単調増加傾向にあったのに対し、被験者 2 の場合では不安定な結果となった。さらに他の被験者に比べ学習回数が多い。これは、被験者 2 に対して本実験で使用した評価関数が適切でないためだと考えられる。パラメータ値の変化に伴い各楽曲の評価値が変化し、楽曲の決定に大きく影響することが確認できた。このパラメータ値の違いが各被験者の特徴であると考えられる。

本システムでは、初期パラメータ値の取り方や、報酬変化幅、テンポバイアス値によって早期に収束することがある。そのため、これらの値を適当な値にすることが必要である。初期パラメータ値を 0 にすると、初期段階の変化が大きくなることが確認された。本システムでは単純な積和の評価関数のみを使用しており十分であるとは言えない。今回 MIDI 楽曲 50 曲を使用し実験を行ったが、実験対象として十分な曲数とは言えない。今後さらに曲数を増やす必要がある。

## 6. まとめ

本研究では、MIDI 楽曲選曲システムの基礎の構築を行った。システムを構築し実験を行った結果、ユーザの評価による強化学習を行う事により、分類空間の収束が確認された。これにより、分類システムによるユーザ毎の分類空間の生成が可能であると考えられる。今後、本システムを改善し、ユーザの感性による楽曲分類システムの構築を行う。今回、本研究では評価関数を一つとし、システムを構築している。今後は、複数評価関数を使用し、有用な評価関数の選択を行う必要がある。さらに、音楽的な特徴には多様な情報が存在するため、楽曲情報から何をコードデータとして扱うかによって、システムの有効性が変化すると考えられる。今後、これらの課題の解決を目指す。本研究では、システムの有用性を評価するため相当数のユーザに使用してもらう必要がある。そのため現在、iPod などのデバイス用アプリケーションとして開発を目指している。

## 参考文献

- [1] 平田圭二, 平賀譲: GTTM に基づく音楽表現手法再考, 情報処理学会研究報告, [音楽情報科学], 2002(40), pp.1-7, 2002.
- [2] 吉井和佳, 後藤真孝: 音楽推薦システム, 情報処理 Vol.50 No.8 Aug.2009, p751-p755
- [3] 黒瀬崇弘, 梶川嘉延, 野村康雄: 感性情報を用いた楽曲推薦システム, 第 19 回ファジィシステムシンポジウム, 2003
- [4] 坂和正敏, 田中雅博: 遺伝的アルゴリズム, 朝倉書店, 1995.
- [5] Richard S. Sutton, Andrew G. Barto (三上貞寿・皆川雅章 共訳): 強化学習, 森北出版, 2000