

感情の表現力をもつ旋法の獲得のための強化学習モデル

田中 翼^{1,a)} 大村 英史^{2,b)} 古川 聖^{1,c)}

概要：音楽における感情表現の手段として旋法の選択は最も重要な要素の一つであろう。感情と旋法の関係性は古代ギリシャの時代から論じられてきており、近年では実証研究が行われている。しかし、想定しうる旋法の種類は極めて多く、ある感情を最も良く表現しうる旋法はどれかを心理実験によって調べ上げることは困難である。そこで本稿では、人間からの評価値のフィードバックを用いた強化学習のアプローチをとることで、目標となる感情の表現力を高めるように旋法を適応変化させていく手法を提案する。

1. はじめに

長旋法の明るさと短旋法の暗さの間の対照からもわかるように、旋法は感情との結びつきが強い。その関係性については古代ギリシャの時代から議論されてきた [1][2]。例えばプラトンの「国家」[1]においてソクラテスは、教育的な観点から、悲しみを帯びた混合リディア旋法や高音リディア旋法、柔弱なイオニア旋法等は排除すべきであり、勇気と節度を表すドリア旋法とフリギア旋法を用いるべきだと論じている^{*1}。

近年では、旋法と感情の対応についての実証的研究が行われている。例えば Kastner らは、長旋法と肯定的な感情、短旋法と否定的な感情の関係は3歳から12歳の幼児でも判別できるという実験結果を得ている [3]。教会旋法に関する研究例として、Hill らは、イオニア旋法とフリギア旋法の曲を被験者に聴かせ、救済 (salvation) と断罪 (condemnation) を表すのはどちらかを被験者に判定させる実験を行い、イオニア旋法が救済、フリギア旋法が断罪に対応するという結果を得た [4]。Ramos らは7種類の教会旋法を被験者実験から valence (快/不快) と arousal (覚醒) の2次元平面にマッピングし、旋法による感情の違いを示した [5]。

旋法と感情の対応づけができれば、作曲や自動作曲における感情表現の手段として旋法の効果的な使用につながり、映画音楽や音楽療法などへの応用が考えられる。

しかし旋法と感情の対応づけを調べるには大きな障害がある。それは、旋法の種類が数学的に膨大にあるため、現実

的には極めて少数の旋法についてしか調べられないことである。また、インドの旋法であるラーガにおいては、多数の旋法を使い分けることで、様々な感情を表現できるとされるが、ラーガは単なるピッチの集合としての音階とは異なり、ピッチ間の旋律的な傾向をも含んだ概念である [6]。したがって本稿で扱う旋法 (mode) の概念には、ピッチだけではなくピッチの旋律的な遷移傾向をも含めるものとし、音階 (scale) の概念は、単なるピッチあるいはピッチクラスの集合のことを指す語として区別する。12音の調律システムにおけるピッチクラス集合 (音階) の作り方は単純計算で $2^{12} = 4096$ 通りもある、さらにピッチの遷移傾向を含めた旋法を扱おうとすれば、バリエーションはほとんど無限にあり、一つ一つ調べつくすのは不可能であろう。

そこで本稿では、固定したいくつかの旋法を評価する心理実験ではなく、強化学習 [7] を用いた適応的なアプローチをとることでこの困難を回避する手法を提案する。それは、目標となる特定の感情に対して一人の人間の評価者からの感情の評価のフィードバックを用いて、旋法の方を適応変化させるというものである。これにより、あらゆる種類の旋法の中から目的の旋法を探索できる可能性がある。

強化学習を音楽に応用する先行研究としては、人間の望む音楽的緊張感の高低に対して、音符の強度や音価のパラメータを適応変化させる研究 [8] や、楽曲の楽譜から音楽スタイルを学習して音楽を生成する研究 [9] 等がある。強化学習を用いて旋法を感情に適応させる研究は、知る限りわれわれの研究が最初である。

このような提案を行うにあたり、次章では提案手法の基礎となる強化学習を解説し、3章では提案手法を述べる。4章では提案手法の挙動を把握するための予備的な実験について報告する。5章ではまとめと今後の課題を述べる。

¹ 東京藝術大学大学院美術研究科 先端芸術表現専攻

² 理化学研究所 脳科学総合研究センター

a) t-tsubasa@y4.dion.ne.jp

b) ohmura@brain.riken.jp

c) kf@zkm.de

*1 ここでの旋法名は教会旋法と同一名称の異なる旋法である。

2. 強化学習

2.1 基本的な枠組み

強化学習は教師無し機械学習法の一つであり、ある環境内に置かれたエージェントに、試行錯誤に基づく経験から適切な行動を学習させる方法論である。強化学習の枠組みでは、環境は状態空間 S 内をタイムステップ t ごとに遷移を繰り返す。エージェントは、状態 $s \in S$ を観測し、遷移ごとにある行動 $a \in A(s)$ ($a \in A(s)$ は状態 s でとることのできる行動の集合) をとって環境に働きかけ、行動の結果の善し悪しに応じて環境から報酬 r を受け取る。エージェントは次式などの形で表される r の関数としての収益 R_t を最大化できるように学習を行う：

$$R_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}, \quad (1)$$

ここで、 γ は $0 \leq \gamma \leq 1$ の割引率と呼ばれるパラメータである。 γ が 0 のときはエージェントは即時報酬 r_{t+1} のみを考慮して学習するが、 γ が 1 に近づくにつれ、は未来の報酬をより多く考慮に入れるようになる。エージェントは r をもとに、状態 s の価値 $V(s)$ や、 s における行動 a の価値 $Q(s, a)$ を推定してゆく。 $V(s)$ および $Q(s, a)$ は次式で定義される：

$$V(s) = E\{R_t | s_t = s\}, \quad (2)$$

$$Q(s, a) = E\{R_t | s_t = s, a_t = a\}. \quad (3)$$

強化学習において、多くの場合、状態は有限かつ離散的であり、環境は基本的に MDP (Markov Decision Process) と呼ばれる形式をとることが仮定される。MDP においては、状態遷移は現在の情報にのみ依存して遷移先が決まる一次のマルコフ過程であり、エージェントが行動 a をとったとき、環境は状態が s であれば状態 s' に確率 $P_{ss'}^a$ で遷移する。また、エージェントの行動 a に対して環境は、期待値 $R_{ss'}^a$ の報酬 r を与える。このように MDP は $P_{ss'}^a$ と $R_{ss'}^a$ で特徴づけられる。

他方、エージェントは方策 (policy) と呼ばれる行動指針 π をもち、状態 s において、ある確率 $\pi(s, a)$ で行動 a をとる。 $\pi(s, a)$ の値は $V(s)$ や $Q(s, a)$ を参照しながら決定することが多く、 $V(s)$ や $Q(s, a)$ は $\pi(s, a)$ と共に学習・改善されてゆき、エージェントは価値の高い状態や行動を選択できるようになると期待される。

もし環境の $P_{ss'}^a$ と $R_{ss'}^a$ が既知の定数でかつ $\gamma < 1$ あれば、 $V(s)$ と $Q(s, a)$ は Bellman 方程式と呼ばれる連立方程式を用いて解くことができる。しかし多くの場合、未知環境の下での学習が問題となる。未知環境の下では、 $\pi(s, a)$ の下での試行錯誤から経験的に学習を行う。

2.2 モンテカルロ法

一般に価値関数の改善は、次式のような $V(s)$ や $Q(s, a)$ を目標値に近づける更新によって行われる：

$$V(s_t) \leftarrow V(s_t) + \alpha[R_t - V(s_t)], \quad (4)$$

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[R_t - Q(s_t, a_t)], \quad (5)$$

ここで α は各ステップの学習の度合いを表す学習率と呼ばれるパラメータである。式 (1) で表される収益は、実際に計算できるようになるまでの時間が無限大であるため厳密には計算ができない。そこで、多くの場合ブートストラップと呼ばれる逐次的な近似計算による学習が行われる。

しかし、状態遷移の数が有限で終了するエピソードの回復によって学習する場合、エピソードの終了を待たば R_t の値が確定する。この R_t を直接用いて、エピソードに出現した各状態の状態価値 $V(s)$ と行動価値 $Q(s, a)$ を式 (4) および (5) で一括更新することができる。この手法はモンテカルロ法と呼ばれる。本研究ではエピソードを学習の単位とし、モンテカルロ法を用いる。

2.3 関数近似、最急降下法

状態と行動の組み合わせ (s, a) 全てに対して $Q(s, a)$ の訓練を十分な回数行うには、多くの時間やデータが必要となる。そこで、経験した (s, a) から経験していない (s, a) へと経験を汎化することが重要となる。つまり、 (s, a) の組み合わせの数よりも小さな要素数 n のパラメータベクトル $\vec{\theta} = (\theta_0, \theta_1, \dots, \theta_{n-1})^T$ の関数として $Q(s, a)$ を近似できれば効率的に学習を行うことができる。 $Q(s, a)$ が $\vec{\theta}$ を用いて近似される場合、最急降下法 (Gradient Descent) を用いて $\vec{\theta}$ を次式で更新することができる：

$$\vec{\theta} \leftarrow \vec{\theta} + \alpha[R_t - Q(s, a)] \nabla_{\vec{\theta}} Q(s, a) \quad (6)$$

関数近似の最もシンプル形態の一つは線形近似である。これはパラメータベクトル $\vec{\theta}$ に対応する特徴ベクトル $\vec{\phi}(s, a) = (\phi_0(s, a), \phi_1(s, a), \dots, \phi_{n-1}(s, a))^T$ を用いて

$$Q(s, a) = \vec{\theta}^T \cdot \vec{\phi} = \sum_{i=0}^{n-1} \theta_i \cdot \phi_i(s, a) \quad (7)$$

のように $Q(s, a)$ を線形に近似することである。この場合、式 (6) は次のようなシンプルな式に書き直せる：

$$\vec{\theta} \leftarrow \vec{\theta} + \alpha[R_t - Q(s, a)] \cdot \vec{\phi}(s, a). \quad (8)$$

3. 提案手法

3.1 目的

本研究の目的はある感情 E に対して「 E を表現する旋律」を得ることである。そこで、まずこの概念を定義しておく。ただし調律は 12 平均律とする*2。

*2 本稿の定式化は一般の 1 オクターブ n 分割 ($n > 0$) の調律システムに容易に拡張できる。ただし音程の扱いの上では、音程のシフトに対して音程幅が厳密に一定である n 平均律が望ましい。

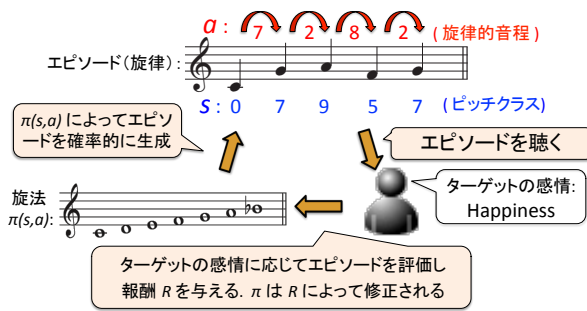


図 1 提案手法の概略図

Fig. 1 Outline of the proposed method.

定義 1 12 平均率の 12 個のピッチクラスの集合 $\mathbf{Z}_{12} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ を状態空間 S として設定する。ここで S 上の 1 次のマルコフ連鎖 M および M から生成される時系列としての旋律 $[s_1, s_2, \dots, s_T]$ を考える。評価者が生成された旋律を聴取して、平均的に E を表していると評価されるマルコフ連鎖 M_E が得られたとする。そのときこの M_E を「 E を表現する旋法」と呼ぶ。

この定義によって、旋法を、出現する可能性のあるピッチクラス集合としての音階と区別し、ピッチクラス間の遷移傾向を含めたものとして扱う。

3.2 学習手法

E を表現する旋法 M_E を獲得するため、モンテカルロ法による強化学習を応用した提案手法を以下に説明する。図 1 は提案手法の概念図である。本手法では、エージェントが旋法に対応する方策 $\pi(s, a)$ (定義は次節で行う) に従って確率的に行動 a の選択を繰り返し、エピソードとしての決まった長さ T の旋律 $[s_1, s_2, \dots, s_T]$ を生成する。ただし、行動 $a \in A(s)$ は前のピッチクラス s から次のピッチクラス s' への旋律的音程の移動を表す ($\forall s \in S, A(s) = \mathbf{Z}_{12}$)。 s' は s と a によって $s' = (s + a) \bmod 12$ で一意に定まり、 $P_{ss'}^a = 1$ である。エピソードを聴き終わった時点 ($t = T$) で評価者の人間が評価値 (報酬 r_T) を与える。ここで、この報酬は旋律全体に対しての評価であるため、 γ を 1 として割引を行わない。そして r_T をもとに、エピソード内に出現した s と a に対応する価値関数を式 (4) および (5) で更新する (モンテカルロ法)。それによって方策 $\pi(s, a)$ も更新される。このようなサイクルを繰り返し、徐々に方策 $\pi(s, a)$ を改善していき、 $\pi(s, a)$ が目的の感情を表現する旋法となるように学習を行う。ここで評価者の人間は強化学習の枠組みにおける環境に相当する。

r_T は、個々のエピソードの旋律からどの程度感情 E を感じるかについての一人の評価者による評価値であり、 $\{-3, -2, -1, 0, 1, 2, 3\}$ の 7 段階の整数とする。各値の意味は、3: 強く感じる, 2: 感じる, 1: 少し感じる, 0: 全く感

じない, -1: 他の感情を少し感じる, -2: 他の感情を感じる, -3: 他の感情を強く感じる, とする。このような報酬のもとで、価値関数 $V(s)$ は、ターゲットの感情 E を表現するためのピッチクラス s 値を、行動価値関数 $Q(s, a)$ はあるピッチクラス s から s' へ音程を a だけ移動することの価値を表すものと考えられる。

3.3 行動価値関数を用いた方策の定義

エピソード生成に用いる方策 π は次のように定義する:

$$\pi(s, a) = \begin{cases} \frac{Q(s, a)}{\sum_{u \in \mathbf{Z}_{12} \text{ s.t. } 0 < Q(s, u)} Q(s, u)} & \text{if } 0 < Q(s, a) \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

このように方策 π を定義することで、 s においてプラスの行動価値関数 $Q(s, a)$ をもつ行動 a は、 $Q(s, a)$ に比例する確率で出現し、マイナスの行動価値をもつ a は全く出現しなくなる。それにより、目標となる感情を表現するのに必要のない状態と行動を排除することができると考えられる。ここで方策 π の意味を考えると $\pi(s, a)$ は状態 s から s' への遷移確率 $Pr(s'|s)$ に等しい。すなわち、方策 π はマルコフ連鎖となる。 $Q(s, a)$ の初期値は報酬の最大値である 3.0 を目安に実際よりも高く設定し (optimistic initial value), 学習の初期において状態行動空間の広い探索を促す。学習が進み、評価者によって平均的に高評価を得られるようになったとき、 π によって得られるマルコフ連鎖は、本章冒頭で定義した「 E を表現する旋法」となる。

3.4 行動価値関数の 4 つの関数近似モデル

$Q(s, a)$ は、状態と行動の空間の直積 $\mathbf{Z}_{12} \times \mathbf{Z}_{12}$ を定義域とする関数であり、 $12 \times 12 = 144$ 個の値をもつ。これら全てを学習するのは非効率で評価者への負担が大きいため、本章では 2.3 節で述べた関数近似を用い、4 通りの近似モデルを構築する。

近似においては、ピッチクラス s とは無関係に旋律的音程 a の価値が定まると仮定し、4 つのモデルに共通するパラメータ $\vec{\theta}$ を設定する。 $\vec{\theta}$ の要素数は 24 とし、 θ_1 から θ_{12} をピッチクラス 0 から 11 の状態の価値に対応するパラメータ、 θ_{13} から θ_{24} は、ピッチクラスの遷移 0 から 11 の行動の価値に対応するパラメータとして用いる。

3.4.1 モデル 1 (ピッチモデル)

最初のモデルは、 a による旋律的音程の遷移の価値を無視し、遷移先のピッチクラス $s' = (s + a) \bmod 12$ の価値のみから $Q(s, a)$ を近似するモデルである。このモデルは式 (7) の線形近似において特徴ベクトルを次のようにとったものに相当する:

$$\phi_i(s, a) = \phi_i(s') = \begin{cases} 1 & \text{if } i = s' \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

このモデルでの行動価値関数 $Q_1(s, a)$ は

$$Q_1(s, a) = \sum_{i=0}^{23} \theta_i \cdot \phi_i(s, a) = \theta_{s'} \quad (11)$$

となる。このとき式 (9) の ϕ によるマルコフ連鎖においては $Pr(s'|s) = Pr(s')$ となり、 M は現在の状態 s に依存せず、 s' に対する単純な確率分布に退化する。

3.4.2 モデル 2 (音程モデル)

2つ目のモデルは、状態 s に依存せず、 $Q(s, a)$ をピッチクラスの遷移 a の価値のみから近似するモデルである。このモデルは線形近似において特徴ベクトルを次のようにとったものに相当する:

$$\phi_i(s, a) = \phi_i(a) = \begin{cases} 1 & \text{if } i = 12 + a \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

このモデルでの行動価値関数 $Q_2(s, a)$ は

$$Q_2(s, a) = \sum_{i=0}^{23} \theta_i \cdot \phi_i(s, a) = \theta_{12+a} \quad (13)$$

となる。このとき式 9 の ϕ によるマルコフ連鎖においては $Pr(s'|s) = Pr(a)$ となり、 M は現在の状態 s に依存せず、 a に関する単純な確率分布に退化する。

3.4.3 モデル 3 (加算モデル)

3番目のモデルは、 $Q(s, a)$ を遷移先のピッチクラス s' およびピッチクラスの遷移 a の価値の両方を考慮し、その和で近似するものである。これは線形近似において特徴ベクトルを次のようにとったものに相当する:

$$\phi_i(s, a) = \begin{cases} 1 & \text{if } i = s' \text{ or } i = 12 + a \\ 0 & \text{otherwise.} \end{cases} \quad (14)$$

このモデルでの行動価値関数 $Q_3(s, a)$ は次のよう表せる:

$$Q_3(s, a) = \sum_{i=0}^{23} \theta_i \cdot \phi_i(s, a) = \theta_{s'} + \theta_{12+a}. \quad (15)$$

3.4.4 モデル 4 (シグモイド乗算モデル)

4番目のモデルは、遷移先のピッチクラス s' およびピッチクラスの遷移 a の価値の貢献度が相乗効果をもつように、それぞれに対応するパラメータをシグモイド関数 $\sigma(\theta) = 1/(1+e^{-\theta})$ で変換したものの積を用いた近似モデルである。シグモイド関数は $\theta \rightarrow -\infty$ で 0、 $\theta \rightarrow \infty$ で 1 に漸近するため、 $\sigma(\theta)$ の範囲をコントロールすることができる (シグモイド無しで直接 θ 同士を乗算すると負の値と負の値をかけると正の値になるなどの不都合が生じる)。これを用いて、このモデルでの行動価値関数 $Q_4(s, a)$ は

$$Q_4(s, a) = 6\sigma(\theta_{s'})\sigma(\theta_{12+a}) - 3 \quad (16)$$

と表される。ここで係数と切片は $Q_4(s, a)$ が区間 $[-3, 3]$ に収まるように設定した。

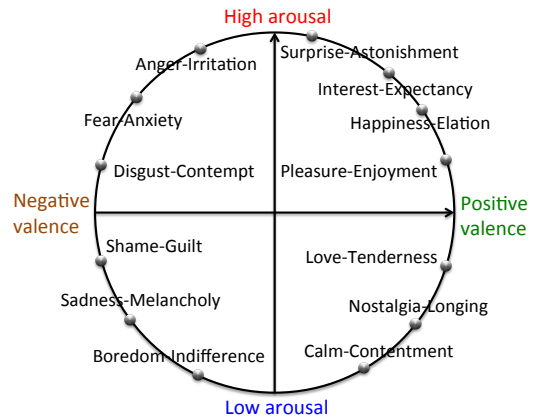


図 2 感情カテゴリーの環状モデル

Fig. 2 Circumplex model of emotion categories.

4. 予備実験

本章では、行動価値関数 $Q(s, a)$ の 4つの関数近似モデルの挙動を確認するために行った、著者の 1人による予備実験について報告する (これを元に、今後より大規模な実験を行う予定である)。

4.1 実験条件

学習対象の感情の選択においては、Juslin らの環状モデル ([10], 図 2) を参照し、各象限からの代表として喜び (happiness)、悲しみ (sadness)、恐れ (fear)、優しさ (tenderness) を選択した。このモデルは、valence (快/不快) と arousal (覚醒度) の張る 2次元平面上に、諸感情を環状に配置したものであり、各象限から代表を 1つずつ選択することで、できるだけ広い範囲から互いにかげ離れた種類の感情を選択できる利点がある。

エピソードの旋律に関しては、音数が長すぎると学習が進まない恐れがあるため、短く設定し 5音 ($T = 5$) とした。12個のピッチクラスは MIDI 番号 60 から 71 までの 12音とし、音色は MIDI のピアノ音とした。発音タイミングは 300ms ごとの等間隔とした。評価の精度を上げるため、評価に確信が持てるまで旋律を反復した。エピソードの生成・評価の回数は、評価者が評価値の変化をこれ以上感じられなくなったと判断した時点で終了とした。旋律はトランスポーズ (移調) したものと同一の旋律になるため、調の混同で学習が進まない恐れがある。そこで、参照点となる主音を MIDI 番号 60 の C として設定し、エピソードの最初の音を主音に固定した。学習率 α は 0.1 に設定した。

4.2 結果

4.2.1 評価値の推移

図 3 に各感情に対する 4つのモデルの学習過程における、直近 10 エピソードの平均報酬 (評価値の平均) を示す。この図から次のようなことがわかる。まず、いずれの

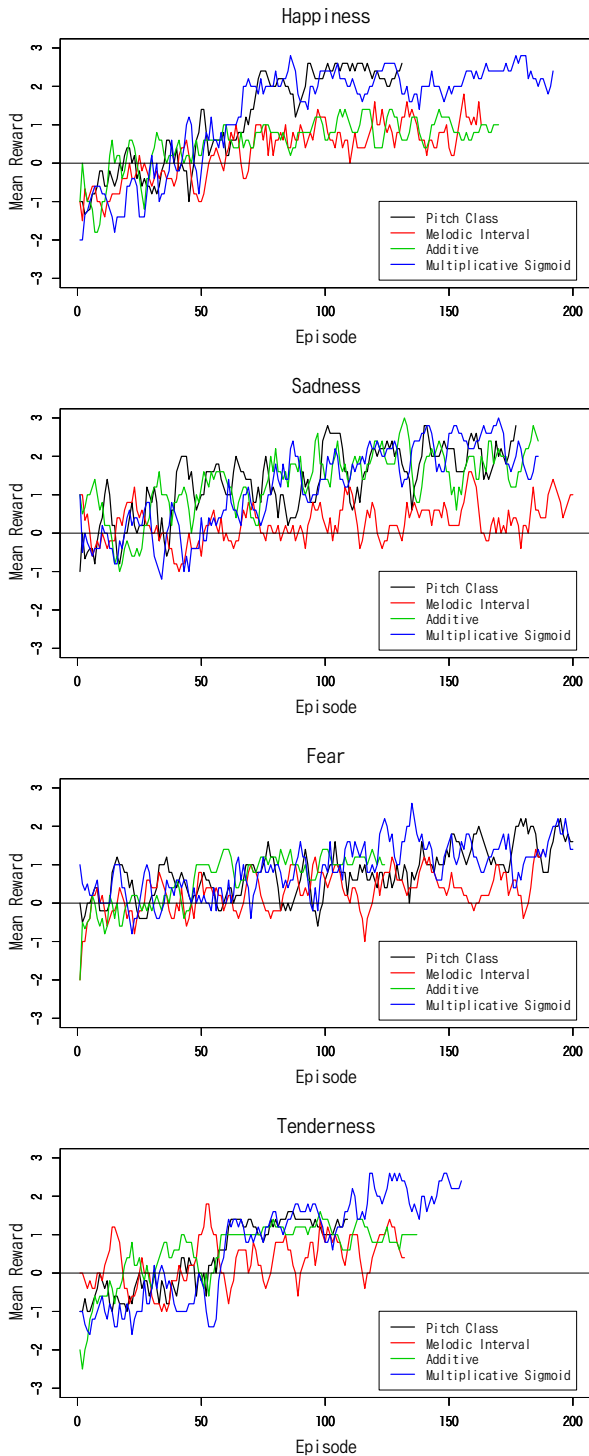


図3 平均報酬の推移
 Fig. 3 Transition of mean reward.

モデルにおいても評価値が向上していく傾向がみられ、期待値がおおむね0を超えるようになってきている。ただし、音程モデルにおいては学習後半の評価値が0から1の間程度と向上が小さい。モデル間の比較から、いずれの感情においても比較的良い性能を示しているのはピッチモデルとシグモイド乗算モデルである。この二つの性能は拮抗しており比較は難しいが、「優しさ」においてシグモイド乗算モデル

ルがピッチモデルを上回っており、4つのモデルの中での最良の結果だと解釈できる。シグモイド乗算モデルにおいては「恐れ」以外では2程度の評価値を、「恐れ」において1程度の評価値を平均的に獲得している。

これらの結果は次のことを示唆すると解釈できる。音程モデルの方がピッチモデルよりも評価値の向上が小さいことから、旋律的音程よりもピッチの方が、より感情の表現にとっては重要であると考えられる。ただし、音程モデルにおいても、ある程度の向上傾向がみられることから、旋律的音程も感情の表現に一定の寄与をしていると考えられる。そして、シグモイド乗算モデルの性能の良さからは、音程とピッチの両方を適切なモデルで学習することで、音程あるいはピッチ単独の場合よりも感情の表現力を高める相乗効果をもたらす可能性があると考えられる。

4.2.2 獲得した旋法

図4に学習後に得られた $\vec{\theta}$ を示す。この $\vec{\theta}$ によって $Q(s, a)$ および旋法としての $\pi(s, a)$ が定まる。この図から、各ピッチクラスの価値に関わる θ_0 から θ_{11} の方が、各音程の価値に関わる θ_{12} から θ_{23} よりも起伏がはっきりしていることがわかる。そして、0より大きい値をもつピッチと0あるいは0以下の値をもつピッチの両者の差の明確さからも、旋法における主要部としてのピッチクラス集合（音階）の重要性が読み取れる。そこで、音階としての学習結果をみるため、シグモイド乗算モデルにおける正の価値をもつ $(V(s) > 0)$ s の集合を音符で図5に示す。この図において、喜びの旋法はミクソリディアンと同一である。悲しみの旋法はロクリアンにDとGを付加したものであり、恐れは悲しみの旋法にさらにEおよびBを付加したものである。優しさの旋法はイオニアン（長旋法）からFを除いたものに一致する。図2を参照しながらそれらの構成音の音数を比較すると、valenceの高い方（右側）が音数が少なく、arosalが高い方（上側）が音数が多い傾向が読み取れる。ここから、構成音の音数と感情の間には法則性があるのではないかという仮説を立てられる。このように、生成した旋法を分析することで、音楽心理学的な知見を得られる可能性をもつ点も、提案手法の利点である。

また興味深いことに、日本の伝統的な旋法[6]において、明るいことを意味する名の陽旋法(C,D,E,G,A,Bb)および、暗いことを意味する名の陰旋法(C,Db,F,G,Ab,Bb)が（これらは上行形と下降形の両者を合わせたものである）、それぞれ喜びの旋法および悲しみの旋法の中に埋め込まれている。このことは、評価者の文化の固有性が適応的な手法によって反映されたか、あるいは逆に、陽旋法と陰旋法の表現する感情の普遍性を表しているのかもしれない。この点も、今後本格的に検証する価値のある問題であろう。

5. まとめ

感情を表現する旋法を獲得するため、人間の評価を用い

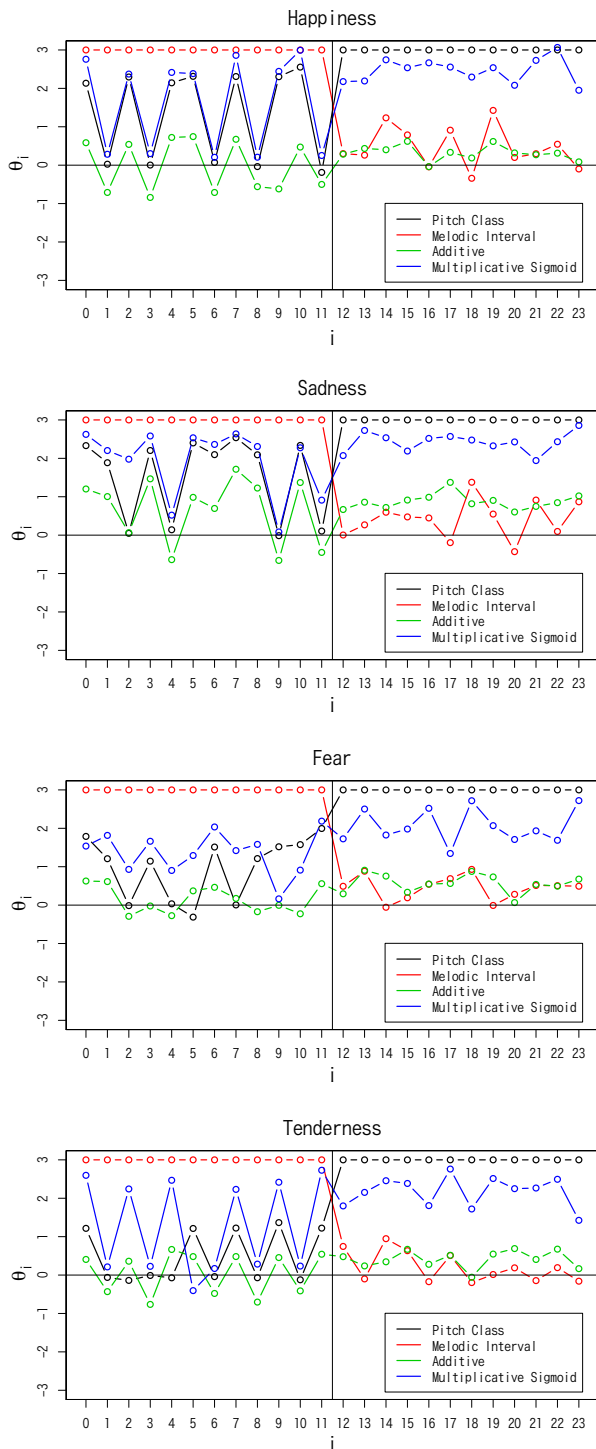


図4 学習後の $\hat{\theta}$
Fig. 4 $\hat{\theta}$ after learning.

た4つの適応的な強化学習モデルを構築し、それらの性能を比較を試みた。4種類の感情について学習を行ったところ、最良のモデルであったシグモイド乗算モデルにおいては、喜び、悲しみ、優しさの表現力をもつ旋法を、恐れに関しては若干の表現力をもつ旋法を獲得することができた。各モデル間の比較からは、旋律的音程よりもピッチの方が感情の表現にとってはより重要であるが、旋律的音程も感



図5 学習後に得られたピッチクラス集合 (シグモイド乗算モデルの学習結果において、 $V(s) > 0$ となる s の集合)

Fig. 5 Pitch class sets after learning (Sets of s such that $V(s) > 0$ after learning by sigmoid multiplicative model).

情の表現に一定の寄与をしていることが示唆された。そして、ピッチと音程の両方を適切に用いた強化学習モデルによって、音程あるいはピッチ単独の場合よりも感情の表現力を持ちうる可能性が示唆された。今後、提案手法から得られる旋法の収集・分析により、旋法の特徴と感情の間の法則性や、旋法によって表現される感情がどの程度普遍的なのか、個人や文化圏の固有性がどの程度反映されるのかなど、音楽心理学的な知見が得られる可能性がある、さらに、本手法を拡張し12音以外の調律システムでの旋法の生成することも興味深い課題である。

謝辞 本研究は日本学術振興会特別研究員奨励費(12J11238)の助成を受けたものである。

参考文献

- [1] プラトン: 国家 (藤沢令夫 訳), pp.209-213, 岩波書店 (2005).
- [2] アリストテレス: 政治学 (山本光雄 訳), pp.368-381, 岩波書店 (1997).
- [3] M. P. Kastner and R. G. Crowder: Perception of the major/minor distinction: IV. Emotional connotations in young children. *Music Perception*, 8, 189-202 (1990).
- [4] D. S. Hill, S. B. Kamenetsky, and S. E. Trehub: Relations among text, mode, and medium: Historical and empirical perspective, *Music Perception*, 14, 3-21 (1996).
- [5] D. Ramos, J. L. O. Bueno and E. Bigand: Manipulating Greek musical modes and tempo affects perceived musical emotion in musicians and nonmusicians, *BRAZILIAN JOURNAL OF MEDICAL AND BIOLOGICAL RESEARCH*, Vol.44(2), pp.165-172 (2011).
- [6] S. Sadie (Eds.): *The New GROVE Dictionary of Music and Musicians 12*, pp.446, Macmillan Publishers Limited, London (1980).
- [7] R. S. Sutton and A. G. Barto: *Reinforcement Learning: An Introduction*, The MIT Press (1998).
- [8] S. Le Groux, and Paul F. M. J. Verschure: Towards Adaptive Music Generation by Reinforcement Learning of Musical Tension, *Proc. SMC*, pp.160-165 (2010).
- [9] M. V. Butz, O. Sigaud, G. Pezzulo and G. Baldassarre (Eds.): *Anticipatory Behavior in Adaptive Learning Systems*: A. Cont, S. Dubnov, and G. Assayag: Anticipatory Model of Musical Style Imitation using Collaborative and Competitive Reinforcement Learning, pp.285-306, Springer (2007).
- [10] P. N. Juslin and J. A. Sloboda (Eds.): *Handbook of Music and Emotion - Theory, Research, Applications*, Oxford University Press (2010).