2R-7

対話型遺伝的アルゴリズムを用いたユーザの嗜好と曲の構成を考慮した自動作曲

高野美央　長名優子

東京工科大学 大学院 バイオ・情報メディア研究科

1　はじめに

人間ではなく，コンピュータに作曲を行わせようとする研究の歴史は古く，1957 年にマルコフ過程を用いた自動作曲の研究が行われて以来，数多くの研究が行われてきた．そのような手法の 1 つとして遺伝的アルゴリズムと N グラムモデルを用いた曲の構成を考慮した自動作曲システム [1] が提案されている．このシステムでは，ユーザが学習に用いる曲を指定することで，ユーザの嗜好をある程度は反映することができる．しかし，学習に用いる曲と類似した特徴を持っているかどうかを評価の基準として用いることしかできないため，ユーザの細かい好みを反映した曲の生成を行うことは難しい．

本研究では，文献 [1] のシステムを拡張し，対話型遺伝的アルゴリズム [2] を用いたユーザの嗜好と曲の構成を考慮した自動作曲システムを提案する．提案システムでは，生成されたメロディに対し，ユーザがフレーズごとに評価を行い，その評価の傾向を適応度の計算に反映することで，ユーザの嗜好を反映したメロディの生成を実現する．

2　遺伝的アルゴリズムを用いた複数の曲の特徴を考慮した自動作曲

ここでは，提案システムにおいて初期個体群の生成に用いる遺伝的アルゴリズムを用いた複数の曲の特徴を考慮した自動作曲システムについて説明する．このシステムでは，リズムと音の高さの情報を遺伝子の形で表現し，選択，交叉，突然変異などの遺伝的操作を繰り返すことで曲を生成する．

2.1　遺伝子による曲の表現

遺伝子はリズムと音の高さの情報で表現する．リズムは音の長さの最小単位 (16 分音符の長さの 1/3) ごとに 0 (休符)，1 (音の始まり)，2 (音が続いている状態) のいずれの状態であるかによって表す．音の高さはリズムと同様に音の長さの最小単位ごとに音名とオクターブの情報を割り当てることで表す．

2.2　初期個体群の生成

このシステムでは，学習に用いたすべての曲の特徴を反映した形で初期個体群の生成を行う．初期個体の生成では，まずはじめに曲の構成ごとの始まりの状態の分布と曲の構成ごとの音符と休符の始まる位置に関するマルコフモデルを用いてリズム系列の生成を行う．各構成の始めを表す状態に続く状態の系列は音休符の始まる位置に関するマルコフモデルにおける状態遷移確率に従ったランダムウォークによって生成する．生成されたリズムの系列に対して，音の高さの分布確率にしたがって音の高さの情報を付与する．このようにして複数のメロディを生成し, それを初期個体とする．

2.3　適応度の計算

このシステムでは，(1) リズムの進行，(2) 音の高さの進行，(3) 音の高さと長さの進行，(4) 1 小節に含まれる音の数の遷移，(5) 4 小節ごとのリズムの類似度，(6) 音階固有の音の割合，(7) 連続する 2 つの音の高さの差分の分布，(8) 休符の長さの分布に関する適応度を用いる．

(1)∼(5) については学習に用いる曲の特徴を学習した N グラムモデルを用いて計算する．(6)∼(8) については，学習に用いる曲における分布確率を利用して計算する．学習に用いた曲と似た特徴を持つ曲を表す遺伝子の適応度が高くなる．

なお，(1)∼(5) の適応度の計算に用いる N グラムモデルは，メロディやサビなどの構成ごとに学習を行う．また，複数の曲を学習に用いる場合には，それぞれの曲の構成ごとに N グラムモデルを作成することになる．複数の曲を学習に用いる場合には，それらの曲のいずれかの特徴と類似していればよいと考え，出現確率の最も高い N グラムモデルの値を適応度計算に用いるものとする．しかし，この方法で適応度計算に用いる N グラムモデルの選択を行うと連続した状態の系列を評価する際に別の曲の特徴を学習した N グラムモデルが選択されてしまい，メロディとして自然でないものの適応度が高くなってしまう可能性があ

Automatic Composition System using Interactive Genetic Algorithm considering User's Evaluation and Melody Blocks

Mio Takano and Yuko Osana (Tokyo University of Technology, osana@stf.teu.ac.jp)

Copyright ©2014 Information Processing Society of Japan.
All Rights Reserved.

る．そこで，出現確率の最も高い N グラムモデルを選択した後で、前後の状態に対して選択された N グラムモデルの情報も含めて再度決定を行う．ここでは，(1) 前後の状態に対して別の曲を学習した N グラムモデルが選択されており，(2) 2 つ前の状態と 2 つ後の状態の両方もしくは一方に対して，前後の状態に対して選択されたものと同じ N グラムモデルが選択されているとき、着目している状態に対して選択されている N グラムモデルが前後の状態に対して選択されている N グラムモデルに変更する．このような修正を行うことで，連続した状態の系列に対してはある程度同じ曲の特徴を学習した N グラムモデルを用いて適応度の計算を行うことが可能となる．

3 対話型遺伝的アルゴリズムを用いたユーザの嗜好を考慮した自動作曲

提案する対話型遺伝的アルゴリズムを用いたユーザの嗜好と曲の特徴を考慮した自動作曲システムでは，**2** で述べた遺伝的アルゴリズムを用いた複数の曲の特徴を考慮した自動作曲を用いて初期個体群の生成を行い，さらに対話型遺伝的アルゴリズムを用いてユーザが評価を行うことでユーザの嗜好を反映した曲の生成を実現する．

3.1 遺伝子による曲の表現

提案システムでは，**2.1** で述べたのと同様の方法でリズムと音の高さの情報により生成する曲を遺伝子として表現する．

3.2 初期個体群の生成

提案システムでは，遺伝的アルゴリズムを用いた複数の曲の特徴を考慮した自動作曲システムを用いて，ユーザがサンプルとして入力した曲に対してメロディを生成し，それを初期個体とする．提案システムではユーザが生成された曲を評価しながら曲の生成を行うため，評価回数が多くなるとユーザの負担が大きくなってしまう．そのため，ユーザが最初に評価を行う段階でもランダムに生成されたメロディではなく，ユーザが指定した曲の特徴を反映して生成したメロディを提示するようにすることでユーザによる評価の回数を少なくできるように工夫している．

3.3 ユーザによるメロディの評価

提案システムでは，システムが生成したメロディの一部を一定世代ごとにユーザに提示して評価してもらい，適応度の計算に利用する．なお，ユーザによる評価は曲全体ではなく，メロディをフレーズに区切り，フレーズ単位で行う．フレーズ単位で評価を行うのは，曲全体に対して評価を行った場合にはユーザがその曲のどの部分を良いまたは悪いと思ったのかを分析することが難しく，ユーザも曲の一部だけを良いと思ったような場合に評価がつけにくいと考えられるためである．ユーザは，提示されたフレーズに対して −2 (嫌い)，−1 (少し嫌い)，0 (どちらでもない)，1 (少し好き)，2 (好き) の 5 段階で評価を行う．

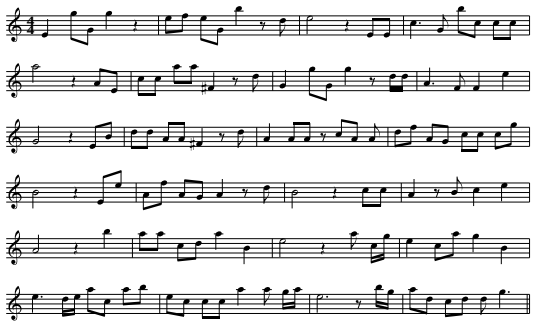

図 1: 生成した曲の例

3.4 適応度の計算

提案システムでは，遺伝的アルゴリズムを用いた複数の曲の特徴を考慮した自動作曲システムと同様に **2.3** で述べた 8 つの項目について適応度を計算する．なお，適応度の計算には初期個体群の生成に用いたシステムで使用した N グラムモデルを用いる．さらに，(1) 2 小節単位のリズム，(2) 最後の 2 小節のリズム，(3) 音の長さ，(4) 最後の 2 小節の音の長さ，(5) 音の高さと長さ，(6) 最後の 2 小節の音の高さと長さの遷移についてユーザによる評価の傾向を分析し，特定の遷移に対し，ユーザの評価がプラスまたはマイナスに偏っている場合にはそれを適応度に反映させる．このようにすることで，どのようなフレーズが含まれる場合にユーザの評価が高く，もしくは低くなるのかを分析し，生成される曲にその傾向を反映することが可能となる．

4 計算機実験

提案システムを用いて,「君をのせて」と「海の見える街」の 2 つの曲を学習させ，曲の生成を行った結果の一部を図 1 に示す．

参考文献

[1] M. Takano and Y. Osana : “Automatic composition system using genetic algorithm and N-gram model considering melody blocks,” Proceedings of IEEE Congress on Evolutionary Computation, Brisbane, 2012.

[2] R. Dawkins : The Blind Watchmaker, Longman, Essex, 1986.

[3] 北研二：確率論的言語モデル, 東京大学出版会, 1999.

Copyright ©2014 Information Processing Society of Japan.
All Rights Reserved.