

LSTM を用いたユーザの嗜好を反映する BGM 自動作曲システム

岡部太亮^{†1} 棟朝雅晴^{†2}

概要：本研究では、ニューラルネットワークを用いた作曲エージェントを対話型進化計算により進化させることで、ユーザの好みを反映した音楽を連続して自動生成する。ニューラルネットワークの学習データとして、楽曲データを事前にクラスタリングすることで、ユーザの好む雰囲気楽曲がどのクラスターに属するかをモデル化し、適切な学習データ集合を各クラスターから選択するため、対話型遺伝的アルゴリズムを用いることでユーザの好みを反映させた自動作曲エージェントを生成する。これにより、単に音楽を生成するのではなく、連続して好みの音楽を自動的に生成するシステムを実現する。

キーワード：対話型遺伝的アルゴリズム, LSTM, クラスタリング

1. はじめに

自動作曲を用いた BGM 提供システムは数多くあり [1], [2], [3], それらはユーザに提供する BGM を作成したり、作曲者が作曲のヒントとするために利用されている。それに伴い、機械学習を用いた自動作曲に関する研究は近年盛んである。しかし、ユーザにとって好ましい音楽を連続して自動生成することに関してはいまだ挑戦的な課題である。

本研究においては、ユーザが好ましく思う音楽を連続して生成する自動作曲エージェントを実現するために、楽曲の構造を学習する方法として深層学習による LSTM (Long short-term memory) ニューラルネットワークを採用する。事前に学習データとなる楽曲データベースをクラスタリングし、各クラスターから学習データとして採用する割合を遺伝子表現として対話型遺伝的アルゴリズム [10] を実行することで、学習データ内のユーザが好む領域を探索する。

2. 関連研究

2.1 遺伝的アルゴリズムを用いた研究

Ping-huan Kuo らの研究 [10] では、Adaptive Partition Evolutionary Genetic Algorithm という手法を用いて、楽曲の作成とクオリティの向上を行なっている。楽曲内の音の情報を遺伝子表現し、個体の評価値が低い箇所を他個体の評価値が高い部分と交換することで旋律を改善していく手法である。また、ユーザの好みに応じて楽曲の雰囲気を調整するために、ラッセルの円環モデル [7] を利用している。このモデルは人間の感情を二次元平面にマップしたものでユーザはこれを使って楽曲の雰囲気を自分の感情に近いものに調整することができる。

2.2 深層学習を用いた研究

深層学習を用いた研究では、楽曲の情報を MIDI や文字列に置き換えてニューラルネットワークに学習させる手法

がある。ABC 記譜法で記述された楽曲を、学習用に正規化して学習させる。ABC 記譜法はコンピュータで楽譜をテキストとして表す時に使われる表現方法である。この記譜法ではタイトルや表紙、基本となる音価なども含めて記される。

2.3 楽曲のクラスタリングに関する研究

Stefano らの研究 [11] では、対象を複数のジャンルから集めたギタリストのソロに絞っている。楽曲の音の遷移からネットワークを作成し、作成されたネットワークのノード数やエッジ数をパラメータとしてクラスタリングを行なっている。この手法を用いることで、同ジャンルのギタリストや同年代のアーティストを同じクラスターに分類することができている。

3. BGM 作曲システム

3.1 システムの実行手順

提案するシステムの実行手順の概略を以下の図 1 に示す。

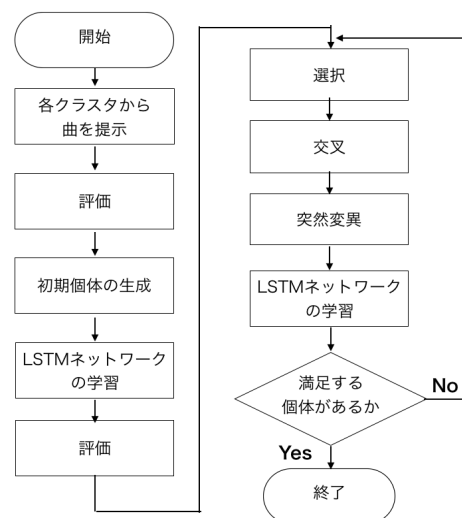


図 1. 提案システムの実行手順

^{†1} 北海道大学大学院情報科学研究科
Graduate School of Information Science and Technology, Hokkaido University
^{†2} 北海道大学情報基盤センター
Information Initiative Center, Hokkaido University

楽曲データのクラスタリングを事前に行い、その結果をもとに初期個体および初期のLSTMネットワークを生成する。その後、対話型遺伝的アルゴリズムをwebアプリケーションとしてユーザに提供して実行し、その好みを反映した遺伝子集団の探索およびLSTMネットワークの学習を、ユーザが満足する個体エージェントが生成されるまで実行する。

3.2 クラスタリングについて

提案手法では k-means+法を用いて、楽曲のクラスタリングを行う。楽曲を音の遷移として表現したネットワークの特徴量からノード数、エッジ数、ネットワークの直径などをパラメータとしてクラスタリングした。クラスタ数はエルボー法を使って調整した結果8とした。

3.3 深層学習について

深層学習には Keras ライブラリを用い、ネットワークは中間層に LSTM ブロックを 128 個配置した 3 層構造とした。出力層の活性化関数にはソフトマックス関数を採用している。ネットワークは、入力として楽曲のシーケンス、出力としてそれに対する次の一音の確率分布を生成する。

また、学習データとしては abcnotation.com[4]から北欧の folk music を用いた。

3.4 遺伝的アルゴリズムについて

データセットの各クラスタから学習データとして採用する割合を遺伝子表現とする。クラスタ数と遺伝子長は等しいので、遺伝子長は 8 である。選択手法はトーナメント法を、交叉手法は BLX- α を用いている。

4. 実験

4.1 学習精度について

用意した学習データに対して 2 割を validation set に設定して、学習精度をプロットした。validation set の学習が test set を上回っていないので、過学習は起きていないと言える。また、学習率は 0.55 程度で収束している。

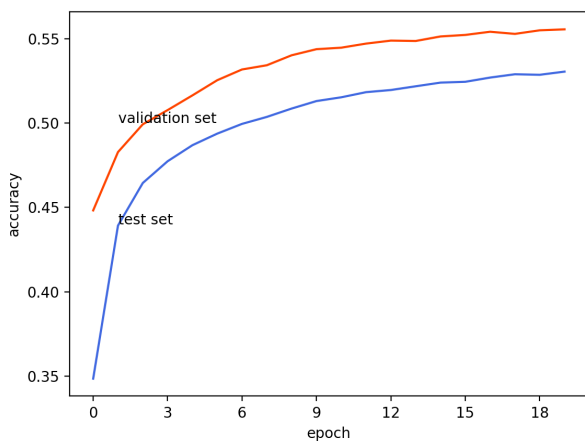


図 2. LSTM の学習精度



図 3. 生成された楽曲例

4.2 生成された楽曲について

生成された楽曲の一例を図 3 に示す。色で囲まれた部分は同じリズムパターンを再現しており、学習データとして採用した folk music に見られるリズムパターンが生成されている。

5. まとめと今後の課題

本研究では、ニューラルネットワーク LSTM を用いた作曲エージェントを、対話型遺伝的アルゴリズムで進化させることで BGM の自動作曲を実現した。学習精度の向上のために、ネットワークのチューニングを行い、結果として学習データの持つ特徴を有する新たな音楽パターンを生成できている。今後の課題としては、ユーザの代わりとして自律的に代理評価を行えるニューラルネットワークを事前に生成することで、ユーザの負荷を減らすことを検討する。

参考文献

- [1] Amper Music, <https://ampermusic.com/>
- [2] SolarBeat, <http://www.whitevinylsdesign.com/solarbeat/>
- [3] BlokDust, <https://blokdust.com/>
- [4] abcnotation.com, <http://abcnotation.com/>
- [5] Glorot, Xavier, and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks." Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. 2010.
- [6] Hochreiter, Sepp, and Jürgen Schmidhuber. "Long short-term memory." *Neural computation* 9.8 (1997): 1735-1780.
- [7] James A. Russell, "A Circumplex Model of Affect", 1980, Journal of Personality and Social Psychology
- [8] KUO, Ping-Huan, et al. Development of an Automatic Emotional Music Accompaniment System by Fuzzy Logic and Adaptive Partition Evolutionary Genetic Algorithm, IEEE Access, 2015, 3:815-824
- [9] Takagi, Hideyuki. "Interactive evolutionary computation: Fusion of the capabilities of EC optimization and human evaluation." Proceedings of the IEEE 89.9 (2001): 1275-1296.
- [10] T. Tieleman and G. Hinton, "Divide the gradient by a running average of its recent magnitude," lecture 6.5 of Coursera "Neural Networks for Machine Learning," 2012.
- [11] Stefano Ferretti, Clustering of Musical Pieces through Complex Networks: an Assessment over Guitar Solos, arXiv:1709.05193v1 [cs.LG] 14 Sep 2017