

機械学習を用いたテレビドラマ音楽の自動生成

遠野 大輔[†] 横山 真男[†] 植木 一也[†][†]明星大学 情報学部

1 はじめに

近年、人工知能は急成長を遂げ、それを支える技術のひとつとしてディープラーニングが注目されている。この技術は医療や自動車など様々な分野で活用されているが、音楽といった芸術分野に適用し、人間が作ったものに近い作曲ができるのではないかと思われる。

本研究では、Magenta[1]というGoogleのライブラリを使いドラマのカテゴリ別また行事や季節などのテーマで分けられている楽曲の学習による自動作曲を行った。生成された楽曲のテレビ音楽らしさについて、「学習回数の違い」「サビのメロディで学習した場合」「メロのメロディで学習した場合」の3点について評価を行った。

2 使用楽曲

使用楽曲はYAMAHAのMIDIデータショップ[2]の特集から選出した。表1に使用した楽曲から抜粋した一部を示す。

表1 使用楽曲

ジャンル	タイトル
ミステリードラマ	ORION
	VOICE
	おとなの掟
	影
	告白
夏のテーマ	RIDE ON TIME
	あー夏休み
	イケナイ太陽
	夏色
	真夏の果実
卒業のテーマ	3月9日
	みんな空の下
	栄光の架橋
	思い出がいっぱい
	贈る言葉
朝ドラマ	カーネーション
	にじいろ
	花束を君に
	春よ、来い
	麦の唄
大河ドラマ	花燃ゆ メインテーマ
	軍師官兵衛 メイン・テーマ
	八重の桜 メインテーマ
	風林火山 -メイン・テーマ
	龍馬伝

Automatic composition of TV drama music using machine learning

[†]Daisuke TONO, Meisei University, [†]Masao YOKOYAMA, [†]Kazuya UEKI

3 Magenta を用いた強化学習

スタンダード MIDI ファイルをMagentaではそのままのファイル形式で使うことができないので、Magentaが扱うNoteSequenceというファイル形式に変換する。変換したファイルを基に学習用のデータセットを作成する。

次に学習を行うと同時にパラメーターの設定も行う。ここでは、一度に学習させるデータの数と層の大きさ、学習回数を指定することができる。本研究ではデータの数を[128]、層の大きさを[64]に設定しこれを二つ重ねたもの学習させる。

学習したモデルを扱いやすいMAGファイルに変換する。

作成したモデルを利用してMIDIファイルを作成する。パラメーターの設定では、楽曲の生成数、楽曲の長さ、楽曲の最初の音を設定することができる。また、生成が完了した際に対数尤度というMagentaの考えるその楽曲の正確性が生成された楽曲毎に数字として表示される。

例えば、楽曲の長さは入力されたstep数が16であれば1小節分で楽曲を生成するので、128であれば8小節分になる。

Magentaは最初にきっかけの音を与えなければ楽曲の生成をすることができない。今回は八十八鍵の中心であり、MIDIの規格であるノートナンバー60のドを出だしの音として与えた。

4 学習回数による音楽らしさの評価

楽曲生成の際のパラメーターは下記の通りである。5曲生成しその中で対数尤度を基に一番数値が良いものを実験に使用する。楽曲は一曲の長さに相当するように3分を目処に生成する。

- 生成数:5曲
- step数:1440(90小節分)
- 出だしの音:60(ド)

表2 生成楽曲に対するアンケート項目

曲として成り立っているか?	1:悪い-5:良い
好きなフレーズはあったか?	1:少-5:多
この曲は機械的な曲か?人よりの曲か?	1:機械的-5:人寄り
ミステリードラマっぽいと感じたか?	1:悪い-5:良い

表3 学習回数での結果

学習回数	評価結果	学習回数	評価結果
100	2.35	600	3.05
200	2.8	700	2.95
300	3.45	800	2.8
400	3.15	900	2.7
500	2.35	1000	3.25

一番曲の雰囲気として分かりやすいミステリードラマの楽曲データを用いて、学習回数を

100, 200, 300, . . . と 100 回刻みで、100 回から 1000 回まで増加させて 10 個の学習モデルを作成し、生成させた楽曲を被験者に聞いてもらった。曲に対する 5 段階評価のアンケート(表 2)で得られた評価結果を学習回数ごとに合計し平均を出した(表 3)。今回のアンケートでは一番良い評価は学習回数 300 回で次いで 1000 回、400 回であった。

5 生成楽曲のテレビ音楽のジャンルらしさの評価

楽曲生成の際のパラメーターは下記の通りである。4 章の結果から学習回数 300 回で 10 曲生成しその中で対数尤度を基に一番数値が良いものを実験に使用する。評価の対象がメロとサビの部分のみなので、その点を考慮して 4 章の時の長さの三分の一の楽曲を生成する。

- 生成数: 10 曲
- step 数: 480 (30 小節分)
- 出だしの音: 60 (ド)

今回の実験では旋律のみを学習に使用し、さらにそれをメロとサビを分け生成し楽曲に違いが出るのかを実験した。

大河ドラマに関しては旋律がメロとサビの区別がないので一つにしてある。なお、ドラムなどでリズムを刻むだけの区間のデータ部分を除外している。

実験方法は 5 つのジャンルのサビとメロをそれぞれ学習、生成した楽曲を聴取してもらい、その楽曲が 5 つのジャンルのどれに聴こえるかを回答してもらった(表 4)。

表 4 各ジャンルの正答率

ジャンル	メロ	サビ
ミステリードラマ	0%	33.3%
夏ソング	16.7%	50%
卒業ソング	16.7%	16.7%
朝ドラマ	16.7%	16.7%
大河ドラマ	0%	

表 4 からミステリードラマのメロと大河ドラマに関してはそのジャンルらしさを感じる人はいなかった。しかし、ミステリードラマのサビと夏ソングのサビに関してはそのジャンルらしさについて少しは感じる人が多い。

6 SD 法における印象評価

5 章と同様に生成した楽曲を被験者に聴取してもらい SD 法により生成楽曲の印象語の評価を行った。表 5 に各ドラマ音楽の印象語をまとめる。

表 5 SD 法による印象評価のアンケート項目

ミステリードラマ	1: 自然な-5: 不思議な
	1: 単純な-5: 複雑な
	1: 落ち着いた-5: 緊張した
夏ソング	1: くだい-5: すっきり
	1: 暗い-5: 明るい
	1: 元気がない-元気な
卒業ソング	1: 嬉しい-5: 悲しい
	1: 賑やかな-5: 寂しい

	1: 充実した-5: 空虚な
朝ドラマ	1: 楽しい-5: 苦しい
	1: 冷めた-5: 情熱的な
	1: 陰気な-5: 陰気な
大河ドラマ	1: 臆病な-5: 勇敢な
	1: 弱々しい-5: たくましい
	1: こじんまり-5: 壮大

被験者に章 5 と同じ楽曲を聴取してもらいジャンルごとに選択した印象評価(表 5)に関して最低が 1、最高が 5 で評価を行ってもらった。例えばミステリードラマであれば、不思議な、複雑な、緊張しているを印象として抱きやすいのでこの三つが高い評価を受けると思われる。集計方法はジャンルごとの三つの印象評価を合わせて%で表したものをまとめた(表 6)。

表 6 SD 法における印象評価とのマッチング

ジャンル	メロ	サビ
ミステリードラマ	58.87%	82.27%
夏ソング	50%	67.8%
卒業ソング	63.27%	62.2%
朝ドラマ	55.53%	55.6%
大河ドラマ	51.07%	

表 6 から全体的に 5 割は超える数値を出せてはいるが、8 割を超えなければ楽曲の完成度は低いと予想する。しかし、表 4 を参照にミステリードラマのサビと夏ソングのサビに注目すると他の数値よりも比較的高い結果が出ていると感じる。

7 考察

表 4 と表 6 の結果を比較すると二つの表ともミステリードラマのサビと夏ソングのサビでは他の楽曲の結果と比べて、5%以上高い数値が出ている。このことからこの二つの楽曲に関してはディープラーニングを使うことで比較的音乐らしさを表現するという実験に関して成功していると言える。しかし、今回は単音のみでしか生成させることができなかった。今後は、和音でも同じように結果が出せるか研究が必要である。

8 まとめ

本研究では、Google のライブラリである Magenta を活用し、ドラマのカテゴリ別また行事や季節などのテーマで分けられている楽曲を使用して楽曲生成を試みた。結果としては全く再現することができないわけではなかったが、実用化できるほど完成度はまだ高くないと思われる。

今後の課題としては、より完成度の高い楽曲の生成、和音での楽曲生成またドラムなどのリズムを刻む楽器の再現などの実現が挙げられる。

9 参考文献

- [1]Magenta, <https://magenta.tensorflow.org/>
- [2]ヤマハミュージックデータショップ, <https://yamahamusicdata.jp>