

# LSTM によるメロディからのドラム及びベーストラックの自動生成

1190310 沖 貴司 【ソフトウェア検証・解析学研究室】

## 1 はじめに

近年、作曲用ソフトウェア等の発展により、音楽制作が身近なものとなっている。メロディの製作であれば、作曲に慣れていない人でも充分可能であるが、そのメロディに伴奏を付けるためには、より高度な専門知識や経験が必要である。そこで、機械学習を用いてメロディから伴奏を自動生成する研究がいくつかなされている。一方、音楽認識や音楽生成の分野では、時系列データを扱うことのできる再帰型ニューラルネットワークの一つである LSTM (Long short-term memory) を用いた研究もいくつかなされている。しかし、現時点では、LSTM を用いてポピュラー音楽の伴奏を生成する研究は知られていない。

そこで、本研究では、伴奏の中でも特に重要なパートであるドラムとベースに着目し、LSTM を用いてメロディからドラム及びベーストラックを自動生成する手法を提案する。

## 2 提案手法

本手法は、メロディの MIDI ファイルを入力として、(i) ベースの音高生成モデル、(ii) ベースのリズム生成モデル、(iii) ドラム生成モデルを使用して、ドラム及びベーストラックを生成する。その後、メロディ、ドラム、ベーストラックを合わせた MIDI ファイルを出力する。

### 2.1 ベースの音高生成

ポピュラー音楽等の一般的な楽曲において、ベースは、コードのルート音を演奏していることが多い。よって、本手法では、メロディからコード進行を生成し、それらのコードのルート音をベースの音高とする。ここでは、LSTM を用いてメロディからコード進行を生成する既存手法を使用する [1]。

### 2.2 ベースのリズム及びドラム生成

**■生成モデルの学習** ベースのリズム及びドラム生成は、まず、最初の 4 小節のリズムを生成し、それ以降のリズムを 1 小節ずつ生成していく。最初のリズムを生成するためのモデルは、4 小節間のリズムを、その 4 小節間のメロディのリズムを入力として学習する。それ以降のリズムを生成するためのモデルは、1 小節間のリズムを、その 1 小節間のメロディのリズムと、直前の 4 小節間のリズムを入力として学習する。

**■学習に使用するデータ** データセットとして、The Lakh MIDI Dataset v0.1 - Clean MIDI subset[2] を使用する。このデータセットは、約 17000 個の MIDI ファイルの集合であり、ポピュラー音楽も多く含まれている。モデルの学習には、これらの MIDI ファイルから、

The figure displays a musical score for 'Dre Mi no Uta' in 4/4 time. It consists of three staves: 'melody' (treble clef), 'bass' (bass clef), and 'drums' (percussion clef). The melody is a simple sequence of notes. The bass line provides a steady accompaniment. The drums feature a consistent rhythmic pattern. The score is attributed to R. Rodgers.

図 1 「ドレミのうた」の生成結果の楽譜

メロディ、ドラム、ベーストラックのそれぞれをリズム情報に変換して使用する。

## 3 結果と考察

5 曲のメロディを用いて生成した結果、ある程度音楽的に良いフレーズが生成可能であったが、音楽的に不必要な音や足りない音が多数含まれるという問題もあった。また、生成されたリズムがワンパターンであり、盛り上がり欠けるものとなった。よって、安定して良いドラム及びベーストラックを生成するためには、メロディから学習するだけでは不十分であり、楽曲の構造やジャンル等、より多くの情報を用いて学習する必要があると考えられる。

一例として、図 1 に「ドレミのうた」のメロディを入力したときの生成結果の楽譜を示す。

## 4 まとめ

再帰型ニューラルネットワークの一つである LSTM を使用して、メロディからドラム及びベーストラックを自動生成する手法を提案した。実際に、学習したモデルを使用して、メロディからのドラム及びベーストラックの自動生成が可能であることを確認した。

今後の課題として、ピアノやギター等、伴奏を担う他の楽器のトラックを生成することや、楽曲の構造やジャンル等のメタ情報を用いて学習することで、より複雑な伴奏が生成可能かどうかを確認することが挙げられる。

## 参考文献

- [1] H. Lim, S. Rhyu, K. Lee. "Chord Generation from Symbolic Melody Using BLSTM Networks," ISMIR 2017, pp. 272–278, 2017.
- [2] C. Raffel. "Learning-Based Methods for Comparing Sequences, with Applications to Audio-to-MIDI Alignment and Matching," PhD Thesis, Columbia University, 2016.