

創作ノート

ライブパフォーマンスにおけるニューラルネットワーク可視化システムの提案

Proposal on the Neural Network Visualization System for live performance

津曲 洸太, 三輪 眞弘

Kota TSUMAGARI, Masahiro MIWA

情報科学芸術大学院大学

Institute of Advanced Media Arts and Sciences

概要

筆者は人間と AI のインタラクションによるアンサンブル演奏を目指して、『LSTM を用いたライブパフォーマンスのためのセッションシステム』を制作した。本システムでは、与えられた人間の演奏データ (MIDI) に対して、それに呼応した自動演奏 (MIDI) をリアルタイムに生成する。また、本システムを用いたライブパフォーマンスにおいては、演奏の発音の要因を映像として提示する必要があると筆者は考える。本稿では、このシステムの概要の概説と、演奏とともに提示する映像について提案する。

定義

AI (artificial intelligence, 人工知能)

一般に明確な AI の定義は存在しないが、本稿では自動演奏を行うシステムを便宜上「AI」と呼ぶ。

1. 研究背景

AI 研究における課題の一つとして、AI による自動作曲が挙げられる。一般には、AI に過去の音楽を学習させた RNN による生成モデルによって、音楽を再構築することが行われている。しかし、AI が行うことは、学習した音楽の共通項や特徴量を見つけ出し、「似通った複製物」を量産することであり、このように AI が生成した音楽に創造性があるかどうかは一考の余地があると言える (徳井 2018)。これは適度な新奇性が求められるアート表現においては、きわめて重要な問題である。以上により、AI 単体で音楽を生成するのではなく、人間との音楽的インタラクションを行うことのできるシステムを構築し、それをリアルタイムに用いてライブパフォーマンスを行うことに目指した。これにより、人間と AI、それぞれ単体では達し得なかった、新たな音楽表現、創造性を目指すことができると考えた。

1.1. 人工知能合奏システムの作品例

ヤマハ株式会社は、人工知能合奏システム「MuEns」を使用した、人間と AI が協奏するプロジェクトを多数発表している (ヤマハ 2018)。これらは、プレイヤーの演奏をリアルタイムに解析し、どのような弾き方やタイミングで伴奏すれば調和の取れた合奏ができるかを瞬時に予測し実行している。しかし、演奏されるのはあらかじめ記録されたスコアの再生であり、AI による自律的な出力ではない。AI が逐次一音ずつ予測するシステムであれば、人間の行う未知のアドリブや即興に対し柔軟に対応した表現が可能となりうる。

2. システム

2.1. 演奏システム

2.1.1. MIDI の整数インデックスへの変換

MIDI ファイルを整数インデックスの時系列データへ変換する。Google の Magenta プロジェクトによる「Performance RNN」(Magenta 2017) を参考にし、以下のように変換した。

表 1: MIDI イベントと整数インデックスの対応

イベント	最小値	最大値	範囲
NOTE_ON	0	127	0 ~ 127
NOTE_OFF	0	127	128 ~ 255
TIME_SHIFT	1	100	256 ~ 356
VELOCITY_CHANGE	1	32	357 ~ 388

ここで TIME SHIFT は、ノートオン・オフの間隔を 0.01 秒から 1 秒までを 100 分割にマッピングしたものであり、VELOCITY_CHANGE は直後のノートオン

のベロシティ情報を 32 段階にマッピングしたものである。

2.1.2. LSTM モデルの学習

前述の方法で整数インデックスに変換された時系列データを用いて、以下の図に示すモデルを学習する。本モデルでは、2つのパートを入力しそれぞれの整数インデックスを推定する、多入力多出力かつシンメトリのモデルとなっている。このようなモデルにしたのは、人間による演奏と AI による演奏の2パートを同等の扱いにするためである。本モデルに2つのパートの整数インデックスが入力されると、それぞれの出力に 388 次元のベクトルが得られる。この各ベクトルは、各パートにおける整数インデックスの確率分布を表す。この値と正解の入力するタイムステップ数は長期的な依存関係を学習するためにできるだけ幅を大きくすることが望ましいが、今回学習に使用したマシンのスペックを考慮し、64 ステップの固定長とした。

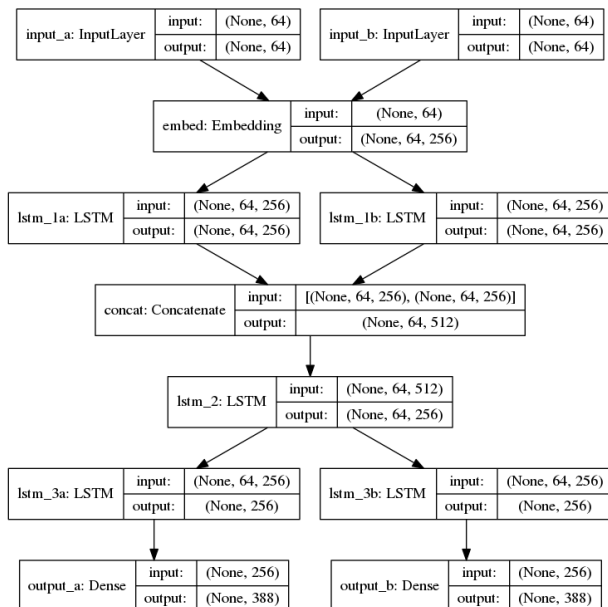


図 1: LSTM モデルの構造

2.1.3. 人間による演奏が与えられた際の生成

前述の学習により得られた LSTM モデルにより、2つのパートの MIDI を入力すると、各パートそれぞれの次のノート情報を推定する。このとき、一方の入力には人間による演奏、もう一方には本システムによる出力を再帰的に入力することにより、人間と AI 両者の演奏を元に計算を行い、MIDI として演奏をリアルタイムに出力する。このシステムにより、人間との音

楽的なインタラクションを行う自動演奏システムの構築を試みた。

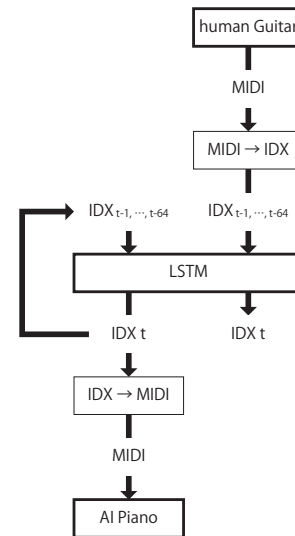


図 2: システムの全体構造

2.2. 可視化システム

ラップトップ PC などコンパクトなシステムを用いた即興演奏では、観客がコンピュータを操作している様子が見れないため、パフォーマンス中に何を行なっているのかわからないという問題がある。また、AI を用いたライブパフォーマンスにおいても、演奏が記録の再生によるものなのか、AI によるリアルタイム出力なのかかわからないといった、同様の問題が発生すると考えられる。この点において、パフォーマンスアートのひとつであるライブコーディングにおいては、パフォーマンス中にラップトップ PC で行なっている操作をすべて観客に見えるようにプロジェクションを行うことにより、この問題を対処している (田所 2018)。本制作においてはライブコーディングを参考に、AI の構造をオープンにして、リアルタイムに生成を行なっている様子を観客に見せることを試みる。

著者は過去に、主にニューラルネットワークの構造をニューロンの重みに着目して可視化させる作品を制作している (津曲 2017)。しかし、本システムにおいてはパラメータ数が膨大なため、可視化を内部構造すべてではなくニューラルネットワークの入力部と出力部に限定した。

3. 考察と今後の展望

現状の結果からは、人間と AI が音楽的インタラクションが行われているようには感じられていない。特に、人間が演奏を終えてもシステムは演奏をし続ける点などにおいて、それは顕著にあらわれている。これは学習データの不足などが起因していると思われる。学習及び演奏をリズムの呼応のみにするなど、研究の課題設定の見直しを検討する必要がある。



この作品は、クリエイティブ・コモンズの表示 - 非営利 - 改変禁止 4.0 国際 ライセンスで提供されています。ライセンスの写しをご覧になるには、<http://creativecommons.org/licenses/by-nc-nd/4.0/> をご覧頂るか、Creative Commons, PO Box 1866, Mountain View, CA 94042, USA までお手紙をお送りください。

4. 参考文献

- 徳井直生 (2018) 『AI による「新しい」音楽表現を実現するために—前編 AI と創造性』ウェブサイトより <https://link.medium.com/DTtnTd8pFU>
- 田所 淳 (2018) 「演奏するプログラミング、ライブコーディングの思想と実践 —Show Us Your Screens」 18-19 ビー・エヌ・エヌ新社

5. 参考作品

- 津曲 洸太 (2017) 『Processing を用いた深層学習の可視化の試み』 <https://youtu.be/YDWjeQzzIjY>
- ヤマハ株式会社 (2018) 『人工知能合奏システム MuEns』 <https://research.yamaha.com/ja/technologies/muens/index.html> (2019 年 2 月 28 日アクセス)
- Google AI Magenta 『Performance RNN』 <https://magenta.tensorflow.org/performance-rnn> (2019 年 2 月 28 日アクセス)

6. 著者プロフィール

津曲 洸太 (Kota TSUMAGARI)

1995 年、愛知県名古屋生まれ。名古屋学芸大学 メディア造形学部 映像メディア学科を卒業後、情報科学芸術大学院大学 (IAMAS) メディア表現研究科に在籍。ニューラルネットワークを主題に作品を制作している。

三輪 眞弘 (Masahiro MIWA)

1958 年東京生まれ。ベルリン芸術大学、ロベルト・シューマン音楽大学で作曲を学ぶ。アルゴリズムックコンポジションと呼ばれる手法で数多くの作品を発表。旧「方法主義」同人。「フォルマント兄弟」の兄。情報科学芸術大学院大学 (IAMAS) 教授