

創作ノート

機械学習モデルとモジュラーシンセサイザによる自己の演奏の演奏 Performing based on personal performance data using machine learning models and modular synthesizers

公文 太一, 小坂 直敏, 高橋 時市郎
Taichi KUMON, Naotoshi OSAKA, Tokiichiro TAKAHASHI
東京電機大学
Tokyo Denki University

概要

本稿は演奏の演奏をコンセプトとしたパフォーマンス作品『improvisation with mimics』についての創作ノートである。本作品の制作においてはとくに演奏に着目し実践を行う。本作品では、著者自身によるモジュラーシンセサイザの演奏を機械学習モデルに学習させる。それを2種類の機械学習モデルを搭載したモジュラーシンセサイザとして楽器化する。このモジュラーシンセサイザをパフォーマンスの中で演奏することにより演奏行為の拡張を試みる。本稿では、本作のコンセプトと利用する機械学習モデルとそのモジュール化について述べる。

1. はじめに

1.1. 生成モデルとアート

近年ではアート作品の創作において生成モデルを活用する事例が増えている。音楽においても活用は進み徳井直生『Emergent Rythm』(徳井 2022)は生成モデルによって合成された音響を人間がDJのようにキュレーションするパフォーマンスとなっている。この例では単に生成モデルに音楽を出力させるのではなく人間による選択が含まれている。一方、生成モデルにを単純に使用した作品の氾濫や、生成モデルを活用し他者の作風を盗用などが起こり議論を呼んでいる。Eslingは、創造性は社会的なコンテキストの中において評価されることから、生成結果のデータとしての誤差を使って学習を行う人工知能の限界を指摘している。データとしての近さだけを目的とする人工知能による生成は「infinite norm-generating(無限の規範の生成)」をおこない*創造性の局所最適化を招くとしている。一方、高度な計算によって人間の知覚能力を超えてしまうような

対象を扱うことが可能である人工知能と人間の行動が、それぞれの限界を補うことができるとしている。

1.2. 計算機による個人の創作の模倣

個人の創作を計算機によって再現する試みは、機械学習、深層学習が発展する以前から行われている。Harold Cohenによって開発されたAARON(Cohen 2006)はCohen自身が絵画を描く際の意思決定を条件分岐として記述したプログラムとなっている。Choenにとって、AARONの開発は計算機科学として絵を描くプログラムを実現する取り組みではなく、自身の芸術に対する意思決定をプログラムとして記述することにより定式化する試みであった。白石によるパフォーマンス作品『Inclusion』では、創作者を近似するAIとアーティストにより共同ライブコーディングを行い、新たなライブコーディングの形態を体現するとともに、新たなAIの領域の可能性を一人称による視点から示した(白石 2020)。岡部らは、人工知能と人間による共働作曲についての実践研究を行った。この中でプロのアーティストが自らの楽曲の特徴を学習した人工知能システムの利用しながら作曲を行い、その観察やインタビュー分析を行っている。この結果、人工知能技術との協働的な活動はアーティストの創造性の資源を底上げし、創作に関する省察の機会と幅を広げることを支援すると結論付けている(岡部 2019)。

1.3. モジュラーシンセサイザ

モジュラーシンセサイザとは電子楽器を構成する機能が個別に実装されたモジュールを自由に組み合わせて演奏する楽器である。現在では、Doepferにより公開されたユーロラック規格に多くのメーカーが追従し人気となっている。オシレータ、フィルタ、エフェクト

など信号処理を行うモジュール、シーケンサ、エンベロープジェネレータなど制御のためのモジュール、キーボード、タッチパネルなどインターフェースを提供するモジュールなどが販売されている。モジュールによって生成される制御信号である Control Voltage(以下 CV) や、音響信号をパッチケーブルにより自由に接続し全体としてのシステムを構築する。ユーザはインターフェースの操作やパッチケーブルのつなぎ替えなどにより演奏を行う。

モジュールの組み合わせと接続方法により多様な電子楽器を構築することが可能なモジュラーシンセサイザにおいて、電子楽器における機械学習モデルの活用の有効性を示すことが広範な電子楽器への機械学習モデルの応用可能性の提示につながると考えられること。発展の中において非音楽用のアルゴリズムの転用など様々なテクノロジーを活用する文化が醸成されおり、その中において機械学習モデルの音楽的活用を進めることにも価値があると考えている。

2. コンセプト -演奏の演奏-

本作のコンセプトとして演奏の演奏をあげる。自身の演奏を学習した機械学習モデルが搭載されたモジュラーシンセサイザを中心として演奏を行うことにより「演奏の演奏」を行う。モジュラーシンセサイザの演奏シンセサイザの状態の選択が含まれているその中で個人の趣向にある状態をモデルに学習させる。このモデルによる生成は、奏者個人の範囲において、Esling が指摘するようにデータに基づいた規範なものであると考える。機械学習モデルによる生成結果をモジュラーシンセサイザにより再度加工したり、本来の用途とは別の利用することにより演奏の演奏を行う。

2.1. 本作における機械学習モデルと演奏者

本作における機械学習モデルの役割は、演奏に基づいたデータの生成を行うことである。人間はその生成のためのパラメータの操作を行ったり、生成出力の利用、加工を行う。学習データは個人の演奏における趣向を反映した生成を行うが演奏者はモデルの生成物をモジュラーシンセサイザを通して新たに作りえる。モデルは奏者による演奏の 1 側面を取り出したコピーとしてふるまうが、その結果の利用方法は再度奏者にゆだねられる。

3. モジュラーシンセサイザの操作

3.1. パッチング

パッチングとはモジュラーシンセサイザのパネル上有る入出力端子同士をケーブルを用いて接続すること

とを指す。接続される信号には、オシレータやサンプラーなどから発せられる音響信号、Low Frequency Oscillator やエンベロープから出力される制御用信号 (CV)，主にシーケンサ類から出力される Trigger や Gate 信号などがある。モジュラーシンセサイザにおいては各モジュール同士をパッチケーブルにより接続しこれらの信号を組み合わせていく。また各信号は自由に接続できるよう設計されるため本来意図された用途と別の方法で利用することも可能である。例えば音響信号を Voltage Controlled Amplifier(VCA) の CVinput に接続することが可能である。この場合 VCA の音響入力につながれた信号と、CVinput に接続された信号の Amplitude Modulation となる。このようにモジュール同士をパッチングすることにより多様な変調、パラメータ同士の接続が可能になる。

3.2. 操作子によるパラメータ操作

モジュールのパネル上に展開されたノブやスライダー、スイッチを操作しパラメータを操作することが可能である。ノブやスライダーによりモジュールのパラメータを直接制御することが可能である。またモジュールによっては CVinput の効果を調整するノブなども存在する。またタッチパネルや、ジョイスティックなど多様なセンサを用いたモジュールも販売されている。

4. モジュラーシンセサイザの演奏

本作における演奏について定義する。ここではモジュラーシンセサイザによる演奏を 2 つに分類し、後述する機械学習モデルを搭載したモジュールにそれぞれ割り当てる。また演奏を学習したモデルの生成に対して CV を使ったコントロールや、生成出力を用いたモジュラーシンセサイザの演奏により「演奏の演奏」を行う。

4.1. 音色の演奏

1 つ目は音色の演奏である。モジュラーシンセサイザは、モジュールの組み合わせとそのパッチングにより多様な音色を合成する楽器であり、その音色の時間的変化を中心とした演奏を行う場合も多い。このためモジュラーシンセサイザの演奏において音色の演奏は重要である。音色の演奏はオシレータ、フィルタ、オーディオエフェクトなど音響信号を扱うモジュールに対して LFO など CV 生成を行うモジュールを使って変化を与えることにより行う。ここでは、これらのモジュールの組み合わせたパッチングと操作子による操作により音色を変化させていくことを音色の演奏と定義した。これに基づいて、オシレータ、フィルタ、オーディオ

エフェクト, LFO, シーケンサを用いた演奏を行い学習データの収録した。

4.2. リズムの演奏

2つ目はリズムの演奏である。モジュラーシンセサイザの演奏においてリズムパターンを扱う場合は、シーケンサを用いるか、アルゴリズムによる生成を行う場合が多い。ここでは、シーケンサを用いたリズムパターンを扱い、ステップシーケンサによる演奏を収録し学習データを収集した。



図 1: 改良版の RAVE モジュール

本作では 2 機種の機械学習モデル搭載モジュールを中心とした演奏を行う。

5.1. RAVE 搭載モジュール

1つ目は RAVE を搭載した音響合成モジュールである。RAVE(Caillon 2021) はリアルタイムに音響合成を行うことが可能な機械学習モデルとなっている。本作では、このモデルに自身によるモジュラーシンセサイザの演奏を学習させ用いる。音色合成用のモデルでありモジュラーシンセサイザによる音色の演奏部分を学習したモデルとして扱う。このモジュールのノブと外部モジュールから入力可能な CVinput を使うことにより音色の演奏を試みる。本機は以前に行ったパフォーマンス作品『improvisation with neural synthesizer』(公文 2023) にて制作したデバイスの改良版となっている。デバイスは Raspberry Pi 4 Model B にノブと CVinput の入力回路実装したものになっている。基本構成同一だが、パネルやノブを再設計し操作性の向上を行っている。またノブと、入力された CV を読み取るための ADC を ADS1115 から ADC1015 に変更した。これによりノブと CV を読み取る速度が 860samples per seconds(sps) から 3300sps と向上している。一方分解能は 16bit から 12bit へと減少しているが今回の用途においては十分であると判断した。音響の合成は Raspberry Pi 内の Pure Data によって行っている。nn~を用いて Pure Data 内にて RAVE を利用した音響合成行っている。

5.2. RAVE モデルの学習

本モジュールに搭載するためのモデルの学習を行った。データセットとして自身によるモジュラーシンセサイザの演奏を 60 分録音した。Raspberry Pi に搭載できる RAVE モデルは容量が限られるため、オシレータとフィルタ、エフェクタとそれらに対する制御用の LFO, CV 生成器のみを用いたシンプルな組み合わせで録

音を行った。データセット作成に含まれる演奏行為は各モジュールのノブの操作とパッチングである。この録音を用いて, Raspberry Pi 用の構成の RAVE を 6Mstep 学習させている。

5.3. RAVE モデルの潜在空間次元数

RAVE はデフォルトの設定では 128 次元の潜在空間を持つ VAE として定義されている。しかし実用上 128 次元の空間を扱うことは困難であり、音響合成を行う上あまり重要でない次元も存在することになる。そこで RAVE では、次元数を調整するために生成への寄与率が高い次元を取り出すように主成分分析を行い重要な次元を取り出すことが可能になっている。学習後にモデルを書き出す場合に fidelity と呼ばれるパラメータを設定することにより使用時の見かけの次元数を調整することが可能である。例えば fidelity を 0.9 と指定すると学習データセットに含まれるデータのうち 90% を説明するのに必要な数だけの潜在次元を持つモデルがモデルが実行される。RAVE を Max8 上にて読み込むためのエクステナルオブジェクトである nn~ を用いる。その際に設定した fidelity に基づく数の次元数のエンコーダからの出力、デコーダへの入力が展開される。図 2 は fidelity=0.9, 0.999 に設定し書き出したモデルを nn~ を利用して読み込んだ図である。0.9 では 4 次元分の inlet, outlet があるのに対し、0.999 では 63 個の inlet, outlet がある。fidelity を調整することによりモジュールの操作性と理解のしやすさが向上している。以前のバージョンでは 4 つのノブと 4 つの CV がそれぞれ 8 次元の入力を持つ RAVE デコーダの入力に接続されていたが、8 つの次元がどのように生成に寄与するのかがわかりづらかった。一方、今回のバージョンでは潜在次元を 4 次元にし入力を各 CV から受け付けるように変更している。ノブは各 CV 入力に対するアッテネータとしての機能を割り振った。潜在次元が 4 次

元になったことと、各次元に対してノブで重みづけが可能になったことにより操作者にとって各入力と生成物の対応関係がわかりやすくなっている。



図 2: nn~を利用して読み込んだ fidelity 調整済みの RAVE モデル

5.4. ドラムパターン生成モジュール

2つ目は演奏情報の演奏としてドラムパターン生成モジュールを活用する。こちらも自身によるパフォーマンス作品『improvisation with machine learning sequencer』において利用した機械学習によるドラム生成を行うモジュール(公文 2024)の改良版である。プロトタイプに対してユーロラック規格に沿ったインターフェースとパネルの実装、外部クロック入力の追加の改良を行った。こちらは自身の演奏あるいは制作したドラムパターンの MIDI を学習したモジュールになっている。キックのパターンを入力として、それに対応するスネア、ハイハットのパターンを生成する基本機能は変更されていないが、クロック入力を設けたことにより外部からステップの進行を制御ことが可能となった、クロックを変化させることによりパターンの演奏を変化させる。



図 3: 改良版のドラムシーケンサ

6. 結果と考察

本実践において、機械学習モデルに自身による演奏を学習させたモデルを新たな音源、パターン源として活用することが可能であった。筆者の主観ではあるが、自身の演奏の規範的な生成を行う存在として機械学習モデルを導入することにより、そのデータを再度演奏するというコンセプトは達成一定程度達成できたと考えている。一方、機械学習モデルや計算機の性能による制約についても見出された。例えば RAVE による音源合成モジュールのモデルを学習する際には、パラメータ数が少ないモデルを利用することになる。そのようなモデルに録音を学習させる場合、元の録音のバリエーションが多すぎると学習がうまくいかない問題が生じた。より幅広い演奏を扱い個人の演奏の全体をカバーすることが今後の課題となる。

7. まとめ

『improvisation with mimics』は生成モデルによる人間の創作の模倣を背景として、自身の演奏を学習させた機械学習モデルをモジュラーシンセサイザを介して演奏することにより演奏の演奏を行うパフォーマンス作品である。モジュラーシンセサイザの演奏を音色の演奏と、構造の演奏に分類し、それぞれのドメインについてに対応するデータを学習した機械学習モデルを搭載した 2 機種のモジュラーシンセサイザを使用しパフォーマンスを行う。自身の演奏を規範とした生成を行う存在として機械学習モデルを導入しその出力を再度演奏するという活用を行った。音色の演奏には使用する潜在次元を調整し演奏性の改良を行った RAVE を搭載したモジュールを使用した。構造の演奏としてはドラムパターンを生成する機械学習モデルを搭載したモジュールを使用する。これには外部クロック入力端子を実装し外部から変化を加えらるよう改良を行った。この 2 モジュールによって演奏の演奏をコンセプトとした即興演奏を行う。

8. 参考文献

IAMAS. 白石覚也. 新たな創造的な関係性をもたらす AI との協奏ライブコーディングシステム (<https://www.iamas.ac.jp/annual/mr/shiraishi-kakuya/>) 2024 年 9 月 19 日閲覧.

Philippe Esling, Ninon Devis. *Creativity in the era of artificial intelligence*.

Harold Cohen. AARON, *Colorist: from Expert System to Expert*.2006. (<https://www.aaronshome.com>).

com/aaron/publications/index.html)2024

年 9 月 19 日閲覧.

岡部 大介, 大谷 紀子. アーティストと人工知能技術の協働作曲にみる創造と省察. 質的心理学研究. 18巻, 1号, p.61-75, 2019.

Antoine Caillon, Philippe Esling. 2021. *RAVE: A variational autoencoder for fast and high-quality neural audio synthesis* (<https://arxiv.org/abs/2111.05011>) 2024 年 9 月 19 日閲覧.

公文太一, 小坂直敏, 高橋時市郎. 2023. 「機械学習によるパターン生成を用いたモジュラーシンセサイザ向けドラムシーケンサ」『先端芸術音楽創作学会会報』 Vol.15, No.2, pp.1-5

公文太一, 小坂直敏, 高橋時市郎. 2024. 「RAVE を用いたモジュラーシンセサイザの制作」『先端芸術音楽創作学会会報』 Vol.16, No.1, pp.25-30

9. 参考作品

徳井直生. 2022. 『Emergent Rhythm』 .

白石覚也. 2024. 『Inclusion』 .

10. 著者プロフィール

公文 太一 (Taichi KUMON)

2000年生まれ, 2023年より東京電機大学未来科学研究所情報メディア学専攻, ビジュアルコンピューティング研究室所属.



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/> をご覧頂くか、Creative Commons, PO Box 1866, Mountain View, CA 94042, USA までお手紙をお送りください。