

創作ノート

# 深層学習のデジタル信号処理を用いた楽器による環境音の変奏 Variation of Environmental Sounds by Musical Instruments using Deep-Learning Digital Signal Processing

大久保 雅基

Motoki OHKUBO

名古屋芸術大学

Nagoya University of the Arts

## 概要

「変奏」とは1つの主題に対し装飾的な音を加えたり、リズムやハーモニーを変更することが本来の意味である。本研究においてはその解釈を広げ、鳥の鳴き声を楽器演奏によって再現するような変換も変奏と呼ぶ。鳥の鳴き声を楽器演奏によって再現することは容易であるが、人々の話し声などの環境音は難しい。そこで、深層学習のデジタル信号処理による音色変換と、振動スピーカーを取り付けた楽器を用いることによって、環境音を楽器で変奏することを可能にした。その仕組みを用いて、サウンドインスタレーション『自販機のゴミ箱に嵌められたプラスチックカップ』として作品化した。本稿では本作のシステムの解説と、考察について述べる。

The word “variation” originally means to add agreement or change the rhythm or harmony of a single theme. In this paper, we broaden the interpretation of “variation” to include transformations such as the reproduction of bird-calls by instrumentation. While it is easy to reproduce birdsong by playing a musical instrument, it is difficult to reproduce environmental sounds such as people’s conversations. Therefore, the author has made it possible to transform environmental sounds with musical instruments using timbre transformation based on deep learning digital signal processing and musical instruments with attached vibrating speakers. The author created a sound installation entitled “Plastic Cups Fitted in a Vending Machine Trash Can” using this system. This paper describes the system of this work and its discussion.

## 1. はじめに

近年様々な深層学習を用いたデジタル信号処理技術が研究されている。その中でも Google が開発した Ma-

genta はメロディ生成や、即興演奏等に特化した様々なモデルが公開されており、機械学習を使った音にまつわるプロジェクトに広く用いられている。本研究では Magenta のライブラリである DDSP(Differentiable Digital Signal Processing) を用いて、環境音をヴァイオリンの音とスネアドラムの音に変換している。そして、その楽器に振動スピーカーを取り付け、変換された楽器音を再生することで、音響を楽器胴体に共鳴させ、楽器から発音されたように聞かせる。

本研究ではこのようにして得られた音響を楽器による環境音の「変奏」と呼ぶ。従来の「変奏」の意味とは1つの主題に対し装飾的な音を加えたり、リズムやハーモニーを変更した音のことである。その解釈を広げ、ある音を再現すること、そしてそれによって起こる変容が付加された音を変奏と呼ぶ。

## 2. DIFFERENTIABLE DIGITAL SIGNAL PROCESSING

DDSP(Differentiable Digital Signal Processing) とは、微分可能な計算式を持つ DSP モジュールにニューラルネットワークを組み込み機械学習を行う手法である。最初の実用例では、ある楽器音を他の楽器音に変換する手法が紹介されている。微分可能なオシレーター、フィルター、残響のパラメーターを用いて、異なる周波数の正弦波と時変フィルタでフィルタリングされたホワイトノイズを組み合わせるとオーディオを生成。生成されたオーディオとソースとなるオーディオを比較することで損失を計算する。バックプロパゲーションと確率的勾配降下法でトレーニングを行っている (Engel, Hantrakul, Gu, Roberts 2020)。

### 3. 振動スピーカーを楽器に取り付けた音楽

フランスの現代音楽グループ、アンサンブル・ルガルでは様々な楽器に対して振動スピーカーを取り付け、更に楽器音をマイクやピエゾ素子でリアルタイムに入力し変調することができる楽器を「ハイブリッド楽器」と呼び制作している。メンバーであるホアン・アロヨは『Smaqra』で弦楽四重奏のハイブリッド楽器を用いており、松宮圭太は『したたり』(松宮 2019)にてピアノをハイブリッド楽器にしている。(松宮 2019)

筆者は過去に演奏再現装置として、スネアドラムに振動スピーカーを取り付けた『sd.mod.live』(大久保 2018)を制作した。ハイブリッド楽器との違いは、マイクなどでリアルタイム入力はしないことである。その代わりに、事前に録音した楽器のサンプルを再生している。これは、ハイブリッド楽器がリアルタイム変調によるライブ性を主目的としていて、筆者の作品はフィクスド・メディアのような音響テクスチャの生成を目指していたことによる違いだと考えられる。

ライブエレクトロニクスのような作品では、楽器の音はマイクで入力・増幅され、スピーカーから変調された音が拡声されてきた。その手法では、楽器音はスピーカーから拡声される音にマスキングされ、生音の響きは得られなくなる。振動スピーカーを楽器に取り付けることによって、変調された音も楽器胴体に共鳴し、そこから発音しているように聞かせることができる。

本研究ではその現象を活かし、DDSP によって生成された合成音そのモデルとなった楽器に振動スピーカーを通じて再生することで、生の楽器の響きを作り出している。

## 4. 作品について



図 1: 展示風景

本研究では楽器に振動スピーカーを取り付け、DDSP による音色変換を用いることで、環境音を楽器で変奏す

ることを目指し作品化した。サウンドインスタレーション『自販機のゴミ箱に嵌められたプラスチックカップ』は宮城県仙台市のギャラリー TURN ANOTHER ROUND にて、2022 年 7 月 9 日から 14 日にかけて行われた筆者の個展「大久保雅基：都市の変奏」にて展示された。

仙台市の商店街を映像で記録し、フィールドレコーディングされた環境音を、DDSP を用いてヴァイオリンとスネアドラムの音に変換することで変奏している。事前にヴァイオリンとスネアドラムの音色モデルを作成しておき、そのモデルを用いて環境音をそれぞれの楽器音に変換している。そうして生成された音声データを、振動スピーカーを取り付けた楽器に再生することによって、楽器が持つ周波数特性に共鳴させて、楽器から鳴らされたかのような音現象を作る。

### 4.1. 収録

収録は宮城県仙台市の商店街にて行われた。撮影機材は、ビデオカメラとして Panasonic Lumix S5 が、スタビライザーに DJI RS 2 Pro コンボが用いられた。音色変換によるオリジナルソースはダイナミクスが十分に取られている必要がなく、コンプレッションがかけられ明瞭に音が聞こえることが重要だったため、特にマイクは用意せず、Lumix S5 にて収録された音声そのまま使用することにした。

### 4.2. モデルの生成

音色モデルを生成するために、まずはヴァイオリンとスネアドラムの収録を行った。ヴァイオリンの収録には通常の弦を弓で擦る演奏以外にも、ピッチカートや強いスル・ポンティチェロ、胴体を叩くなど特殊奏法を含めて 17 分程録音を行った。ヴァイオリンの胴体から鳴らすことが可能な様々な周波数特性の音を抽出するためである。マイクには DPA VO4099 を用いた。スネアドラムの収録では、打面をスティックで叩く通常の奏法のみを採用した。これは振動スピーカーを取り付ける位置が打面であり、胴体を叩くなどの音響は不自然な発音になるからである。マイクには Oktava MK-012 を用いて 15 分ほど収録した。

録音したヴァイオリンとスネアドラムの音源を Magenta の DDSP プロジェクトで公開されている Train Autoencoder Colab を使用して、音色モデルの生成を行った。

### 4.3. 音色変換

同様に Magenta の DDSP プロジェクトで公開されている DDSP Timbre Transfer Demo を用いて、そこに前述した方法で作成した音色モデルをアップロードし環

境音から楽器音への変換を行った。24分ほどの環境音をアップロードするとエラーが起きたため、5分ごとの5つのオーディオファイルに分割して、それぞれのモデルで変換を行った。そして変換されたスネアドラムの音源を左チャンネル、変換されたヴァイオリンの音源を右チャンネルに振り分けたステレオのオーディオデータを作り、映像データの音声部分をそれに置き換え、映像データを書き出した。

#### 4.4. システム

本作のシステムは映像と音声に分岐している。メーカー不明のメディアプレイヤーから HDMI 出力された映像と音声を、Amazon ベーシックのコンバーターを用いて分岐している。映像信号はプロジェクターに送られスクリーンに投影される。音声信号はオーディオアンプ Fostex AP20d に送られ、L 信号をスネアドラムに取り付けられた振動スピーカー Dayton Audio TT25-8、R 信号をヴァイオリンに取り付けられたメーカー不明の振動スピーカーに送ることで、再生を行った。



図 2: 振動スピーカーを取り付けたスネアドラム

### 5. 考察

#### 5.1. 音響メディアとしての装置

谷口、中川、福田は「レコードやラジオ、電話など 19 世紀後半に登場した音を記録・再生する技術によって、人々のコミュニケーションを媒介するもの」を「音響メディア」と呼んだ。その中でも、録音技術を駆使して録音媒体に固定されるかたちで制作される音楽を「レコード音楽」と呼び、その場での演奏に基づく「ライブ音楽」を谷口は区別した(谷口, 中川, 福田 2015)。一方で作曲家の三輪真弘は、前者は従来の音楽とは別物の「録音芸術」もしくは「録楽」と呼び区別して

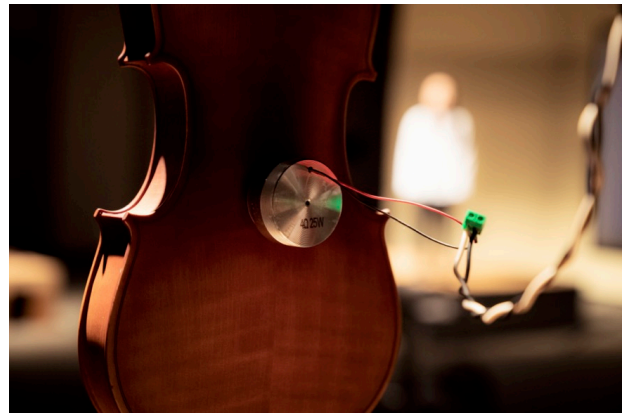


図 3: 振動スピーカーを取り付けたヴァイオリン

いる。そして後者を「音楽」と呼ぶ道を選んだ(三輪 2010)。本研究における装置も音響メディアであり、録音した音を再生することから録楽の性質を持つが、それだけでは含みきれない側面も持っていると考えられる。レコード音楽や録楽は、スピーカーでの再生を前提としているが、本研究における装置の最終出力は楽器である。それはオルゴールや自動オルガン、自動ピアノのような自動演奏楽器の延長にあるものと解釈する方が正しいと言えるだろう。オルゴールや自動オルガンはピアノロール、自動ピアノは MIDI データという入力データを用いるが、本研究における装置はオーディオを使用する。ピアノロールは音高と音価が記述でき、自動ピアノにおける MIDI データは音量も記述できるが、オーディオデータは波形であり時間的解像度が高いことにより、音色も再現することができる。この装置は自動演奏楽器のカテゴリの中で、新たに解像度の高い表現を可能するものである。

#### 5.2. 新たな変奏

その新たに可能となった表現の一つが、本研究にて行った環境音の楽器での変奏であろう。楽器に振動スピーカーを取り付け環境音を再生しただけでは、楽器の胴体が持つ周波数特性でフィルターしたオーディオである。深層学習のデジタル信号処理によって可能になった音色変換を用いて、環境音を楽器の音色に変換し、それを楽器で振動させることによって、その楽器が演奏されているように聞くことができる。

作曲家のヘルムート・ラッヘンマンは楽器の特殊奏法を用いて作曲する自身の作曲法を「楽器によるミュージック・コンクレート」と呼んでいる。しかしそれはノイズ音を用いて作曲することを分かりやすく例えるため、音響テクスチャによって作曲するミュージック・コンクレートの概念を用いただけである。実際、ラッヘンマンの作品における演奏法は、ミュージック・コンクレートで多用される環境音を再現したものではな

い。しかし本研究による方法を用いることで、その環境音を楽器で変奏することが可能となる。ミュージック・コンクレート（現在は「アコースマティック音楽」と呼ぶほうが一般的）をスピーカーではなく、楽器から再生することも可能になるだろう。

しかし、環境音を変奏しているかどうかを理解できるのは、本作のように映像が付随していることによるものである。もしくは、筆者の音楽劇『声のゆくえ』で用いられた表現である、役者の声が発声された後、ヴァイオリンの音色に変換された声の音が再生されるように、オリジナルの音源と変換後の音源を聴き比べられるように提示されていることによって変奏だと理解できる。現在の変換の精度では音によっては、変換されたものと分からないものが多い。

## 6. まとめ

本研究におけるシステムを用いることで、環境音のようなオーディオを楽器で変奏することが可能となった。環境音をスピーカーから聴くことは写実的な表現を受容することであり、その音によって風景のイメージを思い浮かべることができる。その瞬間にスピーカーそのものの存在を認識はしていない。一方で楽器に変奏された環境音とは、音色という楽器のアイデンティティが現前に先立つため、楽器というオブジェクトの存在を強く感じることができる。本研究における手法は、デジタル信号処理が主流な時代において、スピーカーによって失われつつある楽器のような発音する媒体の存在を、再び取り戻すことができる一つの道筋なのかもしれない。

## 参考文献

- Jesse Engel, Lamtham (Hanoi) Hantrakul, Chenjie Gu, Adam Roberts. 2019."DDSP: Differentiable Digital Signal Processing." *ICLR 2020*
- 大久保雅基. 2018.『映像と振動スピーカーを用いた演奏再現装置によるパフォーマンス作品』先端芸術音楽創作学会会報.10(3), 30-34.
- 谷口文和, 中川克志, 福田裕大. 2015.『音響メディア史』2nd ed. まえがき i, 10-11. ナカニシヤ出版.
- 松宮圭太. 2019.『ハイブリッド楽器の研究『したたり』ピアノと電子音響のための(2019)の創作意図と表現方法を巡って』先端芸術音楽創作学会会報.11(3), 34-38.
- 三輪眞弘. 2010.『三輪眞弘音楽藝術一全思考一九九八-二〇一〇』171-172, アルテスパブリッシング.

## 7. 参考作品

- Arroyo, Juan. 2014. *Smaqra pour quatuor à cordes*. Tana Quartet. <https://www.youtube.com/watch?v=IN9E4JmzuQ8> (2023年3月4日アクセス)
- Anderson, Leroy. 1948. *Sleigh Ride*. 『そりすべり』
- Honegger, Arthur. 1924. *Pacific 231*. 『パシフィック231』
- Russolo, Luigi. 1910-1930. *Intonarumori*. 『イントナルモーリ』
- Sparke, Philip. 1987. *Orient Express*. 『オリエント急行』
- Mahler, Gustav. 1888. *Symphony no.1 1st mov.*
- Messiaen, Olivier. 1953. *Réveil des oiseaux*. 『鳥たちの目覚め』
- Messiaen, Olivier. 1955-1956. *Oiseaux exotiques*. 『異国の鳥たち』
- Messiaen, Olivier. 1956-1958. *Catalogue d'oiseaux*. 『鳥のカタログ』
- Messiaen, Olivier. 1960. *Chronochromie*. 『クロノクロミー』
- Vivaldi, Antonio. 1723. *Le quattro stagioni "La Primavera" 1st mov. Allegro*. 『四季 春 第1楽章』
- 大久保雅基. 2018.『sd.mod.live』(初演)2018.3.6 第21回日本電子音楽協会定期演奏会「詠像詩」浦安音楽ホール ハーモニーホール.
- 大久保雅基. 2022.『音楽劇 声のゆくえ』(初演)2022.8.13-14 エル・パーク仙台 ギャラリーホール.
- 松宮圭太. 2019.『したたり』

## 8. 著者プロフィール

### 大久保 雅基 (Motoki OHKUBO)

1988年宮城県仙台市出身。作曲家、サウンドエンジニア。名古屋芸術大学芸術、愛知淑徳大学、相愛大学非常勤講師。洗足学園音楽大学音楽・音響デザインコースを卒業。情報科学芸術大学院大学メディア表現研究科修士課程修了。先端芸術音楽創作学会、日本電子音楽協会、日本AI音楽学会会員。



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。