

ソフトウェアエフェクタを利用した同一機材を必要としない 機械学習によるエレキギター音色の自動再現手法の検討

有山 大地^{1,a)} 安藤 大地¹ 串山 久美子¹

受付日 2020年1月30日, 採録日 2020年9月10日

概要: エレキギターの音作りは一般的にアナログのアンプ, エフェクタなどを用いて行われ, その組合せは膨大な数にのぼり, 演奏者やエンジニアが求めている音色を得るために大きな労力を必要とする. たとえば, 憧れのギタリストの音色を再現したいという場合, まず同一の機材の入手, 加えてステージセッティングの画像などからエフェクタのつまみの位置を推測する必要がある, 手動での再現のプロセスは困難なものである. そこで音色再現の問題を解決するため, ユーザが所持している1つのソフトウェアエフェクタのみを用いて, ユーザが目標とする音になるべく近い音作りができることを目標として, エフェクタの複数のパラメータを機械学習により探索する手法の開発を行った. 本研究は現在までにギターエフェクトの1つである「歪み系エフェクト」を中心に, 開発したシステムによって再現された音色について調査・実験を行い, 再現度の評価において一定の成果を得た. また, 本研究で開発した手法がエレキギター音色の再現において有効であることを確認した.

キーワード: エレキギター, 楽器音響, 遺伝的アルゴリズム

Investigation of Automatic Reproduction Method of Electric Guitar Tones by Machine Learning Using Software Effectors without the Need of the Same Equipment

DAICHI ARIYAMA^{1,a)} DAICHI ANDO¹ KUMIKO KUSHIYAMA¹

Received: January 30, 2020, Accepted: September 10, 2020

Abstract: Though the sound making of electric guitar is generally carried out using analog amplifiers, effectors, etc., the combination of them becomes enormous numbers, and it is difficult to obtain a tone color required by performers and engineers. For example, consider the case where you want to reproduce the tone color of the guitarist you admire. First, it is necessary to obtain the same equipment as the target guitarist, and then to guess the position of the knob of the effector from the image of stage setting, etc. It is proven that manual reproduction of a tone color is difficult, even if only this procedure is considered. In order to solve the problem of a tone color reproduction, we developed a method to search for multiple parameters of the effectors by machine learning, aiming at creating a sound as close as possible to the user's target sound using only one software effector owned by the user. In this research, we investigated and tested a tone color reproduced by the developed system, focusing on "Distortion Effect" which is one of the guitar effects, and obtained a certain result in the evaluation of the reproducibility. And, it was confirmed that the technique developed in this research was effective in the reproduction of a electric guitar tone color.

Keywords: electric guitar, instrumental acoustics, genetic algorithm

1. はじめに

エレキギターは, 弦の振動をボディを通した空気振動經由で発音するのではなく, マグネティックピックアップの

¹ 東京都立大学大学院システム研究科
Graduate School of System Design, Tokyo Metropolitan
University, Hino, Tokyo 191-0065, Japan
^{a)} scarcement0623@gmail.com

電磁誘導の信号を楽器の音として用いるという特徴がある。一般的に電気信号として伝達されるため「エレキギター」と「ギターアンプ」の間に「エフェクタ」と呼ばれる波形を加工する装置をつなぎ、音に様々な変化をもたらせることが多い。

エレキギターの音を決定する大きな要素としては、エレキギターを構成する木材や弦巻・ブリッジなどの「パーツ」、音を増幅する「ギターアンプ」、金属の弦の振動を電気信号に変換する「ピックアップ」、そして電気信号である波形を様々な手法で加工する「エフェクタ」の4つをあげることができる。このうち後ろ2つが電気回路によるものであり、変化を加えることが容易であるが、前2つはそうではない。そして自分が目指す音を再現する作業を一般的に「音作り」と呼び、音作りは上記の4要素の組合せで行われる。

ギターそのもの、あるいはエレキギター用のアンプやエフェクタを製造している企業・工房は、国際的なギターフォーラムを持つ GUITAR LIST^{*1}に登録されている数だけでも3,663件にのぼり、製品化されている4要素はそれぞれ膨大な数が存在していることが分かる。また、「ギターアンプ」と「エフェクタ」は必ず一定のエフェクトを提供するわけではなく、それぞれ無段階で調節できる複雑な、しばしば直感的でない複数のパラメータを持つ。エレキギターの音作りでは一般的に複数台の「エフェクタ」を直列、あるいは並列に接続して行われるため、さらに組合せが複雑化し、演奏者やエンジニアが求めている音色を得るためには非常に労力がかかる。

1.1 現在までの音作りの背景と問題点

音作りの指標として一般的に好まれているのは、「名ギタリスト」と呼ばれる人々の音作りである。多くのユーザは、自らエフェクタを用意し、手探りでパラメータを調節することで、自分の目指すギタリストの音色を模索することになる。しかし、パラメータの組合せやエフェクタの組合せは、1980年代に出現した House^{*2}や Drum'n'bass^{*3}などにルーツを持つ、楽曲制作にコンピュータやデジタルエフェクトを積極的に取り入れた EDM^{*4}と呼ばれる音楽の台頭などを背景に、近現代の音楽シーンにおいて巨大化している[1]。たとえばエフェクタのデジタル化により、従来のアナログ回路では作れなかった音色を作ることができる

ようになっている。

そのため、たとえば憧れのギタリストの音色を再現したいという場合、機材の情報やアンプ、エフェクタの種類、エフェクタ接続順、つまみの位置などの写真によるパラメータ推定が必須であり、これらの情報なしには再現が困難である。多くのギタリストは自ら、あるいは熱心な研究者の推測によって、「アンプ」、「ピックアップ」、「エフェクタ」の組合せを公開している。しかし、実際に公開された組合せを試しても、多くの場合期待した結果は得られない。というのも、ギターの音色とは、演奏者の癖や、使用している機器の個体差によって微妙な違いが生じてくるためである。また、プロのギタリストの使用する機材は非常に高価な品・ワンオフ品・希少品である可能性もあり、そもそも手に入れることが困難な場合も多い。

1.2 関連研究

デジタル楽器であるシンセサイザの音色を再現する試みは Barkan ら [17] によって提案されている。シンセサイザは調節可能な複数のパラメータを持つが、Barkan らの提案する手法では、音源を入力としてとることで、ニューラルネットワークによって自動的に音色を再現することができる。

ただし Barkan らの研究は、あるアーティスト A がシンセサイザ X を用いて、他のアーティスト B が同一のシンセサイザ X を用いて生成したサウンドを複製することを想定している。

我々の手法では再現したい音色で用いられている機材は考慮せず、用意した複数のエフェクタすべての組合せのなかから、類似する音色を多点同時探索である遺伝的アルゴリズムによって探索する手法を採用している。そのため、必ずしも目標とする音源で用いられているエレキギター/エフェクタと同一の機材を用意する必要がない。我々は同一機材を用いない再現実験を行うことでこの主張を検証する。

また、弦楽器の楽器音色の再現に関する関連研究としては、単純回帰型ネットワークで再現する試みが文献 [18] で提案されている。

上記のような物理モデルを用いた楽器音色の再現は、電子音楽制作において広く用いられる合成手法の1つである。

同様に糸山ら [19] の研究では、コンピュータ上で複数のソフトウェア音源とランダムなパラメータセットを生成し、線形回帰分析を用いて自動的に目標の音色と近い音源・パラメータのセットを推定する方法が提案されている。

さらに、瀧田ら [20] は深層学習モデルの1つである Deep Belief Network を用いることで、糸山らの提案する線形回帰分析を用いた推定よりも誤差を縮めることに成功している。

文献 [18], [19], [20] の手法はいずれも学習後の演奏にお

^{*1} イギリス在住の Bromham の運営するエンサイクロペディア・サイト <https://www.guitar-list.com/> (2020年1月22日閲覧)

^{*2} サンプリングやリミックスといった手法を用い、リズムマシンやコンピュータを用いて4つ打ちのリズムを基本に作られる音楽。

^{*3} あらかじめ録音したリズムパターンを切り刻んで並べなおし、150-200BPM ほどの速いビートで作られる音楽。エフェクトを多重に適用する幻想的な音作りが行われる。

^{*4} Electronic Dance Music の略称。しばしばシンセサイザやサンプラーを用いてコンピュータ上で作られる、エンタテインメント用の音楽の総称。

いては学習元の楽器は用いず、再現した音色のデータを主にシンセサイザ用のソフトウェア音源として利用する点で、我々のアプローチと異なる。これらの研究では推定するパラメータが仮想シンセサイザ、仮想エフェクタのものを対象としており、現実のエレキギターで用いる物理的なアナログエフェクタに適用する方法については提示されていない。

加えて、上記の研究はいずれも単一楽器音を対象としている。つまり MIDI ノートオンからノートオフまでの単独音を目標音源とし、目標の単独音にできるだけ類似した音を生成するようなシンセサイザのパラメータを推定することを目的としているが、本研究ではフレーズ演奏を入力・目標音源としている点で異なる。

また、松谷 [21] は異なる特性を持つエレキギターの音色を再現するために、適応信号処理を採用した。松谷は同一の形状をしたエレキギターであっても、使われているパーツ、あるいは木材の差異によって、音色が異なることを指摘している。適応信号処理を用いた音色再現は、広い音域で音色再現が可能であることが示されている一方で、あらかじめ指定した音高でしか再現を行えないこと、そして適応フィルタで発生する遅延によりリアルタイムの演奏が行えないという2つの問題点も指摘されている。

我々の研究ではエレキギターに複数のエフェクタがつながれていると仮定したうえで、ソフトウェアエフェクタのパラメータを変化させることで音色再現を行う。エフェクタのパラメータセットとして音色再現を行うことで、学習と演奏を切り離して扱うことができ、音高の制限や遅延の問題を解決する。

1.3 目的と手法

本研究では、エレキギターの音作りを手助けするため、「パーツ」、「アンプ」、「ピックアップ」、「エフェクタ」などの組合せがいっさい不明である状況においても、自身の所有する機材だけを用いて理想とするギターの音色を再現できるシステム、つまり機材や環境に左右されない音色再現のシステムを開発することを目指している。

本論文ではエレキギターの音色再現に適切な音響特徴用・距離尺度、および機械学習によるエフェクタのパラメータ探索の問題に焦点を当て、まずは所有する機材を「エレキギター本体」と「ソフトウェアエフェクタ」の2つに絞った音色再現を検討する。「ソフトウェアエフェクタ」は音色再現システム用に実装した連続的なパラメータを持つものを扱い、より拡張的な市販されている「ソフトウェアエフェクタ」や、離散的なパラメータを持つ「ソフトウェアエフェクタ」に対する問題は今後の研究のために残しておく。

また、本研究で想定する状況では、前述の「パラメータの組合せが多様化する」という問題が浮上する。音色再現システムを開発するにあたり、このパラメータの多様化に

対応することも大きな課題の1つである。そこで、開発するシステムは可能な限り自動化されることが望ましい。自動化の手法としては、進化論的計算のうちの遺伝的アルゴリズムを採用している。理想とする音色を目標音源として入力し、これを正解としてパラメータを変化させ、自動的な探索を行うという手法である。本研究の同一機材を用いずに音色を再現するという性質上、音色再現におけるパラメータの組合せは1通りではないと予測される。それに加え、エフェクタのパラメータと音色の変化は必ずしもリニアな関係にはなく、音色の変化がある1つの軸のみに沿ったものであるとも限らない。そのため、多峰性を許容し、かつ多点同時探索である遺伝的アルゴリズムが適当であると考え、手法として選択した。

2. 機械学習によるパラメータの自動推定

まず、本研究で実装した機械学習によるパラメータ推定システムについて述べる。このシステムは現段階では14のパラメータを持つ歪み系のソフトウェアエフェクタを備える。つまり、このシステムではまず「自身の所有するギターの音源」にランダムに生成したパラメータのソフトウェアエフェクタを適用し、出力される音色が「目標とする音源」に近くなるパラメータセットを機械学習によって探索する。その際用いる音響特徴量、および距離尺度については3章で述べる。

機械学習の手法については、本研究では遺伝的アルゴリズムを採用した。遺伝的アルゴリズムは多点同時探索であるため、エレキギターのエフェクトパラメータのような多峰性が予測される問題に適していると考えられる。以下にその手順について述べる。また、推定のフローを図1に示す。

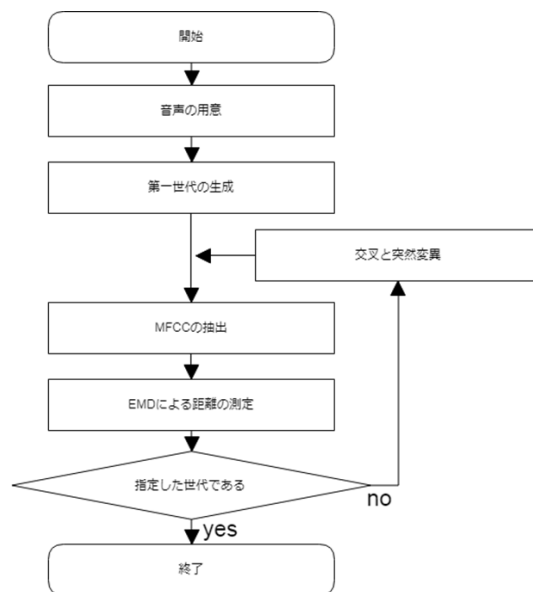


図1 遺伝的アルゴリズムによるパラメータ推定の流れ

Fig. 1 Flow of effect parameter estimation by using Genetic Algorithm.

まず、「目標とする音源」と「自身の所有するギターの音源」の2つの音源を用意する。前者を評価の基準とし、後者にランダムなエフェクトをかけ、距離の測定を行うためである。用意する音源は計算領域を揃えるため、所謂「CD音質」と呼ばれるサンプルレート 44,100 Hz, 16 bit の wav 形式に限る。また、エレキギターの音声はその性質上モノラルであるため、目標とする音声もこれに準拠するものとする。音声の処理には青木による C 言語のライブラリ [11] を利用した。

検証を行うために自由にパラメータを変更できる歪み系エフェクタを C++ にて実装した。実装したエフェクタはハードクリッピング、およびソフトクリッピングによる歪み、音量レベルに加え、10 帯域のイコライザ、コンプレッサをパラメータとして持つ。イコライザはピーキングフィルタを使用し、中心周波数はギターの周波数特性を考慮し、31.25 Hz, 61.5 Hz, 125 Hz, 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, 8 kHz, 16 kHz に設定した。強度は -12.0 dB から 12.0 dB の間を浮動小数点数でとる。

続いて「自身の所有するギターの音源」に対してランダムなパラメータの歪み系エフェクトを適用して第一世代を生成する。今回想定している歪み系のソフトウェアエフェクタのパラメータ調整は、一般的につまみを模した GUI で行われる。この場合パラメータは 128 段階あれば十分であるため、エフェクタの個々のパラメータは精度が高い実数である必要はない。そのため、実数値遺伝的アルゴリズムを用いなくとも、エフェクタのパラメータ探索では遺伝的アルゴリズムでも十分な性能を持つ。またギターのエフェクタの場合、1つのエフェクタを駆動するパラメータの数が少なく、かつ1つのパラメータが取りうる状態数も少ないということは、組合せ爆発が起きないことでもある。したがって遺伝的アルゴリズムは今回実験に用いたパラメータの組合せには十分に対応できる。そこで低精度の組合せ最適化として計算コストが一番低い遺伝的アルゴリズムを採用した。

1 世代を n 個の個体と設定した場合、 n セットのパラメータが生成され、自動的にナンバリングされ、wav 形式で書き出される。ここで生成されるパラメータ 1 つ 1 つを遺伝子、1 音声分のパラメータの集まりを染色体、染色体とそれを用いて出力される音響をまとめて個体と定義する。また、染色体は歪み、レベル、10 帯域のイコライザ、コンプレッサの合計 14 個の遺伝子を持つ。

1 つの遺伝子はそれぞれ 1 byte で表され、整数値で $[0, 255]$ の定義域を持つ。MIDI の大部分の数値情報は 7 bit で表現され、これはその倍であり、「つまみの設定量」としては十分な解像度である。たとえばエフェクトパラメータ x が $[1, 1, 1, 0, 0, 0, 0, 1]$ であり（整数値では 225）、パラメータの上限・下限がそれぞれ $E_{\max} \cdot E_{\min}$ であるとき、適用されるエフェクトのパラメータ $f(x)$ は次式で求められる。

$$f(x) = \frac{x}{255}(E_{\max} - E_{\min}) + E_{\min} \quad (1)$$

第一世代の個体から抽出された音響特徴量について距離の測定を行い、シグネチャ間の距離を求める。求めた距離はテキストデータとして書きだされ、第二世代以降はこれを基に世代の交配を行う。遺伝的アルゴリズムの実装は伊庭による解説 [12] と渡辺らによる実験 [13] を参考に、予備実験の結果も考慮し行った。次世代に残す個体はエリート選択方式とルーレット選択方式の両方を用いて選択する。次世代のうち 1 つは最も距離の近かった個体をそのまま引き継ぐ。続いて全体の 8 割の個体は距離を基に算出したスコアを基準として、ルーレット選択方式で選んでいく。 n 番目のスコア $s(n)$ を求める式は f_{EMD} を EMD によって導かれた距離として、以下のとおり設定した。

$$s(n) = 100/(f_{\text{EMD}} + 1) \quad (2)$$

上式よりスコアの最大値は 100 となり、距離が遠いほどスコアは小さくなる。また、スコアの合計を f_{sum} とした場合に、各個体がルーレット盤面に占める割合 $f(n)$ は次式で求められる。

$$f_{\text{sum}} = \sum_{i=1}^n s(i) \quad (3)$$

$$f(n) = s(n)/f_{\text{sum}} \quad (4)$$

ここで、0 から f_{sum} までの数をランダムに生成し、上式をもとに対応する染色体の選択を行う。染色体の交叉は 2 点交叉を用いて行った。ルーレット選択によって個体を順に 2 つずつ選び、これを親として交配を行った。染色体を交叉する点数は 2 点で固定であるが、交叉する範囲はランダムに決定される。染色体には 14 の遺伝子が含まれるため、交叉する点の候補は 13 カ所から選ぶことができる。そこで 1 から 13 までのランダムな数値を 2 つ生成し、その数値を基に 2 点交叉を行う区切りを決定した。

これを個体の数が第一世代と同じになるまで行う。エリート選択した 1 つの個体を除き、残るすべての個体に対して突然変異の判定を行う。突然変異は各個体の遺伝子の合計からランダムに任意の数の遺伝子を変異させる。この突然変異の確率 f_{mutation} は大向による研究 [14] を参考に、遺伝子のビット長を L として次式で求められた値を使う。

$$f_{\text{mutation}} = 1/L \quad (5)$$

なるべく多くの種類のエフェクタに適用できることを目指し、類似度の多峰性が予想されるエフェクタにも対応するため、交叉と突然変異だけでは局所解に陥る可能性もある。そこで、局所最適解に陥ることを防ぐため、交叉と突然変異だけではなく、新規ランダム生成個体への置き換えも行うこととした。置き換わる個体数 f_{replace} は個体の総数を f_{total} として次式で求まる。

$$f_{\text{replace}} = f_{\text{total}}/5 - 1 \quad (6)$$

突然変異を終えたところで世代の生成は完了とし、再度距離の計算を行う。これを繰り返して最適解を導く。

3. 音響特徴量の設定と距離尺度の検討

続いて本研究が開発した再現手法について、どのような形に録音データを変換し、どのような尺度で距離を測るのかについて述べる。

音色の距離を測定するためには、目標となる音源（CD音源など、エフェクトの適用されたギター音源）と自分の所有するギターの録音（なにもエフェクトをかけていないギター音源）の2つのデータを入力として取り、何らかの形で比較することになる。ここで、再現したいのは波形そのものではなく、その音色「らしさ」であることに注意する。したがって「らしさ」を適当に表す音響特徴量を抽出し、その距離により機械学習を行う必要がある。

3.1 音響特徴量の検討

ギター音源の音響特徴量として、機械学習で繰り返し計算させることを考え、データの圧縮率・楽器に対しての適応事例をもとに5つの特徴量を比較することとした。

3.1.1 Mel Frequency Cepstral Coefficient

Mel Frequency Cepstral Coefficient（以下、MFCC）とは主に人間の喋り声を対象に開発された特徴量で、音源から声道特徴量を強調して取り出し、音源を人間の知覚に特化した形で圧縮できるという特徴がある。通常比較的短い音源を対象に変換が行われ、その特徴量を12次元^{*5}の実数ベクトルで表すことができる。管楽器や弦楽器といった、所謂オーケストラ楽器の識別に用いられた事例が存在する [2], [3]。

3.1.2 Zero-crossing rate

Zero-crossing rate とは音源の振幅が、正から負、あるいは負から正に変化する頻度を記録し、特に音源に含まれるドラムのようなアタックの強い音源を識別する際に利用される [4]。また、Zero-crossing rate が信号のノイズの尺度として解釈できることから、音楽ジャンルの分類などにも利用される [5]。

3.1.3 振幅スペクトル

振幅スペクトルとは、様々な周波数の正弦波を含む音源に対し、周波数ごとに正弦波がどのような振幅で含まれているかを表したものである。2次元のグラフで表すことができるため、人間が特徴量として直感的に理解しやすいというメリットがある。たとえば同一音高の音色から振幅スペクトルを抽出し、直接比較することで、和楽器と西洋楽器の特徴を比較する研究などの実績がある [6]。

3.1.4 スペクトル重心

スペクトル重心は音色の明るさと強い相関があるといわれる音響指標値で、振幅スペクトルの加重平均として計算される。音色に含まれる周波数の概形を表す特徴量と考えることができ、音色の持つイメージを大まかに示すことができる。たとえば異なる響きを持つ楽器間の音色の補間について実績がある [7]。

3.1.5 メル・スペクトログラム

メル・スペクトログラムとは、音源に短時間フーリエ変換を行い、取り出したスペクトログラムに対してメルフィルタバンクを適用したものである。メルフィルタバンクとは人間の聴覚特性に準じた間隔で並べられたバンドパスフィルタで、波形をより人間の知覚に特化した形に整形することを目的としている。主に音声合成の分野で利用され、近年では Deep Neural Network との相性が MFCC よりも優れていることから、MFCC に代わって CNN^{*6}で広く用いられる [8]。

3.2 距離尺度の検討

また、音源の特徴量として録音データの特徴量ベクトルとして圧縮した後は、目標の録音データとの距離を適当に算出する必要がある。ユークリッド空間におけるベクトル間の距離の測り方としてはユークリッド距離があげられるが、ここで注意すべき点は、単純に時系列順にベクトル間のユークリッド距離を1対1で測ると、不適当な対応関係で距離を測定してしまう可能性がある点である。本研究が想定する使用シーンでは、目標となる録音と自分の所有するギターの録音は異なる環境で別々に録音されたものであるため、双方の録音データのサンプル箇所が一致していないという前提がある。もし音源にわずかでずれが生じた場合正しく距離を測れない手法であるならば、実際の使用シーンを考えると合理的ではない。そのため、サンプル箇所が一致していない状況でも録音データを正しく比較できる手法を選定する必要がある。

そこで本研究ではタイミングやピッチの揺れを許容する距離算出の手法として、以下の2つの方法を利用した。

3.2.1 Dynamic Time Warping

Dynamic Time Warping（以下、DTW）とは時系列データどうしの距離を測定する際に利用される手法である [9]。DTW では比較する2つの音源の各点に対し互いを対応づける最短の経路を見つけ、対応する点と点で値の比較を行うことで、タイミングのずれた波形でも適切に距離を測定することができる。また、DTW では対応する点の重複を認めるため、サンプル数が合わない波形どうしも比較することができる。ただし対応する点は時系列順に限られ、順番を入れ替えた対応関係は作られない。

^{*5} 音声認識の分野では通常12次元分を用いるが、楽曲の分類や医療の分野ではその限りではない。

^{*6} Convolutional Neural Network の略称。ディープラーニングの代表的手法で、畳み込みニューラルネットワークとも。

3.2.2 Eearth Mover's Distance

Eearth Mover's Distance (以下, EMD) とは Rubner らによって提唱された手法 [10] で, ある点からある点への輸送量という視点で集合間の距離を測ることができる. つまりベクトルに設定した重みを移動させるとき, その輸送コストが最小になる経路を見つけることで距離を求めるという手法である. その性質上, EMD は時系列順に縛られずに距離の測定を行うことができる点が DTW との大きな違いである.

3.3 エレキギター音色への適用と比較実験

3.3.1 目的

実際に上記 5 種類の音響特徴量がエレキギター音色に対してどの程度有効か検証するため, 実際の使用シーンを想定した音源を用意し, 実験を行った.

3.3.2 素材

実験に際し, 中空構造で音色に低音成分を多く含むセミアコースティックギター (以下, セミアコ) とソリッド構造で音色に高音成分を多く含むストラトキャスター (以下, ストラト), そして BOSS 社^{*7}の歪み系のギターエフェクタ^{*8}である DS-1 の 3 つの機材を用い, 同一のフレーズを演奏して下記の 5 つの音源を用意した.

- A セミアコ + DS-1
- B セミアコ (エフェクトなし)
- C セミアコ + DS-1 (異なる録音環境で弾き直す)
- D ストラト + DS-1
- E ストラト + DS-1 (フレーズをわざと間違える)

フレーズの演奏はエレキギターを 10 年間趣味で演奏する著者のうち 1 人が行い, BPM120 で 2 小節の単音フレーズを録音した. 録音は WindowsOS のコンピュータ, YAMAHA 社のオーディオ・インタフェースである「AG06」, フリーのオーディオ・エディタである「Audacity」を用い, ビットレートは 16 bit, サンプルレートは 44,100 Hz, モノラルで行った.

3.3.3 手法

これらの音源から, 3.1 節で述べた音響特徴量をそれぞれ抽出し, 目標とする音源との距離を 3.2 節で述べた比較手法を用いて適切に測れるか検証した.

尚, DTW と EMD によって算出される数値はいずれも距離であるが, ここでは比較を容易にするため, 各点のユークリッド距離をコストとし, 累積距離を 0.0 から 1.0 に正規化したものを類似度として定義する. 正規化の結果, 値が 1.0 に近いほど類似度が高い傾向にあることを示す. また, ここで言及される類似度の評価は相対的なものであり, 「似ていない」の類似度 < 「あまり似ていない」の類似度 <

「似ている」の類似度 < 「完全一致」の類似度 = 1.0 という関係にしたがう.

3.3.4 分析

手法で述べた「適切に」という指標だが, おおむね以下のように類似度が測定されたなら, 人間の知覚と一致した類似度であると設定した.

A : A 同一音源なので完全一致

A : B 似ていない (エフェクト有無の判定)

A : C 似ている (録音環境の差異の許容)

A : D あまり似ていない (使っているギターの違いの判別)

A : E A : D と同程度の結果 (フレーズ間違いの許容)

A : A の比較は同一の音源どうしの比較であるので, 類似度は完全に一致する. したがって A : A の比較は似ていると判定されることが望ましい.

A : B の比較は同一のエレキギターを用いてはいるが, 片方はエフェクトがかけられ, もう片方はかけられていない. 音色再現にあたって, エフェクトの程度を判定することは重要であるため, A : B の比較は似ていないと判定されることが望ましい.

A : C の比較は同一のエレキギターと同一の歪み系エフェクタを用いてはいるが, 異なる録音環境で録音された音源どうしを比較している. 音色再現にあたっては純粋な音色のみに注目するべきであり, フレーズやピッチ, タイミングのわずかな揺れは無視されるべきである. そのため A : C の比較は似ていると判定されることが望ましい.

A : D の比較は同一の歪み系エフェクタを用いてはいるが, 異なる特徴を持つエレキギターを用いて録音された音源どうしを比較している. 使用しているエフェクタは同一であるが, エレキギター自体の音響特性の違いを判定することは重要であるため, A : D の比較はあまり似ていないと判定されることが望ましい.

A : E の比較は, 使用機材は A : D と同じであるが, 異なるフレーズを演奏した音源どうしを比較している. A : C の比較でも述べたが, 音色再現にあたっては純粋な音色のみに注目するべきであり, A : E の判定は A : D の比較と同程度の結果になることが望ましい.

3.3.5 結果

実際に DTW を用いた実験結果を図 2 に, EMD を用いた結果を図 3 にそれぞれ示す.

図 2 より, DTW を用いた比較では, どの特徴量も A : C の類似度が最も高くなっており, 知覚と一致して最も似た音色を示している. ただし, Zero-crossing rate はすべての比較でほとんど差がなく, 明確に違いを示せてはいない.

A : D, および A : E を比較した類似度が同程度になっているかについては, それぞれの特徴量ごとに算出された類似度の最大値を 1, 最低値を 0 として正規化した場合, スペクトル重心は類似度の差が 0.009 ポイント, MFCC は

^{*7} <https://www.boss.info/jp/>

^{*8} 波形にクリッピングを行うことで音を歪ませ, 信号に倍音成分を含ませる処理を行うエフェクタ.

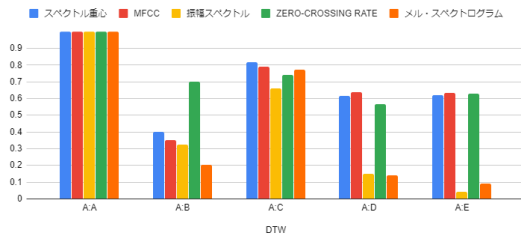


図 2 DTW を用いて特徴量の類似度を比較した図

Fig. 2 Comparison of similarity of feature values using DTW.

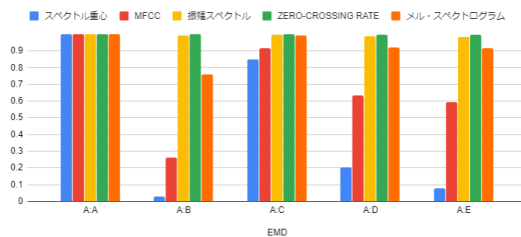


図 3 EMD を用いて特徴量の類似度を比較した図

Fig. 3 Comparison of similarity of feature values using EMD.

0.011 ポイントであり、スペクトル重心、MFCC の両方で、類似度の差が非常に小さかった。

EMD を用いた比較でも、どの特徴量も A : C の類似度が高いと示しているが、こちらは振幅スペクトルと Zero-crossing rate はすべての音源に対してほとんど差を出せていない。

A : D, および A : E を比較した類似度が同程度になっているかについては、同様にそれぞれの特徴量ごとに算出された類似度の最大値を 1, 最低値を 0 として正規化した場合、MFCC は類似度の差が 0.055 ポイント、メル・スペクトログラムが 0.019 ポイントであり、MFCC, メル・スペクトログラムの両方で、類似度の差が非常に小さかった。ただし、メル・スペクトログラムは A : D, A : E ともに高い類似度を示しており、これは知覚と一致しない。

3.3.6 考察

MFCC は 3.1.1 項でも触れたとおり、すでに管楽器や弦楽器の分類で実績があり、エレキギターでも同様に有効であるという結果を示すことができたが、同様に楽器分類で実績のある Zero-crossing rate はエレキギター音色の識別ではほとんど差がでなかった。これは歪み系エフェクタによる波形の変化が振幅方向に作用するため、「フルートとギター」など、異なる周波数特性を持つ楽器間の分類と比べて差が出にくかったものと思われる。

3.4 音響特徴量と距離尺度の決定

3.4.1 音響特徴量のサイズ

3.3 節の実験結果を元に、本研究ではエレキギター音色の距離を測定する音響特徴量として MFCC が適当だと判断した。計算時間とエレキギターの周波数特性を考慮し、サンプリングレート 44,100 Hz の音源データを 4,096 サン

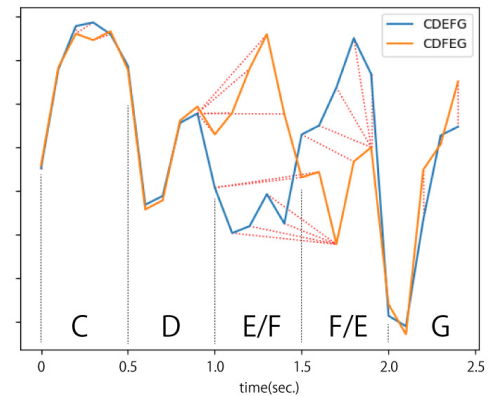


図 4 DTW によるフレーズを間違えた音源の比較経路

Fig. 4 Comparison path of the sound source with the wrong phrase by DTW.

プルずつ (約 0.1 秒分に相当する) のグループに分け、入力音源の中心の 163,840 データ (約 4 秒) 分で MFCC の計算を行った。つまり 1 つの音源データから MFCC として 12 次元のベクトルが 40 本抽出される。

3.4.2 DTW と EMD の比較

目標の録音データとの距離を測定する手法として、3.3 節の実験では MFCC と組み合わせることにより、DTW・EMD 共に距離を適切に測定することができた。

ただし実験の過程で、DTW は時系列順に縛られるという性質から、フレーズの間違いに対して適切に対応できないことが分かった。図 4 は、「ド・レ・ミ・ファ・ソ」と弾いた音源と、同じタイミングで「ド・レ・ファ・ミ・ソ」と、音の順番を間違えて弾いた音源から MFCC を抽出し、DTW によってシグネチャ間の距離を測定した際の経路を疑似的に 2 次元にプロットした図である。

図を見ると分かりますとおり、DTW は順に縛られるという性質と重複を許すという性質から、フレーズを間違えた箇所では特定の 1 点に対して、特徴が異なる複数の点が対応させられてしまっている。このことからタイミングの揺れに関しては DTW は対応可能であるが、ピッチの揺れに対しては弱いことが分かる。

対して図 5 は EMD によってシグネチャ間の距離を測定した際の経路を 2 次元にプロットした図である。前の音の残響成分などに影響を受け、完全に一致はしていないが、EMD では時系列順に縛られないこともあり、25 個のシグネチャの内 23 個 (内 2 個は重みの分散あり) のシグネチャ間の距離を正しい組合せで比較することに成功している。このことから、フレーズに含まれる音の順番が前後した際も、おおむね正しく同一の音高どうしを比較していることが分かる。結果を受け本研究では、DTW より柔軟に音源データを比較できる手法として EMD を選択した。

4. パラメータ自動推定の実験

設定した距離尺度が適切であるか検証するため、実際の

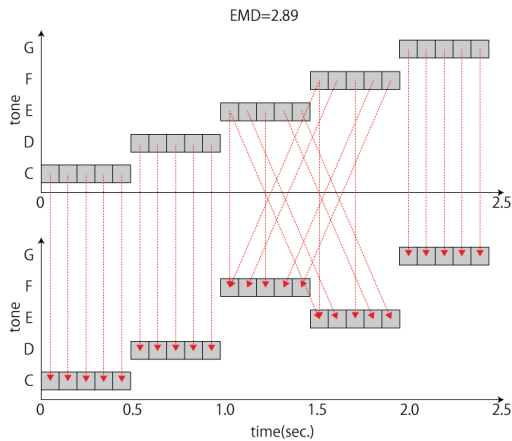


図 5 DTW によるフレーズを間違えた音源の比較経路

Fig. 5 Comparison path of the sound source with the wrong phrase by EMD.

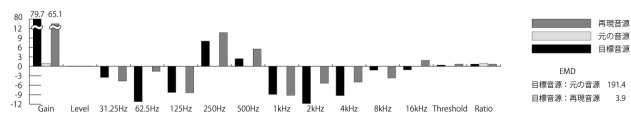


図 6 同一のギターを用いた自動再現の結果

Fig. 6 Results of automatic reproduction using the same guitar.

使用シーンを想定した状況で 3 種類の実験を行った。

4.1 予備実験：同一のギターを用いた再現実験

まず 3.3 節で使用したセミアコを用いて、同一のギターを使用した場合に正しくパラメータの推定が行えるか予備実験を行った。1 世代の個体数を 100 個体に設定し、100 世代まで交配を行った。結果を図 6 に示す。

図 6 における「目標音源」は 2 節で実装した C++ による歪み系エフェクトを適用した音源、「元の音源」はエフェクトを適用する前の音源、「再現音源」は自動推定を行い 100 世代目で最も距離が近いと判定された音源である。EMD による測定では「元の音源」と「目標音源」の距離は 191.4 と計測された。これに遺伝的アルゴリズムによって自動的に推定されたパラメータを適用することにより、距離が 187.5 縮まり、距離が縮まることに対応して求められたエフェクトの各パラメータも「目標音源」と近いものが推定できていることが確認できた。

4.2 実験 1：異なる特性を持つギターでの再現実験

続いて 3.3 節で使用したセミアコとストラトの両方を用いて、異なる特徴を持つエレキギター間でも再現が可能であるか実験を行った。

1 世代の個体数を 100 個体に設定し、100 世代まで交配を行った。結果を図 7 に示す。

図 7 における「目標音源」はストラトのクリーン音源、「元の音源」はセミアコのクリーン音源、「再現音源」は自動

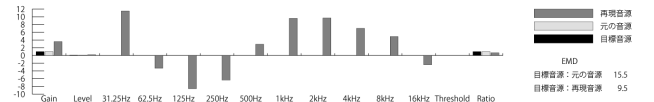


図 7 異なる特徴を持つギターを用いた自動再現の結果

Fig. 7 Results of automatic reproduction using guitars with different characteristics.

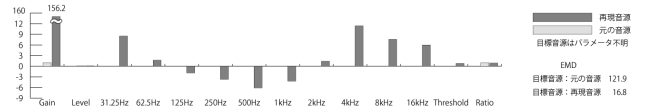


図 8 DS-1 を目標とした自動再現

Fig. 8 Automatic reproduction of sound with DS-1 as target.

推定を行い 100 世代目で最も距離が近いと判定された音源である。図 7 を見ると、再現音源では高音に特徴のあるストラトの音色に合わせ、高音成分の弱いセミアコの 1kHz から 4kHz が強調されるようにイコライザが適用され、距離が縮まっていることが見て取れる。

この実験により、遺伝的アルゴリズムによるパラメータ推定が、異なる特徴を持つギター間の音色の再現においても一定の効果を持つことが確認された。

4.3 実験 2：市販のエフェクタを用いた再現実験

実際に市販され、広く使用されている歪み系エフェクタ 5 種類を目標として、開発した自動再現システムを用いて音色の再現が可能かの実験を行った。なお、エフェクタのパラメータはすべてのつまみを 12 時とした。実験で扱う音源には、「目標音源」、「元の音源」とともにセミアコを用いた。

すべての実験において 1 世代の個体数は 100 個体、100 世代まで交配を行った。

4.3.1 BOSS : DS-1

BOSS 社のエフェクタ「DS-1」*9 を目標とし、再現を行った。結果を図 8 に示す。

DS-1 は、高音成分を強調する歪み系エフェクタでコンプカットの回路も含まれていない。再現音源として、2kHz から 16kHz が強調され、スレッシュホルドの値が高いパラメータが選ばれた。これは DS-1 の傾向と一致している [22]。

4.3.2 BOSS : MT-2

BOSS 社のエフェクタ「MT-2」*10 を目標とし、再現を行った。結果を図 9 に示す。

MT-2 の特徴として、「LOW」、「MIDDLE」、「HIGH」の 3 バンドイコライザを備えていることがあげられる。今回の実験ではすべてのつまみを 12 時に設定したため、イコライザもフラットの状態であるが、再現されたパラメータではイコライザの値はフラットにならなかった。しかし生成された音響の距離は近かった。これは元々 MT-2 の実機

*9 <https://www.boss.info/jp/products/ds-1/>

*10 <https://www.boss.info/jp/products/mt-2/>

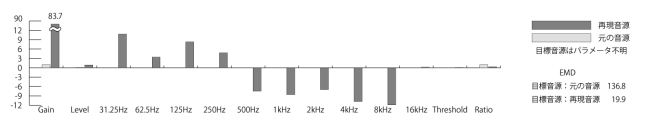


図 9 MT-2 を目標とした自動再現

Fig. 9 Automatic reproduction of sound with MT-2 as target.

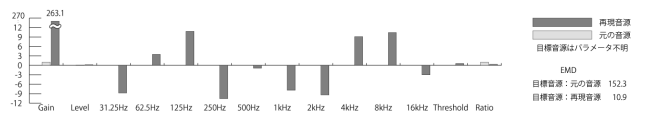


図 10 Big Muff を目標とした自動再現

Fig. 10 Automatic reproduction of sound with Big Muff as target.

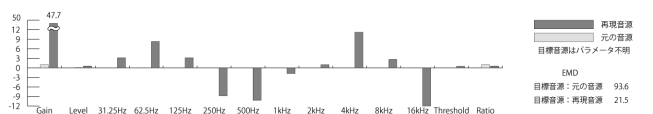


図 11 RAT を目標とした自動再現

Fig. 11 Automatic reproduction of sound with RAT as target.

は、イコライザがフラットの状態であっても中音域を中心として 31.25 Hz から 250 Hz をブーストし、500 Hz から 8 kHz をカットするという傾向があり、それを模倣できたと考えられる。

4.3.3 ELECTRO-HARMONIX : Big Muff

ELECTRO-HARMONIX 社のエフェクタ「Big Muff」^{*11}を目標とし、再現を行った。結果を図 10 に示す。

この再現では低音域と高音域が強調され、250 kHz から 2 kHz の中音域がカットされるという結果が得られた。一般的にこのような音作りは「ドンシャリ」と呼ばれ、Big Muff の特徴と一致する。また、Big Muff は歪み系エフェクタのなかでも比較的歪みが強い「ファズ」と呼ばれる種類であり、再現音源でも Gain の値が大きいものが選ばれている。

4.3.4 Pro Co : RAT

Pro Co 社のエフェクタ「RAT」^{*12}を目標とし、再現を行った。結果を図 11 に示す。

この再現でも RAT の音響特性に近いパラメータセットが選ばれたが、Big Muff の再現と異なり、16 kHz 帯が大きくカットされている。16 kHz 帯はプレゼンスと呼ばれ、音の余韻や繊細さに影響があるとされる。

図 12 は同一の音源にそれぞれ Big Muff と RAT のエフェクタを適用した波形のパワースペクトル示したものである。Big Muff を適用した波形とは異なり、RAT を適用した波形のパワースペクトルは 16 kHz 帯以降の周波数では微弱であり、このことから RAT の特性に近いパラメータセットが選ばれていることが分かる。

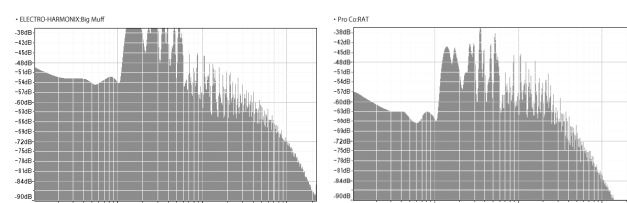


図 12 Big Muff と RAT のパワースペクトルの比較

Fig. 12 Comparison of the power spectrum between Big Muff and RAT.

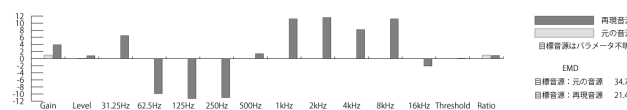


図 13 TS808 を目標とした自動再現

Fig. 13 Automatic reproduction of sound with TS808 as target.

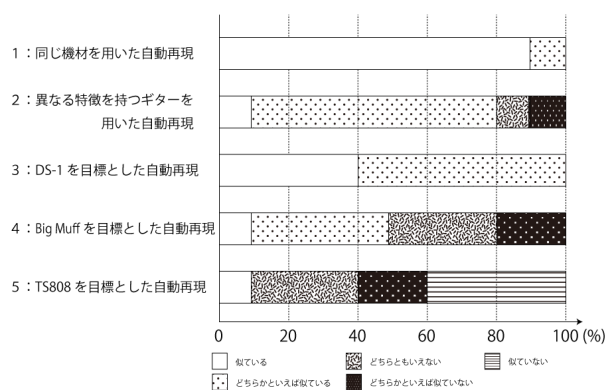


図 14 音色再現度についてのアンケート結果

Fig. 14 Results of questionnaire on tone reproduction.

4.3.5 IBANEZ : TS808

IBANEZ のエフェクタ「TS808」^{*13}を目標とし、再現を行った。結果を図 13 に示す。

この再現では中音域を中心として 62.5 Hz から 250 Hz をカットし、1 kHz から 8 kHz をブーストするような傾向が見られた。TS808 の備えるローパスフィルタのカットオフ周波数は 719.7 Hz であり、再現されたパラメータは実機よりも低いカットオフ周波数であることを示している。また Gain の値が 3.9 と僅かであるが、TS808 は増幅率が低い設計のため、こちらは TS808 の傾向と一致している。

4.4 再現結果の主観評価

実験の結果について主観品質評価を行うため、複数人にアンケートでの聞き取り調査を行ったところ、10 代から 40 代までの 10 名から回答を得た。アンケートの結果を図 14 に示す。

アンケートは対面で行い、同一のノート PC とヘッドフォンを用いて行った。アンケートには 4.1 節、4.2 節の実

^{*11} <https://kcmusic.jp/ehx/big-muff-pi.html>

^{*12} <https://www.ratdistortion.com/>

^{*13} https://www.ibanez.com/jp/products/detail/ts808_99.html

験結果に、4.3 節の実験結果から 3 つを抜き出した合計 5 つの実験結果の音源をもとにあらかじめ作成した動画を使用した。動画では各実験結果ごとに「再現音源」と「目標音源」の音源が交互に 1 フレーズずつ再生され、これを 2 回繰り返したところで一度動画の再生を止め、その間に回答者はアンケートに記入を行う。記入が完了したのち、動画を次の実験結果へ進ませる。評価は「似ている」、「似ていない」、「どちらともいえない」、「どちらかといえば似ていない」、「似ていない」の 5 段階を指定した。

5. 考察

4 章の実験によって、すべての機種について多少の差はあるものの、開発した手法で距離を縮め、同一機材、同一設定を用いずに距離を縮めることができることを示した。EMD を用いることで、フレーズ演奏を入力とすることによって生じる発音タイミングの差や演奏ミスなどの、音色そのものではない違いの影響を最小限にとどめることができた。これは本研究の独自性であり、関連研究の項で述べた他の音色探索の研究に対して優位な点である。

しかし、図 14 を見ると、比較的歪みの量が少ない「TS808」について、60%の人が似ていない、あるいはどちらかといえば似ていないと回答していることが分かる。このほかにも歪みの量が少ない場合に再現が不安定になるケースが確認できており、結果を受けて特徴量について見直しを行った。図 15 はクリーンの音源と歪み系エフェクトを掛けた音源の 2 つについて、抽出した MFCC の 20 次元分を比較したものである。

音声認識の分野では通常抽出したケプストラムのうちの 1 次項から 12 次項までの値を特徴量として用いるため、本研究でもそれに従い同じ条件で比較した。図 15 から MFCC の値は 13 次項以降の数値は双方ほとんど変動もなく、ギターの音源においても 13 次項以降は無視できると著者らは考えた。

一方で一番左の 0 次項は大きく変動している。0 次項は「各周波数帯域の平均エネルギーの集合」と相関があるとされ、パワー係数とも呼ばれる。しかし通常は信頼性の低さから通常は除外される。このパワー係数については音声認識の分野では様々な議論があり、Zheng らによる研究では、パワー係数を考慮することで中国語の音節の誤りを約 10% 低減できるとしている [15]。

また、Ismail らによる研究ではパワー係数に注目することで、心臓の異常を発見することができるとしている [16]。

ギターの音色についても歪みの量が多いほど音圧が上昇し、この音圧と「各周波数帯域の平均エネルギーの集合」には強い相関があることが分かった。そこで実験 3 の「TS808」について、MFCC の 0 次項も加えた状態で再現を行った。その結果を図 16 に示す。

比較するベクトルの次元が増えたため、0 次項を考慮する

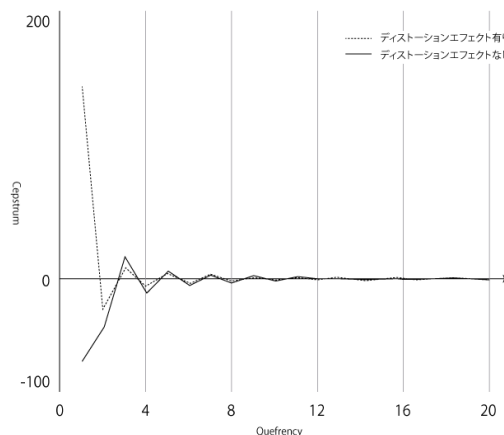


図 15 20 次元分の MFCC の比較

Fig. 15 20 Dimensional MFCC Comparison.

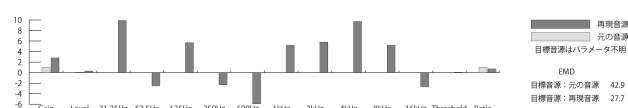


図 16 TS808 を目標とした 0 次項を加えた自動再現

Fig. 16 Automatic reproduction of sound considering 0-order term with TS808 as target.

前と比べ、距離は相対的に離れており、パラメータにも差が出ている。0 次項を考慮した音源について再び聞き取り調査を行ったところ、0 次項を加えていないものと比べ、やや再現度が高まったとの回答が多かった。このことから 0 次項に関して、エレキギター音色の再現において有用である可能性が示された。MFCC の 0 次項の取り扱いについては実験の条件を変え、今後さらなる検証を行う予定である。

6. アプリケーション

Steinberg 社^{*14}の提供する SDK である「VST SDK3.6」を元にコーディングを行った。「VST SDK3.6」は C++ で記述されるため、4 章までに開発したシステムをほぼそのまま統合し、DAW からシームレスに利用可能とした。

大きな違いは入出力デバイスの管理やデータサイズの設定をホストとなるアプリケーションに依存することである。これまでは wav データのサイズ分のバッファをとり、すべての処理を終えた段階で音声を出力していたが、DAW プラグインとして実装することで、音声は最低限の作業バッファを通過する過程で処理され、入力とほぼ同時に出力される。これによってリアルタイムでのエフェクトの適用を実現している。先読みなどの処理が必要なエフェクトの場合、この入力の処理を変更する必要があるが、現段階では 4 章で実装したものと同一の処理で問題はない。

また、「VST SDK3.6」にはデフォルトで GUI^{*15}の機能

^{*14} ドイツのソフトウェア会社、オーディオドライバに関する規格の ASIO や、音楽ソフトウェアに関するプラグイン規格の VST を策定している。

^{*15} Graphical User Interface の略称。

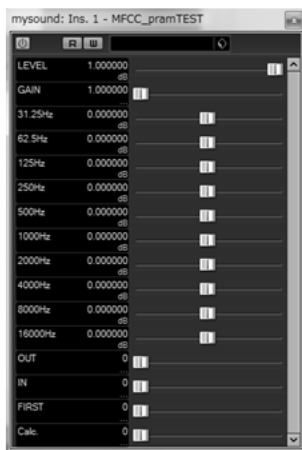


図 17 DAW のプラグインとして実装された本システムの GUI
Fig. 17 GUI of our system implemented as a DAW plug-in.

が備わっている。移植にともない、推定後のパラメータをこの GUI で微調整できるようにした。実装した GUI を用いることで、DAW 上で計算に必要な音源の書き出しから、染色体・個体・世代数の設定をシームレスに行うことができる。実装した GUI のイメージを図 17 に示す。

7. おわりに

本研究では「目標とする音源」と「自身の所有するギターの音源」の 2 種類の音源から MFCC を抽出し、EMD を用いてエフェクトの掛けられたエレキギターの音色間の距離を測定する手法を提案した。また、遺伝的アルゴリズムを手法として選択することで、音色の再現に必要な要素を統合的なパラメータとして設定し、目標とするパラメータの自動推定を行うシステムの開発を通して新たな再現手法を提案した。

現在は歪み系エフェクトを中心としたエフェクトのみを実装しているため再現には限界があるが、今後扱うエフェクトを増やすことでより正確な再現が可能になると考える。ギターの音色という物理的でありながら個人の印象に左右されるものを再現するにあたり、本研究で提案した手法は、組み合わせの多様化や環境などの制約にとらわれない新たな音色再現の手法として提案できると考える。

参考文献

- [1] The History Of EDM, JMC Academy, available from <https://www.jmcacademy.edu.au/news/the-history-of-edm> (accessed 2015-11-21).
- [2] 馬場貴之, 山田武志, 北脇信彦: HMM と MFCC を用いた楽器音の音源同定の検討, 情報処理学会研究報告, Vol.2005, No.82, pp.79–84 (2005).
- [3] 北原聡志, 甲藤二郎: 楽器の階層的分類を考慮した SVM による音源同定, 電子情報通信学会総合大会講演論文集, Vol.2006, p.152 (2006).
- [4] Gouyon, F., Pachet, F. and Delerue, O.: *On the Use of Zero-Crossing Rate for an Application of Classification of Percussive Sounds*, Digital Audio Effects Conference

- (DAFx) (2000).
- [5] Kulkarni, N. and Bairagi, V.: *EEG-Based Diagnosis of Alzheimer Disease*, pp.47–59, Academic Press (2018).
- [6] 徳島達也: 音響解析による和楽器の特徴について, 2010 年度日本海学グループ研究支援事業, 入手先 (<http://www.nihonkaigaku.org/library/group/tokushimadata.pdf>) (参照 2018-12-25).
- [7] 引地孝文, 小坂直敏: スペクトルセントロイドを利用したピアノ音とギター音の音色補間, 日本音響学会平成 11 年度秋季研究発表会講演論文集, Vol.1999, No.2, pp.547–548 (1999).
- [8] Shen, J., Pang, R.M., Weiss, R.J., et al.: Natural TTS Synthesis by Conditioning Wavenet on MEL Spectrogram Predictions, *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2018).
- [9] Müller, M.: *Information Retrieval for Music and Motion*, pp.69–84, Springer (2007).
- [10] Rubner, Y., Tomasi, C. and Guibas, L.J.: A Metric for Distributions with Applications to Image Databases, *Proc. 1998 IEEE International Conference on Computer Vision*, pp.59–66 (1998).
- [11] 青木直史: C 言語ではじめる音のプログラミング, オーム社 (2008).
- [12] 伊庭斉志: C による探索プログラミング 基礎から遺伝的アルゴリズムまで, オーム社 (2008).
- [13] 渡辺勝正事, 池田宜明, 松尾聡ほか: 遺伝的アルゴリズムの改良とその応用, 福井大学工学部研究報告, Vol.40, No.1 (1992) (オンライン), 入手先 (<https://core.ac.uk/download/pdf/59038313.pdf>).
- [14] 大向一輝: 分散遺伝的アルゴリズムにおける最適な交叉率および突然変異率 (オンライン), 入手先 (http://mikilab.doshisha.ac.jp/dia/monthly/monthly99/19990830/ohmukai_21.pdf) (参照 2018-12-25).
- [15] Zheng, F., Zheng, G. and Song, Z.: Comparison of Different Implementations of Mfcc, *J. Computer Science and Technology*, Vol.16, pp.582–589 (2001).
- [16] Ismail, K., Salleh, S.H., Najeb, J.M., et al.: Heart Sound Analysis Using MFCC and Time Frequency Distribution, *3rd Kuala Lumpur International Conference on Biomedical Engineering* (2006).
- [17] Barkan, O. and Tsiris, D.: Deep Synthesizer Parameter Estimation, *Proc. ICASSP* (2019).
- [18] Su, A.W.Y. and Liang, S.F.: A Class of Physical Modeling Recurrent Networks for Analysis/Synthesis of Plucked String Instruments, *IEEE Trans. Neural Networks*, Vol.13, No.5, pp.1137–1148 (2002).
- [19] Itoyama, K. and Okuno, H.G.: Parameter estimation of virtual musical instrument synthesizers, *Proc. Joint ICMC and SMC 2014*, pp.1426–1431 (2014).
- [20] 瀧田寿明, 櫛部義之, 星野准一ほか: 深層学習による入力音響信号からの MIDI 音色パラメータ推定, FIT2014, Vol.40, No.1 (1992) (オンライン), 入手先 (<https://core.ac.uk/download/pdf/59038313.pdf>).
- [21] 松谷 佑: 異なる特性を有するギターの音質再現, 高知工科大学情報システム工学科学学位士論文, pp.1–31 (2001) (オンライン), 入手先 (<http://www.lab.kochi-tech.ac.jp/fukulab/research/Doc/2010/1110286.pdf>).
- [22] Hal Leonard Publishing Corporation (Eds.): *The Boss Book: The Ultimate Guide to the World's Most Popular Compact Effects for Guitar*, Hal Leonard Corp (2002).



有山 大地

東京都立大学大学院（旧，首都大学東京）システムデザイン研究科修了．現在，同大学院博士課程在籍．修士（芸術工学）．修士課程在学中より一般社団法人国際 STEM 学習協会，FabLab Kamakura にて講師として活動．中学・高校生を対象とした STEAM 教育や，工作機械を用いた楽器制作を实践．エレキギターをワークフローの中心とした DTM 制作を支援するデバイスの開発，ディープラーニングを用いたフレーズ提案ソフトウェアの研究を行っている．



安藤 大地

国立音楽大学音楽学部声楽学科を経て音楽デザイン学科卒業．Sweden, Chalmers University of Technology より MSc. 授与．東京大学大学院新領域創成科学研究科基盤情報学専攻修了．博士（科学）．現在，東京都立大学学術情報基盤センター准教授．対話型進化システムや人工生命，群知能等の音楽創作・分析への応用を中心とした研究，音楽認知学の知見を応用した即興演奏等の研究・演奏実践を行っている．



串山 久美子

筑波大学大学院芸術研究科修了．メディア制作業務の株式会社グロースを起業後，JST さきがけ研究員，東京農工大学客員教授を経て 2006 年より東京都立大学（旧，首都大学東京）システムデザイン研究科教授に就任．博士（デザイン学）．日本バーチャルリアリティ学会理事，国際対抗 VR コンテスト (IVRC) 審査委員，情報処理学会 DCON 編集委員．専門はインタラクティブアート，バーチャルリアリティ，触視覚ディスプレイの開発と研究．2000 年より ACM.SIGGRAPH Art Gallery, Emerging Technology や Ars Electronica に数多く出展．