

パーソナライズド空間上での音データの関係性可視化による 音色感覚の訓練

永田 大志^{1,a)} 山西 良典^{1,b)} 辻野 雄大^{2,c)}

概要: 本稿では、音色に対する印象と物理特徴量の関係性 (i.e. 音色感覚) の自己認識を明確化するためのアプローチを提案する。音源選択や音作りの過程で生まれる音色は編集過程で忘れやすく、音楽初学者にとって音色感覚を身につけることは難しい。これは、音色が時間的メディアであり、一過性の性質を持つことに起因する問題である。そこで、時間的メディアである音データを物理特徴量に基づいて空間的に配置し、音色の持続的な認知を可能にすることで問題の解決を図る。提案手法では、物理特徴量として取得した MFCC を t -SNE 法によって次元圧縮して 2 次元空間に各音データを配置する。この可視化結果がユーザごとの音色感覚に適合するように t -SNE 法のメタパラメータを対話的に調整する。これにより、複数の音データの相対的な関係性を視覚的に確認可能になり、音色感覚の訓練ができると考えた。本稿では、個々の音色の定義が曖昧な音源としてエフェクターを用いたエレキギターの音データを対象とする。発話思考法に基づく実験を通して、提案手法の適用によりユーザの音色感覚が訓練されたかについて議論した。

1. はじめに

演奏や作曲などの音楽活動では、音色を編集する場面が多く存在する。音色の編集は、使用する楽器や機材の組み合わせの変更や、楽器・機材のパラメータの設定によって実現される。複数の機材の組み合わせやパラメータの設定を試行していくため、音色の編集は、音色に対する印象と物理特徴量との関係性 (i.e., **音色感覚**) が明確でない初心者には困難な作業である。演奏者の中には、演奏自体ではなく、音色の編集の困難さを理由として音楽活動から遠ざかってしまう人もいる。音色の編集に必要な音色感覚を身につけることは、演奏の技術を習得することと同等以上に困難な課題であると考えられる。

イラストのような空間的メディアとは異なり、音データは時間的メディアであるため、一過性の性質を持つ。それに加えて、音色の編集過程では機材の組み合わせやパラメータの設定を何度も繰り返して変更するため、音色とパラメータ情報もまた一過性の性質を持つと言える。本稿では、音色感覚の習得に対して、時間的メディアである音データを物理特徴量に基づいて空間的に配置することで、

本来一過性の性質を持つメディアを持続的に把握可能にするアプローチを提案する。音色編集の繰り返しの過程の中で生まれた音色サンプルを空間上に配置して可視化することで、一過性の性質を原因とした音色感覚習得の難しさの軽減をねらう。可視化された空間を確認しながら、ユーザの目標とする音色間関係性に適合するようにパラメータを対話的に調整することで、複数の音データの相対的な関係性を視覚的に確認することができ音色感覚を訓練できると考えた。

本稿では、「音色の空間的な配置によって、音色感覚の訓練は可能か?」という仮説に対して、実験用システムを開発し、ユーザ実験を通して仮説の検証を行う。空間上への可視化においては、物理特徴量を反映しつつもユーザそれぞれにとって納得性の高い可視化が得られる必要があると考えた。そこで、高次元の音響特徴をメタパラメータに応じて可視化に適した低次元情報に圧縮し、音データ間の相対的な関係性を 2 次元空間上に可視化する。このとき、ユーザは情報圧縮時のメタパラメータを対話的に操作 (チューニング) することで自身の音色感覚に適合した可視化空間を得る。ユーザごとにチューニングされた空間に音色編集の過程で生まれる音サンプルをユーザが逐次的に追加可能にすることで、音色感覚の訓練をねらう。訓練対象としては、個々の音色の定義が曖昧なエフェクターを用いたエレキギターのモノフォニックな音源を対象とし、音色の一つの要素である「歪み」の強弱についての音色感覚を扱った。

¹ 関西大学
Kansai University, Takatsuki, Osaka 569-1095, Japan

² 立命館大学
Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

a) k447677@kansai-u.ac.jp

b) ryama@kansai-u.ac.jp

c) yudait@fc.ritsumei.ac.jp

2. 関連研究

本研究は、一過性の性質を持つコンテンツに対しての感覚獲得を支援する研究と位置づけられる。デジタル描画における消し跡機能のデザイン・実装を試みる研究 [1] では、描画における試行錯誤の過程で生まれる消し跡をデジタルな環境でもアナログのように持続的に可視化して描画支援に利用している。文書作成シーンで生まれた棄却テキストの活用を検討する研究 [2] では、棄却テキストを提示することで文書作成の参考資料として活用している。本研究で扱う音色に対しても、空間上での持続的な可視化は音サンプル同士の差異や類似性を把握するうえでの参考資料として活用可能であると考えた。

データの分類や可視化を行って、コンテンツに関する人間の認知を支援する先行研究もいくつか報告されている。山西らの研究 [3] では、演技音声の特徴量と感情やコンテキストの関係性を考察するため、*t*-SNE 法を用いて演技音声のデータを可視化することで、特徴量に基づいたデータの集合が形成され、データの感情やコンテキストごとのまとまりや距離が可視化された。高橋らの研究 [4] では、楽曲中の楽器編成に基づいた楽曲検索を行うことを目的とし、楽曲内の楽器編成の情報を円グラフで表示し、*t*-SNE 法を用いて似た楽器編成の楽曲同士を近傍に配置した。戸本らの研究 [5] では、オノマトペの関係性を把握可能にすることを目的とし、オノマトペの音素特徴をニューラルネットワークで次元圧縮を行い、2次元上に可視化することでオノマトペ間の関係性を視覚的に確認できるシステムが提案された。本研究では、音データを持続的に視認可能にすることがユーザの音色感覚の獲得に有効であるかを確認することを目的としているため、ユーザへのデータの見せ方が重要な観点となる。そこで、データ間の相対的な関係性を低次元空間上に可視化するために、音響特徴量に対して *t*-SNE 法による次元圧縮と可視化を試みる。

音楽コンテンツを対象とした研究では、様々な楽器を対象とした練習支援システムが提案されている [6], [7], [8], [9]。練習支援では初学者を対象として、運指 [10] やリズム [11] をはじめとする演奏に必要なスキルを扱っていることが多い。本研究で対象とする音色編集は、正確性が求められる演奏スキルとは異なり、ユーザ自身の感覚に従って音色を自由に作成できるようになる「感覚」の獲得が重要であると考えられる。本稿では、ユーザの感覚に基づいて音色感覚を学び、その感覚をシステム利用後の音楽活動においても活用可能になるような訓練を提供可能にすることを旨とする。

3. 音作りの難しさ

作曲や演奏のために出力する楽器の音色を編集することを「音作り」という。音作りは、使用する楽器の音色に対して様々な種類のエフェクターを組み合わせたり、パラメー

Aさんはエレキギタープレイヤーである。「THE ORAL CIGARETTES」のコピーバンドを披露するために、マルチエフェクターを用いて音作りを行う。

～Step1～

Aさんは、オリジナル楽曲っぽい音作りをすることを目標に決め、歪み系のエフェクター（約20種類）から探し始めた。まずは、歪み系の中でも最初に表示される「TS Drive」というノーマルなエフェクターの初期設定から試行した。その音を聞いたAさんは、歪み具合が足りないと感じた。

～Step2～

歪み具合を変化させるつまみ（Gain）の数値を上げて再度試行した。その際に、アンプリファイア（アンプ）のつまみの調節も行った。Aさんは、それでも歪み具合が足りないと感じ、このエフェクターでは目標とするような音を作れないと考えた。

～Step3～

次に、普段からAさんが好んで使用している「Metal World」というエフェクターを選択した。Step1～2と同様に、Gainとアンプのつまみを変化させて数回試行した。しかし、「歪みが強すぎる。曲の雰囲気には合わない。」と感じ、違うエフェクターを試すことにした。

～Step4～

1つ目のエフェクターより歪みが弱すぎず、2つ目のエフェクターよりも歪みが強すぎない種類のエフェクターを探した。「Red Crunch」というエフェクターを試してみることにした。今までと同様に、初期設定からGainとアンプのつまみを変化させて試行した。「1つ目のエフェクターと2つ目のエフェクターの音やパラメータは、明確に覚えていないけど、今回のエフェクターが一番目標に近い音作りができると感じた」と納得し、そのエフェクターに決めた。

図1 エレキギターとエフェクターを用いて音作りをする様子の一例

タを変更したりと、無数の組み合わせの中から目標とする音色に近づけていく作業である。図1に、音作りが困難なエレキギターとエフェクターを用いた音作り [12] の一例を示す。この例では、演奏者が歪み系エフェクターの選択とパラメータの設定を決める際の様子を示されており、音作りの設定が変更されるごとにStep分けをしている。

エレキギターとエフェクターを用いた音作りは、以下の2点において音楽初学者にとって特に難しい問題であると考えられる。

(1) エフェクターに対する曖昧な定義

エフェクターは、個々のサウンドの特徴・回路の種別などの定義が明確ではなく、エフェクターのメーカーや機種間で曖昧である。演奏者は、定義が曖昧なエフェクターを使用して音作りをしなければならない。例えば、エフェクターには歪み系という役割が存在し、一般的には歪みが強い順に Fuzz, Distortion, Over Drive, Crunch という名前がつけられている。しかし、エフェクターの中には Over Drive という名前であるにも関わらず、Distortion よりも強い歪みを生み出す機種も存在する。標準的なパラメータ設定では歪みの強弱関係が通説どおりであっても、パラメータの設定を変更することで歪みの強弱関係が逆転するケースも存在する。そのため、音作りは多大な時間を消費し、エフェクターの機種とそのパラメータ設定の組み合わせを試行し続ける作業である。

(2) 一過性の性質

音作りをする中で、音データは一過性の性質を持つ。

音データは、時間的メディアであるため、ユーザは目標とする音色に近づけるために試行錯誤の中で作成した音情報を記憶し続けなければならない。このとき、音色とパラメータ情報を紐づけて記憶しなければ、以前作った音色を再現することが難しくなってしまう。音作りの過程で生まれた複数の音データを一覧して比較することはできず、試行錯誤の過程で処理すべき情報は多い。このようなパラメータを調整する中で自身の望む音色を明確化する行為は、逐次的に生まれは消える複数の音データ間の関係性をモデル化しなくてはならないため、音作りの経験が十分でない音楽初学者には難しい。

本稿では、これらの問題の中でも特に、(2)の問題にアプローチすることで、音作りの難しさの解決に取り組む。

4. 提案手法

4.1 デザイン指針

演奏者にとって音作りでは、オリジナル曲や参考動画などの音色の再現そのものだけでなく、演奏者が納得し音色を作成可能になることが重要と考えられる。演奏者自身が音作りの試行錯誤の中で意思決定できるように支援するためには、ユーザごとの音色感覚を訓練可能なしくみが必要と考えた。

図1のモデルケースを考察して提案手法のデザイン指針を定める。まず、Step1で目標とする音色を定め、ノーマルなエフェクターから試行し、音色の歪みが足りないと感じた。次に、Step2では、同じエフェクターのパラメータを変更し、音色の歪みを強くすることを試みた。それでも音色の歪みが足りていないと感じ、Step3で普段から好んで使用しているエフェクターを試行した。しかし、歪みが強すぎると感じたため、Step4で他のエフェクターを試行した。1種類目のエフェクターと2種類目のエフェクターの音色を思い出して比較し、目標に近いと感じたエフェクターを選択した。結果的にAさんは納得してエフェクターを決定したが、「音やパラメータは、明確に覚えていない」という発言から、音色の編集過程での音データは、一過性の性質を持つことがわかる。試行回数が増えるほど、どのような音色であったのかを記憶することは難しくなる。さらに、編集過程でパラメータを変更し続けるため、パラメータの設定と紐づけて記憶することも困難であり、音データを比較するために以前の音データを再現することも困難となるケースも存在する。これらのことから、編集過程の音データを持続的に認知可能にすることで音作り中の音データが持つ一過性の性質の問題点を解決し、複数の音データを比較することができると考えた。

Step4のAさんのように、演奏者は音色の編集過程において試行した音データ同士を比較し、目標の音色に対して最適と判断した設定を選択することが多い。音データの持

表1 MFCCの算出条件

フレーム長	2048
オーバーラップ幅	512
窓関数	hann 窓
メルフィルタバンク数	128
MFCC次元数	13

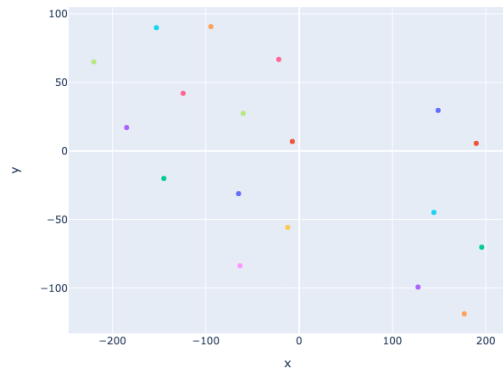


図2 MFCCを入力としてt-SNE法を用いた音データ集合の空間上への可視化結果

続的な認知を可能とするとき、複数の音データの相対的な関係性を提示することで、ユーザが音データを比較する支援になると考えた。

同じ音データに対して、ユーザごとの経験や感性によって受ける印象や感じ方は異なる。複数の音データの相対的な関係性についても同様に、ユーザごとの感覚が存在するものと考えられる。そのため、提示された複数の音データの相対的な関係性は、ユーザごとの感性や理解に従って整理可能であることが望ましい。

以上の考察から、本稿では以下の3つを提案手法のデザイン指針とする。

- 音データの持続的な認知を可能とすること
- 複数の音データの相対的な関係性を提示すること
- 提示された音データの相対的な関係性をユーザごとに整理可能とすること

4.2 音色の空間的可視化

デザイン指針で定めた「音データの持続的な認知を可能とすること」と「複数の音データの相対的な関係性を提示すること」を達成するためにt-SNE法[13]を用いて音響特徴に基づいた空間的な可視化を行う。図2に、聴覚特性に応じた音響特徴量であるMFCCを入力とし、t-SNE法を用いて音データを空間的に可視化した例を示す。表1に、MFCCの算出条件を示す。t-SNE法では、データ入力を確率分布とみなし、確率分布が近い類似データ同士が近傍に配置される。配置されたデータ間の距離で音色の類似度を表現し、ユーザの目的とする音色の関係性で整理されることが期待される。ここで、t-SNE法には調整可能なメタパ

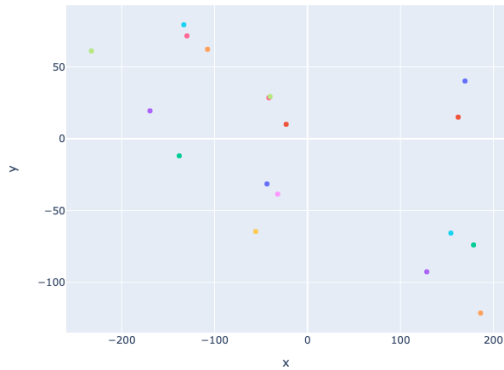


図3 図3に示した t-SNE 法で得られた可視化空間をカーネル回帰によってモデル化した可視化空間

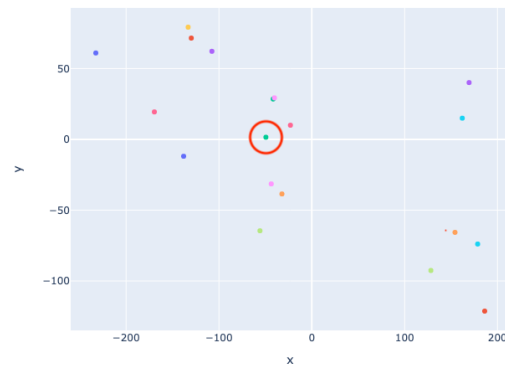


図5 図2で用いた分析対象とする既存のデータ群に対して新たにデータを追加し、図3に示したカーネル回帰によって得られた空間に可視化した様子。赤丸で囲われたデータが新たに追加されたデータを示す。

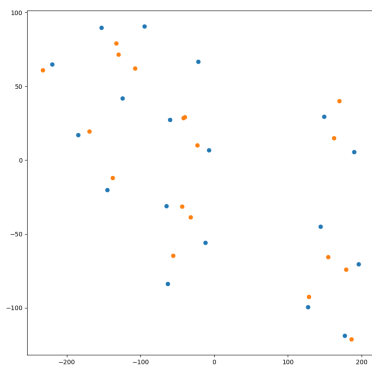


図4 カーネル回帰による可視化空間の再現。図2の点を青色、図3の点を橙色で示す。

ラメータが存在する。このうち、提案手法では Perplexity と Learning rate の 2 種類のパラメータを各ユーザが対話的にチューニング可能にする。これらのメタパラメータを変化させることで低次元に情報圧縮された特徴量が異なるものとなり、結果として可視化空間における音データの配置関係をチューニング可能になる。これにより、ユーザの感覚に沿った複数の音データの相対的な関係性を視覚的に確認可能になり、音色感覚の訓練の実現が期待される。

4.3 カーネル回帰による可視化の保持

t-SNE 法で得られた次元圧縮結果に対して、カーネル回帰によってデータ間の相対関係を数理的に模倣する。入力されたデータ集合に応じて情報圧縮を行う t-SNE 法のみでは、データの追加が行われるたびにデータ集合の次元圧縮を再計算しなければならない。ユーザが対話的にチューニングした可視化空間を保持することはできない。そこで、ユーザのメタパラメータのチューニングによってユーザごとの音色感覚として適した可視化結果が得られた既存データに対する次元圧縮のモデルをカーネル回帰によって数理的に模倣する。チューニングされた空間を模倣したモデル

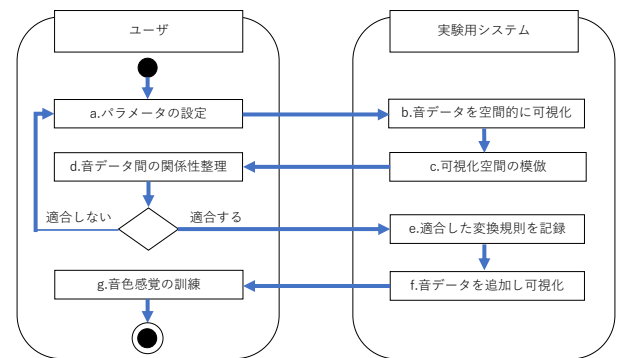


図6 実験用システムについてのアクティビティ図

であるカーネル関数を通すことで、既存のデータ集合内の音データの相対関係が保持された空間内に逐次的にデータを追加し、可視化することが可能になる。図3に、図2に示した t-SNE 法の情報圧縮のモデルをカーネル回帰して得られた既存データの可視化空間を示す。図4に、図2と図3の比較を示す。図中、青色が図2、橙色が図3をそれぞれ示す。同図から、カーネル回帰を行った結果、おおよそ同じ場所にデータが配置されていることがわかる。図5に、図2で分析対象としたデータ群に新たにデータを追加し、図3で得られた空間に可視化した結果を示す。これにより、複数の既存の音データ集合を用いてユーザ自身の音色感覚を 2 次元空間として可視化し、その可視化空間に音色編集の過程で得られる音データを既存の音色との相対関係に応じて配置可能になる。このような可視化によって一過性の音色を空間上に複数配置して参照可能になり、音色感覚の継続的な訓練ができると考えた。

4.4 実験用システムを利用した音色感覚獲得のためのインタラクション

図6に、実験用システムのアクティビティ図を示す。4.2節と4.3節で述べた技術を組み合わせ（図6中bおよ



図 7 実験用システムのメイン画面

び c), ユーザごとに複数の音データの相対的關係を可視化するパーソナライズド空間を作成可能にした。

図 7 に, 実験用システムのメイン画面を示す. 実験用システムでは, ユーザはスライダーによって t -SNE 法に使用するパラメータを調整 (図 6 中 a) でき, 空間内の音データ間の位置關係がユーザの感性ごとに納得できる状態になるまで繰り返し可視化結果を変更 (図 6 中 d) できる. Perplexity は, 下限 5, 上限を (データ数 -1) とし, 分解能は 0.1 で変更可能とした. また, Learning rate は, 下限 150, 上限 250 として, 分解能は 1 で変更可能とした.

5. 実験

訓練対象とする音色感覚として「歪みの強弱」を設定し, 提案手法を検証するための実験を行った. 実験時間は 90 分間とした. ユーザは, ヘッドフォンを装着し音データを試聴する. 発話思考法に基づいた評価を行うため, 実験中は画面録画を行い実験用システムの操作とユーザの音声を同時に記録した. 音色のリストのみを参照するコントロール群と実験用システムを用いる実験群で発話内容を比較した. 確認問題では提示された音データの歪みの強弱關係について, あらかじめ聴取した音データ集合を思い起こしながら回答する. また, システム全体への評価に関するアンケートは, 1~7 の 7 段階評価で実施した.

本研究の目標は, 正しい音色の關係性が身につくということよりも, ユーザ自身が音色の關係性を空間上の位置關係を確認しながら意識することができるようになることに意味があると考えている. そのため, ユーザが実験用システムを使用することによって, 音色の關係性について比較・考察できていると思われる発言が増加することが望ましい.

5.1 実験手順

本実験は以下の手順で行った.

- (1) 事前アンケートの回答
- (2) 実験の説明
- (3) 音色感覚獲得に関する試行

コントロール群 音色リストを用いて音データの相対的關係性を整理

表 2 データセット [14] 内に存在するエフェクターの種類

種類	エフェクト系統	設定の種類数
Over Drive	歪み系	3
Distortion	歪み系	3
Chorus	モジュレーション系	3
Flanger	モジュレーション系	3
Phaser	モジュレーション系	3
Tremolo	モジュレーション系	3
Vibrato	モジュレーション系	3
FeedbackDelay	空間系	3
SlapbackDelay	空間系	3
Reverb	空間系	3
EQ	フィルター系	2
NoFX	—	1

実験群 音色感覚と適合する可視化結果の生成

(4) 確認問題 2 問

コントロール群 口頭でのみ音色間の相対關係を回答

実験群 可視化空間を用いて音色間の相対關係を回答

(5) 事後アンケートの回答

最初に, 被験者は実験参加前での音色の歪みについての理解度を 7 段階の自己申告で回答した. 被験者は, あらかじめ用意されている 18 種類の音データを試聴した. コントロール群では, データ群の相対的な關係性を音源リストをもとに整理した. 一方で, 実験群は実験用システムを用いて, データ群の相対的な位置關係がユーザの音色感覚と適合する可視化結果を作成した. 確認問題には各問 1 種類ずつ 2 種類の音データを用意した. コントロール群と実験群どちらの被験者も用意された問題の音データを試聴して, (3) の試行中に聴取した音データ群と比較して歪みの強弱關係を考察した. このとき, コントロール群の被験者はリスト表示された音データ群を聴取した. 一方で, 実験群の被験者は, 生成した可視化空間を参照しながら音データ群を聴取した. 確認問題を刺激として与えることで, 音色についての発話を誘発するとともに, 被験者の音色に関する考察を確認する狙いがある. 音色感覚獲得のための試行と確認問題への取り組み時の発話を対象として, 音データの空間上への可視化の効果を考察した. 事後アンケートでは, 被験者自身の主観的な評価として, 実験による音色感覚の獲得について評価してもらった.

実験には, 各群で 7 名ずつ計 14 名が参加した. そのうち, コントロール群では 2 名, 実験群では 3 名, が音楽経験者であった.

5.2 実験に用いるデータ

本実験では, Comunita らが作成し公開するエレキギターエフェクターを使用した音源のデータセット [14] から, エレキギターのモノフォニック音源を対象とする. 表 2 に示すように, Comunita らのデータセットには, 11 種類のカ

表 3 実験に用いる各音域の音データの詳細

弦	フレット (音名)	midi 番号
6 弦	0 フレット (E)	40
4 弦	2 フレット (E)	52
1 弦	0 フレット (E)	64

表 4 確認問題で用いる音データのエフェクト条件

問題番号	エフェクト系統	エフェクト設定
確認問題 1	Distortion	2
確認問題 2	Over Drive	3

テゴリに加えて NoFX の音源が wav 形式で用意されており、それぞれ 2 秒間の音源データが収録されている。

実験手順 (3) では、「音域」「エフェクターの種類」「エフェクターの設定」が異なる複数の音データを用いる。音高は、表 3 に示す低域・中域・高域 3 種類それぞれの音域から 1 音ずつオクターブ違いの音データを用意した。エレキギターを用いて音作りをする際にギター本体の変更を行うことは少ないが、エフェクターの種類の変更は多く発生する。そこで、エレキギターは固定とし、エフェクト系統は歪み系の 2 種類 (OverDrive と Distortion) それぞれに対応した 3 種類のエフェクト設定を扱うこととした。データセット内には、奏法ごとの音データが用意されているが、本稿では最も一般的なピック弾きのみを使用する。以上より、計 18 個の wav 形式のデータ (3 音域×2 エフェクト系統×3 エフェクト設定) を用いる。手順 (4) および (5) の確認問題で用いる音データには、6 弦 12 フレット (E) (midi 番号:76) の音高に対して表 4 に示すエフェクト条件の音データを用いた。

5.3 実験結果

手順 (3) と (4) で各群で得られた発話に対して、考察を行う。実験群では、パラメータや可視化結果を理解することが困難な様子が多く見られた一方で、コントロール群では音色の関係性を整理することが困難な様子が見られた。実験用システムはパラメータを自由に設定して可視化結果を試行できる反面、パラメータが可視化結果にどのように影響して変化しているのかということに疑問を持つ様子やパラメータと可視化結果の対応関係を忘れてしまう様子が見られた。

表 5 に、実験中に得られた各群の発話の代表例を示す。コントロール群では、音データの関係性の可視化空間を見ずに視聴のみで比較する環境を用意した。コントロール群の被験者からも、音色の関係性を考察し、音色に対して考察する発言が見られたことから、複数の音データを強制的に比較・聴取させることで、音色感覚を訓練できる可能性が考えられる。一方で、コントロール群では、「音データを忘れてしまう」や「2 択なら評価できるが 18 個のデータを比較することは難しい」という実験群には見られなかった

表 5 実験中に各群で得られた代表的な発話例

被験者群	得られた発話の内容
コントロール群	音データを忘れてしまう。 音の高さが異なるため、歪みの強弱を比較できない。 音色の違いを理解できたが、それが歪みなのかかわからない。 2 択なら評価できるが 18 個のデータを比較することは難しい。
実験群	常に歪みとは何かを考えながらシステムを使用した。 エフェクトのかかり具合や揺れ具合がなんとなくわかった気がした。 パラメータ設定がどのように可視化結果に影響するかわからない。 歪みに関しては理解できなかったが、音色の違いの仮説を立てながら考察しやすかった。 ある程度知識のある状態であれば理解向上につながる。 この音は、こっちにいるべきではない。 音程が固まっている。

表 6 被験者の音色感覚に関する主観的評価

質問 ID	質問内容	平均	
		コントロール群	実験群
1	実験前の歪みに対する理解度	2.29	3.34
2	実験用システムを通して、音データ間の関係性をどれくらいイメージできたか	4.86	4.57
3	実験を通して、歪みの関係性についてどれくらい理解できたか	5.00	4.86
4	実験用システムを今後も使用すれば、さらに音色の関係性について理解できるか	4.29	5.14

一過性の性質に起因する問題点が被験者から発言されていた。実験群では、被験者が音色の関係性を比較し、考察する様子が確認できた。実験用システムを使用し、可視化結果の音データ間の距離から音色の関係性を考察し、歪みの感覚に対して類似しているにも関わらず空間上での距離が離れている音データ同士や異なる感覚の歪みであるにも関わらず距離が近いデータ同士を意識した発言が見られた。被験者の感覚に適合する可視化結果が作成されていない場合でも、可視化を試行するたびに歪みに対する発言の表現が形容詞やオノマトペを用いて明確化されている様子が見られた。これらのことから、可視化空間を参照することで音色感覚の訓練が効率的に実現されたことが示唆された。一方で、「本当にこの感覚は歪みであっているのか」といった被験者の発言も見られ、音色の関係性について考察でき

ても歪みそのものを理解できたとは言えない様子が見られた。音楽未経験者にとっては、エフェクターを用いた歪みという曖昧な定義の音色について思考を巡らせる課題が難しかったものと見られる。これらのことから、音色感覚を訓練する上で、一過性の性質の解決が必要であることが示唆された。

表6に、手順(5)で得られた被験者の主観評価の結果を示す。事前アンケートの質問ID1では、平均的に、実験群に比べてコントロール群の方が歪みに対する理解度が低かったと回答した。実験後の質問ID2と3では、実験群よりもコントロール群の方が平均的に高評価となった。このことから、実験群に比べてコントロール群の方が歪みに対して低い理解度から高い理解度へと成長したようにも見える。しかしながら、表5に示すように、実験中の発言を比較すると、実験群の方が音色を意識した発言が多く音色の関係性について考察できていることがわかる。対して、コントロール群は、理解ができるようになったと事後に自己評価しているものの、実験中には「忘れてしまう」「歪みの強弱を比較できない」など、音色の関係性を十分に把握できているとは考えにくい発言が多く、十分に音色感覚の訓練ができないままに、その状態で音色感覚が習得できたと捉えていると考えられる。また、質問ID4ではコントロール群に対して、実験群の方が高評価であり、被験者は空間的可視化による音色感覚の成長に可能性を感じていることが推察される。

5.4 全体の考察と今後の課題

実験結果から、音データ間の相対的な関係性を可視化することで音色感覚の訓練において、音データの一過性の性質に起因する問題点は解決できる可能性が示唆された。実験用システムを用いた音色感覚の訓練では、コントロール群に比べて複数の音色同士を比較し、それらの違いを考察する発言が多く見られた。

一方で、実験用システムにおける可視化結果を変化させるためのパラメータ調整はユーザビリティの観点から改善点が見られた。パラメータ調整と可視化結果との対応関係についてのアフォーダンスが十分でなかったことで実験用システムの使い方自体に困惑する様子も見られた。また、これに起因して、ユーザが納得する音データ間の関係性を示す可視化空間を得られないことがあった。今後、可視化空間の調整方法と音データそれぞれの音色に対する発言を訓練過程で保存可能にするなど、ユーザ自身の音色に対する感性を空間上に可視化する方法についても検討していく。

6. おわりに

本稿では、音色感覚の訓練にむけて、音響特徴量によって得られた音データ同士の相関関係をユーザごとにパーソナライズした空間上に可視化する方法の有用性を検討し

た。空間上に音データの関係性を可視化する実験用システムを用いた場合と音データを単純に聴取・比較する場合とで音色について比較・考察する過程での発言や音響特徴量間の相対的な関係性を問う課題への回答を比較した。結果として、実験用システムを用いることで、多くの音データの関係性を考察可能になるとともに、音響特徴量間の相対的な関係性を習得可能になることが確認された。

今後は、可視化空間のパーソナライズにおけるパラメータ調整やパラメータと可視化空間の対応関係を示すアフォーダンスの改善などに取り組む。また、音色編集の過程で生まれる音データを音色感覚の習得に向けて利用するうえで有用な音響特徴量についての検討も行う。

参考文献

- [1] 山田大誠, 高島健太郎, 西本一志: 試行錯誤的描画を支援する消し跡機能のデザインに関する基礎的検討, 電子情報通信学会, Vol. 第5回コミック工学研究会予稿集, pp. pp.1-8 (2021).
- [2] 生田泰章, 才記駿平, 西本一志: 文章作成過程における棄却テキスト断片の活用に関する一検討, 情報処理学会インタラクシオン 2016, pp. 302-305 (2016).
- [3] 山西良典, 田中一星, 井本桂右, 山下洋一: 音声エンタテインメントからのウェブ音声マイニングの可能性, 情報処理学会論文誌, Vol. 61, No. 11, pp. 1708-1717 (2020).
- [4] 高橋卓見, 深山 覚, 後藤真孝: INSTRUDIVE: 楽器編成の自動認識に基づく楽曲探索システム, 情報処理学会論文誌, Vol. 61, No. 4, pp. 777-788 (2020).
- [5] 戸本裕太郎, 中村剛士, 加納政芳, 小松考徳: 音素特徴に基づくオノマトペの可視化, 日本感性工学学会論文誌, Vol. 11, No. 4, pp. 545-552 (2012).
- [6] 土井麻由佳, 宮下芳明: 奏法を考慮した箏演奏学習支援手法に関する考察, 情報処理学会論文誌, Vol. 59, No. 3, pp. 912-928 (2018).
- [7] 榊原絵里, 宮下芳明: 磁気ヘッドと磁気テープを用いたヴァイオリン運弓練習支援システム, 情報処理学会インタラクシオン 2013 論文集, pp. 235-237 (2011).
- [8] 元川洋一, 齋藤英雄: 拡張現実表示技術を用いたギターの演奏支援システム, 映像情報メディア学会誌, Vol. 61, No. 6, pp. 789-796 (2007).
- [9] 早川和輝, 長谷川大, 佐久田博司: 主観視点の3D手動作教材提示によるドラム演奏学習支援および熟練者支線情報を利用した教材による学習効果, 日本知能情報ファジィ学会論文誌, Vol. 28, No. 1, pp. 511-521 (2016).
- [10] 竹川佳成, 寺田 努, 塚本明彦: 運指認識技術を活用したピアノ演奏学習支援システムの構築, 情報処理学会論文誌, Vol. 52, No. 2, pp. 917-927 (2011).
- [11] 竹川佳成, 寺田 努, 塚本明彦: リズム学習を考慮したピアノ演奏学習支援システムの設計と実装, 情報処理学会論文誌, Vol. 54, No. 4, pp. 1383-1392 (2013).
- [12] 有山大地, 安藤大地, 串山久美子: ソフトウェアエフェクター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, 情報処理学会論文誌, Vol. 61, No. 11, pp. 1729-1740 (2020).
- [13] van der Maaten, L. and Hinton, G.: Visualizing Data using t-SNE, *Journal of Machine Learning Research*, Vol. 9, pp. 2579-2605 (2008).
- [14] Comunità, M., Stowell, D. and Reiss, J. D.: Guitar Effects Recognition and Parameter Estimation with Convolutional Neural Networks, *The Journal of the Audio Engineering Society*, Vol. 69, No. 7, pp. 594-604 (2020).