

キャラクターの着せ替えイメージによる
DTM(DeskTopMusic)音色プリセット管理手法の提案

2022009 阿部瑞稀

令和7年度提出

目次

第1章	はじめに	4
第2章	関連研究	5
第3章	提案手法	6
3.1.1	音色着せ替え機能	7
3.1.2	コーディネート保存・読み込み機能	8
3.2	想定する利用者と利用シナリオ	8
第4章	システム実装	10
4.1	システム構成	10
4.2	楽曲制作効率向上のための仕様設計	11
第5章	ユーザー実験・評価	13
5.1	アンケート調査	13
5.2	ヒアリング調査	15
5.2.1	調査方法	15
5.2.2	コメントの整理	15

5.3 考察.....	18
5.3.1 ファッションアイテムと音色の対応付けによる音色選択支援の評価.....	18
5.3.2 発展可能性と今後の課題.....	19
第6章 おわりに.....	20
謝辞.....	21
参考文献.....	21

第1章 はじめに

楽曲制作における「音色選び」には、音源プリセットの多さや、プリセット名から音色を想像しにくいといった理由から時間や手間がかかってしまう課題がある。さらに、音色同士の組み合わせの管理は、マルチ音源の「マルチプリセット」やユーザーが保存したセットを名称（文字列）によって識別・管理する方式が一般的である。しかし、名称だけでは内容を直感的に想起しにくく、自分のイメージに合致する音色の組み合わせを探し出すことは困難で、効率的な作曲作業を妨げる要因となる。したがって、音色やその組み合わせを直感的かつ創造的に扱える新たな支援手法を検討することには、音楽制作環境の改善という点で意義がある。

一方で、近年のデジタルコンテンツ分野では、「着せ替え」や「キャラクターメイク」といった要素をもつコンテンツが人気を博している。着せ替え要素があるゲームとして、「ひみつのアイプリ」[1]などの女兒向けのものから、「モンスターハンターワイルズ」[2]などの幅広い層から人気を集めるものまで、多くのコンテンツが存在する。さらに、近年では、VRプラットフォーム「VRChat」で、3D アバターの見た目や衣装を「変更」する文化が活発に広がっており、「キャラメイク」が個性の演出や交流の手段となっているといえる。このように「着せ替え」要素をもつコンテンツの人気が非常に根強いことが分かる。

さらに、現代では、「ボカロ曲」などのインターネット発の音楽が非常に人気を集めている。その中で YouTube やニコニコ動画といった動画投稿サイトに楽曲を投稿する作曲者が多く存在する。そこで「動画」投稿サイトという性質上、視聴者にとって視覚的な情報も重要であり、曲そのものだけでなくミュージックビデオやサムネイルといったビジュアルも再生数や認知度に関わる大きな要素であるといえる。VOCALOID 楽曲投稿イベントである「ボカコレ」のランキングを見ると、ミュージックビデオ内に、合成音声キャラクターやオリジナルキャラクターといった「キャラクター」を登場させる投稿者が非常に多く、またそのような楽曲の人気が高いことがわかる。しかし、全ての作曲者がイラスト、キャラクターデザイン、動画編集といったスキルも持っているわけではないため、ミュージックビデオ制作のハードルが高いという問題がある。

そこで、本研究では、「着せ替えゲーム」の楽しさを「作曲」に新たに取り入れ、音楽制作における「音色選び」の悩みと「ミュージックビデオ作成のハードルの高さ」を同時に解決し、ユーザーが楽しみながら視覚的に音色を管理できることを目指す。作曲者自身の「オリジナル曲とリンクするオリジナルキャラクター」を生み出すことを叶える創作支援ツール『きせおんどール』を提案する。

第2章 関連研究

音色の検索手法に関する先行研究としては、機械学習や可視化を用いて、大量のプリセットから目的の音色に到達しやすくする枠組みが提案されている。SynthScribe[3]では、シンセサイザー音色の検索・探索を支援するため、音声信号やテキストなど複数の情報を共通の潜在空間に埋め込むマルチモーダルモデルを用い、自然言語や例示音から近い音色を検索できるインタフェースを実現している。また、齋藤ら[4]は、DTM環境における音色づくりの効率化を目的として、音色に付随するメタデータと機械学習を組み合わせ、ユーザーの意図や好みに近い音色プリセットを推薦する手法を検討している。これらはいずれも、音色とテキスト情報・メタデータをモデルとして扱うことで、自然言語や例示音といった高レベルな入力から目的のプリセットに到達しやすくする点に特徴がある。

一方、吉岡ら[5]は、電子音色辞書において、音響特徴量に基づく音色間の類似関係を3次元空間にマッピングし、粒子として可視化するインタフェースを提案している。ユーザーは3D空間上を移動しながら音を試聴することで、音色空間の構造や類似度を視覚的に把握しつつ目的の音を探索できる。このように、機械学習や可視化を用いた音色探索手法は提案されているが、いずれも対象は単体の音色プリセットであり、複数音色からなる“セット”全体の管理や想起は扱っていない。

市販のマルチティンバー型音源であるKONTAKT[6]やOmnisphere[7]などでは、単一のプラグイン内で複数の音色をレイヤーしたりマルチティンバー構成で同時に鳴らしたりでき、それらを「Multi」等として保存・読み込みできる。しかし、こうしたプリセット管理は多くの場合文字列のプリセット名称やフォルダ構造のみに依存しており、どの組み合わせがどの音色を含んでいたのかを名称から直感的に想起することが難しい場合がある。ライブラリ規模が大きくなるほど、過去に自分が作成した良い組み合わせを探し出す手間が増え、同様の構成を繰り返し作り直してしまうという問題が起こる可能性がある。また、プリセットブラウザのタグや検索機能も、多くは単一パッチの分類・検索を補助するものであり、複数パートからなるマルチ全体を視覚的に把握しやすくする仕組みは限定的である。

これに対して、本研究の『きせおんどール』は、「音色」と「ファッションアイテム」を対応させ、6種類の音色カテゴリからなる構成を「コーディネート」として扱うことで、マルチ音源の構成を衣装の組み合わせとして記憶・操作できる仕組みを提案する。ユーザーは「この衣装の組み合わせ=この音色セット」という形でマルチ構成を視覚的に想起できるため、従来の文字列中心の管理と比べて、過去の構成の再利用や比較、バリエーションの生成が容易になると期待される。さらに、コーディネートはキャラクターの外見とも紐づいているため、後続のミュージックビデオ制作やキャラクター表現とも一貫した形で扱える点で、従来のマルチティンバー型音源とは異なる創作支援の枠組みを提供できると考える。

第3章 提案手法

本研究では、楽曲制作支援プラグイン『きせおんどール』を提案する。このプラグインを VST3 プラグインに対応した DAW(音楽制作ソフトウェア)に読み込ませることで、楽曲制作における音色選びの場面で使用することができる(図1)。ユーザーが DAW のプラグイン上でドール型のキャラクターに服を着せ替えることで楽器の音色を変更できるようにし、ユーザーが組んだ「コーディネート」のそれぞれの服に紐づいた音色のセットで作曲することを叶える。

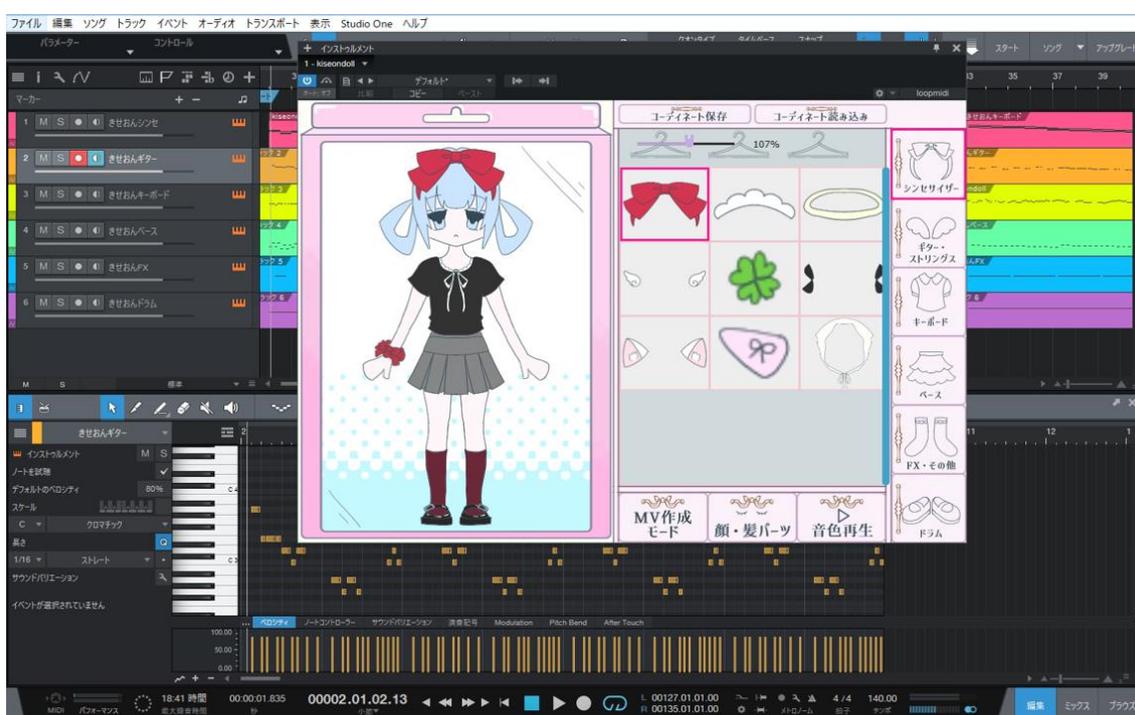


図1: 『きせおんどール』をDAW(Studio One)に読み込んで作曲している画面

3.1 システム概要

本システムは、DAW プラグインとして独立した6種類の音源と着せ替え画面のユーザーインターフェースから構成される。画面は左右で構成されている。左側ではキャラクターの全身図が表示され、衣装が反映される。「音色」を「ファッションアイテム」に、「音色セット」を「コーディネート」に対応させている。右側の「音色アイテム選択画面」ではファッションアイテムを選ぶことができ、様々なアイテムに割り振られた音色を見ながら選択することが可能である。

3.1.1 音色着せ替え機能



図2: 『きせおんドール』VST プラグインの音色着せ替え機能

ファッションアイテムのカテゴリと音色のカテゴリを対応させているため、コーディネートを組むように音色を組み合わせたことが可能である(図2)。

本プラグインでは、ファッションアイテムのカテゴリと音色のカテゴリが対になった全6カテゴリを用意した。ファッションアイテムのカテゴリの「髪飾り」は音色のカテゴリの「シンセサイザー」に、「アクセサリ」は「ギター・ストリングス」に、「トップス」は「キーボード」に、「ボトムス」は「ベース」に、「靴下」は「FX・その他」に、そして「靴」は「ドラム」に対応させた。音色の6カテゴリに対応したアイテムのイラストを制作し、合計69アイテムを着せ替えることができた。

さらに、ファッションアイテムだけではなく「ドール」自体の見た目もカスタマイズできるようになっており、「顔・髪パーツ変更モード」で、ユーザーは髪パーツの選択ができるほか、カラーサークルで髪と目の色を変更することができる。

このプラグインで選択した音色は、カテゴリごとにトラックを分けてMIDI入力できるようにしている。シンセサイザーはch1、ギター・ストリングスはch2、キーボードはch3、ベースはch4、FX・その他はch5、ドラムはch6と、ユーザーがDAW上でトラックのチャンネルを設定することで、選んだ複数の音色を同時に鳴らして作曲することができる。

3.1.2 コーディネート保存・読み込み機能



図 3: 『きせおन्दール』のコーディネート保存・読み込み機能

図 3 に示す通り、「コーディネート保存」ボタン、「コーディネート読み込み」ボタンでコーディネートに紐づいた音色セットの保存・読み込みができる。コーディネートは、選んだ 6 カテゴリのアイテムのサムネイル画像とコーディネート名が表示される。これにより、ユーザーが作成した音色セットの想起が容易になり、楽しみながら視覚的に音色を管理することを可能にする。「コーディネート」という視覚的な情報に紐づいた状態で音色のセットを管理できることで、文字列の名称で管理するよりもユーザーにとって記憶しやすくと考えられる。

3.2 想定する利用者と利用シナリオ

『きせおन्दール』は、ユーザーによって使用方法が選択できるため、作曲初心者から経験者まで幅広く使用されることを想定している。

作曲初心者に向けての『きせおन्दール』

本プラグインは、“着せ替えゲーム”をするように音色を選び、自分で作曲に用いる音色のセットを制作することができる。そのため、作曲初心者は、「どんな音色を使ったらよいかわからない」「音色選びって難しそう」という悩みを抱えることがあるが、ゲームのように手軽に音色セットを作成できることで、作曲を始める際の心理的負担が軽減されると考えられる。さらに、着せ替えゲームが好きな人や可愛いキャラクターが好きな人にとっても、本プラグインは作曲に興味を持つきっかけになり得ると考えられる。

作曲経験者に向けての『きせおんどール』

作曲経験者に向けては、「コーディネート縛り作曲」という使い方を想定している。プラグインでファッションアイテムと対応した音色を選ぶ際、「音色試し聞きボタン」を押すことで対応した音色を聞くことができる。しかしあえてそのボタンを押さず、アイテムを見た目のみから選ぶことで完成する「コーディネート」に対応した「音色セット」の音を確認しないまま作曲に取り組み始める、「コーディネート縛り作曲」も想定している。作曲経験者にとって、普段作曲する際は音色の音を一つ一つ聞いて確認するのが一般的である。しかし、『きせおんどール』で、視覚情報からコーディネートを組み、音を聞くことなく音色セットを作成し作曲を始めることで、思いがけない音色の組み合わせができたり、マッチする音色の組み合わせが見つかったりと、新しい作曲の楽しさが生まれると考えている。

ユーザーが『きせおんどール』を使っていく中で気に入ったコーディネート=音色セットができたら、「コーディネート保存ボタン」で保存し、「コーディネート読み込みボタン」でいつでも呼び出すことができる。コーディネートはプラグイン上に保存されているため、新しいプロジェクトで新たに曲を作り始めた際にも、以前保存したコーディネート=音色セットを呼び出し作曲に用いることができる。

“ボカロ P”に向けての『きせおんどール』

VOCALOID などの合成音声ソフトウェアを用いてインターネット上に楽曲を投稿する、いわゆるボカロ P も重要な利用者像である。動画投稿サイトにおける楽曲では、サムネイルやミュージックビデオにオリジナルキャラクターを用いる事例が多く見られ、キャラクターの存在が作品の世界観や認知度に大きく関わっている。サムネイルやミュージックビデオに登場させる「キャラクター」は音声合成ソフトウェアに付随するキャラクターである場合が多いが、音声合成ソフトウェアキャラクターではない「オリジナルキャラクター」が登場する事例も少なくない。したがって、楽曲に「オリジナルキャラクター」を用いたいという需要が存在すると考えられる。しかし、ボカロ P にとって、自分自身で楽曲一つ一つに対応するオリジナルキャラクターを考えるのはハードルが高い場合がある。本研究で提案する『きせおんどール』を用いると、音色セットにコーディネートが対応しているため、曲に用いた音色セット=コーディネートを身につけた、『曲とリンクするオリジナルキャラクター』が生まれる。本プラグインで作成した「オリジナルキャラクター」を、楽曲のミュージックビデオやサムネイルに登場させるという活用方法はもちろん、曲を作るとキャラクターが生まれるという新しい体験から、新しい楽しさも生まれると考える。

第 4 章 システム実装

4.1 システム構成

本システムは C++ 言語と JUCE フレームワークを利用して VST プラグインとして開発した。ユーザーインターフェースおよびチャンネルごとの MIDI メッセージ処理も JUCE の機能を利用している。音源部については、「シンセサイザー」は減算式シンセサイザーを実装し、「ギター・ストリングス」、「キーボード」、「FX・その他」、「ドラム」はサンプラー方式で、「ベース」については減算式シンセサイザーとサンプラー方式両方を用いて実装した。

ホスト DAW からは、6 つのカテゴリ（シンセサイザー、ギター・ストリングス、キーボード、ベース、FX・その他、ドラム）に対応した MIDI チャンネル入力を与えられる。プラグイン内部では、受信した MIDI メッセージをチャンネル番号に基づいて各カテゴリの音源エンジンに振り分け、カテゴリごとに独立したオーディオバスへ出力する。これにより、ホスト DAW 上でカテゴリごとに別トラックとして扱うことができ、後述するように、トラック単位で音量やエフェクト、パンニングを調整できるパラアウト構成を実現している。

ユーザーインターフェースは、JUCE の `AudioProcessorEditor` クラスを継承して実装しており、左側にはキャラクターの全身図と着せ替え結果を表示する領域を、右側の「音色アイテム選択画面」では各カテゴリのアイテムをアイコン状に並べて表示し、ユーザーがクリックすることで音色アイテムを選択できるようにしている。サンプラー方式の音源で使用するドラムループや楽器音色の WAV ファイルは、JUCE の `BinaryData` 機能を用いてプラグインバイナリに埋め込み、外部ファイルのパス設定を行うことなく利用できるようにしている。

4.2 楽曲制作効率向上のための仕様設計

本システムでは、単に着せ替えによって音色を変更できるだけでなく、DAW 上での楽曲制作を効率のかつ直感的に行えるよう、音源部および出力構成にいくつかの仕様上の工夫を施している。以下に、主な設計方針とその意図を示す。

コーディネート単位のプリセット管理

本プラグインでは、コーディネートとそれに対応する音色セットをプラグイン内に保存し、いつでも呼び出せるようにしている。ユーザーが『きせおんどール』を使用する中で気に入ったコーディネート=音色セットができた場合、「コーディネート保存」ボタンにより保存しておくことで、新しいプロジェクトで楽曲制作を始める際にも、以前作ったコーディネートを即座に呼び出して利用できる。

ユーザーはコーディネート=音色セットに名前を付けて保存することができる。使いたい音色セットを呼び出す際には、名前だけではなく、コーディネートの見た目も表示されるため探しやすいと考えられる。

パラアウト構成によるトラックごとの制御

本プラグインは、6つの音色カテゴリそれぞれに対して独立したオーディオ出力バスを持つパラアウト構成とした。これにより、ホスト DAW 上では、各カテゴリを別トラックとして立ち上げることができ、カテゴリごとに音量バランス、エフェクト、パンニングなどを個別に調整できる(図4)。

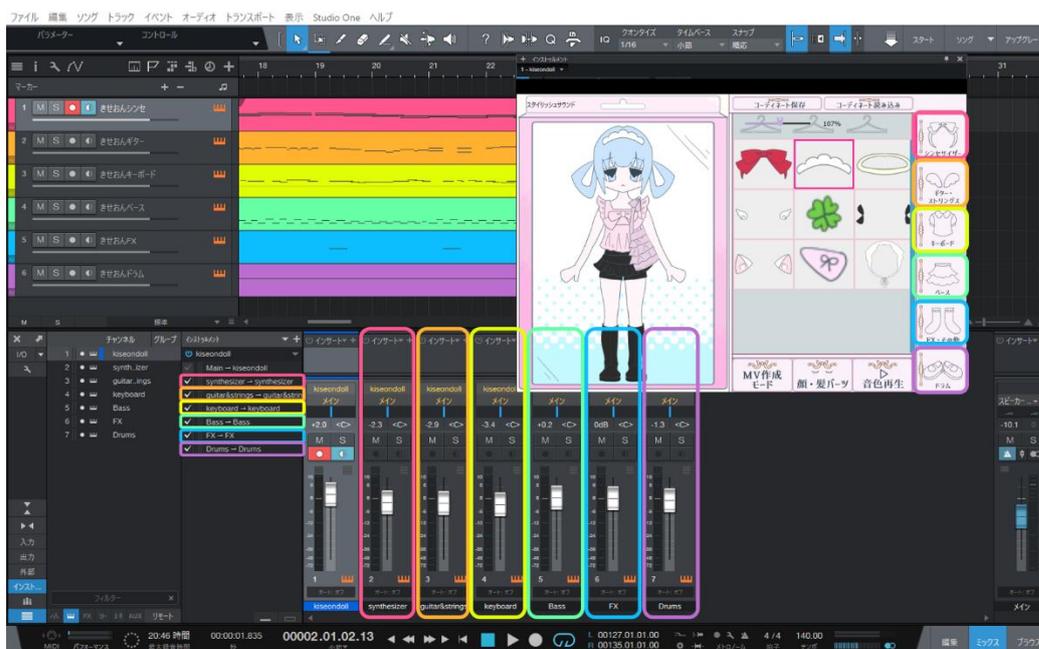


図4:DAW 上で、『きせおんどール』のカテゴリごとにパラアウト出力されている様子

ポリフォニック／モノフォニックの設計

本システムでは、「シンセサイザー」、「ギター・ストリングス」、「キーボード」「ベース」カテゴリについては和音演奏を想定し、ポリフォニックな発音方式で実装している。一方で、ドラムループ音源を実装している「ドラム」カテゴリと WAV ファイルの FX 音源を実装している「FX・その他」のカテゴリについては、一般的なフレーズやループ素材をそのまま再生することが多く、同時に複数のノートを重ねて演奏する場面が少ないためあえてモノフォニックな仕様とした(図5)。

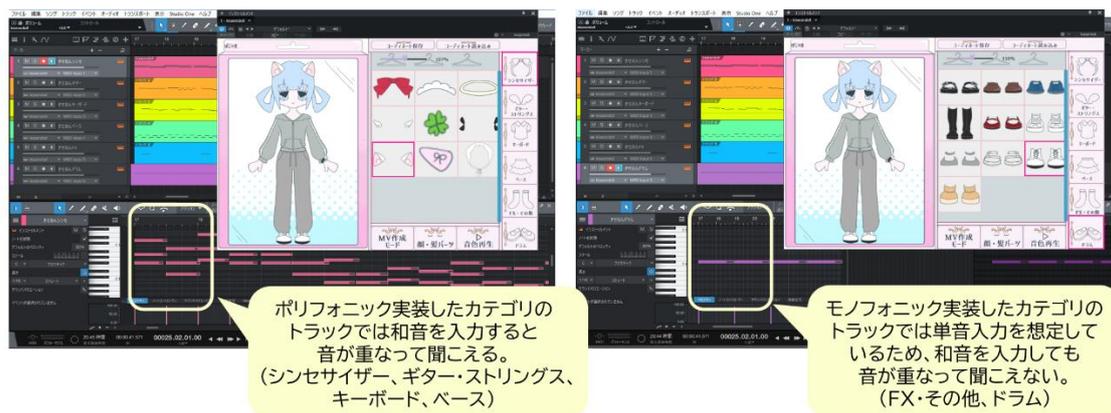


図5:ポリフォニック・モノフォニック実装したカテゴリでそれぞれ作曲している様子

ドラム／FX カテゴリのピッチ固定

「ドラム」と「FX・その他」のカテゴリでは、MIDI ノート番号に応じて発音タイミングのみを制御し、音程（ピッチ）は変化させないように実装している。これは、ドラムループや効果音素材に対してピッチシフトをかけると、音程だけでなく質感も大きく変化してしまい、意図しない結果を招く場合が多いためである。本システムでは、これらのカテゴリを「楽曲の土台となるリズム・雰囲気を提供する役割」と位置づけており、音程よりもリズムや質感が重要であると考えた。そのため、MIDI キーに関わらず常に原音のピッチで再生し、テンポや配置のみをユーザーが DAW 側で調整できるようにしている。

ドラムループのテンポ同期とピッチ保持

ドラムカテゴリでは、初心者が短時間でリズムトラックを構築できるように、単発のドラム音ではなくドラムループ素材を採用した。これらのドラムループは、宿主 DAW の BPM に自動的に同期するように実装している。具体的には、受信したテンポ情報に基づいて再生速度を調整し、ループの長さが小節単位で DAW のグリッドに一致するようにしている。さらに、テンポを変更してもドラムループの音程が変化しないよう、タイムストレッチにより再生速度のみを変更し、サンプルのピッチは維持する方式を採用した。これにより、ユーザーは楽曲ごとに任意の BPM を設定しても、ドラムのピッチが不自然に高くなったり低くなったりすることがないそのままのピッチのリズムパターンを利用することができる。

第5章 ユーザー実験・評価

2025年12月03日(水)～12月05日(金)に行われた「WISS 2025: 第33回インタラクティブシステムとソフトウェアに関するワークショップ」にて、本研究のデモ発表を行った。割り当てられたブースにて、『きせおんどール』の実演展示、説明を行った。ブースを訪れた来場者を対象に、『きせおんどール』を体験してもらい、アンケート調査とヒアリング調査を実施した。

体験の具体的内容としては、VST プラグイン『きせおんどール』を用いて自由にコーディネート=音色セットを組んでもらった後、DAW上でその音色セットと対応した曲を流し、来場者が組んだコーディネートと対応した音色を聴いてもらった。その後、「コーディネート呼び出し機能」を用いて他の来場者が作ったコーディネートを曲が流れているリアルタイムで呼び出し、同じメロディでも楽器が変わることで曲の聞こえ方も変わることを実感してもらった。なお、本章で報告する評価は、イベント会場での短時間の体験に基づくものであり、参加者に十分な時間を与えて楽曲制作そのものを行ってもらったわけではない。したがって、以下の結果は、本システムに対する第一印象や概略的な使い心地を示すものとして位置付ける。

5.1 アンケート調査

本研究では、提案システムに対する定量的な評価を行うため、ブース来場者を対象にアンケート調査を実施した。回答者数は23名である。アンケートでは、まずDAW利用経験年数を尋ね(Q0)、その後、以下の3項目について5段階評価で求めた。

- Q0: DAW 経験年数 (10年以上, 5～9年, 1年未満, 未経験)
- Q1: 音色をイラストと対応させる視覚的管理は有効だと感じたか
- Q2: 音色の組み合わせを管理できる機能は必要だと感じたか
- Q3: きせおんどールの総合的な評価

アンケート結果を表1、図6に示す。全回答者の平均値は、Q1が4.0、Q2が4.3、Q3が4.4(小数第2位を四捨五入)となり、いずれの項目においても5段階中4点以上の評価が得られた。また、Q1およびQ2では23名中19名、Q3では23名中21名が4または5と回答しており、視覚的な音色管理機能、音色セット管理機能ならびにプラグイン全体の使用感に対して、概ね高い評価が得られたといえる。

DAW 経験年数別に見ると、未経験者(15名)および1年未満(4名)の回答者では、3項目とも平均値がおおよそ4.0～4.8と高く、特に総合評価(Q3)については5を選択した回答が多かった。一方で、5～9年の経験者(3名)では、視覚的管理の有効性(Q1)の平均値が3.3とやや低くなる傾向が見られたが、音色の組み合わせ管理の必要性(Q2)

および総合評価(Q3)はいずれも 4.0 程度であり、おおむね好意的な評価が得られている。なお、「10 年以上」の回答者は 1 名であるため、本カテゴリの平均値については解釈に注意が必要である。

以上より、提案システムは特に DAW 経験の浅いユーザーから高く受け入れられている一方で、経験年数の長いユーザーからは視覚的管理の有効性に対してやや慎重な評価も見られることが分かった。この点については、次節のヒアリング調査の結果とあわせて、第 5.3 節で考察を行う。

- Q0：DAW 経験年数（10 年以上，5～9 年，1 年未満，未経験）
- Q1：音色をイラストと対応させる視覚的管理は有効だと感じたか
- Q2：音色の組み合わせを管理できる機能は必要だと感じたか
- Q3：きせおんドールの総合的な評価

表 1: DAW 経験年数別の回答者数と各質問項目の平均評価値

DAW 経験	N	Q1 平均	Q2 平均	Q3 平均
未経験	15	4.20	4.27	4.47
1 年未満	4	4.00	4.75	4.75
5～9 年	3	3.33	4.00	4.00
10 年以上	1	4.00	5.00	4.00

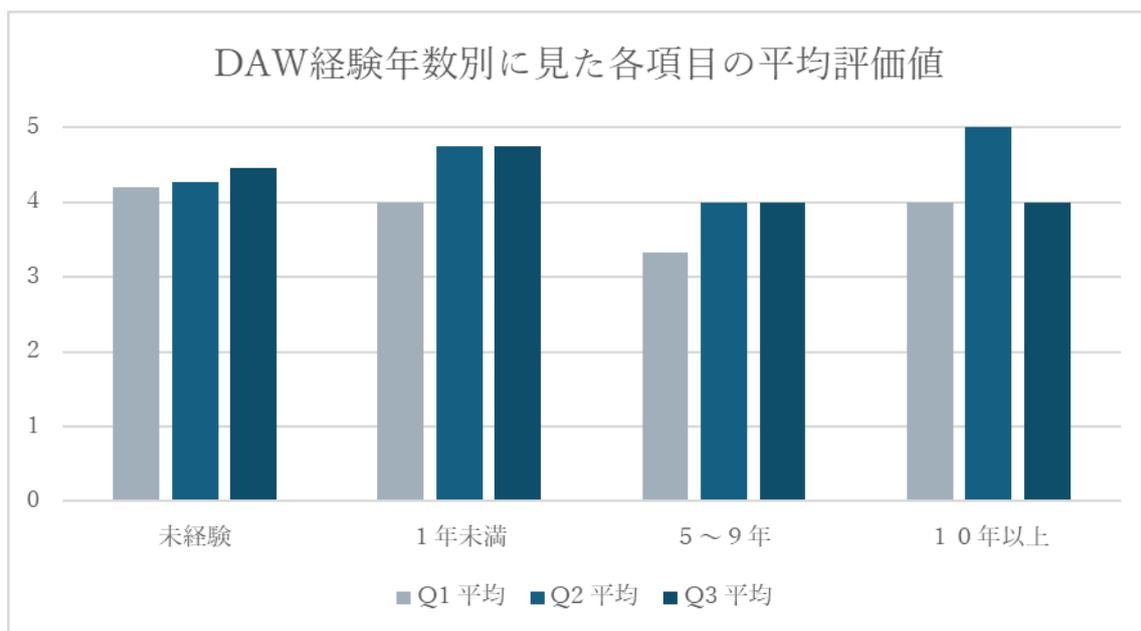


図 6: DAW 経験年数別に見た各項目の平均評価値

5.2 ヒアリング調査

5.2.1 調査方法

本研究では、提案システムに対するユーザーの受容性や改善点を把握することを目的として、ブースに来た来場者を対象にヒアリング調査を実施した。本プラグイン『きせおんドール』を展示し、来場者に実際に操作してもらいながら口頭で意見や感想を収集する形式で行った。

来場者には、プラグインの概要説明を行った後、着せ替え操作およびコーディネート保存・読み込み機能を試用してもらい、さらに来場者が組んだ「コーディネート」に紐づく音色を使用した曲をその場で聴いてもらった。その際、その場で生じた疑問点やアイデア、意見などを口頭で聴取した。内容を記録した参加者は、作曲経験者、作曲未経験者、企業関係者、学生など多様な背景をもつ来場者であり、複数名からはゲームへの応用可能性や、音楽とキャラクター表現の関係についてのコメントも得られた。

5.2.2 コメントの整理

収集したコメントを内容に基づいて分類した結果、主に (1) ユーザーインタフェースと表示方法に関する意見、(2) 音色選択とコーディネートの関係に関する意見、(3) 応用可能性に関する意見、(4) ターゲット層と支援機能に関する意見、(5) 音色と服の対応付けに関する意見、(6) サウンド面の指摘、(7) コンセプトや位置づけに関する質問、に整理できた。以下に、各カテゴリの代表的なコメントを示す。

(1) ユーザーインタフェースと表示方法に関する意見

「コーディネート読み込み」画面に表示するコーディネート一覧について、現在はカテゴリごとのアイテムの並びで構成しているが、「ユーザーはキャラクターが着ている状態のビジュアルとして記憶しているため、サムネイルはキャラクター全体の見た目が分かる形のほうがよい」との指摘があった。また、「作曲する際は楽器の左右のパンや空間上の位置を重視することが多いため、アクセサリーの位置などと連動させて、音の定位をアイテムごとに可視化できるとよい」といった意見も得られた。

さらに、「このキャラクターが動き出したら面白そうだ」といったコメントも複数回挙がっており、キャラクターのアニメーションや動きと音の変化を連動させることへの期待も示された。

(2) 音色選択とコーディネートに関する意見

本プラグインが想定している「コーディネート縛り作曲」に関連して、「作曲経験者だと音の組み合わせを重視するため、どうしても音色試し聞きボタンで音を確認したくなるのではないか」という指摘があった。また、「音にこだわって組み合わせると見た目のコーディネートが崩れてしまったり、逆に見た目だけで選ぶと音の相性が悪くなりそうだ」といった懸念も示された。

一方で、作曲経験者から「楽器名のリストから選ぶよりも、服として見た目から音色を選ぶほうが楽しい」との感想も得られており、音色選択のインターフェースとしてのポジティブな評価も確認された。

(3) 応用可能性に関する意見

複数の来場者からは、「作曲だけでなくゲームの世界にも応用できそう」とのコメントが寄せられた。具体的には、「アクションゲームでキャラの装備を変えると戦闘 BGM が装備に対応して変化する」、「ゲーム冒頭のキャラメイク時にコーディネートを組みると、そのキャラ専用のテーマ曲が生成される」、「レースゲームで乗るカートごとに対応した BGM が流れる」といった、ゲーム内の装備やキャラクターと音楽を連動させるアイデアが挙げられた。また、コーディネート読み込みモードでリアルタイムに音色が切り替わる様子を体験した参加者からは、「見た目が切り替わるのと同時に音が変わるので、DJ イベントなどのライブパフォーマンスでも使えそう」といったコメントもあり、ライブ演出ツールとしての利用可能性も指摘された。

(4) ターゲット層と支援機能に関する意見

来場者からは「どのような層をターゲットにしているのか」という質問があり、作曲初心者と経験者の双方を想定していることを説明したところ、「作曲初心者もターゲットに含めるのであれば、作曲の仕方を教えるようなガイド機能があるとよいのではないか」との提案があった。

また、作曲未経験者からは、「良い音色の組み合わせが分からないので、あらかじめいくつかの音色セットのプリセットが用意されていて、そこから選べると助かる」といった意見も得られた。一方で、「着せ替えゲームとしても楽しく、作曲初心者にとって良い入り口になりそう」、「キャラクターや服がかわいいので、使っていて楽しい」といった肯定的なコメントも多く、キャラクターや着せ替え要素がポジティブに受け止められている反応が確認された。

(5) 音色と服の対応付けに関する意見

「音色カテゴリとファッションアイテムのカテゴリをどのような考え方で対応付けているのか」という質問が複数の参加者から出され、現状は制作者自身の感覚に基づき決定していることを説明した。これに対して、「対応関係に何らかの理論的な根拠があると、今後の発展につながるのではないか」、「AI に音色と対応する服を決めてもらう、あるいは音色に対応する服のデザインを生成してもらうと面白そう」、「どの音色がどの服に合うと感じるかについてアンケート調査を行い、対応付けを決めるのもよいのではないか」といった提案が得られた。

さらに、「ファッションアイテムに限らず、何かを音と対応付けるというアイデアには拡張性がありそう」、「音に対応させる対象は服以外でもよいのではないか」といったコメントもあり、本研究の枠組みを一般化して捉える視点も示された。

(6) サウンド面の指摘

シンセサイザーカテゴリ内の音色について、「アイテムごとに音量差があり、特にサイン波のような波形は他と比べて聞こえにくい印象があるが、音量バランスについてどのようなアプローチをとっているのか」といった、サウンドデザインやミキシングに関する技術的な質問も寄せられた。

(7) コンセプトや位置づけに関する質問

コンセプトに関しては、「これはどのように思いついたのか」といった質問が複数の参加者から寄せられた。

また、「わざわざキャラクターや服を使う意味が分からない。作曲する人にとってキャラクターであるメリットは何か」といった問いもあり、ボーカロイド楽曲ではサムネイルにキャラクターを用いることが多く、オリジナルキャラクターへの需要があること、曲とリンクしたオリジナルキャラクターを作成できる点が本システムの特徴であることを説明したが、必ずしも全員に納得されるわけではなく、本システムの価値の伝え方に課題があることも明らかになった。

さらに、作曲未経験者からは「このプラグインを使えば自動的に曲ができるのか」という質問も複数あり、本システムがあくまで音色選択やコーディネートを支援するツールであり、実際の作曲はユーザー自身が行う必要があることを説明した。このことから、提案システムの機能範囲や位置づけについて、今後さらに明確に伝える工夫が必要であると考えられる。

5.3 考察

本節では、アンケート調査およびヒアリング調査の結果を踏まえ、提案システムの有効性と今後の改善方針、および発展可能性について考察する。

5.3.1 ファッションアイテムと音色の対応付けによる音色選択支援の評価

アンケート調査において、音色をイラストと対応させる視覚的管理の有効性 (Q1)、音色の組み合わせ管理の必要性 (Q2)、きせおんドールの総合的な評価 (Q3) のいずれについても平均値が 4.0 以上であったことから、音色をファッションアイテム、音色セットをコーディネートと対応させて扱うという本システムの枠組みは、参加者全体に概ね良好に受け入れられたといえる。特に DAW 未経験および 1 年未満の回答者で高い評価が得られた点は、「着せ替え」という方法が、従来のプリセット名中心のインタフェースに比べて、音楽制作の経験が浅いユーザーにとって分かりやすく親しみやすい入り口として機能していることを示唆している。

ヒアリング調査でも、「楽器名のリストから選ぶよりも、服として見た目から音色を選ぶほうが楽しい」、「着せ替えゲームとしても楽しく、作曲初心者にとって良い入り口になりそう」といったコメントが得られた。これらの結果は、本研究が目標とした「着せ替えゲームの楽しさを作曲体験に取り入れる」というコンセプトが、少なくとも初心者層に対しては有効に機能していることを裏付けている。「『きせおんドール』を利用することでユーザーが楽しみながら視覚的に音色を管理できるようにする」という目標を達成できる可能性も示唆された。

一方で、5～9 年の経験をもつユーザーでは Q1 の平均値が作曲未経験・1 年未満の回答者と比べて相対的に低いことから、経験の蓄積とともに、視覚的な分かりやすさに加えて、音響的な情報や既存のワークフローとの整合性がより重視される傾向があると考えられる。ただし、この層においても Q2 および Q3 はおおむね高評価であり、音色セット管理機能やプラグイン全体のコンセプト自体は受け入れられていることが分かる。したがって、本システムは特に初心者～中級者に対して「音色選択のハードルを下げるツール」として有効でありつつ、経験者に対しては「コーディネート縛り作曲」や「曲とリンクするオリジナルキャラクターができる」点など、新しい作曲の楽しみ方としての利用を促すツールとして位置づけるのが適切であると考えられる。

5.3.2 発展可能性と今後の課題

ヒアリング調査からは、本システムの枠組みがさまざまな方向に拡張し得ることも明らかになった。特に、「このキャラクターが動き出したら面白そうだ」といったコメントが複数の参加者から寄せられた点は、音色選択だけでなく、キャラクターの動きや演出と組み合わせた表現への期待が大きいことを示している。ゲームへの応用や DJ イベントでの利用可能性に関するコメントは、「見た目の変化と音の変化を結びつける」という発想が、作曲支援にとどまらず、他のエンターテインメントへと展開可能であることを示している。本研究では第 6 章で 3D キャラクター MV 制作ソフトとの連携について述べるが、これらの意見は、提案システムが映像表現やパフォーマンスと結びついた総合的な制作環境へ発展し得ることを後押しするものといえる。

今後の課題として、まず、音色と服装の対応付けをより体系的に扱うことが挙げられる。AI によって音色から服装を提案したり、ユーザアンケートを通じて多数の対応例を収集することで、「どのような音がどのような服に合うと感じられるのか」をモデル化していくことが可能性として考えられる。しかし、服と音色対応の感覚は個人によって異なると考えられるため、多くのユーザーにとって納得感のある対応付けをすることは難しさがあるといえる。

次に、ターゲットごとに利用方法を提示することである。初心者向けには、あらかじめ用意したコーディネートプリセットを提供し、「まずはこのコーディネートに対応した音色セットを使えばいい感じになる」といった体験をしてもらうことが作曲においての音色選びのハードルを下げることに繋がると考えられる。一方で経験者向けには、「コーディネート縛り作曲」など、既存の制作ワークフローに対する“新しい作曲の楽しみ方”としての位置づけを明確にすることで、本システムの独自性を活かしやすくなると考えられる。

さらに、本システムのコンセプトや機能範囲を分かりやすく伝えるための情報設計も重要である。説明文やチュートリアル、デモ動画などを通じて、自動作曲ツールではなく、あくまで音色選びの段階で使うとキャラクターができるという楽曲制作支援ツールであることを明示することで、より適切な文脈で本システムを利用してもらえると考えられる。自動作曲機能を付けない理由としては、作曲をしたことがない人が『きせおんどール』が楽曲制作へ取り組むきっかけとなることも期待しており、ユーザー自身でメロディを考え曲を作っていく楽しさを知る前段階としての“音色選び”への支援ツールとして位置付けたいと考えているからである。

以上のように、提案システムは、特に初心者に対する音色選択支援という点で一定の有効性が確認されると同時に、ユーザーニーズに応じた機能拡張や説明の工夫によって、さらなる発展が期待できることが示された。

第6章 おわりに

本研究では、ユーザーが楽しみながら視覚的に音色を管理できるようにすることを目指し、音色を「ファッションアイテム」に、音色のセットを「コーディネート」に対応させたマルチティンバー型音源プラグイン、『きせおんどール』を開発した。

実装面では、JUICE フレームワークを用いて VST3 プラグインとして開発し、カテゴリごとに独立したオーディオバスを持つパラアウト構成、ポリフォニック・モノフォニックの使い分け、ドラムループのテンポ同期とピッチ保持など、実際の DAW 上での楽曲制作を意識した仕様を導入した。これにより、見た目の楽しさだけでなく、実用的な楽曲制作ツールとして利用できることを目指した。

さらに、本研究では現在、VST プラグイン『きせおんどール』で作成したキャラクターが DAW 上の MIDI 入力に合わせて動く Unity 製の MV(ミュージックビデオ)制作ソフトの開発にも取り組んでいる。『きせおんどール』で作ったコーディネート=音色セットや髪形、髪色の情報を文字コードとしてプラグインから出力し、Unity 製 MV 制作ソフトに読み込むとプラグインで作ったキャラクターが 3D キャラクターとして現れるようにしている。さらに、例えば「足をトントンする」という動作は「C3」という音階に対応、「ふらふら揺れる」という動作は「D3」という音階に対応させるというように、「動作」と「音階」とを対応させている。『きせおんどール』プラグインを開いた DAW 上で音階を MIDI 入力すると、キャラクターが MV 制作ソフト上でその音階に対応した動作を行うようにしている。これによって、作曲者が普段 DAW で作曲するようにミュージックビデオを作ることや考えられると考える。このツールで、“作曲はできるけれど動画編集はできない。しかし曲に MV を付けたい”、“自分でキャラメイクしたオリジナルキャラクターを動かしたい”というニーズに応えることができると考えている。本研究は 2025 年度「北海道 IT クリエータ発掘・育成事業(新雪プログラム)」に採択され、“『きせおんどール』～「キャラメイク」で音色選びができる楽曲制作支援プラグイン・3DMV 作成ツール～”というプロジェクトで開発に取り組んでいる。2026 年 1 月 17 日(土)に行われる「新雪プログラム 2025 Demo Day 北海道 IT クリエータ発掘・育成事業 成果発表会」にて、『きせおんどール』の VST プラグインや MV 制作ソフトについて成果発表を行う予定である。

本研究は、「キャラメイク」で音色選びができる楽曲制作支援プラグインとして『きせおんどール』を提案し、一定の評価と今後の発展可能性を確認した。今後は、ユーザーニーズに応じた機能拡張と、3D キャラクター MV 制作ソフトとの連携を通じて、「曲とリンクしたオリジナルキャラクターが生まれ、そのキャラが 3D になって動く」という新しい創作体験を支援するツールへと発展させていきたい。

謝辞

本研究は、経済産業省 AKATSUKI プロジェクト北海道 IT クリエータ発掘・育成事業（新雪プログラム）の支援を受けている。

参考文献

- [1] タカラトミーアーツ. ひみつのアイプリ. <https://aipri.jp/>, 2024.
- [2] CAPCOM. モンスターハンターワイルズ. <https://www.monsterhunter.com/wilds/>, 2025.
- [3] Stephen Brade, Bryan Wang, Mauricio Sousa, Gregory Lee Newsome, Sageev Oore, and Tovi Grossman. Synthscribe: Deep multimodal tools for synthesizer sound retrieval and exploration. In Proceedings of the 29th International Conference on Intelligent User Interfaces, IUI '24, page 51–65, New York, NY, USA, 2024. Association for Computing Machinery.
- [4] 齋藤 創 and 大場 みち子. メタデータを利用した機械学習による dtm(desktop music)での音色づくりの効率化. In 研究報告音楽情報科学 (MUS) , volume 2016-MUS-110, page 1–6, 2016.
- [5] 吉岡 靖博, 石塚 慎也, and 齋藤 佳紀. 電子音色辞書における音色探索システムおよび3次元音色表示インタフェース. volume 2009, pages 53–58. 東京: 情報処理学会, 2009.
- [6] NativeInstruments. takt. Kon <https://www.nativeinstruments.com/jp/catalog/software/>, 2024.
- [7] Spectrasonics. Omnisphere. <https://www.spectrasonics.net/products/omnisphere/>, 2015.