シンセサイザを用いた音の創作に向けたインタラクティブな音色探索支援

Zefan Sramek[1,a)]　矢谷 浩司[1,b)]

概要：電子音楽用シンセサイザを用いると，ユーザは音色を構成する多数のパラメタを直接制御でき，多様な音を作り出すことが可能である。一方でこの提供されている自由度の高さはシンセサイザのインタフェースを複雑化させる要因となっている。本研究では，ユーザがシンセサイザを通して新たな音色を発見することを支援するインタフェースの構築を目指す。10 人のエレクトロニック・ミュージシャンへの定性的調査の結果，独自の音色を発見することには興味があるものの，プログラミングの知識の不足などにより，事前にシステム内にセットされている音色に依存していることがわかった。我々が提案するインタフェースはシンセサイザに付属して使用するもので，シンセサイザ内の個々のパラメタを直接制御する必要なく音色空間をユーザが探索できるようになっている，これによって，複数のパラメタの組み合わせによって生成される様々な音色を技術的な知識がなくとも即座に試すことができ，ユーザの新しい音色の発見につながることが期待される。本稿では，定性的調査の結果とシステムの実装を報告するとともに，今後の課題について議論する。

1. 序論

電子音楽用シンセサイザは，作曲とサウンドデザインの観点から音楽技術の領域を大きく占めている。シンセサイザとは，アナログ信号またはデジタル信号による電子的な手段によって音を作り出す楽器である。シンセサイザは，ボリュームやピッチのみならず，その音色も調節する能力をユーザに提供している点で他の楽器と異なる。これは，楽器の合成エンジンのパラメタを操作することで達成される。サウンド合成の種類は数多く，それぞれのタイプを実装する方法も多様であり，それによって様々なアーキテクチャを有するシンセサイザや多様なプログラミング方法が存在している。プログラミングはシンセサイザのプログラミング・インタフェースを使用することによって行われるが，これは合成エンジンと楽器のデザインによって様々である。一般的に，インタフェースによってユーザは合成エンジンのパラメタを調節できるようになる。普通，より複雑な合成エンジンの方が広範な音を生み出すことができ，それによってユーザがサウンドデザインを行う上でさらなる柔軟性を与えている。しかし，複雑性が増すことは多くの場合合成パラメタが増えることも意味しており，それによってプログラミングの難易度も高くなっている。

数百ものパラメタを有するシンセサイザは膨大な数の独特の音色を生み出すことができるため，楽器の音色空間を探索する作業は労力が大きい。これは特に，シンセサイザのインタフェースがユーザに対し一度にたった 1 つのパラメタの制御しか提供していなかったためである。その結果，ユーザは多くのパラメタを個別に調節して音を探索または洗練させなければならない。そのため，複雑な合成エンジンを搭載したより強力なシンセサイザはより多くの音を生み出す能力を有しているが，同時にその複雑性と，従来のままのインタフェースによって，限られたユーザビリティしか提供できない。シンセサイザの発明以来合成エンジンは大きく進化してきた一方で，そのインタフェースは進化していない。本研究では，このようなインタフェースの設計における発展の可能性に挑戦することを目指す。シンセサイザの基本のインタフェースは，ユーザがシンセサイザ上で一度に 1 つのパラメタしか扱えないこと（ミクロレベル），または既定の音から選択することを強いられる（マクロレベル）ものである。ユーザが複雑なシンセサイザの音色空間を探索し，パラメタを 1 つずつ編集するという負担なしに独自の新しい音を作り出せるようなミクロとマクロの中間レベルに位置するインタフェースを本研究を通じて実現する。そのような方法が確立すれば，ミュージ

[1] 東京大学大学院 工学系研究科
Interactive Intelligent Systems Laboratory,
Graduate School of Engineering，The University of Tokyo
7-3-1 Hongo，Bunkyo-ku，Tokyo，Japan
[a)] zefanS@iis-lab.org
[b)] koji@iis-lab.org

© 2021 Information Processing Society of Japan

シャンが強力な合成エンジンへより良くアクセスできるようになり，独自の創作やサウンドデザインに活用できることが期待できる。

2. 関連研究

本節では，様々な技術的または人工知能的な手段を用いて一般的なミュージシャンの創作活動と音楽的な目標の達成を支援しようとしているいくつかの研究プロジェクトについて見ていく。この領域における主な課題はマッピング，すなわちユーザによるシステムへの入力値と，音または演奏を制御するパラメタの接続である。

従来のシンセサイザのアーキテクチャでは，すべてのマッピングが一対一で行われる。すなわち，合成エンジンの各パラメタが独自の専用インプットにマッピングされている。しかし，Hunt と Wanderley [3] が指摘している通り，伝統的なアコースティック楽器の場合はこのようなことは滅多に見られない。

代わりに，アコースティック楽器では一般的に少数のインプットしかなく，これらのインプットは非線形的に数多くのパラメタに同時にマッピングされ，楽器の音に影響している。アコースティック楽器にヒントを得た Hunt と Wanderley による複雑なマッピング戦略の実験からは，一対一ではないマッピング戦略がユーザにとってより魅力的である可能性が示された。しかし，特に複雑な合成エンジンの場合は何百ものパラメタを少数の入力値にマッピングすることが考えられることを踏まえ，特定のユースケースに応じてこのようなマッピングをどのように行うのが最善であるかは研究課題として未解決である。

Constantini ら [2] は，人工ニューラルネットワークによるアプローチを用いてこのマッピングの問題に取り組んでいる。彼らは，マルチパラメタ・シンセサイザのインタフェースを 2 つのパラメタによるパフォーマンス・インタフェースにマッピングし，様々な音の合成とそれらの間における遷移を簡素化することを目指している。プロトタイプでは 2 次元の面を入力値として用い，X と Y の座標を 2 つのパラメタとして使用している。マッピングを達成するために，作曲者（システムのユーザ）は平面上で n 点を選び，それらの点に対して関連するシンセサイザのパラメタを設定する。

次に，これらの点を用いて，ニューラルネットワークを設定し訓練する。一旦訓練した後は，ネットワークは平面上のすべての点についてパラメタを計算でき，ユーザが設定した点の間で複雑な補間を行っている。ネットワークでは順伝搬・誤差逆伝搬型ニューラルネットワーク構成を使用している。その結果，これらのネットワークは音と音の間で複雑で非線形な関係を補間できる。Constantini らは，そのような複雑なマッピングが従来の一対一のインタフェースよりも音楽的な表現性を支援できると結論づけている点において Hunt と Wanderley と一致している。しかし，彼らのシステムは未だ 2 次元のマウスベースのインタフェースに制限されており，さらなる研究として，このニューラルネットワークによるアプローチを音楽的に妥当なインタフェースへ適応させることが考えられる。

全体的な視点からは，この研究の大半を，合成システムの複雑性を軽減させる試みとして捉えることができる。大量の合成パラメタを用意していることは，技術的またはサウンドデザインの観点からは有利である可能性があるが，音楽的な観点からは，ミュージシャンが目標を達成することを妨げたり，コンピュータミュージックシステムを使用することを完全に断念させてしまう恐れがある。

現に，サウンド合成における新たな技術によって音の生成と操作のためにより高度な手段がもたらされているが，アコースティック・ミュージシャンや新規ユーザにとっては引き続き近づきにくいものであると Scurto と Bevilacqua [7] は主張している。特にコンピュータミュージックの領域では，新しいシステムを学習する上でシステムの審美的な理解を発達させるだけでなく，技術的な理解も求められることから，新規ユーザには重い認知的な負荷がかかる。そのため，Scurto らは，ユーザの探索を支援する人工知能エージェントを開発している [7]。エージェントはシンセサイザのプログラミングを担当するアシスタントの役割を担い，人間がその結果について知覚的なフィードバックを提供する。強化学習を通じて 2 つのエージェントが協働し，人間のユーザが知覚的に納得のいくものと判断できる音に近づけていく。そのため，最初の探索段階ではユーザはシステムの技術的な面に関する認知的な負荷から解放され，審美的・知覚的な理解を発達させることに集中できる。残念ながら，継続中のこのプロジェクトでは大規模なユーザスタディはまだ実施されていないが，予備調査ではこの共同探索モデルの有用性が示されている。

本節では，様々なコラボレーション・システムを構築することでミュージシャンを支援するいくつかの試みについて議論した。これらのシステムでは人工知能やデータマイニングを活用して，新たなマッピング戦略や，コンピュータミュージックシステムの技術的な複雑性を軽減させるその他の手段を提供している。最終的に，本研究では，複雑なツールを使用してミュージシャンがその創造性に集中し，音楽的な目標を達成することを支援する新たな手段について調査する。この発想は，筆者らの予備調査の方向性の基礎を形成するものである。

3. インタビュー調査

まず，エレクトロニック・ミュージシャンによるシンセサイザに関するこれまでの経験についてより深く理解することを目指し，インタビュー調査を実施した。筆者らのシステム開発では，ユーザが抱えている真の悩みに対処して

© 2021 Information Processing Society of Japan

いなければ，彼らが筆者らのシステムを使う動機はない。筆者らは，シンセサイザ全般の使用に関するミュージシャンの経験や，音を選択しプログラミングすることに関する経験について知りたいと考えた。彼らが自身の音をプログラム化していない場合，その理由を知りたい，自身の音をプログラム化している場合，その体験がどのようなものであり，何が彼らにとって不満であったか知ろうとした。また，シンセサイザ愛好家は，単なる楽器またはサウンドデザインツールとしてシンセサイザを単に利用しているミュージシャンとは違う経験をしている可能性があるため，特有の性質も知りたいと考えた。

3.1 調査手法

本研究では，半構造化インタビューによるアプローチを採用した。まずは標準的な質問群から始め，参加者が自身の体験について語り始めた後，参加者の個々の体験について洞察を得られるように質問した。インタビューは参加者によって 45 分から 90 分までの時間を要した。インタビューは Zoom 上で日本語または英語で行った。参加者には 2000 円相当の報酬を支払った。

3.2 参加者

インタビュー調査には 10 名の参加者を募集した。すべての参加者はシンセサイザを使用した経験がある現役のミュージシャンであった。参加者に関する情報を表 1 にまとめている。

3.3 分析

各インタビューでは音声を録音し，その内容は分析のために転写した。その次に，筆者らの研究における主な焦点がシンセサイザのプログラミング・インタフェースであることを念頭に置き，オープン・コーディングの手法を使ってテーマ分析 [1] を実施した。インタビューはまずコード化し，その後コード化した引用をテーマ別に整理した。最後に，コード化した結果から，インタフェースの設計に関するガイドラインを生成した。筆者らが用いた手法には [4] も参考にしている。

3.4 結果

参加者から様々な体験談を収集し，それらからシンセサイザのプログラミング・インタフェースを設計するためのガイドラインを導出した。ガイドラインの中身どれも複数の参加者が言及していたことであった。全ての参加者は使いやすかった機器に関する肯定的な経験と，使いにくかった機器に関する否定的な経験の両方について語った。また，それらの体験についてなぜ肯定的または否定的に感じたのかについても述べてもらった。以下の節では，ガイドラインについてより詳細に述べる。

3.4.1 ミュージシャンは新しい音を求めているが，もっとプログラミングしたいと必ずしも思っているわけではない

このガイドラインは，今回のインタビュー調査で得られた最も重要な発見と言える。ミュージシャンが自分の音楽のために新しい音を求めていることと，音のプログラミングを求めていることを区別することが重要である。新しい音や独特の音を求めている人は，より高度なプログラミング技術やゼロからの音作りでそれを実現しようとするのは当然のことのように思えるものの，実際にはそうではない。実際，この 2 つは多くのミュージシャンにとって大きく分離している要素のようである。このことは，複数の参加者による回答にも表れている。

P6 は自身のプロセスについて，次のようにはっきり述べている。「私のプロセスは最初から，実験とか，電子音楽に対する理解がまったくなくて，何も調べようとしなかったことから生まれたものなんです。この音はどうやって出すのか，とか，このシンセはどうやって動かすのか，そういったことは考えていませんでした。私はただ電子音楽をやりたいだけで，自分が考えつく最も直接的な方法でそれをやろうとしているのです。」言い換えると，彼の目的は作曲することであって，シンセサイザをプログラミングすることではないのである。P5 がこの発想を最もはっきり表現しており，音を創作する過程と作曲の過程を明確に対比させている。「家を建てる前に，金づちを作っているみたい... いい木を削って，何年もかかって，家を建てるみたいな。なんか，時間がかかる。なかなか家に住めないというか... 最初から自分で作るという（のは）ちょっと目標が違う... 一番は自分の作った音楽を聴きたくて... その，過程が長すぎる，前の。」

3.4.2 個々のパラメタをいじるよりも，プリセットを使って探索できる

このインタビュー調査から得られたもう 1 つの重要な洞察は，参加者の多くがゼロからの大幅なプログラミングは行わなかったものの，独自の方法で音の探索をしていたという点である。多くの場合，そのためにプリセットが用いられていた。多くの参加者が，編集や探索の出発点としてプリセットを活用していると述べている。P1「通常はプリセットから始めますね。気に入ったプリセットを見つけて，それをいろいろと変更してみるんです。」P2「私のサウンドデザインの多くは，ただプリセットをつかってそれをいじるだけです。」他の参加者は，まずプリセットを直接使い，それにエフェクトやその他の処理を施すことによって，求めている音に仕上げている。P5「基本はプリセットでエフェクターとか EQ を触って，編集して，音を作っていく。もう，最初から作るということはまずやらないですね。」P6「ストックされている音やプリセットとかに外部エフェクト処理を施すことで，プログラミングする必要性

© 2021 Information Processing Society of Japan

表 1: インタビュー調査の参加者

#	年齢	ジェンダー	居住地	言語	音楽的な役割	経験年数	経験レベル
1	30-39	M	カナダ	英語	プロデューサー	10+	上級者
2	30-39	M	米国	英語	作曲家，プロデューサー，DJ	10+	上級者
3	20-29	M	日本	英語	キーボーディスト	0-5	初心者
4	40-49	M	日本	日本語	プロデューサー	5-10	初心者
5	20-29	M	日本	日本語	作曲家	0-5	初心者
6	20-29	M	米国	英語	プロデューサー	5-10	中級者
7	30-39	F	日本	日本語	作曲家，ボーカリスト	0-5	初心者
8	30-39	M	日本	日本語	作曲家，DJ	10+	中級者
9	20-29	M	日本	日本語	作曲家	5-10	初心者
10	20-29	M	米国	英語	プロデューサー，サウンドデザイナー	5-10	中級者

を一切回避するというのが私のやり方です。」これらの観察結果は，ミュージシャンがどのようにシンセサイザと関わり，どのように新しい音を探索し生成しているかを理解する上で不可欠である。多くの参加者が多種多様なプリセットから始めるという同じようなワークフローを説明しているので，新しいインタフェースの設計では，この点を考慮することが重要である。

3.4.3 プリセットについては，それだけでは十分とは見なされないことが多い

多くの参加者がプリセットを使って音を探索していたことを踏まえると，そもそもなぜ高度な音の探索・制作ツールが必要なのかという疑問が湧いてくる。膨大な数のプリセットを用意して，ユーザに選んでもらう方が簡単なのではないか？参加者はこの点についても言及しており，独特で適切な音を求めるためには，プリセットでは不十分なことが多いという意見を示していた。

P2 は，Waldorf Blofeld の使用と，プリセットが彼の音楽スタイルにそぐわないことからあまり使用しないことについて述べている。「私があまり使っていなかったのは，どのプリセットもメタルのようなゾーンからはじまり，それを興味深い音にするためには何をどうすればよいかしっかり知っていないといけないからなのかもしれません。」また，シンセサイザのプログラミングの複雑性により，求めている種類の音を作ることが難しく，プリセットに不満があったため，あまり使わなくなったことを語った。

P7 は，ソフトウェア・シンセサイザのプリセットについて，新しい音やプリセットがないことは，その機器に対する興味を次第に失ってしまいうることを意味することを指摘している。「いくつか Ableton の音色を使って，曲を作っていたんですけど... 入っているものも，私が持っている Ableton のやつはそこまで多くないので，音色が。なので作っていると... 自分が好みの音色がだいたい同じなので，そこが毎回同じように使っていて，またこれ使っているなとか，そういう，もっといろんな音色を使えた方がいいなっていう。なんか，似てるというわけじゃないけど，似たようなのばっかり使っちゃうのでそれで OmniSphere の方に変えましたね。」魅力的であった Ableton のプリセットの数が限られていたことは，彼女にとって，同じ種類の音をいつも使っているように感じられることを意味していた。さらに幅広い音を使いたいという願望から，彼女はプリセットのライブラリが膨大であるソフトウェア・シンセサイザのパッケージ Omnisphere を使うようになったのである。プリセットの数が少ないと，ユーザが自分で独特の音を探索したり作ることができる能力が制限され，新しい音源を求めるようになるのではないかと考えられる。

P9 も同じ意見を述べ，彼が使っているプリセットの質を称賛しつつも，実際にはその多くが彼にとって魅力的ではないとも言っている。「プリセット（は）よくできていると思う。全然僕が使わない感じのテイストのプリセットもいっぱいあるけど，これがほとんどだけど。」P10 は最初からプログラミングを自分で行うことが多く，プリセットを出発点として使うことも稀にあるものの，その場合は彼が望む音を実現することがより難しくなると語った。「プリセットの音とかをちょっと変更する使い方とかもたまにするんですけど，やっぱり最初から作ったほうが自分のイメージした音に近いものを作ることができるなという気がしているので，いつもは波形の段階から作っていくようにしていますね。」これは，音を細かく制御できることで，他の参加者が経験した不満を一部解決できる可能性があることを示唆している。

3.4.4 外部エフェクトの処理はシンセサイザのパラメタをいじることと同等，またはそれ以上に重要

多くの複雑なシンセサイザは独自のパラメタを通じてサウンドデザインに十分な機能を提供しているものの，ほぼすべての参加者がシンセサイザの出力に外部エフェクトを適用することによって音をさらにカスタマイズしていると述べた。実際，一部の参加者にとって，これが独特の音を作る主な方法であった。P6 はこの方法を特に活用してい

© 2021 Information Processing Society of Japan

る。「ラフなプリセットから始めて，それを少しいじって，それから録音します。そこから，サウンドデザインの 90 パーセントはエフェクトペダルを使うか... 3 つか 4 つのソフトウェアを経て，それからペダル，サンプラー，そんな感じです。そうやって独自のサウンドを生み出しているんです。」これらの洞察は，ミュージシャンのワークフローを理解する上で重要である。他の処理技術の使用を除外するデバイスを開発することを期待するのではなく，既存のパラダイム内に適合することを目指すべきである。

3.4.5 プログラミングおよび探索の過程は素早く，流動的で，即時的であるべき

数名の参加者がシンセサイザのプログラミングに関する体験について，重要な要素としてスピードや流動性を挙げていた。P2 は，素早く作業したいという願望をはっきり述べていた。「即時性と言えばいいのでしょうか，直感とか，何か創造性のスピードを高めてくれるもの，素早く作業できることが私にとってとても大事なのです。」P5 も，サウンドデザインの詳細にこだわるのではなく，素早く作業を進めたいという願望を表していた。「時間かけたくないというか，そこまで音にこだわってないって言うか... ちょっとスピードは速くどんどんやりたいタイプ。」したがって，筆者らは，音の創作過程においてスピードまたは流動性を高めることを重視することにした。

3.4.6 タンジブルインタフェースは重要で，価値を高める

すべての参加者がハードウェアとソフトウェアの両方を作曲活動で活用していたが，複数の参加者が，ソフトウェアのインタフェースだけではなく触覚的なインタフェースを使うことが彼らの体験にとって重要であると指摘していた。一部の参加者は，触覚的なインタフェースの方が面白い，刺激的，または魅力的であることを強調していた。P1 が述べているように，「ソフトのシンセをプログラミングすることはあまり楽しくないです。 ノブをいじっているときの触感や，テーブルまわりで手を動かしているときに自分が何をやっているのか分かっているという感覚が欠けている気がするのです。」また，P2 も次のように述べている。「見た目がよくて，いじるのが面白いノブが付いていて，そのノブでかっこいい音が出せるなら，それを間違いなくもっと使うと思います。」P4 は面白さと，音をダイナミックに変える能力の両方について触れている。「兄の方がいっぱいノブが付いているアナログシンセを 1 台持っているんです。それをいじってたら楽しいですね。簡単に音が変化するし，ダイナミックにギュイーンってツマミを回していると面白くなってますね。」これらの結果は，多くのミュージシャンにとって，触知できる身体的なインタフェースがユーザ体験の重要な部分であることを示唆している。

3.4.7 いくつかの重要なパラメタ／制御はいつでもすぐ使えるようになっているべき

また，一部の参加者からは，重要なパラメタや制御をいつでも利用できるようにしてほしいという意見もあった。これらは，フィルターのカットオフのように，音にすぐに明確な影響を与えるパラメタであることが多かった。P8 は音全体に影響を与えることができる制御が欲しいと述べている。「全体に対してカットオフやレゾナンスがあったらいいのにと思います。何があっても音全体を操作できるようなものがあれば...」同じように，P2 もすぐに制御できるようにしたいというと考えている。「理想としては Digitone のようなものですが，再生可能なフィルターとかそういったもののための専用ノブがもっと付いてほしいです... ボタンを 2 つ押すなどしてメニューダイブしてフィルターに行くことはできますけど，そんなことはしたくないです...」言い換えれば，筆者らのシステムがサウンドデザインや探索のための新しい手法を提供できたとしても，ミュージシャンはいくつかの重要なパラメタに簡単にアクセスする必要があるのである。そのため，その可能性を排除しないことが重要である。

3.4.8 たくさんのパラメタに簡単にアクセスできることは必ずしも役に立つわけではない

筆者らの主な仮説の 1 つが「パラメタへすぐにアクセスできないことにより，プログラミング体験を難しくもどかしいものにしている」というものであったが，多くの参加者が，パラメタが多すぎるとアクセスしやすくても圧倒されたり苛だったりすると話していた。300 以上ものシンセサイザ・パラメタを搭載しているデジタルシンセサイザである Roland D-50 について，P1 は安定的にプログラミングできないことが多かったことを指摘している。「何が起こるのか正確に予測できるほどよく理解できていません。なので，少し調節してみて，別のパラメタも調節してみるうちに，音がおかしくなってしまうのです。」特に注目したのは，シンセサイザの個々のパラメタに物理的な制御を加え，表向きにはパラメタがアクセスしやすくなっている PG-1000 Programmer を使用している点である。しかし，彼はその使用が引き続き困難であることを指摘している。「... スライダーが 75 個くらいあるんです ... すごいですよね... 音をプログラミングするときは，なんというか，常にシンプルであることが大事だと思います。なので，それでプログラムを組もうとするときも，あまり期待していません。」多くのパラメタに直接アクセスできるにもかかわらず，P1 はプログラミングのプロセスを理解するのに苦労しており，細かい変更を施すにとどめていた。他の参加者も同様の困難について述べており，P4 は次のように語っている。「いっぱいノブがあると，どこがオシレーターでフィルターで，エンベロープでとかは全然わからなくて。最初はわからなかったですね。 」全体的に，パラメタの数

© 2021 Information Processing Society of Japan

が多いと，参加者は苛立ったり混乱するようである。このことは，ミュージシャンにすべてのパラメタを直接制御する機能を与えるだけでは，新しいインタフェースの設計はうまくいかないことを示している。確かに単一のパラメタを選択して編集するインタフェースは創作をつまらなくする可能性があるが，大量のパラメタを同時に操作できる場合は，ユーザが圧倒され混乱してしまう可能性がある。

3.4.9 創造性と効果的に目標を達成するためには制限は重要である

同様に，何人かの参加者は，創造的なプロセスにとって制限が重要であると述べている。P1 は次の通り説明している。「制限があると助かります。既に作っておいたプリセットを設定すればよいだけで，それがカッコいい音だということは既に分かっているんです。それからフィルターをかけるとか，そういうことをいくつかできます。ひっきりなしに延々といじっていれば何も決められないですから，それがあると楽になります。」P4 も，機器の制限の中で作業することが楽しみの 1 つであると述べている。「一番好きなシンセでいえば，やっぱり TG33 は好きですね。フィルターもないし... 結構自由とかできないので，あんまり自由度とかは少ないんですよ... でもそういう制限がある中で面白い音を作るのが楽しいからですからね。」創造的なプロセスには制限が重要であることを理解することは，新しいサウンド創作ツールの設計にとって大事な洞察である。パラメタの数が多すぎると圧倒されてしまうのと同様に，創作における制約はユーザの混乱を防ぐだけでなく，ユーザの創作意欲を喚起することができる。

3.5 考察

筆者らによるこれまでのシンセサイザ開発のレビューや現場での経験からは，高レベルと低レベルの音色操作の中間に位置する音色探索やサウンドデザインツールの必要性が示唆されていたが，本稿で紹介する研究では，シンセサイザ・プログラミング・インタフェースを設計する上で焦点となるいくつかの重要なポイントが明らかになり，設計に向けたガイドラインが完成した。これまでに多くの研究者がシンセサイザのプログラミングや演奏のための新しいインタフェースを開発してきたが，本研究は特にミュージシャンの経験やニーズから出発しており，単に斬新なだけでなく，ミュージシャンの既存のワークフローに適合したインタフェースの開発に向けた重要な一歩となっている。インタビュー調査からは，ミュージシャンのワークフロー，音の創作や選択において彼らが直面する問題，そして彼らと機器との間にある関係性について重要な洞察が得られた。特筆すべき点として，参加者の間では様々なワークフローや音の創作に関する好みが見られた。デザイナーやエンジニアとしてツールやシステムを設計する際には，ユーザが単一の用途や技法に従うことを期待してしまいがちだが，筆者らの結果は，そうではないことをはっきりと示した。さらにもう 1 つの重要な洞察として，参加者は全員が音楽を作曲する際に適切な音を模索していたものの，それを達成する方法が極めて多様であった。シンセサイザはユーザが直接音をカスタマイズすることを可能にするように設計されているものの，多くの参加者がそれをせず，別の（時には癖のある）方法を好んでいた。これらの点をインタフェース設計へ取り込むことは重要となる。

3.6 本事前調査に関する注意点

参加者が 10 名しかいなかったため，ミュージシャンがどのように作業を進め，どのような好みを持っているかについては限られた範囲でしか把握できなかった。音楽の作曲は高度に個人的で癖のある過程であるため，参加者の意見を調整することは難しい。筆者らは複数の参加者が言及したテーマに基づいてガイドラインを特定することのみを試みたものの，1 つのテーマについてすべての参加者が触れたことは稀であり，さらに一部の参加者は，熟練度や個人的な好みの性質により，自然と矛盾した経験をしていた。さらに，10 名の参加者のうち，女性はたった 1 名であった。そのため，筆者らの結果にはジェンダーバイアスがかかっている可能性がある。ジェンダーの観点からバランスの取れた研究を行うことで，バイアスのかかっていない結果をより多く得ることができる可能性がある。また，今回の参加者は北アメリカまたは日本に在住している者たちであるため，文化的または言語に基づいた影響があるとすれば，筆者らの結果は限定的なものである。最後に，筆者らによるインタビューの長さと方法が限られていたため，参加者の経験を事細かく明らかにすることが不可能であった。大規模な研究を行うことで，これらの多くを補える可能性が高い。

4. システムの概要

第 1 節で述べた課題に対処しつつ形成的研究の結果を考慮するために，プロトタイプシステムを開発した。システムはシンセサイザそのものではないが，MIDI [5] を経由して既存のシンセサイザ [*1] に接続するモジュールであり，ユーザにとって新たなプログラミング・インタフェースを提供する。すべての合成パラメタの操作は MIDI システム・エクスクルーシブ・プロトコルを使用して行われ，ユーザは元のインタフェースに依存することなくシンセサイザを制御することができる。基本的なインタフェースは，カラー LCD スクリーン，ジョイスティック，いくつかのノブとボタンで構成されている。中心となるインタフェース要素はジョイスティックであり，これはユーザがシステムを触ってすぐに操作できる方法として選ばれた。このイ

[*1] 筆者らのプロトタイプには Roland D-50 を採用している。

© 2021 Information Processing Society of Japan

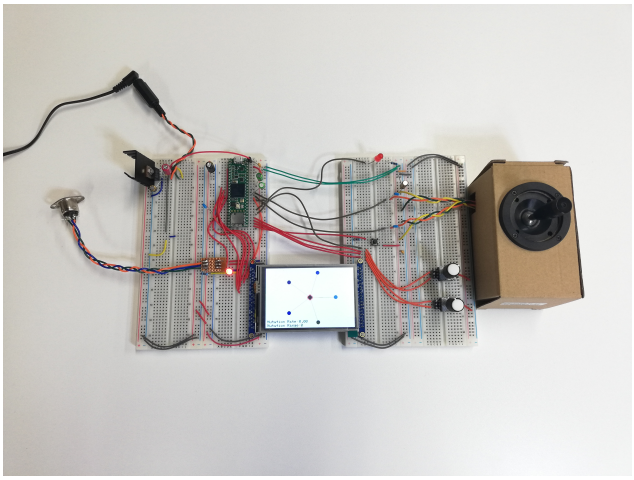

図 1: 最初のプロトタイプデバイス。

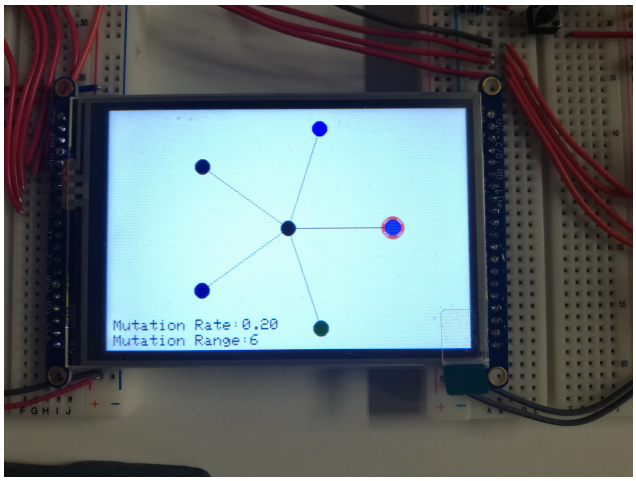

図 2: 進化モードのグラフィクスの詳細

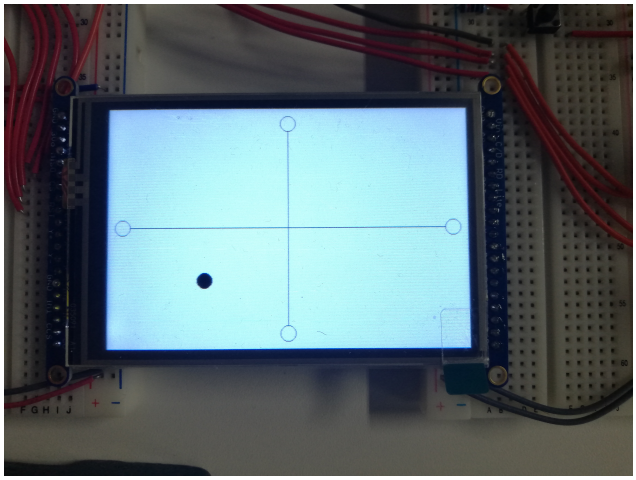

図 3: モーフィング・モードのグラフィックの詳細

ンタフェースの目的は，ユーザが自分の音楽のために独特の音色を探索し，作り出すためのツールを提供することである。形成的研究で明らかになったように，ミュージシャンは自分の曲に独特の音を求めているものの，その多くが自分の音をゼロからプログラミングする動機や専門知識を持っておらず，特に合成パラメタの数が非常に多い複雑なシンセサイザに圧倒されていた。筆者らによるシステムでは，ユーザが音色をより直感的に探索できるように個々のパラメタを抽象化し，代わりに音色を音や他の音色との関係性の観点から概念化することに重点を置いている。

特に，筆者らのインタフェースには音色探索のために 2 つのモードを搭載している。進化的探索モードでは制御したランダム性を用いて，ユーザが選択した単一のシード音色に基づいて音色のコレクションを生成し，モーフィング型探索モードでは，4 つの選択した音色の間でユーザが音色空間を探索できるようになっている。両モードとも，合成エンジンの各パラメタを次元として捉えることによって構成されている，高度に多次元な空間である音色空間の概念を採用している。この空間では，類似した音色は，異なる音色よりも近い距離にある。この概念化により，ユーザが音色領域全体または複数の音色の交差領域を探索することが可能になり，音色の関係性をパラメタではなく音に根ざしたものにしている。以下の節では，各モデルについて詳しく紹介する。

4.1 進化モード

進化モードでは，1 つのシード音色から子音色 [*2]群を生成することで，ユーザは音色空間を反復的に移動することができる。すべての子音色は制御されたランダムな突然変異によって生成され，それぞれの子音色について，どの合成パラメタがどのように変異するかは，システムがランダムに決定する。ただし，ユーザは変異を制御するために，変異させるパラメタの比率と，変化させる範囲という 2 つのパラメタを設定する。このようにして，ユーザは，個々の合成パラメタを抽象化しつつ，結果として得られる音色がシード音色とどれだけ似ているか，あるいは異なっているかを決定することができる。したがって，ユーザは 1 つのプリセット音色から始め，進化過程を制御する 2 つのパラメタの値を決定する。ここからシステムは 5 つの子音色を生成し，ユーザはジョイスティックでそれぞれの子音色を選択して試聴することができる（図 2 を参照）。音色を選択すると，対応するパラメタのデータをシンセサイザに送信し，ユーザが各子音色を素早く試聴できる。ここからは，子音色の中から任意の音色を選び，次の世代の子音色のベースとなる新しいシード音色として，このプロセスを繰り返すことができる。反復を繰り返すことで，ユーザは個々のパラメタを直接変更することなく音色空間を移動し，元のシード音色から次第に異なる音を生成することができる。このようにして，インタフェースはユーザと聴こえてくる音との間により直接的な関係性を築く。

[*2] ‘offspring timbre’

© 2021 Information Processing Society of Japan

4.2 モーフィングモード

モーフィング・モードでは，4 つの音色が交差している音色空間を自由に探索することができる。4 つの音色のパラメタ値は，音色の交差空間を表す 2 次元平面上で，ユーザが選んだ位置に関連して補間される（図 3 参照）。各音色の影響度（重み）は，平面上の位置の近さによって決まる。したがって，ユーザがカーソルを特定の音色に近づけると，その音色がよりはっきりと聞こえるようになる。ジョイスティックを使って，ユーザはモーフィングの位置を流動的に制御でき，位置が選択されると，パラメタのデータがシンセサイザに送信される。100 × 100 ポイントの解像度では，たった 4 つの音色が交わるだけで，特定の領域に大別できる非常に多くの新しい音色が生まれる。ユーザは空間を探索することで，これらの領域を発見し，より正確に音色を磨き上げることができる。最初の 4 つの音色を慎重に選択し，結果として得られる音色の交点を探ることで，ユーザは個々のパラメタを変更することなく，多種多様な独特の音色にアクセスできるようになる。ユーザがジョイスティックを使って音色を探索し，その体験がそのまま音に結びつくことで，音色探索の過程がより直接的なものになる。

5. 実装

5.1 ハードウェア仕様

プロトタイプは ARM Cortex M-7 マイクロプロセッサをベースとした Teensy 4.1 マイクロコントローラ [8] を中心に構成されている。すべてのファームウェアはマイクロコントローラ上で実行しているため，プロトタイプはスタンドアローンのシステムになっている。マイクロコントローラはユーザに GUI を表示するために 3.5”LCD カラーディスプレイに接続しており，ジョイスティックやロータリーエンコーダ，ボタンなど，すべてのインタフェース要素からの入力値を読み取る。シンセサイザとの接続は，標準的な 5-pin DIN 式 MIDI インタフェースを通じて行われる。

5.2 データの送信

すべてのデータは，電子音楽器のために開発されたシリアルデータ送信プロトコルである MIDI 経由でシンセサイザに転送した。筆者らのプロトタイプは MIDI システム・エクスクルーシブ（Sysex）メッセージを用いてシンセサイザのパラメタを遠隔制御する。これらのメッセージはマルチバイトのコマンドストリングであり，特定のパラメタを提供した値に設定するようにシンセサイザに指示を出す。その名前が示唆しているように，これらのコマンドストリングはメーカーと対象となるシンセサイザにとって専用である。Sysex の実装に関する情報は D-50 MIDI 実装文書 [6] に記載されている。

5.3 アルゴリズム

5.3.1 進化モード

4.1 節で述べたように，進化モードでは制御されたランダム性を用いて子音色を生成する。まず，ユーザが設定した変異の比率を用いて，対応する数のパラメタを全パラメタの一覧から無作為に導出する。次に，各パラメタにつき，元の設定値に対してユーザが設定した変異の範囲内で乱数化する。この過程を個別に 5 回繰り返すことにより，5 つの子音色を生成する。

5.3.2 モーフィングモード

4.2 節で述べたように，モーフィング・モードでは，図 3 に示した平面上の対応する距離を決定することによって，4 つの元の音色のそれぞれの影響する比率が決まる。結果として得られる音色の各パラメタは，4 つの元の音色の対応するパラメタの加重平均を計算することによって決定する。X-Y 座標から対応する 4 つの重みに変換する作業は，明確な方法がないという意味で容易ではない。今回のケースでは，4 つの音色に対応する 4 つのポイントのそれぞれが，純粋にその音色だけを生み出すことが重要だった。つまり，その音色に対する重みを 1 とし，他のすべての重みを 0 とすることである。

筆者らのアルゴリズムは以下のように動作する。まず，現在のカーソル位置と 4 つの点のそれぞれの間の距離が計算される。次に，この距離の逆数を，4 つの固定点のいずれかの間の最短距離に基づいて正規化する。次に，重みが 0 以上になるように固定される。最後に，得られた 4 つの重みをすべて 0〜1 の範囲で正規化する。この方法では，純粋な単一の音色に対応するすべての点が正しい重みを持ち，中心点などの様々な重要な点も正しい重みを持つという望ましい結果が得られる。

6. 結論と今後の研究

本論文では，インタラクティブな音色探索システムの開発に関するインタビュー調査の結果について述べた。ミュージシャンが創作にあたりどのようにシンセサイザを使い，音色探索に関する経験がどのようなものかをより良く理解するために，インタビュー調査を実施した。調査結果を用いて，今後のシンセサイザ・プログラミング・インタフェースの開発のためのガイドラインを作成した。筆者らの研究により，ミュージシャンが音楽を創作するワークフローにおいてどのようにシンセサイザを活用しているか強固な洞察を得ることができ，彼らのニーズに対応できる可能性が高いソリューションの開発が可能になった。

その次に，インタビュー調査による結果に基づいて設計した音色探索インタフェースのプロトタイプについて述べた。個々のパラメタを編集したりプリセットを試聴するのではなく，筆者らによるシステムでは，ユーザが局所的な領域において音色空間を定義し探索することが可能であ

© 2021 Information Processing Society of Japan

る。これは，モーフィング型と進化型という 2 つのモードによって実現した。モーフィングモードでは，ユーザは 4 つのプリセットした音色の交差で領域を定義し，その領域を探索することができる。進化モードでは，ユーザは単一のプリセットした音色から始め，音色空間を移動する反復的な過程によってバリエーションを生成する。両方のモードにおいて，個々のパラメタがどのように変化しているのかを理解する必要なく何百もの独特の音色を生成することが可能である。

今後の研究では，より完成度の高いプロトタイプを実装し，複数のミュージシャンを対象にユーザスタディを実施する予定である。参加者には各自の音楽のプロダクションにおいてプロトタイプを使用する機会を与え，筆者らが目標をどれだけ達成できているか評価してもらう。この長期的な調査により，筆者らが提案する新たな音色探索パラダイムがどのようにしてミュージシャンによる作曲に影響を与えるかについて情報を収集することができる。筆者らは，本研究がミュージシャンの活動に好影響を及ぼし，独特な音を産み出す新たな音源をもたらすだけでなく，彼らの音楽にとって新たなインスピレーションにもなる可能性があると考えている。

謝辞

本研究，および本論文に助言を頂いた Arissa J. Sato，Carla F. Griggio，下島銀士，桝田拓磨，および IIS Lab の研究員の皆様に感謝申し上げます。また実験参加者の皆様に感謝申し上げます。

参考文献

[1] Braun, V. and Clarke, V.: Using thematic analysis in psychology, *Qualitative Research in Psychology*, Vol. 3, pp. 77–101 (online), DOI: 10.1191/1478088706qp063oa (2006).
[2] Costantini, G., Todisco, M. and Carota, M.: A Neural Network Based Interface to Real Time Control Musical Synthesis Processes (2007).
[3] Hunt, A. and Wanderley, M. M.: Mapping performer parameters to synthesis engines, *Organised Sound*, Vol. 7, No. 2, p. 97–108 (online), DOI: 10.1017/S1355771802002030 (2002).
[4] Maudet, N., Jalal, G., Tchernavskij, P., Beaudouin-Lafon, M. and Mackay, W. E.: *Beyond Grids: Interactive Graphical Substrates to Structure Digital Layout*, p. 5053–5064 (online), available from ⟨https://doi.org/10.1145/3025453.3025718⟩, Association for Computing Machinery (2017).
[5] MIDIAssociation: MIDI, `https://www.midi.org/` (2021). Accessed: 2021-05-07.
[6] Roland: D-50 MIDI Implementation, `https://archive.org/details/synthmanual-roland-d-50-midi-implementation` (2012). Accessed: 2021-05-07.
[7] Scurto, H. and Bevilacqua, F.: Appropriating Music Computing Practices Through Human-AI Collaboration, *Journées d'Informatique Musicale (JIM 2018)*, Amiens, France, (online), available from ⟨https://hal.archives-ouvertes.fr/hal-01791504⟩ (2018).
[8] Stoffregen, P. J.: Teensy USB Development Board, `https://www.pjrc.com/teensy/` (2021). Accessed: 2021-01-21.

© 2021 Information Processing Society of Japan