

MoSE : Wekinator を用いたモーションコントロールによるワンチップCPUギターエフェクタの開発

毛利 数馬¹ 平野 砂峰旅² 片寄 晴弘¹

概要: ライブ演奏においてギタリストの多くがエフェクタによる音色の操作を行うが、通常用いられるペダルやノブでの操作では視覚的な演出が制限されるという課題がある。本研究ではステージでの視覚的な演出を考慮し、モーションによりエフェクタの切り替えやパラメータ制御を行うエフェクタ、「MoSE」を作成する。ライブ演奏用途においては、ミュージシャンの思考の妨げとならない UI の提供と時間的に厳格なジェスチャ認識が求められる。そこで、ジェスチャの予備動作への着目や Wekinator の利用によって、これらの課題の解決に当たる。

1. はじめに

ロックやポップスのライブにおいてパフォーマーは、楽器の演奏だけではなくエフェクタによる音色の変化、さらには、誇張した演奏動作や必ずしも演奏動作とは関係ない身体表現のような視覚的パフォーマンスによってステージを盛りあげている。視覚的パフォーマンスと出音を一致させることはパフォーマンスにおける重要な要求事項であるが、通常エフェクタは足元に置かれているコントローラを用いて操作されることが多く、演奏者の動きに制約が生まれてしまう。そこで本研究ではステージでの視覚的な演出を考慮し、演奏者のモーションによってエフェクタ操作を行うエフェクタ、「Motion Sound Effector (以下 MoSE)」の設計と実装を行う。音楽デバイスの作成においては、その操作が明瞭で遅延が少ないことが求められる。しかしモーションを用いた操作では、その設定手順や処理が複雑化し、ユーザビリティが低下する恐れがある。そこで本研究では機械学習ソフトである Wekinator を用いたパラメータマッピングの自動補完機能や、予備動作に着目したジェスチャの早期認識手法を搭載し、操作性の向上を図る。また、これらのシステムを全てワンチップCPUに集約することで、頑健性や携帯性の向上を目指す。

2. 関連研究

モーションによって操作するエフェクタはすでに多く存

在する。平井らは「竹管の宇宙」プロジェクト [1] において、モーションによってリアルタイムに音や映像にエフェクタをかける作品を制作した。しかし、この研究で使われるデバイスはどれも大掛かりで、セッティングに時間がかかるという問題点がある。K. Konovalovs らの研究 [2] では、足と楽器に装着する無線通信デバイスを用いることで、ステージ上のどこからでも音にエフェクタをかけることができるデバイスを制作した。しかしこれは、足の動きでエフェクタを操作するため、演奏者が自由に動くことができなくなるという問題点がある。

機械学習を用いた楽器デバイスの例として、McLean J. Macionis らが作成した Sansa [3] が挙げられる。Sansa は Sunsula と呼ばれる楽器に加速度センサや圧力パッド、無線通信機器などを取り付けた楽器である。各センサによってエフェクタ操作やジェスチャによる音楽トラックの制御などが行われ、センサデータの処理には機械学習が用いられている。しかしこのデバイスはユーザがパラメータやジェスチャをカスタマイズすることができず、使用する場面が限られる。

3. MoSE の設計

システムが満たすべき要件と、それを実現するために用いた手法を挙げる。

3.1 モーションによる操作

通常のエフェクタは、演奏者が操作中エフェクタ付近に拘束されることや、手での操作時に演奏を中断する必要があり、ライブパフォーマンスの低下を招く可能性があった。そこで本研究では、視覚的パフォーマンスと演奏の両方の

¹ 関西学院大学理工学部
Kwansei Gakuin University
² 京都清華大学芸術学部
Kyoto Seika University

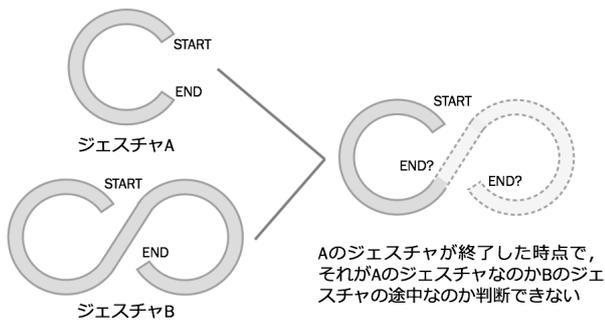


図 1 ジェスチャ A がジェスチャ B に含まれる場合

妨げにならないよう、モーションによる操作を採用した。具体的には、無線通信可能なセンサを用いて楽器の動きを計測し、角度によってエフェクトのパラメータ操作を、ジェスチャによってエフェクトの種類を切り替える。ここで、MoSE を用いた様々なパフォーマンスに対応するために、楽器の角度とパラメータのマッピングやエフェクトの切り替えに用いるジェスチャは、演奏者が自由に設定できることが望まれる。しかしパラメータマッピングには膨大な作業時間がかかる恐れがあるため、本研究では機械学習ソフトである Wekinator[4] を用いたマッピングの自動補完機能を搭載する。Wekinator に搭載された Neural Network によるマッピング機能を用いて、角度とパラメータを数点对応づけると、残りの部分を自動的に補完する機能を搭載した。これによってセッティングの時間を大幅に短縮したほか、非線形で複雑なパラメータ変化を可能とした。

3.2 ジェスチャの早期認識

ジェスチャ認識は通常、ジェスチャが完了してから認識を開始し結果を出力するが、音楽においてはわずかな時間のズレも演奏クオリティの低下を招く可能性がある。そこで本研究では、エフェクト切り替えの際のジェスチャ認識において、ジェスチャの予備動作に着目し、ジェスチャ終了前に認識結果を判別、ジェスチャが終了するタイミングに合わせて結果を出力するシステムを搭載する。

ここで、登録された複数のジェスチャの中で、あるジェスチャ (A) が他のジェスチャ (B) の一部に含まれている場合、A のジェスチャが完了した時点で、それが A のジェスチャであるのか、B のジェスチャの一部であるのか判断することができない (図 1)。そこで本研究では、ユーザがジェスチャを登録する際、あるジェスチャが他のジェスチャの一部に含まれている場合、他のジェスチャを登録するようユーザに警告するシステムを搭載する。

3.3 ワンチップ CPU による制御

システムに PC を用いると、バックグラウンドで動作している他のシステムの影響で、PC がフリーズしたり動作が重くなる可能性がある。そこで MoSE はシステムをワン

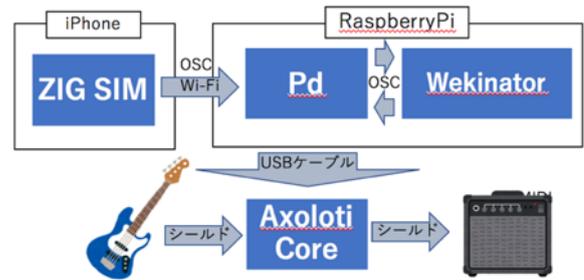


図 2 システムの流れ



図 3 デバイスの外装

チップ CPU で動作させ、頑健性を確保するほか、デバイス小型化により携帯性を向上させる。

4. 実装

4.1 システム構成

楽器のモーションを取得するセンサは、小型で無線通信可能なものであることが求められる。そこで本研究では楽器に iPhone を取り付けセンサ情報を取得する。iPhone 内では、機器のセンサ情報を取得するアプリケーションとして ZIG SIM[5] を動作させる。取得したセンサデータは、Wi-Fi 経由でエフェクタに送信される。エフェクタ部分のワンチップ CPU は、コンパクトかつ Wekinator が動作するほど処理能力が高いことが求められる。そこで本研究では Raspberry Pi[6] を採用した。Raspberry Pi 内では Pure Data[7] による処理プログラムで、センサデータの受信や Wekinator との伝達を行う。楽器の音にエフェクトをかける部分は、小型であることやプログラム可能であることが求められる。そこで楽器の音にエフェクトをかける部分として Johannes Taelman によって開発された Axoloti Core[8] を用いる。ここで楽器の音にエフェクトをかけるとともに、Wekinator の機械学習結果をもとにパラメータ操作やエフェクトの切り替えを行う。システム全体の流れを図 2 に、デバイスの外装を図 3 に示す。

4.2 モーションによる操作

エフェクトのパラメータ操作では、楽器に取り付けた重

力加速度センサの xyz の 3 軸を入力とし、エフェクトのパラメータ 2 軸の操作を行う。従来のエフェクトのパラメータ操作は 1 軸変化のみであるため、このような 2 軸のパラメータ操作により、従来得られなかったエフェクト変化を得ることができると考えられる。パラメータマッピング設定として、まずユーザはエフェクトのパラメータの値を指定し、楽器を構えながらその時の楽器の角度を登録する。これを数回繰り返して、データを Wekinator に登録、その後学習を行うことで、登録した箇所以外が自動的に補完される。

エフェクト切り替えのジェスチャは、重力加速度センサの xyz と角加速度センサの xyz の計 6 軸の値を入力とする。ジェスチャの登録として、ユーザは iPhone の画面をタッチしながらジェスチャを行う。すると、画面をタッチした瞬間を開始点、画面から指を離れた瞬間を終了点としてジェスチャが登録される。その後、3.2 節で述べた、あるジェスチャが他のジェスチャの一部に含まれていないかのチェックをシステムが行う。この時、類似したジェスチャであってもその行われるスピードが異なる可能性があるため、ジェスチャ検出の際に時間的伸縮を加味して識別を行う必要がある。そこで登録されたジェスチャ同士を、Wekinator に搭載されている DTW を用いた距離計算アルゴリズムを用いて比較する。ここで類似のジェスチャが検出された場合、ユーザに新しいジェスチャを登録し直すよう警告する。

4.3 ジェスチャの早期認識

エフェクト切り替え時にジェスチャが行われる時、曲や場面によってその速さが異なる可能性がある。そこでジェスチャ認識に、類似ジェスチャの検出と同じく Wekinator の DTW を用いた距離計算を行う。しかし、DTW による認識にはジェスチャの開始から終了までのデータが必要であり、早期認識のためにジェスチャの途中までで認識結果を出すことができない。塚本らの研究 [9] では、途中までの入力データから距離計算を行うことができる DTM アルゴリズムを作成した。しかしこの研究では、ジェスチャの識別結果をジェスチャ終了前に出力することを目的としており、本研究のようなジェスチャ終了と同時にトリガーを出力する場合を考慮していない。塚本らは他にジェスチャの特定タイミングを推定する研究 [10] を行なっているが、これはリアルタイム処理ではなく本研究には不向きである。そこで本研究では塚本らの研究を参考に、ジェスチャの早期認識に加え、ジェスチャの時間的伸縮率や処理に要する時間を加味した、ジェスチャの終了と同時にエフェクトが切り替わるシステムを作成する。

5. 今後の課題

現在は iPhone を用いて楽器の動きを計測しているが、エ

フェクタとの無線通信に Wi-Fi を利用しており、Wi-Fi 環境下でしか使用できないことや、通信強度によってシステムのレイテンシが左右されるという問題点がある。またデバイスが比較的大きく、楽器に取り付けられる箇所が限られてしまう。そこで、Xbee[11] のような専用の無線通信経路を持った小型のセンサモジュールを作成し、使用環境に左右されず、かつギターヘッドのような楽器の動きをより大きく取れる部分に装着できるようにする。また、ユーザが MoSE を用いたパフォーマンスをさらにカスタマイズできるようにするため、市販の MIDI コントロールに対応したエフェクタを操作する機能の追加を検討している。これにより、ユーザの任意のエフェクタをモーションによって操作することが可能となり、さらなるパフォーマンスの幅が広がると考えられる。また、完成したデバイスの有用性や従来のデバイスよりも有為であることを、実験を通して調査する。具体的には、ジェスチャの認識率やタイミングの正確さなどを定量的に計測するほか、実際に楽器経験者に MoSE を使用してもらいアンケートを用いて評価を行う。

参考文献

- [1] 平井重行, 志村哲, 金森森, 片寄晴弘: インタラクティブコンピュータミュージックの制作と作品の検討-「竹管の宇宙」を題材として-, 情報処理学会論文誌, Vol. 43, No. 2, pp. 310-319 (2002).
- [2] Konovalovs, K., Zovnercuka, J., Adjorlu, A. and Overholt, D.: A Wearable Foot-mounted/Instrument-mounted Effect Controller Design and Evaluation, *NIME*, pp.354-357 (2017).
- [3] M. J. Macionis, A. Kapur.: 'Sansa' a modified sansula for extended compositional techniques using machine learning, *NIME*, pp. 78-81 (2018).
- [4] Fiebrink, R., Trueman, D., and Cook, P. R.: A meta-instrument for interactive, on-the-fly machine learning, *NIME*, pp. 280-285 (2009).
- [5] ZIGSIM : <https://zig-project.com>
- [6] Raspberry Pi : <https://www.raspberrypi.org>
- [7] Pure Data : <https://puredata.info>
- [8] Axoloti : <http://www.Axoloti.com/>
- [9] 出田 怜, 村尾和哉, 寺田努, 塚本昌彦: 加速度センサによるジェスチャの早期認識手法の提案とその応用, 情報科学学会研究報告, No. 15, pp 1-7 (2014).
- [10] 村尾和哉, 山田浩史, 寺田努, 塚本昌彦: ウェアラブルセンサを用いた中の特定動作発生タイミング推定手法, 情報処理学会インタラクシオン, pp 11-19 (2019).
- [11] XBee : <http://www.digi-intl.co.jp/products/wireless-wired-embedded-solutions/zigbee-rf-modules>