

EWI wind-synthesizer のための トーキングボックス的ヴォコーダの設計とセンサのアサイン

安藤 大地^{1,a)}

概要：本稿では、AKAI 社の EWI という管楽器シンセサイザ製品でヴォコーダをトーキングモジュレーションのように身体的演奏表現と直結させて使用することを目的とした、パラメータとセンサのアサインについて述べる。

トーキングモジュレーションは、現代ではそれをフィルタによって行うヴォコーダに取って代わられているが、マイクを通した声の入力音のみならず演奏者の身体的動作が出力音に直結していることからライブパフォーマンスとして有効性が高い。また、EWI は、音量と音高を決定する管楽器としての最低限のセンサの他に、元となった管楽器では存在し得ないアサインになっているモジュレーション量を取るためのリップセンサと、やはり元となったサクソフイドラーには存在しない右手親指のピッチベンドセンサを持つ。これらのセンサを用いて、実際に音楽表現として用いることができる、動作に直結したヴォコーダすなわちトーキングモジュレーションを EWI のために設計・実装を行った。そのためにリップセンサのヴォコーダのパラメータへのアサイン手法、ピッチベンドの有効活用について、実験を元に考察を行った。

1. はじめに

1.1 木管型ウィンドシンセサイザのセンサ

木管楽器型のシンセサイザーコントローラ (ウィンドシンセサイザ) では、現在 AKAI 社の EWI という製品群が独占的な状態にある^{*1}。現行の EWI 製品である 4000, 5000 はサクソフイドラー型の運指で操作する静電タッチ検知型のピッチ決定センサ列と、管を通過する息の量を検出するブレスセンサを備え、管楽器のようにシンセサイザをコントロールすることができる。

ウィンドシンセサイザは一般的には専用の非 PCM 音源 (もしくはアナログシンセサイザのようにコントロールできる PCM 音源) が使用され、鍵盤型シンセサイザが音量の代わりに使っているペダリティではなくブレスコントロールという専用のコントロールチェンジを用いて、管楽器特有の演奏表情を演出できるようになっている。

しかしながら、現実的にはブレスコントロールだけでは演奏表情の演出には不十分であり、それを補完する意味で、1 次元リップセンサと、クラリネットを除く木管楽器では基本的に操作がアサインされていない右手親指にピッチベンドセンサが用意されていることが多く、EWI もその例に

則っている。

実際のウィンドシンセサイザの演奏はこれらの追加センサをフル活用して行われており、アコースティック楽器に匹敵する音楽表現力を求めるために、シンセサイザの音作りそのものが追加センサの存在を前提に行われていることがわかる。

右手親指のピッチベンドへの固定のアサインが一般的なのに対して、リップセンサの方は音源の特性に応じてアサイン方法が異なる。例えば YAMAHA のウィンドシンセサイザコントローラ WX シリーズはシングルリード型の 1 次元リップセンサを持ち、これがピッチベンドとして機能する^{*2}。すなわち YAMAHA WX のリップセンサは、シングルリード楽器の演奏そのまま、唇でリードを締めることによりピッチが上がり、緩めることによりピッチが下がる、という形であり、通常の楽器と同様に唇を締めたり緩めたりを短時間で繰り返すことでビブラートをかける。

それに対して AKAI EWI のリップセンサは 1 次元のセンサなのは同様だが、ピッチの上下ではなくモジュレーションデプスにアサインされている^{*3}。しかし通常の MIDI コ

¹ 首都大学東京システムデザイン学部

^{a)} dandou@tmu.ac.jp

^{*1} 他に Roland が Aerophone という製品を販売しているが、本格的な演奏用とは考えられていない

^{*2} 特許が同じであることから、初の市販木管楽器型ウィンドシンセサイザである Lyricon も同じ方式を取っている。また同時期の 1975 年に初のプロトタイプが完成した金管楽器型ウィンドシンセサイザである Steiner EVI は、トランペットの小指によるピッチ調整をモチーフにしたピッチベンド操作が実装されている。

^{*3} モジュレーションデプスは鍵盤型シンセサイザではホイールにアサインされている

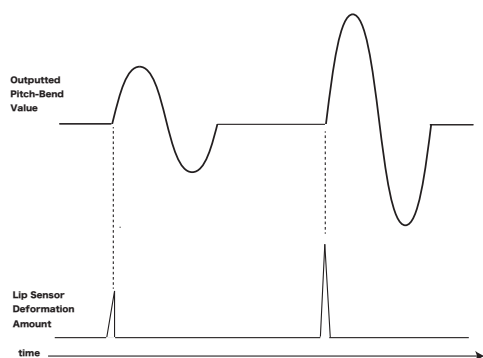


図 1 AKAI EWI のリップセンサの変形量とモジュレーションのかかり方。YAMAHA のように変形量がそのままピッチベンド量にアサインされるのではなく、ピッチベンドを駆動する LFO の振幅にアサインされる。

ントロールのモジュレーションデプスとは異なり、リップセンサを噛んで緩めるパルス 1 回で LFO が自動的に駆動され 1 周期分のビブラート（ピッチビブラート）がかかるようになっている（図 1）。リップセンサの変形量は LFO の振幅にアサインされる。

そのような歴史的背景からか、デジタル化されリップセンサの値を楽器外に取り出すことができるようになった EWI4000s 以降でもリップセンサそのものの値は独立した値としては取り出せず、ピッチベンドの値かブレスセンサの値を 1 周期分波打たせる形で出力される。^{*4}

リズムに合わせてリップセンサを噛む動作を繰り返せばビブラートをかけ続けることができるため、演奏が楽である反面「あらかじめ音源にプログラムされたビブラートしか使用することができない」と言うことができる。アコースティック楽器並みの音楽表現力を得るためには YAMAHA WX シリーズのリップセンサをピッチベンドに補助的に使う方が有効であり、EWI のリップセンサのモジュレーションデプスへのアサインは、クリティカルに音楽表現に結びついているとは言い難かったと著者は考えている。

1.2 EWI とヴォコーダ

2008 年に Seamus Blake や Mike Philipps が AKAI EWI を使用する際にヴォコーダを使用したことが話題になった [1], [2]。この手法はブレスコントローラは使用せず、ピッチ検出タッチセンサのみを使用して、声の入力はマイクで行なっている（図 2）。

当然ながらウィンドシンセサイザはブレスコントローラを使用する関係上、喋りながら演奏することはできない^{*5}。

^{*4} ブレスセンサの値を用いた場合には、はピッチビブラートではなく、音量ビブラートすなわち Amplitude Modulation 的なビブラートになる

^{*5} 過去の EWI はブレスセンサに吹き込んだ息が抜けない構造になっていたため、吹き込みながら喋ることも原理的には可能ではあるが、その奏法は非常に難しいと推測される。またサクソには Growl 奏法と呼ばれる母音を発音しながら演奏する奏法がある。

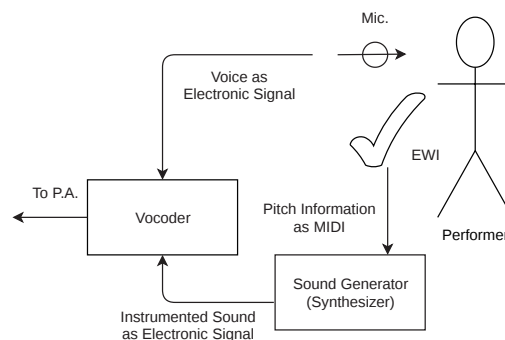


図 2 「管楽器を演奏している、管楽器が喋る」感じの演奏にはならない点で問題がある、Seamus Blake や Mike Philipps の EWI ヴォコーダパフォーマンス。EWI のピッチ決定センサだけを使い、鍵盤系のシンセサイザを駆動してサウンドを発生させ、マイクから入力した声とともにヴォコーダに入力する。

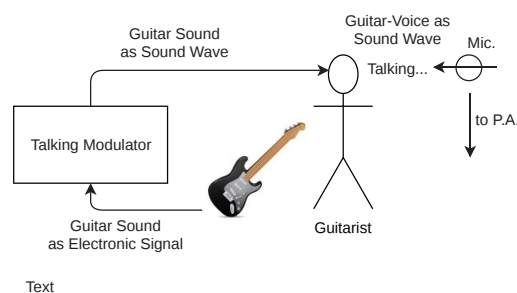


図 3 エレキギターを弾きながら口を動かすという身体表現が「演奏表現」に直結しているエレキギターのトーキングボックス。ギターからの電気信号出力がトーキングボックスで音波に変換され、ホースを通じてギタリストの口の中へ送り込まれ、その状態でギタリストが喋ると、ギターが喋るような音響を得られる。

上記の Philipps のパフォーマンスも EWI をヴォコーダのピッチにのみ使用しており管楽器の使用とは言い難く、さらにこの使用法では、EWI が持つウィンドシンセサイザ独特のアナログシンセサイザ的な音を活かすことができないため、ライブパフォーマンスの飛び道具的な意図であったと著者は推測している。

1.3 トーキングボックス

一方エレキギターでは、トーキングボックスというヴォコーダに似たエフェクトが古くから用いられてきた。トーキングボックスは人間の口腔にホースを通じてエレキギターのスピーカからの空気振動を送り込み、その状態で喋るように口を動かすことで周波数フィルタを働かせ「人間がギターの音で喋っている」ような印象を与えるエフェクトである（図 3）。

トーキングボックスは、マイクに向かって喋り DSP 内部で処理されるヴォコーダに対して、人間の動作が直接音響の変化を生み出すという点で強い印象を与えることができる。これは「演奏による音楽表現」として非常に強力であると著者は考えている。

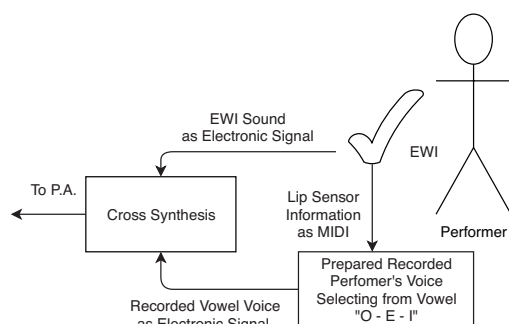


図4 提案する「身体表現と演奏表現が結びつきやすい」EWI用の新たなヴォコーダ。あらかじめ演奏者の母音「え、い、お」を録音しておき、リップセンサからのMIDI信号で出力する母音を切り替え、EWIのサウンドとクロスシンセシスを行う。

2. 提案する EWI トーキングボックス的ヴォコーダ

2.1 概要

著者は、トーキングボックスをモチーフにし、EWIの1次元リップセンサを使ってヴォコーダの周波数フィルタを変化させる手法を考案した(図4)。これによりEWIを演奏しながら喋っているような効果を出すことが可能であり、既存手法の問題点である「パフォーマンスの身体性」の問題を解決することができる。

EWIは通常のアナログシンセサイザ的なウィンドシンセサイザの音を出力する(図4中の「EWI Sound」)他に、ピッチベンド情報をMIDIで出力する(図4中Lip Sensor Information)。ピッチベンド情報はあらかじめ演奏者の声の母音を録音しておいたサンプルに回され、ピッチベンド値に応じて、「お」「え」「い」の母音のどれかを出力する。EWIのサウンドとサンプルの出力でクロスシンセシスを行い、最終的な出力とする。

クロスシンセシスは、FFT-逆FFTを用いるものの単なる乗算を行うだけで、ほぼチャンネルヴォコーダと同じ動作をする演算である。

この手法では、演奏者が口を動かすことが直接音の変化に結びつくことから、ライブパフォーマンスとしてはヴォコーダというよりトーキングボックスに近いものであり、演奏者にとってのライブパフォーマンスの音楽表現になると著者は考えている。

2.2 サンプルからの母音の出力とリップセンサの値のマッピング

サンプルは、演奏家の母音を録音しておいたものを、グラニューラーサンプリングで鳴らし続けている。ピッチベンド値でこの出力母音をクロスフェードして出力する。

図5に、1次元リップセンサの値から生成されるLFO的なピッチベンド波形から出力する母音のマッピング、クロスフェード幅を示す。前述のようにEWIピッチベンド波

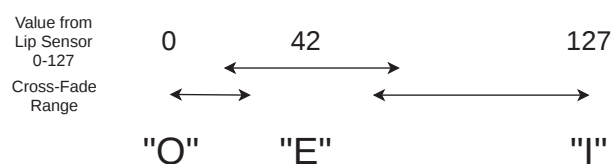


図5 リップセンサへの刺激を元にLFO的に動くピッチベンド値から、母音を出力する際の値のマッピングとクロスフェード。演奏家の実験を元に値を変更可能に設計してあるが、著者の場合「え」をかなり低い値に、クロスフェード幅をかなり狭く(5から10)に設定すると、扱いやすいことがわかった。

形は、リップセンサへの1回の刺激で刺激量を振幅とした1周期分のLFO的な出力を行うため、「え」を中心として「お」と「い」の母音をLFOに応じて短時間で移動することになり、歌っているというよりも「喋っている」という印象を与えることができる。

著者がEWI4000sにて実験を行った所、著者の場合、噛みしめる方向すなわちリップセンサの値が大きくなる方向、すなわちLFOの振幅が大きくなる方向にはコントロールがしづらく、「え」の母音を中央値64に設定すると「え」と「い」の間の切り替えが難しくなってしまうことがわかった。そのため「え」の母音は中央よりかなり低い値に設定した。

さらにクロスフェードについて実験を行ったところ、クロスフェードする幅は狭い方がコントロールがしやすい、LFOの移動によりはっきり母音の移り変わりが聴こえることがわかった。現状ではクロスフェードの値は5から10程度にして調整を行っている。

2.3 右手親指ベンドセンサのアサイン先

母音を変化させるのと同時にピッチベンド操作を行うことで、母音の変化がより強く印象付けられる表現を行うことができ、非常に有効であることがわかった。

さらに、クロスシンセシスを行うため、右手親指ベンドセンサによるピッチベンドを母音出力サンプル側すなわちモジュレーション側にも適用するかの検討を行った。

クロスシンセシスと同様の操作を行うチャンネルヴォコーダの運用では、一般的にモジュレーション側波形にはピッチの変化を求めない。人間の話し言葉がそのまま抑揚を持ちフォルマントが変化することでピッチ変化と同じような効果を出すことができることがその最大の理由である。またキャリア側とモジュレーション側の双方の基底周波数が完全に同一でなければ、ピッチが定められている楽器での音楽的效果は薄いと言われている。

しかし、本提案手法ではグラニューラーサンプリングを用いて伸ばし続けている母音をモジュレーション側波形に用いており、またキャリア側の基底周波数もリップセンサ・ベンド情報とともにMIDIの形で既に得られているため、フォルマントの変化などにも用いることが可能である。

ように思われる。

そこで、グラニューサンプリグ自体のピッチ変化を利用して、バンドセンサの情報をサンプラ側にもアサインしてみた所、現状では効果が薄いことがわかった。一般的にフォルマント変化がないまま歌声を低域高域間で移動させると違和感を感じることがあり、その違和感をヴォコーダを音楽表現としてポジティブに利用する方法があるが*6、今回の実験手法ではそのような違和感を演出することができなかった。これは、(1) ピッチベンドの幅が狭いこと、(2) サンプラが現状では母音のみを出力していることからピッチの高低によるフォルマント変化が元々あまりないこと、の2つが原因ではないかと著者は推測している。

ピッチベンドに関しては、音楽表現として有効に機能する形でより良いアサインを今後検討していく。

3. ライブパフォーマンスとしての考察と展望

3.1 噛みしめる身体動作に対する音色変化への違和感

一般的に管楽器奏者は、「良い音を出す」ことを目的として、ビブラートを除いてはアンブシュアと呼ばれるマウスピースの咥え方や吹き込み方を演奏中一定に保つ訓練を受けており、特に持続音の最中に積極的にアンブシュア変化させることには違和感を持つ。実際にリップセンサを積極的に大きく変化させる提案手法は慣れないとかなり難しい動作である。

ただし、アコースティック管楽器と電子楽器では条件が異なると著者は考えている。アコースティック管楽器はリードへの入力情報量が非常に膨大な次元となり、身体表現としては小さい変化でも大きな音色変化動きがついた演奏表情を得ることができる。EWIのリップセンサは1次元であるため、積極的に大きく頻繁に身体を動かしていかないと大きな動きを得ることができない。

この点について、今後は提案ヴォコーダの利用者を増やし、口の身体動作の違和感について、様々なフィードバックを得ていく必要がある。

3.2 母音マッピングとクロスフェード

今回は、「お」「え」「い」の3つの母音についてのみ検討を行った。リップセンサを元に出力されるピッチベンドが1次元であることから、音色が似ている「あ」「う」を落とした形になったため、母音というよりもフィルタの変化であるように聞こえるようになった。中間母音（今回の「え」）の位置やクロスフェード幅については、LFOをコントロール可能な個人差が大きいため、今後複数の演奏家に演奏してもらい、フィードバックを集めていく。

3.3 子音が存在しないことによる独特な音楽表現

提案手法では、ヴォコーダと言いつつも子音が存在しないことから、ヴォコーダというよりどちらかといえば「持続音のジューズハープ」のような「発音体+口腔フィルタ変調」の演奏に近い表現が得られた。これは管楽器の演奏表現としては非常に独特であり、さらに「本来なら口腔操作で行うことを、1次元のリップセンサのみで行う」ことによりピッチベンドの併用も含め、大きな音色変化を「明確な操作」でLFOにより頻繁に行うことが可能になった。

提案手法が新たな音楽表現の可能性を示したことから、今後子音の単純なセンサへのアサインも含め、積極的に検討する価値があると著者は考えている。

4. おわりに

本項では「身体表現として実際に口を動かす」と「管楽器を演奏する」ことの両立を目指したヴォコーダを、リップセンサの情報により発生させた母音とEWI本体の音色のクロスシンセシスにより実現した。またピッチベンドセンサを併用して様々な表現ができることを確認した。現在は母音だけであるが、かえてこれまでの管楽器にない演奏表現を得ることができた。これは提案手法の可能性の高さを示している。

今後は多くのEWI奏者に利用してもらい、フィードバックを得ていく必要がある。

謝辞 本研究は、ジャズサックス演奏家の尾崎朝子氏の依頼によって行われた。

参考文献

- [1] Blake, S.: seamus blake boris kozlov and jonathan blake in japan, <https://www.youtube.com/watch?v=2nUWXXsnC1o> (2018年12月20日閲覧).
- [2] Phillipps, M.: EWI used to control a vocoder-Mike Phillips, <https://www.youtube.com/watch?v=w3ftNajVHcg> (2018年12月20日閲覧).

*6 いわゆる「ケロケロ声」