



平成28年度研究助成 【サウンド技術振興部門】より

音程感覚の習得とより良い 歌唱体験の補助

会津大学 コンピュータ芸術学講座
准教授

Julián Villegas

1. 序論

歌うことが健康に良い影響をもたらすことは長い間議論されてきた。歌うことが不安、落ち込み等の治療に導入された場合の効果については、一部の研究者は確かな証拠に欠くとしているが¹⁾、歌うことには、快適さと健康増進をもたらす効果とともに、社会的インタラクション、感情表出、認知的発達を促進するという魅力的な特徴がある。

音痴と呼ばれる音楽的な障害に苦しんでいる一部の人々は、音のピッチを処理することが出来ないが、このような全体の人口に占める割合は少ない(2%~5%)²⁾。したがって、大部分の人間は歌うことが身体的に出来るが、我々の多くは、異なる理由(自意識過剰や、早まった批判等)で私的な時でさえ歌うことを控える。結果として、しばしば自己申告により“下手な歌手”とされる個人は、歌う幸せを逃しているのかもしれない。

歌唱能力が劣るということには多くの要因があるが、ピッチの不正確さが最も目立つ点である。ピッチの不正確さの他には、タイミングのズレや表現不足などがある。本研究では、ピッチの不正確さを改善することに焦点を当てて、ここで提示する内容と同様の技術が他の要因の改善にも応用できると考えている。

いくつかの歌唱トレーニングをするための製品がすでに存在している。いくつかのソルフェージュ(読譜や聴音)のツールは歌手のパフォーマンスが終わった後で音によるフィードバックを提供するが、これでは、ユーザーが誤ったピッチの位置を把握することを制限してしまう。また、レコーディング・スタジオやライブで広く使われているリアルタイムの自動チューニング・システム(例えば、Antares Auto-Tune Live)があるが、それらはふつうピッチを修正するためのもので、歌唱力の向上に使用されるものではない。さらに、著者の知る限り、これらのツールのどれにも、チューニングの調整機能(異なるチューニングで歌う場合に必要)を具備しておらず、声を自己ハーモナイズさせる機能、つまり独りでも和声を楽しむことができるという機能もない。そこで、本研究では今までのべた欠点を克服するためのツールを構想した。このツールは誰もが入手可能であり(オープンなハードウェアとソフトウェアを使用)、汎用的である。例えば、それは歌唱力の判断に有用だけでなく、訓練された歌手に対しても異なるチューニングをマスターすることを補助してくれる。

聴覚における粗さ感(ラフネス)³⁾との関係に基づいて、そして弦楽器の弦間のうなりを除去する一般的なチューニング手法に触発され、本システムのユーザーは現在のピッチと

望ましいピッチとの間の干渉（現在の声の調のずれ）を最小化するために、音調を整えることができるかと仮定した。このような方法は長期的な歌唱力の改善にもつながることが期待される。

この進捗報告では、歌うことに遠慮がちな歌い手の歌唱力向上のためのツールの実装方法、提示方法の可能性、および今後の評価法に関して議論する。

2. 実装

このシステムは、音響系リアルタイム・ソフトウェア・アプリケーションのために広く使わ

れるヴィジュアルプログラミング言語のPure-data⁴⁾によって構築されている。Pure-dataコーディングの移植性を活かすことにより、携帯性のあるマイクロプロセッサであるRaspberry Piや、デスクトップコンピュータのためのアプリケーションをつくることができる。図1は、このアプリケーションの異なるモードとその構成要素を示している。以下に記すように、2つのオペレーションモードは同じDSP技術と画像表示系を使用する。いくつかのプラットフォームのためのハードウェア実装要領、ソースコード、実行ファイルは、以下のURLから入手可能である。<https://bitbucket.org/julovi/>

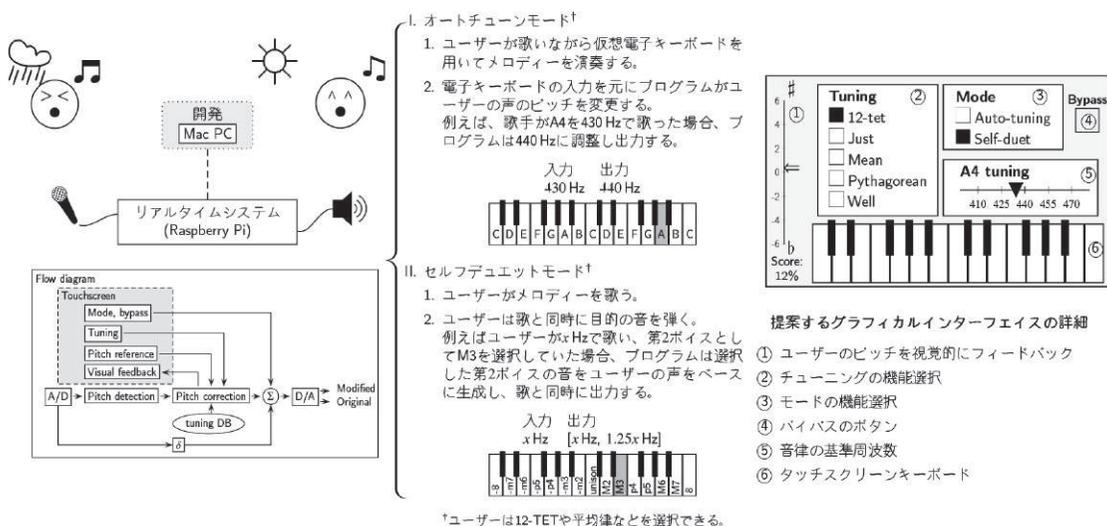


図1 プロジェクトの全体図

self-duet.

2.1 構成

歌い手のピッチ（ここでは、ピッチとは基本周波数と同義）は母音箇所（すなわち、信号が調和的であるとき）からそのサイン波成分を検出することによって追跡される。5)で提案されているように、現在のピッチを決定するのに尤度関数が用いられる。この追跡では最新の4096の音声サンプルが考慮され、デフォルトでは半音の10分の1までのずれを検出可能である。

ピッチ調整は‘Pitch-Synchronous Overlap and Add’ (PSOLA) アルゴリズム⁵⁾を使用して実行される。これはピッチ周期を検出し、その周期間で信号を重ね合わせる量を調整してピッチを上げたり（重なりを増やす）下げたり（重なりを減らす）する手法に基づいている。他の方法と比較して、PSOLAは音声フォルマントを保つという長所があり、処理された声をより自然に保つことができる。

このプロジェクトで使われたPSOLAの実装はshifter ~⁷⁾と呼ばれているMax/MSPオブジェクトから移植している。この実装例ではピッチ周期は1000のサンプルから計算され、希望値と現在のピッチ (Hz) との比率 r が $0.1 < r \leq 6.0$ である場合のみ調整される。ピッチ検出の段階と同様に、これらの調整は音声の母音箇所だけ

で行われる。

音律系はいくつかの12音律から選ぶことができる。これらの音律は‘Scala’データベース⁸⁾からピタゴラス音律、ブトレマイオス音律、中全音律、平均律のような歴史的なものを含めて抽出した。これらの音律の基準周波数はデフォルトで $A4 = 440\text{Hz}$ としている。

干渉と雑音を避けるために、耳覆い型のヘッドフォンで音を再生するとともに、頭部に取りつけた高指向性マイクで歌い手の音声を捉えることを推奨している。

2.2 オペレーション

ユーザーには図1の右側に示したようなグラフィック・インターフェイスが提供される。ラジオボタンによって、‘オートチューニング’か‘セルフデュエット’モードの選択、音律の選択、基準ピッチ（例えばバロック音楽では $A4 = 415\text{Hz}$ 、クラシックでは $A4 = 430\text{Hz}$ など）の選択を行うことができる。バイパスボタンによって、ユーザーは出力信号に処理された信号を加えたり、除いたりすることが出来る。聴覚系のフィードバックの他に、ユーザーは現在のピッチの望まれる値からのズレを視覚的に得ることも出来る。ピッチの理想値と実現値との差の累積距離値が数値スコアの計算に使われる。このスコアがユーザーの上達度合をたどるために使われる。

2.3 オートチューニングモード

オートチューニングモードが選択されたときは、ユーザーは仮想MIDIキーボードを通してメロディーを演奏することができ、選ばれた音律に対応する譜の周波数に合わせて歌手のピッチを調整する。そのままの音声と、プログラムによって調整された音声と同時に提示されれば、うなりを除去することで音調の修正を促進することが出来る。

初期設定では12音平均律のチューニングが提供されるが、ユーザーの目的によりふさわしい1つ（中全音律、ピタゴラス音律など）に変えることができる。メロディーがない場合、ソフトウェアは歌手のピッチを選択された音律の中で最も近い音高に合わせる。

2.4 セルフデュエットモード

このモードのデフォルトでは、ユーザーはメロディーを歌いながら、同時にユーザーが望む音程を仮想キーボードで弾くことができる。ソフトウェアはオリジナルの音声に加えて、その音声をユーザーが選んだ音程に移調する。セルフデュエットモードでは、それぞれのキーは仮想キーボードの中心のC4からの音程を表す。例えば、E4を押下するとそのままの音声よりも長三度（M3）上の音を生じさせる。このモードでは実際の歌手のピッチと基準のピッチ

とは無関係であるため、音高を合わせるプレッシャーからは解放されるが、和声で歌う事の利点は維持される。また、協和音の音程を選ぶことは音楽の初心者には難しいように思われるかもしれないが、提案されたインターフェースが新たな探求心を招き、それらの懸念が払しょくされることを信じている。

3. 最後に

システムのプロトタイプはすでに動作を確認済みであり、現在マイクロプロセッサ Raspberry Pi による組み込みシステム開発の最終段階に至っている。PCによるプロトタイプでは、この方法が歌うことに不安を感じる人たちの自信とスキル向上のために有用であるかどうかの事例調査を実施している。リアルタイムの視覚的フィードバックのみによる方法に対し、オーディオとビジュアルのフィードバックを使用することの利点を比較するために一連の実験を準備している。また、芸術的表現のためのツールとして、この装置をライブパフォーマンスで使う方法を探求している。

謝意

この文章に貴重なコメントと校正をしてくれた小泉宣夫教授、西村 憲教授と石原陸子氏に対して感謝したい。

参考文献

- 1) I. Clark and K. Harding, "Psychosocial outcomes of active singing interventions for therapeutic purposes : a systematic review of the literature," *Nordic Journal of Music Therapy*, vol. 21, no. 1, pp. 80 – 98, 2012.
- 2) I. Peretz and K. L. Hyde, "What is specific to music processing? insights from congenital amusia," *Trends in cognitive sciences*, vol. 7, no. 8, pp. 362 – 367, 2003.
- 3) H. Fastl and E. Zwicker, *Psychoacoustics : Facts and Models*. Berlin: Springer, 3rd ed., 2006.
- 4) M. Puckette et al., "Pure data : another integrated computer music environment," *Proc. 2nd. Intercollege computer music concerts*, pp. 37 – 41, 1996.
- 5) M. Puckette, T. Apel, and D. Zicarelli, "Real-time audio analysis tools for pd and msp," in *Proc. Int. Computer Music Conf.*, 1998.
- 6) U. Zölzer, ed., *DAFX-Digital Audio Effects*. New York, NY, USA : John Wiley & Sons, 2nd ed., 2011.
- 7) T. Jehan, "Tristan Jehan's Home Page." [Software]. Retrieved October 21, 2016. Available from web.media.mit.edu/~tristan, 2008.
- 8) M. O. de Coul, "Scala Home Page." [Software]. Retrieved October 21, 2016. Available from www.huygens-fokker.org/scala, 2006.