

演奏情報処理を応用した教育と芸術のためのソフトウェア

Software aided education through pitch estimation technology.

新井 俊一¹⁾ 小畑 幹²⁾
Shunichi ARAI Miki OBATA

1) 有限会社メロートーン (E-mail: arai@mellowtone.org)

2) 有限会社アンブレラ

ABSTRACT. In this project, we developed an education software which assists to learn to sing. Our innovative pitch-estimation technology enables realtime very robust high-sensitivity broad-range pitch extraction from human voices. We used that technology to make the song training software. Also we showed many applications which are made possible by our technology in this article.

1 背景

現在のパーソナルコンピュータでは、音楽を再生したり、音楽を作成したりするためのソフトウェアは数多く存在している。しかし音楽演奏を楽しんだり、音楽演奏を練習したりするためのソフトウェアは遅れをとっており、まだ殆ど存在していない。コンピュータの性能増加と、音楽情報科学の目覚ましい発展によって、音楽や音を扱ったソフトウェアは大きな可能性と需要を秘めている。そこで、我々は、音楽や音を扱ったソフトウェアを研究開発していくことに大きな意義を見いだしている。

新井はすでに高精度のリアルタイム音程抽出技術を開発しており、これを応用することで、いままでに無かったような新しい技術やソフトウェアを数多く開発することができる見込みである。リアルタイム音程抽出技術とは、歌唱音や楽器音を入力データとして、その音程をリアルタイムに抽出する技術である。これまでカラオケの採点機能などに利用されることがあったが、我々の技術はより高い精度を持ち、これまでに製品の存在しなかった新しい分野に応用することが可能である。

また、これまでに歌唱の練習用ソフトウェアなどは研究レベルのものが試作されたにとどまり、実用となるようなソフトウェアが存在していなかった。

世の中では多くの人がコーラスや声楽を趣味としているが、彼らの練習手段は教師に習ったりグループで練習することに限られ、一人で手軽に安価に練習する手段が存在しない。声楽のレッスンなどは遠隔地の教室を訪問して、教師に一万円弱の費用を払って受講する必要があるため、気軽に受けることができないのである。またコーラスでは、自分のパートを正確に歌う必要があるのだが、集団レッスンでは一人一人の音程の正しさなどを確実にフォローすることができなかった。

さらに多数の人が調子外れ（いわゆる音痴）に悩んでいるが、彼らの悩みを解消する手段は驚くほど少ないのが現状である。少数の教則本があるが、歌うということを文章で説明することは難しく、発声の方法を指導したりできるだけである。また数名の治療専門家が存在しているが、専門家の数は少なく、治療や通うための時間も多いため、多数の人が利用するには困難を伴うのが現状となって

いる。

2 目的

前章で述べた音程抽出技術を利用して、我々は歌唱練習用ソフトウェア「うたうたう」を開発した。これまでに本格的な歌唱練習用ソフトウェアは公開されておらず、我々のソフトウェアは声楽の練習や調子外れの治療などのさまざまな用途で利用されることが期待される。

但し、歌唱練習用ソフトウェア「うたうたう」は音程抽出技術の応用の一例として作成するものであり、我々の音程抽出技術はより広い応用範囲を持っている。今回の開発の目的の一つとして、音程抽出技術のデモンストレーションを行い、より多様な可能性を多くの人に実感してもらうことがあげられる。

歌唱練習用ソフトウェア「うたうたう」の狙いは、声楽・コーラスの練習や、音痴の治療を、一人で手軽かつ安価に行うことができるようにすることである。また一般のユーザも楽しく家庭で歌を楽しめるようなソフトウェアにすることも目標となる。

これによって、多くの声楽愛好家や調子外れのユーザが音楽を手軽に練習することができるようになり、コンピュータの多様な利用法の一つとして音楽を楽しむということが認識されれば望ましい。

本プロジェクトでは、歌唱練習用ソフトウェア「うたうたう」の試作版を完成させた。現在、ウェブページ上で公開を行っており、近い将来にシェアウェアとしての販売を開始する予定である。

本稿では、まず3章で、歌唱練習用ソフトウェア「うたうたう」の概要について述べる。その後、4章で、リアルタイム音程抽出技術の概要を述べる。

3 歌唱練習用ソフトウェア「うたうたう」

歌唱練習用ソフトウェア「うたうたう」は、個人が歌を練習するための Windows 用アプリケーションソフトウェアである。

パーソナルコンピュータに市販の安価なマイクロフォン（1000 円以下で購入可能）を接続するだけで、コンピュー

タを歌の個人教師とすることができる。

新井が開発した高精度のリアルタイム音程抽出技術を用いることで、ユーザのうたった歌を即座に解析し、その音程や音量をビジュアルに画面へと表示するものである。それによってユーザは自分の歌声について、音が外れていないか、音量は適切かなどを瞬時に判断することができる。

「うたうたう」の機能として、標準 MIDI ファイル (SMF) で記述された楽曲の伴奏を再生し、ユーザが歌うべきメロディの譜面データを線分として画面上に提示することができる。それにより、ユーザは自分の歌声が可視化された軌跡と、譜面をあらゆる線分を照らし合わせて、歌唱が適切かどうかを判断する。

ゲーム的に利用するさいには、ソフトウェアが自動で歌唱の適正度を評価して、画面に指示や評価をリアルタイムに提示することができる。たとえば、OK、GOOD、もっと高く、もっと低く、などである。そして最終的な適合率をパーセンテージやスコアとして表示することができる。

「うたうたう」が歌唱の練習ツールとして優れているのは、客観的な評価が得られることである。自分のうたった歌のよしあしは自分では分かりにくいものである。また、人による評価では、どこがどのように外れているかを指摘するのは容易ではない。そのために声楽の専門家による評価と指導を受ける必要があった。

しかし「うたうたう」を利用することで、音程の外れている箇所だけでなく、歌の細かい表現までも可視化して表示することができる。そのため、ひとたび優れた歌唱のデータを教師データとして提示すれば、自分の歌との違いなどを客観的に比較することができるものである。

このように「うたうたう」は歌唱の練習や評価の世界に、いままでになかった客観性を付加できる可能性を秘めているのである。音楽学校などの専門家の現場では、そのパラダイムシフトが受け入れられにくいとも思われるが、愛好家個人にとっては一人で歌を練習するという新たな価値を提供できるものである。

(1) 声楽・コーラスの個人練習での使用

これまで声楽やコーラスの愛好家にとっては、練習した歌がうまく歌えているかどうかの判断を指導者に任せるしかなかった。指導を受けるには時間や金銭などのコストを費やす必要があり、いつでも気軽に受けられるわけではない。

またコーラスなどの合唱では、ただ上手に歌うだけでなく、自分のパートを正確な音程で歌う必要がある。指導者なしで正確な音程を実感するのは容易ではない。

こうした問題は「うたうたう」を利用して、音程や音量の客観的なデータを利用すれば解決する。MIDI の曲データとの比較を行ってもよいが、指導者やプロ歌手の吹き込んだ音声データを利用すれば最善であろう。

(2) 調子外れの治療での個人利用

多数の人が調子外れ (いわゆる音痴) に悩んでおり、カラオケなどを恐れていたたり、もっとうまく歌うことができればよいのに、と思っている。また子供を育てるとき、幼児に聞かせる語り歌 (マザーリーズ) などにおいても、調子外れが気になって十分な歌がつけられないことがある。

こうした人たちのために、簡単な音合わせ等からはじまり、易しい曲などにあわせて歌っていくことで上手に歌えるようになるレッスンプログラムを用意していきたい。

我々の最終的な目標は、調子外れに悩むすべての人の解決となるレッスンプログラムを提供することである。

(3) 音楽教室・遠隔教育での使用

「うたうたう」は子供向けの音楽教育プログラムの一環として使用することができる。これは歌った音程が画面に表示されることにより、「音の高さ」「音の長さ」「音の大きさ」といった概念を体感的に学ぶことができるためである。

歌うことにゲーム性を付加することで、子供にとってより楽しく練習することができる。これまでの集団練習と違い、子供一人一人に PC が与えられれば、調子外れなどの問題を早期に発見して、早めに治していくといったアプローチをとることが可能である。

また「うたうたう」を利用して、ネットワーク経由で送られた教師の音声データにあわせて生徒が歌うといったシステムを構築することにより、歌唱の遠隔教育の補助とすることができる。

4 リアルタイム音程抽出技術

本プロジェクトは、新井の開発したリアルタイム音程抽出技術をベースとしている。この技術は入力された音声データからリアルタイムかつ高精度で安定した音程抽出を行うことができるものである。

(1) 本技術の特長・利点

これまでにも歌唱音から音程を抽出するという技術はいくつか存在していたが、さまざまな問題点があり、歌唱の練習用として利用するには至っていなかった。

既存の技術の問題点:

- ・ オクターブの違いを認識することができなかった。
- ・ 半音単位での音程識別しかできず、半音以下の細かい分別ができない。
- ・ 多くの製品では認識誤りが多く、安定して多様な環境で動作しなかった。

我々の技術:

- ・ 三オクターブの範囲を認識することができる。
- ・ 10cent (半音の 1/10) の高い精度での認識が可能である。
- ・ 多様なマイクロフォンとサウンドカードやユーザに対応して、安定した動作が可能である。

本末踏事業において、我々はリアルタイム音程抽出技術のプログラムコードの MMX 対応作業を行い、大幅な性能向上を実現した。すなわちアセンブラルーチンを Intel MMX Technology に対応し、並列して 4 つの減算と乗算を行えるようにした。これにより約 4 倍の高速化を達成することができた。そのため、より低速の PC でも動作するようになった。

このように現在の技術の実装は Intel Pentium Processor を対象としているが、他のプロセッサへの移植は容易である。とくに MMX のようなマルチメディア機能をもった CPU や DSP であれば高速に動作するものと思われる。

(2) 本技術の応用領域

このように我々の技術はすぐれた特長をもっており、そのため多様な分野での応用の可能性を秘めている。本節では、応用の可能性について述べていきたい。

a) 外国語の発音練習

中国語のような言語では、声調といわれる声の調子 (音高の変化) がとても重要な意味をもっている。同じ子音と母音の単語であっても、声調によって意味が全く変化してしまうのである。

そのため中国語学習者は母音、子音の発音だけでなく、声調を学ぶことがひつようである。しかし口の動きで表現することができる母音や子音の発音と異なり、声調は自覚的に練習することがより難しいと思われる。

これを解決するために中国語の声調練習ソフトウェアを現在、開発中である。子音や母音の練習支援技術についても研究中であるが、その実現には時間がかかるため、まずは声調練習のみを実装することを考えている。

ちなみに声調によって意味の変化する言語は中国語だけでなく、日本語、タイ語など多様な言語が存在している。

b) 楽器の練習

我々のリアルタイム音程抽出技術は、現時点では単音のみしか認識できず、和音など複数の音が同時にしている場合は音程を抽出することができない。そのため、ピアノやギターなどの練習にはすぐには応用することができない。

しかしヴァイオリンのような楽器では、入門者はまず単音の音程を正しく発音するところから練習することがあり、そうした分野の練習用としては十分な可能性がある。もちろん、フルートやトロンボーンのような単音楽器であれば少々の改良で利用することができる。

今後の研究課題として、和音の認識や、楽器の音色の評価などを行えるようにすることがあげられる。

c) メロディ入力装置として

マイクから入力された音程を、MIDI データとして出力することができる。これによって、鼻歌の譜面化や、楽器音への変換などを実現できる。作編曲の支援や、新たな音楽表現が可能となる。

これまでも同種の製品は販売されていたが、抽出技術が不安定であったり、ハードウェアが必要であったために、あまり普及することがなかった。しかし我々の技術はニーズを満たすことが可能であると考えている。

今後の研究課題としては、CD などの複数音が存在する環境でのメロディ抽出がある。これが成功すれば、既存の演奏から譜面を作る作業はぐっと簡単になるであろう。また既存の音楽家の演奏を分析し、研究に供することができる。

d) 聴覚障害者のための発声練習

生まれつきの聴覚障害者は、さまざまな方法で言葉を発音するためのトレーニングをすることがある。ただし、耳からのフィードバックが得られない状況下では、なかなか正しい発音を認識することが困難である。

前出の外国語発音の問題でも、大人になってからでは分別できない外国語の母音子音高低などを発音することが困難な問題である。これを解決する手法を開発すれば、外国語のみならず聴覚障害者の発声練習を支援するソフトウェアが開発可能となる。

我々の技術を利用すれば、現時点においても、音の高低の発音練習支援ソフトウェアを開発することができる。

5 おわりに

本稿では、我々の開発した歌唱練習ソフトウェア「うたうたう」と、その基礎となるリアルタイム音程抽出技術についての説明を述べた。

3章で「うたうたう」が音楽愛好者の練習や、調子外れの治療にとって有用なソフトウェアであることを示した。

4章では、リアルタイム音程抽出技術の技術的特長とその有用性、とくに多様な応用の可能性について示した。

今後の課題としては、本稿で述べた多様な応用製品を試作することができる。また「うたうたう」の最終的な製品化についてのプランや協力企業や資金についても模索中である。

現在の「うたうたう」試用版は以下の URL で公開している。サーバ環境が不安定のために一時的にアクセスできない場合もあるが、そのさいは期間をおいてご再訪願えれば幸いである。

<http://www.mellowtone.com/uta.html>

最後に、著者について簡単に述べることをお許し願いたい。

新井は代表開発者として曲データの作成以外のすべての業務を行った。リアルタイム音程抽出技術は、余暇を利用して 1998 年に開発した技術がベースとなっている。1978 年生まれ、中学校卒業後より独学でソフトウェアの研究開発の仕事に従事してきた。

本末踏事業終了後もソフトウェアの研究開発を行っていきたくと考えている。とくに興味があるのは音声や音楽の分野ではあるが、それに限らずに広く研究開発を行う予定である。研究開発業務を依頼したり、雇用することに興味をお持ちの方は、お気軽にご連絡頂ければ幸いである。

大学での研究も検討してみたが、大学院で研究をするために学部を卒業しなければならぬため、コストが大きくなり思案中である。もしも学士号なしで修士課程に受け入れてくれる研究室があれば、ぜひご一報願いたい。

小畑は有限会社アンブレラで音楽制作者として勤務しており、休暇を利用して「うたうたう」のための作編曲を行った。ゴスペルコーラスなどを愛する音楽の愛好者であるため、「うたうたう」の音楽アドバイザーでもある。有限会社アンブレラは音声認識を利用するユニークなゲームの開発元として有名である。

6 参加企業及び機関

有限会社メロートーンは代表開発者である新井の設立した企業であり、本稿で述べたソフトウェアや技術の共同実施権を保持している。

社名: 有限会社メロートーン

設立: 2002 年

代表: 取締役社長 新井 俊一

資本: 300 万円

URL : <http://www.mellowtone.com/>

7 参考文献

- ・コンピュータと音楽の世界 bit 別冊, 長島 洋一 et al, 共立出版
- ・調子外れを治す 音楽指導ハンドブック, 村尾 忠広著, 音楽之友社