



## 2009 年度上期未踏 IT 人材発掘・育成事業 採択案件評価書

### 1. 担当PM

安村 通晃 PM(慶應義塾大学 環境情報学部 教授)

### 2. 採択者氏名

チーフクリエイター: 中野 皓太(立命館大学 情報理工学部 メディア情報学科)  
コクリエイター : なし

### 3. プロジェクト管理組織

株式会社ゴーガ

### 4. 委託金支払額

2,870,066 円

### 5. テーマ名

誰でも好みの曲を手軽に歌える歌唱支援システム

### 6. 関連Webサイト

なし

### 7. テーマ概要

本テーマでは、使用者が入力した音声をリアルタイムで分析・加工し、誰でもプロ並の歌い方を体験できる新たなアプリケーションの開発を行う。

これは、カラオケにおいて歌詞の時間情報が既知である条件を利用し、事前に収録した他者の歌唱データを使用者の歌唱に転写することで、使用者の歌唱を補正する

機能を有するアプリケーションである。

従来の VoiceCoder(通称 Vocoder)と呼ばれる音声分析再合成の技術は、実時間で動作を行うが出力される音声は肉声にはとても聞こえない、いわばロボットボイスのようなものであった。高品質な音声合成では、Text-To-Speech (TTS)のように、文章・単語ごとあるいは母音子音ごとに用意された素片を繋げて再生する方法が一般的であった。

一方、入力された音声そのものの声質を変化させる技術の用途として、カラオケに代表されるようなエンターテインメント産業が挙げられる。このような応用では、使用者が手軽に遊べるもの、楽しめるものとして導入されるため、消費者が歌いながら声質を変化させるリアルタイム性が求められる。また、自身が歌う歌声という特性上、声質・音質が悪く、ロボットボイスのようであると使用者は不快であろう。そのため、肉声と何ら変わりのない十分な品質で、実時間で動作するシステムが必要とされる。

申請者は入力された歌唱に事前に分析した他者の歌唱データを転写することにより、高い品質を保持したまま歌唱力の補正を行うアプリケーションを提案する。本アプリケーションを実現するために、高品質 Vocoder により実現された音声モーフィングという技術を活用する予定である。

## 8. 採択理由

Vocoder 方式の音声分析合成技術を用いて、自分の歌声を他人の声質で表現し、歌唱支援を行うことでカラオケをもっと楽しくしようとする提案で非常に面白いと思う。開発者が属している研究室は、音情報処理研究室でそのスタッフとして、Vocoder 技術に基づく Straight で実績のある先生がおられるのでその指導は受けやすいだろう。

また、今回の提案は Straight を踏まえつつ、そのリアルタイム化に自ら挑んだ点で開発者の独自性も高く、評価したい。ぜひ、品質の高い高速のシステムを実現するように頑張ってもらいたい。

また、自分の声質が少しずつ他人のものに変えられていったとき、果たして違和感を感じないものか、どこまでを自分の声と認識するのか、認知や心理の問題としてもこのプロジェクトは興味深い。

非常に楽しみなプロジェクトの一つであり、その成果を大いに期待している。

## 9. 開発目標

カラオケは日本を起源とする歌唱を用いたエンターテインメントであり、有効なコミュニケーションのツールとしても用いられてきた。しかし一方では、歌唱を用いるという特性上、歌唱が苦手な人々にとっては敬遠されるものであることが知られている。その

ような層に対しカラオケでは円滑なコミュニケーションが図れないといった問題があった。また場合によっては敬遠する使用者に対し精神的な負荷を与えてしまう、コミュニケーションがより悪化するといった深刻な問題もあった。一方で、カラオケを楽しむ層にとっては楽しい時間を他者と共有することが目的であり、歌唱の得手不得手は敬遠する層が気にするほど重要視はされていない。そこで本プロジェクトでは歌唱を苦手とする層が気軽にカラオケを楽しめるように歌い方を加工するシステムを提案し、カラオケが苦手な層に対してもカラオケが有効なコミュニケーションのツールとなることを目指す。また提案するシステムは、特定の歌手の歌い方を事例として用い加工を行なうため、歌唱の得意な層に対しては自身の歌い方を他者に配布するといった楽しみ、コミュニケーション手段を提供することも期待される。

## 10. 進捗概要

中野君は、提案時には、システムはまだまったくできてはいなかった。所属する研究室が、ボコーダーなどに関する高い技術を持っているので、スタート時には、ある意味安心していた。ところが、1 回目のプロジェクトレビューでは、ほんの少しだけ、作りかけのシステムのデモを見せてもらっただけで、最終目標にはかなり遠いとの印象を受けた。ところが本人は、むしろハードウェアまで手を広げようとしていたので、さすがにそれは止めてもらった。また、学会発表論文の執筆や学会発表の準備、さらには卒論の執筆なども忙しいようで、開発の遅れが気になった。2回目のレビューでも、まだ目的とする音のサンプルにはたどり着いてはいなかった。年末になって、ようやく、音楽家による歌唱の収録と、それをを用いた音の加工結果のサンプルが届いたが、まだノイズが残るものであった。最終成果報告会までは、実はヒヤヒヤであったが、いざ蓋を開けてみると、無事に目的とした音声加工がきちんとできていて、ホッと胸を撫で下ろした。もう少しだけ、早くこの結果が得られていたら、その先の改良までできたのと思い、PM のプロジェクト進捗管理の甘さを反省した。

## 11. 成果

本システムの全体概念図を図 2-4-1 に示す。ユーザの声や他の歌手の声は、音声分析部 (Analyze part) で分析された後、転写部 (Morphing part) で変形され、最後に音声合成部 (Synthesis part) で合成される。これをリアルタイムで行なうのが特徴である。

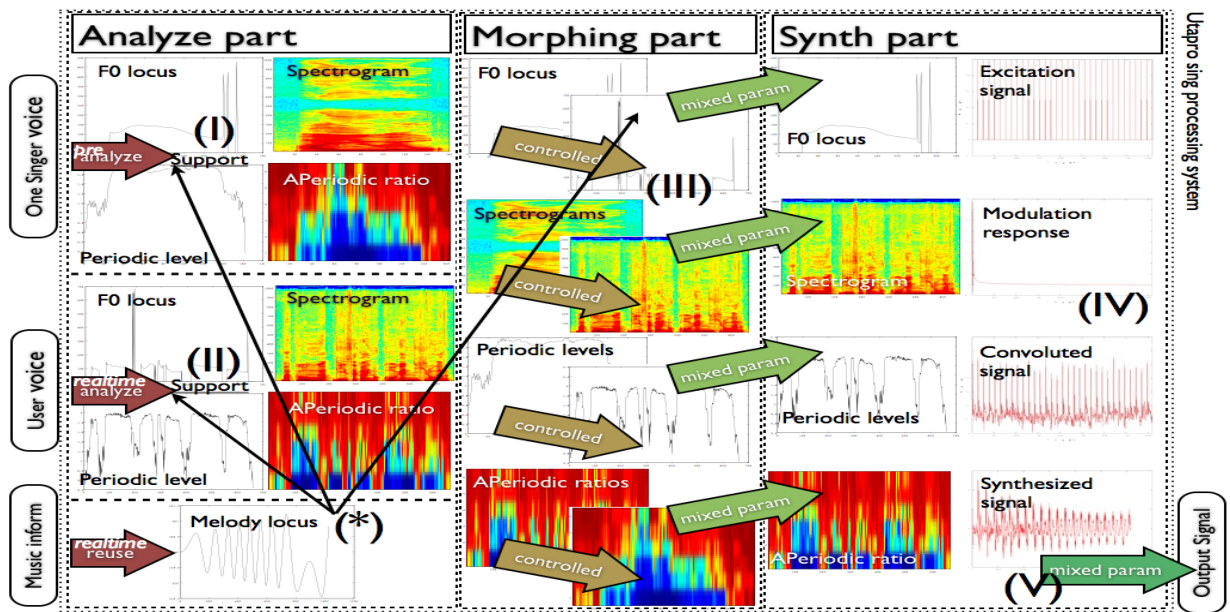


図 2-4-1 システム概念図

今回用いている技術は、ボコーダー(Vocoder)と言われる基本的な枠組みもので、入力された音声信号から基本周波数はスペクトル包絡といったパラメータを推定し、そのパラメータを用いて音声合成する。ここではボコーダーとして、和歌山大学の河原英紀らが開発した STRAIGHT という方式を用いている。ただ、この STRAIGHT は一般には処理に時間がかかるため、リアルタイム処理には向かないと思われていたが、今回、インテルの命令セット中の SIMD 命令を用いることで高速化を図った。そのため、信号処理ライブラリをこの命令を用いて書き換えを行なった。次に、ボコーダー部のうち、STRAIGHT では、基本周波数、スペクトル包絡、非周期性パラメータを推定するものである。このうち、特に非周期性パラメータは、計算量が多いが、分析には必要だが、本プロジェクトのような音声合成には必要ではないことが分かったため、非周期性パラメータと無声音は直接分離することにより、計算量を大幅に削減した。この提案方式で音声合成した結果を、従来方式の STRAIGHT などと比較評価(主観評価)したところ、STRAIGHT とほとんど変わらない品質が得られることが分かった。

今回の開発の要である、歌唱支援システムは、事前に分析したプロ歌手の分析パラメータと利用者の歌声をリアルタイムで分析したパラメータを用いて、利用者のパラメータをプロ歌手のパラメータに置き換えることで、転写を行ない、そうして得られたパラメータを復号化して音声合成する。

実際に、プロの歌手に依頼して、8曲(\*)ほど収録し、その合成音声を評価したところ、STRAIGHT とあまり変わらないほどの品質であることが分かった。

また、付随的な開発ではあるが、オーディオ入出力の遅れを軽減するための工夫や、カラオケ的に動かすために歌詞提示 GUI の製作、さらに PC を使わずに操作するため

のリモコン部などを作成した。

注(\*1): 収録したのは次の通り:(1) 美空ひばり「川の流れのように」、(2) 和田アキ子「あの鐘を鳴らすのはあなた」、(3) 森山直太郎「さくら」、(4) 石川さゆり「天城越え」、(5) 浜田省吾「もう一つの土曜日」、(6) さだまさし「無縁坂」、(7) エルベス・プレスリー「I can't stop loving you」、(8) ドリス・デイ「ケセラセラ」。

## 12. プロジェクト評価

今回の開発システムは、自分の声で歌いつつそれを人の歌い方としてリアルタイムに合成できる、まさに画期的なカラオケ支援機能であり、歌が下手なためにカラオケでは充分楽しい思いができなかった人たちへの大きな朗報となるものである。クリエイターである中野君は、音声の分析と合成におけるボコーダー技術の最先端を十分に理解した上で、これを高速化するために、一つはインテルプロセッサの並列命令を用いて、より高速に音声分析をするライブラリを開発する一方で、もう一つは、ボコーダーにおいて分析には必要となるが合成には必ずしも必要と無い非周期パラメータの分離を直接行なうことに着目した高速化により、従来時間はかかるが品質が良いSTRAIGHT というボコーダー方式と品質的にはあまり変わらないもので、しかもリアルタイムで分析と合成ができるシステムを作り上げた。学部の4年生とは思えないほどの力量である。成果報告会の場では、生のデモを行なってくれたが、元の歌声と合成した歌声の両方を聞くことができ、その違いは驚愕するほどのものであった。STRAIGHT という最新の音声技術を理解しつつ、それを高速化するためのプログラミング技術を駆使する、さらに、こういう基礎技術をカラオケへと応用するための柔軟さと三拍子揃ったクリエイターである。

## 13. 今後の課題

収録曲を増やして、かつ、本システムをベータ公開し、多くの人に使ってもらって、そのフィードバックにより、さらに良いシステムを目指して欲しい。