

MIKUTYPE : リアルタイムで初音ミク に歌わせるタイプソングシステム

竹本 拓真^{1,a)} 馬場 隆¹ 片寄 晴弘¹

概要: 初音ミクに代表される歌唱合成エンジン VOCALOID は世界的に注目されており、多くの派生アプリケーションも開発されている。その中で、「VOCALOID キーボード」のようにリアルタイムで歌唱生成をするシステムも提案されているが、専用のキーボードが必要で、習得に鍛錬が必要である。本稿では、PC キーボードによる日本語入力と鼻歌によるピッチ操作によって、実時間で歌唱生成を実施するシステムについて提案する。

MIKUTYPE : Real time Type Song System by the voice of Hatsune Miku

TAKEMOTO TAKUMA^{1,a)} BABA TAKASHI¹ KATAYOSE HARUHIRO¹

Abstract: VOCALOID represented by Hatsune Miku, It has attracted attention worldwide. It has been developed derived many applications. Among them, System for singing generation in real time has been proposed as VOCALOID Keyboard. However, special equipment is required for this. and it is necessary to discipline to learn. In this study will introduce the system without the use of special equipment for implementing the singing generation in real-time by pitch operation by humming and Japanese input on the PC keyboard

1. はじめに

YAMAHA が開発した歌唱合成技術 VOCALOID[1] は、CRYPTON より発売された初音ミクの登場以降、急速な普及が進んだ。VOCALOID では、ピアノロール上で音符を入力した音符に歌詞を付与することで歌唱の生成がなされている、ビブラートなどの歌唱表現に関するパラメータをマニュアルで修正することで、歌唱の微妙なニュアンスが付与される。最近では世界各地で VOCALOID のライブステージが催されており、歌唱のリアルタイム生成が求められる場面が多くなってきている、その要求に応えるものとして VOCALOID キーボード [2] が開発されている。VOCALOID キーボードは、右手で鍵盤を演奏しながら左手で同時に歌詞を入力することで歌唱が生成される。ユーザに演奏と歌詞入力を両手で同時に行わせるため、歌唱表

現に関する細かなパラメータの操作行えないことと、操作の習熟に時間がかかるという問題がある。本稿では、その問題を解決するものとして、リズム・歌詞入力を PC キーボードで指定し、鼻歌によってピッチ・ピッチに関する歌唱表現の付加を行うシステムについて述べる。

2. 関連研究の問題点と本研究の位置づけ

VOCALOID キーボードは、YAMAHA が開発した VOCALOID 用のリアルタイム歌唱生成システムである。これは、ユーザが、右手の鍵盤操作により、ピッチ・リズムの指定し、歌詞の指定を左手で専用の母音・子音が並べられた文字キーボードを操作することでリアルタイムでの歌唱合成を行っている。このシステムには、表情付けの機能は実装されておらず、人間らしい歌唱に必要とされる表現を付加することはできていない。また、VOCALOID の歌唱生成機能がハードウェア音源化されており、特殊な鍵盤やキーボードが使用されているので、一般ユーザが入手することができない。そのため操作も特殊なものとなり、

¹ 関西学院大学
Kwansei Gakuin University
^{a)} takumarakan@kwansei.ac.jp

習熟に時間がかかるなどの問題がある。

歌唱表現に関する表情付けを手作業による付与方法以外にも提案されている。VocaListener[3]では、VOCALOIDでの歌唱表現に関するパラメータ修正を、従来の手作業によるマニュアル入力から人間の歌声解析による自動入力としている。これにより人間の歌唱に含まれる歌唱表現に関するパラメータを歌声からVOCALOIDに反映させることでより人間的な歌唱生成を可能にしている。しかし、歌声解析した結果をリアルタイムに反映させることはできず、リアルタイムによる歌唱表現の付与は行われていない。

本研究の狙いは、誰でも簡単に操作でき、より人間性の高い歌唱の実現である。

MIKUTYPEでは、ピッチ・表情付けを鼻歌、リズム・歌詞入力をPCキーボードで行うことで、リアルタイムでの歌唱合成を実現している。これによりキーボード入力とマイク入力しか使用しておらず、一般的なノートPC1台で動作可能である。

鍵盤によるピッチ指定では、機械的な制御による正確なピッチが指定できるが、ピッチ変化などによる人間性の付加は行われておらず、表情付けを別途行わなければならない。鼻歌は人間のピッチ制御そのものであり、しゃくり・フォール・ビブラートなどのピッチに関する歌唱技法がすでに含まれている。従来手作業で人工的に表情付けを行っていたが、鼻歌により自然な表情付けをピッチ指定と同時に付加することが可能である。また、ユーザが思いついたフレーズや流れている音楽を鍵盤で再現するのに絶対音感などの高度な音楽スキルが必要だが、鼻歌であれば、音楽経験がないユーザであっても再現することが容易にでき、より直感的に操作することが可能である。特殊なキーボードによる片手での歌詞入力では、操作方法を習得する時間が必要であり、入力速度も限界がある。鼻歌によるピッチ指定により、鍵盤入力に使用していた片手を使用することができるようになり、両手による歌詞入力が可能になった。これにより、文字入力において最も普及しているqwerty配列キーボードによる文字入力が可能になる。qwerty配列キーボードを使用することにより、操作の習熟時間の削減とタイピング速度の飛躍的向上が期待できる。

3. MIKUTYPEの実装

3.1 システム概要

MIKUTYPEはMax/MSPにより実装されている。ユーザの鼻歌をリアルタイムに録音・解析してピッチを取得、あらかじめ用意されたVOCALOID2初音ミクの音声データのピッチを鼻歌のピッチに変換させる。音声データには、初音ミクの日本語音韻113種について音長1[s]でサンプリングしたものを使用する。ユーザは鼻歌と同時にPCキーボードを用いて両手で歌詞を入力する。母音入力時を、各

音声データの発音のタイミングとしている。

3.2 鼻歌のピッチ解析

ピッチ鼻歌の f_0 を推定し、フェイズボコーダを用いて音声データのピッチを鼻歌のピッチ(推定 f_0)に変更する。フェイズボコーダにおけるモジュレータ部には音声データを、キャリア部には鼻歌の推定 f_0 を基本波としたノコギリ波を、それぞれ使用する。これにより初音ミクの音色で任意のピッチを生成できるだけでなく、鼻歌に含まれるしゃくり・フォール・ビブラートなどのピッチに関する音楽表現を初音ミクの音声に付加することができる。

3.3 PCキーボードによる歌詞入力及び旋律各音のオンセットオフセット指定

文字入力において最普及しているqwerty配列キーボードによるローマ字入力により歌詞を入力する。子音入力後母音を入力すると発音を開始し、“a”、“i”、“u”、“e”、“o”、“n”キーを押し続けている間継続して発音し続け、キーを離すと発音を停止する。リズムは、母音を押し始める時間と、離す時間を調節することでコントロール可能である。

4. 考察

実装の結果をテストするため、童謡「かえるのうた」、[花]、合唱曲「翼をください」の3楽曲について歌唱を行った結果、全ての楽曲について十分な歌唱が行っていた。また、鼻歌によるしゃくり・フォール・ビブラートの抽出も行えており、リアルタイムで音声に反映させることができていた。

5. おわりに

今回、PCキーボードでの日本語入力と鼻歌によるピッチ操作によって、実時間での歌唱生成を実施するシステムを作製した。鼻歌によるピッチ指定により、鍵盤では表現が困難であるビブラートや滑らかなピッチ変更を可能にし、リアルタイムでの歌唱表現の付加を実現した。しかし、現段階では童謡などのテンポが比較的遅い楽曲の歌唱しかできない状況である。今後高速化により初音ミクの楽曲の歌唱実現と、歌唱システムだけではなく会話システムとしての使用を念頭に開発を進めていく。

参考文献

- [1] ヤマハ株式会社：“VOCALOID”
<http://www.VOCALOID.com/>.
- [2] 加々見 翔太，“リアルタイム日本語歌唱鍵盤楽器“VOCALOIDキーボード”の開発”，IPSJ,2012
- [3] 産業技術総合研究所：“VocaListener”
<http://staff.aist.go.jp/t.nakano/VocaListener/index-j.html>