

# 音楽情報処理 最前線!

未来の音楽の楽しみ方、作り方はどう変わるのか？  
コンピュータは音楽を理解できるようになるのか？  
コンピュータを使って音楽を研究する「音楽情報処理」  
という研究分野が、世界的に注目を集めています。  
本連載では、そうした最先端の研究事例を紹介していきます。

## VOCALOIDに歌を真似させる VocaListener (ばかりす)

話題の初音ミクや鏡音リン・レンを買ってみたけれど、思いどおりに歌わせるのは大変。  
こういう風に歌って欲しいというニュアンスを、自分で歌って伝えたい！  
そんな望みをかなえる技術、VocaListener (略して、ばかりす)を紹介します。

### 1. 誰でも使いやすいVOCALOIDへ

初音ミク発売以降の VOCALOID の人気はすごい。これまで作曲をしたことがなくても、購入して歌わせてみた人は多いのではないだろうか。そこには、作曲のためだけではなく、自分の好きな歌を VOCALOID で歌わせたい、VOCALOID がこういう風に歌ったらどうなるのか聴いてみたい、といったさまざまな想いもあるだろう。

VOCALOID で歌を作る

には、マウスやMIDIキーボードを使って、音符と歌詞を入力する。一番単純にはこれで、入力した通りの音符と歌詞を、初音ミクや鏡音リン・レンの声で歌わせることができる。さらにコントロールパラメータを調整すれば、歌い方に自分なりの表情を付けたり、より人間らしい声で歌わせたりもできる(図1)。

しかし、この作業はとても大変で“自分なりの初音ミク”を作り出すためには、細かなパラメータ調整と、実際に合成して聴いてみるという作業を何度も繰り返さなければならない。しかも、時間をかけてようやく調整終わった初音ミク用のパラメータは、例えば鏡音リン等の別の歌手ではニュアンスが微妙に変わってしまい、パラメータの再調整が必要となる。

このような問題は、VOCALOID を実際に使った人ならば経験したことがあるだろう。そしてこう思ったことがないだろうか、「歌を作っているのだから、歌ってニュアンスを伝えたい」と。以下では、そのような望みを叶える技術 VocaListener [1] を紹介する。

### 2. VOCALOIDに歌を真似させたい

VocaListener は歌声からその歌い方をそっくり真似て、音高と音量に関するコントロールパラメータ(PIT, PBS と DYN)を自動的に調整する。歌詞の入力も、「こういう歌詞ですよ」とテキストを事前に与えれば、音符ごとに割り当てる作業はほとんど自動で行える。歌詞の入力ではミスが発生する可能性があるが、その場合は「ここが間違っていますよ」と「ダメ出し」すると、新しい候補を選び直してくれる(図2)。それでも直らないような箇所は、手作業で微調整する。

では、歌を真似させるためにはどうするか。単純には、元の歌声から声の高さ(ピッチ)や声の大きさを抽出して、それをそのままパラメータにしてしまえば良いのでは、と思える。しかし、実はこれではうまくいかない。それには二つの原因があって、一つ目にはVOCALOIDは人間が歌った声を“サンプリング”して、それをつなぎ合わせて合成しているということがある。人は正確無比なピッチで歌えるわけではないし、歌詞によって音量にもばらつきがある。この結果、同じパラメータを与えても合成結果が変わってしまう(図3)。これは、初音ミクと鏡音リン・レンを切り替えた時に、パラメータを再調整しなければならない原因でもある。もう一つの原因は、自然な歌声を合成するために、VOCALOIDが自動的に音を滑らかに繋ぐことにある。これは、滑らかな歌声を合成するためには必要な処理だが、歌を真似させたい場合には、パラメータの調整結果が思いどおりにならない場合がある。

そこでVocaListenerは、実際に合成してみて、それが元の歌声を本当に真似できているかを比較して確かめる。真似できていなければ、パラメータを調整し直し、もう一度合成して比較する(図4)。このような処理を繰り返すことで、楽譜や歌詞が変わったり、歌手を切り替えたりしても、歌を真似ることができる。

図1 VOCALOID を歌わせるためには、音符やコントロールパラメータをマウスやMIDIキーボードによる打ち込みで入力していたが、VocaListenerを使うことで“歌って”入力することが可能となる。

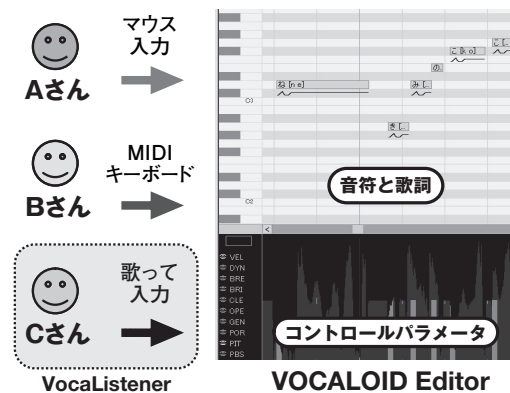


図2 歌詞の自動入力で、二つの繋がった音符の境界位置が間違っていた場合、“ダメ出し”をすると VocaListener は自動的に新しい候補を選びなおしてくれる。この例では、歌詞の中で「想い」の「も」が間違っていた(長くなってしまっていた)

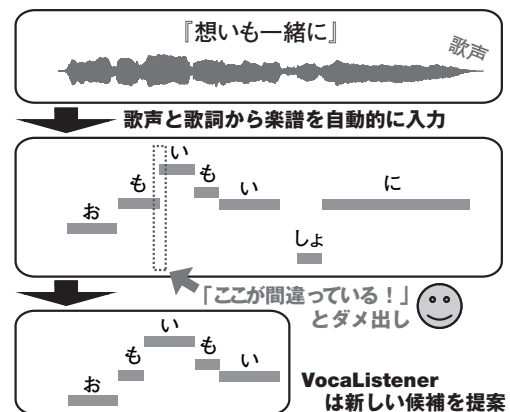
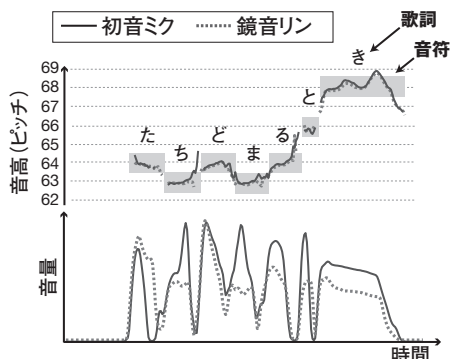


図3 全く同じ楽譜、歌詞、コントロールパラメータで、歌手を初音ミクとした場合と鏡音リンとした場合の合成結果の違い。特に音量にばらつきがあることが分かる。



## 中野 倫靖

(なかの ともやす)

2008 年筑波大学大学院図書館情報メディア研究科博士後期課程修了。博士(情報学)。現在、産業技術総合研究所特別研究員。口(くち)ドラム認識や歌唱力評価等、歌声情報処理に関する研究に取り組んでいる。歌声を用いた表現や歌声に関する人間の理解を解明し、その応用システムを計算機上に実現することを目指している。

## 後藤 真孝

(ごとう まさたか)

1998 年早稲田大学大学院理工学研究科博士後期課程修了。博士(工学)。現在、産業技術総合研究所主任研究員。計算機によって実世界の音楽・音声コンテンツを自在に扱える技術の確立を目指し、音楽・音声の音響信号の自動理解と、それに基づくユーザインタフェースの研究を中心に、様々な研究課題に取り組んでいる。

「音楽情報科学研究会」へ参加してみませんか？

情報処理学会 音楽情報科学研究会(SIGMUS)は、コンピュータと音楽とが関わり合うあらゆる場面を活動対象とする学際的研究会で、年5回の研究発表会を開催しています。研究会に会員登録すると、参加できなかった研究発表会の論文集の郵送、過去の全研究発表会の論文のダウンロードなどの特典があります。研究会の登録方法や研究発表会の開催に関する最新情報などは <http://www.ipsj.or.jp/sigmus/> をご覧ください。

### 3. でも、思いどおりに歌えない

歌を真似させることでニュアンスを細かく伝えられるのは分かったけれど、それでは歌がうまい人しか使えないのではないかな。それに、音が高くて歌えない場合はどうするのか。そのような場合のために、VocaListener はへたな歌をうまくしたり、声の高さを変えたり、歌い方を変えたりするための機能も備えている(図5)。

#### 「音高変更機能」

へたな歌をうまく変換するために特に重要なのは「音程のずれを直す」ことだろう。そのようなずれた音は、音高変更機能を使うことで修正できる。音高変更機能では、音高の変化が半音単位となるように音高をずらす「調子はずれ補正」と、部分的もしくは歌唱全体の音高をずらす「音高トランスポーズ」がある。これらの機能により、音程のずれを補正したり、男性が歌った歌唱を女性の声域に変更したりできる。

#### 「歌唱スタイル変更機能」

VocaListener では、声の高さと声の大きさを変更する歌唱スタイル変更機能を備えている。この機能では、ビブラート(音高と音量の周期的な変動)を自動的に検出し、ビブラートとそれ以外を分けて個別に歌唱スタイルを変更できる。この機能を使うことで、歌い方を誇張したり、のっぺりとした歌い方にしたりできる。

自分で歌ってそれをミクやリン・レンの声で客観的に聴くことは、実は効果的な歌唱トレーニングにもなっている。いい歌を作ろうとしていたら、自分も歌がうまくなっていた。こんな嬉しい副作用によって、補正機能なしでも魅力的な歌声が作れるようになるかもしれない。今後は、そうした「歌を作りながら歌がうまくなる」支援機能も研究していきたい。

### 4. VocaListener から始まる新しい音楽の作り方、楽しみ方

作曲や演奏を自分で経験してみることは、音楽のより深い理解につながり、素晴らしい音楽を聴いた時の感動を大きくする。これは、単にBGMとして音楽を聴くだけでなく、その細部に凝らされた技巧や、曲の意味、曲を引き立てる歌い方に意識を向け、理解できるようになった結果だろう。

VOCALOID や VocaListener を通じて、普段音楽を作ろうとは思っていなかった人でも手軽にボーカルパートを作ることができれば、新しい音楽の作り方や楽しみ方を生み出せるのではないだろうか。例えば、一から作曲するのは難しくても、「ここはこう変えたい」というアレンジ案を出すことは、比較的簡単だろう。また、歌は楽器と違って特別な道具を必要としない。うまく歌うのは必ずしも簡単ではないが、歌うこと自体には特別な訓練は必要ない。つまり、VocaListener を使うことで、音楽を聴くだけだった人もより積極的に音楽制作に関わっていける可能性があると考えている。作曲が出来る人、歌がうまい人、音楽を聴くだけだった人、様々な人が関わりあって新しい音楽を生み出し、そして楽しめる未来。我々はそんな場面を支援したい。そのために、VocaListener という技術を何らかの形で世の中に出して、多くの人達に使ってもらえるよう検討を進めている。

### 5. 研究としての VocaListener

これまで、VocaListener を VOCALOID のパラメータ調整支援システムという側面から紹介してきた。しかし、我々は元々「歌声情報処理」として以前から歌声に関する様々な研究を展開していて、VocaListener はその一つの成果に位置付けられる。そこでは、VocaListener は歌声を自在に加工・出力するための研究用ツールでもあり、今後は歌声知覚の研究にも役立てる等、様々な観点から歌声の秘密を探っていく予定である。

図4 歌を真似るために、VocaListener は実際に合成して元の歌声と比較してパラメータを調整し直し、元の歌声に近付ける。ここでは音高を例に挙げているが、音量も同じようにして比較を繰り返す。

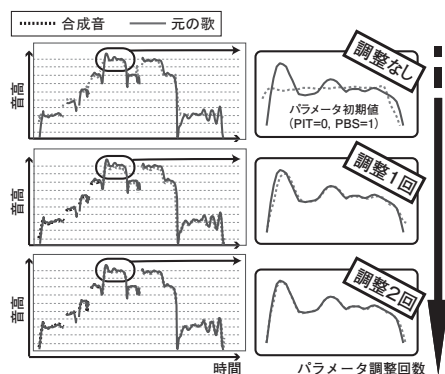


図5 音高変更機能の「調子はずれ補正」と歌唱スタイル変更機能を利用した結果の音高と音量。歌唱スタイル変更機能では、ビブラート区間を強調し、それ以外の区間を抑制している。

