

v.morish: 歌唱モーフィングのインタフェースデザイン

森 勢 将 雅[†] 大 西 壮 登^{††}
河 原 英 紀^{†††} 片 寄 晴 弘[†]

v.morish: Interface Design of Singing Morphing

MASANORI MORISE,[†] MASATO ONISHI,^{††} HIDEKI KAWAHARA^{†††}
and HARUHIRO KATAYOSE[†]

1. はじめに

音楽コンテンツにおいて歌唱の役割は極めて重要である。その分求められる品質も高く、他の楽音の電子的な合成法が早期に実用化されているのに対し、合成歌唱アプリケーションの普及は進んでいなかった。ところが、2007年に VOCALOID のキャラクター・ボカール・シリーズとして発売された初音ミク¹⁾ は、動画共有サイトのコンテンツとしての人気も相まって、DTM ソフトウェア全体の中でも異例のヒットとなった。また、このことをきっかけとして、歌唱合成やその支援技術²⁾ に対する注目が高まっている。

筆者らは、人間の歌唱を対象とした変換技術として、2名による歌唱を、声質と歌い回しの2軸で記述・モーフィングを行う歌唱モーフィング³⁾ を提案している。歌唱モーフィングには、2名の歌唱の中間歌唱を生成したり、一方の歌手の歌い回しを、もう一方の歌手の歌唱に転写したりというような使用方法があり、歌唱デザイン支援の1つとして期待されている。ただし、豊田ら³⁾ の実装例では、歌唱再生中に声質・歌い回しの各モーフィング率を変更し、即時応答させることはできないという問題があった。

本稿では、歌唱モーフィングを、歌唱デザインシステムとして実用化していくためのインタフェース設

計について取り扱う。ここでは、歌唱モーフィングに基づく歌唱デザインを実時間でを行うインタフェース“v.morish”を紹介する。

2. 歌唱モーフィングによる歌唱デザイン

歌唱モーフィングは、高品質な音声分析変換合成法 STRAIGHT⁴⁾ を用いて、歌声から音高の遷移と音色に相当するスペクトル情報を取り出し、他人のものと混ぜ合わせ再合成することにより実現される。使用者は、声質と歌い回しのモーフィング率を操作することで、歌唱デザインを行う。

2.1 歌唱デザインインタフェースの設計方針

歌唱デザインを支援するインタフェースでは、使用者の操作とレンダリング結果とを直結させる工夫がなされてきた。また、歌唱デザインを詳細に行う場合では、エディタを使い、オフラインで編集する手法が一般的である。VOCALOID や Melodyne といったソフトウェアでは、歌唱デザインの各パラメタをオフラインで編集するためのエディタが用意されている。

歌唱モーフィングにおいても、モーフィング率の操作と歌唱の変化とを直結させることで、モーフィングによる印象の変化を直感的に把握することができる。また、声質と歌い回しのモーフィング率を編集するエディタを組み合わせることで、効率的な歌唱デザインが期待できる。

3. v.morish: リアルタイム歌唱モーフィングインタフェースの実装

筆者らは、前節で示した要望を満たす歌唱デザインの支援インタフェースとして v.morish を実装した。図1に GUI の概観を示す。

[†] 関西学院大学理工学部

School of Science and Technology, Kwansei Gakuin University

^{††} 和歌山大学大学院システム工学研究科

Graduate School of Systems Engineering, Wakayama University

^{†††} 和歌山大学システム工学部

Faculty of Systems Engineering, Wakayama University

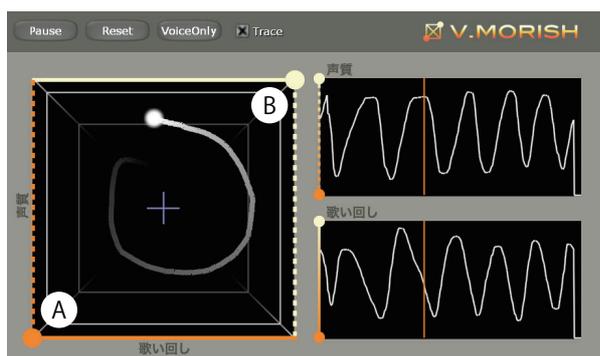


図 1 実時間歌唱モーフィングインタフェース「v.morish」
Fig. 1 “v.morish”: Real-time singing morphing interface.

v.morish には、声質と歌い回しを 2 次元平面へマッピングさせたインタフェース (図 1 左) が用意されている。使用者は、このインタフェースを用いることで、モーフィング率の操作と歌唱の変化とを直結させることができる。

図 1 右の上下に示されるグラフは、声質と歌い回しのモーフィング率を編集するためのエディタである。使用者は、このエディタを用いることにより、オフラインでモーフィング率を精密に編集できる。使用者は、楽曲再生中にリアルタイムで操作するインタフェースで大まかな特徴を決め、エディタを用いて微調整する手順で歌唱デザインを行う。以下では、それぞれの方法について説明する。

3.1 マウスを用いた実時間の歌唱デザイン

図 1 左の横軸は歌い回しのモーフィング率、縦軸は声質のモーフィング率に対応する。2 名の歌手を A, B とした場合、図 1 に示されるように、左下と右上が歌手の元の歌声となる。v.morish では、歌唱の再生中にマウスカーソルを移動させることで、リアルタイムでそのマウスカーソルから得られたモーフィング率の歌唱となる。また、再生中におけるマウスの移動によるモーフィング率の変化は、図 1 右上、右下のエディタにも反映される。つまり、エディタを用いた詳細な編集は、このインタフェースを用いて作った軌跡に基づいて実施することができる。

3.2 エディタを用いた詳細な歌唱デザインとトレース再生

図 1 左のインタフェースを用いて大まかに作られたモーフィング率の時系列は、図 1 右上、右下のエディタを用いて、さらに細かく調整できる。また、v.morish では、エディタにより編集されたモーフィング率の時系列に基づいて再生を行う「トレース再生」機能が用意されている。図 1 左のインタフェースでは、モーフィング率の細かい調整を行うことができず、同じ条件を再現することも不可能であった。トレース再生により、何度でも同じモーフィング率の歌唱を再生する

ことが可能となる。トレース再生中は、図 1 左のインタフェースに基づくモーフィング率の変化は反映されず、エディタの時系列から得られるモーフィング率に基づいて歌唱が再生される。

3.3 インタフェースの利用

v.morish は、一般に開放されたイベント、企業向けの講演、研究者を対象とした展示により、試用されている。定性的なものながら、観客からユーザビリティに関する意見を集約した。

声質・歌い回しを 2 次元空間へとマッピングし、実時間で歌唱を操作するインタフェースや、2 名の中間的な声質というこれまでに体験したことが無い歌唱が合成できることは、多くの利用者に好意的に受け止められた。その一方で、エディタの時系列に歌詞が表示されていないため、部分的に編集を行うことが困難であるという指摘もあった。

4. おわりに

本稿では、歌唱モーフィングによる実時間歌唱デザインを支援するインタフェース v.morish について述べた。今後は、本インタフェースをフリーソフトウェアとして配布し、モーフィングに基づく歌唱デザインの有効性の検証を行う。また、歌唱モーフィングに伴う品質劣化の改善についても検討する。

謝辞 本研究の一部は、科学研究費 若手 (スタートアップ)20800062 と、科学技術振興機構による戦略的創造研究推進事業のデジタルメディア領域 CrestMuse プロジェクトの支援を受けて行われた。

参 考 文 献

- 1) 佐々木 渉: 仮想楽器をリアルにする「未来 (ミク) の記号」と、VOCALOID で注目される「人の形」「声の形」について、音楽情報科学研究会, Vol.2008, No.50, pp.57-60 (2008).
- 2) 中野倫靖, 後藤真孝: VocaListener: ユーザ歌唱を真似る歌声合成パラメータを自動推定するシステムの提案, 音楽情報科学研究会, Vol.2008, No.50, pp.49-56 (2008).
- 3) 豊田健一, 片寄晴弘, 河原英紀: STRAIGHT による歌声モーフィングの初期的検討, 音楽情報科学研究会 (2006).
- 4) Kawahara, H., Masuda-Katsuse, I. and de Cheveigne, A.: Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction, *Speech Communication*, Vol.27, No.3-4, pp.187-207 (1999).