

| | |
|--------------|---|
| Title | 声区表現を可能とする歌声合成を目的としたARX-LFモデルの制御法に関する研究 |
| Author(s) | 元田, 紘樹 |
| Citation | |
| Issue Date | 2013-03 |
| Type | Thesis or Dissertation |
| Text version | author |
| URL | http://hdl.handle.net/10119/11326 |
| Rights | |
| Description | Supervisor: 赤木正人, 情報科学研究科, 修士 |

声区表現を可能とする歌声合成を目的とした ARX-LF モデルの制御法に関する研究

元田 紘樹 (1110061)

北陸先端科学技術大学院大学 情報科学研究科

2013年2月6日

キーワード: 声区, 歌声合成, ARX-LF モデル, 音声生成機構.

計算機上で人工的に歌声を生成・加工する歌声合成の分野は、音声科学における重要な分野の一つである。より高品質かつ多様な歌声合成システムを構築することは、音楽情報処理分野への貢献のみならず、音声の生成・知覚に関する新たな知見を与える上でも、重要な役割を担っている。これに対し、人のように自然で多様な歌声合成は、未だ実現に至っていない。その原因の一つとして、‘声区’の表現が挙げられる。

声区とは、人の声域を発声法と声質の相違によって区分したものである。人は、声区ごとの声帯振動様式の違いを歌唱訓練によって習得することで、広い音域を自然な声質で歌うことができる。一方で、歌声合成の分野では、そのような声区表現には十分に対応できていないため、高音域及び低音域で不自然な合成音を生じる。高音域及び低音域における合成音の自然性を向上させる方法として、声区ごとの声帯音源特性を付加することが考えられる。そのためには、声帯音源特性を記述できるモデルが必要となる。

本研究では、声区表現を可能とする歌声合成に向けた、声帯音源特性の制御法の検討を目的とする。目的を遂行するため、音声生成過程を模擬することで、声区表現のための声帯音源特性の制御が可能である ARX-LF モデルを適用する。ARX-LF モデルが持つ、声帯音源特性に対応する複数の ARX-LF パラメータを、音高の変化に伴い適切に変化させるように制御モデルを構築することで、声区ごとの声帯音源特性を付加できるようになる。声区表現を可能とするための歌声合成システムの枠組みを提案し、ARX-LF パラメータ制御モデルを構築した。そして、歌声合成音を作成し、客観評価と主観評価を実施することで ARX-LF パラメータ制御モデルの評価を行なった。これらの結果を報告する。

まず、ARX-LF モデルによる分析・制御・合成を行うことで、声区ごとの声帯音源特性を付加できる歌声合成システムを提案した。次に、声区表現に対する ARX-LF モデルの有効性を検証するために、声区ごとの ARX-LF パラメータを分析したところ、先行研究の声帯音源特性の知見に合致した結果が得られた。分析結果に基づいて、それぞれの ARX-LF パラメータ制御モデルを構築した。各声区内を線形で補間することで、作成する歌声合成音の音高ごとに、適切に各パラメータが制御されるようにした。

そして、提案システムによって歌声合成音を作成し、客観評価と主観評価を行なった。客観評価のために、低周波数域におけるスペクトル傾斜を分析し、声区ごとに比較を行なったところ、falsetto では急峻な傾き、vocal fry では緩やかな傾きが得られた。分析結果の妥当性を検証するために、人の歌声についても分析を行なったところ、歌声合成音と同様の傾向が得られ、歌声合成音におけるスペクトル傾斜の再現性が示された。主観評価のために、聴取実験を実施して高音域及び低音域における聴取印象の比較を行った結果、falsetto では氣息性、vocal fry では粗慥性といった、声区ごとの典型的な聴取印象が得られた。さらに、音域が広いほど、声区ごとの声帯音源特性の付加が重要である可能性が示唆された。これらの結果から、声区表現が可能な歌声合成に対する、ARX-LF パラメータ制御モデルの有効性が示された。

本研究で提案した、音声生成機構からのアプローチに基づく歌声合成は、人のような歌声合成システムの実現だけでなく、音声生成機構・音響的特徴・知覚の相互関係性の解明にも繋がるものであると考える。