

報告番号	※甲	第	号
------	----	---	---

主 論 文 の 要 旨

論文題目 多様な歌唱様式を予測・説明する歌声音響信号の
分析合成モデルとその応用に関する研究
氏 名 大石 康智

論 文 内 容 の 要 旨

人間の口から発せられる音響事象の一つである歌声は、歌詞に込められた感情や想いを旋律に乗せて伝達する最も身近な音楽演奏手段である。歌声は、これまで、音響学、生理学、解剖学、心理学、歌唱学といった多角的な立場から研究され、歌声知覚や生成機構の解明に向けた様々な知見が得られた。例えば、歌声の旋律を表す基本周波数 (F_0) には、通常の話声に見られない、ビブラートなどの動的な変動成分が存在する。また、歌声の音色を表すスペクトルの 3kHz 付近には、顕著なピーク（歌声特有の響き成分）が存在し、このピークは音声振幅やビブラートに同期して振幅変調される。さらに、これらの変動成分の動きは歌唱者ごとに多様であり、個人性知覚に寄与する特徴である。しかし、歌声を工学的に利用する実用的なシステムでは、このような音響的特徴の時間的な動きが精緻にモデル化されていなかった。そのため、ハミング検索システムでは歌唱者によってその検索性能は低く、歌声合成システムでは多様な歌唱様式による歌声の合成が困難であった。同じ歌声であっても、人間は歌唱様式や技術の差を容易に区別できるように、システムも歌唱者による歌声の多様性を理解し特徴づけた上で、目的となる歌声の認識や合成を行うべきである。

そこで本研究では、多様な歌唱様式を予測・説明する歌声音響信号の分析合成モデルを構築する。そのために、従来のように特定の歌唱者や歌唱法に限定するのではなく、一般人を含めて、様々なレベルの歌声を対象とした大規模データベースを用いる。

まずはじめに、聴取実験と自動識別実験に基づいて、歌声と通常の話声を人間と同程度の性能で識別しうる信号特徴尺度を構成する。通常の話声との識別の観点から、あらゆる歌声に共通する音響的特徴として、スペクトルや音高の時間的な変動成分の重要性を示す。

次に、歌声の音高 (F_0) に焦点を当て、その時間的な動きをモデル化し、歌唱者ごとにどのように動きが異なるか、歌唱様式を特徴づけることに取り組む。本研究では、 F_0 が歌唱様式を特徴づけるなんらかの微分方程式に従って生成されると想定する。しかし、この微分方程式が実際どのような形をしているのかは未知である。そこで、観測される F_0 から、その背後にある微分方程式を明らかにする問題と位置づけ、2つのモデルを提案する。1つ目のモデルでは、 F_0 とその時間微分によって構成される相平面もしくは相空間上に F_0 軌跡を描き、その同時確率分布によって F_0 の動きをモデル化する。相平面は、複雑で解けない微分方程式の解の性質を調べるための便利な手段であり、この平面の同時確率分布が近似的に微分方程式を表現すると考えた。このモデルの有効性を2つの観点から評価した。まず提案モデルによって F_0 軌跡から階段状の目標音高軌跡を推定し、これをハミング検索の検索クエリに利用したところ、従来法と同等以上の性能が得られた。次に、提案モデルによって特徴づけられる動的変動成分から、歌唱者ごとの歌唱様式を分析した。そして、提案モデルが歌唱様式を大まかに自動分類できることを評価実験から確認した。

2つ目のモデルは、 F_0 制御モデルである。旋律を表す階段状の音高軌跡を入力とし、これに様々な動的変動成分が複雑に重ね合わされた状態で F_0 が出力されると想定し、この入出力関係を全極モデルで表現する。この全極モデルのパラメータ (逆フィルタのインパルス応答) によって、 F_0 動的変動成分が表現される。そして、観測される F_0 だけから、反復法によって、階段状の音高軌跡と逆フィルタのインパルス応答を推定するアルゴリズムを提案する。逆特性のインパルス応答を推定することは、線形微分方程式の係数を近似的に求めていると考えることができる。評価実験では、提案手法によって、 F_0 軌跡から階段状の音高軌跡と動的変動因子を表すインパルス応答を適切に推定できること、また推定されたこれらの信号から F_0 軌跡を再合成できることを正解率の観点から評価した。最後に、ある歌唱者の F_0 動的変動成分を別の歌唱者の動的変動成分と取り換え、声質と音高を変化させることなく歌唱様式だけを自由に変換可能な歌声合成手法を提案する。