

論文内容の要旨

博士論文題目 Techniques for Improving Voice Conversion Based on Eigenvoices
(固有声に基づく声質変換のための改善技術)

氏名 大谷 大和

(論文内容の要旨)

音声は、言語情報のみでなく、話者性などの非言語情報も同時に伝達可能なコミュニケーション媒体である。より多様な音声コミュニケーションをもたらす技術として、言語情報を保ったまま非言語情報を変換する声質変換技術が研究されており、様々な応用が期待されている。近年、統計的手法に基づく声質変換技術の話者変換分野における発展が顕著である。中でも、変換モデル適応技術を導入した固有声変換法の提案により、特定の元話者から任意の目標話者への変換を可能とする一対多声質変換と、任意の元話者から特定の目標話者への変換を可能とする多対一声質変換が実現され、声質変換技術の実用性は大幅に改善された。今後、声質変換技術の実用化を目指す上で、変換処理のさらなる性能改善と、より柔軟な変換モデル構築の枠組みの実現が望まれている。

本論文では、固有声変換における変換性能改善と、任意の話者対を対象とした声質変換の実現に取り組んでいる。まず、変換処理の基本性能を改善するために、変換音声合成時に用いる音源モデルを見直し、音源特徴量の変換を実現した。次に、固有声変換のための適応学習法を提案し、適応処理性能を改善した。上記2つの提案技術と、既に高い有効性が確認されている系列内変動を考慮した変換法を、一対多固有声変換に統合することで、大幅な性能改善効果が得られることを示した。さらに、多対一／一対多固有声変換技術を拡張し、任意の元話者から任意の目標話者への変換を可能とする多対多固有声変換法を提案した。また、本変換法を拡張することで、発話内容の制約を一切必要としない多対多固有声変換モデル学習法も提案し、その有効性を示した。

以上の理論提案およびその実験的検証により、より変換性能および実用性に優れた固有声変換技術の実現が示された。

(論文審査結果の要旨)

言語や身体的な制約の壁を越えて、より円滑な音声コミュニケーションをもたらす技術として、言語情報を保ったまま話者性等の非言語情報を変換する声質変換技術が注目されている。確率モデルの適応処理を導入した固有声変換技術により、特定話者と任意の話者間の声質変換が実現されたが、その変換性能は十分とは言い難く、さらなる改善が望まれている。また、より利便性の高い技術として、任意の話者間における声質変換の実現が望まれている。本論文では、以下に述べる幾つかの手法を提案し、その有効性を実験的に検証している。

- (1) 主に声道特徴量変換の向上による性能改善がなされてきた声質変換処理に対し、さらなる改善を行うため、音源特徴量変換の向上に取り組んだ。高品質な音声分析合成方式STRAIGHTの混合励振源モデルを導入し、そのモデルパラメータの統計的変換処理を提案することで、変換性能を改善した。
- (2) 固有声変換において、モデル適応性能を改善するために、適応元モデルの最適化処理に取り組んだ。音声認識分野で高い有効性が認められている適応学習を固有声変換に導入することで、変換音声の音質改善に成功した。
- (3) 上述のSTRAIGHT混合励振源を用いた音源特徴量変換と固有声変換のための適応学習、さらに、既に高い有効性が確認されている系列内変動を考慮した変換法を統合することで、一对多固有声変換の性能を大幅に改善した。
- (4) 固有声変換技術を応用することで、任意の話者間における声質変換を可能とする多対多固有声変換技術を提案した。また、本技術を拡張して、任意の音声データを用いた多対多固有声変換モデル改善法を提案し、有効性を示した。

これらの手法は、従来技術では達成できなかった変換性能の向上と、柔軟性に優れた変換モデル構築を実現したものとして、有用かつ画期的なものであったと高く評価できる。本研究成果は、3編の学術論文や、14編の国際会議論文として発表され、国際的にも高い評価を得ている。声質変換デモソフトウェアを構築する力もあり、非常に貴重な人材である。以上より、平成21年12月24日に開催した公聴会の結果も参考にして、本博士論文の審査を行い、本論文は博士論文(工学)の学術論文として十分な価値があるものと判断した。