

論文内容の要旨

博士論文題目 Augmented speech production beyond physical constraints using statistical voice conversion —Alaryngeal speech enhancement and singing voice quality control—

(統計的声質変換を用いた身体的制約を超えた音声生成

—無喉頭音声強調及び歌声の声質制御—)

氏名 土井 啓成

(論文内容の要旨)

音声は、言語情報だけでなく、感情や話者性といったパラ言語／非言語情報までも、同時に伝達できる。また、リズムや音楽にのせることで歌声となり、楽曲を構成する要素としての役割も果たすことができる。このように、人は多種多様な情報を伝達できる声を生成できるが、その声質は個人の身体的特徴に強く依存するため、自身の身体的制約を超えた声質を生成することはできない。このことは、現状の音声コミュニケーションにおいて、様々な障壁を生み出している。例えば、事故や病気等の理由で声帯を失った喉頭摘出者は、声帯振動による音源生成機能の消失により、通常の発声は困難となる。代用発声法による無喉頭音声の生成が必要となるが、その音質は低く、円滑な音声コミュニケーションを妨げる要因となる。別の例では、歌声における表現の限界がある。個人が扱える声質(音色)は限られるため、音楽表現に大きな制限がかかる。

本論文では、無喉頭音声強調および歌声の声質制御を対象として、統計的声質変換技術を用いた身体的制約を超えた音声生成の実現に取り組んだ。喉頭摘出者の無喉頭音声を健常者の通常音声へと変換することで、その自然性を大幅に改善する音声強調技術を構築した。また、固有声変換技術を導入することで、変換後の声質制御を可能とし、音声により伝達される話者性までも改善する技術を実現した。さらに、本技術を歌声における声質制御にも応用し、任意の歌手間における声質変換処理を実現し、物理的制約を超えた歌唱表現を実現した。以上の理論提案およびその実験的検証により、無喉頭音声強調および歌声の声質制御において、身体的制約を超えた音声生成技術の実現が示された。

(論文審査結果の要旨)

人間は、調音器官を巧みに用いて、多種多様な情報を伝達可能な音声や歌声を生成することができる。一方で、生成される声の声質は、身体的制約の影響を大きく受けるため、各個人に限定されたものとなる。この音声生成の物理的制約は、音声コミュニケーションにおいて様々な障壁をもたらす。例えば、後天的に声帯を失った喉頭摘出者は、声帯振動による音源生成が不可能となるため、代用発声法を用いて無喉頭音声を生成するが、その音質は大きく劣化する。また、歌唱において、生成できる声質が限られることは、自由な音楽表現の妨げとなる。本論文では、統計的声質変換による音声生成機能増強という枠組みを提案し、以下の通り、複数の応用技術に対して、その有効性を検証している。

- (1) 無喉頭音声の一つである食道音声に対して、統計的声質変換技術に基づく音声強調処理を提案し、食道音声の品質を大幅に改善できることを示した。さらに統計的声質制御技術を導入し、声質の手動制御・自動適応を可能とすることで、食道音声の話者性強調処理を実現した。これにより、失われた声と類似した声で発声できる(失われた声を取り戻せる)可能性を見出した。
- (2) 他の無喉頭音声(電気音声および微弱電気音声)に対しても、話者性強調処理を提案し、その有効性を示した。さらに、各種無喉頭音声に対して提案技術により得られる音声品質を比較することで、その利点・欠点を明確にした。
- (3) 歌声に対して、統計的声質変換・制御技術に基づく声質制御法を提案し、任意の歌手が自身の歌声を所望の声色へとリアルタイムに手動制御・自動適応可能とするボイスエフェクターを構築し、その有効性を示した。

これらの成果は、従来技術では本質的に解決困難であった問題に対して、一つの明確な解決策を示しており、有用かつ画期的なものであると高く評価できる。本研究成果は、1編の学術論文や、8編の国際会議論文として発表され、その内2編の国際会議論文には論文賞が授与されており、国際的にも高い評価を得ている。国内においても、最優秀発表賞を2度授与している。以上より、平成25年1月16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)として十分な価値があるものと判断した。