

要 旨 (和文)

(1, 000字程度)

専攻名	電気・化学専攻	氏名	井野 竣介
学籍番号	2181210		
主 題	構音障がい者による発話内容の認識と変換を目的とした音響モデルの検討		

要 旨

音声信号処理技術は、これまで健常者を対象としたものがその多くを占めてきた。しかし我が国において2016年に「障害者差別解消法」が施行され、ハードウェア・ソフトウェアの両面から各分野における情報技術支援の重要性は非常に高まっている。本研究では、脳梗塞による半身麻痺による構音障がい者の音声認識の検討を行う。半身麻痺による構音障害者の発話スタイルは、神経麻痺による筋肉の弛緩が原因で健常者と大きく異なり不安定であるため、特定話者モデルでの音声認識には限界がある。これに対し本研究では、構音障がい者の音声認識精度の改善を目的とし、音声の特徴量に加えて動画から口唇の動きを特徴量として用いる手法を検討する。

はじめに、本研究の構想システムにおける認識率の目標を設定するにあたり、構音障がい者の音素認識誤りの傾向を調べた先行研究を参考にした。この研究では、構音障がい者3名を対象とした音素認識実験により正解率が低下している音素が母音、子音ともに類似しており、特に母音/a/, /i/, /u/, /e/, /o/の正解率が低いことが分かった。3名の母音5種の正答率の平均は86.78%であった。本研究の目標としては、音声認識システムの改良のため、本研究の構想システムにおける認識率の目標は有意差が出たといえる90%とした。次に、読唇術システムを作成し、そのデータを分析することで構音障がい者の発話時における口唇の動き方の特徴の分析や、母音5種の分類を行った。モデルには、サポートベクターマシンとニューラルネットワークを用いた。構音障がい者について、ニューラルネットワークを用いたクラス識別の精度は88%となり、母音5種ではあるものの目標値を上回った。

学習モデルの精度評価では、健常者については組み合わせた特徴量、音声のみの特徴量、読唇術のみの特徴量の順で精度が高かった。構音障がい者については組み合わせた特徴量、読唇術のみの特徴量、音声のみの特徴量の順で精度が高かった。音声の精度が悪かったが、読唇術の特徴量と組み合わせることで高い精度を出すことができた。

最後に、本研究の構想システムについて性能評価実験を行った。構音障がい者については音声のみで行ったテキスト化と比較して、読唇術の特徴量と組み合わせることで高い性能を出すことに成功した。