

論文審査の結果の要旨

学位申請者 張 兆峰

本論文は、「A study on speech signal processing for noise robust speaker and speech recognition (雑音に頑健な話者および音声認識のための音声信号処理に関する研究)」と題され、5章から構成されている。雑音に頑健な話者認識と音声認識を目的としており、マルチチャンネル最小平均二乗法と深層学習に基づく残響(乗算性雑音)の除去方法を提案している。また、マルチチャンネル最尤法に基づく特徴変換を利用することで、乗算性雑音に対して頑健な話者の認識方法、および、加算性雑音に頑健な音声の認識方法を提案している。

第1章では、雑音に頑健な話者認識と音声認識を目的とした種々の従来法を調査し、雑音除去と認識手法との関連をまとめ、本研究の目的と範囲、各章間の関係を説明している。

第2章では、第3章以降に記述される共通事項として、特徴抽出法、話者認識法、音声認識法(音響モデル、言語モデル、デコーダ)についてまとめている。

第3章では、複数のマイクを設置して情報量を増やすことで、マルチチャンネル最小平均二乗法に基づく残響の推定方法、および、一般化スペクトル減算法に基づく残響の除去方法を提案している。シミュレーション実験では、深層学習に基づく特徴変換の手法、および、残響除去の手法を、話者識別のタスクに適用している。結果、従来の遅延和ビームフォーミング法の性能を、提案法により大幅に改善できることが確認された。

第4章では、マルチチャンネル最尤法に基づく特徴変換の手法を提案している。加算性雑音により音声の認識性能が劣化する問題に対し、マルチチャンネルを介して取得された複数の観測信号に共通する変換行列を推定することで、クリーンな音声を得る手法を提案している。その際、クリーンな音声に対する学習モデルを用い、EMアルゴリズムにより最適な変換行列を自動的に推定している。音声認識のタスクに適用し、様々な加算性雑音を印加して実験を行った結果、提案手法により従来手法を凌ぐ性能が達成された。

第5章では、本論文の結論と今後の課題について述べている。本論文により得られた成果は、乗算性雑音と加算性雑音が強い劣悪な環境下であっても、適切な雑音除去が可能となり、高精度な話者および音声の認識システムが実現できると結論付けられている。

よって、本論文は工学上及び工業上貢献するところが大きく、博士(工学)の学位論文として十分な価値を有するものと認める。

審査委員主査 岩橋 政宏 印