



HOKKAIDO UNIVERSITY

Title	音響的類似単語の雑音ロバストフレーズ音声認識に関する研究 [論文内容及び審査の要旨]
Author(s)	宮崎, 善行
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第15559号
Issue Date	2023-03-23
Doc URL	https://hdl.handle.net/2115/89606
Rights(URL)	https://creativecommons.org/licenses/by/4.0/
Type	doctoral thesis
File Information	Yoshiyuki_Miyazaki_review.pdf, 審査の要旨



学位論文審査の要旨

博士の専攻分野の名称 博士 (情報科学)

氏名 宮崎 善行

審査担当者 主 査 准教授 筒井 弘
副 査 教 授 大鐘 武雄
副 査 教 授 齊藤 晋聖
副 査 教 授 西村 寿彦

学位論文題名

音響的類似単語の雑音ロバストフレーズ音声認識に関する研究
(A Study on Robust Phrase Speech Recognition for Noisy Acoustically Similar Words)

音声認識技術は、音声入力インターフェースとして機器の音声制御、特にハンズフリー操作などの実現に必須の技術である。近年、音声認識を搭載したスマートフォン、タブレット端末、ロボットなどが商品化され、音声により制御可能な製品が普及しつつある。一般に、スマートフォンなどで利用される音声認識は、連続音声認識である。一方、機器制御などを行う音声認識には、耐雑音性能を向上し、特定フレーズのみを認識するようなフレーズ音声認識方式が用いられる。本論文ではこの雑音ロバストフレーズ音声認識を対象とする。

連続音声認識方式では大語彙音声認識が可能であり、カーナビゲーションシステムなどにおいて地名などの固有名詞を認識できるとして高い有効性を示す。盛んに研究開発が推進されているが、近距離マイクの利用や、スイッチや特定フレーズ発話など音声認識のトリガが必要となることが多い。これは、雑音環境下での認識率向上が困難なためである。そのため、音響処理段階において雑音除去や高性能指向性マイクを利用するなどの対策が必要になることが多い。

一方フレーズ音声認識方式は、認識対象フレーズ数が数 100 程度と制限があるが、雑音環境下でのロバスト音声認識を効率良く実現可能であり、認識トリガ不要でマイクから 2~3 m 離れていても音声認識が可能である。実用性を考慮する場合、雑音環境下でかつマイクから離れていても高い認識率を維持する必要がある。このようなシステムの需要は極めて高く、著者はその音声認識エンジンモジュールの開発を通じて、雑音ロバストフレーズ音声認識の実用化を推進した。

しかし、雑音ロバストフレーズ音声認識を利用しても雑音環境下で判別困難なフレーズが存在する。このようなフレーズは音響的類似単語であり、「財布」と「ライス」など、母音の並びが同じで、異なる子音を含むものである。そこで本論文では、音響的類似単語の雑音ロバストフレーズ音声認識手法を提案する。

提案雑音ロバストフレーズ音声認識は、音声信号から抽出した特徴量に基づき HMM (隠れマルコフモデル, hidden Markov model) を用いて発話フレーズを推定する。音響的類似単語については、雑音環境下で認識精度の低い単語のペアに対して、雑音ロバスト音声認識を行ったのち、音響的に類似していない区間を抽出する。その後その区間のみを、再度雑音ロバスト音声認識をする。

適切に各種ノイズを付加した音声によるマルチコンディション学習を利用し、雑音環境下にて評価した結果、提案手法により難認識語の認識率向上が可能であることを示した。

これを要するに、著者は、フレーズ音声認識に関して、雑音に強いフレーズ音声認識システムの実用化を推進するとともに、音響的に類似する難認識語について、その認識率改善可能な手法およびシステムを提案・開発・実現・評価した。これにより、雑音環境下で認識率が低下する難認識語のフレーズ音声認識について、その認識率を提案手法により改善可能であることを明らかにした点で、音声認識システムの分野に貢献するところ大なるものがある。

よって著者は、北海道大学博士 (情報科学) の学位を授与される資格あるものと認める。