

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Robust Speech Recognition Based on Non-extensive Statistics
著者(和文)	ヒルマンパルデッド
Author(English)	Hilman Pardede
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第9248号, 授与年月日:2013年6月30日, 学位の種別:課程博士, 審査員:篠田 浩一
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第9248号, Conferred date:2013/6/30, Degree Type:Course doctor, Examiner:Koichi Shinoda
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	審査の要旨
Type(English)	Exam Summary

論文審査の要旨及び審査員

報告番号	甲第		号	学位申請者氏名		Hilman Ferdinandus Pardede	
		氏名		職名		氏名	職名
論文審査 審査員	主査	篠田浩一		教授	審査員	藤井敦	准教授
	審査員	佐藤泰介		教授			
		村田剛志		准教授			
		杉山将		准教授			

論文審査の要旨 (2000 字程度)

本論文は「Robust Speech Recognition Based on Non-extensive Statistics」と題し、7章より構成されている。

第1章「Introduction」では、本研究の背景、目的、概要について述べている。現在の音声認識技術は、静かな環境では高い性能を示すものの、話者や発声スタイルの違いや、背景雑音や回線歪みの存在など、様々な要因によりその性能が劣化することを述べ、本論文では特に背景雑音や回線歪みに対して頑健な音声認識に焦点をあてることを述べている。さらに、従来の耐雑音音声認識手法が Boltzman-Gibbs 統計に基づく示量性の統計理論を基礎としており、その場合、複雑な事象である音声のモデル化には限界があることを述べている。そして、本論文では、この問題の解決のために非示量性の統計理論を導入し、スペクトル平均正規化とスペクトル減算の2つの耐雑音処理において新しい方式を提案することを述べた上で、本論文全体の構成について述べている。

第2章「Statistical speech recognition: an overview」では、統計的音声認識手法について概観している。特に代表的な特徴量であるメル周波数ケプストラム係数(MFCC)と代表的な認識方式である隠れマルコフモデル(HMM)について、より詳細に説明している。

第3章「Noise robust speech recognition」では、第2章で述べた音声認識手法に用いられる耐雑音技術について、その代表的な手法を紹介している。音声モデルを雑音下音声に適応させるモデル適応、雑音を除去するスペクトル減算、雑音の影響を軽減する特徴量正規化などについて述べた上で、それらの実用において、理論的には生じないはずの音声と周囲雑音や回線歪との相関を表す交差項がしばしば出現し、その扱いが課題となっていることを述べている。

第4章「Non-extensive statistics: a review」では、まず、ある種の複雑な事象では、その構成要素やそれらの間の依存関係が不明であり、従来の Boltzman-Gibbs 統計に基づく示量性の統計理論の適用には限界があることを述べている。そして、この問題に対し、近年、Tsallis らにより非示量性の統計理論が提案されていることを述べ、その理論的枠組みについてその概略をまとめた上で、宇宙物理学、核物理学、株式市場分析、タンパク質構造解析などの様々な分野でその有効性が示されていることを述べている。

第5章「Feature normalization based on non-extensive statistics」では、特徴量正規化手法として、Tsallis らの非示量統計に基づいた q -対数スペクトル正規化法を提案し、それにより音声と回線歪みの間の相関を取り込んだ定式化が可能であることを述べている。雑音下音声の認識により評価した結果、特徴量正規化の代表的な手法であるケプストラム正規化と比較して大幅な改善が見られたことを報告している。さらに、相関の度合いを制御する q パラメータの設定方法が今後の課題としてあげられることを述べている。

第6章「Q-Gaussian based spectral subtraction」では、雑音下音声为非示量性統計に基づく q -ガウス分布に従うことを仮定した新しいスペクトル減算法を提案している。また、従来の非線形スペクトル減算法は提案法の特別な場合であり、従来意味づけが困難だったパラメータを非示量性統計の枠組みの中で自然に説明できることを述べている。さらに、様々な種類や大きさの雑音を用いた雑音下音声の認識による評価において、提案法が、大多数の条件下で、従来法に比べ顕著に高い性能を示したことを述べている。

第7章「Conclusions and future work」では、本論文の結論及び貢献について述べ、今後の課題についてまとめている。

以上で述べたように、本論文では、非示量性統計を用いた、背景雑音や回線歪に対し頑健な耐雑音音声認識技術を提案し、従来手法に比べ効果があることを比較評価実験で確認している。本論文で得られた成果は、雑音下音声に対する音声認識技術の性能向上に寄与し、その応用範囲を大きく広げるものであり、工学上、寄与するところが大きい。よって本論文は博士(工学)の学位論文として十分な価値があるものと認める。