

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Efficient Voice Activity Detection and Speech Enhancement Algorithms based on Spectral Features
著者(和文)	MaYanna
Author(English)	Yanna Ma
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9632号, 授与年月日:2014年9月25日, 学位の種別:課程博士, 審査員:西原 明法,國枝 博昭,山田 功,府川 和彦,篠田 浩一,入野 俊夫
Citation(English)	Degree:, Conferring organization: Tokyo Institute of Technology, Report number:甲第9632号, Conferred date:2014/9/25, Degree Type:Course doctor, Examiner:,,,,,
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		Ma Yanna	
論文審査 審 査 員		氏 名	職 名		氏 名	職 名	
	主査	西原明法	教授	審査員	篠田浩一	教授	
	審査員	國枝博昭	教授		入野俊夫	和歌山大学 教授	
		山田 功	教授				
		府川和彦	教授				

論文審査の要旨（2000 字程度）

本論文は「Efficient Voice Activity Detection and Speech Enhancement Algorithms based on Spectral Features(スペクトル特徴を用いた効率的な音声区間検出と音声強調アルゴリズム)」と題し、英文 5 章よりなっている。

第 1 章「Introduction(序論)」では、音声コミュニケーションにおける基本的な信号処理として、音声区間検出と音声強調の重要性を述べ、過去に多くの研究がなされてきたが、特に低 SN 比の環境では未だ十分な特性が得られておらず、これらを改善するアルゴリズムを提案することが本研究の目的であると述べている。

第 2 章「Literature Review (先行研究調査)」では、まず音声検出に関して、規格化された 3 種類の手法、統計モデルに基づく手法、そして最近提案された長期信号変動性(Long-Term Signal Variability)に基づく手法を概説し、その特徴を述べている。次に音声強調に関して、スペクトル差分法、部分空間法、統計モデル法、ウィーナー型の 4 種類の手法の概要とそれぞれの特徴を述べている。これらから、それぞれが利点のみでなく欠点や制約をもち、いずれも改善の余地があることを示している。また、本研究の実験で使用する音声信号や雑音のデータベースとして、TIMIT、AURORA、NOISEX-92、NOIZEUS を簡単に紹介している。

第 3 章「Efficient Voice Activity Detection Algorithm using Long-term Spectral Flatness (長期スペクトル平坦性を用いる効率的な音声区間検出)」では、まず音声のスペクトル特徴に基づき、長い窓内で短時間スペクトルを平均したスペクトル平坦性があることを述べている。平坦性の評価にはスペクトルの幾何平均と算術平均の比を用い、判別の閾値は対象としているフレームより前のフレームでの音声と無音声それぞれの長期スペクトル平坦性を用いて適応的に決定している。データベース上の音声信号に 12 種類の雑音を、-10dB から 10dB の SN 比で加え、精度と誤差率の観点で既存の手法と比較し、提案手法で実際に雑音に埋もれた音声と音声の無い雑音とを的確に区別でき、既存手法に対して優位性があることを確認している。特に非定常なマシンガン雑音や多くの音声からなっている人混み雑音等の、これまで困難であった雑音に対して良い検出特性が得られると述べている。

第 4 章「A Modified Wiener Filtering Method Combined with Wavelet Thresholding Multitaper Spectrum for Speech Enhancement (修正ウィーナーフィルタとウェーブレット閾値処理した複数窓スペクトルによる音声強調)」では、ウェーブレット閾値処理した複数窓スペクトルを真の音声とみなし、それを制約としてウィーナー利得関数に課した手法を提案している。音声強調においては、雑音抑圧と音声信号歪みの増大が二律背反するため、強調された音声スペクトルを真の音声スペクトルの 2 倍以下に収めることで、音声了解度が最大化すると述べている。その制約をウィーナー利得関数に課すことで、了解度の高い音声強調が可能になると述べている。音声了解度は人間による聞き取りテストで判定されるべきものであるが、複数の客観指標の重み付け和や SN 比減少で良く推定できることから、8 種類の異なる雑音に埋もれた音声信号を提案手法で強調した結果を複合指標と SN 比減少の観点から既存手法と比較し、特に低 SN 比の(雑音が多い)信号に対して良好な結果を得ている。

第 5 章「Conclusion and Discussions (結論と議論)」では、本論文で得られた結果を総括し、議論をまとめている。

以上を要するに、本論文は、雑音に埋もれた音声信号の有無をより正確に検出し、音声の歪みを抑えつつ雑音を低減する効率的な手法を提案したもので、種々の音声信号処理の前処理として有用であり、工業上並びに工学上貢献するところが大きい。よって、我々は本論文が博士(工学)の学位論文として、十分に価値あるものと認める。

注意：「論文審査の要旨及び審査員」は、東工大リサーチリポジトリ(T2R2)にてインターネット公表されますので、公表可能な範囲の内容で作成してください。