

氏名	Atanu Saha
博士の専攻分野の名称	博士（工学）
学位記号番号	博理工甲第 869 号
学位授与年月日	平成 24 年 3 月 22 日
学位授与の条件	学位規則第 4 条第 1 項該当
学位論文題目	Robust Speech Enhancement in Adverse Environments (不利な環境でのロバスト音声強調)
論文審査委員	委員長 教授 島村 徹也 委員 教授 久野 義徳 委員 教授 前川 仁 委員 准教授 橋口 博樹

論文の内容の要旨

Many speech enhancement algorithms have been proposed over the last few decades. Most of these algorithms improve the quality of speech, but at the expense of introducing speech distortion, which in turn may impair intelligibility and perceptuality of speech at various degrees. The solution to this common problem of speech enhancement may also depend on the application at hand and the characteristics of the noise source. The noise could be stationary or nonstationary. In nonstationary environments, where the spectral characteristics of the noise might be changed dramatically, to solve the above problem is one of the most challenging tasks. It is therefore imperative to design effective speech enhancement algorithms to remove the noise without introducing any perceptible distortion in the signal of speech degraded by highly nonstationary noises.

To minimize speech distortion, we propose a constrained minimization approach that considers the speech and noise distortion as constraints. The approach exploits the masking properties to determine the audible spectral components. These spectral components are attenuated in the proposed approach as a predefined attenuation factor. For better perceptual speech quality, we propose generalized Gamma distributed Bayesian estimators under speech presence probability (SPP). The SPP is also computed based on signal-to-noise ratio (SNR) discrepancy measure.

We further propose a family of Bayesian estimators for speech enhancement based on a bivariate Gamma model where it is assumed that the spectral components are correlated with each other. Three different kinds of Bayesian estimators are derived from the statistics of the bivariate Gamma model. The proposed algorithms require the estimation of noise spectrum. We therefore introduce a noise spectrum estimation technique to better deal with the nonstationary noises.

The experimental results, based on objective and subjective measures, show the superiority of the proposed algorithms over state-of-the-art algorithms.

論文の審査結果の要旨

当学位論文審査委員会は、平成 24 年 2 月 28 日に論文発表会を公開で開催し、活発な質疑とともに論文内容の審査を行った。以下に、審査結果の要約を示す。

本論文は、雑音により品質劣化した音声信号から元のクリーンな音声信号を強調する音声強調問題において、特に一つのマイクロフォンのみを用いて問題解決方法を検討したものである。我々の日常的な音声信号の利用時には、しばしば環境雑音が混入する。本論文では、様々な混入雑音を考慮し、耐性のある四つの音声強調アルゴリズムを提案し、それらの有効性をそれぞれ検証している。

まず第 1 章では、一つのマイクロフォンのみを用いた音声強調の従来研究を整理し、代表的なアルゴリズムを五つのクラス：スペクトル引き算、ウィナーフィルタ、統計モデル、部分空間、カルマンフィルタに大別している。それぞれのクラスの長所短所から、本研究での方向性に言及し、以降に続く 2, 3, 4, 5 章での提案アルゴリズム導出の目的を述べている。

第 2 章では、音声強調アルゴリズムに共通して利用可能な雑音推定アルゴリズムを検討している。従来の最小統計法の概念、再帰型平均法の概念、最適平滑化の概念を統合する新しい雑音推定アルゴリズムを導出し、それを代表的な音声強調アルゴリズムであるスペクトル引き算と組み合わせることにより、時変雑音追従型の音声強調アルゴリズムを実現している。計算機シミュレーション実験では、従来手法に比べ提案アルゴリズムが大幅な特性改善を与える事実が示されている。

第 3 章では、ウィナーフィルタの改良アルゴリズムに取り組んでいる。人間の聴覚が有するマスキング効果をウィナーフィルタ設計に組み入れるアルゴリズムを検討し、聴覚的に良好な雑音抑圧と音声品質の確保を実現する音声強調アルゴリズムの導出に成功している。試聴実験を行い、提案アルゴリズムの有効性が確認されている。

第 4 章では、音声強調のための確率モデルが検討されている。第 1 章および 2 章で取り上げたアルゴリズムは基本的には線形フィルタ処理であるが、非線形フィルタ処理へ拡張し、ベイズ推定器として音声強調を実現するとき、残留雑音が耳障りでなくなる性質が報告されている。本章では、このベイズ推定の概念を拡張し、ガンマ分布モデルを利用することにより、音声強調アルゴリズムの特性改善を図っている。従来アルゴリズムとの比較実験を行い、本章でのガンマ分布アルゴリズムが、多くの計算量を必要とする反面、良好な音声強調を実現できる事実を明らかにしている。

第 5 章では、第 4 章でのガンマ分布アルゴリズムのさらなる特性改善に取り組んでいる。各パラメータを確率密度関数から統計的に最適化させる工夫をし、改良ガンマ分布アルゴリズムを導出している。実験では、ガンマ分布アルゴリズムよりより聴覚的に好ましい結果が得られている。

第 6 章は、本論文のまとめである。提案する四つの音声強調アルゴリズムの特性を整理し、実際の音声処理システムにおいてそれぞれのアルゴリズムをどのように利用可能であるかの示唆を与えている。

本論文は、以上に述べたように、雑音混入音声信号からの音声強調問題に対し、従来方式に特性改善を与えるアルゴリズムを提案し、その実験的評価を行っている。本論文の結果は、2編のレフリー付学術雑誌に採録され、また国際学会での発表で公表されている。

以上のように、本論文は新しい音声強調アルゴリズムの提案と、その有効性を検証した論文であり、音声情報処理分野への貢献が大である。したがって、博士（工学）の学位にふさわしい内容を持つものと判断し、審査委員会として「合格」の判定を行った。