

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 18 日現在

機関番号：55503

研究種目：研究活動スタート支援

研究期間：2014～2015

課題番号：26880027

研究課題名(和文) 高次統計量追跡に基づく音声品質を自動制御する雑音抑圧手法に関する研究

研究課題名(英文) Study of noise reduction which can automatically control speech quality based on higher-order statistics pursuit

研究代表者

宮崎 亮一 (Miyazaki, Ryoichi)

徳山工業高等専門学校・情報電子工学科・助教

研究者番号：40734728

交付決定額(研究期間全体)：(直接経費) 2,100,000円

研究成果の概要(和文)：本研究では、様々な条件における目的音声信号の高次統計量の推定精度について評価した。また、信号の複素スペクトルが一般化ガウス分布に従うと仮定した場合における、目的音声信号の高次統計量推定の有効性についても検証した。

客観評価実験より、入力SNRが0 dBより良い場合に関しては、高次統計量の推定精度(正規化誤差)が0.2以下であり、どの雑音の場合にも高精度に推定できることを確認した。

研究成果の概要(英文)：In this study, I evaluated the accuracy of speech higher-order statistics estimation in various environments. Also I inspected the efficacy of speech higher-order statistics estimation with the assumption that the complex spectra obey the generalized gaussian distribution. From the objective evaluation, I confirm that the proposed method can estimate the higher-order statistics with high accuracy in each noise environment.

研究分野：音響信号処理

キーワード：雑音抑圧 高次統計量 補聴システム

1. 研究開始当初の背景

聴覚は我々の生活の中で、コミュニケーションを取り合う上で非常に重要である。しかし、人間の聴力は年齢を重ねると共に低下し、また、日本では高齢化が急速に進展しているため、難聴者の聴覚をサポートする補聴器の開発が急務である。しかし、日本国内では難聴者の4人に1人しか補聴器を使用しておらず、視覚をサポートするメガネやコンタクトレンズと比べて、圧倒的に普及率が低い。補聴器の普及率が低い原因は価格、デザイン等様々な要因が考えられるが、その中でも現在の補聴器は、単純に音圧を増幅させるだけのものが多く、難聴者の聞き取りやすさをサポートするように設計されていない。そのため、申請者は背景雑音を抑圧し、目的信号を強調する雑音抑圧技術を補聴器に導入し、補聴器使用者の聞き取りやすさを向上させる事を目指す。

雑音抑圧後の音声の品質に関しては、従来、ケプストラム歪みや信号対歪み比という評価尺度を用いることが一般的である。しかし、これらの評価尺度は観測信号の目的音声成分と、雑音抑圧信号の目的音声成分が必要となる。実環境下で雑音が混入していないクリーン音声信号を用意することは不可能である。すなわち、実環境下で雑音抑圧後の音声の品質を評価することはできない。また、現在の雑音抑圧技術を導入したシステムの多くは、ある固定のパラメータを用いて雑音抑圧を行う。そのため、システムを使用する環境が変われば、背景雑音のレベルや雑音の種類が変わり、雑音抑圧後の音声信号が極端に歪む、または、ほとんど背景雑音を抑圧できないという問題が発生する。以上より、様々な環境下でシステムを使用しても、雑音抑圧後の音声の歪み量が一定以下になるように自動制御する枠組みが望まれる。

2. 研究の目的

本研究の目的は、雑音抑圧後の目的音声の品質を自動制御可能な補聴器の開発である。これを実現するためには、教師なしで雑音抑圧後の目的音声の品質を評価できる新たな尺度が必要である。そこで本研究では、雑音抑圧処理前後の目的音声の高次統計量の変化に着目し、どのような統計量が人間の聴覚や、従来の教師あり評価尺度と相関があるのかを明らかにする。次に、音声の高次統計量の変化を追跡することによって、音声歪み量を自動的に制御可能な雑音抑圧手法を提案する。そして、提案手法を評価し、提案技術は補聴器の開発に貢献できることを確認する。

本研究の特色は、教師なしの音声品質の評価尺度を新たに考案する点及び、雑音抑圧後の音声信号の品質を自動制御する点である。本研究によって、あらゆる環境下で雑音抑圧処理を行っても、目的音声の品質が一定以上に制御でき、補聴器使用者がコミュニケーション

を取りやすくなることが期待できる。また、これまでの雑音抑圧手法は目的音声に従う分布を事前情報として用い、出力信号に従う分布と仮定した事前分布の距離を最小化するものであった。統計学の分野ではこれが最適解であるが、人間の聴覚に最適な信号が出力される根拠はない。しかし、本研究では雑音抑圧前後の音声信号の統計量を追跡し、聴覚に相関のある客観評価値に基づいて直接音質を制御できるため、聴覚に適した信号が出力されることが期待でき、今後の雑音抑圧手法の研究にも繋がる大きな学術的意義を持つと言える。

3. 研究の方法

雑音抑圧後の音声の品質に関しては、従来、ケプストラム歪みや信号対歪み比という評価尺度を用いることが一般的である。しかし、これらの評価尺度は観測信号の目的音声信号と、雑音抑圧処理後の目的音声信号が必要となるため、実環境下で音声の品質をリアルタイムに測ることは不可能である。そこで、このような教師情報を必要としない実環境で測定可能な新たな音声品質評価尺度を提案する事を目指す。

一般に、目的音声信号と雑音信号のような加法信号の確率密度分布から、目的音声信号の統計量を分解することは図1に示すように逆畳み込み問題を解くことに相当し、最尤推定などの手法では解析的に解くことができない。しかし、申請者の近年の研究において、モーメントとキュムラントを交互に用いることにより、目的音声信号の高次統計量をブラインドに推定することが可能であると報告されている[Miyazaki, et al., HSCMA2011]。

本研究では、様々な条件における目的音声信号の高次統計量の推定精度について検証した。また、信号の複素スペクトルが一般化ガウス分布に従うと仮定した場合における、目的音声信号の高次統計量推定の有効性についても検証した。

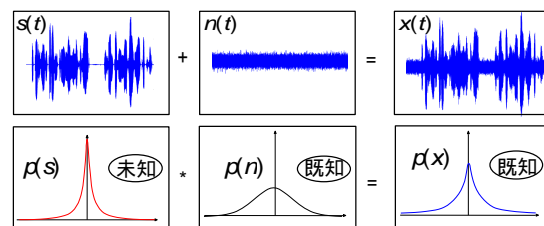


図1 問題設定

4. 研究成果

研究成果の一つとして、様々な雑音環境下における目的音声のパワースペクトルの高次モーメント推定精度を検証した。入力 SNR は -10, 0, 10 dB の三種類、雑音信号には白色ガウス雑音、パブル雑音、実収録した駅の雑音を用いた。また、高次モーメントの評価には正規化誤差を用いた。

実験結果より、入力 SNR が 0 dB 以上と良

い条件の場合においては，正規化誤差が 0.2 以下であり，雑音の種類によらず高い推定精度であることを明らかにした．一方，入力 SNR が -10 dB の場合においては，モーメント推定精度が極端に悪くなる（図 2～図 4 参照）．

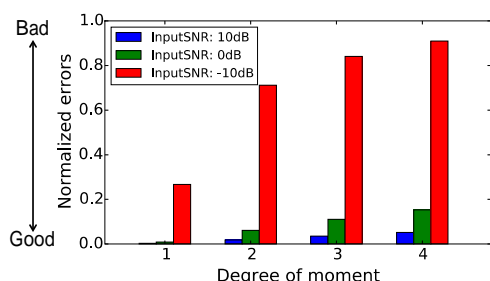


図 2 目的信号のパワースペクトルの高次モーメント推定精度（白色ガウス雑音）

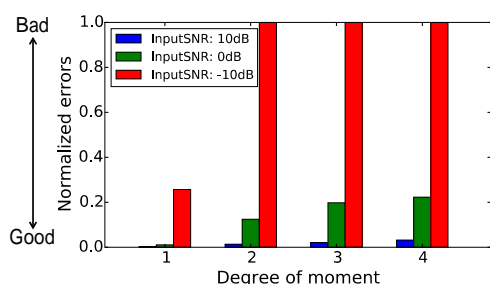


図 3 目的信号のパワースペクトルの高次モーメント推定精度（バブル雑音）

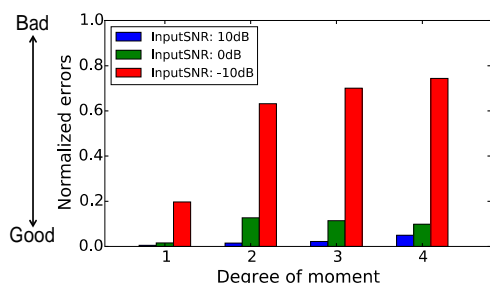


図 4 目的信号のパワースペクトルの高次モーメント推定精度（駅雑音）

次に，上記で述べたモーメントの推定精度を向上させるために，信号の複素スペクトルを一般化ガウス分布でモデリングする手法を提案した．一般化ガウス分布と一般化ガウス分布の m 次モーメントはそれぞれ以下の式で表される．

$$P(a) = \frac{p}{2\alpha\Gamma(1/p)} e^{-(|a|/\alpha)^p}$$

$$\mu_m(a) = \int_{-\infty}^{\infty} a^m P(a) da = \alpha^m \frac{\Gamma((m+1)/p)}{\Gamma(1/p)}$$

ここで， α は尺度母数， p は形状母数であり，一般化ガウス分布におけるモーメントは形状母数 p を決めることによって，求めることができる．信号から形状母数 p を推定するために以下の式を満たす p をニュートン法に寄り推定した [Roenko, 2014] ．

$$\frac{\mu_{r_1}}{(\mu_{r_2})^{r_1/r_2}} = \frac{\Gamma((r_1+1)/p)}{\Gamma^{r_1/r_2}((r_2+1)/p)\Gamma^{(1-r_1/r_2)}(1/p)}$$

ここで， r_1 ， r_2 はそれぞれモーメントの次数を表す．

一般化ガウス仮定におけるモーメント推定精度は，予想に反して向上しなかった．その原因として，モーメントの推定精度は r_1 ， r_2 の設定に大きく依存することが実験より明らかになっている．どのような値に設定することがモーメントの推定精度を向上させるかの議論は今後も続けて行きたいと考えている．

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕(計 0 件)

〔学会発表〕(計 11 件)

久保真太郎，宮崎亮二，様々な雑音抑圧手法による音声認識率と内部パラメータの関係，第 18 回日本音響学会関西支部若手研究者交流研究発表会，2015 年 12 月 13 日，関西大学 100 周年記念会館（大阪府吹田市）．

垣元宏太，宮崎亮二，様々な条件での音声モーメント推定精度の評価とその改善，第 18 回日本音響学会関西支部若手研究者交流研究発表会，2015 年 12 月 13 日，関西大学 100 周年記念会館（大阪府吹田市）．

垣元宏太，宮崎亮二，雑音中からの音声モーメントの推定精度に関する評価，日本高専学会第 21 回年会，2015 年 8 月 29 日～30 日，徳山工業高等専門学校（山口県周南市）．

久保真太郎，宮崎亮二，音声認識性能を最大限に引き出すスペクトル減算法のパラメータに関する調査，日本高専学会第 21 回年会，2015 年 8 月 29 日～30 日，徳山工業高等専門学校（山口県周南市）．

宮崎亮二，MAP 推定器におけるミュージカルノイズフリー理論の確立，日本音響学会講演論文誌，2014 年 9 月 3 日～5 日，北海学園大学（北海道札幌市）．

Shunsuke Nakai, Hiroshi Saruwatari, Ryoichi Miyazaki, Satoshi Nakamura, Kazunobu Kondo, Theoretical Analysis of Biased MMSE Short-Time Spectral Amplitude Estimator and Its Extension to Musical-Noise-Free Speech Enhancement,

4th Joint Workshop on Hands-free Speech Communication and Microphone Arrays (HSCMA2014), May 12-14 2014, Nancy (France).

その他 5 件

〔図書〕(計 1 件)

Hiroshi Saruwatari, Ryoichi Miyazaki,
“Statistical Analysis and Evaluation of Blind
Speech Extraction Algorithm (in book
chapter),” *Advanced in Modern Blind Source
Separation Techniques: Theory and
Applications* (Ganesh R. Naik and Wenwu
Wang Eds.), Springer, pp.291--322, 2014.

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

〔その他〕

APSIPA Newsletter 執筆

“Musical-Noise-Free Noise Reduction Theory
Based on Higher-Order Statistics”

6 . 研究組織

(1) 研究代表者

宮崎 亮一 (MIYAZAKI, Ryoichi)

徳山工業高等専門学校・情報電子工学科・
助教

研究者番号 : 40734728