

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 15 日現在

機関番号：10106

研究種目：挑戦的萌芽研究

研究期間：2012～2014

課題番号：24656239

研究課題名(和文)サブバンド間の相互推定による雑音抑圧

研究課題名(英文)speech enhancement using parameter estimation between frequency subbands

研究代表者

中垣 淳(NAKAGAKI, Atsushi)

北見工業大学・工学部・講師

研究者番号：90250539

交付決定額(研究期間全体)：(直接経費) 2,000,000円

研究成果の概要(和文)：本研究では騒音環境下で円滑な音声対話を維持するために、周波数サブバンド間で特徴パラメータの推定を行う雑音抑圧法の開発を行った。開発手法は音声信号の帯域を二分割し、SNRの高い低域サブバンド信号は従来法で雑音抑圧を行うが、SNRの低い高域サブバンド信号は低域サブバンドの情報を用いて雑音抑圧を行う。SNRの高いサブバンドの情報を使うことで騒音レベルに対して頑健な雑音抑圧をめざした。開発手法では雑音抑圧に用いる情報として線スペクトル対を採用し、学習した低域・高域サブバンド間の関係を用いて、低域のLSPは高域のLSPに変換される。

研究成果の概要(英文)：In order to maintain speech communication in noisy environments, a new speech enhancement method using parameter estimation between frequency subbands is developed. This method divides a speech signal into a low subband signal and a high subband signal. While the low subband signal which has higher SNR is enhanced by the conventional method, the high subband signal which has lower SNR is enhanced using information obtained from low subband signal. The line spectrum pairs (LSPs) are employed as feature parameters of subband signals. The low subband LSPs are transformed into the high subband LSPs by using the relation between two subbands which was learned preliminarily.

研究分野：音声信号処理

キーワード：雑音抑圧 音声強調 線スペクトル対 サブバンド信号

1. 研究開始当初の背景

(1) 携帯電話などのモバイル音声サービスでは、利用者の使用環境により様々な雑音が存在する。環境雑音は会話を阻害する要因のひとつであり、特に高雑音環境で円滑な会話を維持するためにはノイズサプレッサの導入が不可欠である。

(2) スペクトルサブトラクションなどの従来の雑音抑圧法は、雑音レベルが一定以上に大きくなると雑音抑圧性能が著しく劣化する。従来の雑音抑圧法には、音声信号の抑圧により生じるひずみと、雑音成分の引き残しによる残差を同時に小さくすることはできないトレードオフが存在する。そのため、高雑音環境では音声信号のひずみも大きくなり使用に限界がある。そのため高雑音環境下でも高品質の通話を可能にする新しい雑音抑圧法の開発が要求されている。

2. 研究の目的

(1) 本研究の目的は、観測信号を帯域分割して得られたサブバンド信号を独立に処理するのではなく、各サブバンド信号の情報を相互にうまく利用する雑音抑圧方式を開発することである。音声信号や雑音の周波数領域での偏在により、各サブバンド信号のSNRにはばらつきが生じる。高SNRのサブバンド信号の情報のみを使って、低SNRのサブバンド信号の雑音を抑圧することができれば、音声のひずみと残差のトレードオフを解決できる有効な手段となる。

(2) サブバンド信号間で相互に利用する情報として線スペクトル対 (LSP) を検討する。線スペクトル対は線形予測分析に基づく周波数領域の特徴パラメータであり、優れた量子化特性を有する。観測信号から推定した線スペクトル対を用いた雑音抑圧法を開発し、その性能を評価する。

(3) サブバンド信号間で情報のやり取りを行うための変換方式を検討する。事前に用意された音声データを用いて、各サブバンド信号の線スペクトル対を求め、一方のサブバンドの線スペクトル対から他方のサブバンドの線スペクトル対を想起する写像関係を学習する。その写像関係を未知入力音声の低SNRサブバンド信号の雑音抑圧に適用する。

3. 研究の方法

(1) 線スペクトル対を用いた雑音抑圧法を開発する。この雑音抑圧法では信号は線形予測係数の線スペクトル対による偏微分行列によって変換される。偏微分行列は、線スペクトル対を中心周波数とするフィルタバンクとして機能する。そのため、同じ帯域幅で信号を帯域分割するのではなく、スペクトルのピークの位置、鋭さに応じて帯域幅が決まることになる。また、偏微分行列は観測信号

の相関行列を対角化するという、雑音抑圧において望ましい性質を有している。この方式による処理手順を図1に示す。

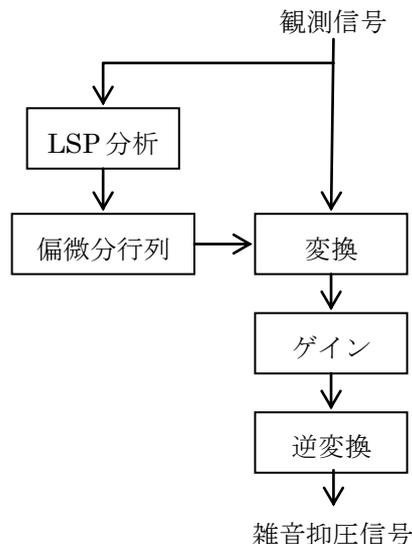


図1 線スペクトル対を用いた雑音抑圧

(2) サブバンド間での線スペクトル対の変換を用いた雑音抑圧法を開発する。この方式による雑音抑圧の手順を図2に示す。観測信号を低域・高域二つのサブバンド信号に分離する。比較的高SNRな低域サブバンド信号は、(1)の線スペクトル対を用いた雑音抑圧法を適用する。また、線スペクトル対は事前に学習した変換行列を用いた写像変換により高域サブバンド信号の線スペクトル対に変換される。低域サブバンド信号から推定した高域サブバンド信号の線スペクトル対を用いて低域サブバンドの雑音は抑圧される。

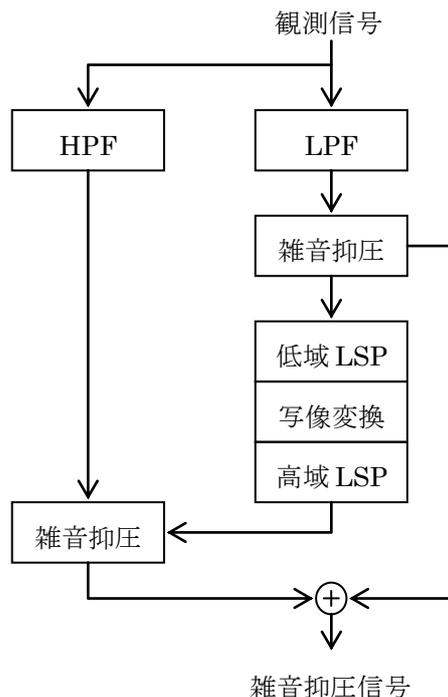


図2 写像変換を用いた雑音抑圧

4. 研究成果

(1) 線スペクトル対を用いた雑音抑圧では、雑音の白色化処理の有無による二種類の手法を検討した。雑音の白色化処理を行わない手法を LSP1、行う手法を LSP2 とする。評価実験には ATR 日本語音声データベースの男性、女性各 5 名の音声データを使用した。雑音は、白色雑音、有色雑音、複合正弦波信号の三種類を用いた。有色雑音は 400Hz にピークをもつ 2 次の全極型システムに白色雑音を入力して生成した。複合正弦波信号は、200, 400, 600, 800, 1000Hz の正弦波信号を加算して生成した。

評価実験の結果を表 1~3 に示す。雑音抑圧した音声品質の評価は、客観的評価手法である PESQ を用いた。PESQ は 0.5 から 4.5 の値を取り、数値が大きいほど品質がよいと評価される。また、比較のために信号部分空間法 (SSA) による雑音抑圧結果も示した。評価実験の結果は、白色雑音では、LSP1、LSP2 とともに従来法である SSA とほぼ同程度のスコアを示している。一方、有色雑音に対しては、LSP2 が、SSA や LSP1 と比べて高いスコアを示している。複合正弦波信号に対しては、女性話者音声では LSP2 が高いスコアを示しているが、男性話者音声では LSP1 と同程度である。

以上より、線スペクトル対は雑音抑圧に利用する特徴パラメータとして有効であることが確認できた。

表 1 白色雑音に対する雑音抑圧性能
女性話者

SNR	未処理	SSA	LSP1	LSP2
20dB	2.45	3.16	3.17	3.16
15dB	2.10	2.82	2.83	2.82
10dB	1.78	2.46	2.46	2.45

男性話者

SNR	未処理	SSA	LSP1	LSP2
20dB	2.56	3.43	3.43	3.43
15dB	2.23	3.10	3.11	3.10
10dB	1.92	2.75	2.77	2.76

表 2 有色雑音に対する雑音抑圧性能
女性話者

SNR	未処理	SSA	LSP1	LSP2
20dB	3.16	3.23	3.28	3.37
15dB	2.77	2.83	2.89	2.98
10dB	2.41	2.43	2.50	2.57

男性話者

SNR	未処理	SSA	LSP1	LSP2
20dB	3.51	3.65	3.67	3.69
15dB	3.21	3.34	3.36	3.42
10dB	2.88	2.99	3.01	3.12

表 3 複合正弦波に対する雑音抑圧性能
女性話者

SNR	未処理	SSA	LSP1	LSP2
20dB	2.90	3.24	3.21	3.33
15dB	2.55	2.88	2.85	2.98
10dB	2.24	2.56	2.49	2.64

男性話者

SNR	未処理	SSA	LSP1	LSP2
20dB	3.08	3.61	3.60	3.57
15dB	2.79	3.37	3.37	3.34
10dB	2.54	3.08	3.07	3.09

(2) サブバンド間での変換を用いた雑音抑圧では、線スペクトル対の学習について検討した。まず、線スペクトル対を雑音抑圧に使用するには、40 次程度の高い分析次数に設定する必要があり、線形写像のための変換行列の学習がうまくいかず、周波数サブバンド間での写像を行うには問題があった。そこで、音声信号のスペクトル包絡を近似できる程度まで線スペクトル対の次数を下げ学習を行った。この学習した結果を用いて、低域サブバンドから高域サブバンドへの線スペクトル対の変換を行い、雑音抑圧に適用した。しかし、評価実験では雑音抑圧の効果は認められるが、写像変換を用いない場合よりも低い性能しか得られなかった。この原因は、次数を低く設定したため、線スペクトル対にスペクトル微細構造の情報が反映されていないためである。そこで、高域サブバンドのスペクトル微細構造は、低域サブバンドの微細構造を外挿して生成する処理を追加した。これにより、低次の線スペクトル対のみを利用する場合と比べて雑音抑圧性能が向上した。しかし、写像変換を用いない場合の性能には達していない。

(3) 現状では、予想した雑音抑圧性能を得られていない。この原因としては、サブバンド間の情報のやり取りがうまくいっていないことが挙げられ、これを解決できれば雑音抑圧性能が大幅に改善する余地はある。今後は、現在採用している線形写像に代えて非線形写像を導入するなど、サブバンド間の写像方式を検討する予定である。

<引用文献>

- ① Y. Ephraim and H. L. V. Trees, A subspace approach for speech enhancement, IEEE Speech & Audio Process., vol. 3, no. 4, pp. 251-266, July 1995
- ② ITU-T recommendation P. 862, evaluation of speech quality (PESQ): an objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs, Feb. 2001.

5. 主な発表論文等

〔雑誌論文〕（計 0件）

〔学会発表〕（計 3件）

- ① 中垣淳、LSP による変換行列を用いた雑音抑圧法、電子情報通信学会総合大会、2015年3月13日、滋賀県・草津市
- ② 中垣淳、LSP に基づく変換による有色雑音の抑圧、IEICE 北海道支部学生インターネットシンポジウム、2015年2月17日～23日、北海道・札幌市
- ③ 中垣淳、LSP に基づく変換を用いた音声強調、IEICE 北海道支部学生インターネットシンポジウム、2013年2月6日～15日、北海道・札幌市

〔図書〕（計 0件）

〔産業財産権〕

○出願状況（計 0件）

○取得状況（計 0件）

〔その他〕

ホームページ等

6. 研究組織

(1) 研究代表者

中垣 淳 (NAKAGAKI, Atsushi)

北見工業大学・工学部・講師

研究者番号：90250539