

令和 2 年 6 月 17 日現在

機関番号：15101

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00272

研究課題名(和文)リアルタイム処理可能な突発性雑音除去システムの開発

研究課題名(英文)Development of impulsive noise suppression with real-time processing

研究代表者

笹岡 直人(SASAKA, Naoto)

鳥取大学・工学部・准教授

研究者番号：80432607

交付決定額(研究期間全体)：(直接経費) 2,800,000円

研究成果の概要(和文)：ヒアラブルデバイスや補聴器において、マイクに入力される周囲環境雑音の除去は必須である。また、技術要件として低遅延でリアルタイム処理が要求される。ただし従来法では突発性雑音の打撃部には対応していたが、固有振動を伴う残響部の抑圧は困難であった。そこで、本研究では高次統計量を用いる適応アルゴリズムおよび打撃部検出により、打撃部および残響部のリアルタイム雑音除去システムを実現した。

研究成果の学術的意義や社会的意義

残響を伴う突発性雑音のモデル化並びに統計的性質を明らかにすることにより、雑音除去システム発展の一助となる。また、ヒアラブルデバイスや補聴器に必須となる低遅延リアルタイム処理可能な雑音除去システムを開発することで、ヒアラブルデバイスの応用範囲を広げ、ひいては難聴の方が生活しやすい環境をつくる一助となる。

研究成果の概要(英文)：For a hearable device and hearing-aid, noise suppression is a key technology. Besides, these devices require real-time processing with low delay. The impulsive noise converts an initial peak sound to a ringing sound due to the natural mode radiation. The initial peak sound has a wideband component, and the ringing sound is composed of a narrowband component. Unfortunately, a conventional method suppresses only the wideband component. Therefore, we propose to suppress not only the wideband but also the narrowband component. A linear predictor using the least mean fourth algorithm can suppress the narrowband component in noisy speech because the fourth-order moment of speech is about zero in the short-time analysis. Besides, the proposed noise detector, which uses fourth-order cumulant, can detect a wideband noise frame, and then the wideband noise can be suppressed. From computer simulations, we can see the proposed method improves the noise reduction ability.

研究分野：デジタル信号処理

キーワード：信号処理 騒音抑圧 音声強調 高次統計量 ヒアラブルデバイス

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

IoT (Internet of Things)のためのセンサデバイスの1つとして、マイクロホンも注目を浴びている。一例として、マイクロホンから得られる音声から、会話状況を判定し、会社組織の活性化につなげる試みがなされている。また、ヒアラブルデバイスの開発が進められている。これは、補聴器のような機器に、様々なセンサを装備し、スマートホンなどと連動することで、移動した距離、速度、心拍数などを計測するとともに周囲の音、自分が発した声なども検出するものである。IoTデバイスとして期待されており、市場は50億ドルに成長するとの試算もされている。

ヒアラブルデバイスは、補聴器と同様に雑音除去手法の導入による会話環境向上も必要となる。雑音除去法はこれまで様々な手法が提案されてきた。スペクトルサブトラクションなどがよく知られている。しかし、周囲環境雑音が定常、もしくは特性変動が緩やかであることを仮定しているため、突発的に発生する衝突音、破裂音など発生時間が未知であり、電力変動の激しい雑音の除去はこれまで困難であった。なお、突発性雑音は、鳴り始めはその成分は広帯域に存在し(衝撃部)その後、物体の固有振動に対応する狭帯域成分(残響部)に変化していく。

このような突発性雑音に対応した手法として、事前に学習した騒音と音声の統計モデルを利用した手法などが提案されている。しかし、音声認識向けに開発されており、従来手法ではヒアラブルデバイスや補聴器が要求するリアルタイム処理が困難である。さらに、事前学習していない騒音に対しては抑圧が困難である。事前情報を用いない手法として、インパルス騒音の位相スペクトルの線形性を基にする手法、ゼロ位相信号を基にする手法が提案されているが、インパルス騒音の周波数スペクトルが白色となることを用いているため、抑圧可能な騒音の種類が限定される。また、ゼロ位相信号による手法は残響部にも対応しているが、残響部が音声の帯域と重なっていない場合のみ対応可能となっている。

### 2. 研究の目的

本研究は、音声に混在する突発性雑音の雑音除去システムの開発を目的とする。具体的な研究項目は、4次キュムラントを利用する衝撃部検出アルゴリズムの開発、衝撃部に対する雑音除去システムの開発、残響部を含めた突発性雑音に対する雑音除去システム開発、実装である。

### 3. 研究の方法

本研究では、まず突発性雑音のモデル化を行う。それに基づき、突発性雑音の残響部の除去手法について検討を行い、突発性雑音抑圧回路に必要な残響部の推定回路をFPGA(Field Programmable Gate Array)により実装した。その後、突発性雑音の衝撃部と残響部の抑圧手法を検討した。

#### (1) 残響部のための線形予測

##### 突発性雑音のモデル化

突発性雑音は初期の衝撃部から残響部に移行する。このとき、打撃部の周波数特性は広帯域に分布し、残響部はいくつかの線スペクトルを持つ狭帯域成分となる。打撃部の正規化4次キュムラントは50~150、またはそれ以上の値を持つことが知られている。そこで、本研究では残響部は、複数の共振点を持つ全極フィルタを非ガウス性白色雑音により励起することにより発生すると仮定した。その仮定により、全極フィルタは線形予測器により推定できるため、残響部は線形予測器の出力信号として得られる。

雑音重畳音声 $x(n)$ を

$$x(n) = s(n) + d(n) \quad (1)$$

とする。ここで、 $s(n)$ は音声信号である。 $d(n)$ は雑音である。線形予測器出力 $\hat{d}(n)$ は次式で表わせる。

$$\hat{d}(n) = \mathbf{x}^T(n) \mathbf{h}'(n) \quad (2)$$

$$\mathbf{x}(n) = [x(n-1), \dots, x(n-M)]^T \quad (3)$$

$$\mathbf{h}'(n) = [h'_1(n), \dots, h'_M(n)]^T \quad (4)$$

ここで、 $h'_j(n) (j = 1, \dots, M)$ は $j$ 番目のタップ係数、 $M$ はタップ数である。また、予測誤差は、

$$e(n) = x(n) - \hat{d}(n) \quad (5)$$

により表せる。

#### 4次モーメントを基にする適応アルゴリズム

線形予測器では、雑音重畳音声から残響部のみを推定する必要がある。しかしながら、2次統計量を利用した適応アルゴリズムでは実現は困難である。これはタップ入力に含まれる音声成分が強い相関を持つためである。そこで本研究では、4次統計量である4次モーメントを基にするLMF(Least Mean Fourth)アルゴリズムを利用する。音声は長時間観測では非ガウス性であることが知られているが、短時間観測では4次モーメントは零付近の値を持つことが知られている。本研究ではこの特性を利用する。

LMF アルゴリズムは次式により与えられる。

$$\mathbf{h}'(n+1) = \mathbf{h}'(n) + \mu e^3(n-M)\mathbf{x}(n) \quad (6)$$

ここで、 $\mu$ はステップサイズである。LMF アルゴリズムにおける音声の影響を解析する。予測誤差を

$$e(n) = e_d(n) + e_s(n) \quad (7)$$

とおく、ここで、 $e_d(n)$ および $e_s(n)$ はそれぞれ誤差に含まれる雑音、音声成分である。LMF アルゴリズムの評価関数は次式で表現できる。

$$\begin{aligned} E[e^4(n)] &= E[(e_d(n) + e_s(n))^4] \\ &= E[e_d^4(n)] + 6E[e_d^2(n)]E[e_s^2(n)] \end{aligned} \quad (8)$$

ここで、すべての奇数次モーメントは零とする。また、騒音の4次モーメントは音声の4次モーメントと比較して十分大きいと仮定する。音声が発生していないとき、評価関数は $E[e_d^4(n)]$ のみとなり、騒音のみが線形予測器で推定できる。一方、音声が発生しており、騒音が発生していないときは、音声の4次モーメントは零付近となるため、評価関数は零となり、騒音推定に影響を与えない。また、音声と騒音が同時に発生しているときは、音声の電力より、騒音の4次モーメントが十分大きいとき音声成分は線形予測器に影響を与えない(引用文献)。これにより騒音のみの推定が可能となる。

#### 計算機シミュレーション評価

提案線形予測器の有効性を確認するため、適応アルゴリズムとして2次統計量を基にするLMS(Least Mean Square)アルゴリズムとしたときとLMFアルゴリズムとしたときの騒音抑圧性能を比較する。図1(a)にSignal to Noise Ratio (SNR)の改善量を、図1(b)にPerceptual Evaluation of Speech Quality (PESQ)の改善量を示す。それぞれの適応アルゴリズムで得られるSNRおよびPESQの最大値はLMFアルゴリズムを用いた方が改善しており、本手法の有効性が確認された。

#### 実装

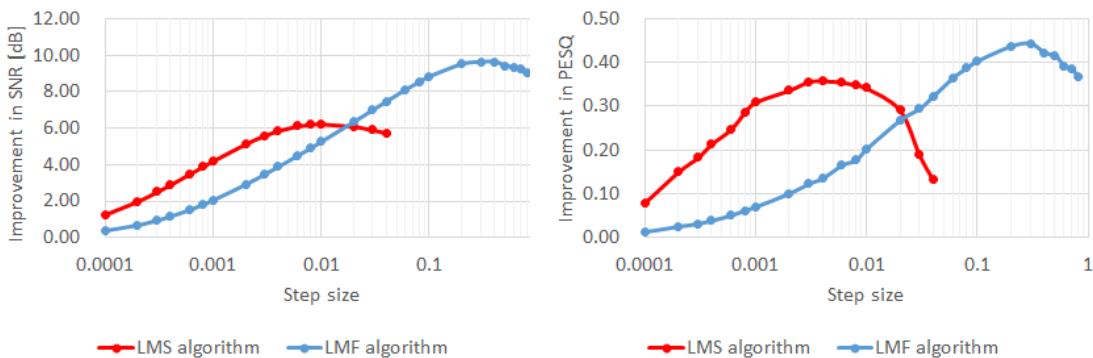
本手法の有効性を確認するため、FPGAによる残響部推定回路の実装を行った。サンプリング周波数は16kHz、入出力信号は24bit量子化されており、固定小数点演算により実装した。使用したFPGAボードはZynq-7020である。また、入出力波形を図2に示す。上段が入力信号、下段が残響部推定波形である。実装結果より本システムが有効に動作していることが確認できる。

#### (2) 突発性雑音抑圧システム

図3に衝撃部および残響部を抑圧する提案回路を示す(引用文献)。残響部抑圧には前節で説明した線形予測器を利用する。線形予測器による残響部抑圧後の信号 $x'(n)$ を、

$$x'(n) = s(n) + d_w(n) \quad (9)$$

とする。ここで、 $d_w(n)$ は雑音の衝撃部である。離散フーリエ変換された雑音重畳音声 $X'_m(k)$ は、



(a)SNR 評価 (b)PESQ 評価

図1 線形予測器による雑音除去性能評価

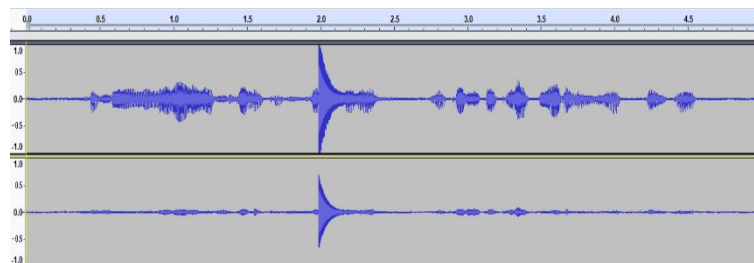


図2 線形予測器実装による処理波形(上段:入力信号、下段:推定残響部)

$$X'_m(k) = S_m(k) + D_m(k) \quad (10)$$

となる。ここで、音声を $S_m(k)$ 、突発性騒音を $D_m(k)$ とし、 $m$ はフレーム番号、 $k$ は周波数番号を表す。そして1フレーム内の高周波数帯域電力 $P_{h,m}$ を次式により求める。

$$P_{h,m} = \sum_{k=(N-1)/8}^{(N-1)/2} |X'_m(k)|^2 \quad (11)$$

ここで、 $N$ はフーリエ変換のフレームサイズを表す。次に、高帯域成分の電力スペクトルに対する瞬時4次キュムラントを求めることにより突発的な信号の変化を検出する。そこで高帯域電力 $P_{h,m}$ を確率変数としたとき、 $l$ 次モーメント $M_{l,m}$ 及び正規化4次キュムラント $C'_{4,m}$ を次式で定義する。

$$C'_{4,m} = C_{4,m}/M_{1,m} \quad (12)$$

$$M_{l,m} = \frac{1}{T} \sum_{n=0}^T P_{h,m-n+1}^l \quad (13)$$

$$C_{4,m} = M_{4,m} - 4M_{3,m}M_{1,m} - 3M_{2,m}^2 + 12M_{2,m}M_{1,m}^2 - 6M_{1,m}^4 \quad (14)$$

また、騒音フレーム検出精度を改善するため、高帯域電力 $P_{h,m}$ も騒音フレーム検出に併用する。次式により、過去 $N'$ フレームに対する高周波数帯域電力の最大値を算出する。

$$P_{max} = \max(P_{h,m-n}) \quad (0 \leq n < N') \quad (15)$$

ここで、 $N'$ は騒音の継続時間を考慮した任意のフレームサイズである。以上より、式(12)で算出した4次キュムラント $C'_{4,m}$ に対する閾値 $\alpha$ 、式(15)で算出した $P_{max}$ に対する閾値 $\beta$ を設定し、次式により騒音フレームの検出を行う(引用文献 )。

$$C'_{4,m} \geq \alpha \quad \text{且つ} \quad P_{max}$$

$\geq \beta$  のとき、騒音フレーム、

$C'_{4,m} < \alpha$  又は  $P_{max} < \beta$  のとき、非騒音フレーム

次に騒音フレーム検出による結果を基に音声並びに騒音の電力スペクトルを推定する。音声の電力スペクトル $|\tilde{S}_{m'}(k)|^2$ は、騒音発生フレームの直前の過去 $M$ フレームの騒音重畳音声 $X'_{m'}(k)$ から次式により得られる。

$$|\tilde{S}_{m'}(k)|^2 = \frac{1}{M'} \sum_{m=0}^{M'} |X'_{m''-n}(k)|^2 \quad (16)$$

ここで $m''$ は騒音発生フレームの直前のフレームを表す。一方、騒音の電力スペクトルは、推定された音声電力スペクトルを用いて、

$$|\hat{D}_{m'}(k)|^2 = |X_{m'}(k)|^2 - |\tilde{S}_{m'}(k)|^2 \quad (17)$$

により得られる。最後に騒音発生フレームに対し、ウィーナーフィルタによる音声強調処理を行う。強調音声 $\hat{S}_{m'}$ は、

$$\hat{S}_{m'}(k) = G_{m'}(k) |X'_{m'}(k)| e^{j\theta_{m'}(k)} \quad (18)$$

により得られる。ここで、騒音が発生しているフレーム番号を $m'$ 、 $G_{m'}(k)$ をゲイン関数、 $\theta_{m'}(k)$ を騒音発生フレームの位相とする。ゲイン関数は、事前SNR $\xi_{m'}(k)$ を用いて次式で定義される。

$$G_{m'}(k) = \sqrt{\frac{\xi_{m'}(k)}{1 + \xi_{m'}(k)}} \quad (19)$$

$$\xi_{m'}(k) = |\tilde{S}_{m'}(k)|^2 / |\hat{D}_{m'}(k)|^2 \quad (20)$$

また、突発性騒音の広帯域成分が発生していないフレームに対しては、強調音声の音質を維持するため、ウィーナーフィルタによる処理は行わない。

### (3) 計算機シミュレーション評価

本提案法の有効性を確認するため、計算機シミュレーション実験を行った。性能評価には、Segmental SNR (Seg. SNR)並びに PESQ を用いた。図4 (a)及び(b)に入力 SNR に対する Seg. SNR および PESQ の改善量を示す。なお、従来法として、ゼロ位相信号による突発性騒音抑圧手法(引用文献 )を用いる。また、適応アルゴリズムの更新停止に4次キュムラントによる騒音フレーム検出を併用する手法を提案法1、併用しない手法を提案法2とする。図4より、提案法は従来法と比較して Seg. SNR を維持しつつ、PESQ を改善していることが確認される。

#### 4. 研究成果

本研究において得られた成果を下記に挙げる。

- ・ 4次モーメントを基にする適応アルゴリズムを利用する線形予測器により残響部の推定が可能である。つまり、残響部除去が可能であることが評価関数の解析からも確認された。
- ・ LMF アルゴリズムを用いる線形予測器を FPGA により実装し、本提案法の有効性を確認した。
- ・ 打撃部および残響部の両方を除去するシステムを提案し、雑音除去性能の改善が確認された。

本研究で開発した手法はいずれもリアルタイム処理が可能であり、フレーム設定により低遅延で実現可能である。これによりヒアラブルデバイスや補聴器への実装が可能となる。

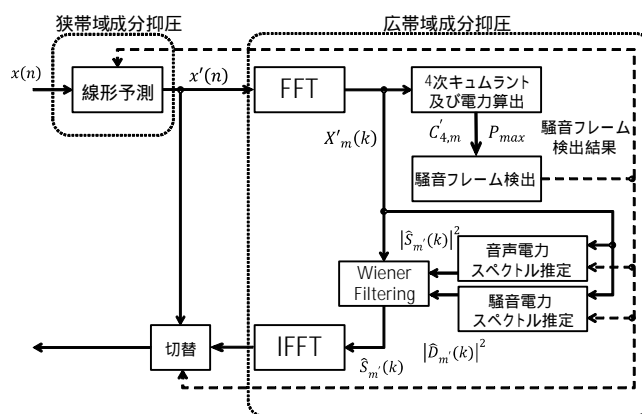
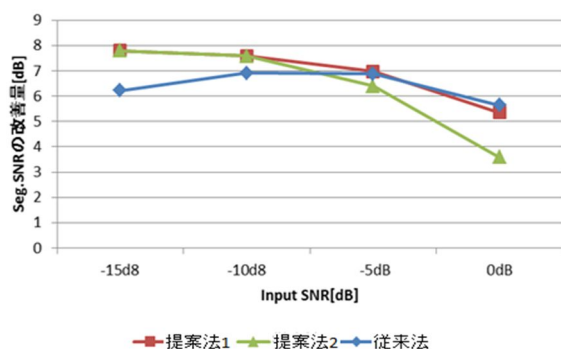
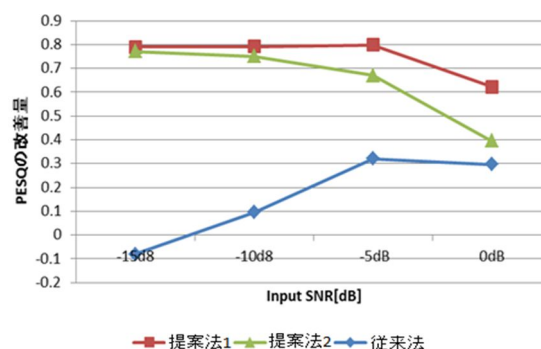


図3 提案手法構成回路



(a) Seg. SNR 評価



(b) PESQ 評価

図4 提案手法による雑音除去性能評価

#### < 引用文献 >

- N. Sasaoka, E. Akamatsu, A. Kawamura, N. Hayasaka, Y. Itoh, "4th Order Moment-Based Linear Prediction for Estimating Ringing Sound of Impulsive Noise in Speech Enhancement," IEICE Trans. Fundamentals., Vol.E103-A, No.10, Oct. 2020.
- Y. Yoneyama, N. Sasaoka, Yoshio Itoh, "Transient noise reduction based on high order statistics for hearable devices," Proc. 2017 International Workshop on Smart Info-Media Systems in Asia, pp.142-147, Sep. 2017.
- N. Sasaoka, N. Hamahashi, Yoshio Itoh, "Speech Enhancement with Impact Noise Activity Detection Based on the Kurtosis of an Instantaneous Power Spectrum," IEICE Trans. Fundamentals, Vol.E100-A, No.9, pp.1942-1950, Sep. 2017.
- A. Kawamura, N. Hayasaka, N. Sasaoka, "Impact and High-Pitch Noise Suppression Based on Spectral Entropy," IEICE Trans. Fundamentals, Vol.E99-A, No.4, pp.777-787, Apr. 2016.

5. 主な発表論文等

〔雑誌論文〕 計4件（うち査読付論文 2件 / うち国際共著 0件 / うちオープンアクセス 1件）

1. 著者名 Naoto Sasaoka, Naoya Hamahashi, and Yoshio Itoh	4. 巻 E100-A
2. 論文標題 Speech enhancement with impact noise activity detection based on the kurtosis of an instantaneous power spectrum	5. 発行年 2017年
3. 雑誌名 IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences	6. 最初と最後の頁 1942-1950
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transfun.E100.A.1942	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 米山雄統, 笹岡直人, 伊藤良生	4. 巻 117
2. 論文標題 高次統計量を基にする騒音フレーム検出及び線形予測による突発性騒音抑圧	5. 発行年 2017年
3. 雑誌名 電子情報通信学会技術研究報告	6. 最初と最後の頁 77-82
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Naoto Sasaoka, Eiji Akamatsu, Arata Kawamura, Noboru Hayasaka, Yoshio Itoh	4. 巻 E103-A
2. 論文標題 4th Order Moment-Based Linear Prediction for Estimating Ringing Sound of Impulsive Noise in Speech Enhancement	5. 発行年 2020年
3. 雑誌名 IEICE Trans. Fundamentals.	6. 最初と最後の頁 -
掲載論文のDOI（デジタルオブジェクト識別子） 10.1587/transfun.2020EAL2005	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計3件（うち招待講演 0件 / うち国際学会 2件）

1. 発表者名 赤松英治, 笹岡直人, 伊藤良生
2. 発表標題 突発性騒音における残響成分抑圧に関する一検討
3. 学会等名 2019年電子情報通信学会総合大会
4. 発表年 2019年

1. 発表者名 Yuuto Yoneyama, Naoto Sasaoka, and Yoshio Itoh
2. 発表標題 Transient noise reduction based on high order statistics for hearable devices
3. 学会等名 Proc. 2017 International Workshop on Smart Info-Media Systems in Asia (国際学会)
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	伊藤 良生  (ITOH Yoshio)  (70263481)	鳥取大学・工学部・教授    (15101)	