

擬似白色化信号を用いたハウリングキャンセラに関する検討 II

A study on howling canceller using quasi-whitened input signals II

○藤村達弘[†] 工藤憲昌 田所嘉昭[‡]

○Tatsuhiko Fujimura[†] Norimasa Kudoh[†] Yoshiaki Tadokoro[‡]

[†]八戸工業高等専門学校 [‡]豊橋技術科学大学

[†]NIT, Hachinohe College. [‡]Toyohashi Univ. of Tech.

キーワード：適応フィルタ，周波数推定，補聴器

連絡先：〒039-1192 青森県八戸市田面木上野平 16-1
八戸高専 産業システム工学専攻 電気情報工学コース
tel:0178-27-7281, e-mail:kudohk-e@hachinohe-ct.ac.jp

1. はじめに

補聴器を使用した際に、補聴器内のマイクに入力された音は増幅され、スピーカから出力される。Fig. 1 は補聴器の構造を簡便に表したものである。スピーカとマイクの距離が非常に近い為に、スピーカから出力された音の一部は外耳道を伝わり、再びマイクの入力となる。これが続くことにより、補聴器内に音響フィードバックループが形成され発振することで、ハウリングが発生する。ハウリングは、補聴器の使用者にとって非常に耳障りな音になる。

本稿では、このハウリングの除去を学習によって自動的に行うことを目的とした適応ハウリングキャンセラを用い、補聴器の品質改善について検討する。また、[1]で提案されている修正 LMS 法について、従来の FIR 型ノッチフィルタ (NF) と新たに提案する IIR 型 NF との比較を行う。

以下に本稿の構成を示す。2.では LMS 法を用いた従来法と、[1]で提案されている FIR 型 NF を用いた修正 LMS 法、およびそれらの問題点について簡単に述べる。3.では 2.で説明した問題点を低減する、IIR 型 NF を用いた提案法について述べる。4.は数値例およびシミュレーション条件である。5.はまとめである。

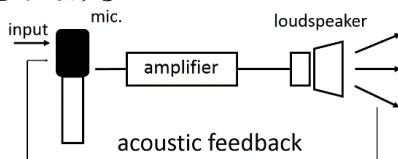


Fig.1 Hearing aids

2. ハウリングキャンセラ

2.1 LMS 法

音響フィードバック経路をモデル化し、ハウリングを除去するために、補聴器内に適応フィルタを設ける。単純に適応フィルタを設けた補聴器のブロック図を Fig. 2 に示す。

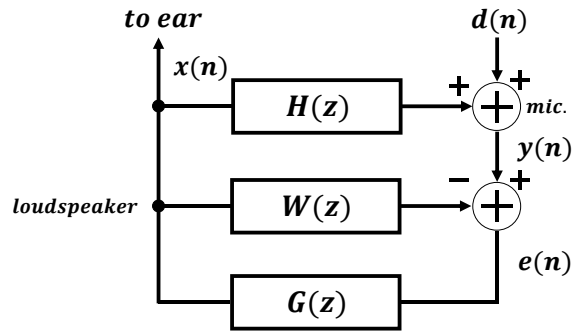


Fig.2 Hearing aids using a plain LMS

ここで $d(n)$ は入力信号、 $H(z)$ および $W(z)$ はそれぞれ音響フィードバック経路および適応フィルタの伝達関数である。 $H(z)$ は補聴器の使用者の外耳道(耳の穴から鼓膜まで)の構造によって異なるため、 $W(z)$ において LMS (Least mean square) 法によって $H(z)$ を同定する。また、 $H(z)$ および $W(z)$ は FIR フィルタとして設計する。 $G(z)$ は増幅部、 $x(n)$ は補聴器からの出力信号を表し、同時に $W(z)$ への入力信号にもなる。ここで出力信号の一部は $H(z)$ を通り、入力に戻る。 $y(n)$ は音響フィードバックを含んだマイクからの信号となる。定常状態において $W(z)$ が $H(z)$ を完全に同定した場合、 $e(n)$ は入力信

号 $d(n)$ と等価になり、ハウリングが完全に除去される。これ以降、Fig. 2 のような構成でハウリングを除去する方法を「プレーン LMS 法」と呼ぶ。Fig. 2 内の各信号の式を示す。

$$x(n) = G(z)e(n), \quad (1)$$

$$y(n) = d(n) + H(z)x(n), \quad (2)$$

$$e(n) = y(n) - W(z)x(n) \\ = d(n) + (H(z) - W(z))G(z)e(n). \quad (3)$$

$H(z)$, $W(z)$ はそれぞれ、

$$H(z) = \sum_{l=0}^{L-1} h_l z^{-l}, \quad (4)$$

$$W(z) = \sum_{l=0}^{L-1} w_l z^{-l}, \quad (5)$$

で表され、適応フィルタ $W(z)$ の l 番目の係数の更新式は、

$$w_l(n+1) = w_l(n) + \mu e(n)x(n-l), \quad (6)$$

となる。ここで、 μ ($0 < \mu < 1$) はステップサイズであり、この値で更新量が決まる。

通常の学習システムの面から考えると、 $x(n)$ は学習の入力信号であり、 $d(n)$ は補聴器への入力信号であると同時に、学習の観測雑音としても作用する。よく知られているように、学習において入力信号と観測雑音に相関がある場合、 $W(z)$ の推定値にバイアスが発生する[2]。Fig. 2 において、 $d(n)$ と $x(n)$ とには明らかに相関があることがわかる。

2.2 信号の擬似白色化

2.1 で述べたように、プレーン LMS 法では $d(n)$ と $x(n)$ との間に相関があるため場合、 $W(z)$ の推定値にバイアスが生じ、これにより $H(z)$ の同定が難しくなる。そのため $d(n)$ と $x(n)$ との相関を少なくすることで、推定値に生じるバイアスを低減させる方法が提案されており、これ以降、修正 LMS 法と呼ぶ[1]。Fig. 3 に修正 LMS 法のブロック図を示す。

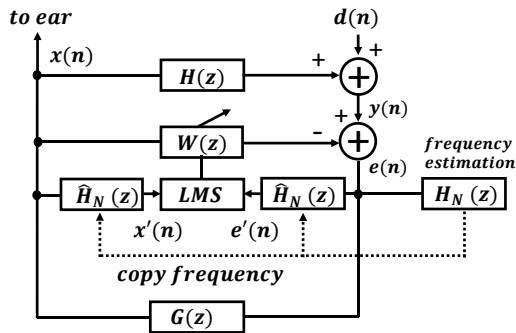


Fig. 3 The block diagram proposed in [1]

ここで、 $\hat{H}_N(z)$ は FIR 型 NF であり、正弦波と白色ガウス雑音との和のような AR 過程の信号としてモデル化された入力信号に対して、正弦波を除去するために使用される。

適応フィルタ $W(z)$ の l 番目の係数の更新式は、

$$w_l(n+1) = w_l(n) + \mu e'(n)x'(n-l), \quad (7)$$

で表され、 $x'(n)$ および $e'(n)$ はそれぞれ、

$$x'(n) = \hat{H}_N(z)x(n), \quad (8)$$

$$e'(n) = \hat{H}_N(z)e(n), \quad (9)$$

と定義される。 $H_M(z)$ は入力信号から正弦波を除去するため、 $H_M(z)$ の出力信号 $e(n)$ の電力が小さくなるように LMS 法を用いて更新される。 $H_M(z)$ の l 番目の係数 $w_{pl}(n)$ の更新式は、

$$w_{pl}(n+1) = w_{pl}(n) - \mu_p e(n-l)\varepsilon(n), \quad (10)$$

となる。したがって、 $H_M(z)$ の伝達関数は定常状態において、式(11)で示される 2 次の NF の連続接続型と等価になる。

$$H_N(z) = \sum_{l=0}^{2K} w_{pl} z^{-l} = \prod_{i=0}^K (1 - 2 \cos \omega_i z^{-1} + z^{-2}) \quad (11)$$

ここで、 K は入力信号に含まれる正弦波の数の最大値、 ω_i は角周波数を表す。 $H_M(z)$ において更新された係数は $\hat{H}_N(z)$ に代入され、 $x'(n)$ および $e'(n)$ を生成するために使用される。以上のように、修正 LMS 法は学習に擬似白色化した信号 $x'(n)$ および $e'(n)$ を使用する以外は、プレーン LMS 法と同様の動作を行う。

しかしながら FIR 型 NF は、正弦波が同じ周波数帯域に集中しているときに問題を引き起こすことが知られている。2 次の NF の周波数特性を見ると、ノッチ周波数において振幅は 0 となるが、ノッチ周波数から離れた領域では振幅は 1 より大きくなってしまふ。したがって、ノッチ周波数が同じ周波数帯域に集中している場合、その帯域での振幅はさらに小さくなり、離れた帯域での振幅はさらに大きくなってしまふ。その結果、学習において入力される信号と未知系 $H(z)$ に入力される信号が大きく異なってしまふ。さらに、本研究における最終目標は、人間の肉声を補聴器の入力信号とするものであるため、ノッチアウトする帯域幅を制御することができる、IIR 型 NF を使用する方法を提案する。

3. 提案法

本稿では、修正 LMS 法において、FIR 型 NF に代えて IIR 型 NF を使用する方法を提案する。 $e(n)$ に含まれる正弦波の周波数を推定するために、[4] で提案されている周波数推定法を基本的使用する。Fig. 4 にブロック図を示す。提案法における $H_M(z)$ は Fig. 4 の破線で囲まれた部分である。

Fig. 4 において、適応 NF $H_{Mk}(z)$ および、適応バンドパスフィルタ (BPF) $H_{Sk}(z)$ はトリー型に接続され、伝達関数はそれぞれ

$$H_{Mk}(z) = \frac{1 - \alpha_k z^{-1} + z^{-2}}{1 - \gamma \alpha_k z^{-1} + \gamma^2 z^{-2}}, \quad (12)$$

$$H_{Sk}(z) = \frac{-(1-\gamma)z^{-1} + \gamma(1-\gamma)z^{-3}}{1 - \gamma \alpha_k z^{-1} + \gamma^2 z^{-2}}, \quad (13)$$

と表される. ここで, $\gamma (0 < \gamma < 1)$ は極半径を制御する値である. μ_{af} は周波数推定におけるステップサイズ, α_k はタップ係数でありその真値は $2\cos\omega_k$ とした. さらに, 周波数推定の高速化のため, 積分ループ $(1 - \beta z^{-1})^{-1}$ を更新式に組み込むと, タップ係数の更新式は式(14)で表される[5]. ここで, $\beta (0 < \beta < 1)$ は高速化係数である.

$$\hat{\alpha}_k(n+1) = (1 + \beta)\hat{\alpha}_k(n) - \beta\hat{\alpha}_k(n-1) - \mu_{af}\varepsilon(n)s_k(n) \quad (14)$$

以降, 2.2 と同様に, 推定した周波数成分を学習における入力信号 $x(n)$ および出力信号 $e(n)$ から除去し, 擬似白色化を行う. そして, 擬似白色化された信号を用いて, 式(7)にしたがい $W(z)$ の係数を更新する.

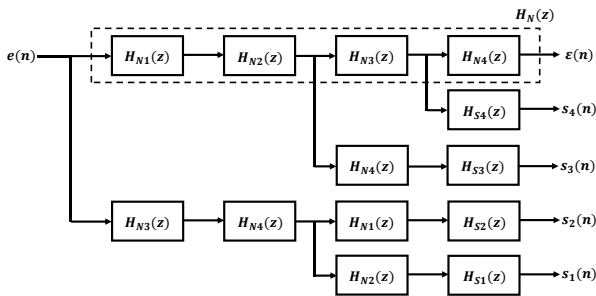


Fig. 4 Block diagram of frequency estimation (in case of four frequencies)

4. 数値例

ここでは, 通常の LMS 法, 修正 LMS 法, および提案法について, 各フィルタ係数の収束値に発生するバイアスと, 後述する推定誤差の正規化電力 ζ_h の観点から比較を行う. また, Fig. 5 に示すように人間の肉声の周波数特性は, ピークを中心に少し広がった共振周波数(フォルマント)を持ち, モデル化の次数をあげていくと, いくつかの帯域にフォルマントが集中するようになる[6]. しかし, 今回は前段階として正弦波と加法性白色ガウス雑音の和を入力信号に用いる.

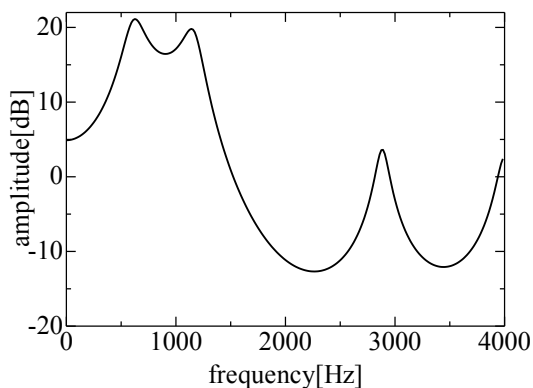


Fig. 5 Frequency characteristics of human voice

4.1 シミュレーション条件

入力信号 $d(n)$ は式(15)で表される.

$$d(n) = \sum_{i=1}^4 \left\{ A_i \cos\left(\frac{2\pi f_i}{f_s} n\right) + B_i \sin\left(\frac{2\pi f_i}{f_s} n\right) \right\} + \phi(n) \quad (15)$$

ここで A_i および B_i は信号の振幅, f_i は入力周波数, f_s はサンプリング周波数, $\phi(n)$ は分散 σ_ϕ^2 をもつ平均 0 の加法性白色ガウス雑音をそれぞれ表す.

シミュレーション条件は以下のように設定する:

- (1) $A_i = B_i = 1.0, f_s = 8000$ [Hz];
- (2) $f_i (i = 1 \sim 4) = 700, 1000, 1600, 2000$ [Hz];
- (3) $\mu = 0.05 (n = 0 \sim 800), \mu = 0.005 (n = 800 \sim)$;
- (4) $\mu_p = 0.01$;
- (5) $\mu_{af} = 0.01 (n = 0 \sim 200), \mu_{af} = 0.001 (n = 200 \sim)$,
 $\gamma = 0.98, \beta = 0.5$;
- (6) $H(z) = -0.01745 + 0.2151z^{-1} + 0.2849z^{-2} + 0.2151z^{-3} - 0.01745z^{-4}$;
- (7) $G(z) = 2.0z^{-12}, \sigma_\phi^2 = 0.5$;
- (8) $K = 6$;

シミュレーション条件(2)で設定したように, 入力信号に含まれる正弦波の周波数は低周波数帯域に集中させた.

また, シミュレーションの評価基準として, 式(16)で表される推定誤差の正規化電力 ζ_h を用いる.

$$\zeta_h = \frac{\sum_{i=1}^{L-1} (h_i - w_i)^2}{\sum_{i=1}^{L-1} h_i^2} \quad (16)$$

4.2 シミュレーション結果

3 つの方法について収束特性を比較した結果を Fig. 6~8 に示す. また, シミュレーション結果は 10 回の独立な試行の集合平均を取っている.

Fig. 6 に示すように, プレーン LMS 法による推定では, フィルタ係数の推定が収束したあともフィルタ係数の推定値にバイアスが生じていることがわかる.

修正 LMS 法は, 正弦波の周波数分布が等間隔の場合, バイアスが小さくなり適切な推定を行うことができる. しかしながら, Fig. 7 に示すように, 今回のシミュレーションでは正弦波の周波数分布を集中させたため, フィルタ係数の推定値にバイアスが生じている.

上記 2 つの方法に対して提案法では, Fig. 8 に示すようにどのフィルタ係数も正しい推定を行うことができた.

Fig. 9 に推定誤差の正規化電力 ζ_h を比較した結果を示す. ζ_h の値が小さいほど, フィルタ係数の推定値のバイアスが小さく, 良い特性であることを意味している. シミュレーション結果より, 提案法は他の二つの方法に比べて良い特性を有していることが分かる. さらに提案法は, プレーン

LMS法に加法性白色雑音のみを入力した時に匹敵する特性を有している。つまり、入力信号に対する周波数推定および擬似白色化が正確に行われていることを意味する。

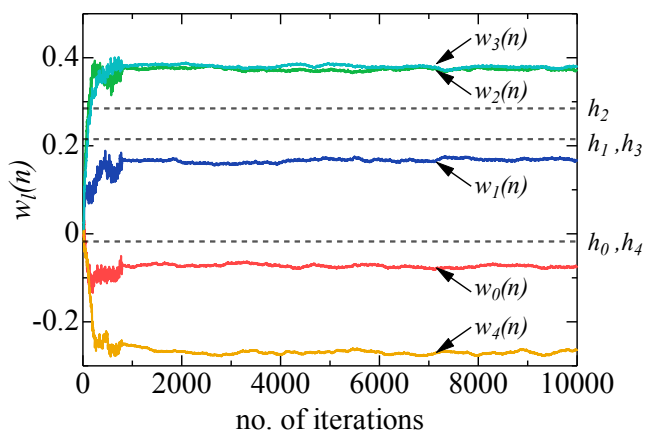


Fig. 6 Convergent curves of the plain LMS method

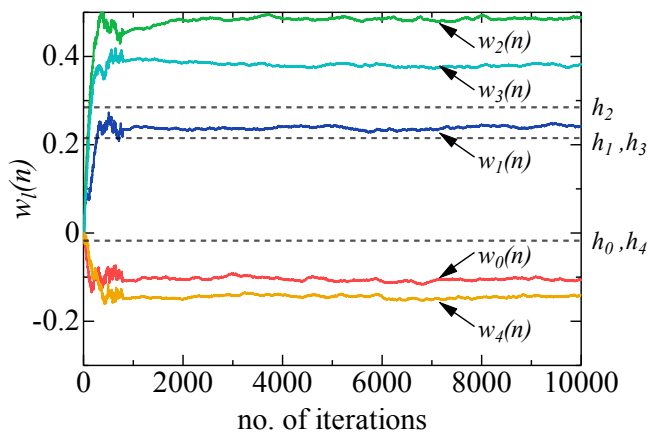


Fig. 7 Convergent curves of the modified LMS method

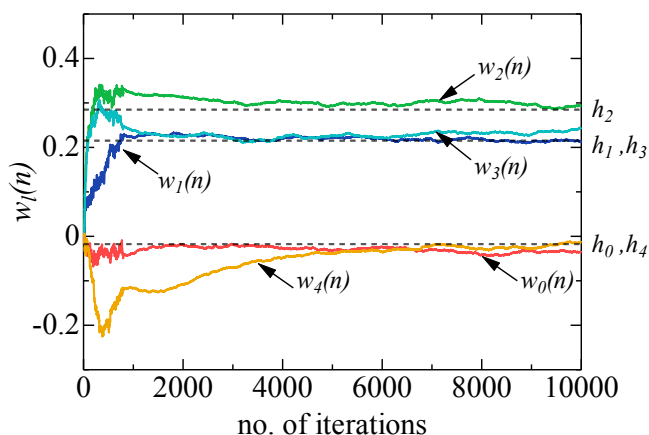


Fig. 8 Convergent curves of the proposed method

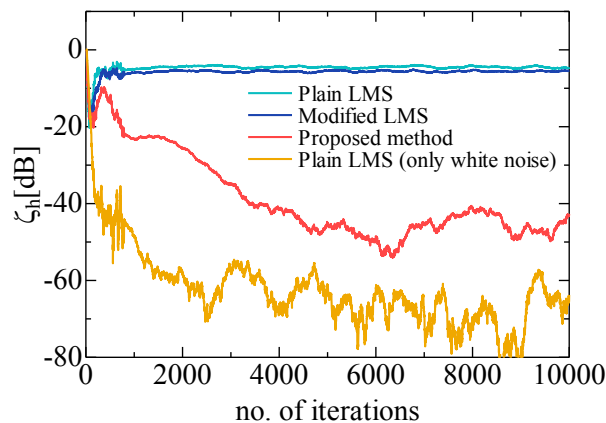


Fig. 9 Comparison of ζ_h in dB

5. まとめ

本稿では、はじめに、プレーン LMS 法を用いた従来法と[1]で提案されている修正 LMS 法による補聴器のハウリングキャンセラ、そして二つの方法の問題点について示した。問題は以下の二つによるものである。

- (1)学習における適応フィルタへの入力信号と観測雑音との間の相関
- (2)入力信号に含まれる正弦波の周波数分布の偏り

これらの問題を解決するため、IIR 型 NF を用いたハウリングキャンセラを提案した。

シミュレーションにより提案法における有効性を確認することができた。今後の課題として、人間の肉声への適用が挙げられる。

参考文献

- [1] H.Sakai, "Analysis of an adaptive algorithm for feedback cancellation in hearing aids for sinusoidal signals", IEICE Technical Report, SIP 2007-130, pp. 43-47, 2007
- [2] S.Haykin, "Adaptive filter theory", 2nd ed., Englewood, NJ: Prentice-Hall, 1991
- [3] T.Fujimura, N.Kudoh, Y.Tadokoro, "A study on howling canceller using quasi-whitened input signal", Tohoku-section joint Convention Record of Institute of Electrical and Information Engineers, 2A09, Aug., 2015
- [4] J.F. Chicharo, T.S.Ng, "Gradient-based adaptive IIR notch filtering for frequency estimation", IEEE Trans. on Acoust., Speech, Signal processing, 38(5), pp. 769-777, 1990
- [5] N.Kudoh and Y.Tadokoro, "Performance analysis of a new LMS-based Fourier Analyzer", CD-ROM Proceedings, IEEE TENCON '03, pp.1-4, Bangalore, India, Oct., 2003
- [6] 安居院猛, 中嶋正之, "コンピュータ音声処理", 産報出版