

前処理を用いたステレオエコーキャンセラの収束解析 Convergence Analysis of a Stereo Echo Canceller with Pre-Processing

武部 歌津衣 平野 晃宏 中山 謙二
Kazue Takebe Akihiro Hirano Kenji Nakayama

金沢大学大学院 自然科学研究科 電子情報システム専攻
Dept. of Electrical and Computer Eng., Faculty of Eng., Kanazawa Univ.
E-mail: takabe@leo.ec.t.kanazawa-u.ac.jp

あらまし

前処理を用いたステレオエコーキャンセラの収束解析を行なう。適応フィルタの係数が最適値に収束する速度は、入力信号の相関行列の固有値拡がりに関係がある。固有値拡がり大きいと収束が遅く、固有値拡がり小さいと収束が速い。本論文では、ランダム時変する前処理を用いたステレオエコーキャンセラを用いて、固有値拡がりと収束速度についての解析を行なう。また、前処理を用いない場合と用いる場合との固有値拡がりについて解析を行なう。

ABSTRACT

This paper analyzes convergence of a stereo echo canceller with pre-processing. Convergence speed relates to eigen value spread of correlation matrix of input signals. If eigenvalue spread is large, convergence speed is slow, and vice versa. This paper analyzes the relation between eigenvalue spread and convergence speed of a stereo echo canceller with randomly time-varying pre-processing. This paper also compares eigenvalue spread in case of using pre-processing with that not using pre-processing.

1 まえがき

現在、TV会議システムなどの音声通信においては、スピーカからマイクへの回り込み信号である音響エコーを除去するために音響エコーキャンセラが用いられている。TV会議システムは音声伝達に関してシングルチャネルシステムがほとんどであるが、より空間的なりアリズムの追求のためには音声を多チャンネルで伝送する必要がある。そのため、音響エコーキャンセラも多チャンネルに対応したものが必要となる。

SECでは、入力音響信号の両チャンネル間の強い相互相関の影響が大きな問題となる[1-3]。これにより適応フィルタ係数は最適値に収束せず、また、収束速度も遅くなってしまう。即ち、エコーパス特性を正しく同定することができない。そこでこの問題を解決するためにSECの2つの入力信号に前処理を行なう方法が試みられている[4,5]。前処理を用いたSECの収束解析も試みられているが[6,7]、前処理フィルタと収束速度の関係は完全に解決されていない。今回は、ランダム時変する前処理フィルタと収束速度の関係について解析する。

2 ステレオエコーキャンセラ

2.1 基本構成

図1に前処理を用いたSECの構成図を示す。SECは4つの適応フィルタから成り、各適応フィルタは対応するエコーパス特性を同定し、エコーを除去する。

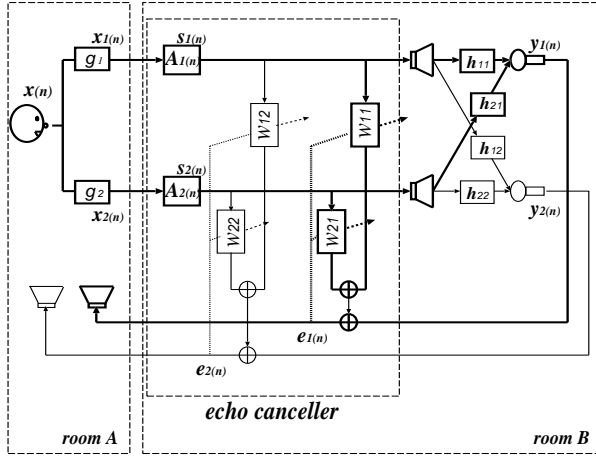


図 1: ステレオエコーキャンセラシステムの構成

2.2 係数不確定性の問題

図 1 において、第 j チャンネルの送信側の部屋の伝達特性は $G_j(z)$ 、入力信号は $X(z)$ 、部屋 B の第 j 番目のスピーカーから第 i 番目のマイクに至るエコーパスの伝達関数は $H_{ji}(z)$ 、適応フィルタの伝達関数は $W_{ji}(z)$ となる。ただし、周波数領域で表記する。前処理フィルタである $A_j(z, n)$ が無いとき、部屋 A の第 i 番目のスピーカーから出力される誤差信号 $E_i(z)$ は

$$E_i(z) = \sum_{j=1}^2 (H_{ji}(z) - W_{ji}(z)) G_j(z) X(z) \quad (1)$$

となり、 $E_i(z) = 0$ を満たすフィルタ係数は、 $W_{ji}(z) = H_{ji}(z)$ で一意に決まるとは限らない。そのため、エコーは消えても必ずしもフィルタ係数がエコーパス特性を同定できるとは限らない。これが SEC における係数不確定性の問題である。

3 前処理を用いた SEC

前述した問題を解決するため入力信号に前処理を行なう方法が試みられている [4,5]。ここでは、時間変化する APF $A(z, n)$ を用いて入力信号を前処理する。入力信号に前処理を行なう SEC の構成を図 2 に示す。この場合、誤差信号 $E_i(z, n)$ は次のようになる。

$$E_i(z, n) = \sum_{j=1}^2 (H_{ji}(z) - W_{ji}(z)) A_j(z, n) G_j(z) X(z) \quad (2)$$

$A_j(z, n)$ を時間変化させるとすると、多くの異なる $A_j(z, n)$ に対して $E_i(z, n) = 0$ を満たす必要がある。よって、適応フィルタのフィルタ係数は $W_{ji}(z) = H_{ji}(z)$ で一意に決定することができる。

4 ステレオエコーキャンセラの収束解析

部屋 A において、第 i 番目のスピーカーから出力される誤差信号 $e_i(n)$ は、前処理フィルタを通った後の信号 $s_j(n)$ 、部屋 B の第 j 番目のスピーカーから第 i 番目のマイクに至るエコーパスのインパルス応答を h_{ji} 、適応フィルタの係数を $w_{ji}(n)$ とすると、次の式で表される。

$$e_i(n) = \sum_{j=1}^2 (h_{ji} - w_{ji}(n)) s_j^T(n) \quad (3)$$

ただし、 T は、行列の転置を表す。 h_{ji} 、 $w_{ji}(n)$ 、 $s_{ji}(n)$ は、

$$h_{ji} = \begin{bmatrix} h_{ji,0} & h_{ji,1} & \cdots & h_{ji,N-1} \end{bmatrix} \quad (4)$$

$$w_{ji}(n) = \begin{bmatrix} w_{ji,0}(n) & w_{ji,1}(n) & \cdots & w_{ji,N-1}(n) \end{bmatrix} \quad (5)$$

$$s_j(n) = \begin{bmatrix} s_j(n) & s_j(n-1) & \cdots & s_j(n-N+1) \end{bmatrix} \quad (6)$$

とする。

$$s(n) = \begin{bmatrix} s_1(n) & s_2(n) \end{bmatrix} \quad (7)$$

$$h_i = \begin{bmatrix} h_{1i} & h_{2i} \end{bmatrix}^T \quad (8)$$

$$w_i(n) = \begin{bmatrix} w_{1i}(n) & w_{2i}(n) \end{bmatrix}^T \quad (9)$$

を用いて、式 (3) を表すと、

$$e_i(n) = s(n)(h_i - w_i(n)) \quad (10)$$

で表される。

LMS アルゴリズムでフィルタ係数の収束を考えると、フィルタ係数更新式は、

$$w_i(n+1) = w_i(n) + \mu s^T(n) e_i(n) \quad (11)$$

となる。また、

$$c_i(n) = w_i(n) - h_i \quad (12)$$

とし、式 (11) に式 (10) を代入し式 (12) を用いて両辺の期待値をとると、

$$E[c_i(n+1)] = (I - \mu R) E[c_i(n)] \quad (13)$$

となる。 $s(n)$ は、非定常な前処理を行なった信号であるので、 $s(n)$ は、定常ではないと考えられるが、問題を簡単にするため $s(n)$ を定常として考える。 R は、 $s(n)$ の相関行列であり、

$$\mathbf{R} = E[\mathbf{s}^T(n)\mathbf{s}(n)] \quad (14)$$

となる. 相関行列 \mathbf{R} を固有値 \mathbf{A} と固有ベクトル \mathbf{Q} を用いて表すと,

$$\mathbf{A} = \mathbf{Q}^T \mathbf{R} \mathbf{Q} \quad (15)$$

固有値 \mathbf{A} , 固有ベクトル \mathbf{Q} は,

$$\mathbf{A} = \text{diag} \left[\lambda_0 \quad \lambda_1 \quad \cdots \quad \lambda_{2N-1} \right] \quad (16)$$

$$\mathbf{Q} = \left[\mathbf{q}_0 \quad \mathbf{q}_1 \quad \cdots \quad \mathbf{q}_{2N-1} \right] \quad (17)$$

となり, 式(13)を式(15)を用いて表すと,

$$\mathbf{v}_i(n+1) = (\mathbf{I} - \mu \mathbf{A}) \mathbf{v}_i(n) \quad (18)$$

となる. ここで, \mathbf{v}_i は,

$$\mathbf{v}_i(n) = \mathbf{Q}^T E[\mathbf{c}_i(n)] \quad (19)$$

$\mathbf{I} - \mu \mathbf{A}$ は, 対角行列であるから,

$$v_{ik}(n) = (1 - \mu \lambda_k)^n v_{ik}(0) \quad (20)$$

となり, 式(13), 式(20)より,

$$\begin{aligned} E[\mathbf{w}_i(n)] &= \mathbf{h}_i + \mathbf{Q} \mathbf{v}(n) \\ &= \mathbf{h}_i + \sum_{k=1}^{2N-1} v_{ik}(n) \mathbf{q}_k \end{aligned} \quad (21)$$

となる. $E[\mathbf{w}_i(n)]$ の収束速度は $v_{ik}(n)$ の収束速度で決まる. $v_{ik}(n)$ の中で最も収束の遅いものは λ_{min} に対する成分である. 今, 固有値の総和, すなわち $\mathbf{s}_1(n), \mathbf{s}_2(n)$ の電力和が一定であるとする, 固有値拡がり大きいほど λ_{min} が小さくなり, 収束速度が遅くなる. さらに全体としては $v_{ik}(0)$ を考慮する必要がある. $\mathbf{v}_i(0)$ は,

$$\mathbf{v}_i(0) = \mathbf{Q}^T E[\mathbf{c}_i(0)] = -\mathbf{Q}^T \mathbf{h}_i \quad (22)$$

となる. ここで \mathbf{h}_i はエコーパスのインパルス応答であり, 問題によって異なるため予め解析することは難しい.

5 計算機シミュレーション

5.1 前処理フィルタ

用いる前処理フィルタは1次APFをランダム時変させる方式[5]と, 2次APFをランダム時変させる方式3種類[8]の合計4種類のフィルタを用いる.

1次APFは,

$$A(z, n) = \frac{z^{-1} - b(n)}{1 - b(n)z^{-1}} \quad (23)$$

で表され, の変数 $b(n)$ を以下のように更新する.

$$\begin{aligned} b(n+1) &= b(n) + \text{rand}(n) \\ b(n+1) &= b_{max} \text{ if } b(n+1) > b_{max} \\ b(n+1) &= b_{min} \text{ if } b(n+1) < b_{min} \end{aligned} \quad (24)$$

$\text{rand}(n)$ は一様乱数を取り, $-a \leq \text{rand}(n) \leq a$ の範囲で動くものとする.

2次のAPFは, 次式で表される.

$$\begin{aligned} A(z) &= \frac{b_2(n) - b_1(n)z^{-1} + z^{-2}}{1 - b_1(n)z^{-1} + b_2(n)z^{-2}} \\ &= \frac{(r(n))^2 - 2r(n) \cos \theta(n) z^{-1} + z^{-2}}{1 - 2r(n) \cos \theta(n) z^{-1} + (r(n))^2 z^{-2}} \end{aligned} \quad (25)$$

$b_1(n) = 2r(n) \cos \theta(n), b_2(n) = (r(n))^2$ の更新を次の3通りの方法で行なった.

- (a) θ は固定とし, r をランダム変化させる.
- (b) r は固定とし, θ をランダムに変化させる.
- (c) θ が小さいと r も小さくし, θ が大きいと r も大きくなるように, r と θ を両方ランダムに変化させる.

(a) のとき, $r(n)$ は,

$$\begin{aligned} r(n+1) &= r(n) + \text{rand}(n) \\ r(n+1) &= r_{max} \text{ if } r(n+1) > r_{max} \\ r(n+1) &= r_{min} \text{ if } r(n+1) < r_{min} \end{aligned} \quad (26)$$

(b) のとき $\theta(n)$ は,

$$\begin{aligned} \theta(n) &= 180 * Z_b(n) \\ Z_b(n+1) &= Z_b(n) + \text{rand}(n) \\ Z_b(n+1) &= Z_{bmax} \text{ if } Z_b(n+1) > Z_{bmax} \\ Z_b(n+1) &= Z_{bmin} \text{ if } Z_b(n+1) < Z_{bmin} \end{aligned} \quad (27)$$

(c) のとき $r(n), \theta(n)$ は,

$$\begin{aligned} r &= R * Z_c(n) \\ \theta &= 180 * Z_c(n) \\ Z_c(n+1) &= Z_c(n) + \text{rand}(n) \\ Z_c(n+1) &= Z_{cmax} \text{ if } Z_c(n+1) > Z_{cmax} \\ Z_c(n+1) &= Z_{cmin} \text{ if } Z_c(n+1) < Z_{cmin} \end{aligned} \quad (28)$$

となる. R は定数とする.

5.2 シミュレーション条件

2つのチャンネルが対称で、それぞれが独立に動作することから、1つのチャンネルについてのみシミュレーションを行なう。入力信号は白色雑音、適応フィルタは $N = 64$ タップの FIR フィルタ、学習アルゴリズムは NLMS、ステップサイズは $\mu = 0.5$ とした。部屋 A の伝達特性、部屋 B のエコーパス特性 $G_i(z), H_{ji}(z)$ は、それぞれ 20 次のバターワースフィルタを用いてローパスフィルタで構成した。 G_1, H_{11} はカットオフを $0.3f_s$, G_2, H_{21} はカットオフを $0.25f_s$ とした。フィルタ係数誤差 (NCEV) は、

$$NCEV(n) = 10 \log_{10} \frac{\sum_{j=1}^2 \|h_{j1} - w_{j1}(n)\|}{\sum_{j=1}^2 \|h_{j1}\|} \quad (29)$$

で評価する。次の条件でシミュレーションを行なった結果を図 2 に示す。

1 次 APF: $-0.9 \leq b(n) \leq 0$

2 次 APF の (a): $0.2 \leq r(n) \leq 0.9, \theta = 180$

2 次 APF の (b): $0.2 \leq Z_b(n) \leq 0.9, r = 0.2$

2 次 APF の (c): $0.2 \leq Z_c(n) \leq 0.9, R = 0.35$

$rand(n)$ は、すべての前処理フィルタで $-0.05 \leq rand(n) \leq 0.05$ の範囲をとる。このとき、入力を音声にした時の前処理フィルタを通した後の音声はどれもほぼ同等であり、激しい音質劣化は無い。

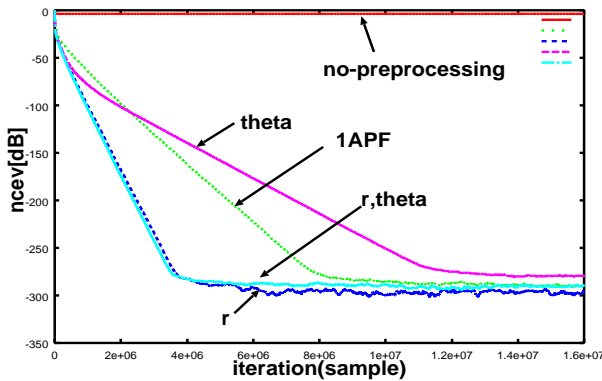


図 2: フィルタ係数誤差 (NCEV[dB])

この時の相関行列 R を、

$$\begin{aligned} R &= E[s^T(n)s(n)] \\ &\cong \frac{1}{M} \sum_{n=0}^{M-1} s_n s_n^T \end{aligned} \quad (30)$$

として、固有値拡がり ($\lambda_{max}/\lambda_{min}$) を求めると、

表 1: 固有値拡がり ($\lambda_{max}/\lambda_{min}$)

no-preprocessing	$3.0761e + 13$
1APF	4629.7
(a) r	1356.7
(b) θ	$1.0128e + 04$
(c) r, θ	3605.6

となる。前処理によって固有値拡がりを大幅に改善できることがわかる。前処理フィルタを用いた時のそれぞれの固有値拡がりの大きさは収束速度と関係があると思われる。このときの各固有値 ($\lambda_0, \lambda_1, \dots, \lambda_{2N-1}$) の値は図 3 のようになる。ただし、 $10 \log_{10} \lambda_{ik}$ で評価する。

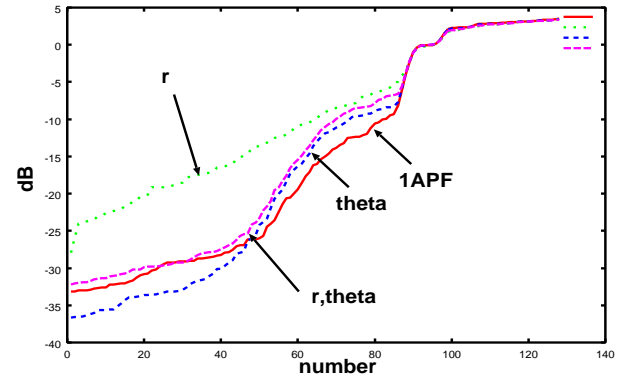


図 3: 各固有値の値 [dB]

図 4 は、2 次の APF (b) r で θ を 0, 45, 90, 110, 135, 180 と変化させたときのシミュレーション結果である。その他の条件は、前のシミュレーションと同じ条件をとる。

この時の固有値は、

表 2: 固有値拡がり ($\lambda_{max}/\lambda_{min}$)

$\theta = 0$	$1.6982e + 04$
$\theta = 45$	$1.7513e + 04$
$\theta = 90$	$1.3564e + 04$
$\theta = 110$	3050.8
$\theta = 135$	1236.2
$\theta = 180$	1356.5

となる。

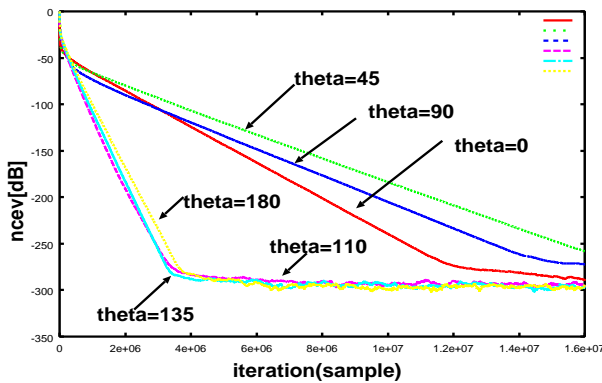


図 4: フィルタ係数誤差 (NCEV[dB])

図 4 で更新の初期の収束特性を図 5 に示す。

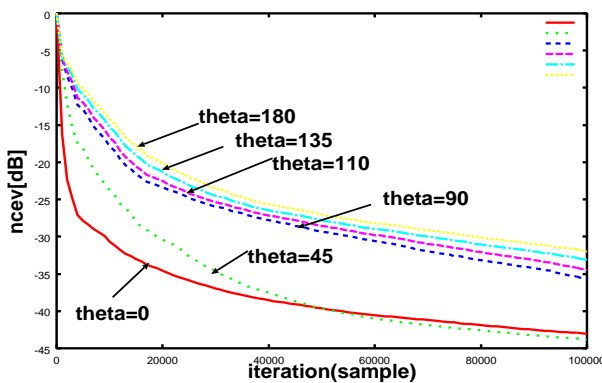


図 5: フィルタ係数誤差 (NCEV[dB])

フィルタ係数誤差 (NCEV) が小さい部分と大きい部分では収束速度が異なる。これは、収束は $v_{ik} = (1 - \mu\lambda_k)^n v_{ik}(0)$ で決まり、前半では $v_{ik}(0)$ の影響があり、後半では $(1 - \mu\lambda_k)^n$ の影響が強いためである。

6 むすび

前処理を用いたステレオエコーキャンセラの収束解析を行なった。ステレオエコーキャンセラのフィルタ係数が最適値をとるまでにかかる時間は入力信号 s の相関行列 R の固有値拡がりに関係がある。固有値拡がり大きい時、収束速度は遅く、固有値拡がり小さい時、収束速度は速くなる。今回は、ランダム時変する 1 次、2 次の APF を用いた前処理フィルタ 4 種類を用いて固有値拡がり収束速度の解析を行なっ

た。今回のシミュレーションでは、誤差が非常に小さい領域では収束速度と固有値拡がりの関係が見られた。しかし、必ずしも比例関係ではない。誤差が -40dB ぐらいまでの領域では固有値拡がり以外の要因が働いている。さらに、時変回路の動作を定常過程を前提とした手法で解析していることによる問題も考えられるので、このあたりも今後さらに検討したい。

参考文献

- [1] A. Hirano and A. Sugiyama, "Convergence Characteristics of a Multi-Channel Echo Canceller with Strongly Cross-correlated Input Signals - Analytical Result -," Proc. of 6th DSP Symposium, pp. 144-149, November 1991.
- [2] M. M. Sondhi and D. R. Morgan, "Stereophonic Acoustic Echo Cancellation - An Overview of the Fundamental Problem," IEEE SP Letters, vol.2, no.8, pp.148-151, August 1995
- [3] A. Hirano and S. Koike, "Convergence Analysis of a Stereophonic Acoustic Echo Canceller Part I: Convergence Characteristics of Tap Weights," Proc. of 11-th DSP Symposium, pp. 569-574, November 1996.
- [4] Y. Joncour and A. Sugiyama, "A Stereo Echo Canceller with Pre-Processing for Correct Echo-Path Identification," ICASSP, pp.3677-3680, 1998.
- [5] M. Ali, "Stereophonic Acoustic Echo Cancellation System Using Time-varying All-pass Filtering for Signal Decorrelation," ICASSP, pp.3689-3692, 1988.
- [6] 渡辺和伸, 中山謙二, 平野晃宏: "付加回路をもちいるステレオエコーキャンセラの収束性解析" 第 13 回デジタル信号処理シンポジウム, pp177-181, 1998.
- [7] 平野晃宏, 中山謙二, 渡辺和伸: "前処理を用いたステレオエコーキャンセラの収束解析" 電子情報通信学会論文誌 2001/5 Vol.J84-A No.5
- [8] 武部歌津衣, 平野晃宏, 中山謙二: "2 次ランダム可変付加回路を用いたステレオエコーキャンセラ" 第 15 回デジタル信号処理シンポジウム, pp.541-545, 2000.11