

音響エコーキャンセラにおける 相互相関を考慮した可変ステップ制御法

A Variable Step size Acoustic Echo Canceler: A Step Size Control Method Based on Cross-correlation between Echo and Near-end Signal

牛丸 晋一郎 中山 謙二 平野 晃宏
Shinichiro Ushimaru Kenji Nakayama Akihiro Hirano

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: ushimaru@leo.ec.t.kanazawa-u.ac.jp

あらまし

近端音声に遠端音声のエコーが加わった信号からエコーを消去する方法として適応フィルタを用いた音響エコーキャンセラ [1] がある。双方向通話時 (ダブルトーク) において、遠端と近端音声の相関性が強い場合に学習が不安定になる。本稿では、遠端と近端音声の相互相関に基づいてステップサイズを制御する方法において、相互相関の新しい計算法を提案し音響エコーキャンセラに応用している。相互相関の強い信号と弱い信号を用いた計算機実験によりその有効性を確認する。

ABSTRACT

In an acoustic echo canceler, learning coefficients during double talk causes unstable behavior. Especially, when cross-correlation between the echo and the near-end signal is high, the adaptation is hard to converge. The far-end and near-end signals are usually both speech signals, therefore, the cross-correlation is high. In this paper, a variable step size adaptive filter is proposed. The step size is controlled by the cross-correlation. The cross-correlation cannot be directly estimated, because the echo and the near-end speech are mixed. An indirect estimation method is proposed. Instead of the echo, the echo replica is used. Furthermore, the additional adaptive filter with a fixed step size is em-

ployed to stabilize the step size control. The proposed method is evaluated through computer simulation using the far-end and near-end signals, whose cross-correlations are high and low. Compared with the fixed step size method, the proposed method can always guarantee fast convergence and small residual error.

1 まえがき

TV 会議などの音声伝達システムにおいて回線に遠端音声のエコーが含まれる。このエコーは回線を単独で伝わる場合 (シングルトーク) と近端音声と混在して伝わる場合 (双方向通話:ダブルトーク) がある。シングルトーク時におけるエコー除去は適応フィルタを用いて比較的容易に行われる。しかしダブルトーク時においては、混在するエコーと他の音声の相互相関により学習の収束性が悪くなる。すなわち、他の音声と混在しているため、エコーのみの特性をとらえることが難しくなる [2]-[4]。相互相関が強い場合にはステップサイズを小さくすることにより収束特性を改善することができる。しかし、収束速度と残留誤差におけるトレードオフが問題となる [5]-[7]。

本研究では、遠端音声のエコーと近端音声の相互相関に基づいてステップサイズを制御する方法を提案する。この方法では、ダブルトークの有無にかかわらず、常に一定の法則に基づいてフィルタ係数を学習できる。相互相関の強い信号と弱い信号を用いた計算機シミュレーションにより、その有効性を示す。

2 音響エコーキャンセラの基本構成

音響エコーキャンセラの一例として、遠隔会議システムにおける音響エコーキャンセラを図1に示す。この方式は、室内音響(エコーパス)をシミュレートする適応フィルタ AF (Adaptive Filter)を挿入することにより、遠端音声を通して疑似エコーを生成し、減算によって実際のエコーを除去することを目的としている。

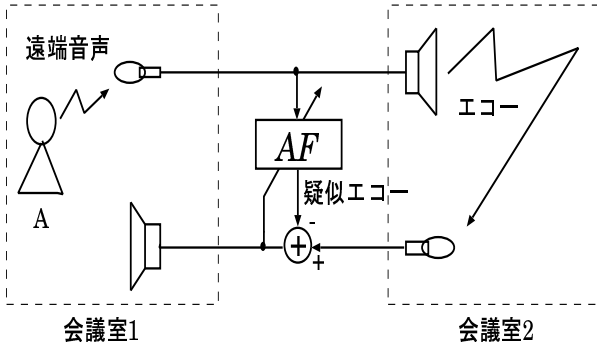


図1: 遠隔会議システムにおける音響エコーキャンセラ(シングルトーク)

3 ダブルトーク時の収束特性

3.1 ダブルトーク問題

図2に上記と同様な遠隔会議システムにおけるダブルトークの様子を示す。

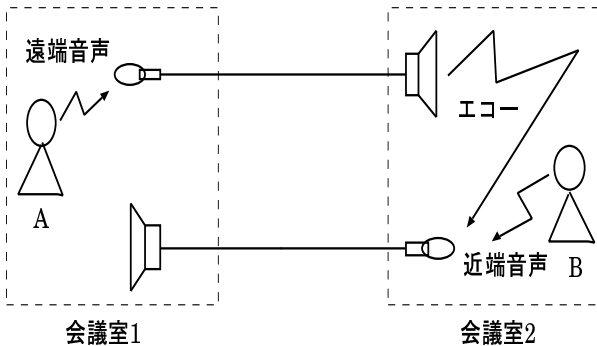


図2: 遠隔会議システムにおけるダブルトーク

ダブルトークとは図に示すように遠端音声と近端音声と同時に送信される状態をいう。これは会議システムなどにおいては当然考えられ、避けられない状況である。

ダブルトーク時の問題は回線中にエコーと近端音声が入り混じることである。このシステムにおいては近端音声のみ会議室1へ伝えるためにエコーだけを除去しなければならない。しかし、適応フィルタに対する参照信号がエコーと近端音声からなる混在信号なのでエコーパスの推定が困難となる。従って、適応動作にも相応の考慮が必要となってくる。そこで、本研究では遠端、近端音声間の相互相関に着目した。

3.2 相互相関と収束特性

図3にダブルトークを考慮した音響エコーキャンセラシステムを示す。

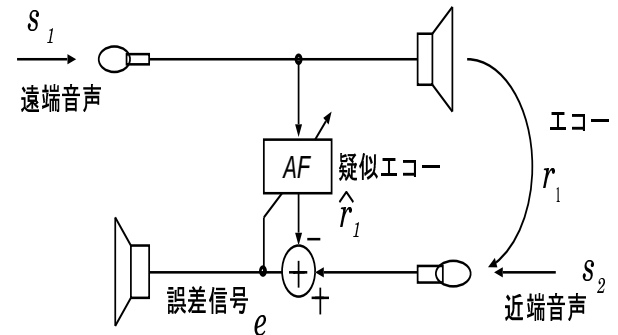


図3: 音響エコーキャンセラシステム

遠端音声 s_1 が回線を通り、近端側のスピーカから出力される時エコー r_1 が発生する。このエコーを適応フィルタ AF によって生成された疑似エコー \hat{r}_1 により差し引きし除去する。この時、ある値を持った近端音声 s_2 が入力されるとダブルトークとなる。このシステムにおいて、ある時刻 n における誤差信号 $e(n)$ は以下の式で与えられる。

$$e(n) = r_1(n) + s_2(n) - \hat{r}_1(n)$$

適応フィルタは $e(n)$ の二乗平均 $E[e^2(n)]$ を最小にするように係数を学習する。 $E[e^2(n)]$ は次式のように求められる。

$$\begin{aligned} E[e^2(n)] &= E[(r_1(n) - \hat{r}_1(n))^2 + 2(r_1(n) - \hat{r}_1(n))s_2(n) + s_2^2(n)] \\ &= E[(r_1(n) - \hat{r}_1(n))^2] + 2E[\underline{(r_1(n) - \hat{r}_1(n))s_2(n)}] \\ &\quad + E[s_2^2(n)] \end{aligned}$$

上式において右辺第3項は適応フィルタの特性と独立しているので一定とみなすことができる。右辺第2項に着目すると $r_1(n) - \hat{r}_1(n)$ と $s_2(n)$ との間に相関がなければ0となる。従って $E[e^2(n)]$ を最小にする

には右辺第1項の最小化を考えればよい。つまりこれは適応フィルタの伝達関数がエコーパスの伝達関数に近づくことを意味している。しかし、 $r_1(n) - \hat{r}_1(n)$ と $s_2(n)$ との間に相関がある場合、第2項は0にはならない。この状態で $E[e^2(n)]$ の最小化を行なうと、エコーパスの推定が理想値からずれることは明らかである。従って、ダブルトーク時には、遠端音声と近端音声の相互相関の度合に応じた学習法を考慮する必要がある。

4 可変ステップサイズ

4.1 可変ステップサイズの提案

本稿では、勾配法を対象とし、具体的には以下の式で示されるNLMSアルゴリズムを考える。

$$w(n+1) = w(n) + \mu e(n)x(n) / \{x^T(n)x(n)\}$$

係数更新の過程で変化の度合を制御するのがステップサイズ μ である。この μ の設定が収束速度と安定性に大きく影響する。相互相関が強い場合においても小さいステップサイズを用い、十分な収束時間があればエコーを除去することが可能である。しかし、収束時間も重要なので安易にステップサイズを小さくすることは実環境には適さない。

そこで効率の良い学習を行うために、遠端と近端の音声の相互相関の強弱によってステップサイズを変化させる可変ステップサイズを提案する。これは、相互相関が弱い時はステップサイズを大きく、逆に強い時はステップサイズを小さくすることを目的としている。これにより、小さいステップサイズを終始用いる場合に比べ収束時間の短縮がはかれ、また相互相関の強いエコーも除去できる。

4.2 基本構成

提案する可変ステップサイズを用いた音響エコーキャンセラを図4に示す。ここで、 $AF1$ のステップサイズは固定、 $AF2$ のステップサイズは可変である。

4.3 相互相関の導出

提案した可変ステップサイズを用いるにあたって、信号間の相互相関の値を求める必要がある。遠端音声 s_1 と近端音声 s_2 の相互相関はそれらの積の二乗平均で示される。しかし、回線を伝わるのは近端音声 s_2 とエコー r_1 の混在信号なので s_2 のみを検出することはできない。従って直接この相互相関を求めることは不可能である。そこで以下のようにして相互相関を求めた。

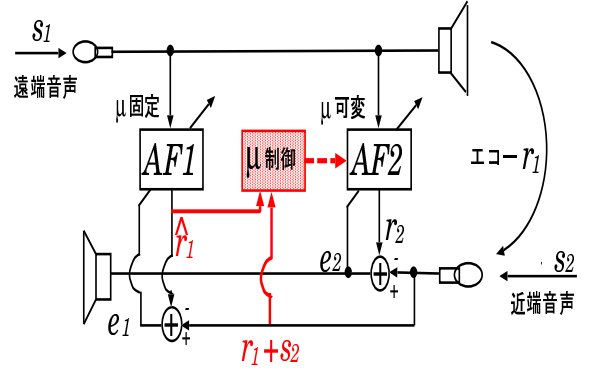


図4: 可変ステップサイズを用いた音響エコーキャンセラシステム

まず、 $AF1$ により生成される疑似エコー \hat{r}_1 と混在信号 $r_1 + s_2$ の相互相関を求める。 E は平均を表す。

$$\alpha = E[\hat{r}_1(r_1 + s_2)] \quad (1)$$

次に、疑似エコー \hat{r}_1 の自己相関を求める。

$$\beta = E[\hat{r}_1^2] \quad (2)$$

ここで、

$$E[\hat{r}_1^2] \simeq E[\hat{r}_1 r_1] \quad (3)$$

と近似できると仮定する。式(1)、(2)、(3)より α と β の差 γ をとると近似的に疑似エコー \hat{r}_1 と近端音声 s_2 の相互相関が得られる。

$$\begin{aligned} \gamma &= \alpha - \beta \\ &= E[\hat{r}_1 r_1] + E[\hat{r}_1 s_2] - E[\hat{r}_1^2] \\ &\simeq E[\hat{r}_1 s_2] \end{aligned} \quad (4)$$

適応動作が進み、疑似エコー \hat{r}_1 がエコー r_1 の特性に近づくにつれ、 γ の値はエコー r_1 と近端音声 s_2 の相互相関に近づく。

4.4 可変ステップサイズの更新式

前で述べた様に、この提案モデルの目的は、相互相関が弱い時はステップサイズを大きく、逆に強い時はステップサイズを小さくすることである。従って、基本的にはステップサイズ μ を γ の逆数になるように設定すればよい。そこでまず、 γ を以下の式によって大きさが1より小さくなるように正規化した。

$$\Gamma = \frac{\gamma}{E[(r_1 + s_2)^2]} \quad (5)$$

この Γ を用いて考えられる可変ステップサイズ更新式の基本形を式(6)に示す。

$$\mu = \frac{K_1}{K_2 |\Gamma|^\xi + \delta} \mu_0 \quad (6)$$

ここで K_1 、 K_2 、 ξ 、 δ 、 μ_0 は定数である。

4.5 適応フィルタの二重化

この提案モデルでは適応フィルタを2つ用いているが、この目的は適応動作の安定化である。可変ステップサイズシステムを適応フィルタ1つで実現した場合、正常な適応動作が行われない可能性がある。相互相関が強い場合、適応フィルタによって生成される疑似エコーはエコーのみならず所望信号までも除去対象とし、それにより疑似エコー自身のパワー β が増加する。この結果、式(4)より β に比例し γ も増加する。つまりこれは式(6)から判断できるように、ステップサイズが極めて小さい値に陥ってしまい、学習が停止することを意味する。これにより、所望の疑似エコーが得られなくなり適切な相互相関が導き出せなくなる。この不安定な相乗効果を避けるために固定ステップサイズで学習される適応フィルタ(AF1)をエコーキャンセル用(AF2)とは別に設けてステップサイズの制御に用いる。

5 計算機実験

提案する可変ステップサイズと固定ステップサイズの場合に対して計算機実験を行い比較した。

遠端、近端それぞれに2次のARモデルを用い、白色信号を通して得られる有色信号を入力信号とした。相関が弱い場合として、遠端側のARモデルの極の角度を 0.2π 、大きさを0.95とし、同様に近端側では 0.8π 、0.98とした。また、相関が強い場合として、遠端側のARモデルを 0.2π 、0.95、近端側のARモデルを 0.2π 、0.97とした。その様子を図5に示す。エコーパスにはバターワースフィルタ、適応フィルタにFIRフィルタを用い、そのタップ数を64とした。係数学習にはNLMSを用い、可変ステップサイズの更新式を式(7)のように設定した。この条件の基で、エコーパスのインパルス応答 h と適応フィルタのフィルタ係数 w を比較し、式(8)で示されるような係数誤差をノルムを用いてdB評価した。

$$\mu = \frac{1}{10^3|\Gamma|^2 + 1} \times 0.05 \quad (7)$$

$$D = 10 \log_{10} \frac{\|h - w\|^2}{\|h\|^2} \quad (8)$$

まず、図6に示すように相関が強い場合において、式(7)で示される可変ステップサイズに対し、式(8)で評価される収束特性が同じになるように固定ステップサイズを決める。これらのステップサイズに対し、相関が弱い場合において同様な評価を行った。その結果を図7に示す。結果からも分かるように、可変ス

テップサイズ方式により収束が速くなっている。つまり、収束特性の精度が同じであっても、可変ステップサイズ方式を用いることにより収束時間が速くなることが確認できる。

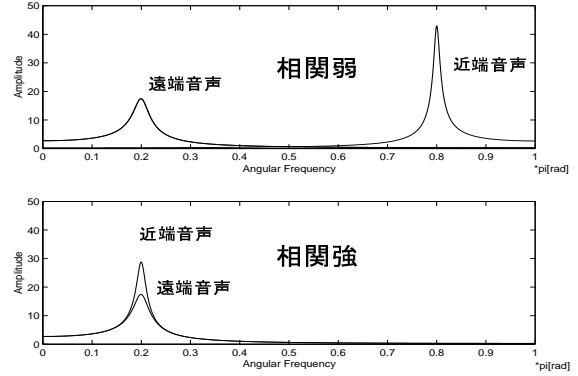


図5: 遠端音声と近端音声の周波数特性

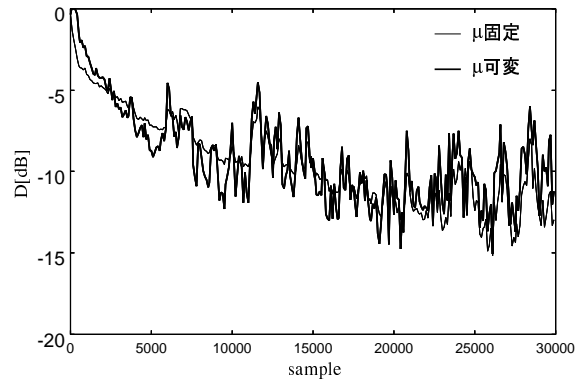


図6: 係数誤差の学習曲線(相関強)

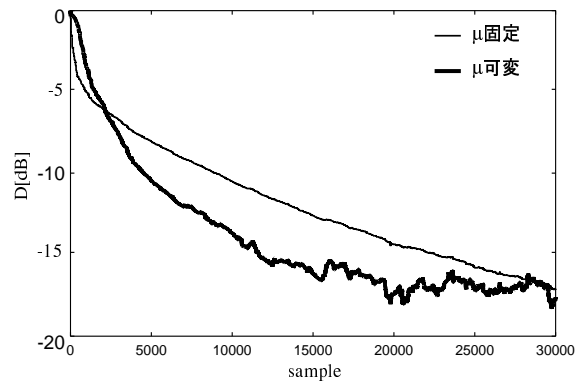


図7: 係数誤差の学習曲線(相関弱)

次に同様な実験を相関の強弱を逆にして行った。相関が弱い場合において図8に示すように、収束特性が同じになるように固定ステップサイズを決め、それらの値を用いて相関が強い場合に対する収束特性を調べた。その結果を図9に示す。多少ではあるが収束精度が改善されている。

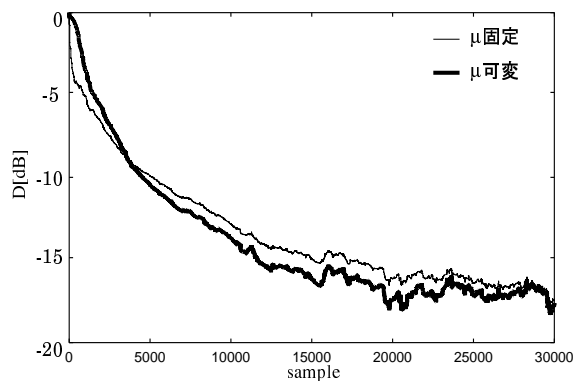


図 8: 係数誤差の学習曲線 (相関弱)

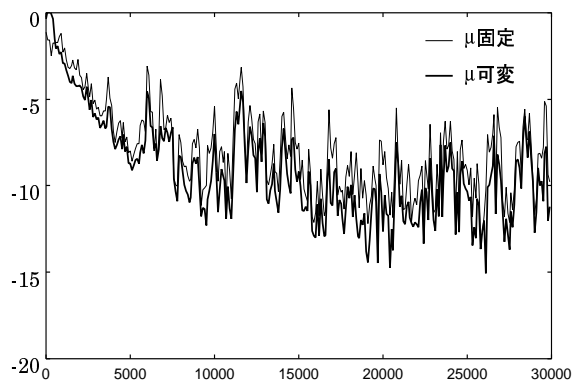


図 9: 係数誤差の学習曲線 (相関強)

6 むすび

音響エコーキャンセラにおいて、双方向通話時の収束特性を改善することとして、遠端音声と近端音声の相互相関に応じてステップサイズを可変する方式を提案した。相互相関の計算では、エコーの代わりにエコーレプリカを用いた。また、ステップサイズ制御の安定化のために適応フィルタを二重化した。可変ステップサイズ方式により、固定のステップサイズを用いる場合と比べて相互相関の強弱に関係なく常に良好な収束特性が得られることが確かめられた。

参考文献

- [1] Simon Haykin, "Adaptive Filter Theory" PRENTICE HALL,1996
- [2] 佐藤 明, 田村潤三, 来山征士: "相関を利用した重畳通話検出方式", 信学技報, CS75-136(1975).
- [3] 来山征士, 田村潤三, 山本誠一, 石上彦一: "共通の適応制御部をもつ多重エコーキャンセラ", 信学技報, CS78-23(1978).
- [4] 古川博基: "適応フィルタの実用技術(3)", 音響誌, 48, 7, pp.513-516(1992).
- [5] Yamamoto S. and Kitayama S.: "An Adaptive Echo Canceler with Variable Step Gain Method", Trans.IECE Japan,E65,1,pp.1-8(1982).
- [6] Akihiko Sugiyama: "An Interference-Robust Stochastic Gradient Algorithm with a Gradient-Adaptive Step-Size", Proc. of ICASSP '93, vol.3, pp.539-542(1993).
- [7] Hector Perez Meana, et. al.: "A Time Varying Step Size Normalized LMS Echo Canceler Algorithm", Proc. of ICASSP '94, vol.2, pp.249-252(1994).