

A-10 エコーキャンセラの双方向通話を考慮した学習法

A Learning Method of Echo Canceller in Double Talk

電子情報システム専攻 牛丸 晋一郎

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。エコーと近端音声が存在する双方向通話時には、適応フィルタの係数学習が不安定になる。本稿では、遠端音声のエコーと近端音声の相互相関を近似的に導出しステップサイズを可変する方法を提案している。さらに、適応フィルタを2つ用いたダブルフィルタ方式を提案し、双方向通話による特性劣化を防ぐと同時にエコーパス変動への対応を可能にした。音声信号を用いて計算機シミュレーションによりその有効性を確認する。

1 まえがき

TV 会議などの音声伝達システムにおいて、自分の発した音声で室内で反響し、数十ミリ～数百ミリ秒遅れて再び聞こえる現象はしばしば起こる。このように反響される音声を一般的に「エコー」と呼んでおり、遅れが約 30 ミリ秒以上になるとエコーとして聞こえる。例えば会議室の反響や衛星回線では 0.1～0.5 秒程度の遅れが生じる。このエコーを除去する為に適応フィルタを用いた音響エコーキャンセラシステムが最も一般的となっている。通信の形態としては、片方が話す場合（シングルトーク）と双方が同時に話す場合（ダブルトーク）がある。シングルトーク時にはエコーのみが検出できるのでその除去は比較的容易である。しかし、ダブルトーク時にはエコーがもう一方の音声と混在する為に、エコーの特性を捉えることが困難となり、適応フィルタの収束特性が劣化する。この問題はステップサイズの値を小さくすることによって解決できるが、これにより収束速度が遅くなるという問題も生じる。このように残留誤差と収束速度は相互に密接な関係を持っており、このトレードオフをいかに効率良く解消するかが最も重要な課題である。

さらなる問題として、エコーパス特性の急激な変動がある。実環境においてはエコーパスは常に一定とは限らず、エコーパスが変化すれば適応フィルタは係数を素早く学習し直さなければならない。よってエコーパス変動の検出が必要となるが、この場合において

も、エコーパス変動とダブルトークはいずれも残留誤差の増加を招くという意味で密接に関係しており、これらの見分けが重要となる。

本研究ではまず、エコーと近端音声の間の相互相関によってステップサイズを可変する方法を提案する。この方法により、固定ステップサイズを用いる場合と比べ、エコーと近端音声の関係によらず常に良好な収束特性と残留誤差を得ることができる。さらに、適応フィルタを2つ用いたダブルフィルタ (DF) 方式を提案する。これは可変ステップサイズで学習する適応フィルタとその結果を受けて動作するフィルタから構成され、単一フィルタでは回避困難であった収束特性の劣化を改善している。最後に、エコーパス変動に対する検討を行い、2つの出力信号のパワーを比較することにより、エコーパス変動への対応を可能にした。計算機シミュレーションにより提案法の有効性を確認する。

2 ダブルトーク問題

図 1 に遠隔会議システムにおけるダブルトークの様子を示す。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の混在することである。このシステムにおいては近端音声のみを会議室 1 へ伝えるためにエコーのみを除去しなければならない。適応フィルタ AF は残留誤差の電力を下げるように学習されるが、ダブルトーク時には、この残留誤差にエコーと近端音声が含まれるためエコーパスの推定が困難となる。これがダブルトークの問題点である。

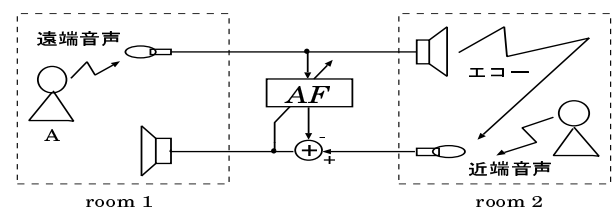


図 1: 遠隔会議システムにおけるダブルトーク

3 可変ステップサイズ

3.1 可変ステップサイズ方式

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。

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

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。ダブルトークの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、様々な条件における収束速度を考慮すると実用的ではない。

そこで、安定で収束が速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が提案されている [3]~[6]。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、小さい固定ステップサイズを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。

3.2 相互相関の導出

前項で述べ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 2 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 2 参照)

まず、AF 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。本研究では、この γ を次式のように遠端音声で正規化した Γ の値を用いた。

$$\Gamma = \frac{\gamma}{E[s_1^2]} \quad (5)$$

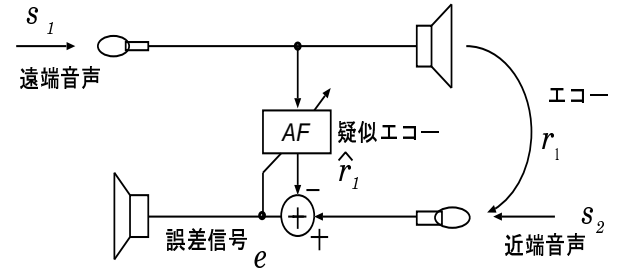


図 2: 音響エコーキャンセラシステム

3.3 短時間平均と長時間平均の利用

文献 [3],[4] では、式 (6) のように Γ に反比例するステップサイズの更新式を提案し、ステップサイズが固定の場合よりも収束特性が改善されることが確かめられた。

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

($K_1, K_2, \xi, \delta, \mu_0$ は定数)

しかし、相関の急激な変化に対して十分に追従できない為、相互相関の変化の大きい箇所では効率良く学習できないという問題がある。

この問題を解決する為に、式 (5) で求めた相互相関 Γ の短時間平均と長時間平均を組み合わせる方法を提案する。この方法では、長時間平均の値を基本として用いるが、ある時点で短時間の平均値が長時間の平均値を上回った場合は、短時間平均の値をその時点での相関値とする。図 3 にこの方法の例を示す。

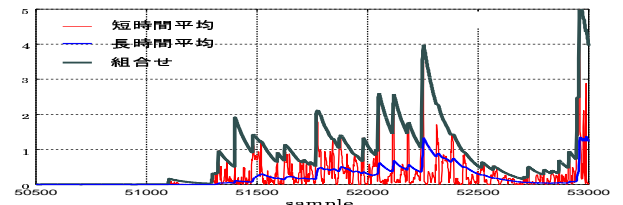


図 3: 短時間、長時間平均を組み合わせた相互相関

短時間、長時間平均を組み合わせた相互相関に基づいて、可変ステップサイズを導出する。まず閾値を設け、この閾値に対して相互相関の値が大きい場合はステップサイズを小さく、逆に相互相関の値が小さい場合はステップサイズを大きくする。相互相関の急激な増加によって閾値を越えた瞬間、ステップサイズを急峻に下げ、学習を抑制する。その後、長時間平均によ

り相互相関は緩やかに変化するのでダブルトーク期間では小さなステップサイズを維持し、安定した学習を行う。閾値よりも小さくなった時点でステップサイズを大きくし学習を進める。相互相関の追従性を高める為にステップサイズは2値のみとした。

4 ダブルフィルタ方式

4.1 可変ステップサイズ方式の問題点

固定ステップサイズを用いる場合と比べ、ステップサイズを可変にすることによって収束特性の改善が確かめられた [5][6]。しかし、実際の相互相関に対し、近似で求めた相互相関では急激な変化に対する追従が完全ではないため、ステップサイズを下げる動作が遅れ、収束特性の劣化を招く。以下ではこの問題を適応フィルタを2つ用いることにより改善する。

4.2 ダブルフィルタ方式の提案

本章では可変ステップサイズを応用したダブルフィルタ (DF) 方式を提案する。図 4 に提案モデルを示す。

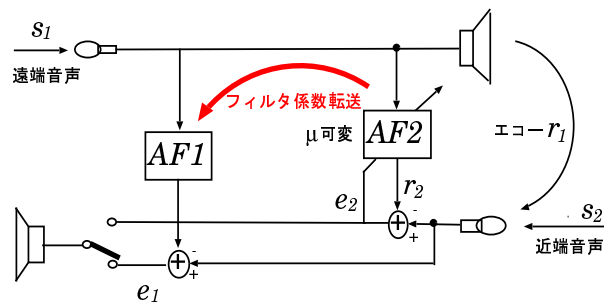


図 4: ダブルフィルタ (DF) モデル

基本的な考え方は、良好な条件で AF2 が学習されたとき、そのフィルタ係数を AF1 に転送して使用することである。良好な条件としてシングルトークに近い状態を想定する。すなわち、エコーと近端音声の相関が低く、遠端音声のレベルが比較的高いことを条件とする。係数更新の学習は AF2 のみで行い、可変ステップサイズを用いる。適応フィルタ AF1 は学習は行わず、以下のフィルタ係数転送条件を満たした場合に AF2 のフィルタ係数の値を AF1 に転送する。

【フィルタ係数転送条件】

- 相互相関が θ_1 以下
- 遠端音声のパワーが θ_2 以上
- これらを同時に N_1 サンプル以上満たす

ここで θ_1 , θ_2 , N_1 は経験的に設定する数値である。可変ステップサイズ方式では、相互相関の急激な変化に対する追従性に限界があるため特性の劣化は避けられなかったが、この方式では、追従が遅れた場合でも AF2 のフィルタ係数が最適値からずれる直前の値を AF1 に保持することが可能となるため、良好な収束特性が維持できる。

ここで初期状態について述べておく。システムが起動してフィルタ係数が転送されるまでは適応フィルタ AF1 は係数が 0 であるため、出力信号として e_2 を用いる。転送が行われた後は出力信号として e_1 を用いる。

4.3 エコーパス変動の検出

DF 方式を用いたエコーパス変動の検出とそれに対する対処法について説明する。エコーパス変動が起こった瞬間、どちらの適応フィルタも収束特性が劣化する。その後、AF2 はエコーパス推定の学習を再開するが、AF1 は転送条件が満たされるまでは劣化した収束特性を保ち続ける。この結果、AF1 側の残留エコーが大きくなり e_1 のパワー P_1 が e_2 のパワー P_2 よりも大きくなると考えられる。このことから【 $P_1 > P_2$ 】という条件が満たされた時、エコーパス変動が起こったと考えることができる。さらに、エコーパス変動の検出後、ステップサイズを大幅に大きくすることにより素早く変動後のエコーパス推定が可能となる。これらをふまえて以下の2つの条件を設けた。

【出力切替条件】

- $P_1 > P_2$ が N_2 サンプル以上続いた場合には出力を e_2 に切替える

【ステップサイズ切替条件】

- エコーパス変動後、相互相関が θ_3 以下の間は可変ステップサイズを μ_{max} に上げる

ここで N_2 , θ_3 , μ_{max} は経験的に設定する数値である。出力切替えにより AF1 の未学習による収束特性の劣化を回避できる。今回は N_2 を 2000 サンプルとしたが、この値が小さいとエコーパス変動時のみではなく、シングルトークからダブルトークに変化した場合にも条件を満たし出力切替えが行われる。ダブルトーク時には係数を保持しなければならないので逆効果となる。従って、 N_2 の設定には注意が必要である。ステップサイズ切替えにより高速なエコーパス推定が実現できるので、シングルトークにおけるエコーパス変動には効果的である。

5 計算機シミュレーション

入力信号には図 6 に示す音声を用いる。最初はシングルトークの状態から始まり、50000 サンプルを過ぎたあたりからダブルトークに変わる。その後、再びシングルトークに戻り 110000 サンプル付近でエコーパスが変動する。最後に 200000 サンプル付近においてダブルトーク中にエコーパスが変動する。図中の点線はエコーパス変動のタイミングを表している。ダブルトーク、エコーパス変動について、それぞれ単独に

発生した場合と同時に起こった場合の影響について検証する。

2つの適応フィルタにはFIRフィルタを用いて、それぞれタップ数を64とし、AF2の学習アルゴリズムにはNLMSを用いた。相互相関をとる範囲は、短時間の平均幅を10サンプル、長時間の平均幅を500サンプルとした。エコーパスには15次のバターワースフィルタを用いた。P1, P2のパワー計算には式(7)の方法によりN=200サンプルの幅で行った。ステップサイズは大小2つの値のみとし、最大値0.3, 最小値0.01とする。図5にこの様子を示す。また、前章で述べた各条件のパラメータは $\theta_1 = 0.02$, $\theta_2 = 0.05$, $N_1 = 100$, $N_2 = 2000$, $\mu_{max} = 1.0$ とした。

$$P_i = \left[\frac{1}{N} \sum_{n=1}^N e_i(n)^2 \right]^{\frac{1}{2}} \quad (7)$$

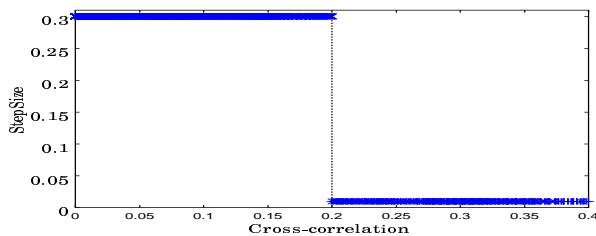


図5: 可変ステップサイズ

この条件の下で、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。但し、AF2の係数更新は遠端音声のパワーが0.05以上で学習を行うこととした。結果を図7に示す。

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

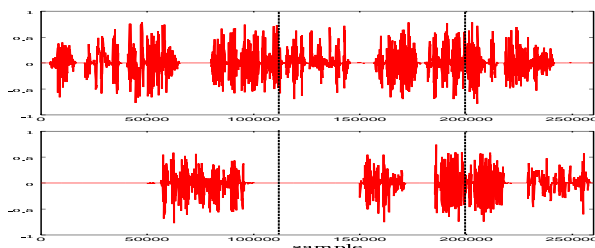


図6: 入力音声 (上:遠端音声, 下:近端音声)

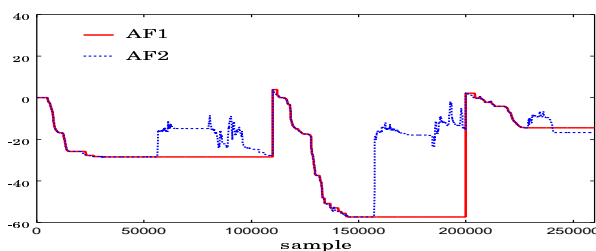


図7: 係数誤差の学習曲線

50000サンプルからのダブルトークに対して、適応フィルタAF2は劣化しているが、劣化直前のフィルタ係数をAF1に転送している為に良好な収束特性を維持している。また、110000サンプル付近でのエコーパス変動では特性は一時劣化するが、初期化条件とステップサイズ条件により素早く最適値に達していることが確認できる。200000サンプル付近では、ダブルトーク期間ではあるがステップサイズを小さくし、徐々にではあるが学習が進んでいる。この結果からも可変ステップサイズの有効性が確認できる。このシミュレーションの場合では、繰り返し回数260000回のうちAF2からAF1へのフィルタ係数の転送回数は23回であった。

6 むすび

音響エコーキャンセラにおいて、双方向通話時における収束特性の改善を目的として、遠端音声のエコーと近端音声の相互相関を近似的に導出する方法を基に、可変ステップサイズを提案した。さらに、ダブルトークによる特性劣化を避ける為に適応フィルタを2つ用いたDF方式を提案し、可変ステップサイズを組み合わせることで、高速かつ安定な学習が可能となった。また、2つの出力信号のパワーを比較することでエコーパス変動への対応が実現できた。計算機シミュレーションにより提案法の有効性を確認した。

参考文献

- [1]S.Haykin, “Adaptive Filter Theory” PRENTICE HALL,1996.
- [2]Y.Wang, K.Nakayama, Z.Ma: “A New Structure for Noise and Echo Cancelers Based on A Combined Fast Adaptive Filter Algorithm” IEICE Trans, vol.E78-A, pp845-853,July 1995.

口頭論文

- [3]牛丸 晋一郎, 中山謙二, 平野晃宏: “音響エコーキャンセラにおける双方向通話時の可変ステップ学習法”, 平成11年度電気関係学会北陸支部連合大会

- [4]牛丸 晋一郎, 中山謙二, 平野晃宏: “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”, Proc. of 14th DSP Symposium pp.683-687(1999).

- [5]牛丸 晋一郎, 中山謙二, 平野晃宏: “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”, 平成12年度電気関係学会北陸支部連合大会 pp.119.

- [6]牛丸 晋一郎, 中山謙二, 平野晃宏: “エコーと近端音声の相互相関に基づく可変ステップサイズ音響エコーキャンセラ”, Proc. of 15th DSP Symposium pp.683-687(2000).