

SA-7-11 エコーキャンセラ用ES射影アルゴリズム

ES Projection Algorithm for Echo Cancellers

牧野 昭二 金田 豊

Shoji Makino Yutaka Kaneda

NTT ヒューマンインタフェース研究所

NTT Human Interface Laboratories

1. まえがき

筆者らは、音響エコー経路のインパルス応答の変動量がインパルス応答と同じ減衰率で指数減衰することに着目し、従来のNLMS法(学習同定法)の収束速度を2倍以上に改善できるES(Exponentially Weighted Step Size)アルゴリズムを既に提案した[1][2]。また、最近では、筆者らと同様の考え方に基いたアルゴリズムもいくつか提案されている[3][4]。

一方、アルゴリズム内部で入力信号の自己相関を取り除くことにより、音声のように相関のある信号に対する収束速度を改善できる射影アルゴリズムが既に提案されている[5][6][7]。

本報告では、ES法と射影法のそれぞれの利点を生かしたES射影法を提案する。以下、本手法を定式化し、計算機シミュレーションおよびDSPで構成した実験装置の両者を用いて有効性を確認する。

2. ES射影アルゴリズム

音響エコーキャンセラの構成を図1に示す。ES法は、インパルス応答変動の大きい係数は大きなステップで、変動の小さな係数は小さなステップで修正するという考え方のもとに、従来のNLMS法のステップサイズ(スカラ量)に代わってステップサイズ行列A(対角行列)を導入したものである。このステップサイズ行列Aの対角成分 α_i は、インパルス応答と同じ減衰率で指数減衰させる。ES法により、収束速度を従来のNLMS法の2倍以上に改善できる。

ES法が音響エコー経路の変動特性のみに着目した手法であるのに対して、射影法は入力信号の性質のみに着目した手法である。射影法は、アルゴリズム内部で入力信号の自己相関を取り除くことにより、音声のように相関のある信号に対する収束速度を改善するという考え方に基いている。従来のNLMS法では、修正後の適応フィルタ $\hat{\mathbf{h}}(k+1)$ が入力信号ベクトル $\mathbf{x}(k)$ に対してのみ正しい出力 $y(k)$ を得るように $\hat{\mathbf{h}}(k)$ を修正するのに対して、2次の射影法では、過去の2個の入力信号ベクトル $\mathbf{x}(k)$, $\mathbf{x}(k-1)$ に対して正しい出力 $y(k)$, $y(k-1)$ を得るように $\hat{\mathbf{h}}(k)$ を修正する。すなわち、

$$\mathbf{x}(k)^T \hat{\mathbf{h}}(k+1) = y(k) \quad (1)$$

$$\mathbf{x}(k-1)^T \hat{\mathbf{h}}(k+1) = y(k-1) \quad (2)$$

ただし、

$$\mathbf{x}(k) = (x(k), x(k-1), \dots, x(k-L+1))^T$$

$$\hat{\mathbf{h}}(k) = (\hat{h}_1(k), \hat{h}_2(k), \dots, \hat{h}_L(k))^T$$

$y(k)$: エコー *^T: ベクトルの転置

L: タップ数 k: 離散化時間

射影法により、実音声入力に対する収束速度を従来のNLMS法の

約2倍に改善できる。

従来の射影法による適応フィルタ $\hat{\mathbf{h}}(k)$ の逐次修正式は

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \alpha [\beta(k) \mathbf{x}(k) + \gamma(k) \mathbf{x}(k-1)] \quad (3)$$

ただし、

α : ステップサイズ(スカラ量) $\beta(k), \gamma(k)$: 定数

と表わされる。射影法にES法の考え方を適用するために、従来の射影法のステップサイズ α (スカラ量)に代わってステップサイズ行列Aを導入して本手法を

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mathbf{A} [\beta(k) \mathbf{x}(k) + \gamma(k) \mathbf{x}(k-1)] \quad (4)$$

ただし、

$\mathbf{A} = \text{diag}[\alpha_1, \alpha_2, \dots, \alpha_L]$: ステップサイズ行列

$\alpha_i = \alpha_0 \lambda^{i-1}$ ($i=1, 2, \dots, L$)

λ : 音響エコー経路のインパルス応答の指数減衰率($0 < \lambda < 1$)

と表わす。(4)式を(1)式に代入すれば

$$\mathbf{x}(k)^T \hat{\mathbf{h}}(k) + \beta(k) \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k) + \gamma(k) \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k-1) = y(k) \quad (5)$$

となる。

$$\mathbf{x}(k)^T \hat{\mathbf{h}}(k) = \hat{y}(k) \quad (6)$$

であるから(5)式は

$$\beta(k) \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k) + \gamma(k) \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k-1) = y(k) - \hat{y}(k) = e(k) \quad (7)$$

となる。同様に(4)式を(2)式に代入すれば

$$\mathbf{x}(k-1)^T \hat{\mathbf{h}}(k) + \beta(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k) + \gamma(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k-1) = y(k-1) \quad (8)$$

となる。 $\hat{\mathbf{h}}(k)$ は1時刻前の(1)式を満たすので

$$\mathbf{x}(k-1)^T \hat{\mathbf{h}}(k) = y(k-1) \quad (9)$$

であるから(8)式は

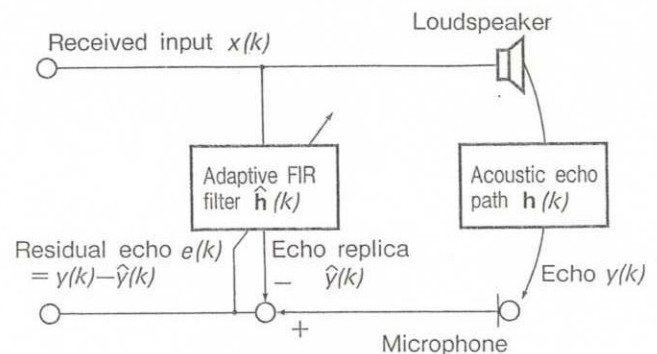


Fig.1 Configuration of an acoustic echo canceller.

$$\beta(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k) + \gamma(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k-1) = 0 \quad (11)$$

となる。

連立方程式(7)(11)を解いて定数 $\beta(k)$ 、 $\gamma(k)$ を求め、(4)式に代入すれば(1)(2)式を満足する $\hat{\mathbf{h}}(k+1)$ が求まる。

以上を基にして新しいステップサイズ μ (スカラー量)を導入すると、 $\mu=1$ の場合以外は(10)式は満足されなくなることに注意して本手法を書き下せば、

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mu \mathbf{A} [\beta(k) \mathbf{x}(k) + \gamma(k) \mathbf{x}(k-1)] \quad (11)$$

$$\beta(k) \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k) + \gamma(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k) = e(k) \quad (12)$$

$$\beta(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k) + \gamma(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k-1) = (1-\mu)e(k-1) \quad (13)$$

$$e(k) = y(k) - \hat{y}(k) \quad (14)$$

$$\hat{y}(k) = \hat{\mathbf{h}}(k)^T \mathbf{x}(k) \quad (15)$$

ただし、

μ : ステップサイズ(スカラー量)

となる。

エコーキャンセラを複数のDSPチップで構成する場合には、ステップサイズ α_i の指数減衰曲線を階段状に近似し、各チップ毎に一定の α_i を設定する。これにより、従来の射影法とほぼ同等の演算量と記憶容量で本手法を実現することができる。

3. 計算機シミュレーション

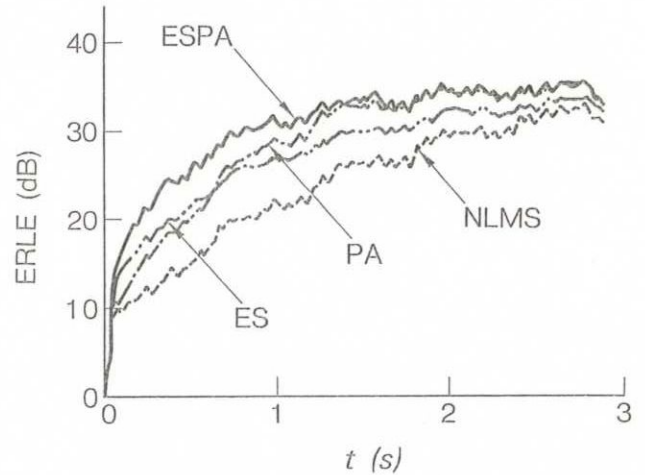
実測したインパルス応答(512タップ、サンプリング周波数8kHz)を用いて計算機シミュレーションを行った。受話入力には実音声を用い、エコー信号にはS/N比が35dBとなるように近端雑音を加えた。エコー消去量(ERLE)の収束特性の50回の平均値を図2に示す。実線は本手法(ESPA, $\mu=0.5$)、一点破線は射影法(PA, $\mu=0.5$)、二点破線はES法($\mu=1$)、破線はNLMS法($\mu=1$)を用いた場合である。ここでは、定常エコー消去量がほぼ等しくなるように、各手法のステップサイズ μ を設定した。図2より、エコー消去量20dBに達する収束速度は従来法に比べて、ES法および射影法で約2倍、本手法で約4倍になることがわかる。

4. 実時間評価実験

DSPで構成した実験装置を用いて会議室内(残響時間300ms(500Hz))で本手法の実時間評価実験を行った。実験に使用した音響エコーキャンセラは、7kHz帯域を2分割し、それぞれの帯域(サンプリング周波数8kHz)で3072タップ(エコー消去時間384ms)とした。入力信号には実音声を用いた。収束特性の実験結果を残留エコーレベル(= $e(k)$ の電力のレベル)を用いて図3に示す。収束特性の実時間評価結果(図3)は、計算機シミュレーション結果(図2)とほぼ同様であることが分かる。

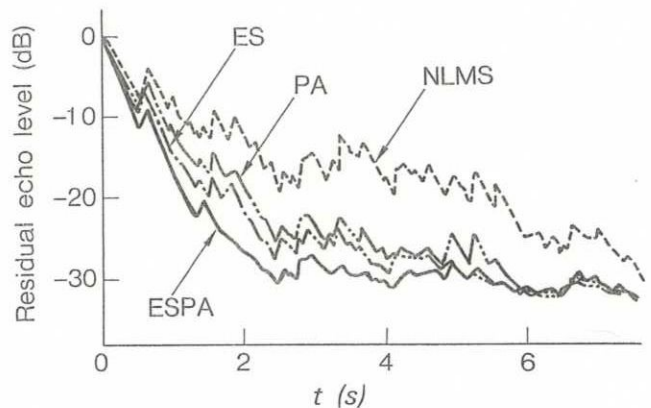
5. あとがき

音響エコー経路の変動特性と入力信号の相関除去を共に反映させたES射影アルゴリズムを提案し、音響エコーキャンセラの収束性能改善を試みた。室内音場のインパルス応答データを用いた計算機シミュレーションおよびDSPで構成した実験装置を用いた実時間評価実験を行い、NLMS法を用いた従来のエコーキャンセラに比べて、実音声入力に対する収束速度を約4倍にできることを明らかにした。



Input signal : speech (male)

Fig.2 Computer simulation results on ERLE convergence. The number of taps is 512 and sampling frequency is 8 kHz. The input signal is speech. Ambient noise with a fixed SNR of 35 dB is added.



Input signal : speech (male)

Fig.3 Real-time experimental results on residual echo level convergence using an acoustic echo canceller constructed with multiple DSP chips. The number of taps is 3072 and sampling frequency is 8 kHz in two bands. Room reverberation time at 500 Hz is 300 ms.

[参考文献]

- [1] 牧野昭二, 小泉宣夫, 音講論, 1-5-13, p. 355(1988. 3).
- [2] S. Makino, Y. Kaneda, ICASSP90, PP. 1133-1136(1990. 4).
- [3] 丸山唯介, 信学全大, B-557, p. 3-263(1989. 3).
- [4] 宇佐川, 松尾, 森田, 江端, 信学技報EA91-35, PP. 25-32(1991. 7).
- [5] 尾関和彦, 梅田哲夫, 信学論(A), J67-A, 2, pp. 126-132(1984. 2).
- [6] 雛元孝夫, 前川禎男, 電学論(C), 95, 10, pp. 227-234(1975. 10).
- [7] H. Yasukawa, I. Furukawa, Y. Ishiyama, ICASSP89, PP. 2041-2044(1989. 5).