

# 音響エコーキャンセラ用ES射影アルゴリズム

Exponentially Weighted Step-Size Projection Algorithm for Acoustic Echo Cancellers

牧野 昭二\* 金田 豊\*  
Shoji MAKINO Yutaka KANEDA

## あらまし

本論文は、従来のNLMSアルゴリズム(学習同定法)と同程度の演算量と記憶容量で、音声入力に対して収束速度が約4倍の、新しい音響エコーキャンセラ用適応アルゴリズム<sup>†</sup>を紹介するものである。ES射影(exponentially weighted step-size projection)アルゴリズムと名付けたこの適応アルゴリズムは、室内インパルス応答の変化量の統計的性質を活用し、入力信号の白色化を行っている。本手法では、トランスバーサル形FIRフィルタのそれぞれの係数に対して、異なるステップサイズを用いている。これらのステップサイズは時不变で、室内インパルス応答の変化量の期待値に比例して重み付けられている。これにより、変化の大きな係数は大きなステップで、変化の小さな係数は小さなステップで修正される。本手法は、室内インパルス応答の変化量の期待値はインパルス応答と同じ減衰率で指數減衰するという知見に基づいている。また、本手法では隣合う入力信号ベクトルの間の相関を取り除いている。これは言い換れば、入力信号を白色化していることに相当する。この操作は、音声のような有色信号に対して効果的である。演算量を削減するために高速算法を導入し、複数のDSPを用いる実用的な構成に対してステップサイズを修正することにより、従来のNLMSアルゴリズムと同程度の演算量と記憶容量で本手法を実現できる。本手法は複数のDSPで構成された音響エコーキャンセラに適用され、優れた収束特性が実時間動作においても確認された。

## Abstract

The algorithm we propose has, for speech input, a convergence speed four times that of the normalized LMS (NLMS) but its computational load is almost the same. This algorithm uses a different step size for each coefficient of an adaptive transversal filter, and these step sizes are time-invariant and weighted in proportion to the expected variation of a room impulse response. They are weighted this way because this expected variation becomes progressively smaller (along the series) by the same exponential with which the energy of the impulse response decays. This algorithm also reflects the whitening of the input signal (that is, it removes the correlation between input vectors) and is therefore especially for speech, whose spectrum is highly nonwhite. To reduce the computational complexity, this algorithm incorporates a fast projection routine modified for a practical DSP structure. The computational load is thus only 2L multiply-add operation. The algorithm is implemented in an echo canceller constructed with multiple DSP chips, and its fast convergence is demonstrated.

\*NTT ヒューマンインターフェース研究所 NTT Human Interface Laboratories

<sup>†</sup>論文中のゴシック文字は、本論文末に解説があります。

©日本電信電話株式会社 1995

## 1 まえがき

テレビ会議システムをはじめとした拡声通話系において、ハウリングやエコーを防止し、自然な通話を実現するために、音響エコーキャンセラが盛んに検討されている。音響エコーキャンセラは、スピーカとマイクロホンの間の伝達特性を適応フィルタを用いて推定して、擬似エコーを作成し、真のエコーから差し引いて消去するものである。

音響エコーキャンセラに適用できる適応アルゴリズムとして、多くのアルゴリズムが提案されている。LMS (Least-mean-squares) アルゴリズム<sup>(1)</sup>は、安定で演算量が少ない。NLMS (Normalized LMS: 学習同定法) アルゴリズム<sup>(2)</sup>は、収束特性が入力信号の大きさに依存しないため、商用の音響エコーキャンセラ<sup>(3)(4)</sup>に広く用いられている。

多くの勾配形適応アルゴリズムにおいて使用されているステップサイズは、フィルタ係数の収束速度と推定精度の間のトレードオフの関係を制御する。このため、収束の途中で収束速度を高速化し、定常状態における推定精度を上げるために、時変ステップサイズや行列形式の時変ステップサイズ<sup>(5)</sup>が検討されている。しかし、これらの時変ステップサイズは、ステップサイズの制御が複雑となる。さらに、これらのアルゴリズムの収束速度は、白色雑音入力に対してステップサイズを1とした場合が上限であり、これを超えることはできない。

著者らは、従来のNLMSアルゴリズムの収束速度を2倍以上に改善できるES (exponentially weighted step-size NLMS) アルゴリズム<sup>(6)(7)</sup>をすでに提案した。ESアルゴリズムでは、トランスバーサル形FIRフィルタのそれぞれの係数に対して、異なるステップサイズを用いている。これらのステップサイズは時不变で、室内インパルス応答の変化量の期待値に比例して重み付けられている。ESアルゴリズムは、室内インパルス応答の変化量の期待値はインパルス応答と同じ減衰率で指数減衰するという知見に基づいている。

LMS, NLMSおよびESアルゴリズムの欠点は、隣合う入力信号ベクトルの相関が大きい音声のような有色信号が入力された場合に、収束速度が遅いことである。一例を挙げれば、8 kHzサンプリングの4000次のフィルタに対して、NLMSアルゴリズムの平均2乗誤差が収束するのに要する時間は、白色雑音入力のとき約2秒、音声入力のとき約10秒である。このため、LMS, NLMSおよびESアルゴリズムの収束速度を向上させることができ強く望まれている。この問題を解決するために、(アフィン)射影アルゴリズム<sup>(8)~(10)</sup>が提案されている。射影アルゴリズムは、隣合う入力信号ベクトル間の相関を取り除くことにより、音声入力に対する収束速度をNLMSアルゴリズムに比べて約2倍に改善できる。

本論文では、ESアルゴリズムと射影アルゴリズム<sup>†</sup>のそれぞ

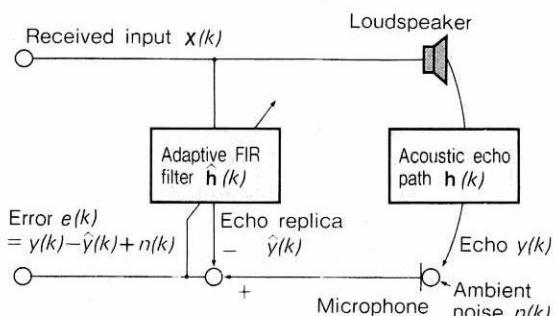
れの利点を共に生かして著者らが提案したES射影(exponentially weighted step-size projection)アルゴリズム<sup>(11)</sup>を紹介する。そして、本手法により音声入力に対する収束速度を従来のNLMSアルゴリズムに比べて約4倍に改善できることを示す。さらに、演算量を削減するために高速算法を使用し、複数のDSPを用いる実用的な構成に対してアルゴリズムを修正することにより、従来のNLMSアルゴリズムと同程度の演算量と記憶容量で本手法を実現できることを示す。

## 2 音響エコーキャンセラと従来の適応アルゴリズム

### 2.1 音響エコーキャンセラの構成

音響エコーキャンセラの構成を図1に示す。音響エコーキャンセラは、音響エコー経路の伝達特性、すなわちスピーカ・マイクロホン間のインパルス応答 $\mathbf{h}(k)$ を推定する。ここで、 $\mathbf{h}(k) = [h_1(k), h_2(k), \dots]^T$ であり、 $h_1(k), h_2(k), \dots$ は離散時間 $k$ におけるインパルス応答係数である。 $T$ はベクトル(行列)の転置を表す。インパルス応答 $\mathbf{h}(k)$ は、人の移動や周囲環境の変化に伴って変化するために、 $\mathbf{h}(k)$ の推定には適応フィルタ $\hat{\mathbf{h}}(k)$ が使用される。安定な実時間動作が可能である、多くの適応アルゴリズムが提案されている、などの理由から、通常、 $\hat{\mathbf{h}}(k)$ としてはトランスバーサル形FIRフィルタが用いられている。

図1において、 $\hat{\mathbf{h}}(k)$ と受話入力信号ベクトル $\mathbf{x}(k)$ が畳み込まれて擬似エコー $\hat{y}(k)$ が生成される。ここで、 $\mathbf{x}(k) = [x(k), x(k-1), \dots, x(k-L+1)]^T$ であり、 $L$ はFIRフィルタのタップ数である。次に、擬似エコー $\hat{y}(k)$ は真のエコー $y(k)$  ( $= \mathbf{x}(k)^T \mathbf{h}(k)$ )から差し引かれ、誤差 $e(k) = y(k) - \hat{y}(k) + n(k)$ が生成される。ここで、 $n(k)$ は周囲騒音である。近端話者の送話音声がマイクロホンに加わる場合(ダブルトーク状態)



音響エコーキャンセラは、音響エコー経路のインパルス応答 $\mathbf{h}(k)$ を、適応フィルタ $\hat{\mathbf{h}}(k)$ により推定して、擬似エコー $\hat{y}(k)$ を作成し、真のエコー $y(k)$ から差し引くことにより、エコー $y(k)$ を消去する。

図1 音響エコーキャンセラの原理

には、推定誤差を生じ性能の劣化をきたす。このため、通常、ダブルトークの期間中はエコーキャンセラの適応動作を停止させる。したがって、ここでは近端話者の送話音声は周囲騒音  $n(k)$  に含まないものとする。適応フィルタ  $\hat{\mathbf{h}}(k)$  は、誤差のパワーが小さくなるようにサンプリングごとに係数を修正する。適応アルゴリズムには実時間動作、速い収束速度、高い定常エコー消去量が要求される。ここで、エコー消去量 (ERLE: echo return loss enhancement) は真のエコーのパワーと周囲騒音を除いた誤差のパワーの比として定義される量である。

## 2.2 従来の適応アルゴリズム

### 2.2.1 NLMSアルゴリズム

NLMS (normalized least-mean-squares) アルゴリズム<sup>(2)</sup> は学習同定法とも呼ばれ、フィルタ係数ベクトル  $\hat{\mathbf{h}}(k)$  を次のように修正する。

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

$$e(k) = y(k) - \hat{y}(k) + n(k) \quad (2)$$

$$\hat{y}(k) = \mathbf{x}(k)^T \hat{\mathbf{h}}(k) \quad (3)$$

ただし、

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

$\hat{h}_i(k)$  ( $i=1, \dots, L$ ) : FIRフィルタ係数、

$\mathbf{x}(k) = [x(k), x(k-1), \dots, x(k-L+1)]^T$  : 受話入力信号ベクトル、

$n(k)$  : 周囲騒音、

$\alpha$  : スカラ・ステップサイズ ( $0 < \alpha < 2$ )。

フィルタ係数ベクトル  $\hat{\mathbf{h}}(k)$  は、受話入力信号ベクトル  $\mathbf{x}(k)$  の方向にのみ修正される。したがって、入力信号が音声のように有色で、隣合うベクトル  $\mathbf{x}(k)$ ,  $\mathbf{x}(k-1)$  の相関が大きい場合には、 $\hat{\mathbf{h}}(k)$  はほとんど修正されない。その結果、音声入力に対しては、白色雑音入力のときに比べて約5倍程度収束速度が遅くなる。

### 2.2.2 指数重み付け (ES) アルゴリズム

指数重み付け (ES: exponentially weighted step-size NLMS) アルゴリズム<sup>(6)(7)</sup> は、次のように表される。

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + A \frac{e(k)}{\mathbf{x}(k)^T \mathbf{x}(k)} \mathbf{x}(k) \quad (4)$$

ただし、

$$A = \begin{bmatrix} \alpha_1 & & 0 \\ & \alpha_2 & \\ 0 & \ddots & \alpha_L \end{bmatrix} \quad (5)$$

また、

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

$\gamma$  : 音響エコーエネルギー応答の指數減衰率 ( $0 < \gamma < 1$ )。ESアルゴリズムは、インパルス応答の変化量の期待値がインパルス応答と同じ減衰率で指數減衰するという経験則に基づいている。この知見を従来のNLMSアルゴリズムに取り入れるために、対角のステップサイズ行列  $A$  [式(5)] を導入し、トランスマルチポーラル形FIRフィルタのそれぞれの係数に対して、異なるステップサイズ  $\alpha_i$  を用いている。これらのステップサイズ  $\alpha_i$  は時不変で、 $\alpha_1$  から  $\alpha_L$  までインパルス応答  $\mathbf{h}(k)$  と同じ減衰率  $\gamma$  で指數減衰するように設定される。これにより、変化の大きな係数は大きなステップで、変化の小さな係数は小さなステップで修正される。ESアルゴリズムにより、従来のNLMSアルゴリズムに比べて収束速度を約2倍に改善できる。

指數減衰率  $\gamma$  は、同一室内のインパルス応答であれば同一である。したがって指數減衰率は、室内の音響条件から求めるか、あるいはインパルス応答を1回測定することにより求めることができる。複数のDSP (digital signal processor) チップを用いて構成する実際のシステムにおいては、指數減衰曲線を階段状に近似し、ステップサイズを各DSPチップごとに1つの定数として与える。これにより、従来のNLMSアルゴリズムと同等 ( $2L$ ) の演算量と記憶容量で本手法を実現できる。NLMSアルゴリズムと同様に、フィルタ係数ベクトル  $\hat{\mathbf{h}}(k)$  は受話入力信号ベクトル  $\mathbf{x}(k)$  の方向にのみ修正される。したがって、ESアルゴリズムにおいても入力信号が音声のような有色信号である場合には、白色雑音入力のときに比べて約5倍程度収束速度が遅くなる。

### 2.2.3 (アフィン) 射影アルゴリズム

(アフィン) 射影アルゴリズム<sup>(8)~(10)</sup> では、隣合う入力信号ベクトルの間の相関を取り除いている。これは言い換えれば、入力信号を白色化していることに相当する。射影アルゴリズムにより、音声入力に対する収束速度をNLMSアルゴリズムに比べて約2倍に改善できる。

2次の射影アルゴリズムは、フィルタ係数ベクトル  $\hat{\mathbf{h}}(k)$  を次のように修正するものである。

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \alpha [\beta_1(k) \mathbf{x}(k) + \beta_2(k) \mathbf{x}(k-1)] \quad (6)$$

ただし、

$\alpha$  : スカラ・ステップサイズ ( $0 < \alpha < 2$ )、

$\beta_1(k)$ ,  $\beta_2(k)$  : 時変パラメータ。

ここで、 $\tilde{\mathbf{h}}(k+1)$  を次のように定義する。

$$\tilde{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \beta_1(k) \mathbf{x}(k) + \beta_2(k) \mathbf{x}(k-1) \quad (7)$$

式(6)と式(7)を比較すれば、式(6)において  $\alpha = 1$  のとき、 $\tilde{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k+1)$  が成り立つことが分かる。パラメータ  $\beta_1(k)$ ,  $\beta_2(k)$  は、 $\tilde{\mathbf{h}}(k+1)$  が次式を満足するように定める。

$$\mathbf{x}(k)^T \tilde{\mathbf{h}}(k+1) = y(k) \quad (8)$$

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

式(6), (7)を用いれば式(8), (9)は

$$\beta_1(k)\mathbf{x}(k)^T\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k)^T\mathbf{x}(k-1) = e(k) \quad (10)$$

$$\beta_1(k)\mathbf{x}(k-1)^T\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1)^T\mathbf{x}(k-1) \quad (11)$$

$$= (1-\alpha)e(k-1) \quad (11)$$

ただし、

$$e(k) = y(k) - \hat{y}(k) + n(k) \quad (12)$$

$$\hat{y}(k) = \mathbf{x}(k)^T \hat{\mathbf{h}}(k) \quad (13)$$

となる。パラメータ  $\beta_1(k)$ ,  $\beta_2(k)$  は、連立 1 次方程式(10)(11)を解いて求める。このようにして求めた  $\beta_1(k)$ ,  $\beta_2(k)$  を式(6)に代入して  $\hat{\mathbf{h}}(k+1)$  を求める。フィルタ係数ベクトル  $\hat{\mathbf{h}}(k)$  は、 $\mathbf{x}(k)$  と  $\mathbf{x}(k-1)$  で張られる平面上の方向へ修正される。その結果、隣合う入力信号ベクトル  $\mathbf{x}(k)$ ,  $\mathbf{x}(k-1)$  の相関が大きい有色信号に対して、収束速度を改善することができる。

特に、 $\alpha=1$  の場合に、連立方程式(10)(11)を解き  $\beta_1(k)$ ,  $\beta_2(k)$  を求めたうえで、 $c(k)$  を

$$c(k) = -\frac{\beta_2(k)}{\beta_1(k)} = \frac{\mathbf{x}(k-1)^T \mathbf{x}(k)}{\mathbf{x}(k-1)^T \mathbf{x}(k-1)} \quad (14)$$

と定義する。さらに  $\mathbf{u}(k)$  を

$$\begin{aligned} \mathbf{u}(k) &= \mathbf{x}(k) - c(k)\mathbf{x}(k-1) \\ &= \mathbf{x}(k) - \frac{\mathbf{x}(k-1)\mathbf{x}(k-1)^T}{\mathbf{x}(k-1)^T \mathbf{x}(k-1)} \mathbf{x}(k) \end{aligned} \quad (15)$$

と定義する。式(15)第 2 項を幾何学的に解釈すれば  $\mathbf{x}(k)$  の  $\mathbf{x}(k-1)$  への射影を表している。したがって、 $\mathbf{x}(k-1)$  のうち  $\mathbf{x}(k)$  と相関のある部分は  $\mathbf{x}(k)$  から取り除かれ、その結果、 $\mathbf{u}(k)$  は  $\mathbf{x}(k-1)$  に直交していることが分かる。

式(6)に式(14), (15)を代入して、さらに、 $\mathbf{u}(k)^T \mathbf{u}(k) = \mathbf{x}(k)^T \mathbf{u}(k) - c(k)\mathbf{x}(k-1)^T \mathbf{u}(k) = \mathbf{x}(k)^T \mathbf{u}(k)$  を用いれば、2 次の射影アルゴリズムは次のように表される。

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \frac{e(k)}{\mathbf{u}(k)^T \mathbf{u}(k)} \mathbf{u}(k) \quad (16)$$

式(16)は、射影アルゴリズム(6), (10), (11)の  $\alpha=1$  の場合の特別な式である。フィルタ係数ベクトル  $\hat{\mathbf{h}}(k)$  は、“相関除去された”あるいは“白色化された”ベクトル  $\mathbf{u}(k)$  の方向へ修正される。

2 次の射影アルゴリズムは、フィルタ係数ベクトル  $\hat{\mathbf{h}}(k)$  を、次のように修正するものである。

$$\begin{aligned} \hat{\mathbf{h}}(k+1) &= \hat{\mathbf{h}}(k) + \alpha[\beta_1(k)\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1) + \dots \\ &\quad + \beta_p(k)\mathbf{x}(k-p+1)] \end{aligned} \quad (17)$$

再び、式(17)において  $\alpha=1$  とおいて  $\hat{\mathbf{h}}(k+1)$  を次のように定義する。

$$\begin{aligned} \hat{\mathbf{h}}(k+1) &= \hat{\mathbf{h}}(k) + \beta_1(k)\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1) + \dots \\ &\quad + \beta_p(k)\mathbf{x}(k-p+1) \end{aligned} \quad (18)$$

パラメータ  $\beta_1(k)$ ,  $\beta_2(k)$ , …,  $\beta_p(k)$  は、 $\hat{\mathbf{h}}(k+1)$  が次式を満足するように決定する。

$$\mathbf{x}(k)^T \hat{\mathbf{h}}(k+1) = y(k) \quad (19)$$

$$\mathbf{x}(k-1)^T \hat{\mathbf{h}}(k+1) = y(k-1) \quad (20)$$

…

$$\mathbf{x}(k-p+1)^T \hat{\mathbf{h}}(k+1) = y(k-p+1) \quad (21)$$

一般に射影アルゴリズムの次数が高くなるほど、収束速度が速くなるが、演算量も増加する。ここでは実用的な観点から、2 次の射影アルゴリズムを考える<sup>(12)</sup>。

### 3 新しい適応アルゴリズム

#### 3.1 ES 射影アルゴリズム

ES アルゴリズムは、インパルス応答の変化量の統計的性質のみに着目した手法である。これに対して射影アルゴリズムは、入力信号の白色化のみに着目した手法である。これら 2 つは互いに独立と考えられる。そこで、これら 2 つを共に反映させることを考える。ES アルゴリズムと射影アルゴリズムのそれぞれの利点を共に生かすために、式(6)のスカラ・ステップサイズ  $\alpha$  を式(5)のステップサイズ行列  $\mathbf{A}$  と定数 (scaling factor)  $\mu$  で置き換える。すると、式(6)は、

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mu \mathbf{A} [\beta_1(k)\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1)] \quad (22)$$

となる。再び、 $\mu=1$  と置いて、 $\tilde{\mathbf{h}}(k+1)$  を次のように定義する。

$$\tilde{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mathbf{A} [\beta_1(k)\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1)] \quad (23)$$

パラメータ  $\beta_1(k)$ ,  $\beta_2(k)$  は、 $\tilde{\mathbf{h}}(k+1)$  が式(8)(9)を満足するようになる。

式(23)を式(8)に代入すれば

$$\begin{aligned} \mathbf{x}(k)^T \hat{\mathbf{h}}(k) + \beta_1(k)\mathbf{x}(k)^T \mathbf{A}\mathbf{x}(k) \\ + \beta_2(k)\mathbf{x}(k)^T \mathbf{A}\mathbf{x}(k-1) = y(k) \end{aligned} \quad (24)$$

となる。

$$\mathbf{x}(k)^T \hat{\mathbf{h}}(k) = \hat{y}(k) \quad (25)$$

であるから、式(24)は

$$\begin{aligned} \beta_1(k)\mathbf{x}(k)^T \mathbf{A}\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1)^T \mathbf{A}\mathbf{x}(k) \\ = y(k) - \hat{y}(k) = e(k) \end{aligned} \quad (26)$$

となる。

同様にして、式(23)を式(9)に代入すれば、

$$\begin{aligned} \mathbf{x}(k-1)^T \hat{\mathbf{h}}(k) + \beta_1(k)\mathbf{x}(k-1)^T \mathbf{A}\mathbf{x}(k) \\ + \beta_2(k)\mathbf{x}(k-1)^T \mathbf{A}\mathbf{x}(k-1) = y(k-1) \end{aligned} \quad (27)$$

となる。一方、式(22)(23)から

$$\hat{\mathbf{h}}(k+1) = \mu \tilde{\mathbf{h}}(k+1) + (1-\mu) \hat{\mathbf{h}}(k) \quad (28)$$

の関係が得られる。式(28)は、 $k$  を  $k-1$  で置き換えて成り立つから、

$$\hat{\mathbf{h}}(k) = \mu \tilde{\mathbf{h}}(k) + (1-\mu) \hat{\mathbf{h}}(k-1) \quad (29)$$

となる。式(29)を式(27)に代入すれば

$$\begin{aligned} \beta_1(k)\mathbf{x}(k-1)^T \mathbf{A}\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1)^T \mathbf{A}\mathbf{x}(k-1) \\ = y(k-1) - \mu \mathbf{x}(k-1)^T \tilde{\mathbf{h}}(k) \\ - (1-\mu) \mathbf{x}(k-1)^T \hat{\mathbf{h}}(k-1) \end{aligned} \quad (30)$$

となる。式(8)(25)は、 $k$ を $k-1$ で置き換えると成り立つから、

$$\mathbf{x}(k-1)^T \hat{\mathbf{h}}(k) = y(k-1) \quad (31)$$

$$\mathbf{x}(k-1)^T \hat{\mathbf{h}}(k-1) = \hat{y}(k-1) \quad (32)$$

となる。式(31)(32)を式(30)に代入すれば

$$\begin{aligned} \beta_1(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k) + \beta_2(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k-1) \\ = (1-\mu) y(k-1) - (1-\mu) \hat{y}(k-1) \\ = (1-\mu) e(k-1) \end{aligned} \quad (33)$$

となる。連立方程式(26)(33)を解いてパラメータ $\beta_1(k)$ ,  $\beta_2(k)$ を求める。このようにして求めた $\beta_1(k)$ ,  $\beta_2(k)$ を式(22)に代入して $\hat{\mathbf{h}}(k+1)$ を求める。

以上をまとめると、ES射影(exponentially weighted step-size projection)アルゴリズムは、次のように表される。

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mu \mathbf{A} [\beta_1(k) \mathbf{x}(k) + \beta_2(k) \mathbf{x}(k-1)] \quad (34)$$

$$\beta_1(k) \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k) + \beta_2(k) \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k-1) = e(k) \quad (35)$$

$$\beta_1(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k) + \beta_2(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k-1) \\ = (1-\mu) e(k-1) \quad (36)$$

$$e(k) = y(k) - \hat{y}(k) + n(k) \quad (37)$$

$$\hat{y}(k) = \mathbf{x}(k)^T \hat{\mathbf{h}}(k) \quad (38)$$

ただし、

$\mu$ :定数 (scaling factor)。

ES射影アルゴリズム(34)～(38)を従来の射影アルゴリズム(6)(10)～(13)と比較すれば、式(6)におけるスカラ・ステップサイズ $\alpha$ が、式(34)ではステップサイズ行列 $\mathbf{A}$ と定数 $\mu$ で置き換えられている。式(10)の $\mathbf{x}(k)^T \mathbf{x}(k)$ と $\mathbf{x}(k)^T \mathbf{x}(k-1)$ は、式(35)では $\mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k)$ と $\mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k-1)$ に置き換えられている。さらに、式(11)の $\mathbf{x}(k-1)^T \mathbf{x}(k)$ と $\mathbf{x}(k-1)^T \mathbf{x}(k-1)$ は、式(36)では $\mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k)$ と $\mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k-1)$ に置き換えられている。式(11)のスカラ・ステップサイズ $\alpha$ は、式(36)では定数 $\mu$ に置き換えられている。本手法は、

$$0 < \mu < 2 \quad (39)$$

のとき収束する<sup>(11)</sup>。

### 3.2 演算量の削減

ES射影アルゴリズムの演算量は、中間変数 $z(k)$ <sup>(13)</sup>を導入することにより次のように削減できる。

$$\mathbf{z}(k+1) = \mathbf{z}(k) + \mu \mathbf{A} [\beta_1(k-1) + \beta_2(k)] \mathbf{x}(k-1) \quad (40)$$

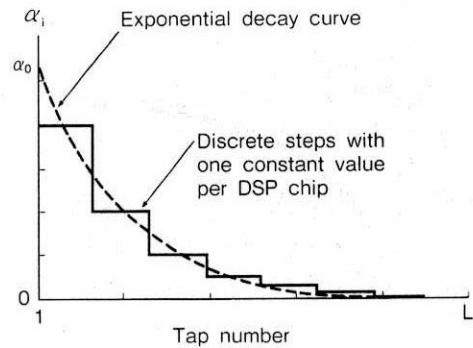
$$\beta_1(k) \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k) + \beta_2(k) \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k-1) = e(k) \quad (41)$$

$$\beta_1(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k) + \beta_2(k) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k-1) \\ = (1-\mu) e(k-1) \quad (42)$$

$$e(k) = y(k) - \hat{y}(k) + n(k) \quad (43)$$

$$\hat{y}(k) = \mathbf{z}(k)^T \mathbf{x}(k) + \mu \beta_1(k-1) \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k) \quad (44)$$

式(34)の計算には、各 $\alpha_i$ が $\mu \beta_1(k) \mathbf{x}(k) + \mu \beta_2(k) \mathbf{x}(k-1)$ に掛けられるため、 $3L$ 回の積と $2L$ 回の和が必要となる。一方、式(40)は、 $2L$ 回の積と $L$ 回の和しか必要としない。また、 $\mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k)$ は式(41)ですでに計算されているため、式(38)と比べたと



複数のDSPチップを用いて構成する実際のシステムにおいては、指指数減衰曲線を階段状に近似し、各DSPチップごとに1つの定数として与える。

図2 ステップサイズ $\alpha_i$ の与え方

き、式(44)における演算量の増加は無視できる。

中間変数 $z(k)$ はフィルタ係数ベクトル $\hat{\mathbf{h}}(k)$ と次のように関係付けられていることを付記しておく。

$$\mathbf{z}(k) = \hat{\mathbf{h}}(k) - \mu \mathbf{A} \beta_1(k-1) \mathbf{x}(k-1) \quad (45)$$

### 3.3 複数DSP構成に対する実用的な修正

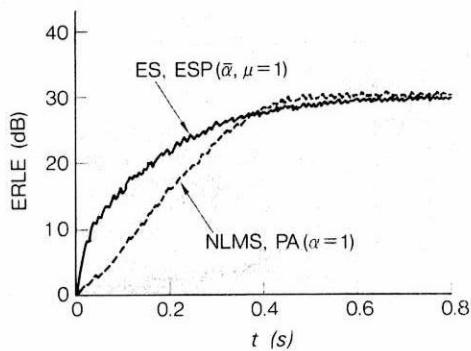
複数のDSP (digital signal processor) チップで構成される実際の音響エコーキャンセラに本手法を適用する場合には、指指数減衰曲線を図2に示すように階段状に近似し、ステップサイズ $\alpha_i$ を各DSPチップごとに1つの定数として与える。

$\alpha_i$ は各DSPチップで1つの定数であるから、式(40)の計算には $L$ 回の積と演算しか必要としない。 $\mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k)$ は各DSPチップで計算した値の総和として求められる。各DSPでの値は、 $\mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k-1)$ に最新の値 $[\mathbf{x}(k) \alpha_i \mathbf{x}(k)]$ を加え、最も古い値を差し引くことにより求めることができる。同様の計算方法が $\mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k-1)$ の計算にも利用できる。したがって、連立方程式(41)(42)を解くのに必要とする演算量は少ない。また、式(44)は $L$ 回の積と演算しか必要としない。このようにして、実用的な修正により、本手法の演算量と記憶容量は、従来のNLMSアルゴリズムと同程度(約 $2L$ 回の積と演算)となる。

### 4 本手法の収束特性

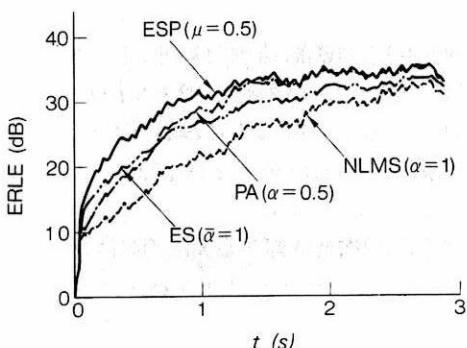
#### 4.1 計算機シミュレーション

ES射影アルゴリズムのエコー消去量の収束特性の計算機シミュレーション結果を図3に示す。計算機シミュレーションに用いたインパルス応答は、残響時間150msの室内で実測したものである。スピーカ・マイクロホン間距離は1mである。タップ数は512タップ、サンプリング周波数は8kHzである。フィルタ係数の初期値は0とした。受話入力には白色雑音および音声を用



ESアルゴリズムとES射影アルゴリズム(図中ESP)の収束速度は、NLMSアルゴリズムと射影アルゴリズム(図中PA)の約2倍である。

(a) 白色雑音入力



ESアルゴリズムと射影アルゴリズムの収束速度は、共に、NLMSアルゴリズムの約2倍である。ES射影アルゴリズムの収束速度は、NLMSアルゴリズムの約4倍である。

(b) 音声入力

タップ数：512，サンプリング周波数：8 kHz，SH比：白色雑音入力時30dB，音声入力時35dB，スピーカ・マイクロホン間距離：1 m，残響時間(500Hz)：150ms。

図3 ES射影アルゴリズムの収束特性の計算機シミュレーション結果

いた。白色雑音にはSN比30dB、音声にはSN比35dBの周囲騒音(白色雑音)を加えた。

エコー消去量(ERLE)は広く用いられており、次のように定義される。

$$ERLE = 10 \log_{10} \frac{p_y}{p_e} (\text{dB}) \quad (46)$$

ただし、

$$p_y = E[y(k)^2],$$

$$p_e = E[\tilde{e}(k)^2] = E[(y(k) - \hat{y}(k))^2],$$

$E[\cdot]$ ：期待値。

計算機シミュレーションにおけるエコーのパワー $p_y$ と残留エコーのパワー $p_e$ は、白色雑音に対しては10サンプル、音声に対しては256サンプルの2乗和から求めた。収束曲線は50回の試行の平均値である。

入力信号として音声を用いた場合には、NLMSアルゴリズムのスカラ・ステップサイズを1.0、射影アルゴリズムのスカラ・ステップサイズを0.5、ESアルゴリズムの平均ステップサイズを1.0、ES射影アルゴリズムの定数(scaling factor)を0.5に設定した。入力信号として白色雑音を用いた場合には、それらすべてを1.0と設定した。これらのステップサイズの値は、定常エコー消去量がほぼ等しくなるように設定した。

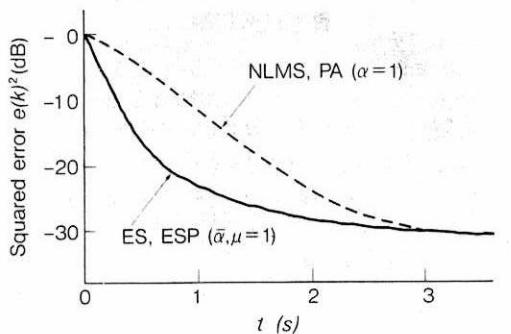
入力信号として白色雑音を用いた場合[図3(a)]には、射影アルゴリズム(図中PA)の収束速度は、NLMSアルゴリズムとほぼ等しく、ES射影アルゴリズム(図中ESP)はESアルゴリズムとほぼ等しいことが分かる。これらから、射影アルゴリズムの白色化の効果は白色雑音に対しては効果がないことが分かる。エコー消去量20dBに達する時間の逆数として収束速度を定義すれば、ESアルゴリズムとES射影アルゴリズムの収束速度は、NLMSアルゴリズムと射影アルゴリズムの約2倍である。これは、ESアルゴリズムとES射影アルゴリズムが音響エコー経路の変動特性を活用しているためである。

入力信号として音声を用いた場合[図3(b)]には、ESアルゴリズムと射影アルゴリズムの収束速度は共に、NLMSアルゴリズムの約2倍である。これはそれぞれ、ESアルゴリズムによる音響エコー経路の変動特性活用の効果と、射影アルゴリズムによる音声信号の白色化効果のためである。ES射影アルゴリズムの収束速度は、NLMSアルゴリズムの約4倍である。これは、ESアルゴリズムと射影アルゴリズムの効果が独立に働いたためである。

#### 4.2 実時間評価実験

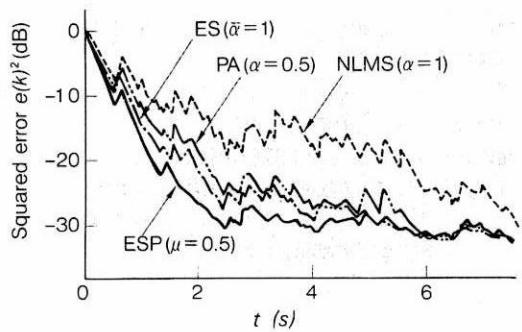
ES射影アルゴリズムを複数のDSPチップ<sup>(14)</sup>で構成した音響エコーキャンセラに組み込んで、実時間評価実験を行った。帯域分割手法を用いて7kHz帯域を2分割し、それぞれの帯域で8kHzサンプリングした。タップ数はそれぞれの帯域で3072タップとした。入力信号には、白色雑音と音声を用いた。ステップサイズの値は計算機シミュレーションの場合と一緒に設定した。スピーカ・マイクロホン間距離は2.5m、実験に使用した室内の残響時間は300msである。

2乗誤差レベルの収束曲線を図4に示す。実験結果は4.1節の計算機シミュレーション結果とほぼ同様である。 $-20\text{dB}$ の2乗誤差レベルに達する時間の逆数として収束速度を定義すれば、入力信号として白色雑音を用いた場合[図4(a)]には、ESアルゴリズムとES射影アルゴリズム(図中ESP)の収束速度は、NLMSアルゴリズムと射影アルゴリズム(図中PA)の約2倍



ESアルゴリズムとES射影アルゴリズム(図中ESP)の収束速度は、NLMSアルゴリズムと射影アルゴリズム(図中PA)の約2倍である。

(a) 白色雑音入力



ESアルゴリズムと射影アルゴリズムの収束速度は、共に、NLMSアルゴリズムの約2倍である。ES射影アルゴリズムの収束速度は、NLMSアルゴリズムの約4倍である。

(b) 音声入力

複数のDSPチップで構成した帯域2分割の音響エコーキャンセラに適用。タップ数：低域、高域共に3,072、サンプリング周波数：8 kHz、スピーカ・マイクロホン間距離：2.5 m、実験に使用した室内の残響時間(500Hz)：300 ms。

図4 ES射影アルゴリズムの収束特性の実時間評価結果

である。

入力信号として音声を用いた場合[図4(b)]には、ESアルゴリズムと射影アルゴリズムの収束速度は共に、NLMSアルゴリズムの約2倍である。ES射影アルゴリズムの収束速度は、NLMSアルゴリズムの約4倍である。

このように、ES射影アルゴリズムは容易にNLMSアルゴリズムに置き換わることができ、商用の音響エコーキャンセラの音声入力に対する収束特性を改善することができる。

## 5 あとがき

音響エコーキャンセラ用の新しい適応アルゴリズムを紹介した本手法では、音響エコー経路のインパルス応答変動特性の統計的性質を活用し、かつ音声のような有色信号の白色化を行っている。本手法では、適応フィルタのそれぞれの係数を、異なる値のステップサイズで修正する。これらのステップサイズは時不变で、室内インパルス応答の変化量の期待値に比例して重み付けられている。さらに本手法では、隣合う入力信号ベクトル間の相関を取り除くことにより、NLMSアルゴリズムに比べて音声入力に対する収束速度を改善している。演算量を削減するために高速算法を導入し、複数のDSPを用いる実用的な構成に対してステップサイズを修正することにより、従来のNLMSアルゴリズムと同程度(約2L)の演算量と記憶容量で本手法を実現できる。本手法を複数のDSPで構成された音響エコーキャンセラに適用し、実際の室内で実時間評価実験を行った。その結果、本手法は従来のNLMSアルゴリズムに比べて、白色雑音入力に対して約2倍、音声入力に対して約4倍の収束速度を有することを示した。

## 文 献

- (1) B. Widrow and S. Stearns : Adaptive Signal Processing. Englewood Cliffs, NJ : Prentice-Hall, 1985.
- (2) J. Nagumo and A. Noda : A learning method for system identification, IEEE Trans. Automat. Contr., AC-12, No.3, pp.282-287, Jun., 1967.
- (3) Y. Itoh, Y. Maruyama, N. Furuya and T. Araseki : An acoustic echo canceller for teleconference, Proc. ICC85, pp.1498-1502, Jun., 1985.
- (4) H. Oikawa, S. Minami and T. Saeki : A new echo canceller realized by high performance digital signal processor, Proc. ISCAS88, pp.1329-1332, Jun., 1988.
- (5) R. Harris, D. Chabries and F. Bishop : A variable step (VS) adaptive filter algorithm, IEEE Trans. Acoust., Speech, Signal Processing, ASSP-34, No. 2, pp. 309-316, Apr., 1986.
- (6) S. Makino and Y. Kaneda : Acoustic echo canceller algorithm based on the variation characteristics of a room impulse response, Proc. ICASSP90, pp.1133-1136, Apr., 1990.
- (7) S. Makino, Y. Kaneda and N. Koizumi : Exponentially weighted step-size NLMS adaptive filter based on the statistics of a room impulse response, to be published in IEEE Trans. Speech and Audio, 1, No.1, Jan., 1993.
- (8) 尾関・梅田：アフィン部分空間への直交射影を用いた適応フィルタ・アルゴリズムとその諸性質、信学論(A), J67-A, pp.126-132, Feb., 1984.
- (9) 離元・前川：拡張された学習同定法：電学論(C), 95, pp.227-234, Oct., 1975.
- (10) H. Yasukawa, I. Furukawa and Y. Ishiyama : Acoustic echo

- control for high quality audio teleconferencing, Proc. ICAS-SP89, pp. 2041-2044, May, 1989.
- (11) S. Makino and Y. Kaneda : Exponentially weighted stepsize projection algorithm for acoustic echo cancellers, Trans. IEICE Japan, E75-A, pp. 1500-1508, Nov., 1992.
- (12) 田中・牧野・金田：音声入力に対する射影法の次数と収束特性について，音講論集1-4-14, pp. 489-490, Oct., 1992.
- (13) 丸山：射影アルゴリズムの高速算法，信学全大，B-744, pp. 3-322, Mar., 1990.
- (14) T. Kaneko, H. Yamauchi and A. Iwata : A 50ns floating-point signal processor VLSI, Proc. ICASSP86, pp. 401-404, Apr., 1986.



牧野 昭二

ヒューマンインターフェース研究所主幹研究員  
昭和56年入社 電気音響変換器、拡声電話器、および音響エコーキャンセラなどの音響信号処理の研究に従事。  
昭和54年東北大工学部機械工学第二学科卒業 56年同大学院機械工学修士課程修了。平成3年工学博士(東北大)。  
IEEE・電子情報通信学会・日本音響学会会員。



金田 豊

ヒューマンインターフェース研究所主幹研究員  
昭和52年入社 主に雑音抑圧マイクロホンアレーの研究に従事。現在、音響信号処理の研究に従事。  
昭和50年名古屋大学工学部電気工学科卒業 52年同大学院工学研究科修士課程修了。平成2年工学博士(名古屋大学)。  
日本音響学会・電子情報通信学会・米国音響学会・IEEE会員。  
平成元年3年日本音響学会論文賞受賞、平成元年IEEE ASSP誌論文賞受賞。

## ■用語解説■

### エコーキャンセラ

エコーキャンセラは、エコー経路の伝達特性を適応フィルタにより推定して擬似エコーを作成し、真のエコーから差し引いて消去する装置である。電話回線のハイブリッド・トランスのインピーダンス不整合により発生する回線エコーを対象とした回線エコーキャンセラや、スピーカ・マイクロホン間の音響エコーを対象とした音響エコーキャンセラなどがある。

### 適応アルゴリズム

適応アルゴリズムとは、各時刻において観測される入力信号と、その入力信号に対する出力誤差に基づいて、適応フィルタの係数を逐次的に修正していく処理手順である。エコーキャンセラに使用される適応アルゴリズムとしては、学習同定法(NLMSアルゴリズム)が最も一般的であるが、音声に対する収束速度が遅いという欠点を持つ。近年検討されている高速RLSアルゴリズムは、数値的に不安定であり、まだ実使用に耐えない。現状では、射影アルゴリズムが最も有望である。

### 射影アルゴリズム

射影アルゴリズムは、入力信号を白色化することにより音声のように有色な信号に対する収束速度を改善することを特徴とする適応アルゴリズムである。射影の次数は白色化フィルタの次数に相当する。1次の射影アルゴリズムは学習同定法(NLMSアルゴリズム)に相当し、無限次の射影アルゴリズムはRLS(逐次最小2乗)アルゴリズムに相当する。近年、演算量を削減した高速射影アルゴリズムが開発され、演算量の増加を伴わずに、収束速度の高速化が可能となつた。