Third International Workshop on
Acoustic Echo Control
7-8 Sept. 1993 Plestin les Grèves France

Exponentially Weighted Step-Size Projection Algorithm for Acoustic Echo Cancellers

Shoji Makino and Yutaka Kaneda

NTT Human Interface Laboratories
3-9-11, Midori-cho, Musashino-shi, Tokyo, 180 Japan
Tel. +81 422 59 3992 Fax. +81 422 60 7811
E-mail: makino@ac-lab. ntt. jp (internet)

This paper proposes a new adaptive algorithm for acoustic echo cancellers with four times the convergence speed for a speech input, at almost the same computational load, of the normalized LMS (NLMS). This algorithm reflects both the statistics of the variation of a room impulse response and the whitening of the received input signal [1].

This algorithm, called the ESP (exponentially weighted step-size projection) algorithm, uses a different step size for each coefficient of an adaptive transversal filter. These step sizes are time-invariant and weighted proportional to the expected variation of a room impulse response. As a result, the algorithm adjusts coefficients with large errors in large steps, and coefficients with small errors in small steps. The algorithm is based on the fact that the expected variation of a room impulse response becomes progressively smaller along the series by the same exponential ratio as the impulse response energy decay [2].

This algorithm also reflects the whitening of the received input signal, *i.e.*, it removes the correlation between consecutive received input vectors. This process is effective for speech, which has a highly non-white spectrum.

A geometric interpretation of the proposed algorithm is derived and the convergence condition is proved. A fast projection algorithm is introduced to reduce the computational complexity and modified for a practical multiple DSP structure so that it requires almost the same computational load, $2L$ multiply-add operations, as the conventional NLMS. The algorithm is implemented in an acoustic echo canceller constructed with multiple DSP chips, and its fast convergence is demonstrated.

References

[1] S. Makino and Y. Kaneda, " Exponentially weighted step-size projection algorithm for acoustic echo cancellers, " *Trans. IEICE Japan*, vol. E75-A, no. 11, pp. 1500-1508, Nov. 1992.

[2] S. Makino, Y. Kaneda, and N. Koizumi, " Exponentially weighted step-size NLMS adaptive filter based on the statistics of a room impulse response, " *IEEE Trans. Speech and Audio*, vol. 1, no. 1, pp. 101-108, Jan. 1993.

Exponentially Weighted Step-Size Projection Algorithm for Acoustic Echo Cancellers

Shoji Makino and Yutaka Kaneda

NTT Human Interface Laboratories
3-9-11, Midori-cho, Musashino-shi, Tokyo, 180 Japan
Tel.+81 422 59 3992 Fax.+81 422 60 7811
E-mail: makino@ac-lab.ntt.jp (internet)

1. ABSTRACT

This paper proposes a new adaptive algorithm for acoustic echo cancellers with four times the convergence speed for a speech input, at almost the same computational load, of the normalized LMS (NLMS). This algorithm reflects both the statistics of the variation of a room impulse response [1] and the whitening of the received input signal [2]. It is called the ESP (exponentially weighted step-size projection) algorithm. The algorithm is implemented in an acoustic echo canceller constructed with multiple DSP chips, and its fast convergence is demonstrated.

References

[1] S. Makino, Y. Kaneda, and N. Koizumi, *IEEE Trans. Speech and Audio*, vol. 1, no. 1, pp. 101-108, Jan. 1993.

[2] S. Makino and Y. Kaneda, *Trans. IEICE Japan*, vol. E75-A, no. 11, pp. 1500-1508, Nov. 1992.

2. CONVENTIONAL ADAPTIVE ALGORITHMS

2.1 NLMS Algorithm

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

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

where

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

$\hat{h}_i(k)(i = 1, \cdots, L)$: coefficients of an FIR filter,

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

: received input vector,

L: number of taps,

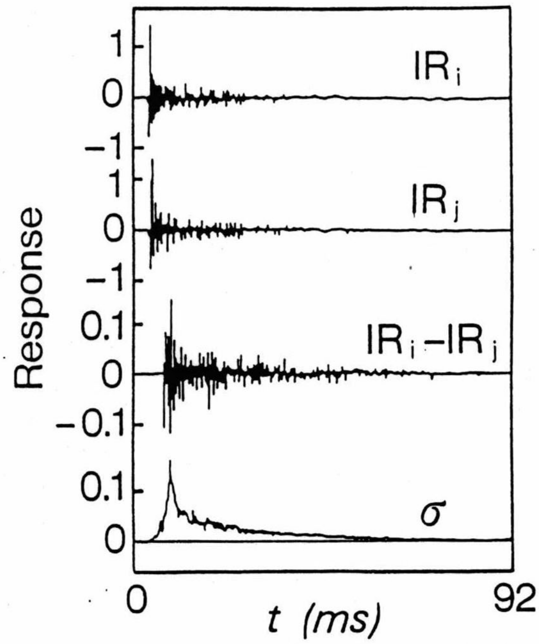
α : scalar step size $(0 < \alpha < 2)$.

- **In the conventional NLMS, convergence is fastest at α=1.**

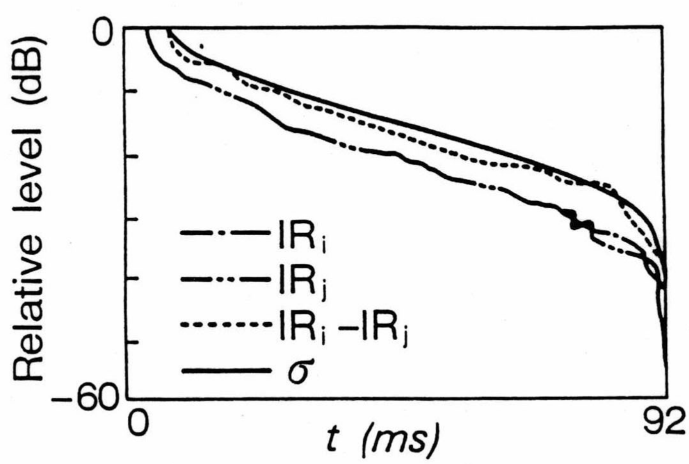
- **Filter coefficient vector $\hat{\mathbf{h}}(k)$ is adjusted only in the direction of the received input vector $\mathbf{x}(k)$.**

- **The NLMS algorithm converges very slowly when the received input signal is a colored signal such as speech.**

2.1 Variation of a Room Impulse Response



(a) Impulse responses and their variation



(b) Reverberent energy decay curves

• Impulse responses attenuate exponentially, and the expected variation of the impulse response also attenuates by the same exponential ratio.

• The exponential attenuation ratio is common to all impulse responses in the same room. We can estimate it from the room conditions, or determine it by measuring one impulse response.

2.3 Exponentially Weighted Step-Size NLMS (ES) Algorithm

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mathbf{A}\,\frac{e(k)}{\mathbf{x}(k)^T\mathbf{x}(k)}\mathbf{x}(k) \qquad (3)$$

where

$$\mathbf{A} = \begin{pmatrix} \alpha_1 & & & 0 \\ & \alpha_2 & & \\ & & \ddots & \\ 0 & & & \alpha_L \end{pmatrix} \qquad (4)$$

and

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

γ : exponential attenuation ratio $(0 < \gamma < 1)$.

- **The ES algorithm uses a different step size for each coefficient of an adaptive transversal filter by introducing a step size matrix A having diagonal form.**

- **Elements α_i are time-invariant and decrease exponentially from α_1 to α_L with the same ratio γ as the impulse response $\mathbf{h}(k)$.**

- **This algorithm updates coefficients with large errors in large steps, and coefficients with small errors in small steps [1].**

- **The algorithm doubles the convergence of the NLMS.**

2.4 Projection Algorithm

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \alpha[\beta_1(k)\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1)] \quad (5)$$

where

α : scalar step size$(0 < \alpha < 2)$,

$\beta_1(k), \beta_2(k)$: time-varying parameters.

Parameters $\beta_1(k)$ and $\beta_2(k)$ are determined so that $\hat{\mathbf{h}}(k+1)$ satisfies

$$\mathbf{x}(k)^T\hat{\mathbf{h}}(k+1) = y(k) \quad (6)$$

$$\mathbf{x}(k-1)^T\hat{\mathbf{h}}(k+1) = y(k-1). \quad (7)$$

Combining (6), (7), and (5) gives

$$\beta_1(k)\mathbf{x}(k)^T\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k)^T\mathbf{x}(k-1) = e(k) \quad (8)$$

$$\beta_1(k)\mathbf{x}(k-1)^T\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1)^T\mathbf{x}(k-1) = (1-\alpha)e(k-1). \quad (9)$$

- **Parameters $\beta_1(k)$ and $\beta_2(k)$ are obtained by solving simultaneous equations (8) and (9). $\hat{\mathbf{h}}(k+1)$ is obtained using $\beta_1(k), \beta_2(k)$, and (5).**

- **Filter coefficient vector $\hat{\mathbf{h}}(k)$ is adjusted in an arbitrary direction on the plane produced by $\mathbf{x}(k)$ and $\mathbf{x}(k-1)$.**

- **Convergence can be doubled for a speech input.**

• Special formula for $\alpha = 1$

$$c(k) = -\frac{\beta_2(k)}{\beta_1(k)} = \frac{\mathbf{x}(k-1)^T\mathbf{x}(k)}{\mathbf{x}(k-1)^T\mathbf{x}(k-1)} \tag{10}$$

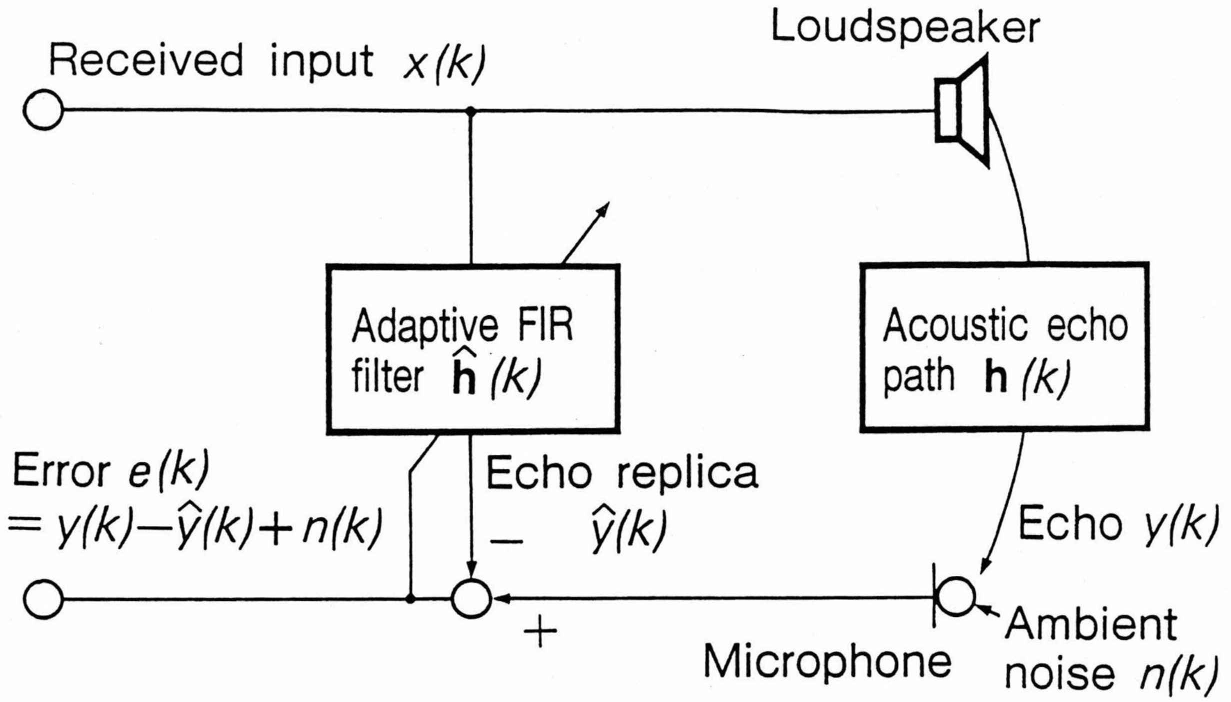
$$\mathbf{u}(k) = \mathbf{x}(k) - c(k)\mathbf{x}(k-1) \tag{11}$$

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \frac{e(k)}{\mathbf{u}(k)^T\mathbf{u}(k)}\mathbf{u}(k) \tag{12}$$

• **Correlated components of $\mathbf{x}(k-1)$ are subtracted from $\mathbf{x}(k)$, *i.e.*, $\mathbf{u}(k)$ is orthogonal to $\mathbf{x}(k-1)$.**

• **Filter coefficient vector $\hat{\mathbf{h}}(k)$ is adjusted in the direction of the " decorrelated " or " whitened " vector $\mathbf{u}(k)$.**

3. NEW ADAPTIVE ALGORITHM



- The ES algorithm only reflects the statistics of the variation of a room impulse response.
- The projection algorithm only reflects a whitening of the received input signal.
- These are independent of each other, so we aim to reflect both of them.

3.1 Exponentially Weighted Step-Size Projection (ESP) Algorithm

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mu\mathbf{A}[\beta_1(k)\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k-1)] \tag{13}$$

Parameters $\beta_1(k)$ and $\beta_2(k)$ are determined so that $\hat{\mathbf{h}}(k+1)$ satisfies (6) and (7).

Combining (6), (7), and (13) gives

$$\beta_1(k)\mathbf{x}(k)^T\mathbf{A}\mathbf{x}(k) + \beta_2(k)\mathbf{x}(k)^T\mathbf{A}\mathbf{x}(k-1) = e(k) \tag{14}$$

$$\beta_1(k)\mathbf{x}(k-1)^T\mathbf{A}\mathbf{x}(k)+\beta_2(k)\mathbf{x}(k-1)^T\mathbf{A}\mathbf{x}(k-1) = (1-\mu)e(k-1). \tag{15}$$

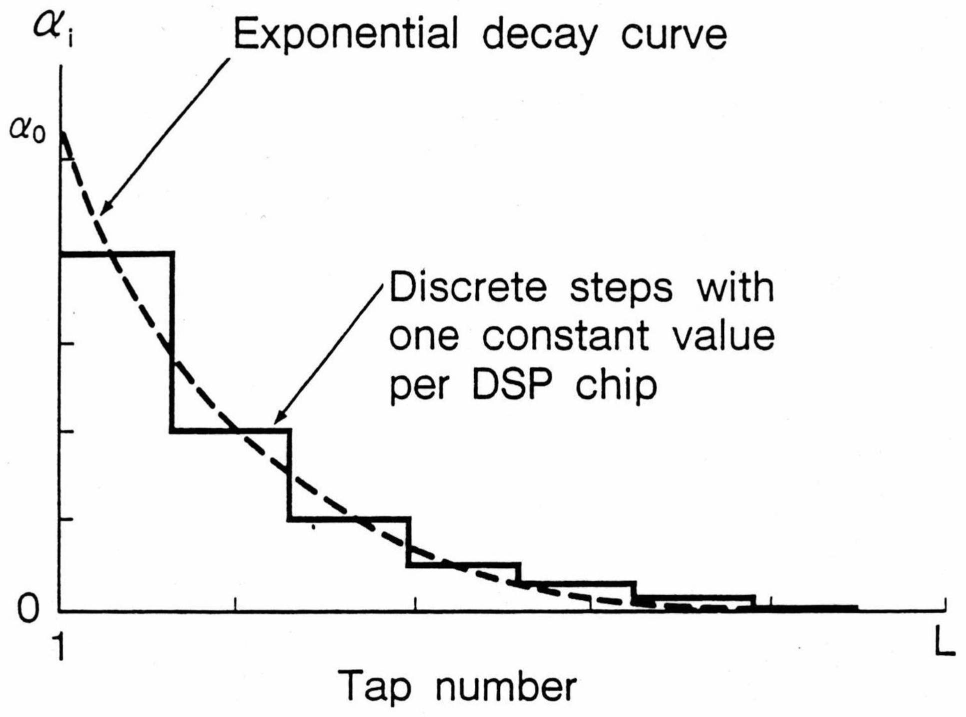
where

μ : step size matrix scaling factor ($0 < \mu < 2$).

• **Parameters $\beta_1(k)$ and $\beta_2(k)$ are obtained by solving simultaneous equations (14) and (15). $\hat{\mathbf{h}}(k+1)$ is obtained by using $\beta_1(k), \beta_2(k)$, and (13).**

3.2 Reduction of Computational Complexity

- The computational complexity of the ESP algorithm can be reduced by introducing an intermediate variable [2].



3.3 Practical Modification for a Multiple DSP Structure

- The exponential decay curve is approximated stepwise, and step size α_i is set in discrete steps.

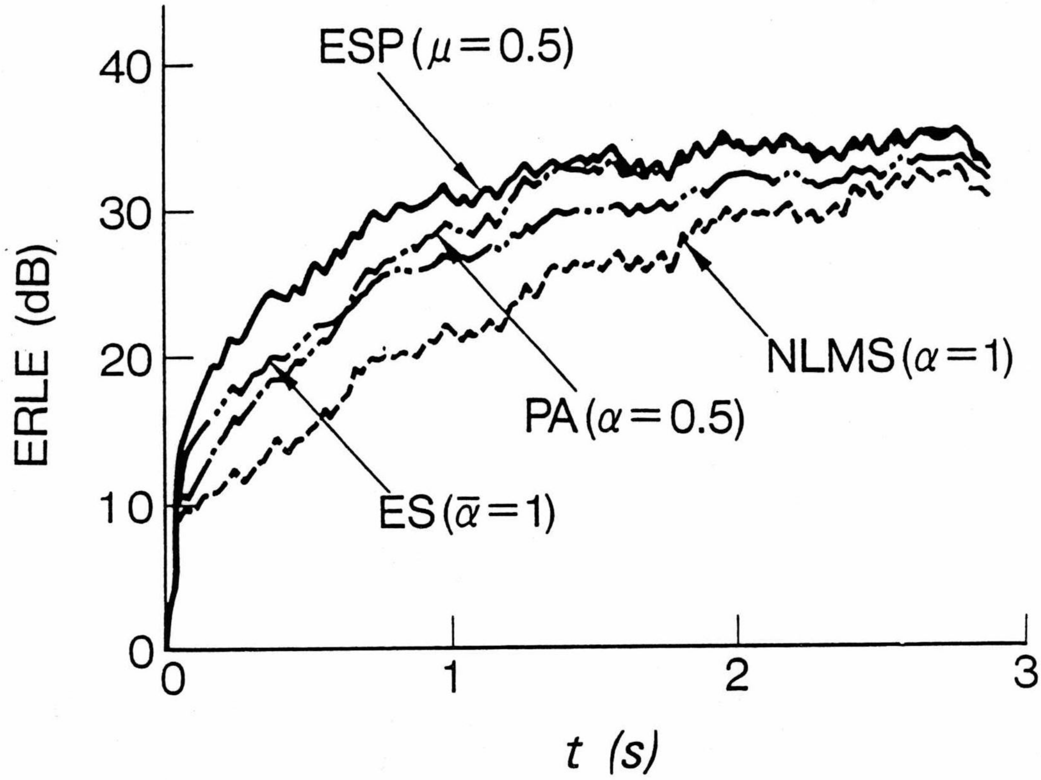
- This modification allows the proposed algorithm to have almost the same computational load as the NLMS.

4. CONVERGENCE COMPARISON

4.1 Computer simulations

- With a speech input, the ES and projection algorithms both converged twice as fast as the NLMS.

- Our ESP algorithm converged four times as fast as the NLMS.



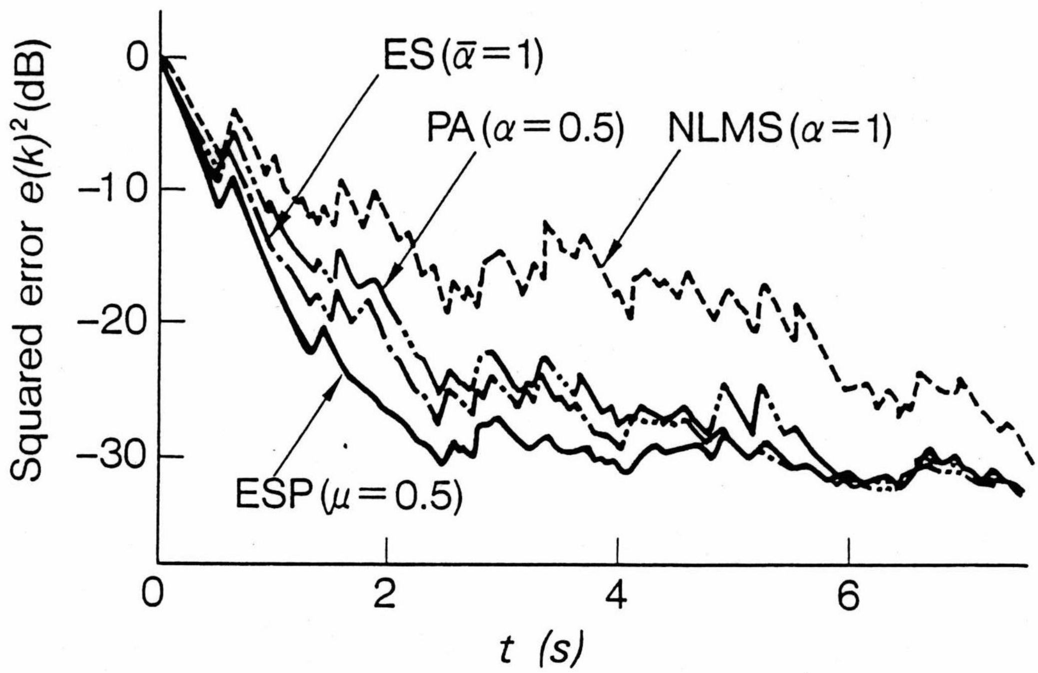
Simulation conditions

- number of taps: 512,
- sampling frequency: 8 kHz,
- received input: speech,
- SNR: 35 dB,
- step sizes were set so that steady-state ERLEs were almost equal,
- average of 50 independent results.

4.2 Real-time Experiments

• With a speech input, the ES and projection algorithms both converged twice as fast as the NLMS.

• Our ESP algorithm converged four times as fast as the NLMS.



Experimental conditions

· implemented in an acoustic echo canceller constructed with multiple DSP chips,

· 7-kHz frequency range was separated into two bands,

· sampling frequency: 8 kHz,

· number of taps: 3072 in each band,

· received input: speech,

· step sizes were set so that steady-state ERLEs were almost equal.

5. CONCLUSIONS

- Our new adaptive algorithm ESP reflects both the statistics of a room impulse response and the whitening of the received input signal.

- This algorithm has four times the convergence speed at almost the same computational load as the conventional NLMS.