

室内インパルス応答の統計的性質に基づく 指数重み付けNLMS適応フィルタ

Exponentially Weighted Step-Size NLMS Adaptive Filter
Based on the Statistics of a Room Impulse Response

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

あらまし 本論文は、従来のNLMSアルゴリズムと同等の演算量と記憶容量で収束速度が約2倍の、新しい音響エコーキャンセラ用適応アルゴリズムを提案するものである。ES (exponentially weighted step size) アルゴリズムと名付けたこの適応アルゴリズムは、適応形トランスバーサル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 (修正係数) を用いる。これらのステップサイズは時不変で、室内インパルス応答の変動量の期待値に比例して重み付けられている。本手法は、室内インパルス応答の変動量の期待値はインパルス応答のエネルギー減衰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。本手法により、誤差の大きな係数は大きなステップで、誤差の小さな係数は小さなステップで修正される。本手法による係数誤差の2乗期待値を定式化することにより、ステップサイズの平均値が収束条件、収束速度、定常状態における過剰平均2乗誤差を決定することを明らかにした。複数のDSPを用いる実用的な構成に対して修正することにより、従来のNLMSアルゴリズムと同等の演算量と記憶容量で本手法を実現できる。本手法は商用の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 This paper proposes a new normalized LMS (NLMS) adaptive algorithm with double the convergence speed, at the same computational load, of the conventional NLMS for an acoustic echo canceller. This algorithm, called the ES (exponentially weighted step size) algorithm, uses a different step size (feedback constant) for each weight of an adaptive transversal filter. These step sizes are time-invariant and weighted proportional to the expected variation of a room impulse response. 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. As a result, the algorithm adjusts coefficients with large errors in large steps, and coefficients with small errors in small steps. A transition formula is derived for the mean-squared coefficient error of the proposed algorithm. The mean step size determines the convergence condition, the convergence speed, and the final excess mean-squared error. The algorithm is modified for a practical multiple DSP structure, so that it requires only the same amount of computation as the conventional NLMS. The algorithm is implemented in a commercial acoustic echo canceller and its fast convergence is demonstrated.

1 まえがき

テレビ会議システムや拡声通信において発生するハウリングやエコーを防止し、自然な拡声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的に推定して、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。RLS (Recursive least-squares) アルゴリズム [1] は、収束速度は早いが演算量が多い。演算量の削減をはかった高速 RLS アルゴリズム [2][3] が近年検討されているが、まだ演算量が多く実時間動作は難しい。一方、LMS (Least-mean-squares) アルゴリズム [4][5] は、安定で演算量が少ない。

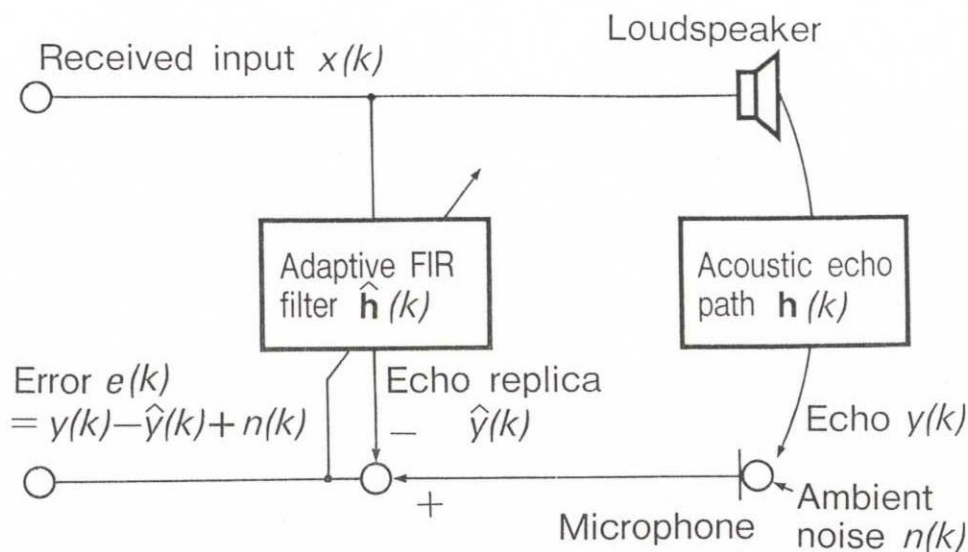


図 1: Configuration of an acoustic echo canceller.

NLMS (Normalized LMS: 学習同定法) アルゴリズム [6] は、収束速度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 [7][8] に広く用いられている。しかし、LMS および NLMS アルゴリズムの欠点は、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、白色雑音入力の時約 2 秒、実音声入力の時約 10 秒である。このため、LMS および NLMS アルゴリズムの収束速度を向上させることが強く望まれている。

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（修正係数）は、フィルタ係数の収束速度を制御するとともに、ウィナー解からの最終的な過剰平均 2 乗誤差を制御する。このため、収束の途中で収束速度を高速化し、定常状態において過剰平均 2 乗誤差を小さくするために、時変ステップサイズや行列形式の時変ステップサイズ [9] 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。一方、時不変ステップサイズを用いる従来の NLMS 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、これを越えることはできない。

さて、従来の適応アルゴリズムには、室内インパ

ルス応答の性質はほとんど利用されていない。音響エコー経路の変動特性に適した適応アルゴリズムが得られれば、収束特性を改善することが期待できる。そこで、まず多数の室内インパルス応答を実測し、インパルス応答の変動特性について検討し、その統計的性質を明らかにする。そして、その結果に基づいて、ES (exponentially weighted step size) アルゴリズムと名付けた新しい適応アルゴリズムを提案する。

以下、本報告は次のように構成されている。音響エコーキャンセラと従来の適応アルゴリズムの概要を第 2 章で述べる。新しい適応アルゴリズムを第 3 章で提案する。第 4 章では提案手法の諸特性について討論し、第 5 章では実用的な音響エコーキャンセラのための修正方法について説明する。実時間評価結果については第 6 章で述べる。第 7 章はまとめである。

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 $h(k)$ 、を推定する。ただし、 $h(k) = [h_1(k)$,

$h_2(k), \dots]^T$ 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。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動するために、 $\mathbf{h}(k)$ 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 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。ただし、 $\mathbf{x}(k)=[x(k), x(k-1), \dots, x(k-L+1)]^T$ 、 L はタップ数である。次に、擬似エコー $\hat{y}(k)$ は真のエコー $y(k)$ から差し引かれ、誤差 $e(k) = y(k) - \hat{y}(k) + n(k)$ が生成される。ただし、 $n(k)$ は周囲騒音である。近端話者の送話音声はマイクロホンに加わるダブルトーク状態では、通常、エコーキャンセラの適応動作を停止させる。従って、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 $\hat{\mathbf{h}}(k)$ は誤差のパワーを小さくするように、毎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。ただし、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。

2.2 従来の適応アルゴリズム

2.2.1 RLS アルゴリズム

RLS (recursive least-squares) アルゴリズム [1] はフィルタ係数ベクトル $\hat{\mathbf{h}}(k)$ を次式に従って修正する。

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \mathbf{k}(k)e(k) \quad (1)$$

$$\mathbf{k}(k) = \frac{\nu^{-1}\mathbf{P}(k)\mathbf{x}(k)}{R + \nu^{-1}\mathbf{x}(k)^T\mathbf{P}(k)\mathbf{x}(k)} \quad (2)$$

$$\mathbf{P}(k+1) = \nu^{-1}\mathbf{P}(k) - \nu^{-1}\mathbf{k}(k)\mathbf{x}(k)^T\mathbf{P}(k) \quad (3)$$

$$e(k) = y(k) - \hat{\mathbf{h}}(k)^T\mathbf{x}(k) + n(k) \quad (4)$$

ただし

$$\begin{aligned} \hat{\mathbf{h}}(k) &= [\hat{h}_1(k), \hat{h}_2(k), \dots, \hat{h}_L(k)]^T, \\ \hat{h}_i(k) (i &= 1, \dots, L): \text{FIR フィルタ係数}, \end{aligned}$$

L : タップ数,

$\mathbf{k}(k)$: L 次元ゲインベクトル,

$\mathbf{P}(k)$: $L \times L$ 行列,

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

ν : 忘却係数 ($0 < \nu < 1$),

R : $n(k)$ の分散.

$\mathbf{P}(k)$ は入力信号の共分散行列の逆行列として定義され、インパルス応答が時不変で $\mathbf{h}_0$ と表わされる場合には、係数誤差の自己相関行列

$$\mathbf{P}(k) = E[\{\mathbf{h}_0 - \hat{\mathbf{h}}(k)\}\{\mathbf{h}_0 - \hat{\mathbf{h}}(k)\}^T] \quad (5)$$

とも見なせる。ただし、 $E[\cdot]$ は期待値である。

RLS アルゴリズムの最も重要な性質は、収束速度が入力信号の性質に依存しないことである。しかし、演算量が多く、(1)–(4) 式の計算に $O(L^2)$ の積和を必要とする。演算量の削減をはかった高速 RLS アルゴリズムが近年検討されているが、 $7L$ 以上の積和を必要とする [2][3]。このため、サブバンドエコーキャンセラの試作機 [11] に使用されたことはあるが、商用の音響エコーキャンセラに導入されたことはない。

2.2.2 NLMS アルゴリズム

NLMS (normalized least-mean-squares) アルゴリズム [6] 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。

$$\hat{\mathbf{h}}(k+1) = \hat{\mathbf{h}}(k) + \alpha \frac{e(k)}{\|\mathbf{x}(k)\|^2} \mathbf{x}(k) \quad (6)$$

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

ただし、

α : スカラストップサイズ ($0 < \alpha < 2$),

$\|\cdot\|$: ユークリッドノルム.

フィルタ係数ベクトル $\hat{\mathbf{h}}(k)$ は、誤差 $e(k)$ と受話入力 $\mathbf{x}(k)$ の両方に比例して修正される。平均 2 乗誤差は $0 < \alpha < 2$ のとき収束し、白色雑音入力に対して $\alpha = 1$ のとき収束速度が最大となる。

NLMS アルゴリズムの最も重要な性質は、演算量が少ないことであり、(6)(7) 式の計算に $2L$ 回の積和しか必要としない。このため、NLMS アルゴリズムは、ほとんど全ての商用エコーキャンセラに用いられている。しかし、収束速度が入力信号の性質に依存することが欠点である。入力信号が音声である場合の収束速度は、白色雑音入力のとくに比べて、約 5 倍程度遅くなる。

NLMS アルゴリズムは RLS アルゴリズムを簡略化したものであるとも解釈できる。実際、(6) 式は (1)–(3) 式において $R = 0$, $\nu = 1$, $P(k) = I$ とし、ステップサイズ α を導入することによって導き出せる。ただし、 I は単位行列である。他の多数のアルゴリズムは、RLS アルゴリズムと NLMS アルゴリズムの間に位置付けることができる。言い替えれば、他の多数のアルゴリズムは、RLS アルゴリズムを簡略化したものであるが NLMS アルゴリズムほど簡単ではなく、その結果、収束速度は NLMS アルゴリズムより早い RLS アルゴリズムより遅く、また、演算量は NLMS アルゴリズムより多いが RLS アルゴリズムより少ない。

3 提案する適応アルゴリズム

3.1 室内インパルス応答の変動特性

音響エコーキャンセラに適した適応アルゴリズムを導くために、室内インパルス応答の変動特性、すなわち室内インパルス応答波形の変化量、について検討した。室内インパルス応答は、いろいろな要因により変化するが、ここでは、一例として、通信会議を模擬した 2 つの場合について検討する。

最初は、スピーカ・マイクロホン間距離が変化した場合である。残響時間が 350 ms の会議室において、スピーカ・マイクロホン間距離が 1 m (IR_1) と 0.6 m (IR_2) の場合のインパルス応答を測定した。これらのインパルス応答波形 IR_1 , IR_2 、およびインパルス応答の変化量 ($IR_2 - IR_1$) を図 2(a) に示す。また、 IR_1 , IR_2 , $IR_2 - IR_1$ の残響エネルギー減衰曲線 [12] を図 2(b) に示す。

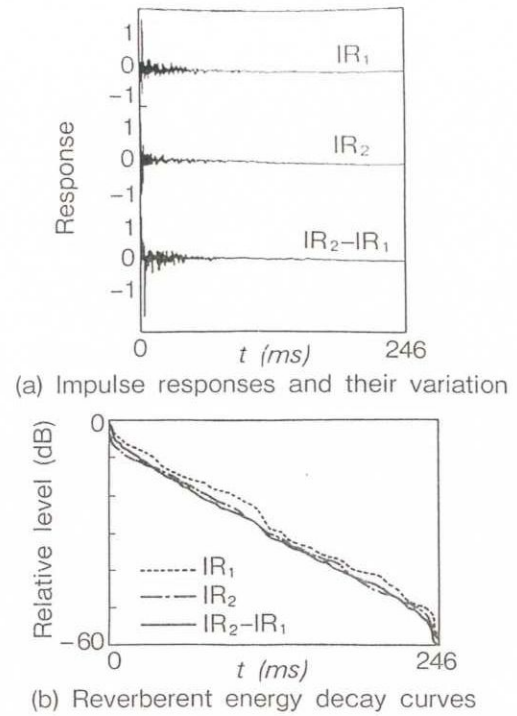


図 2: Variation of a room impulse response when the loudspeaker-microphone distance is changed. IR_1 and IR_2 are impulse responses for loudspeaker-microphone distances of 1 m and 0.6 m respectively. Room reverberation time at 500 Hz is 350 ms.

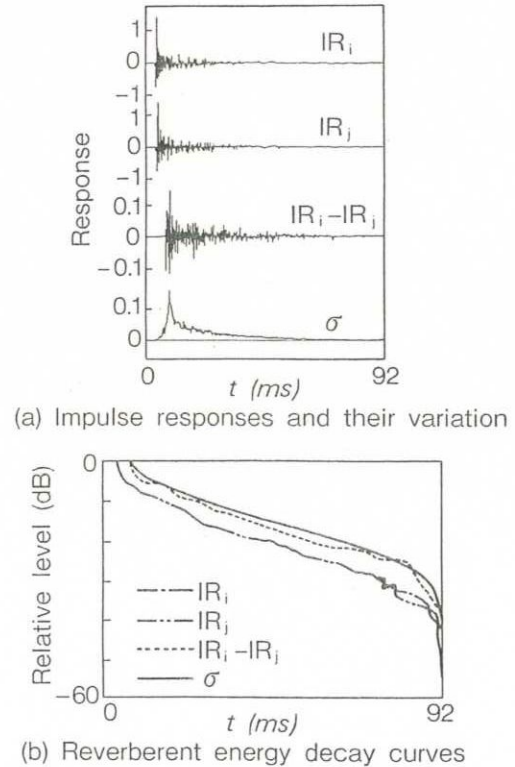


図 3: Variation of a room impulse response when the participants move. IR_i and IR_j are impulse responses. σ is the standard deviation of the variation ($IR_i - IR_j$) ($i, j = 1, \dots, 21$) of the impulse response. Loudspeaker-microphone distance is 1 m. Room reverberation time at 500 Hz is 280 ms.

次に、3人の会議参加者がマイクロホンの前で椅子に座ったまま上半身を動かした場合について調べた。21個のインパルス応答 $IR_i (i = 1, \dots, 21)$ を残響時間 280 ms の会議室内で測定した。スピーカ・マイクロホン間距離は 1 m である。これらのうちの2つのインパルス応答 $IR_i, IR_j (i \neq j)$ 、それらの変化量 $(IR_i - IR_j)$ 、変化量の標準偏差 σ を図 3(a) に示す。また、 $IR_i, IR_j, IR_i - IR_j, \sigma$ の残響エネルギー減衰曲線 [12] を図 3(b) に示す。

図 2, 3 から、インパルス応答は指数減衰し、これらのインパルス応答の変化量もまた同じ減衰率で指数減衰することが分かる。この指数減衰率は、同一室内のインパルス応答に対しては同じであり、部屋の寸法や吸音率など室内の音響条件によって決まる残響時間から求めることができる。従って指数減衰率は、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。

3.2 指数重み付け (ES) アルゴリズム

インパルス応答の変動量が指数減衰するため、係数誤差ベクトル (インパルス応答の推定誤差) 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。この性質を従来の NLMS アルゴリズムに取り入れて、誤差の大きな係数は大きなステップで、誤差の小さな係数は小さなステップ修正することを考える。このために、ステップサイズ行列 A (対角行列) を導入する。

$$A = \begin{pmatrix} \alpha_1 & & 0 \\ & \alpha_2 & \\ & & \ddots \\ 0 & & & \alpha_L \end{pmatrix} \quad (8)$$

ただし、

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

γ : 指数減衰率 ($0 < \gamma < 1$).

対角要素 α_i は時不変で、 α_1 から α_L までインパルス応答 $h(k)$ と同じ減衰率 γ で指数減衰する。ES (exponentially weighted step size) アルゴリ

ズムと名付けられた新しい適応アルゴリズムは、次のように表わされる [13]。

$$\hat{h}(k+1) = \hat{h}(k) + A \frac{e(k)}{\|x(k)\|^2} x(k) \quad (9)$$

$$e(k) = y(k) - \hat{h}(k)^T x(k) + n(k) \quad (10)$$

(6) 式におけるスカラーステップサイズ α は、(9) 式ではステップサイズ行列 A に置き換えられている。本手法は RLS アルゴリズムから、(2) 式分子の行列 $P(k)$ を定数 A に、分母の $P(k)$ を単位行列 I に、 ν を 1 に置き換え、 R を無視することにより導き出すこともできる。ベクトル $e(k) x(k) / \|x(k)\|^2$ の各要素が各 α_i に掛けられるため、本手法は NLMS アルゴリズムに比べて、 L 回の積が余分に必要となる。しかし、この問題は 5 章で述べる修正を施すことにより避けることができる。その結果、本手法は $2L$ 回の積和 (NLMS アルゴリズムと同等の演算量) で実現することができる。

指数的に重み付けた対角のステップサイズ行列は、適応ラインエンハンサへの適用を目的として文献 [14] でも提案されている。しかし、文献 [14] のアルゴリズムは、過去の入力信号の影響を軽減することを目的としたものであり、入力信号の変化の程度を反映させたものである。文献 [14] のアルゴリズムにおける指数減衰率は、入力信号がどの程度非定常であるかにより決定される。

4 提案手法の諸特性

この章では ES アルゴリズムの諸特性について述べる。エコーキャンセラは、時刻 $k = 0$ 以前にはインパルス応答を精度よく推定できていたものと仮定する。さらに、インパルス応答が時刻 $k = 0$ で変化したものと仮定する。すなわち、

$$h(k) = \begin{cases} h'_0 & (k < 0) \\ h_0 & (k \geq 0) \end{cases} \quad (11)$$

$$\hat{h}(0) \approx h'_0 \quad (12)$$

4.1 係数誤差の変化

係数誤差ベクトル $\mathbf{v}(k)$ を次のように定義する。

$$\mathbf{v}(k) = \mathbf{h}_0 - \hat{\mathbf{h}}(k) \quad (13)$$

(13) 式と (9) 式を組み合わせ、 $e(k) = \mathbf{v}(k)^T \mathbf{x}(k) + n(k)$ を用いれば

$$\begin{aligned} \mathbf{v}(k+1) = & \mathbf{v}(k) - \frac{\mathbf{v}(k)^T \mathbf{x}(k)}{\|\mathbf{x}(k)\|^2} \mathbf{A} \mathbf{x}(k) \\ & - \frac{n(k)}{\|\mathbf{x}(k)\|^2} \mathbf{A} \mathbf{x}(k) \end{aligned} \quad (14)$$

となる。ベクトル $\mathbf{v}(k)$ の第 i 番目の要素は

$$\begin{aligned} v_i(k+1) = & v_i(k) \\ & - \frac{\sum_{j=1}^L v_j(k) x(k-j+1)}{\sum_{j=1}^L x(k-j+1)^2} \alpha_i x(k-i+1) \\ & - \frac{n(k)}{\sum_{j=1}^L x(k-j+1)^2} \alpha_i x(k-i+1) \end{aligned} \quad (15)$$

となる。ここで、 $\mathbf{x}(k)$, $\mathbf{v}(k)$, $n(k)$ は平均値が0で互いに統計的に無相関な定常信号であると仮定する。また、ベクトル $\mathbf{x}(k)$ の要素は互いに無相関であると仮定する。すなわち、

$$1) E[x(k-i+1)x(k-j+1)] = \begin{cases} p_x & (i=j) \\ 0 & (i \neq j) \end{cases}$$

タップ数 L が十分に大きいと仮定すれば、次の仮定が成り立つ。

$$2) \|\mathbf{x}(k)\|^2 \approx L \cdot p_x \quad \text{for all } k,$$

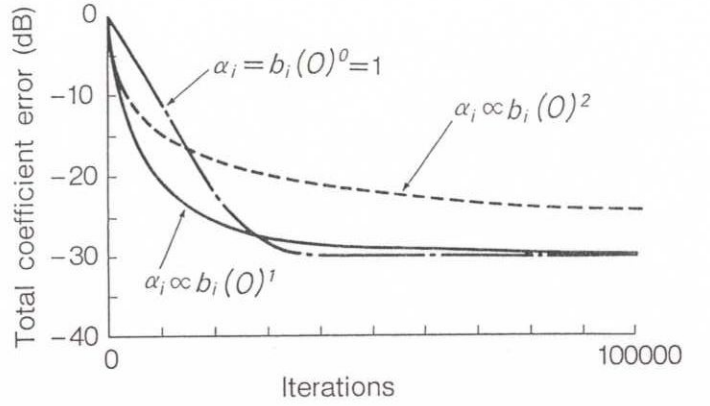


図 4: Convergence of the mean-squared coefficient error for various step sizes $(\alpha_1, \dots, \alpha_L)$. Here, $\alpha_i = b_i(0)^0 = 1$ corresponds to the conventional NLMS with $\alpha = 1$ and $b_i(0)^1$ represents the standard deviation of the room impulse response variation. The number of taps is 3840 and the ambient noise with a fixed SNR of 30 dB is added.

$$\begin{aligned} 3) \sum_{i=1}^L E[v_i(k)^2] E[x(k-i+1)^2 x(k-j+1)^2] \\ \approx p_x^2 \sum_{i=1}^L E[v_i(k)^2]. \end{aligned}$$

これらの仮定に基づき、 $\mathbf{x}(k)$, $\mathbf{v}(k)$, $n(k)$ に関して (15) 式の 2 乗期待値をとれば

$$\begin{aligned} E[v_i(k+1)^2] &= b_i(k+1)^2 \\ &= b_i(k)^2 - 2 \frac{\alpha_i}{L} b_i(k)^2 + \frac{\alpha_i^2}{L^2} \sum_{j=1}^L b_j(k)^2 \\ &\quad + \frac{\alpha_i^2 p_n}{L^2 p_x} \end{aligned} \quad (16)$$

となる。ただし、 $b_i(k)^2$, p_n は次のように定義されている。

$$\begin{aligned} b_i(k)^2 &= E[v_i(k)^2], \\ p_n &= E[n(k)^2]. \end{aligned}$$

(16) 式において、周囲騒音に起因する最後の項が無視できると仮定すれば、

$$\begin{pmatrix} b_1(k+1)^2 \\ b_2(k+1)^2 \\ \vdots \\ b_L(k+1)^2 \end{pmatrix} = \begin{pmatrix} (1 - \alpha_1/L)^2 & (\alpha_1/L)^2 & \cdots & (\alpha_1/L)^2 \\ (\alpha_2/L)^2 & (1 - \alpha_2/L)^2 & \cdots & (\alpha_2/L)^2 \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_L/L)^2 & (\alpha_L/L)^2 & \cdots & (1 - \alpha_L/L)^2 \end{pmatrix} \begin{pmatrix} b_1(k)^2 \\ b_2(k)^2 \\ \vdots \\ b_L(k)^2 \end{pmatrix} \quad (17)$$

あるいは行列形式で

$$\mathbf{b}(k+1) = \mathbf{Q} \mathbf{b}(k) \quad (18)$$

と書き表わせる。

(16)–(18) 式が ES アルゴリズムにおける係数誤差の 2 乗期待値の変化を表わしている。

4.2 α_i の与え方

係数誤差の 2 乗期待値の総和

$$E[\mathbf{v}(k+1)^T \mathbf{v}(k+1)] = \sum_{j=1}^L b_j(k+1)^2 \quad (19)$$

を最小にする最適ステップサイズ α_i は、(16) 式を α_i で偏微分して 0 と置くことにより求めることができる。周囲騒音が無視できる場合には、(16) 式の最後の項が無視できる。このとき、最適ステップサイズは

$$\alpha_i = \frac{L b_i(k)^2}{\sum_{j=1}^L b_j(k)^2} \quad (i = 1, \dots, L) \quad (20)$$

となる。上式は、最適ステップサイズは対応する係数の誤差の 2 乗期待値に比例して与えられることを示している。

$[b_1(k)^2, \dots, b_L(k)^2]$ は、時刻 $k = 0$ においてインパルス応答の変動量の指数減衰特性と一致しているが、係数が収束するに伴って変化する。このことは、全ての時間 k について $b_i(k)^2$ を推定し α_i を修正し直す必要があることを意味している。RLS アルゴリズムなどは、多くの演算量を費やして $b_i(k)^2$ を推定している。

ここでは、実用的な観点から演算量の増加を防止するために、時不変のステップサイズを考える。ES アルゴリズムにおける係数誤差の 2 乗期待値の収束過程を、いろいろな指数減衰率 r のステップサイズ $\alpha_i = b_i(0)^r$ について、(16) 式を用いて計算した。計算結果を図 4 に示す。タップ数は 3840 タップであり、SN 比 30 dB の周囲騒音を加えた。ただし、SN 比は周囲騒音のパワーに対する真のエコーのバ

ワーの比として定義される。ここで、 $\alpha_i = b_i(0)^0 = 1$ は、従来の NLMS アルゴリズムにおいてスラストステップサイズ $\alpha = 1$ に対応しており、このとき NLMS アルゴリズムの収束速度は最も速くなる。

α_i を $b_i(0)^2$ ($r = 2$) に比例して与えた場合には、(20) 式は時刻 $k = 0$ のときのみ満足される。その結果、収束速度は収束の初期においてのみ速く、その後遅くなる。これに対して、 α_i を $b_i(0)^0 = 1$ ($r = 0$) に比例して与えた場合には、白色雑音入力に対して定常状態における全ての係数誤差の 2 乗期待値が等しいため、(20) 式は定常状態において満足される。しかし、(20) 式は、収束の初期においては満足されない。その結果、収束初期の収束速度は遅い。これらの 2 つの間の場合として、 α_i を $b_i(0)^1$ ($r = 1$) に比例して与えた場合には、(20) 式は過渡状態を通じて近似的に満足される。実用的な音響エコーキャンセラにおいては、ERLE が約 20 dB に達するまでの収束時間が重要である。このため、時不変のステップサイズ α_i を $b_i(0)^1$ ($r=1$)、すなわち、インパルス応答の変動量の期待値に比例させて与える。

$$\alpha_i \propto b_i(0) \quad (i = 1, \dots, L) \quad (21)$$

インパルス応答の変動量の期待値 $b_i(0)$ は、インパルス応答と同じ減衰率で指数減衰し、 $i = T_R/T_S$ サンプルの時間で -60 dB 減衰する。ただし、 T_S はサンプリング間隔、 T_R は残響時間であり、音のエネルギーが -60 dB 減衰するのに要する時間として定義される。一方、 $\alpha_i = \alpha_0 \gamma^{i-1}$ ($i = 1, \dots, L$) であるから、指数減衰率 γ は次のように計算できる。

$$\gamma = \exp(-6.9 \frac{T_S}{T_R}) \quad (22)$$

ただし、-6.9 は $\log_e 10^{-3}$ を計算したものである。

実際のスピーカ・マイクロホン間の音響エコー経路のように、インパルス応答に既知の固定伝搬遅延がある場合には、対応する α_i は 0 とする。

4.3 定常エコー消去量

定常状態において、次式が成り立つものと仮定する。

$$b_i(k+1)^2 = b_i(k)^2 = b_i^2 \quad (23)$$

このとき、(16) 式は次のようになる。

$$b_i^2 = \frac{\alpha_i}{2L} \left(\sum_{j=1}^L b_j^2 \right) + \frac{\alpha_i p_n}{2L p_x} \quad (24)$$

(24) 式の b_i^2 をすべての i について足し合わせれば、

$$\sum_{i=1}^L b_i^2 = \frac{1}{2} \bar{\alpha} \sum_{j=1}^L b_j^2 + \frac{1}{2} \bar{\alpha} \frac{p_n}{p_x} \quad (25)$$

となる。ただし、

$$\bar{\alpha} = \frac{1}{L} \sum_{i=1}^L \alpha_i = \frac{\alpha_0}{L} \frac{1 - \gamma^L}{1 - \gamma} \quad (26)$$

である。従って、

$$\sum_{i=1}^L b_i^2 = \frac{\bar{\alpha}}{2 - \bar{\alpha}} \frac{p_n}{p_x} \quad (27)$$

となる。残留エコー $\tilde{e}(k) = y(k) - \hat{y}(k)$ の 2 乗期待値は

$$\begin{aligned} p_e &= E[\tilde{e}(k)^2] = E[\{\mathbf{v}(k)^T \mathbf{x}(k)\}^2] \\ &= \sum_{i=1}^L b_i^2 p_x = \frac{\bar{\alpha}}{2 - \bar{\alpha}} p_n \end{aligned} \quad (28)$$

となる。(28) 式の両辺を $p_y = E[y(k)^2]$ で割り対数をとれば、定常エコー消去量 $ERLE_\infty$ が次のように得られる。

$$ERLE_\infty = SNR + 10 \log_{10} \left(\frac{2}{\bar{\alpha}} - 1 \right) \text{ (dB)} \quad (29)$$

ただし、

$$\begin{aligned} ERLE_\infty &= 10 \log_{10}(p_y/p_e), \\ SNR &= 10 \log_{10}(p_y/p_n). \end{aligned}$$

である。

(29) 式は従来の NLMS アルゴリズムの定常エ

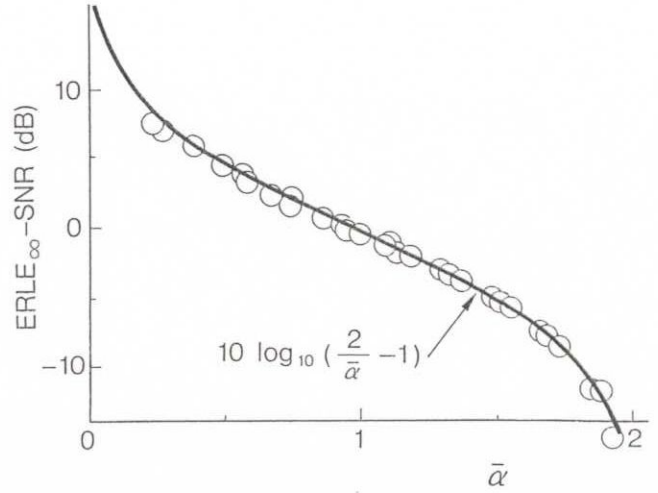


図 5: Steady-state echo return loss enhancement ($ERLE_\infty$). $\bar{\alpha}$ is the mean step size. The number of taps is 500 and the ambient noise with a fixed SNR of 30 dB is added.

コー消去量の式において、スカラーステップサイズ α を平均ステップサイズ $\bar{\alpha}$ で置き換えたものとなっている。

4.4 収束条件

(18) 式の行列 $\mathbf{Q}$ の固有値 λ_i が $|\lambda_i| < 1$ を満たせば、ES アルゴリズムにおいて係数誤差の 2 乗期待値は収束する。収束の必要十分条件は

$$0 < \bar{\alpha} = \frac{1}{L} \sum_{i=1}^L \alpha_i < 2 \quad (30)$$

で与えられる。証明を付録に記す。

4.5 計算機シミュレーション結果

4.1–4.4 章で述べた結果は、入力信号が白色雑音でタップ数 L が比較的大きい場合には、計算機シミュレーション結果とよく一致する。ここでは、一例として、いろいろな平均ステップサイズ $\bar{\alpha}$ に対する定常エコー消去量の計算機シミュレーション結果を図 5 に示す。タップ数は 500 タップであり、SN 比 30 dB の周囲騒音を加えた。(29) 式で与えられる理論曲線も同図に示してある。図 5 から、計算機シミュレーション結果は (29) 式の理論曲線とよく一致していることが分かる。

ES アルゴリズムでは、平均ステップサイズ $\bar{\alpha}$ が重要な役割を果たしており、収束条件を決定し、収束速度と過剰平均 2 乗誤差とのトレードオフを制御する。

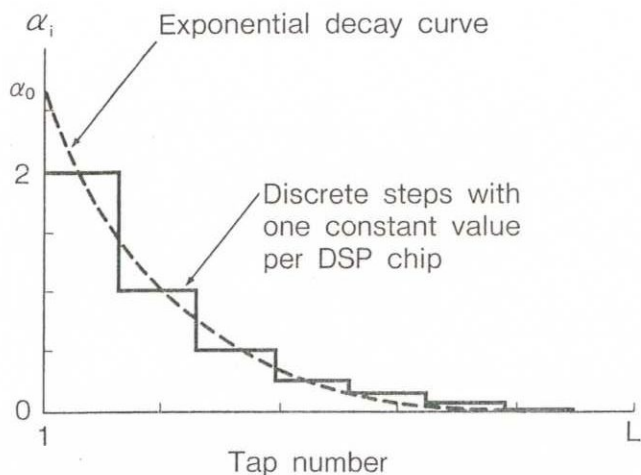


図 6: Step size α_i of matrix A when α_i is set in discrete steps with one constant value per DSP chip.

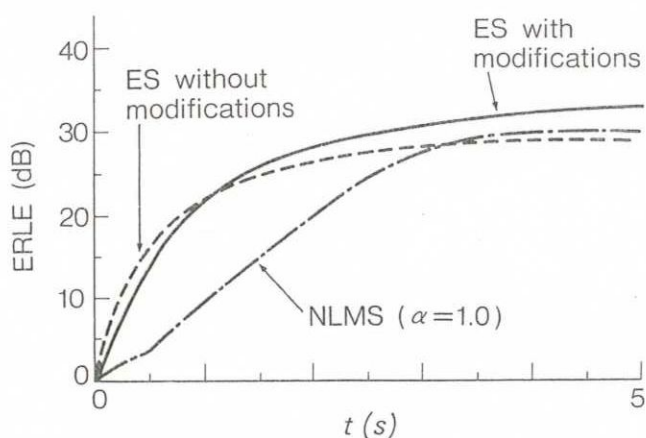


図 7: Convergence of the ERLE in the ES algorithm with and without the practical modifications. The number of taps is 3840 and sampling frequency is 8 k Hz. The input signal is white noise. Ambient noise with a fixed SNR of 30 dB is added.

5 実用的な修正

5.1 音声入力と複数 DSP 構成に対する実用的な修正

入力信号が音声であり、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、これまで述べてきた ES アルゴリズムに実用的な修正を行なう。

まず最初に、ステップサイズの最大値を制限する。収束条件 (30) 式は 4.1 章 1)–3) の仮定を満足する信号を仮定して導き出された。しかし、これら

の仮定が満足されない場合、すなわち、入力信号が音声のように非定常な場合には、収束条件 (30) 式は必要条件ではあるが十分条件ではなくなる。そこで、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、収束を保証するためにより強い十分条件

$$0 < \alpha_i < 2 \quad (i = 1, \dots, L) \quad (31)$$

を用いる。

次に、ES アルゴリズムの $2L$ 回の積和と L 回の積は $2L$ 回の積和に削減できる。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、指数減衰曲線を図 6 に示すように階段状に近似し、ステップサイズ α_i を各 DSP チップごとに一つの定数として与える。この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 アルゴリズムと同等 ($2L$) となる。

5.2 修正アルゴリズムの収束特性

実用的な修正の有無による収束特性の計算機シミュレーション結果を図 7 に示す。タップ数は 3840 タップ、サンプリング周波数は 8 kHz である。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 500 ms の室内で実測したものである。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を用い、SN 比 30 dB の周囲騒音を加えた。曲線は 10 回の試行の平均値である。エコー消去量の計算に用いたエコーと残留エコーの信号パワーは、100 サンプルの 2 乗値から求めた。実用的な修正を加えた場合の本手法の平均ステップサイズ $\bar{\alpha}$ は 0.45、修正を加えない場合は 1.0 である。従来の NLMS アルゴリズムのステップサイズ α は 1.0 である。

図 7 から、実用的な修正を加えても収束特性に大きな変化はなく、実用的な修正の有無によらず従来の NLMS アルゴリズムに比べて、収束速度が約 2 倍であることがわかる。(詳細に比較すれば、実用的な修正を加えると初期の収束速度は低下するが、定常エコー消去量は増加することが分かる。これは、収束速度と過剰平均 2 乗誤差のトレードオフの関係のためである。)

6 実時間評価実験

ES アルゴリズムを複数の DSP チップ [15] で構成した商用の音響エコーキャンセラ [8] に組み込んで、実時間評価実験を行なった。帯域分割手法を用いて 7 kHz 帯域を 2 分割し、それぞれの帯域で 8 kHz サンプルングした。表 1 に実験に使用したエコーキャンセラの諸元を示す。

本手法を実時間動作のためにさらに簡略化し、1 番目の DSP チップで $\alpha_{1-256} = 2.0$ 、2 番目の DSP チップで $\alpha_{257-512} = 1.0$ 、3 番目以降の DSP チップで $\alpha_{513-last} = 0.3$ と設定した。低域と高域の平均ステップサイズ $\bar{\alpha}$ は、それぞれ 0.46 と 0.64 である。スピーカ・マイクロホン間距離は 2.5 m、実験に使用した室内の残響時間は 300 ms である。

2 乗誤差レベルの収束曲線を図 8 に示す。入力信号として白色雑音を用いた場合 [図 8(a)] には、-20 dB の 2 乗誤差レベルに達する収束速度は、従来の NLMS アルゴリズム ($\alpha = 1.0$) の約 3 倍になることが分かる。入力信号として音声を用いた場合 [図 8(b)] には、-20 dB の 2 乗誤差レベルに達する収束速度は、従来の NLMS アルゴリズム ($\alpha = 1.0$) の約 2 倍になることが分かる。定常エコー消去量は 7 kHz 帯域にわたって 30 dB 以上得られている。

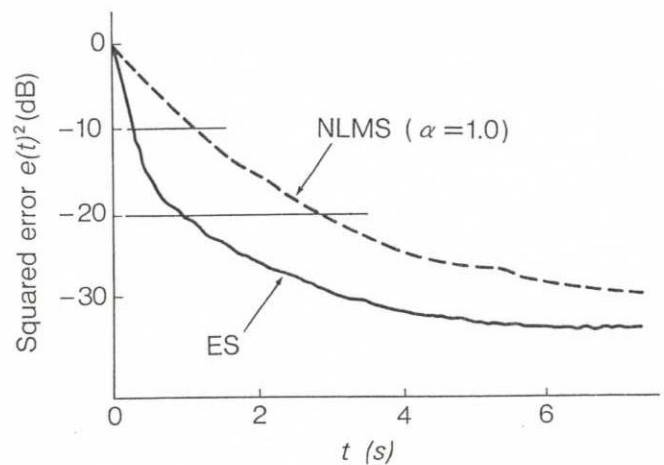
このように、ES アルゴリズムは容易に NLMS アルゴリズムに置き変わることができ、実際の音響エコーキャンセラの収束特性を改善することができる。

7 あとがき

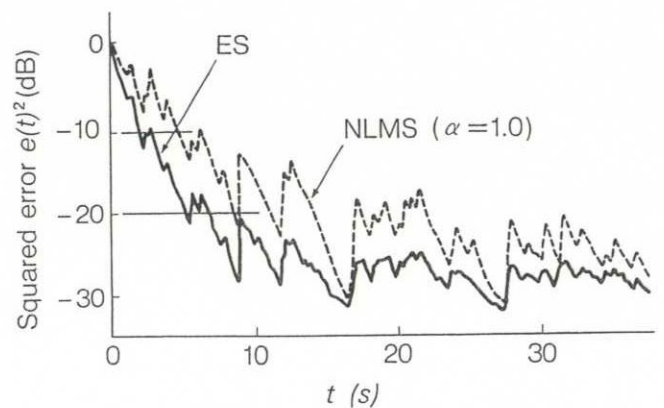
音響エコーキャンセラ用の新しい NLMS 適応アルゴリズムを提案した。本手法は、適応フィルタのそれぞれの係数を、異なる値のステップサイズで修正する。これらのステップサイズ α_i は時不変で、室内インパルス応答の変動量の期待値に比例して重み付けられている。実用的な複数 DSP 構成に対してステップサイズ α_i を修正することにより、従来の NLMS アルゴリズムと同等の演算量と記憶容量で本手法を実現することができる。本手法の収束過程を定式化することにより、係数誤差の 2 乗期待値の変

表 1: 実験に使用したエコーキャンセラの諸元

帯域	周波数範囲 (kHz)	タップ数
低域	0.15-4	3840
高域	4-7	1792



(a) Input signal: white noise



(b) Input signal: speech (male)

図 8: Real-time experimental results on squared error level convergence using a commercial acoustic echo canceller constructed with multiple DSP chips. The number of taps is 3840 in the lower band and 1792 in the higher band. Sampling frequency is 8 kHz in both bands. Room reverberation time at 500 Hz is 300 ms.

化を明らかにした。また、平均ステップサイズが収束条件、収束速度、過剰平均2乗誤差を制御することを明らかにした。本手法は、複数のDSPで構成された商用の音響エコーキャンセラに採用された。室内で行なった実時間評価実験の結果、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3倍、音声信号入力に対して約2倍の収束速度を有することを明らかにした。

付録 収束条件の証明

(30)式が成り立つとき、(18)式の行列 $\mathbf{Q}$ の固有値 λ_i は $|\lambda_i| < 1$ を満たすことは、次のように証明される。

$\mathbf{Q}$ は次のように書き表わせる。

$$\mathbf{Q} = \begin{pmatrix} (1-a_1)^2 & a_1^2 & \cdots & a_1^2 \\ a_2^2 & (1-a_2)^2 & \cdots & a_2^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_L^2 & a_L^2 & \cdots & (1-a_L)^2 \end{pmatrix} \quad (32)$$

ただし、

$$a_i = \alpha_i / L > 0$$

である。

次の行列式を考える。

$$\begin{aligned} |\mathbf{I} - \mathbf{Q}| &= \begin{vmatrix} 2a_1 - a_1^2 & -a_1^2 & \cdots & -a_1^2 \\ -a_2^2 & 2a_2 - a_2^2 & \cdots & -a_2^2 \\ \vdots & \vdots & \ddots & \vdots \\ -a_L^2 & -a_L^2 & \cdots & 2a_L - a_L^2 \end{vmatrix} \\ &= 2^{L-1} a_1 a_2 \cdots a_L (2 - \sum_{i=1}^L a_i) \end{aligned} \quad (33)$$

もし

$$2 - \sum_{i=1}^L a_i > 0 \quad (34)$$

であれば、 $(\mathbf{I} - \mathbf{Q})$ のprincipal minorがすべて正となる。従って $(\mathbf{I} - \mathbf{Q})$ は正定値、すなわち、 $(\mathbf{I} - \mathbf{Q})$ の固有値はすべて正となる。従って、 $\mathbf{Q}$ の固有値 λ_i は

$$\lambda_i < 1 \quad (i = 1, \dots, L) \quad (35)$$

となる。

一方、 $\mathbf{Q}$ は正行列、すなわち、 $\mathbf{Q}$ の要素はすべて正であるから、ペロン・フロベニウスの定理[16]により

$$-1 < \lambda_i < 1 \quad (36)$$

となる。

参考文献

- [1] S. Haykin, *Introduction to Adaptive Filters*. NY: Macmillan Publishing, 1984.
- [2] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*. Cambridge, MA: M.I.T. Press, 1985.
- [3] D. Slock and T. Kailath, "Numerically stable fast transversal filters for recursive least squares adaptive filtering," *IEEE Trans. Signal Processing*, vol. 39, no. 1, pp.92-114, Jan. 1991.
- [4] B. Widrow et al., "Adaptive noise cancelling: Principles and applications," *Proc. IEEE*, vol. 63, no. 12, pp.1692-1716, Dec. 1975.

- [5] B. Widrow and S. Stearns, *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1985.
- [6] J. Nagumo and A. Noda, "A learning method for system identification," *IEEE Trans. Automat. Contr.*, vol. AC-12, no. 3, pp.282-287, June 1967.
- [7] Y. Itoh, Y. Maruyama, N. Furuya, and T. Araseki, "An acoustic echo canceller for teleconference," *Proc. ICC85*, pp.1498-1502, June 1985.
- [8] H. Oikawa, S. Minami, and T. Saeki, "A new echo canceller realized by high performance digital signal processor," *Proc. ISCAS88*, pp.1329-1332, June 1988.
- [9] R. Harris, D. Chabries, and F. Bishop, "A variable step (VS) adaptive filter algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, no. 2, pp.309-316, Apr. 1986.
- [10] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.
- [11] B. Hatty, "Recursive least squares algorithms using multirate systems for cancellation of acoustic echoes," *Proc. ICASSP90*, pp.1145-1148, Apr. 1990.
- [12] M. R. Schroeder, "New method of measuring reverberation time," *J. Acoust. Soc. Amer.*, vol. 37, pp.409-412, 1965.
- [13] 牧野 昭二, 小泉 宣夫, "エコーキャンセラの室内音場における適応特性の改善について," 音講論集, pp.355-356 (1988.3).
- [14] J. Chen and J. Vandewalle, "A μ -vector LMS adaptive system for enhancing non-stationary narrow-band signal," *Proc. ISCAS88*, pp.771-774, June 1988.
- [15] T. Kaneko, H. Yamauchi, and A. Iwata, "A 50 ns floating-point signal processor VLSI," *Proc. ICASSP86*, pp.401-404, Apr. 1986.
- [16] F. R. Gantmacher, *The Theory of Matrices*. New York, NY: Chelsea Publishing, 1959.