

エコーキャンセラの室内音場における 適応特性改善について

Improvement on Adaptation of an Echo Canceller in a Room

牧野 昭二 小泉 宣夫

Shoji MAKINO Nobuo KOIZUMI

NTT ヒューマンインタフェース研究所

NTT Human Interface Laboratories

あらまし 室内音場のインパルス応答の変動特性について検討し、これらの特性に共通に見られる性質を明らかにした。これをもとに、こう配形適応アルゴリズムにおける逐次修正ベクトルの大きさと方向に修正を加えることにより、エコーキャンセラの適応特性を改善する方法を提案した。

LMS法、学習同定法、アフィン射影法などの従来のこう配形適応アルゴリズムは、入力信号系列および推定誤差の瞬時値から平均自乗誤差曲面のグラジエントを推定し最適解に向かって降下してゆく。これに対して、本手法は室内音場のインパルス応答が変動した場合の最適解と推定値の関係を利用し、従来グラジエント方向に向いていた逐次修正ベクトルの大きさと方向に修正を加えたものである。

本手法の有効性を室内で実測した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により検討した。その結果、従来の学習同定法においてステップゲイン $\alpha = 1$ とした場合に比べて、同等の定常エコー消去量に達する収束速度を白色信号に対して約3倍、実音声に対して約2倍にできることを明かにした。また、本手法は階段近似することにより従来法と同等の演算量と記憶容量で実現できることを明らかにした。

1. まえがき

近年、通信会議用拡声装置におけるハウリングやエコーの防止手段として、音響エコーキャンセラがさかんに検討されている。音響エコーキャンセラは、スピーカ、マイクロホン、音場からなる音響信号の伝達経路（音響エコー経路）のインパルス応答を模擬するデジタルフィルタの出力を、真の音響エコー経路の出力から差引きこれを消去するものである⁽¹⁾。音響エコーキャンセラの構成として様々なものが提案されているが⁽²⁾⁽³⁾、現時点において実時間動作が可能であるのは、適応形FIR(Finite Impulse Response)フィルタを使用しLMS法⁽⁴⁾、学習同定法⁽⁵⁾あるいはアフィン射影法⁽⁶⁾に基づきフィルタ係数を逐次修正するものに限られる⁽⁷⁾⁽⁸⁾⁽⁹⁾。これらのアルゴリズムは、真の音響エコー経路を介した信号と模擬音響エコー経路を介した信号の推定誤差を最小とするように、模擬音響エコー経路のインパルス応答を実時間で修正するための簡略化されたこう配形適応手法である。

音響エコーキャンセラが適用される音場の特性は人の移動などにより複雑に変動し、これにより音響エコーキャンセラのエコー消去量は大きく低下する⁽¹⁰⁾。⁽¹¹⁾⁽¹²⁾。従来の室内音場の変動モデルは一様分布す

る状態量（インパルス応答）に加法性雑音加わるものとして表わされ、このモデルを基に各種の適応アルゴリズムが提案されてきた。ところが実際の室内音場の変動モデルはこのようなものとは限らない。従来の適応アルゴリズムに利用されていなかった室内音場の変動特性についての知見を取り入れることにより音場の変動特性に合致した適応アルゴリズムが得られれば、適応特性の改善が期待できる。そこで、まず室内音場のインパルス応答の変動特性について検討した。さらにこう配形適応アルゴリズムを代表して学習同定法を取り上げ、逐次修正ベクトルの大きさと方向に上記変動特性を考慮した修正を加えることにより、適応特性の改善を試みた。このようなベクトル方向の修正を各種の考え方に基づいて行なう試みは、筆者らをはじめとして近年報告がなされはじめている⁽¹³⁾⁽¹⁴⁾⁽¹⁵⁾⁽¹⁶⁾。これらの中で本手法の特徴は、室内音場のインパルス応答の指数減衰特性を用いて逐次修正ベクトルの大きさと方向を修正している点、および従来法と同等の演算量と記憶容量で従来法の2倍～3倍の収束速度が得られる点にある。ここでは、本手法の特徴と適応特性の改善効果について述べる。

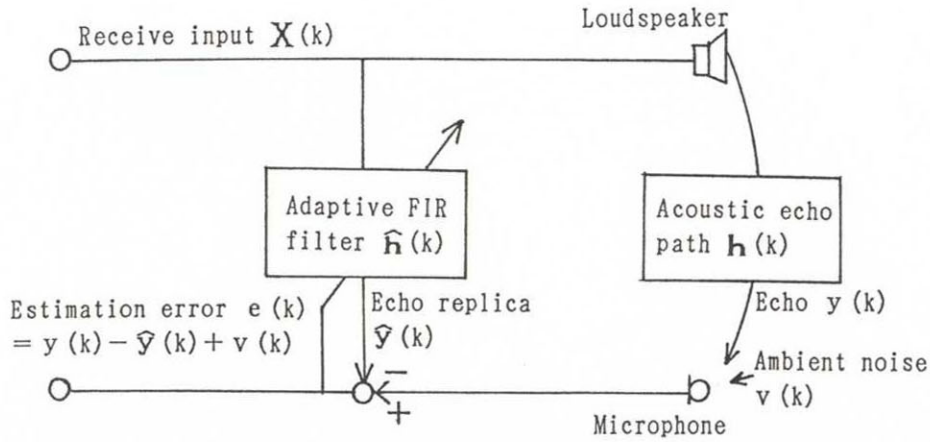


図1 音響エコーキャンセラの構成
Fig.1 Configuration of an acoustic echo canceller.

2. 音響エコーキャンセラの構成

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。以下すべての変数をサンプル値表現し、サンプル時間間隔 T の整数 k 倍における関数値を $f(k)$ と表す。音響エコーキャンセラは、受話入力信号 $X(k)$ と受話入力信号が音響エコー経路を介した回り込み信号 $y(k)$ とから、音響エコー経路のインパルス応答 $h(k)$ を推定し、推定されたインパルス応答 $\hat{h}(k)$ と受話入力信号 $X(k)$ から模擬回り込み信号 $\hat{y}(k)$ を生成し、真の回り込み信号 $y(k)$ から差し引きこれを消去するものである。スピーカ、マイクロホン、音場からなる音響エコー経路は、人の移動や周囲の環境によって変化するため、FIRフィルタは適応形のもが使われ、適応アルゴリズムには実時間動作・高速・高精度が要求される。

2. 2 エコー消去量

音響エコーキャンセラの性能を表す量としてエコー消去量 (ERLE; Echo Return Loss Enhancement) を用いる。

$$ERLE = 10 \log_{10} \frac{y(k) \text{ の電力}}{e^*(k) \text{ の電力}} \quad (\text{dB})$$

ここで $e^*(k)$ は $e(k)$ から図1に示す雑音 $v(k)$ を除いた値である。回り込み信号 $y(k)$ と雑音 $v(k)$ との無相関性を仮定すれば、学習同定法⁽⁵⁾を用いた場合の定常エコー消去量 (ERLE $_{\infty}$) は、回り込み信号 $y(k)$ に加わる雑音 $v(k)$ によって劣下し⁽¹⁷⁾、次式のように表わされる。

$$ERLE_{\infty} = S - N + 10 \log_{10} (2 / \alpha - 1) \quad (\text{dB})$$

S, N : エコー信号 $y(k)$ および雑音 $v(k)$ のレベル
 α : ステップゲイン (スカラー量, $0 < \alpha < 2$)

3. インパルス応答の変動と適応アルゴリズム

3. 1 インパルス応答の変動

室内音場のインパルス応答は様々な理由により変動すると考えられるが、ここでは簡単のため、スピーカ・マイク距離が変化した場合について考える。会議室 (残響時間 (500Hz) 160ms) にスピーカとマイクロホンを設置し、その距離を変えて測定した2つのインパルス応答 IR_1 (スピーカ・マイク 0.6m), IR_2 (スピーカ・マイク 1m) とこれらの差 ($IR_2 - IR_1$) を図2(a)に、また、インパルス応答を自乗積分することによってエネルギー減衰波形の集合平均を求めて算出⁽¹⁸⁾した IR_1 , IR_2 , $IR_2 - IR_1$ のエネルギー減衰波形を図2(b)に示す。図2から音場のインパルス応答 IR_1 , IR_2 は指数減衰し、これらのインパルス応答の変動量と考えられる両者の差 ($IR_2 - IR_1$) もまた同じ減衰率で指数減衰することが分かる。そこで、音場のインパルス応答 $h(k)$ の変動を次のように表す。

$$h(k+1) = h(k) + A \xi(k) \quad (1)$$

ただし、

$$h(k) = (h_1(k), h_2(k), \dots, h_N(k))^T$$

$$A = \text{diag}[a_1, a_2, \dots, a_N]$$

$$a_i = \exp(-6.9(i-1)T_s/T_R)$$

$$(i = 1, 2, \dots, N)$$

k : 離散化時刻 T_s : サンプリング周期

N : タップ数 T_R : 残響時間

T : ベクトルの転置

$$\xi(k) = (\xi_1(k), \xi_2(k), \dots, \xi_N(k))^T$$

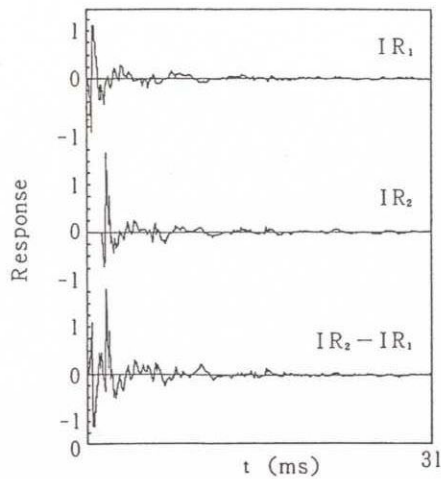
$\xi_1(k), \xi_2(k), \dots, \xi_N(k)$: 平均0、分散 σ^2 の

不規則変動量

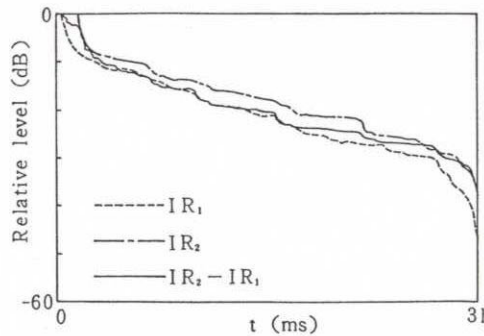
ここで、 a_i は $(j-1)T_s = T_R$ となる j に対して

$$20 \log_{10} (a_j / a_1) = -60 \quad (\text{dB})$$

となるように与えられている。



(a) Impulse response



(b) Reverberent decay curve

図2 インパルス応答とエネルギー減衰波形
IR₁, IR₂: インパルス応答

Fig.2 Impulse response and reverberent decay curve. IR₁ and IR₂ are impulse responses for loudspeaker-microphone distance 0.6m and 1m respectively.
Reverberation time (500Hz) 160ms.

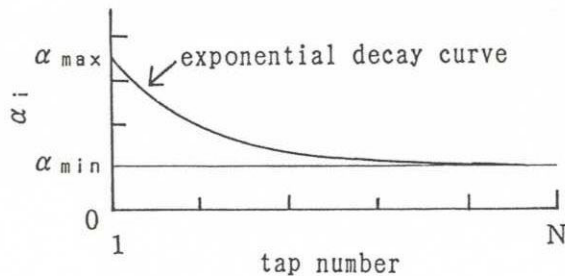


図3 ステップゲイン行列 α の対角成分 α_i ((3)式)
Fig.3 Diagonal component α_i in Eq.(3) of step gain matrix α .

3. 2 適応アルゴリズム

未知の系の伝達特性を同定するためのアルゴリズムは制御理論の分野で種々提案されているが^{(19),(20)}、ここではこう配形適応アルゴリズムを代表して、アルゴリズムが簡単で比較的良好な消去特性が得られる学習同定法⁽⁵⁾を用いる。室内音場のインパルス応答 h (

k)を模擬するエコーキャンセラのタップ係数 $\hat{h}(k)$ は、以下の(2)式に従って推定誤差 $e(k)$ と相手話者からの受話入力 $X(k)$ の大きさに比例して逐次修正される。幾何学的に言えば、推定誤差 $e(k)$ と受話入力 $X(k)$ から平均自乗誤差曲面のグラジエントを推定し最適解に向かって降下してゆく。

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

ただし、

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

α : ステップゲイン (スカラー量, $0 < \alpha < 2$)

$e(k)$: 推定誤差 ($= y(k) - \hat{y}(k)$)

$$y(k) = h(k)^T \cdot X(k)$$

$$\hat{y}(k) = \hat{h}(k)^T \cdot X(k)$$

$\|X(k)\|$: $X(k)$ のノルム

ここで、(2)式における $h(k)$ の形状はいかなるものでもよい。

さて3. 1で明らかにされた、室内音場のインパルス応答が指数減衰しその変動量もまた同じ減衰率で指数減衰するという点を考慮することにより、適応特性を改善することを試みる。 k 時点におけるエコーキャンセラのタップ係数 $\hat{h}(k)$ が k 時点における室内音場の真のインパルス応答 $h(k)$ に収束しており($\hat{h}(k) \approx h(k)$)、 $k+1$ 時点で真のインパルス応答が変動して $h(k+1)$ となったとすると、最適解 $h(k+1)$ と推定値 $\hat{h}(k)$ の関係は(1)式から

$$h(k+1) - \hat{h}(k) \approx h(k+1) - h(k) = A \xi(k)$$

となり、各係数の誤差の大きさは室内音場のインパルス応答と同じ減衰率で指数減衰する。そこで、各係数の誤差の大きさに応じて、誤差の大きい係数は大きなステップで、誤差の小さい係数は小さなステップで修正することを考える。そのために、従来の学習同定法ではスカラー量として与えられていたステップゲイン α を対角行列に拡張したステップゲイン行列 α

$$\alpha = \text{diag}[\alpha_1, \alpha_2, \dots, \alpha_N] \quad (3)$$

ただし、

$$\alpha_i = (\alpha_{\max} - \alpha_{\min}) \exp(-6.9(i-1) \times T_s / T_R) + \alpha_{\min} \quad (i = 1, 2, \dots, N)$$

を新たに導入し、本手法における逐次修正式を

$$\hat{h}(k+1) = \hat{h}(k) + \alpha \frac{e(k)}{\|X(k)\|^2} X(k) \quad (4)$$

とする。ステップゲイン行列 α の対角成分 α_i を図3に示す。

このように従来のステップゲイン α （スカラー量）を対角行列化して逐次修正ベクトルに乗ずることは、従来グラジエント方向に向いていた逐次修正ベクトルの大きさと方向に修正を加えることを意味する。

4. 計算機シミュレーション

室内音場のインパルス応答 h が図2に示した IR_1 から IR_2 に変動した場合の収束特性の計算機シミュレーション結果を図4に示す。タップ数は500タップ、サンプリング周波数は8kHz、送話入力点でのS/N比は30dBであり、受話入力には音声の自己相関の特徴

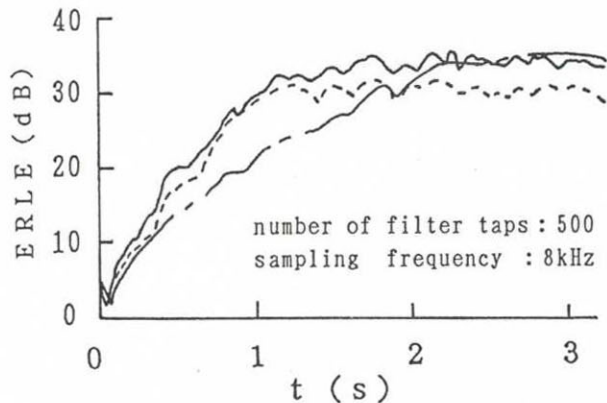
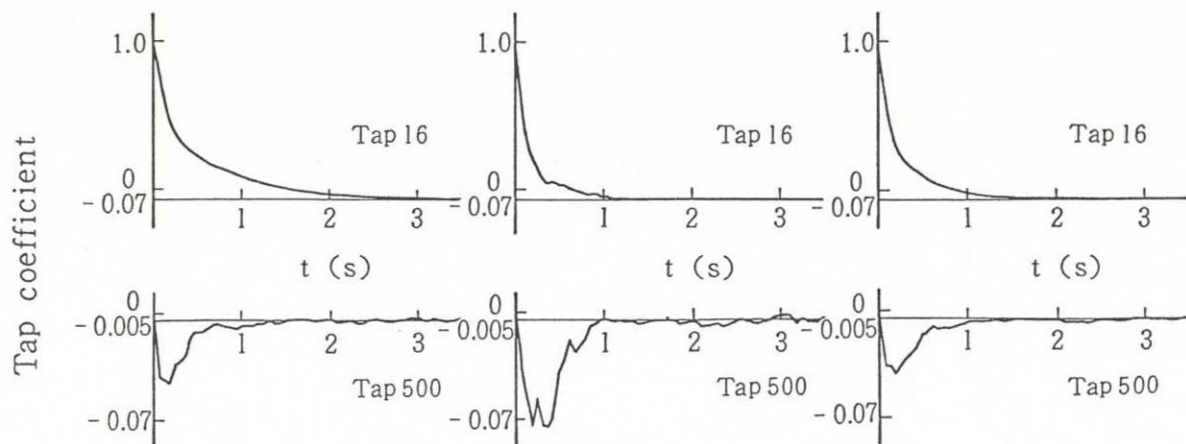


図4 本手法と従来法の収束特性の比較
Fig.4 Convergence performance. $\alpha=1.0$ (---) and $\alpha=0.5$ (...) in learning identification algorithm. Step gain matrix with $\alpha_{\max}=1.0$ and $\alpha_{\min}=0.5$ (—) in proposed method.

を考慮して、白色信号に一次の再帰フィルタ（フィルタ係数0.8）をかけた信号を用いた。従来法では $\alpha=1.0$ で収束速度が最大であり、 α が1.0より小さくなるに従い収束速度が遅くなり定常エコー消去量が大きくなることが知られている⁽¹⁷⁾。図4より、本手法において(3)式の $\alpha_{\max}=1.0$ 、 $\alpha_{\min}=0.5$ としたステップゲイン行列 α を用いた場合には、収束速度は従来法において $\alpha=1.0$ とした場合を上回り、定常エコー消去量は従来法において $\alpha=0.5$ とした場合と同程度となることがわかる。また、従来法と本手法を用いた場合のタップ係数の収束過程の一例として、変動の大きなエコーキャンセラの16番目のタップおよび変動の小さな500番目のタップの学習曲線を図5に示す。本手法において(3)式の $\alpha_{\max}=1.0$ 、 $\alpha_{\min}=0.5$ としたステップゲイン行列 α を用いた場合には、変動の大きいインパルス応答の初期の係数に対しては従来法において $\alpha=1.0$ とした場合と同程度に収束が早く（図5(b)(c)上図）、変動の小さい後期の係数に対しては従来法において $\alpha=0.5$ とした場合と同程度に誤調整が少ない（図5(a)(c)下図）ことが分かる。

次に、(3)式のステップゲイン行列 α において $\alpha_{\min}=1.0$ の時 $1.0 \leq \alpha_{\max} \leq 6.0$ 、 $\alpha_{\min}=0.5$ の時 $0.5 \leq \alpha_{\max} \leq 8.0$ 、 $\alpha_{\min}=0.25$ の時 $0.25 \leq \alpha_{\max} \leq 9.0$ の範囲で $\alpha_{\min}$ 、 $\alpha_{\max}$ を変化させた場合の定常エコー消去量（dB）の計算機シミュレーション結果を図6に示す。図6から定常エコー消去量は(3)式の α_i の相加平均値



(a) Learning Identification method ($\alpha=0.5$ in Eq.(2)) (b) Learning Identification method ($\alpha=1.0$ in Eq.(2)) (c) Proposed method (step gain matrix α where $\alpha_{\max}=1.0$, $\alpha_{\min}=0.5$ in Eq.(3))

図5 タップ係数の収束過程
Fig.5 Convergence process of tap coefficients.

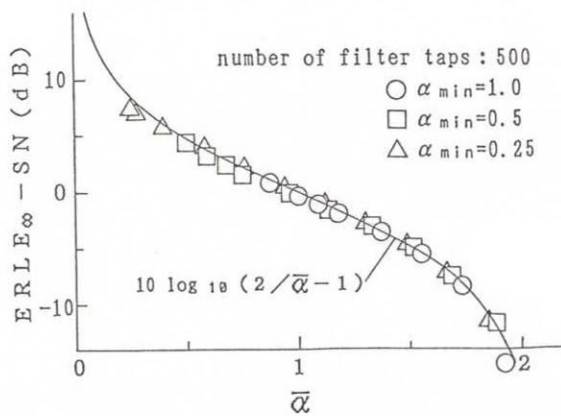


図6 定常エコー消去特性
Fig.6 Echo return loss enhancement at steady state.

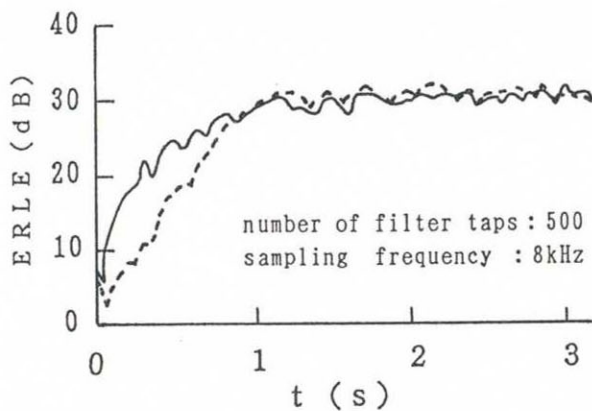


図7 本手法と従来法の収束特性の比較
Fig.7 Convergence performance. Step gain matrix with $\alpha_{\max} = 5.46$ and $\alpha_{\min} = 0.0$ (—) in proposed method. $\alpha = 1.0$ (---) in learning identification algorithm.

$$\bar{\alpha} = \sum_{i=1}^N \alpha_i$$

を用いて

$$\text{ERLE}_{\infty} = S - N + 10 \log_{10}(2 / \bar{\alpha} - 1) \text{ (dB)}$$

と表わせることがわかる。また、収束条件は

$$0 < \bar{\alpha} < 2$$

であり、 $\bar{\alpha} = 1$ のときに収束速度が最大となる。

本手法において(3)式の $\alpha_{\max} = 5.46$, $\alpha_{\min} = 0.0$ ($\bar{\alpha} = 1.0$) としたステップゲイン行列 α を用いた場合の収束特性の計算機シミュレーション結果を図7に示す。定常エコー消去量は従来法において $\alpha = 1.0$ とした場合と等しく、10dBの収束速度は約4倍、20dBの収束速度は約2倍になることがわかる。

以上述べてきた本手法の演算量は $O(3N)$ であるが、Digital Signal Processor などを用いてハードウェアを構成する場合には(3)式の α_i を図8に示すようにチ

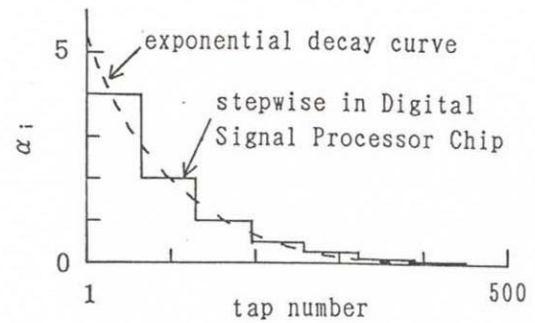


図8 ステップゲイン行列 α の対角成分 α_i ((3)式) Digital Signal Processor 毎に階段状に与える場合
Fig.8 Diagonal component α_i in Eq.(3) of step gain matrix α in the case of α_i changes stepwise in Digital Signal Processor Chip.

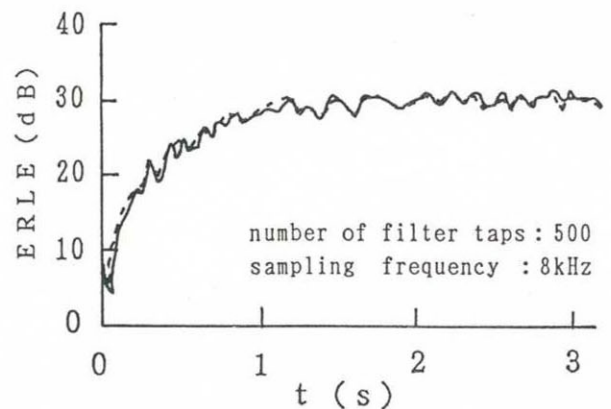


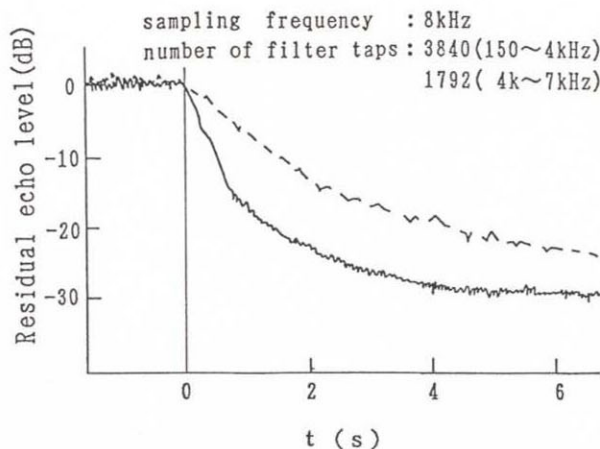
図9 α_i ((3)式)を階段近似した場合の収束特性
Fig.9 Convergence performance. Step gain matrix with $\alpha_{\max} = 5.46$ and $\alpha_{\min} = 0.0$ (—) and α_i changes stepwise in Digital Signal Processor Chip from $\alpha_{\max} = 4.0$ and $\alpha_{\min} = 0.0$ (---) in proposed method. Both $\bar{\alpha} = 1.0$.

ップ毎に階段状に与えることにより演算量を $O(2N)$ に減らすことができる。 α_i を連続的に与えた場合とチップ毎に階段状に与えた場合の収束特性の計算機シミュレーション結果を図9に示す。図9から α_i を階段近似しても連続的に与えた場合と同程度の特性が得られることがわかる。

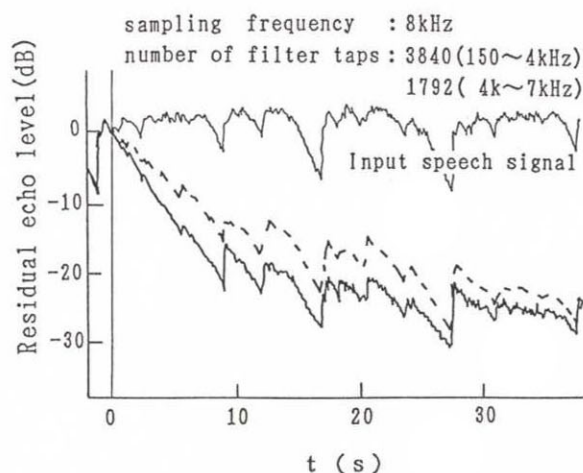
以上の結果より、本手法を階段近似することにより従来法と同等の演算量と記憶容量で、従来法に比べて収束速度を2倍以上に改善できることが明かとなった。

5. 実時間評価実験

本手法を実時間動作のためにさらに簡略化し1~256タップを担当する先頭のDSPチップで $\alpha_i = 2.5$, 257~512タップを担当する2番目のDSPチップで $\alpha_i = 1.5$, 513タップ以降を担当する3番目以降のDSPチ



(a) White noise input



(b) Speech (male) input

図10 本手法と従来法の収束特性の
実時間評価実験結果

Fig.10 Real time experimental results on convergence performance. Step gain matrix with $\alpha_{1-256}=2.5$, $\alpha_{257-512}=1.5$, $\alpha_{513-1792,3840}=0.3$ (—) and $\alpha=1.0$ (---) in learning identification algorithm. Reverberation time (500Hz) 300ms

ップで $\alpha_i = 0.3$ として会議室内（残響時間（500 Hz）300 ms）で収束特性の実時間評価実験を行なった。実験に使用した音響エコーキャンセラは、7 kHz帯域を2分割し、0～4 kHzを3840タップ（エコー消去時間 480 ms）、4～7 kHzを1792タップ（エコー消去時間 224 ms）とした。本手法および従来法において $\alpha = 1.0$ とした場合の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10(a)(b)に示す。受話入力には白色信号および実音声（男声）を用いた。白色信号を用いた場合（図10(a)）には10dB～20dBの収束速度が約3倍になることがわかる。また、実音声（男声）を用いた場合（図10(b)）には

10dB～20dBの収束速度が約2倍になり、本手法の有効性が実時間動作によっても確認された。

6. むすび

室内音場のインパルス応答の変動特性について検討し、音場の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動量と考えられる両者の差もまた同じ減衰率で指数減衰することを示した。これをもとに、こう配形適応アルゴリズムを代表して学習同定法を取り上げ、逐次修正ベクトルの大きさと方向に修正を加えることによりエコーキャンセラの適応特性を改善する方法を提案した。

本手法の有効性を室内で実測した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により検討した。その結果、従来の学習同定法においてステップゲイン $\alpha = 1$ とした場合に比べて、同等の定常エコー消去量に達する収束速度を白色信号に対して約3倍、実音声に対して約2倍にできることを明らかにした。また、本手法は階段近似することにより従来法と同等の演算量と記憶容量で実現できることを明らかにした。

謝 辞

貴重な御助言を頂いた当所金田主幹研究員に感謝いたします。

文 献

- (1) C. W. K. Gritton, D. W. Lin: "Echo Cancellation Algorithms", IEEE ASSP MAGAZINE, APRIL, pp.30-38 (1984).
- (2) 超音輝, 辻井重男: "ダブルトーク期間において制御可能な3端子対エコーキャンセラの提案", 信学論(A), J 71-A, 2, pp.379-386 (1988).
- (3) 谷萩隆嗣, 齊藤成利: "インパルス応答の分割推定に基づいた並列形カルマンフィルタによるエコーキャンセラ", 信学論(A), J 69-A, 1, pp.138-149 (1986).
- (4) B. Widrow, et al.: "Adaptive Noise Cancelling: Principles and Applications", Proc. IEEE, 63, 12, pp.1692-1716 (1975).
- (5) 野田淳彦, 南雲仁一: "システムの学習的同定法", 計測と制御, 7, 9, pp.597-605 (1968).
- (6) 尾関和彦, 梅田哲夫: "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), J 69-A, 10, pp.126-132 (1984).

- (7) 安川博, 古川功, 石山康蔵: “帯域分割構成による相関除去形音響エコーキャンセラの特性”, 信学技報, C A S 8 7 - 2 5 9 (1987).
- (8) H. OIKAWA, S. MINAMI, T. SAEKI: “A New Echo Canceller realized by High Performance Digital Signal Processor”, ISCAS, PP.1329-1332 (1988).
- (9) 及川弘, 小泉宣夫, 牧野昭二: “複数反響路エコーキャンセラを用いた音声会議装置”, 通研実報 37, PP.191-197 (1988).
- (10) 牧野昭二, 小泉宣夫: “直方体ブース内の障害物によるインパルス応答の変動について”, 音講論, p.295 (1987-3).
- (11) 牧野昭二, 小泉宣夫: “音響エコーキャンセラの室内環境における消去特性について”, 信学技報, E A 8 7 - 4 3 (1987-8).
- (12) 金千徳, 安倍正人, 城戸健一: “波形打消しによるハウリング防止に関する二三の検討”, 信学技報, E A 8 3 - 4 (1983).
- (13) 牧野, 小泉, “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, p.355, (1988, 3).
- (14) 牧野, 小泉, “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”, 信学論(A), J 7 1 - A, 1 2, pp.2212-2214 (1988).
- (15) 丸山, “重み付け適応制御を用いたエコーキャンセラの一検討”, 信学全大, B-557, (1989, 3).
- (16) R.W.Harris, D.M.Chabries, F.A.Bishop: “A Variable Step(VS) Adaptive Filter Algorithm”, IEEE Trans. ASSP-34, 2, pp.309-316 (1986).
- (17) 板倉秀清, 西川偉一: “学習同定法を用いたエコーキャンセラのエコー消去特性について”, 信学論(A), J 6 0 - A, 1 1, pp.1015-1022 (1977).
- (18) M.R.Schoeder: “New Method of Measuring Reverberation Time”, J. Acoust. Soc. Am., 37, pp.409-412 (1965).
- (19) 鈴木: “システム・アイデンティフィケーション-I”, 制御工学, 14, 7, pp.423~432 (1970).
- (20) 鈴木: “システム・アイデンティフィケーション-II”, 制御工学, 14, 8, pp.483~494 (1970).