

ES射影アルゴリズムを用いた デュオフィルタ構成のエコーキャンセラの検討

◎羽田 陽一、 牧野 昭二、 小島 順治、 島内 末廣

NTT ヒューマンインタフェース研究所

1 はじめに

TV会議や音声会議などの拡声通話系においては、適応フィルタを用いた音響エコーキャンセラが不可欠である。

音響エコーキャンセラを実環境で安定に動作させるためには、(1) 極く短い時間でダイナミックレンジが変化する音声入力信号に対し、適応フィルタ係数の収束が安定であること、(2) 受話信号と近端話者信号の両方が存在するダブルトーク時に対しても適応フィルタが安定であること、(3) 適応フィルタが音響系を同定していない状態においてもハウリングが起こらないこと、等が重要となる。

本報告では、適応フィルタとしてES射影アルゴリズム[1]を用いたFG/BG方式[2, 3]のエコーキャンセラ（我々はこれをデュオフィルタと呼ぶ）について検討を行なったので報告する。

2 ES射影アルゴリズムの改良

ES射影アルゴリズムは、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を活用し、かつ、音声のような有色信号の白色化を行なうものである。この手法では、NLMSアルゴリズムとほぼ同等の演算量と記憶容量で、音声入力に対する収束速度を約4倍にできる。

エコーキャンセラでは、入力信号が非定常な音声信号であるため、近端話者信号が存在しない状態でも、入力信号が小さくなると適応フィルタの収束が劣化する場合がある。

このような状況に対応するように改良したES射影アルゴリズムを以下に示す。

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

$$\beta_1(k) = \frac{e(k)r_{11} - (1-\mu)e(k-1)r_{10}}{r_{00}r_{11} - r_{10}r_{10} + \delta} \quad (2)$$

$$\beta_2(k) = \frac{(1-\mu)e(k-1)r_{00} - e(k)r_{10}}{r_{00}r_{11} - r_{10}r_{10} + \delta} \quad (3)$$

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

$$\hat{y}(k) = \mathbf{z}(k)^T \mathbf{x}(k) + \mu \beta_1(k-1)r_{10}$$

但し、

$$r_{00} = \mathbf{x}(k)^T \mathbf{A} \mathbf{x}(k)$$

$$r_{10} = \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k)$$

$$r_{11} = \mathbf{x}(k-1)^T \mathbf{A} \mathbf{x}(k-1)$$

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

また、 α_i は音響エコー経路のインパルス応答の指数減衰率を与えたスカラー量である。

式(2)、(3)の分母に正の定数 δ を加えているが、この量を導入することにより、受話信号の微小区間での零に近い値での除算を防止し、安定な動作を実現できる。図1に、定数 δ を加えた場合と、加えない場合とのエコー消去量（ERLE）の違いを示す。

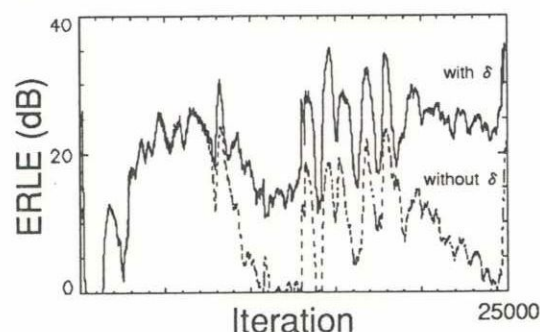


図1 収束特性（ δ の効果）

3 デュオフィルタ構成

適応フィルタは、誤差信号を最小化するように学習を行なうため、目的信号であるエコー信号に近端話者信号が加わるとフィルタ係数が大きく乱れてしまう。

このようなダブルトーク時の対策として、このようなダブルトーク時の対策として、FG/BG方式[2, 3]に着目する。この方式は、適応フィルタと半固定フィルタとを組み合わせ、適応フィルタ係数の誤差が小さく、かつ適応フィルタの誤差の方が半固定フィルタの誤差よりも小さいと判断された場合に、適応フィルタ係数

*Study on the echo canceller based on the duo filter system using the ES Projection algorithm.

By Yoichi Haneda, Shoji Makino, Junji Kojima, Suehiro Shimauchi. (NTT Human Interface Laboratories)

を半固定フィルタ係数に転送するというものである。適応フィルタは実際のエコー消去には使われず、エコー消去は半固定フィルタにより行なう。ダブルトーク時に適応フィルタ係数が乱れた場合には、フィルタ係数は転送されず、半固定フィルタはそれ以前のフィルタ係数を保持するので、ダブルトーク時にも良好なエコー消去が行なえる。

本報告では、このFG/BG方式に、適応アルゴリズムとしてES射影アルゴリズムを適用したデュオフィルタ構成のエコーキャンセラについて収束速度とダブルトーク性能を検討した。図2にデュオフィルタの構成を示す。

図3にデュオフィルタ構成の収束について実験した結果を示す。フィルタ長1000タップで、マイクロホン位置が変化してから約1.2秒で収束していることが分かる。図4にダブルトーク時の送受信信号を示す。ダブルトークの最中やその後においても安定にエコー消去がなされていること、及び近端話者信号 $s(k)$ が影響を受けずに送出されていることが分かる。

4 適応型挿入損失回路の併用

電源投入直後や、系が大きく変動した場合のハウリング発生防止のために、適応損失型の挿入損失回路を併用する。

挿入損失回路では、通話品質の観点から、現在の音響結合量に基づいてハウリングが起こらない最低の損失量を計算して挿入することが重要である。しかし、音響結合量は、受信信号のみが存在する場合に測定を行なわなければならない。本方式では、デュオフィルタの収束状態に基づいて受信状態を判定し、音響結合量と収束量の予測を行なうことにより損失量の制御を行なっている。その結果、的確な挿入損失量が求められ、良好な通話性能が得られるとともに、初期学習も不要となる。

5 まとめ

デュオフィルタ構成のエコーキャンセラでは、音声信号が小さくなる区間やダブルトーク状態でも適応フィルタを安定に動作させることができる。さらに、デュオフィルタと連動して的確な挿入損失量を与える適応型の挿入損失回路を併用することにより、初期学習が不要で、良好な会話性能を持つ音響エコーキャンセラが実現できる。

【謝辞】日頃ご指導頂く、北脇音声情報研究部長、金田主幹研究員、貴重なご意見を頂いた田中研究員に深謝する。

【参考文献】

- [1]S. Makino, Y. Kaneda, "Exponentially weighted step-size projection algorithm for acoustic echo cancellers," Trans. IEICE, Vol. E75-A, no. 11, p.1500, 1992.
- [2]K. Ochiai, T. Araseki, T. Ogihara, "Echo canceller with two echo path models," IEEE Trans. Commun., COM-25, P.589,1977.
- [3]中原、羽田、牧野、吉川、"音響エコーキャンセラにおけるダブルトーク制御方式の検討"、音講論、3-5-7, p. 503, 1992.

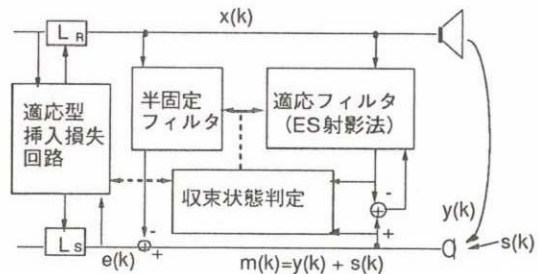


図2 デュオフィルタの構成

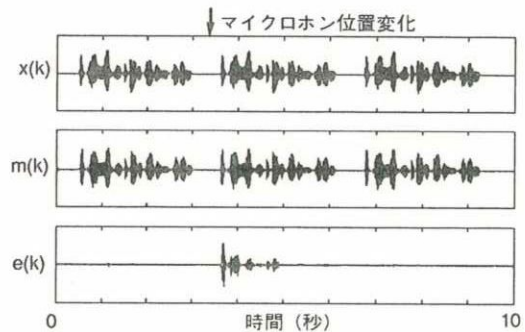


図3 シングルトーク時の送受信信号

$x(k)$: 受信信号、 $m(k)$: マイクロホン受信信号 ($x(k)$ のエコー)、 $e(k)$: 送信信号

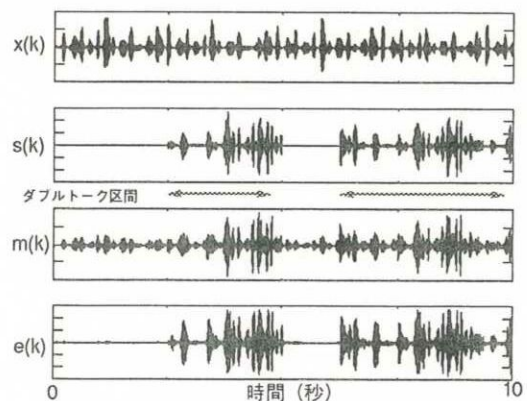


図4 ダブルトーク時の送受信信号

$x(k)$: 受信信号、 $s(k)$: 近端話者信号、 $m(k)$: マイクロホン受信信号 ($x(k)$ のエコー+ $s(k)$)、 $e(k)$: 送信信号