RLS Algorithm and its Stabilization in Alternative Learning Algorithm for Stereophonic Acoustic Echo Canceller
分割学習法に基づくステレオエコーキャンセラにおける RLS アルゴリズムの適用と安定化

Akihiro HIRANO Kenji NAKAYAMA Jun'ya MIZUNO
平野晃宏 中山謙二 水野潤哉
Kanazawa University
金沢大学
E-mail: {hirano, nakayama}@t.kanazawa-u.ac.jp

ABSTRACT

This paper introduces a recursive least squares (RLS) algorithm into the alternative learning algorithm for stereophonic acoustic echo canceller, which can identify correct echo paths without pre-processing. Stabilization techniques including a periodic reset and an adaptive forgetting factor are also proposed. Computer simulation results show faster convergence and improved echo return loss enhancement.

あらまし

前処理なしでエコーパスを同定できる分割学習法に基づくステレオエコーキャンセラにおいて、フィルタ係数更新に RLS アルゴリズムを使用する方法を検討する。学習を安定化させるために、周期的な初期化や時変忘却係数などの係数更新制御を導入する。計算機シミュレーションによって、収束速度が改善されていること、エコー消去量が増加することを示す。

1 Introduction

Echo cancellers are used to reduce echoes in a wide range of applications, such as TV conference systems and hands-free telephones. To realistic TV conferencing, multi-channel audio, at least stereophonic, is essential. For stereophonic teleconferencing, stereophonic acoustic echo cancellers (SAEC's) [1–5] have been studied.

SAEC's have a fundamental problem in which their filter coefficients cannot have an unique solution [1]. Though SAEC's with pre-processing [2] are good candidates for solving this problem, audible sound distortion caused by the pre-processing arises. An SAEC without pre-processing which can estimate correct echo path has also been proposed [4]. While no sound distortion are introduced, its convergence speed is not so fast compared with SAEC's with pre-processing.

This paper proposes a recursive least squares (RLS) algorithm [6] and its stabilization for the alternative learning algorithm. Section 2 reviews the SAEC and its fundamental problem, followed by the alternative learning algorithm. Introduction of the RLS algorithm and its stabilization are shown in Section 4. Computer simulation results show the performance of the proposed algorithm.

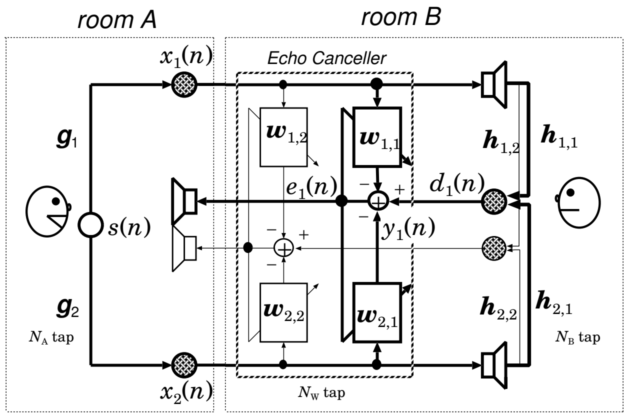

Figure 1: Teleconferencing using SAEC

2 Stereophonic Acoustic Echo Canceller and Uniqueness Problem

Figure 1 shows a teleconferencing using an SAEC. This echo canceller consists of four adaptive filters corresponding to four echo paths from two loudspeakers to two microphones. Each adaptive filter estimates the corresponding echo path.

The far-end signal $\boldsymbol{x}_i(n)$ in the i-th channel at time index n is generated from a talker speech $s(n)$ by passing room A impulse response $\boldsymbol{g}_i$ from the talker to the i-th microphone. $\boldsymbol{x}_i(n)$ passes an echo path $\boldsymbol{h}_{i,j}$ from the i-th loudspeaker to the j-th microphone and become an echo $d_j(n)$. Similarly, adaptive filters $\boldsymbol{w}_{i,j}(n)$ generates an echo replica $y_j(n)$. $\boldsymbol{w}_{i,j}(n)$ is so updated as to reduce the residual echo $e_j(n)$

SAEC's have a fundamental problem in which their filter coefficients cannot have an unique solution [1]. SAEC's may have infinite number of solutions other than the optimum solution $\boldsymbol{w}_{i,j}(n) = \boldsymbol{h}_{i,j}$.

3 Alternative Learning Algorithm

Analyses show that SAEC's may have unique and optimum solution when the number of taps N_W for SAEC and the impulse response length N_A in room A satisfy $N_W < N_A$ [5, 7]. For echo cancellation performance, $N_B < N_W$ is preferable where N_B is the impulse response length in room B. Therefore, if $N_B < N_W < N_A$, SAEC in room B achieves both perfect echo cancellation and optimum solution. Such a condition, however, cannot be satisfied for SAEC's in both room A and B.

In order to satisfy the uniqueness condition for both SAEC's in room A and room B, the number of taps for SAEC N_W is so chosen as to satisfy $N_W/2 < N_A < N_W$ and $N_W/2 < N_B < N_W$. If the size of both rooms are similar, which is usual case, such N_W may exist. In adaptation, $N_W/2$ taps are updated at a time; thus the effective number of taps for SAEC $N_W/2$ is smaller than the impulse response length in the far-end room N_A. To avoid the performance degradation caused by the tap shortage, another $N_W/2$ taps will also update at the other time.

The filter coefficient vector $\boldsymbol{w}_{i,j}(n)$ is divided into two sub-vectors $\boldsymbol{w}_{i,j,f}(n)$ and $\boldsymbol{w}_{i,j,b}(n)$ show by

$$\boldsymbol{w}_{i,j,f}(n) = [w_{i,j,0}(n), \cdots, w_{i,j,N_W/2-1}(n)]^T \tag{1}$$

$$\boldsymbol{w}_{i,j,b}(n) = [w_{i,j,N_W/2}(n), \cdots, w_{i,j,N_W-1}(n)]^T. \tag{2}$$

The superscript ^{T} denotes the transpose of a matrix or a vector. In the first stage, $\boldsymbol{w}_{i,j,f}(n)$ is updated while $\boldsymbol{w}_{i,j,b}(n)$ is fixed. This stage is repeated until $\boldsymbol{w}_{i,j,f}(n)$ converges. As the second stage, $\boldsymbol{w}_{i,j,b}(n)$ is updated while $\boldsymbol{w}_{i,j,f}(n)$ is fixed. This stage is also repeated until $\boldsymbol{w}_{i,j,b}(n)$ converges. These two stages are repeated one after another.

An adaptive step-size and a convergence detection are introduced. An adaptive step-size and a convergence detection are introduced for fast convergence with a small computational cost. The adaptive step-size and the convergence detection are carried out based on the coefficient modification amount defined by

$$D(m) = \frac{\sum_{i=1}^{2} \|\boldsymbol{w}_{i,j,p}(mK) - \boldsymbol{w}_{i,j,p}((m-1)K)\|^2}{\sum_{i=1}^{2} \|\boldsymbol{w}_{i,j,p}(mK)\|^2} \tag{3}$$

where p is either f or b. To avoid the increase of the computational cost, (3) is calculated once in a K iterations. Coefficient adaptation is stopped when (3) is calculated.

The filter coefficients are considered to be converged if $D(m-1) < D(m)$ is satisfied. The step-size is controlled by

$$\mu(m) = \mu_{max} \left(\frac{D(m)}{D_{max}} \right)^{1/4} \tag{4}$$

where D_{max} is a maximum value of $D(m)$ in a same stage. Usually, $D(1)$ is used as a D_{max}. $\mu(n)$ is used within $mK < n < (m+1)K$.

The overview of the adaptation control is as follows:

1. Update filter coefficients with $\mu(0) = \mu_{max}$ for first K iterations.
2. calculate $D(1)$. $D_{max} = D(1)$.
3. Update filter coefficients with $\mu(m)$ by (4) for next K iterations.
4. calculate $D(m)$.
5. If $D(m-1) < D(m)$, then proceed to the next stage.
6. If $D_{max} < D(m)$, then $D_{max} = D(m)$.
7. Goto 3.

4 RLS Algorithm and Its Stabilization

The RLS algorithm is introduced in order to improve the convergence characteristics. Some stabilization techniques are also introduced.

By using combined input signal vectors

$$\begin{aligned}\boldsymbol{x}_f(n) = [&x_1(n), x_2(n), \cdots, \\ &x_1(n - N_W/2 + 1), x_2(n - N_W/2 + 1)]^T\end{aligned} \tag{5}$$

$$\begin{aligned}\boldsymbol{x}_b(n) = [&x_1(n - N_W/2), x_2(n - N_W/2), \cdots, \\ &x_1(n - N_W + 1), x_2(n - N_W + 1)]^T\end{aligned} \tag{6}$$

and also combined filter coefficient vectors

$$\begin{aligned}\boldsymbol{w}_{j,f}(n) = [&w_{1,j,0}(n), w_{2,j,0}(n), \cdots, \\ &w_{1,j,N_W/2-1}(n), w_{2,j,N_W/2-1}(n)]^T\end{aligned} \tag{7}$$

$$\begin{aligned}\boldsymbol{w}_{j,b}(n) = [&w_{1,j,N_W/2}(n), w_{2,j,N_W/2}(n), \cdots, \\ &w_{1,j,N-1}(n), w_{2,j,N-1}(n)]^T,\end{aligned} \tag{8}$$

the RLS algorithm can be applied to the alternative learning algorithm. The RLS based adaptation for the alternative learning algorithm is summarized as

$$y_j(n) = \boldsymbol{x}_{j,f}^T(n)\boldsymbol{w}_{j,f}(n) + \boldsymbol{x}_{j,b}^T(n)\boldsymbol{w}_{j,b}(n) \tag{9}$$

$$e_j(n) = d_j(n) - y_j(n) \tag{10}$$

$$\boldsymbol{k}_p(n) = \frac{\lambda^{-1}\boldsymbol{P}_p(n-1)\boldsymbol{x}_p(n)}{1 + \lambda^{-1}\boldsymbol{x}_p^T(n)\boldsymbol{P}_p(n-1)\boldsymbol{x}_p(n)} \tag{11}$$

$$\boldsymbol{w}_p(n) = \boldsymbol{w}_p(n-1) + \boldsymbol{k}_p(n)e(n) \tag{12}$$

$$\boldsymbol{P}_p(n) = \lambda^{-1}\boldsymbol{P}_p(n-1) - \lambda^{-1}\boldsymbol{k}_p(n)\boldsymbol{x}_p^T(n)\boldsymbol{P}_p(n-1) \tag{13}$$

where the subscript p is either f or b. Either $\boldsymbol{w}_{j,f}(n)$ or $\boldsymbol{w}_{j,b}(n)$ is updated.

For stable and fast convergence, the following techniques are used in the adaptation.

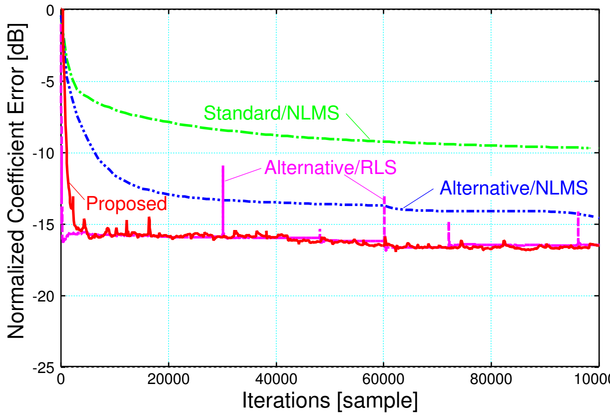

Figure 2: NCE in early stage.

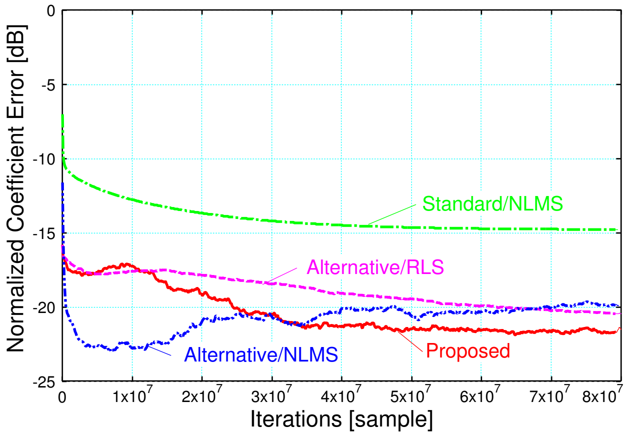

Figure 3: NCE in latter stage.

- Periodic reset
  Reset $\boldsymbol{P}_p(n)$ to zero once in L_1 samples. Such reset avoids unstability.

- Freeze coefficients just after reset
  Coefficient update in (12) is not carried out for L_2 samples after the reset.

- Adaptive forgetting factor
  Adaptive forgetting factor defined by

$$\lambda(m) = \lambda_0 + (1 - \lambda_0)\left\{1 - \left(\frac{D(m)}{D_{max}}\right)\right\} \quad (14)$$

  is used, where $D(m)$ is same as that in Section 3. D_{max} is a maximum value of $D(m)$. $\lambda(m)$ is updated once in K samples.

- Convergence detection as in Section 3

5 Computer Simulations

Simulations have been carried out to show the performance of the proposed algorithm. Far-end room impulse responses $\boldsymbol{g}_i$ are 60-tap FIR filters while those for near-end room $\boldsymbol{h}_{i,j}$ are 64-tap FIR filters. In this case, SAEC's do not have an unique solution. Adaptive filters are 64-tap FIR filters. Colored noise by second-order AR model is used as a talker signal. The pole locations are $\theta = \pm 45, r = 0.9$. Additive white Gaussian noise signal is introduced. The echo-to-noise ratio is 60dB.

The proposed algorithm is compared with the standard SAEC [1], the alternative learning algorithm [4], and the input sliding algorithm [2]. The step-size μ for the standard SAEC is 1.0. The parameters for the alternative algorithm with NLMS are $\mu = 0.5$, $K = 5000$. A smaller step-size is used for stability reason. For the proposed algorithm, $L_1 = 128$, $L_2 = 2000$, $K = 6000$, $\lambda_0 = 0.9999$ are used. The parameters for the input sliding algorithm are chosen as $\mu = 1$, $Q = 60$, and $L = 6$.

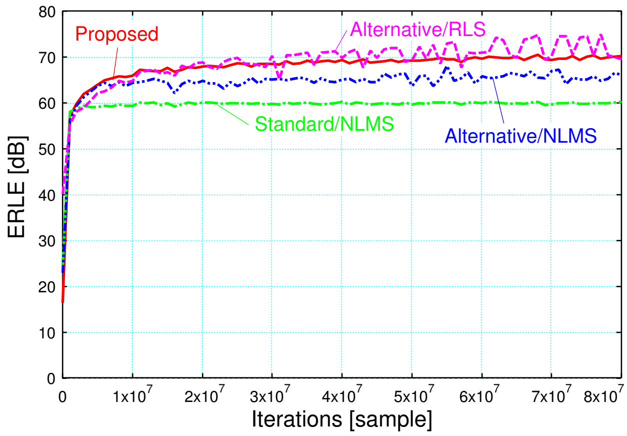

Figure 4: ERLE.

Figure 2 compares the normalized coefficient error (NCE) defined by

$$NCE(n) = \frac{\sum_{i=1}^{2} \|\boldsymbol{w}_{i,j}(n) - \boldsymbol{h}_{i,j}\|^2}{\sum_{i=1}^{2} \|\boldsymbol{h}_{i,j}\|^2}. \quad (15)$$

in early stage. Though the RLS algorithm without stabilization (Alternative/RLS) converges fastest in early periods, impulsive increases of the NCE have been occurred. Such increases are much smaller for the proposed algorithm. The convergence time to -15dB of the NCE for the proposed algorithm is almost 1/20 compared with the alternative learning algorithm with NLMS (Alternative/NLMS).

The NCE for latter stage is compared by Fig. 3. The alternative learning algorithms achieve almost -20dB of the NCE while that of the standard SAEC is -15dB. the proposed algorithm reduces the NCE by 2dB compared with that of Alternative/NLMS. The convergence time of the proposed algorithm to -20dB of the NCE is reduced by almost 60% compared with Alternative/RLS.

The echo return loss enhancement (ERLE) is depicted by Fig 4. Introduction of the RLS algorithm improves the ERLE by almost 5dB.

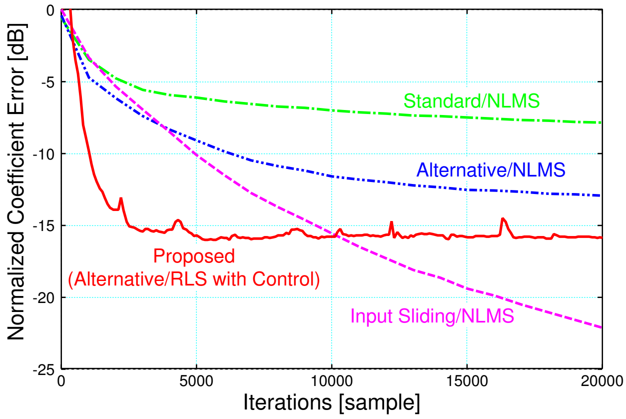

Figure 5: NCE compared with SAEC using pre-processing.

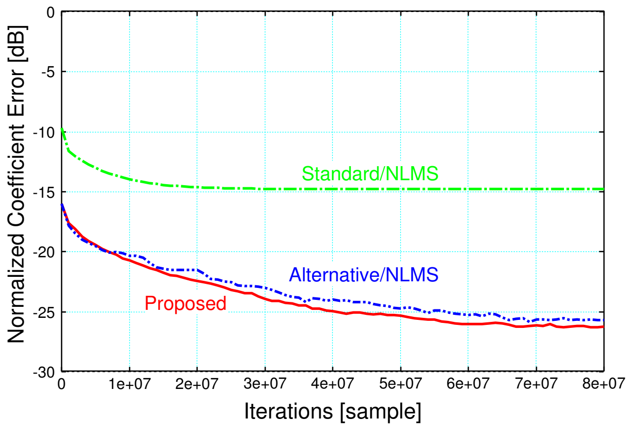

Figure 6: NCE for white Gaussian inputs.

Figure 5 compares the NCE for the alternative learning algorithms with that for the input sliding algorithm (Input Sliding/NLMS). The convergence time to -10dB of the NCE for Alternative/NLMS and Input Sliding/NLMS are comparable. The proposed algorithm reduces the convergence time to -15dB by 60% compared with Input Sliding/NLMS. The convergence to the final value for Input Sliding/NLMS is fastest.

Convergence for white Gaussian inputs has also been compared. The NCE is shown by Fig. 6. The convergence speed and accuracy are slightly improved. A possible reason of convergence speed improvement for white signals by introducing the RLS algorithm might be the influence of the inter-channel cross-correlation.

6 Conclusions

This paper introduces an RLS algorithm and its stabilization into the alternative learning algorithm for SAEC. Stabilization techniques including a periodic reset and an adaptive forgetting factor are also proposed. Computer simulation results show faster convergence and improved echo return loss enhancement.

References

[1] M. M. Sondhi and D. R. Morgan, “Stereophonic acoustic echo cancellation — an overview of the fundamental problem,” *IEEE SP Letters*, vol. 2, no. 8, pp. 148–151, Aug. 1995.

[2] A. Sugiyama, Y. Joncour, and A. Hirano, “A stereo echo canceler with correct echo-path identification based on an input-sliding technique,” *IEEE Trans. SP*, vol. 49, no. 11, pp. 2577–2587, Nov. 2001.

[3] A. Hirano, K. Nakayama, and K. Watanabe, “Convergence analysis of stereophonic echo canceller with pre-processing — relation between pre-processing and convergence —,” *Proc. of ICASSP ’99*, pp. 861–864, Mar. 1999.

[4] A. Hirano, K. Nakayama, D. Someda, and M. Tanaka, “Stereophonic acoustic echo canceller without pre-processing,” *Proc. of ICASSP 2004*, vol. 4, pp. 145–148, 5 2004.

[5] J. Benesty, D. Morgan, and M. Sondhi, “A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation,” *IEEE Trans. Speech and Audio*, vol. 6, no. 2, pp. 156–165, Mar. 1998.

[6] S. Haykin, *Adatptive Filter Theory, Third Edition*, Prentice Hall, 1996.

[7] M. Tanaka, A. Hirano, and K. Nakayama, “On uniqueness of filter coefficients in stereophonic acoustic echo canceller,” *Proceedings of 14th DSP Symposium*, pp. 31–36, Nov. 1999 (In Japanese).