Signal Processing 131 (2017) 422–433

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Tap-length optimization of adaptive filters used in stereophonic acoustic echo cancellation

Asutosh Kar [a,*], M.N.S. Swamy [b]

[a] Department of Electronic Systems, Aalborg University, Denmark
[b] Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

ARTICLE INFO

Article history:
Received 12 February 2016
Received in revised form
24 August 2016
Accepted 4 September 2016
Available online 5 September 2016

Keywords:
Adaptive filter
Tap-length
Stereophonic acoustic echo cancellation
Multiple sub-filters
Convergence
Signal-to-noise ratio

ABSTRACT

An adaptive filter with a large number of weights or taps is necessary for stereophonic acoustic echo cancellation (SAEC), depending on the room impulse response and acoustic path where the cancellation is performed. However, a large tap-length results in slow convergence and increases the complexity of the tapped delay line structure for FIR adaptive filters. To overcome this problem, there is a need for an optimum tap-length-estimation algorithm that provides better convergence for the adaptive filters used in SAEC. This paper presents a solution to the problem of balancing convergence and steady-state performance of long length adaptive filters used for SAEC by proposing a new tap-length-optimization algorithm. The optimum tap length and step size of the adaptive filter are derived considering an impulse response with an exponentially-decaying envelope, which models a wide range of acoustic echo paths. The tap-length optimization is applied to a single long adaptive filter with thousands of coefficients to decrease the total number of weights, which in turn reduces the computational load. To further increase the convergence rate, the proposed tap-length-optimization algorithm is applied to an existing multiple sub-filter-based echo canceller, for which we present a convergence analysis. Computer simulations are also presented, comparing the proposed approach with related work.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Adaptive filtering techniques have been actively used in acoustic echo cancellation for mitigating the echo that arises due to a coupling between the loudspeaker and the microphone in teleconferencing and hands-free telephony. Dual-channel stereo acoustic devices used for better sound quality in desktop conferencing and hands-free telephony provide telepresence to end-users by enabling listeners to locate speakers in meetings, where more than one person may speak at the same time. However, these are more complicated as compared to monophonic devices [1]. Stereophonic arrangements normally comprise a two-way full-duplex channels and dual speakers and microphones at both ends of a communication channel as shown in Fig. 1 [2]. A speaker with speech $a(n)$ is active in the transmission room. Two microphones are used to pick up signals from the speaker via two acoustic paths characterized by two room impulse responses $a_1(n)$ and $a_2(n)$, respectively. The speech signals are convolved with the room impulse responses to produce stereophonic signals $x_1(n)$ and $x_2(n)$. These stereophonic signals are then transmitted to loudspeakers present in the receiving room, which in turn are acoustically coupled to the receiving room microphones through paths characterized by the impulse responses $b_1(n)$ and $b_2(n)$. This coupling produces the echo signal $d(n)$. If nothing is done to suppress the echo, the microphone signal would pass back to the loudspeaker in the transmission room. This would be recirculated again and again, resulting in multiple undesirable echoes [2]. Two adaptive filters $\mathbf{w}_1$ and $\mathbf{w}_2$ use the input signals to produce an estimate of the echo signal. The difference between $d(n)$ and filter output $y(n)$ produces the residual echo signal $e(n)$. The main purpose of the system in Fig. 1 is to minimize the variance of the echo signal $e(n)$. The whole process is referred to as adaptive SAEC.

The filter length increases in accordance with the length of the impulse response of the echo return path [3]. In SAEC, modeling of the impulse response of the echo return path is more challenging than single-channel echo models, as it is required to model two different paths at the same time [4]. In general, these acoustic echo cancellers are based on an adaptive finite impulse response (FIR) filter with thousands of filter coefficients to model the acoustic echo path [3]. A large number of filter coefficients leads to a high computational burden and slow convergence [3]. In order to overcome this problem, the following schemes have been proposed:

* Corresponding author.
E-mail addresses: asuka@es.aau.dk (A. Kar), swamy@ece.concordia.ca (M.N.S. Swamy).

http://dx.doi.org/10.1016/j.sigpro.2016.09.003
0165-1684/© 2016 Elsevier B.V. All rights reserved.

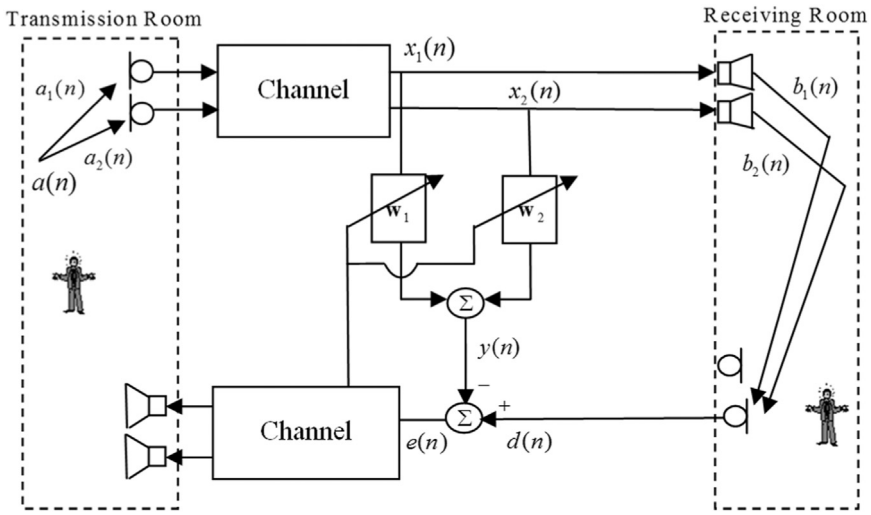

Fig. 1. SAEC using adaptive filters for both near-end and far-end users.

i. Decomposition based on subdividing high-order adaptive filters to multiple lower-order adaptive filters to improve the convergence [5–7].
ii. Optimizing the total number of filter coefficients necessary for echo cancellation using an adaptive tap-length-selection algorithm [8–11].

If we use a single adaptive filter with thousands of coefficients for SAEC without decomposition, then it generally results in slow convergence [2]. The idea of multiple sub-filters (MSF) is based on the observation that the least-mean-square (LMS) algorithm with fewer coefficients converges faster [6], due to the fact that the eigenvalue spread decreases in most scenarios, except in the case of white noise [12,13]. Out of the various existing sub-filter algorithms discussed in [6], only the modified MSF-based combined-error algorithm (COEA) [11], significantly improves the convergence without affecting the steady-state error. However, the filter order remains high in echo cancellation algorithms [6]. This brings the necessity of a tap-length-optimization algorithm.

There is a need to optimize the filter length in echo cancellers to avoid the above discussed issues of complexity and to further improve the convergence. Several tap-length-optimization adaptive algorithms, such as the fractional tap-length LMS (FT-LMS) algorithm [14], variable-leakage-factor FT-LMS algorithm [15], and variable-tap variable-step LMS (VT-VSLMS) algorithm [16], have been introduced to diminish the complexity of the adaptive designs. However, a generalization of these algorithms to SAEC is not available. These existing algorithms also bring their own limitations that have been highlighted in [17]. The variable-tap mean-square-deviation (VT-MSD) algorithm [8] initiates the need for tap-length optimization of adaptive filters used in acoustic echo cancellation. However, that paper does not present the convergence and steady state performance analysis, and in some cases it results in a higher tap-length requirement compared to that of the FT-LMS. The work presented in [11] develops the concept of finding variable-tap-length adaptive filters for sub-filter-based stereophonic and monophonic echo cancellation without any mathematical convergence analysis for the MSF-based echo cancellation algorithms discussed therein. The algorithms proposed in those papers also do not consider an exponentially-decaying impulse response for the unknown room impulse response of the acoustic echo path, which is important in preliminary evaluation of echo cancellers. The variable-tap-length variable-step NLMS with variable error width (VT-VSNLMS$_{VE}$) presented in [17] tries to remedy the downside of existing algorithms in both high and low noise scenarios. However, that work also does not present any convergence analysis, and the performance has only been tested against low-order adaptive filters for non-exponential impulse responses. The existing variable-tap-length algorithms to model an exponentially-decaying impulse response [18] are neither specifically designed nor analyzed for acoustic echo cancellation applications of adaptive filters. Therefore, an advanced tap-length optimization algorithm is needed for improving convergence speed in signal processing applications like echo cancellation, active noise control and equalization cannot be ignored [19,20]. The most advanced tap-length optimization algorithm for feedforward active noise control using the filtered-x LMS algorithm is proposed in [21]. This paper proposed a two-sided formulation and is a useful method for active noise control applications. On the other hand, it does not employ the MSF-based algorithms for improving the convergence. The SAEC scenario is different in many ways compared to feedforward active noise control and finding an advanced variable-tap-length algorithm for SAEC has its own challenges [19]. Hence, there exists a need to derive an improved tap-length-optimization algorithm for finding the optimum tap length and step size for high-order adaptive filters used in single-long-filter (SLF)- and MSF-based SAEC.

The objective of this paper is to propose an improved solution to mitigate the slow convergence and computational intensiveness of long-length adaptive filters used for SAEC. In particular, we provide the following:

i. An improved tap-length selection algorithm for adaptive filters used in SAEC for better performance over existing variable-tap-length algorithms.
ii. A convergence and steady-state analysis of the proposed algorithm.
iii. A convergence analysis of an existing MSF-based COEA for SAEC.

The tap-length-optimization algorithm is specifically designed to model an exponentially-decaying envelope in the echo return path. We apply the proposed algorithm to adaptive SLF and MSF SAECs to avoid the downside of too-short or too-long filters due to tap-length mismatch. This also mitigates the problem of random initialization of filter length and undermodelling. In each iteration, the optimum step size and tap length are derived by minimizing the deviation between the actual and estimated filter weights. In this proposed algorithm the values of the adaptive parameters are not randomly set to a particular value like in FT-LMS; instead, it

finds the optimum filter length for high-order adaptive filters, resulting in faster convergence compared to the existing algorithms.

In this paper, we focus on the adaptive filtering part of the SAEC and do not take into account other aspects such as double-talk detection, multipath estimation, signal decorrelation, and residual echo cancellation since these are beyond the intended scope. The rest of the paper is organized as follows. Section 2 includes the proposed algorithm for optimizing the tap length of a high-order adaptive filter, which finds an application in SAEC and comes as a major contribution of this paper. In Section 3, we present a convergence analysis of the advanced MSF-based COEA used for SAEC. The simulation study is provided in Section 4.

Throughout the paper, the following notation is adopted: $E[\,.\,]$ for the expected value, $(.\,)^T$ for the transpose of a vector or matrix, $\mathrm{tr}\,(\,.\,)$ for the trace operation, $\mathrm{nint}(.\,)$ for converting a fractional value to the nearest integer, $\mathrm{sgn}(.\,)$ for finding the sign, $\|\,.\,\|_2$ for the l_2 norm, $\det[\,.\,]$ for the determinant of a matrix and $\mathrm{diag}[\,.\,]$ for the diagonal elements of a matrix.

2. Proposed tap-length-optimization algorithm

In this section, a new tap-length-optimization algorithm is derived that iteratively attains the optimum filter order to model an exponentially-decaying impulse response. In a variable-tap-length algorithm, the step size is dependent on the tap length [12]. Here, we derive a variable-step-size algorithm and perform a convergence analysis. Later, a steady-state analysis of the proposed tap-length-adaptation algorithm is provided. Following the conventional LMS approach, the taps are updated by the LMS algorithm

$$\mathbf{w}_{L(n)}(n+1) = \mathbf{w}_{L(n)}(n) + \mu(n)\, e(n)\, \mathbf{x}_{L(n)}(n)\,, \tag{1}$$

where $L(n)$ is the tap length, $\mu(n)$ is the step size and

$$e(n) = d(n) - y(n)\,, \tag{2}$$

is the error signal at the nth iteration where $d(n)$ is the desired signal and the filter output

$$y(n) = \mathbf{w}_{L(n)}^T(n)\mathbf{x}_{L(n)}(n)\ . \tag{3}$$

The input vector to the adaptive filter is $\mathbf{x}_{L(n)}(n) = \left[x(n), x(n-1), \ldots, x(n-L(n)+1)\right]^T$ and weight vector pertaining to filter length is $\mathbf{w}_{L(n)}(n) = \left[w_0(n), w_1(n), \ldots, w_{L(n)-1}(n)\right]^T$. The desired signal is observed as the sum of the transmitted signal convolved with the echo path impulse response and an additive noise:

$$d(n) = \mathbf{p}_N^T\, \mathbf{x}_N(n) + \zeta(n), \tag{4}$$

where $\mathbf{p}_N = \left[p_0, p_1, \ldots, p_{N-1}\right]^T$ is the channel response modelled by $p_k = u(k)\, e^{-\delta(k-1)}$, $k = 0, 1, \ldots, N-1$, $\mathbf{x}_N(n) = \left[x(n), x(n-1), \ldots, x(n-N+1)\right]^T$ is the input vector to the unknown channel with N coefficients, $u(k)$ is a Gaussian zero-mean stochastic process with unit variance, δ is a positive constant for the decay rate and $\zeta(n)$ is the additive noise, assumed to be statistically independent of the input sequence $\mathbf{x}_N(n)$. For the analysis here, we assume that both $\mathbf{x}_N(n)$ and $\zeta(n)$ are independent zero-mean Gaussian white noise signals, identically distributed over time. So in this case, the adaptive filter with tap length $L(n) \le N$ is used to estimate the unknown impulse response. Here, we present an analysis to find a procedure for updating the tap length $L(n)$ and the step size $\mu(n)$ at each iteration.

Following the formulation in [18], the exponentially-decaying impulse response $\mathbf{p}_N$ is partitioned into two parts as $\left[\mathbf{p}',\ \mathbf{p}''\right]^T$, where $\mathbf{p}'$ is modelled by $\mathbf{w}_{L(n)}(n)$ and $\mathbf{p}''$ is the un-modelled part. Accordingly, we have two representations for the coefficient error vector $\xi(n)$ as [18]

$$\xi'_{L(n)}(n) = \mathbf{w}_{L(n)}(n) - \mathbf{p}' \tag{5}$$

and

$$\xi_N(n) = \mathbf{w}_N(n) - \mathbf{p}_N\ , \tag{6}$$

where $\xi'_{L(n)}(n)$ is the estimation error of $L(n) \le N$ partial coefficients and $\xi_N(n)$ is the total coefficient estimation error. The vector $\mathbf{w}_N(n)$ is obtained by padding $\mathbf{w}_{L(n)}(n)$ with $N - L(n)$ zeros as,

$$\mathbf{w}_N(n) = \begin{bmatrix} \mathbf{w}_{L(n)} \\ 0_{N-L(n)} \end{bmatrix}. \tag{7}$$

Combining (3) with (4) and substituting into (2), the error signal $e(n)$ can be rewritten as

$$e(n) = \left[\mathbf{p}_N - \mathbf{w}_{L(n)}(n)\right]^T \mathbf{x}_N(n) + \zeta(n)\ . \tag{8}$$

Padding $\mathbf{w}_{L(n)}(n)$ with $N - L(n)$ zeros, as shown in (7), and replacing it with $\mathbf{w}_N(n)$, we have

$$e(n) = -\,\mathbf{x}_N^T(n)\, \xi_N(n) + \zeta(n)\,. \tag{9}$$

Substituting (9) for $e(n)$ in the weight update Eq. (1) and then subtracting $\mathbf{p}_N$ from both the sides and replacing $\mathbf{x}_N(n)$ with $\begin{bmatrix} \mathbf{x}_{L(n)}(n) \\ 0_{N-L(n)} \end{bmatrix}$, we get

$$\xi_N(n+1) = \xi_N(n) - \mu(n) \begin{bmatrix} \mathbf{x}_{L(n)}(n) \\ 0_{N-L(n)} \end{bmatrix} \left[\mathbf{x}_N^T(n)\, \xi_N(n) - \zeta(n)\right]\ . \tag{10}$$

Simplifying the above equation by substituting

$$\mathbf{A}(n) = \mathbf{I}_N - \mu(n) \begin{bmatrix} \mathbf{x}_{L(n)}(n) \\ 0_{N-L(n)} \end{bmatrix} \mathbf{x}_N^T(n)\,, \tag{11}$$

we can write (10) as

$$\xi_N(n+1) = \mathbf{A}(n)\, \xi_N(n) + \mu(n)\zeta(n) \begin{bmatrix} \mathbf{x}_{L(n)}(n) \\ 0_{N-L(n)} \end{bmatrix}\,, \tag{12}$$

where $\mathbf{I}_N$ is the $N \times N$ identity matrix. The MSD is formulated as the figure of merit for finding the optimum step size and tap length parameters. The cost function $f(n)$ is expressed as the expected square of the l_2 norm of the coefficient error vector $\xi_N(n)$as

$$f(n) = E\left[\,\|\xi_N(n)\|_2^2\right]\ . \tag{13}$$

Let the input $x(n)$ and the additive noise $\zeta(n)$ have variances σ_x^2 and σ_ζ^2, respectively. Following the evaluation of MSD made in [18], we formulate

$$\begin{aligned} f(n+1) &= \beta\, E\left[\,\|\xi_N(n)\|_2^2\right] + (\eta - \beta)\|\mathbf{p}''\|_2^2 + \gamma \\ &= \beta\, f(n) + (\eta - \beta)\|\mathbf{p}''\|_2^2 + \gamma\ , \end{aligned} \tag{14}$$

where, as mentioned earlier, the un-modelled part $\mathbf{p}''$ is represented by $\mathbf{w}'_{N-L(n)}(n)$ and

$$\beta = 1 - 2\,\mu(n)\,\sigma_x^2 + \left\{\left[L(n)+2\right]\mu^2(n)\right\}\sigma_x^4\ , \tag{15}$$

$$\eta = 1 + L(n)\,\mu^2(n)\,\sigma_x^4\ , \tag{16}$$

$$\gamma = L(n)\,\mu^2(n)\,\sigma_x^2\,\sigma_\zeta^2\ . \tag{17}$$

Substituting (15)–(17) into (14), we get

$$f(n+1) = \left\{1 - 2\,\mu(n)\sigma_x^2 + \left[L(n)+2\right]\mu^2(n)\,\sigma_x^4\right\} f(n) + \left[2\,\mu(n)\,\sigma_x^2 - 2\,\mu^2(n)\,\sigma_x^4\right] E\left[\|\mathbf{p}''\|_2^2\right] + L(n)\,\mu^2(n)\,\sigma_x^2\,\sigma_\zeta^2 \ . \tag{18}$$

2.1. Variable step size

Here, the objective is to minimize the MSD cost function with respect to the step size $\mu(n)$ by equating $\frac{\partial f(n+1)}{\partial \mu} = 0$, where

$$\frac{\partial f(n+1)}{\partial \mu} = 2\,f(n)\,\sigma_x^2\left[(L(n)+2)\,\mu(n)\,\sigma_x^2 - 1\right] + 4\,\sigma_x^2\left[0.5 - \mu(n)\,\sigma_x^2\right] E\left[\|\mathbf{p}''\|_2^2\right] + 2\,\mu(n)\,\sigma_x^2\,\sigma_\zeta^2\,L(n) \ . \tag{19}$$

Hence, we get

$$\mu(n) = \frac{f(n) - E\left[\|\mathbf{p}''\|_2^2\right]}{f(n)\,\sigma_x^2\left[L(n)+2\right] - 2\,\sigma_x^2\,E\left[\|\mathbf{p}''\|_2^2\right] + L(n)\,\sigma_\zeta^2} \ . \tag{20}$$

The squared norm of the un-modelled part of $\mathbf{p}_N$, having an exponentially-decaying impulse response, can be expressed as [18]

$$\|\mathbf{p}''\|_2^2 = \frac{e^{-2L(n)\delta} - e^{-2N\delta}}{1 - e^{-2N\delta}}\,\|\mathbf{p}_N\|_2^2 \ . \tag{21}$$

Taking the expectation of both sides, we get

$$E\left[\|\mathbf{p}''\|_2^2\right] = \frac{e^{-2L(n)\delta} - e^{-2N\delta}}{1 - e^{-2N\delta}} E\left[\|\mathbf{p}_N\|_2^2\right] \ , \tag{22}$$

where

$$E\left[\|\mathbf{p}_N\|_2^2\right] = \frac{1 - e^{-2N\delta}}{1 - e^{-2\delta}} \ . \tag{23}$$

Substituting (23) into (22), we have

$$E\left[\|\mathbf{p}''\|_2^2\right] = \frac{e^{-2L(n)\delta} - e^{-2N\delta}}{1 - e^{-2\delta}} \ . \tag{24}$$

Substituting (24) into (20), we get

$$\mu(n) = \frac{f(n) - (1 - e^{-2\delta}) - e^{-2L(n)\delta} + e^{-2N\delta}}{\left\{f(n)\,\sigma_x^2\left[L(n)+2\right] + L(n)\,\sigma_\zeta^2\right\}(1 - e^{-2\delta}) - 2\,\sigma_x^2(e^{-2L(n)\delta} - e^{-2N\delta})} \ . \tag{25}$$

The above equation says that the step-size variation depends on the tap length $L(n)$, the decay rate δ, the cost function $f(n)$, and the variances of the noise and input signal. Considering the filter input and filter coefficients to be independent, the MSE of the LMS adaptive filter can be expressed as [19]

$$E\left[e^2(n)\right] = f(n)\,\sigma_x^2 + \sigma_\zeta^2 \ . \tag{26}$$

Moreover, the statistical squared average $E\left[e^2(n)\right]$ can be approximated recursively by its time average estimate as [20,21]

$$E\left[e^2(n)\right] \approx \widehat{e}^{\,2}(n) = \rho\,\widehat{e}^{\,2}(n-1) + (1-\rho)e^2(n) \ , \tag{27}$$

where ρ is the smoothing constant [17]. Combining (13), (26), and (27) and substituting into (25), we get

$$\mu(n) = \frac{\left(\frac{E\left[e^2(n)\right] - \sigma_\zeta^2}{\sigma_x^2}\right) - (1 - e^{-2\delta}) - e^{-2L(n)\delta} + e^{-2N\delta}}{\left\{\left[E\left[e^2(n)\right] - \sigma_\zeta^2\right]\left[L(n)+2\right] + L(n)\,\sigma_\zeta^2\right\}(1 - e^{-2\delta}) - 2\,\sigma_x^2(e^{-2L(n)\delta} - e^{-2N\delta})}$$

$$= \frac{\rho\,\widehat{e}^{\,2}(n-1) + (1-\rho)\,e^2(n) - \sigma_\zeta^2 - \left[(1 - e^{-2\delta}) - e^{-2L(n)\delta} + e^{-2N\delta}\right]\sigma_x^2}{\left\{\left[\rho\,\widehat{e}^{\,2}(n-1) + (1-\rho)\,e^2(n) - \sigma_\zeta^2\right]\left[L(n)+2\right] + L(n)\,\sigma_\zeta^2\right\}(1 - e^{-2\delta})\,\sigma_x^2 - 2\,\sigma_x^4(e^{-2L(n)\delta} - e^{-2N\delta})} \ . \tag{28}$$

It is difficult to evaluate the closed-form solution of (28) for the variable step size. If the MSD is assumed to be a convex function of the step size, then the stationary points of (27) may give rise to a global minimization of the cost function $f(n+1)$. It is also difficult to find the convexity of the cost function due to the resulting Hessian solution [19,22]. If we assume the filter input and filter coefficients to be independent, then the step size shown in (28) is the optimum solution and depends on the tap length at the n th iteration. Here, we consider an approximate solution of $\mu(n)$ by assuming $L(n) \approx L(n-1)$. With this assumption, (28) can be modified as

$$\mu(n) = \frac{\rho\,\widehat{e}^{\,2}(n-1) + (1-\rho)\,e^2(n) - \sigma_\zeta^2 - \left[(1 - e^{-2\delta}) - e^{-2L(n-1)\delta} + e^{-2N\delta}\right]\sigma_x^2}{\left\{\left[\rho\,\widehat{e}^{\,2}(n-1) + (1-\rho)\,e^2(n)\right]\left[(L(n-1)+2)\right] - 2\,\sigma_\zeta^2\right\}(1 - e^{-2\delta})\,\sigma_x^2 - 2\,\sigma_x^4(e^{-2L(n-1)\delta} - e^{-2N\delta})} \ . \tag{29}$$

Now, for simplicity, we consider the assumptions mentioned in [18,22] to find the step-size convergence in this case: zero background noise with $\sigma_\zeta^2 = 0$ and the tap length of the filter perfectly modelled such thatb $\|\mathbf{p}''\|_2^2 = 0$. Again, we assume $L(n) \approx L(n-1)$ to avoid the dependencies between step size and tap length in every n th iteration. Otherwise, a pair of stationary points for both $\mu(n)$ and $L(n)$ will be obtained at each iteration and it will be difficult to find a closed-form solution for convergence. Using these assumptions, (20) can be expressed as

$$\mu(n) = \frac{1}{\left[L(n-1)+2\right]\sigma_x^2} \ . \tag{30}$$

We can use similar assumptions for large decay rate $\delta >> 1$ and $f(n) - 1 \approx f(n)$ in (29) to get (30). This signifies that if the tap length $L(n)$ is large, the step size is small, which results in slow convergence, and vice versa. Hence, for convergence, the step size should be in the range.

$$0 < \mu(n) \le \frac{1}{\left[L(n-1)+2\right]\sigma_x^2} \ . \tag{31}$$

2.2. Optimum tap length

We can find the optimum tap length by minimizing the cost function $f(n+1)$ of (18) with respect to tap length $L(n)$, which requires that

$$\frac{\partial f(n+1)}{\partial L(n)} = f(n)\,\mu^4(n)\,\sigma_x^2 + \mu^2(n)\,\sigma_x^2\,\sigma_\zeta^2 + 2\,\mu(n)\,\sigma_x^2\Big[1-\mu(n)\,\sigma_x^2\Big]\frac{d\,E\Big[\|\mathbf{p}''\|_2^2\Big]}{d\,L(n)} = 0\,. \tag{32}$$

To evaluate (32), we need to find the derivative of $E\Big[\|\mathbf{p}''\|_2^2\Big]$ with respect to $L(n)$. Hence, considering (24) and finding the derivative, we get

$$\frac{d\,E\Big[\|\mathbf{p}''\|_2^2\Big]}{d\,L(n)} = \frac{-2\,\delta e^{-2\delta L(n)}}{1-e^{-2\delta}}\,. \tag{33}$$

Substituting this into (32), we get

$$e^{-2\delta L(n)} = \frac{\mu(n)(1-e^{-2\delta})\Big[f(n)\,\sigma_x^2+\sigma_\zeta^2\Big]}{4\,\delta\Big[1-\mu(n)\sigma_x^2\Big]}\,, \tag{34}$$

which implies that

$$L(n) = 0.5\,\delta^{-1}\,\ln\frac{4\,\delta\Big[1-\mu(n)\sigma_x^2\Big]}{\Big[f(n)\,\sigma_x^2+\sigma_\zeta^2\Big]\mu(n)(1-e^{-2\delta})} \tag{35}$$

Then, we replace the step size $\mu(n)$ in the above expression with (30) and get

$$L(n) = 0.5\,\delta^{-1}\ln\frac{4\,\delta\Big[1-\frac{1}{L(n-1)+2}\Big]}{\Big[\Big(f(n)\,\sigma_x^2+\sigma_\zeta^2\Big)(1-e^{-2\delta})\Big]^{\frac{1}{L(n-1)+2}}}$$
$$=0.5\,\delta^{-1}\ln\frac{4\,\delta\;L(n-1)\sigma_x^2}{\Big[f(n)\sigma_x^2+\sigma_\zeta^2\Big](1-e^{-2\delta})} \tag{36}$$

From (36), $L(n-1)$ can be obtained as

$$L(n-1) = 0.5\,\delta^{-1}\,\ln\frac{4\,\delta\;L(n-2)\sigma_X^2}{\Big[f(n-1)\sigma_x^2+\sigma_\zeta^2\Big](1-e^{-2\delta})}\,, \tag{37}$$

We define $\Delta L = L(n) - L(n-1)$ as the difference between the tap lengths at any two consecutive iterations so we have

$$\Delta L = L(n) - L(n-1) = 0.5\,\delta^{-1}\ln\frac{L(n-1)\Big[f(n-1)\,\sigma_x^2+\sigma_\zeta^2\Big]}{L(n-2)\Big[f(n)\,\sigma_x^2+\sigma_\zeta^2\Big]}$$
$$= \frac{1}{\delta}\,\ln\sqrt{\frac{L(n-1)\Big[f(n-1)\,\sigma_x^2+\sigma_\zeta^2\Big]}{L(n-2)\Big[f(n)\,\sigma_x^2+\sigma_\zeta^2\Big]}} \tag{38}$$

Hence, we get the tap-length adaptation equation

$$L(n) = L(n-1) + \Delta L$$
$$= L(n-1) + \frac{1}{\delta}\,\ln\sqrt{\frac{L(n-1)\Big[f(n-1)\,\sigma_x^2+\sigma_\zeta^2\Big]}{L(n-2)\Big[f(n)\,\sigma_x^2+\sigma_\zeta^2\Big]}} \tag{39}$$

Substituting (26) into (39) and iterating, we have

$$L(n) = L(n-1) + \frac{1}{\delta}\,\ln\sqrt{\frac{L(n-1)\;\widehat{e}^{\,2}(n-1)}{L(n-2)\;\widehat{e}^{\,2}(n)}}$$
$$= L(0) + \frac{1}{\delta}\,\ln\sqrt{\frac{L(n-1)\;\widehat{e}^{\,2}(n-1)}{L(0)\;\widehat{e}^{\,2}(0)}}\,. \tag{40}$$

Replacing the squared estimated error in (40) with (27), we have

$$L(n) = L(n-1) + \frac{1}{\delta}\,\ln\sqrt{\frac{L(n-1)\Big[\rho\;\widehat{e}^{\,2}(n-2) + (1-\rho)\,e^2(n-1)\Big]}{L(n-2)\Big[\rho\;\widehat{e}^{\,2}(n-1) + (1-\rho)\,e^2(n)\Big]}} \tag{41}$$

This shows the proposed tap-length adaptation, which follows from the optimization criteria mentioned earlier for an impulse response having an exponentially-decaying envelope. The filter length is an integer, but evolving the optimum filter order will usually result in a fractional value, which should be rounded to an integer, i.e., nint$\big[L(n)\big]$.

2.3. Steady-state analysis

The steady-state analysis of the tap-length adaption can be obtained from (40) as follows,

$$L(\infty) = L(0) + \frac{1}{\delta}\,\ln\sqrt{\frac{L(\infty)\,\widehat{e}^{\,2}(0)}{L(0)\widehat{e}^{\,2}(\infty)}}\,, \tag{42}$$

where $L(0)$ is the initial value of the tap-length and $L(\infty)$ is the steady-state final value. Similarly, $\widehat{e}^{\,2}(0)$ and $\widehat{e}^{\,2}(\infty)$ are the initial and steady-state values of the estimated square error. If the desired signal $d(n)$, having variance σ_d^2, is considered statistically independent of the noise $\zeta(n)$, then, $\widehat{e}^{\,2}_{(0)}$ and $\widehat{e}^{\,2}_{(\infty)}$ can be obtained as [19,22]

$$\widehat{e}^{\,2}(0) = \sigma_d^2 + \sigma_\zeta^2\,, \tag{43}$$

where

$$\sigma_d^2 = \|\mathbf{p}_N\|_2^2\sigma_x^2 \tag{44}$$

and

$$\widehat{e}^{\,2}(\infty) = \sigma_x^2 f(\infty) + \sigma_\zeta^2\,. \tag{45}$$

After combining (18) and (30), we have

$$f(\infty) = \Big[1-\tfrac{2}{L(\infty)+2}+\tfrac{1}{L(\infty)+2}\Big]f(\infty) + \Big\{\tfrac{2}{L(\infty)+2}-\tfrac{2}{\big[L(\infty)+2\big]^2}\Big\}E\Big[\|\mathbf{p}''\|_2^2\Big] + \tfrac{\sigma_\zeta^2\,L(\infty)}{\big[L(\infty)+2\big]^2\sigma_x^2}$$

$$= \frac{2\;E\Big[\|\mathbf{p}''\|_2^2\Big]\big[L(\infty)+1\big]\sigma_x^2 + \sigma_\zeta^2\,L(\infty)}{\big[L(\infty)+1\big]\sigma_x^2}\,, \tag{46}$$

where $\mathbf{p}''$ is the un-modelled part of $\mathbf{p}$, which is of small value. We assume for simplicity that $L(\infty) \approx L(\infty+1) \approx L(\infty+2)$ and $\sigma_\zeta^2 >> 2\;E\|\mathbf{p}''\|_2^2\;\sigma_x^2$. With these approximations, (46) is simplified as

$$f(\infty) \approx \frac{\sigma_\zeta^2}{\sigma_x^2}\,. \tag{47}$$

Combining (43)–(45), (47) and substituting into (41), we get

$$L(\infty) = L(0) + \frac{1}{\delta}\,\ln\sqrt{\frac{L(\infty)\,\Big(\|\mathbf{p}_N\|_2^2\sigma_x^2+\sigma_\zeta^2\Big)}{2\;L(0)\sigma_\zeta^2}}\,. \tag{48}$$

Again, considering $\sigma_\zeta^2 >> \|\mathbf{p}_N\|_2^2\sigma_x^2$, we have

$$L(\infty) = L(0) + \frac{1}{\delta}\,\ln\sqrt{\frac{L(\infty)}{2\,L(0)}}\,. \tag{49}$$

From the above equation, it can be observed that, under the assumptions discussed in this section, the steady-state tap length is dependent on the initial tap length $L(0)$ and the decay rate δ.

Table 1
Computational complexity of the proposed algorithm.

Equations	Multiplications	Additions
1. $e(n) = d(n) - \mathbf{w}_{L(n)}^T(n)\,\mathbf{x}_{L(n)}(n)$	L	L
2. $\mu(n) = \dfrac{\rho\,\widehat{e}^{\,2}(n-1) + (1-\rho)e^2(n) - \sigma_\zeta^2 - \left[(1-e^{-2\delta}) - e^{-2L(n-1)\delta} + e^{-2N\delta}\right]\sigma_x^2}{\left\{\left[\rho\,\widehat{e}^{\,2}(n-1) + (1-\rho)e^2(n)\right]\left[L(n-1)+2\right] - 2\sigma_\zeta^2\right\}(1-e^{-2\delta})\sigma_x^2 - 2\sigma_x^4(e^{-2L(n-1)\delta} - e^{-2N\delta})}$	21	12
3. $\mathbf{w}_{L(n)}(n+1) = \mathbf{w}_{L(n)}(n) + \mu(n)\,e(n)\,\mathbf{x}_{L(n)}(n)$	$2L$	$2L-1$
4. $\widehat{e}^{\,2}(n) = \rho\,\widehat{e}^{\,2}(n-1) + (1-\rho)\,e^2(n)$	4	2
5. $L(n) = L(n-1) + \delta^{-1}\,\ln\sqrt{\frac{L(n-1)\,\widehat{e}^{\,2}(n-1)}{L(n-2)\,\widehat{e}^{\,2}(n)}}$	6	1
Total	$3L+31$	$3L+14$

2.4. Computational complexity

The computational complexity of the proposed algorithm in terms of the total number of multiplications and additions is shown in Table 1.

3. Convergence analysis of MSF-based SAEC algorithm

The modeling of stereophonic acoustic echo paths based on the process segmentation method, where a long adaptive filter is decomposed into smaller MSFs, is presented in [5–7]. However, these papers do not present any mathematical convergence analysis of the advanced COEA, which we now present. The concept of decomposing the input and weight vector and partitioning a long adaptive filter into smaller MSFs has been presented in [5,6]. The adaptive algorithm for updating each sub-filter is designed depending upon the generation of the error signal. When we obtain the error signals at each stage of the sub-filters in the MSF design, then it is called the individual-error sub-filter-based adaptive algorithm. On the other hand, when the error signal is obtained at the last stage then it is called the common-error sub-filter-based adaptive algorithm [5]. These algorithms each have their own advantages in either convergence rate or steady-state error. Hence, the COEA [11] was introduced to maintain a trade-off between these two algorithms.

Considering the SAEC configuration shown in Fig. 2, the input signal vector of the two stereo channels can be written as $\mathbf{x}_1(n) = [x_1(n), x_1(n-1), \ldots]^T$ and $\mathbf{x}_2(n) = [x_2(n), x_2(n-1), \ldots]^T$. The long adaptive filter is divided into $2S$ sub-filters, (Ssub-filters per channel) as shown in Fig. 2 for improving the convergence, where AF_idenotes the i th sub-filter, $i = 0, 1, \ldots, 2S-1$. The individual sub-filter outputs of the $2S$ sub-filters are generated as

$$y_i(n) = \mathbf{w}_i^T(n)\mathbf{x}_{p,K}(n) \quad , \quad i = 0, 1, \ldots, 2S-1\,, \tag{50}$$

where $K = 0, \ldots, S-1 \;\forall\; i = 0, \ldots, S-1$ for $p = 1$ and $K = 0, \ldots, S-1 \;\forall\; i = S, \ldots, 2S-1$ for $p = 2$. Thus, the sub-filter outputs $\left[y_0(n), \ldots, y_{S-1}(n)\right]$ are obtained from the inputs of Channel 1 $[x_{1,0}(n), \ldots, x_{1,S-1}(n)]$ and the next S outputs $[y_S(n), \ldots, y_{2S-1}(n)]$ are obtained from the Channel 2 sub-filter inputs $[x_{2,0}(n), \ldots, x_{2,S-1}(n)]$. The COEA for SAEC is shown in Fig. 2, where the acoustic echo path output can be described as

$$d(n) = \mathbf{g}_i^T\,\mathbf{x}_{p,K}(n) + \varepsilon(n)\,, \tag{51}$$

where $\mathbf{g}_i = [g_{i,1}, g_{i,2}, \ldots, g_{i,L_i}]^T$ is the Wiener optimum solution of the i th sub-filter, L_i is the length of the corresponding sub-filter, $\mathbf{x}_{p,K}(n)$ is the input sequence, and $\varepsilon(n)$ is the additive noise, which is assumed to be independent of the input with variance $J_{\min} = E\left[\varepsilon^2(n)\right]$. The error signal generated by the individual-error approach applied to sub-filters is obtained as

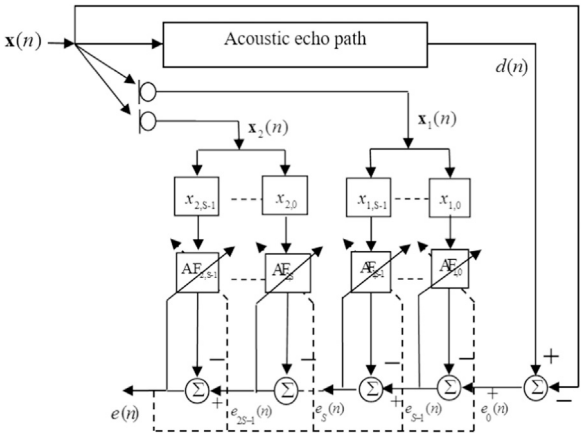

Fig. 2. MSF-based COEA for SAEC [11].

$$e_i(n) = e_{i-1}(n) - \mathbf{w}_i^T(n)\mathbf{x}_{p,K}(n) \quad , \quad i = 1, \ldots, 2S-1\,, \tag{52}$$

where the conditions on $\mathbf{x}_{p,K}(n)$ are the same as in (52), and for $i = 0$,

$$e_0(n) = d(n) - \mathbf{w}_0^T(n)\mathbf{x}_{p,0}(n)\,. \tag{53}$$

Similarly, for the common-error approach applied to the sub-filters, the error signal is given as

$$e(n) = d(n) - \mathbf{w}_i^T(n)\mathbf{x}_{p,K}(n) \quad , \quad i = 0, 1, \ldots, 2S-1, \tag{54}$$

where the conditions on $\mathbf{x}_{p,K}(n)$ are again the same as in (52). In the COEA, the weight adaptation is obtained by combining (52) and (54) as [10]

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \mathbf{x}_i(n)\left[\alpha\,\mu_i\,e_i(n) + (1-\alpha)\mu\,e(n)\right], \quad i = 0, 1, \ldots, 2S-1, \tag{55}$$

where α, $0 \le \alpha \le 1$, is a trade-off parameter that can be chosen as per [9] to achieve the desired performance. In (62) the weight adaptation step size $\mu_i << \mu$ and $e_i(n) << e(n)$. Hence, for the convergence analysis, the error signal is approximated as the common error $e(n)$. Considering the trade-off parameter $\alpha \approx 0$, the above weight update equation can be modified as

$$\mathbf{w}_i(n+1) = \mathbf{w}_i(n) + \mu\,\mathbf{x}_i(n)e(n)\,. \tag{56}$$

We define the weight error vector for the ith sub-filter as

$$\mathbf{H}_i = \mathbf{w}_i(n) - \mathbf{g}_i \tag{57}$$

Then, the MSE under known assumptions is obtained as [20]

$$J = J_{\min} + \sum_{\substack{i=0 \\ i\neq j}}^{2S-1}\sum_{j=0}^{2S-1} E\left[\mathbf{H}_i(n)\right]\mathbf{R}_{i,j}E\left[\mathbf{H}_j(n)\right] + \sum_{j=0}^{2S-1}\mathrm{tr}\left\{R_{i,i}E\left[\mathbf{H}_i(n)\mathbf{H}_i^T(n)\right]\right\}, \tag{58}$$

where $\mathbf{R}_{i,i}$ is the input autocorrelation matrix and $\mathbf{R}_{i,j}$ is the cross-correlation matrix of inputs corresponding to the ith and j^{th} sub-filters. For the MSE to converge, it is necessary that $E\left[\mathbf{H}_i(n)\right]$ and $\mathrm{tr}\left\{\mathbf{R}_{i,i}E\left[\mathbf{H}_i(n)\mathbf{H}_i^T(n)\right]\right\}$ converge. To check the convergence, we subtract $\mathbf{g}_i$ from both sides of (56), substitute (54) for $e(n)$, and take expectation to get

$$E\left[\mathbf{H}_i(n)\right] = (I-\mu_i\mathbf{R}_{i,i})E\left[\mathbf{H}_i(n)\right] - \mu_i\sum_{\substack{i,j=0 \\ i\neq j}}^{2S-1}\mathbf{R}_{i,j}E\left[\mathbf{H}_j(n)\right], \tag{59}$$

which implies that

$$0<\mu_i<\frac{2}{\mathrm{tr}(\mathbf{R}_{i,i})},\quad i=1,2,\ldots,2S-1. \tag{60}$$

This equation shows the essential condition for convergence. To confirm stability, tighter bounds on the step size can be obtained as follows. The autocorrelation matrix of the weight error vector is obtained by post-multiplying (59) by its transposed vector and finding the expectation. Then, considering the analysis in Appendix A, we can ensure the stability and convergence if the step size is chosen such that

$$0<\mu_i<\frac{2}{3\,\mathrm{tr}(\mathbf{R}_{i,i})}. \tag{61}$$

The sub-filters in the MSF-based approach discussed in this section are considered to be of fixed and equal length. If the total number of adaptive coefficients to model the echo path is fixed at L and there are S sub-filters, then each sub-filter will have $\frac{L}{S}$ coefficients. So, the sub-filters converge faster compared to that of a single filter with thousands of coefficients [9]. Though MSF-based decomposition increases the convergence rate, it does not reduce the net filter coefficient requirement to perform the echo cancellation. Normally L is fixed randomly and always at a higher value to avoid undermodelling, and that creates the standing problem of random tap-length initialization in acoustic echo cancellation. Hence, there is a need for an efficient tap-length-optimization algorithm, as discussed earlier, that can be applied to each sub-filter of the MSF framework to find its optimum tap-length. Incorporating the optimum tap-length into each sub-filter, we can then find the net optimum tap length of the overall adaptive filter used in echo cancellation.

4. Simulation and results

For the simulations, only the adaptive filtering portion of the echo canceller is considered in this section. The aim is to examine the proposed optimum tap-length selection algorithm and compare it to existing algorithms, checking the convergence under different SNRs and comparing the results of SLF- and MSF-based SAEC designs with and without application of the variable-tap-length algorithm. As mentioned earlier, following [8–11], some aspects like double-talk detection, multipath estimation, signal decorrelation, and residual echo cancellation are not taken into account for simulating the proposed and existing algorithms. The simulation setup to compare variable-tap-length algorithms is similar to that in [17], and the parameter settings for the MSF-based combined-error SAEC algorithm are similar to those in [9–11]. We examine the derived tap-length-optimization algorithm and compare its performance to that of the existing FT-LMS [14], VT-MSD [8], and VT-VSNLMS$_{\text{VE}}$ [17] algorithms via these simulations. The proposed algorithm is applied to the SLF with a higher number of coefficients and to the MSF with correspondingly smaller sub-filters to compare the difference in convergence, tracking, and echo return loss enhancement (ERLE).

The room impulse response, which is sampled at 8 kHz, is obtained for an enclosure of dimension $10\times12\times10$ ft^3 using a normal loudspeaker and microphone kept at a distance of 1ft. The response of the channel is achieved by applying a white Gaussian stochastic stationary signal with zero mean and unit variance as an input to obtain the measured response. The room impulse response with exponentially-decaying envelope is shown in Fig. 3.

For the adaptive filters used in the SAEC setup, we fix the tap length of the SLF at 1200, and for the MSFs, the number of sub-filters is 4 with each sub-filter containing an equal number of coefficients, namely 300, both with the decay parameter set to 0.005. The noise signal is a white zero-mean random Gaussian sequence with variance 0.1. The results are obtained by averaging over 200 independent runs.

For the SLF design, the proposed and existing tap-length-optimization algorithms are applied with 1200 weights, and the comparison is presented in Fig. 4. It can be noticed that with an increase in the number of iterations, the tap-length requirement increases and reaches a relatively steady value after a few thousand iterations. The proposed variable tap-length SLF algorithm finds the optimum tap length just below 800 taps in comparison to the existing SLF tap-length optimization algorithms, which require higher tap-length, as shown in Fig. 4(a). Similarly, we apply the proposed and existing variable tap-length algorithms for the MSFs and the comparison is presented in Fig. 4(b). We notice that the tap-length requirement of the proposed variable tap-length MSF algorithm is just below 200, which is nearly 50 coefficients less than that in the existing algorithms. Among the existing algorithms, the VT-MSD needs the largest number of coefficients, followed by FT-LMS and VT-VSNLMS$_{\text{VE}}$, respectively. The VT-VSNLMS$_{\text{VE}}$ has improved tap-length convergence and steady-state performance compared to its counterparts according to the analysis made in [17].

For MSE convergence analysis of the SLF, we apply tap-length optimization using the proposed algorithm and the existing VT-VSNLMS$_{\text{VE}}$ algorithm, and compare the results with those of an adaptive SLF with prefixed length of 1200. The results for two different SNR values at 20 dB and 0 dB are shown in Fig. 5(a) and

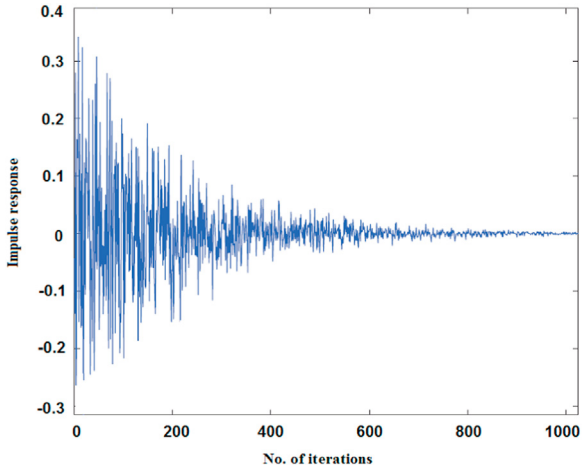

Fig. 3. Room impulse response.

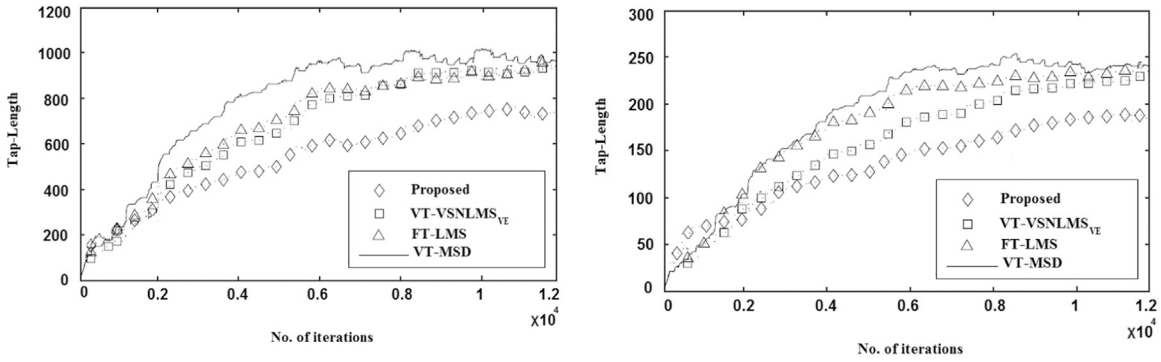

Fig. 4. Tap-length vs. number of iterations for proposed and existing algorithms at SNR = 20 dB. (a) SLF with 1200 taps. (b) MSF with 300 taps per sub-filter.

(b), respectively. The VT-VSNLMS$_{VE}$ achieves faster convergence compared to the adaptive SLF and proposed algorithms, but exhibits poorer steady-state performance. The reason for this is the larger tap length of the adaptive filter used here, whereas VT-VSNLMS$_{VE}$ is designed for filters with a relatively smaller order [17]. The adaptive SLF converges to a low steady-state value, but it takes 8000 and 6000 iterations at 20 dB and 0 dB SNR, respectively. On the other hand, the proposed algorithm converges after only 3000 and 2000 iterations, and also achieves better steady-state performance compared to VT-VSNLMS$_{VE}$.

Similarly, for the MSF design, the proposed variable-tap-length algorithm is used for each 300-tap sub-filter in the COEA to optimize their respective tap-lengths. Above, we had applied the proposed algorithm to the adaptive SLF with 1200 coefficients. To check the difference in convergence in both cases, we plot the MSE with respect to the total number of iterations in Fig. 6(a) and (b) for SNR = 20 dB and 0 dB, respectively. Long length adaptive filters used in echo cancellation are infamous for improper tracking with enclosure dislocations [7].

In Fig. 7, we compare the tracking of the adaptive SLF with and without tap-length optimization, and the MSF-based COEA with optimum tap length, where an enclosure dislocation occurs after 8 s. It can be noticed that tap-length optimization enhances the tracking capability.

Fig. 8 presents the transmitted mixed acoustic signal $d1 + d2$ at the far-end for separate channels which contain signals $d1$ and $d2$, combined in a single-channel, as per the SAEC design shown in Fig. 2. After SAEC, the received separated acoustic signal $r1$and $r2$ along with the error signal $e(n)$ is shown in Fig. 9. This is a simple demonstration of the working of SAEC design. The ERLE is a measure of the amount of echo suppressed by the echo canceller and is defined as the ratio of the original echo power over the residual echo signal power after cancellation, i.e.,

$$\text{ERLE} = 10\ \log_{10}\frac{\text{power of microphone signal}}{\text{power of residual acoustic echo signal}}\ \text{dB}\,.$$

A higher ERLE value indicates better adaptive echo canceller performance. Here, we assume that there is no near-end signal, so only the echo signal is present for ERLE measurement. We check the ERLE for all the existing and proposed variable-tap-length algorithms applied to the adaptive SLF echo canceller for the received signal shown in Fig. 9(b), and the results are shown in Fig. 10. We list the maximum, minimum, and average ERLE in Table 2 along with the average execution time. This clearly depicts comparatively better echo cancellation performance of the proposed variable-tap-length algorithm with only a slightly higher execution time.

Finally, we compare the ERLE of the MSF-based COEA with that of the adaptive SLF using the proposed tap-length optimization, and the result is shown in Fig. 11. The ERLE performance of the COEA-based echo canceller with tap-length optimization is less than that of a SLF echo canceller. On the other hand, it provides faster convergence, which is observed from Fig. 6, and is a better choice for stereophonic echo cancellers employing long-tap-length adaptive filters. Hence, it can be said that the COEA with optimum tap length achieves better convergence compared to the variable-tap-length adaptive SLF. We achieve this faster convergence with a minimum tap-length requirement that brings

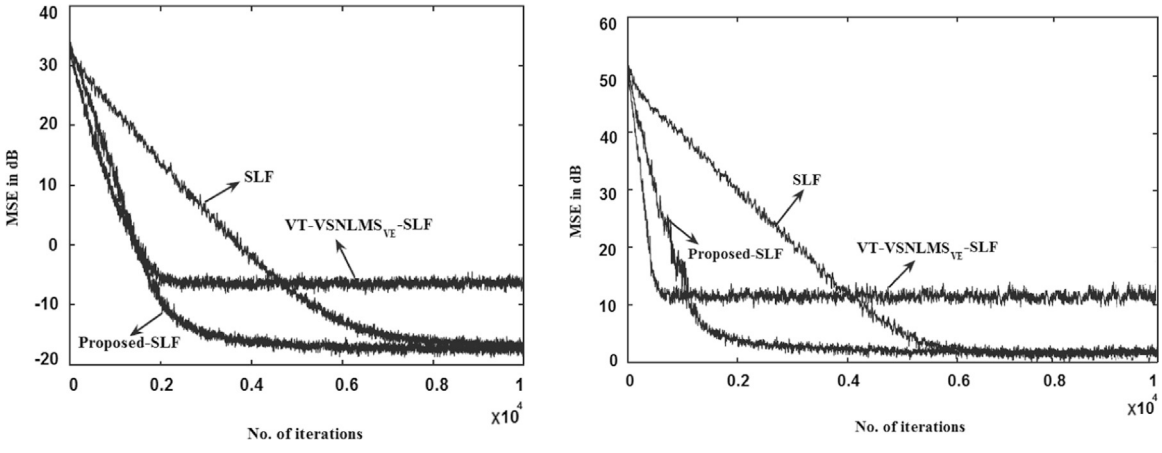

Fig. 5. MSE vs. number of iterations for adaptive SLF convergence analysis with and without tap-length optimization. (a) SNR = 20 dB. (b) SNR = 0 dB.

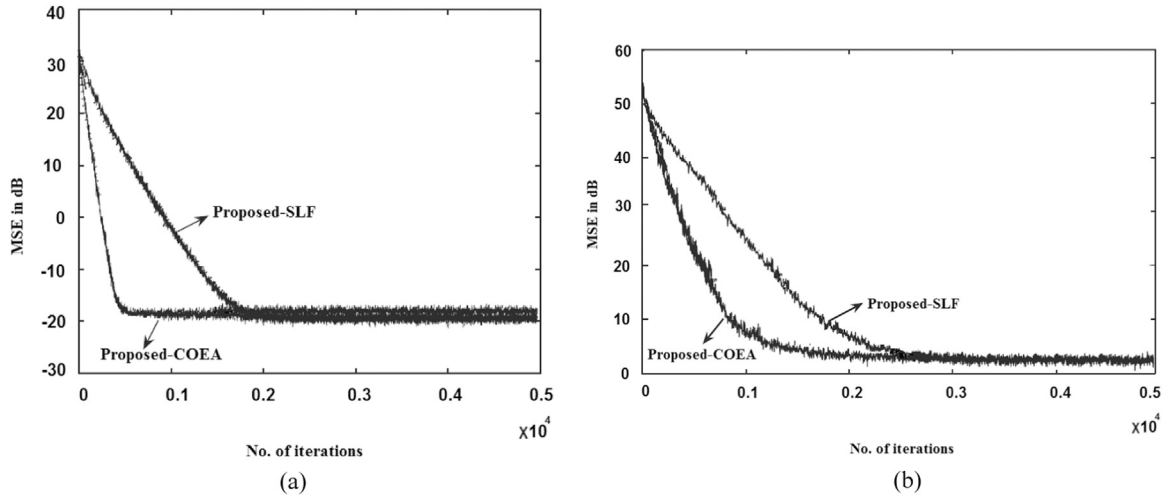

Fig. 6. MSE vs. number of iterations for adaptive SLF and MSF-based COEA using the proposed tap-length optimization. (a) SNR=20 dB. (b) SNR=0 dB.

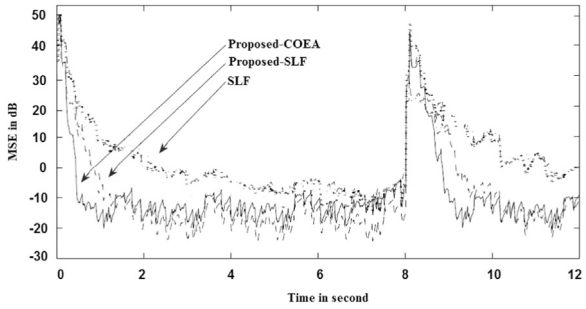

Fig. 7. Tracking performance comparison between adaptive SLF with and without the proposed tap-length optimization, and COEA with the proposed tap-length optimization.

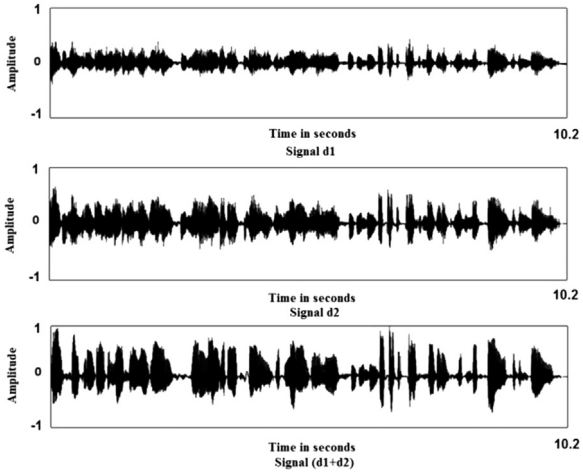

Fig. 8. Mixed signals at the far end. (a)$d1$. (b) $d2$. (c)$d1 + d2$.

down the TDL design complexity using FIR adaptive filters for echo cancellation. Fewer coefficients result in fewer computations, so the computational complexity also decreases when dealing with fewer coefficients.

5. Conclusion

As compared to existing variable-tap-length algorithms, the proposed tap-length-optimization algorithm models a system

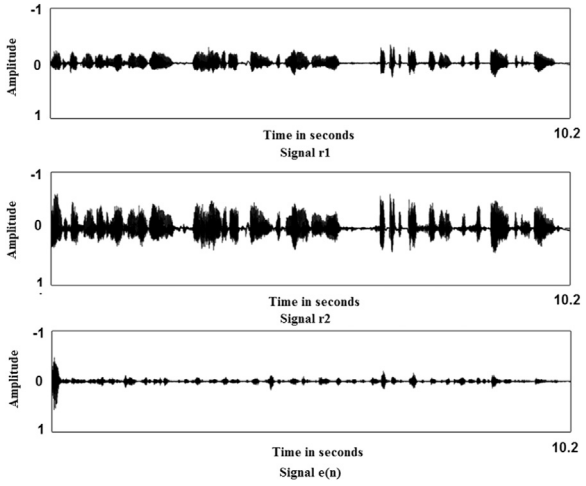

Fig. 9. Separated received signals at the near end. (a) $r1$. (b) $r2$. (c) Error signal, $e(n)$.

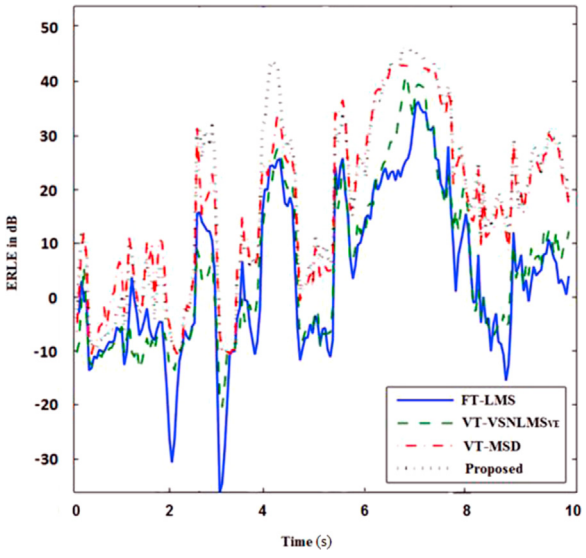

Fig. 10. ERLE performance comparison of proposed and existing tap-length-optimization algorithms for SLF echo cancellation.

with fewer filter coefficients and shows improved convergence speed and better steady-state behaviour under high and low SNRs. The new tap-length optimization has been applied in the SAEC

Table 2
Comparison of ERLE in dB and execution time for all discussed tap-length-optimization algorithms applied to SLF-SAEC.

Variable tap-length algorithm	Maximum ERLE (dB)	Minimum ERLE (dB)	Average ERLE (dB)	Execution time (sec)
FT-LMS [14]	35.5	−35	11.8	138.0
VT-MSD [8]	38.5	−20.02	18.6	142.1
VT-VSNLMS$_{VE}$ [17]	41.7	−10.8	21.2	147.5
Proposed	46.4	−10	24.5	155.4

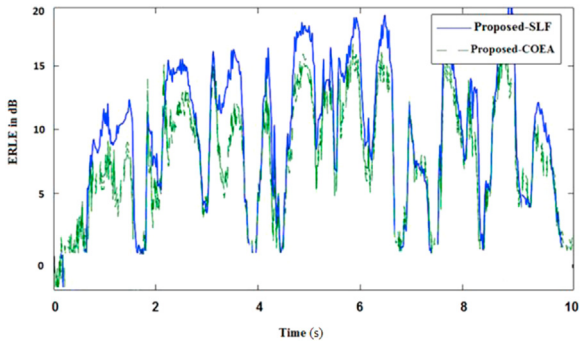

Fig. 11. ERLE performance comparison of adaptive SLF and MSF-based COEA using the proposed tap-length-optimization algorithm.

framework with and without MSF, and comparisons made, which suggest selection of the MSF-based COEA with variable tap length for better convergence of stereophonic echo cancellers. The tap-length-optimization algorithm may further be analyzed for applications like feedback cancellation, channel equalization, and noise reduction. A new analysis with notable performance can also be carried out considering the effects of various acoustic artefacts affecting the echo cancellation and removing the often used assumption of white Gaussian noise and signal decorrelation. The algorithm can also be made robust for applications severely affected by a high noise environment where the SNR is below −20 dB.

Acknowledgment

This work is partially funded by the European Union's Seventh Framework Programme (FP7/2007–2013) under Grant agreement no. ITN-GA-2012-316969. The authors would like to thank the anonymous reviewers for their constructive comments that have allowed improving the proposed approach and the quality of the work. The authors would also like to thank Prof. Søren Holdt Jensen and Prof. Jan Ostergaard of Department of Electronic Systems, Aalborg University, Denmark for their valuable suggestions during this work.

Appendix A

Derivation of step-size stability condition

If the input autocorrelation matrix $\mathbf{R}_{i,i}$ is considered to be symmetric, then we can have a unitary transform $\mathbf{\Lambda}$ with $\mathbf{\Lambda}^{-1} = \mathbf{\Lambda}^T$ that expresses the eigenvalues of $\mathbf{R}_{i,i}$ as

$$\mathbf{\Lambda} R_{i,i} \mathbf{\Lambda}^T = \mathbf{\Upsilon} = \mathrm{diag}\left[\lambda_1, \lambda_2 \ldots, \lambda_L\right], \tag{A1}$$

where $\left[\lambda_1, \lambda_2, \ldots, \lambda_L\right]$ are the eigenvalues of $\mathbf{R}_{i,i}$. We consider

$$\mathbf{M}(n) = \mathbf{\Lambda}\mathbf{H}_i = \mathbf{\Lambda}\left[\mathbf{w}_i(\mathbf{n}) - \mathbf{g}_i\right], \tag{A2}$$

$$\mathbf{V}(n) = E\left[\mathbf{M}(n)\right], \tag{A3}$$

and

$$\mathbf{Q}(n) = E\left[\mathbf{M}(n)\mathbf{M}^T(n)\right]. \tag{A4}$$

Assuming that the input signal is white Gaussian noise with zero mean and all data vectors are independent of each other, we can apply the Horowitz theorem [23] to get

$$\mathbf{Q}(n+1) = \mathbf{Q}(n) + 8\,\mu^2\,\mathbf{\Upsilon}\,\mathbf{Q}(n)\,\mathbf{\Upsilon} + 4\,\mu^2\,\mathrm{tr}\left[\mathbf{\Upsilon}\,\mathbf{Q}(n)\right]\mathbf{\Upsilon} + 4\,\mu^2\,\varpi\,\mathbf{\Upsilon} - 4\,\mu\,\mathbf{\Upsilon}\,\mathbf{Q}(n), \tag{A5}$$

where

$$\varpi = E\left[d^2(n)\right] - \mathbf{P}^T\mathbf{R}^{-1}\mathbf{P}, \tag{A6}$$

and

$$\mathbf{P} = E\left[d(n)\mathbf{x}(n)\right], \tag{A7}$$

where we use $\mathbf{x}_{p,K}(n) = \mathbf{x}(n)$ and $\mathbf{R}_{i,i} = \mathbf{R}$ for notational convenience. Following the analysis in [23], we modify (A5) to scalar equation form as

$$Q_{a,a}(n+1) = \phi_{a,a}\,Q_{a,a}(n) + 4\,\mu^2\left[\lambda_a \sum_{r=1}^{L} \lambda_r\,Q_{r,r}(n) + \varpi\,\lambda_a\right], \tag{A8}$$

and for $a \neq b$, we have

$$Q_{a,b}(n+1) = \phi_{a,b}\,Q_{a,b}(n) \tag{A9}$$

and

$$\phi_{a,b} = 1 + 8\,\mu^2\lambda_a\,\lambda_b - 2\,\mu\,(\lambda_a + \lambda_b), \tag{A10}$$

where $Q_{a,b}(n)$ and $\phi_{a,b}$ are the $(a, b)^{\text{th}}$ elements of matrices $\mathbf{Q}(n)$ and $\boldsymbol{\phi}(n)$, respectively. The non-diagonal elements of $\mathbf{Q}(n)$ are said to converge only after the diagonal elements have converged. Concatenating (A8) into vector form we obtain

$$\boldsymbol{\chi}(n+1) = \mathbf{D}\,\boldsymbol{\chi}(n) + 4\mu^2\varpi\bar{\boldsymbol{\lambda}}, \tag{A11}$$

where $\boldsymbol{\chi}(n) = \left[Q_{1,1}(n), Q_{2,2}(n), \ldots, Q_{L,L}(n)\right]^T$, $\mathbf{D} = \left[\phi_{1,1}, \phi_{2,2}, \ldots, \phi_{L,L}\right]$, and $\bar{\boldsymbol{\lambda}} = \left[\lambda_1, \lambda_2, \ldots, \lambda_L\right]^T$. The convergence of $\boldsymbol{\chi}(n)$ depends on the eigenvalues of $\mathbf{D}$: if all the eigenvalues of $\mathbf{D}$ lie inside the unit circle, then $\boldsymbol{\chi}(n)$ can be said to be stable. To find the eigenvalues of $\mathbf{D}$, we solve

$$\det\left[\mathbf{D} - \beta\mathbf{I}\right] = 0, \tag{A12}$$

which implies that

$$\left[\sum_{b=1}^{L}(\phi_{b,b} - \beta)\right]\left\{1 + 4\mu^2\,\bar{\boldsymbol{\lambda}}\;\mathrm{diag}\left[\frac{1}{\phi_1 - \beta}, \frac{1}{\phi_2 - \beta}, \ldots, \frac{1}{\phi_L - \beta}\right]\bar{\boldsymbol{\lambda}}^T\right\} = 0, \tag{A13}$$

i.e.,

$$\left[(\phi_{1,1}-\beta)(\phi_{2,2}-\beta)\ \ldots\ (\phi_{L,L}-\beta)\right]\left[1+4\mu^2\sum_{b=1}^{L}\frac{\lambda_b^2}{\phi_{b,b}-\beta}\right] = 0\ . \tag{A14}$$

Approximating the above equation for simplification, we obtain

$$\det\left[\mathbf{D}-\beta\mathbf{I}\right] = 1+4\mu^2\sum_{b=1}^{L}\frac{\lambda_b^2}{\phi_{b,b}-\beta} = 0\ . \tag{A15}$$

The poles of $\det\left[\mathbf{D}-\beta\mathbf{I}\right]$ are at $\phi_{b,b}$, and from (A10) we obtain [24]

$$\begin{aligned}\phi_{b,b} &= 1+8\mu^2\lambda_b^2-4\mu\lambda_b\\ &=(1-2\mu\lambda_b)^2+4\mu^2\lambda_b^2\ .\end{aligned} \tag{A16}$$

From (A16), it is observed that all the poles $\phi_{b,b}$ are positive and

$$\frac{d\left\{\det\left[\mathbf{D}-\beta\mathbf{I}\right]\right\}}{d\beta} = 4\mu^2\sum_{b=1}^{L}\frac{\lambda_b^2}{(\phi_{b,b}-\beta)^2} > 0\ . \tag{A17}$$

Between each pair of consecutive poles $\phi_{b,b}$ of $\det\left[\mathbf{D}-\beta\mathbf{I}\right]$, there is a zero β in-between, i.e., $0<\phi_{1,1}<\beta_1<\phi_{2,2}<\beta_2<\ldots<\phi_{L,L}<\beta_L$. If we require $\beta<1$, then all eigenvalues of $\mathbf{D}$ will be located inside the unit circle. But $\beta<1$ if and only if $\phi_{L,L}<1$ and

$$1+4\mu^2\sum_{b=1}^{L}\frac{\lambda_b^2}{\phi_{b,b}-\beta} > 0 \tag{A18}$$

We substitute (A16) into (A18) to get the necessary and sufficient condition on μ

$$\sum_{b=1}^{L}\frac{\mu\lambda_b}{1-2\mu\lambda_b} < 1\ . \tag{A19}$$

If we determine μ as the solution of

$$\sum_{b=1}^{L}\frac{\mu\lambda_b}{1-2\mu\lambda_b} = 1\ , \tag{A20}$$

then, assuming $\lambda_1<\lambda_2<\ldots<\lambda_L$, we have [24]

$$0<\mu<\frac{1}{2\lambda_{L'}},\quad L'=1,2,\ldots,L\ . \tag{A21}$$

Let $\mu_1<\mu_2<\ldots<\mu_L$ be the solutions of (A20), so that we need to find a lower bound on μ_1. For that, we express (A20) as

$$\mu^{-L}-m_1\mu^{1-L}+m_2\mu^{2-L}+\ldots+(-1)^L m_L = 0\ . \tag{A22}$$

By comparing (A22) to (A20), we obtain

$$m_1 = 3\sum_{b=1}^{L}\lambda_b, \tag{A23}$$

and

$$m_2 = 4\sum_{a}^{L}\sum_{b}^{L}\lambda_a\lambda_b\ . \tag{A24}$$

As per the analysis mentioned in [25], μ_1^{-1} is the largest root of (A22) and can be bounded as

$$\frac{1}{\mu_1}\le\frac{1}{\mu}=\frac{g_1+\sqrt{(L-1)(Lg_2-g_1^2)}}{L}\ , \tag{A25}$$

where

$$g_1=\sum_{b=1}^{L}\frac{1}{\mu_b}\ , \tag{A26}$$

and

$$g_2=\sum_{b=1}^{L}\left(\frac{1}{\mu_b}\right)^2\ . \tag{A27}$$

Following the analysis in [25], we obtain

$$g_1=m_1=\sum_{b=1}^{L}\frac{1}{\mu_b}\ , \tag{A28}$$

and

$$g_2=m_1^2-2m_2\ . \tag{A29}$$

Substituting (A28) and (A29) into (A25), we get

$$\mu_1\ge\frac{L}{\left\{3\sum_{b=1}^{L}\lambda_b+\left(\sum_{b=1}^{L}\lambda_b\right)^2\left[(3L-3)^2-8L^2+8\right]\right\}^{\frac{1}{2}}}\ . \tag{A30}$$

So, the choice of step size to ensure convergence should be

$$0<\mu\le\mu_1\ . \tag{A31}$$

From the analysis of the LMS algorithm in [20,21], we know that

$$0<\mu<\frac{1}{\lambda_{\max}}\ . \tag{A32}$$

Now, comparing (A30) and (A32) and adopting the assumptions made in [20], we have

$$\mu_1\ge\frac{1}{3\sum_{b=1}^{L}\lambda_b}\ . \tag{A33}$$

But,

$$\sum_{b=1}^{L}\lambda_b\approx\mathrm{tr}\left(\mathbf{R}_{i,i}\right)\ . \tag{A34}$$

So, finally we have

$$0<\mu\le\frac{1}{3\ \mathrm{tr}\left(\mathbf{R}_{i,i}\right)}\ . \tag{A35}$$

References

[1] J. Benesty, Y. Huang, Adaptive Signal Processing Applications to Real World Problems, Springer, Berlin, 2003.

[2] J. Benesty, D.R. Morgan, M.M. Sondhi, A better understanding and an improved solution to the specific problems of stereophonic acoustic echo cancellation, IEEE Trans. Speech Audio Process. 6 (2) (1998) 156–165, http://dx.doi.org/10.1109/89.661474.

[3] C. Breining, P. Dreiscitel, E. Hansler, A. Mader, B. Nitsch, H. Puder, T. Schertler, G. Schmidt, J. Tilp, Acoustic echo control—an application of very high order adaptive filters, IEEE Signal Process. Mag. 16 (4) (1999) 42–69, http://dx.doi.org/10.1109/79.774933.

[4] M.M. Sondhi, D.R. Morgan, J.L. Hall, Stereophonic acoustic echo cancellation - an overview of the fundamental problem, IEEE Signal Process. Lett. 2 (8) (1995) 148–151, http://dx.doi.org/10.1109/97.404129.

[5] R.N. Sharma, A.K. Chaturvedi, G. Sharma, Tracking behavior of acoustic echo canceller using multiple sub-filter, in: Proceedings of the 14th European Signal Processing Conference, Florence, Italy, 2006.

[6] A. Barik, G. Murmu, T.P. Bhardwaj, R. Nath, LMS adaptive multiple sub-filters based acoustic echo cancellation, in: Proceedings of IEEE International Conference on Computer and Communication Technology 2010, pp. 824–827.

[7] R. Nath, Adaptive echo cancellation based on a multipath model of acoustic channel, Circuit Syst. Signal Process. 32 (4) (2013) 1673–1698, http://dx.doi.org/10.1007/s00034-012-9529-4.

[8] C. Schüldt, F. Lindstromb, H. Li, I. Claesson, Adaptive filter length selection for acoustic echo cancellation, Signalprocessing 89 (1185–1194) (2009), DOI: /10.1016(/j.sigpro.2008.12.023).
[9] A. Kar, A. Barik, R. Nath, An improved order estimation of MSF for stereophonic acoustic echo cancellation, in: Proceedings International Conference on Information System Design and Intelligent Applications, 2012, pp. 319–327.
[10] A. Kar, M. Chandra, A novel variable tap-length learning algorithm for low complexity, fast converging stereophonic acoustic echo cancellation, Int. J. Inf. Commun. Technol. 6 (3) (2014) 309–325, http://dx.doi.org/10.1504/IJICT.2014.063219.
[11] A. Kar, M. Chandra, An improved variable structure adaptive filter design and analysis for acoustic echo cancellation, Radio Eng. 24 (1) (2015) 252–261, http://dx.doi.org/10.13164/re.2015.0252.
[12] Z. Pritzker, A. Feuer, Variable length stochastic gradient algorithm, IEEE Trans. Signal Process. 39 (4) (1991) 997–1001, http://dx.doi.org/10.1109/78.80936.
[13] A.H. Sayed, Fundamentals of Adaptive Filtering, Wiley, New York, 2003.
[14] Y. Gong, C.F.N. Cowan, An LMS style variable tap-length algorithm for structure adaptation, IEEE Trans. Signal Process. 53 (7) (2005) 2400–2407, http://dx.doi.org/10.1109/TSP.2005.849170.
[15] L. Li, J.A. Chambers, A novel adaptive leakage factor scheme for enhancement of a variable-taplength learning algorithm, in: Proceedings of the IEEE International Conference on Acoustics, (Speech and Signal Processing), 2008, pp. 3837–3840.
[16] A. Kar, M. Chandra, Pseudo-fractional Tap-length Learning Based Applied Soft Computing for Structure Adaptation of LMS in High Noise Environment. Soft Computing Techniques in Engineering Applications, Springer-Verlag, Germany 2014, pp. 115–129, http://dx.doi.org/10.1007/978-3-319-04693-8_8.
[17] A. Kar, M. Chandra, Performance evaluation of a new variable tap-length learning algorithm for automatic structure adaptation in linear adaptive filters, AEU-Int. J. Electron. Commun. 69 (1) (2015) 253–261, http://dx.doi.org/10.1016/j.aeue.2014.09.010.
[18] Y. Gu, K. Tang, H. Cui, W. Du, Convergence analysis of a deficient-length LMS filter and optimal-length sequence to model exponential decay impulse response, IEEE Signal Process. Lett. 10 (1) (2003) 4–7, http://dx.doi.org/10.1109/LSP.2002.806704.
[19] B. Widrow, S.D. Stearns, Adaptive Signal Processing, Prentice Hall Inc, Englewood Cliffs, NJ, 1985.
[20] B. Farhang-Boroujeny, Adaptive Filters: Theory and Applications, Wiley, New York, 1998.
[21] K. Shi, X. Ma, G. Tong Zhou, A variable step size and variable tap-length LMS algorithm for impulse response with exponential power profile, in: Proceedings of IEEE International Conference on Acoustics, (Speech and Signal Processing), 2009, pp. 3105–3108.
[22] D.C. Chang, F.T. Chu, Feedforward active noise control with a new variable tap-length and step-size filtered-x LMS algorithm, IEEE/ACM Transact Audio, Speech Lang. Process. 22 (2) (2014) 542–554, http://dx.doi.org/10.1109/TASLP.2013.2297016.
[23] L.I. Horowitz, K.D. Senne, Performance analysis of complex LMS for controlling narrow-band adaptive arrays, IEEE Trans. Circuits Syst. 28 (6) (1981) 562–576, http://dx.doi.org/10.1109/TCS.1981.1085024.
[24] A. Feuer, E. Weinstein, Convergence analysis of LMS filters with uncorrelated Gaussian data, IEEE Trans. Acoust., Speech Signal Process. 33 (1) (1985) 222–230, http://dx.doi.org/10.1109/TASSP.1985.1164493.
[25] O.L. Jagerman, Non-stationary blocking in telephone traffic, Bell Syst. Tech. J. 54 (3) (1975) 625–661, http://dx.doi.org/10.1002/j.1538-7305.1975.tb02858.x.