PAPER Special Section on Digital Signal Processing

A New Class of Acoustic Echo Cancelling by Using Correlation LMS Algorithm for Double-Talk Condition

Rui CHEN[†a)], ***Student Member***, **Mohammad Reza ASHARIF**[†], ***Member***, **Iman TABATABAEI ARDEKANI**[†], ***Nonmember***, ***and*** **Katsumi YAMASHITA**[††], ***Member***

SUMMARY The conventional algorithms in the echo canceling system have drawback when they are faced with double-talk condition in noisy environment. Since the double-talk and noise signal are exist, then the error signal is contaminated to estimate the gradient correctly. In this paper, we define a new class of adaptive algorithm for tap adaptations, based on the correlation function processing. The computer simulation results show that the Correlation LMS (CLMS) and the Extended CLMS (ECLMS) algorithms have better performance than conventional LMS algorithm. In order to implement the ECLMS algorithm, the Frequency domain Extended CLMS (FECLMS) algorithm is proposed to reduce the computational complexity. However the convergence speed is not sufficient. In order to improve the convergence speed, the Wavelet domain Extended CLMS (WECLMS) algorithm is proposed. The computer simulation results support the theoretical findings and verify the robustness of the proposed WECLMS algorithm in the double-talk situation.
key words: *echo canceling, LMS algorithm, double-talk, correlation function, frequency domain, wavelet domain*

1. Introduction

Adaptive FIR filters by using the conventional LMS or NLMS algorithms [1] are very popular for their simplicity and predictable, and therefore these adaptive filter algorithms are utilized for echo canceling. However, in the noisy double-talk environment when both the near-end and the far-end signals are presented, the error signal used for tap adaptations will be uncorrelated with the echo signal and therefore, tap adaptations processes are severely damaged.

The conventional algorithm usually stops adaptation whenever double-talk sensor detects this condition and it keeps freezing the tap coefficient data during the double-talk condition. Stopping the tap adaptation is just a passive action to handle the double-talk condition and it causes lowering speed of adaptations and/or totally mislead when the echo path changed in the period of halting tap adaptation. Other works for challenging the problem of double-talk situation in the echo canceling can be found in [2], [3] and [4] that cause much more complexity adding to a simple LMS algorithm. In this paper, we introduce a new class of algorithm to continue the adaptation even in the presence of double-talk without freezing taps and/or misleading the performance. The proposed method is called correlation LMS (CLMS) algorithm [5], which utilizes the correlation functions of the input signal instead of the input signal itself, to process and find the echo path impulse response. The idea behind this is that we suppose the far-end signal is not correlated with the near-end signal. So the gradient for tap adaptation that is obtained from autocorrelation function does not carry the undesired near-end signal to misadjust the adaptive digital filter for echo path identification. The simulation results show that the CLMS algorithm outperforms the LMS algorithm when the double-talk signal is existing. In this class, the Extended CLMS (ECLMS) [6] is defined to improve the performance of the CLMS in which the MSE is obtained by the sum of lagged squared errors. And also, the Frequency domain Extended CLMS (FECLMS) algorithm [7] is defined to reduce the computational complexity of the ECLMS algorithm. Although the results of computer simulation show the improvement of the performance, the convergence speed is not enough. In order to improve the convergence speed, the Wavelet domain Extended CLMS (WECLMS) [8] algorithm is proposed. The computer simulation results support the theoretical findings and verify the robustness of the proposed WECLMS algorithm in the double-talk situation. This paper is organized as follows. In Sect. 2 a class of adaptive algorithms based on the correlation function are proposed. In Sect. 2.1 the noisy double-talk condition in echo canceller is explained. In Sect. 2.2 CLMS algorithm is presented. In Sect. 2.3 the extended CLMS algorithm is presented. The frequency domain extended CLMS algorithm is derived in Sect. 2.4. Section 2.5 presents the wavelet domain structure for the proposed new class of algorithm. Section 2.6 shows the computational complexity of the proposed algorithms. Then the simulation results appear in Sect. 3. The conclusion to this work is given in Sect. 4.

2. The Proposed Algorithm

2.1 Double-Talk Condition

In the echo canceling system shown in Fig. 1, the acoustic impulse response of the teleconference room is estimated by an adaptive algorithm such as LMS algorithm. The output of the FIR filter, $\widetilde{y}(n)$, is presented by:

Manuscript received November 27, 2003.
Manuscript revised March 2, 2004.
Final manuscript received April 16, 2004.
[†]The authors are with the Department of Information Engineering, Faculty of Engineering, University of the Ryukyus, Okinawa-ken, 903-0213 Japan.
[††]The author is with the Department of Electrical and Electronic Systems, Graduate School of Engineering, Osaka Prefecture University, Sakai-shi, 599-8531 Japan.
a) E-mail: k038656@eve.u-ryukyu.ac.jp

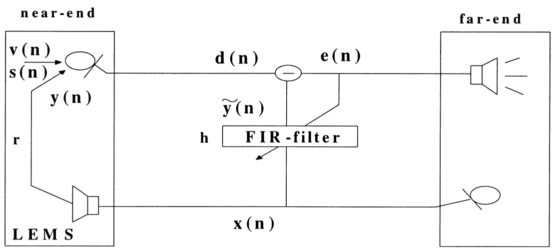

Fig. 1 Echo canceller system.

$$\widetilde{y}(n) = \sum_{i=0}^{N-1} h_i x(n-i) \tag{1}$$

where N is the number of tap, h is the tap coefficient of the adaptive FIR filter and $x(n)$ is the far-end signal at sample n.

The echo signal is obtained from echo impulse response, r, as follows (N is the acoustic impulse response length):

$$y(n) = \sum_{i=0}^{N-1} r_i x(n-i) \tag{2}$$

The error signal, $e(n)$, is calculated as below:

$$e(n) = d(n) - \widetilde{y}(n) \tag{3}$$

where $d(n)$ is microphone signal that usually contains the echo signal. The LMS algorithm is as follows:

$$h_i(n+1) = h_i(n) + 2\mu_0 e(n)x(n-i) \tag{4}$$

where μ_0 is the step size for tap coefficients adaptation. If the near-end signal $s(n)$ and the noise signal $v(n)$ from the near-end, are also presented during the echo canceling, then the microphone signal contains both the echo and the noisy near-end signals:

$$d(n) = y(n) + s(n) + v(n) \tag{5}$$

We call this condition as double-talk condition in the noisy environment. It is well known that the error signal in this case contains uncorrelated component with input and echo signals. Therefore, the algorithm in (4) is failed to track the correct echo impulse response.

2.2 The CLMS Algorithm

In Fig. 2, our new structure is shown [5]. In this structure, we assume the double-talk exists. Since the new structure is based on the processing of autocorrelation function of the input signal (Loudspeaker in the near-end) and the cross-correlation of the input and microphone signal, therefore we should, first, estimate them. The autocorrelation function for the input signal data, $x(n)$ with time-lag k, is defined as below:

$$R_{xx}(n,k) = \sum_{j=0}^{n} x(j)x(j-k) \tag{6}$$

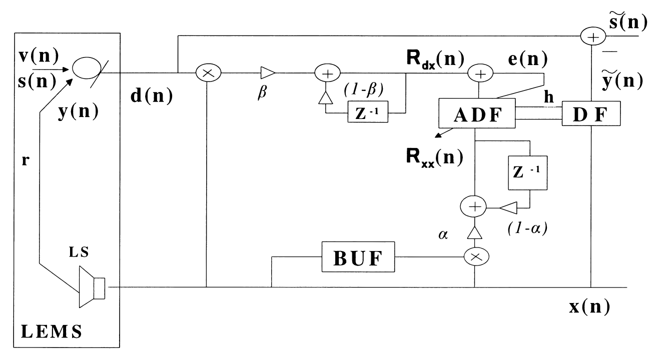

Fig. 2 Correlation LMS algorithm.

Also the cross-correlation between the desired and the input signal is calculated as follows:

$$R_{dx}(n,k) = \sum_{j=0}^{n} d(j)x(j-k) \tag{7}$$

Substituting from (2), (5) and (6) into (7) and assuming that there is no correlation between the far-end and the near-end signals [5] $R_{sx}(n,k) \approx 0$, and also there is no correlation between the far-end signal and the near-end noise signal $R_{vx}(n,k) \approx 0$, the cross-correlation will be obtained as follows:

$$R_{dx}(n,k) \cong \sum_{i=0}^{N-1} r_i R_{xx}(n,k-i) \tag{8}$$

To estimate $R_{dx}(n,k)$, we need to process the autocorrelation values of the input by an adaptive filter. It can be defined as follows:

$$\widetilde{R}_{dx}(n,0) \cong \sum_{i=0}^{N-1} h_i(n) R_{xx}(n,i) \tag{9}$$

where, $\widetilde{R}_{dx}(n,0)$ is the output of the filter which is estimation of the cross-correlation for time-lag $k=0$. The mean squared error (MSE) between the desired cross-correlation function $R_{dx}(n,0)$ and its estimated value $\widetilde{R}_{dx}(n,0)$ (assuming only for the mean component $k=0$) is defined as:

$$J = E[e^2(n)] \tag{10}$$

where

$$e(n) = R_{dx}(n) - \widetilde{R}_{dx}(n) \tag{11}$$

The gradient vector of MSE is:

$$\begin{aligned} \widehat{\nabla} J &= \frac{\partial J}{\partial h} = -2E \begin{bmatrix} e(n)R_{xx}(n,0) \\ e(n)R_{xx}(n,1) \\ \vdots \\ e(n)R_{xx}(n,N-1) \end{bmatrix} \\ &= -2E[e(n)P_{xx}(n)] \end{aligned} \tag{12}$$

Then we obtain the steepest descent algorithm as follows:

$$h(n+1) = h(n) + 2\mu\, E[e(n)P_{xx}(n)] \tag{13}$$

where

$$h(n) = [h_0(n), h_1(n), \cdots, h_{N-1}(n)]^T$$
$$P_{xx}(n) = [R_{xx}(n,0), R_{xx}(n,1), \cdots, R_{xx}(n,N-1)]^T$$

As with LMS algorithm, here we substitute the instantaneous MSE instead of its statistical expectation. The correlation LMS (CLMS) algorithm, which is normalized to the power of the input correlation function to ensure sufficient conditions for convergence, then becomes:

$$h(n+1) = h(n) + \frac{2\mu_0}{1 + P_{xx}^T(n)P_{xx}(n)} e(n)P_{xx}(n) \qquad (14)$$

where μ_0 is the step size for tap coefficients adaptation. It will be shown in simulation section that the CLMS algorithm has good performance comparing with the LMS algorithm in the noisy double-talk situation. However, the CLMS algorithm does not give a sufficient convergence characteristic yet. Then, we extend the CLMS algorithm in order to obtain a sufficient convergence characteristic.

2.3 The Extended CLMS Algorithm

In extended CLMS algorithm [6], we assume the double-talk condition exists. The autocorrelation function and the cross-correlation function are given by Eqs. (6) and (7), respectively. Also we assume that there is no correlation between the far-end and the near-end signals. In the extended CLMS algorithm, we estimate all components of the cross-correlation. Therefore, based on Eq. (8), the output of the adaptive filter is defined here by:

$$\widetilde{R}_{dx}(n,k) = \sum_{i=0}^{N-1} h_i(n) R_{xx}(n, k-i) \qquad (15)$$

where $\widetilde{R}_{dx}(n,k)$ is the estimation value of $R_{dx}(n,k)$. In contrast with Eq. (9) that only the main component of the cross-correlation was estimated, in Eq. (15), we try to estimate all lags up to N. In contrast with the cost function in the CLMS algorithm, the cost function in the ECLMS algorithm is defined by the sum of the lagged squared errors as follows:

$$J = E[\mathbf{e}^T(n)\mathbf{e}(n)] \qquad (16)$$

where the error signal vector is shown by:

$$\mathbf{e}(n,k) = [e(n,0), e(n,1), \cdots e(n,N-1)]^T \qquad (17)$$

with

$$\mathbf{e}(n,k) = R_{dx}(n,k) - \widetilde{R}_{dx}(n,k) \qquad (18)$$

The gradient vector of MSE is:

$$\begin{aligned} \widehat{\nabla} J &= \frac{\partial}{\partial h} E\left[\mathbf{e}^T(n)\mathbf{e}(n)\right] \\ &= -2E\left[Q_{xx}(n)\mathbf{e}(n)\right] \end{aligned} \qquad (19)$$

where

$$Q_{xx}(n) = \begin{bmatrix} R_{xx}(n,0) & R_{xx}(n,1) & \cdots & R_{xx}(n,N-1) \\ R_{xx}(n,1) & R_{xx}(n,0) & \cdots & R_{xx}(n,N-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_{xx}(n,N-1) & R_{xx}(n,N-2) & \cdots & R_{xx}(n,0) \end{bmatrix}$$

Here $Q_{xx}(n)$ is a Toeplitz matrix. Therefore we obtain the steepest descent algorithm as follows:

$$h(n+1) = h(n) + 2\mu E\left[Q_{xx}(n)\mathbf{e}(n)\right] \qquad (20)$$

As like as the LMS algorithm, here we substitute the instantaneous MSE instead of its statistical expectation. The adaptation for ECLMS algorithm, which is normalized to the power of the input correlation function to ensure sufficient conditions for convergence, then becomes:

$$h(n+1) = h(n) + \frac{2\mu_0 Q_{xx}(n)\mathbf{e}(n)}{1 + tr[Q_{xx}(n)Q_{xx}(n)]} \qquad (21)$$

where μ_0 is the step size for tap coefficients adaptation and $tr[\cdot]$ means the trace operator. In order to adapt the tap coefficients according to the ECLMS algorithm, we need to compute $R_{xx}(n,k)$ and $R_{dx}(n,k)$. Then we have used the following recursion formulas for these computations:

$$R_{xx}(n,i) = (1-\alpha)R_{xx}(n-1,i) + \alpha x(n)x(n-i) \qquad (22)$$
$$R_{dx}(n,i) = (1-\beta)R_{dx}(n-1,i) + \beta d(n)x(n-i) \qquad (23)$$

where α and β are limited to $0 < \alpha, \beta < 1$.

In the CLMS algorithm the gradient search algorithm is simply obtained by the correlation function of the input signal. In order to achieve the Wiener solution, in the ECLMS algorithm we estimate all components of the cross-correlation function by using the Toeplitz matrix of the auto-correlation function, therefore the cost function in the ECLMS algorithm can be defined by the sum of the lagged squared errors. The computer simulation results have shown the improvement of the performance than the CLMS algorithm as it was expected. However, for large number of tap coefficient, the ECLMS algorithm is very complex in the computation. So that, here we propose a fast implementation of this algorithm in the frequency domain.

2.4 The Frequency Domain ECLMS Algorithm

In order to reduce the computational complexity of the ECLMS algorithm, we propose the frequency domain ECLMS (FECLMS) algorithm [7], which has been shown in Fig. 3. First, we take N-point of the fast Fourier transform (FFT) of Eqs. (6) and (7) based on the time-lag, k, in the fast Fourier transform kernel as below:

$$F_{xx}(n,p) = \sum_{k=0}^{N-1}\left[\sum_{j=0}^{n} x(j)x(j-k)\right] W^{kp} \qquad (24)$$

$$F_{dx}(n,p) = \sum_{k=0}^{N-1}\left[\sum_{j=0}^{n} d(j)x(j-k)\right] W^{kp} \qquad (25)$$

where W shows complex exponential $e^{-j(2\pi/N)}$, $F_{xx}(n,p)$

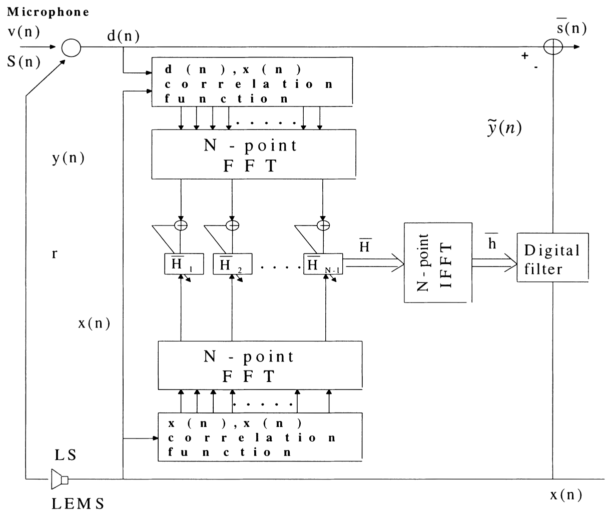

Fig. 3 Echo canceller by using FECLMS algorithm.

shows the FFT of $R_{xx}(n,k)$ at the sample-time n, and p is the frequency variable of the FFT. $F_{dx}(n,p)$ shows the FFT of $R_{dx}(n,k)$.

After the necessary substitutions from Eqs. (2) and (5) into Eq. (25) using Eq. (24), the FFT of the cross-correlation function, $F_{dx}(n,p)$, between $d(n)$ and $x(n)$ signal will be obtained as follows:

$$F_{dx}(n,p) \cong H_p F_{xx}(n,p) \tag{26}$$

where H_p is p^{th} element of the FFT of the echo impulse response vector $r = [r_0, r_1, \cdots, r_{N-1}]$. Then on the basis of Eq. (26), the adaptive filter in which the input signal is the FFT of the autocorrelation function of the far-end signal is defined by:

$$\widetilde{F}_{dx}(n,p) = \widetilde{H}_p(n) F_{xx}(n,p) \tag{27}$$

where $\widetilde{H}_p(n)$ is the adaptive filter tap coefficient in the frequency domain and $\widetilde{F}_{dx}(n,p)$ is the estimation value of $F_{dx}(n,p)$. Next, we define the cost function for adapting tap coefficients as follows:

$$J(n,p) = E[\boldsymbol{\mathcal{E}}^*(n,p)\boldsymbol{\mathcal{E}}(n,p)] \tag{28}$$

where

$$\boldsymbol{\mathcal{E}}(n,p) = F_{dx}(n,p) - \widetilde{F}_{dx}(n,p) \tag{29}$$

The superscript * shows the Hermitian transposition. To obtain the gradient value of Eq. (28), we differentiate Eq. (28) with respect to tap coefficient $\widetilde{H}_p(n)$:

$$\begin{aligned} \nabla J(n,p) &= \frac{\partial}{\partial \widetilde{H}_p(n)} E[\boldsymbol{\mathcal{E}}^*(n,p)\boldsymbol{\mathcal{E}}(n,p)] \\ &= -2E[\boldsymbol{\mathcal{E}}(n,p)F^*_{xx}(n,p)] \end{aligned} \tag{30}$$

From Eq. (30) we derive the steepest descent Frequency Domain ECLMS algorithm (FECLMS) as follows:

$$\widetilde{H}_p(n+1) = \widetilde{H}_p(n) + \frac{2\mu_f \boldsymbol{\mathcal{E}}(n,p)F^*_{xx}(n,p)}{1 + tr[F_{xx}(n,p)F_{xx}(n,p)]} \tag{31}$$

where μ_f is convergence parameter and $tr[\cdot]$ means the trace operator.

As we can see, the structure of the FECLMS algorithm is similar to the ECLMS algorithm, but we process the algorithm in the frequency domain. In FECLMS algorithm, the fast Fourier transform (FFT) of the correlation function is obtained corresponding to the lag-time, not sampling time in the FFT kernel as usually used in conventional methods. And also, we do not need a Toeplitz matrix to estimate all the components of the cross-correlation function, such as in ECLMS algorithm. So that, the computational complex is reduced. The computer simulation results have shown that the FECLMS algorithm has almost same robustness with the ECLMS algorithm, however the computational complexity is reduced.

2.5 Discrete Wavelet Domain Algorithm

As we know, the Fourier transform has been used very popular. However there is a problem with the Fourier transform, when very short-duration and high-frequency bursts occur, it will be hard to detect. The wavelet transform provides a solution to the problem by using an analysis window, which depends on both time and frequency. That means the wavelets can keep track of time and frequency information. The wavelets can be used to “zoom in” on the short bursts mentioned previously, or to “zoom out” to detect long [9], slow oscillations. Therefore, we propose a new implementation of ECLMS algorithm in the wavelet domain called wavelet transform extended correlation LMS algorithm (WECLMS) [8]. The discrete wavelet transform can be defined as [10]:

$$W_\phi(j_0,q) = \frac{1}{\sqrt{M}} \sum_x f(x)\phi_{j_0,q}(x) \tag{32}$$

$$W_\varphi(j_0,q) = \frac{1}{\sqrt{M}} \sum_x f(x)\varphi_{j_0,q}(x) \tag{33}$$

for $j \geq j_0$ and

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{M}} \sum_q W_\phi(j_0,q)\phi_{j_0,q}(x) \\ &\quad + \frac{1}{\sqrt{M}} \sum_{j=j_0}^{\infty} \sum_q W_\varphi(j,q)\varphi_{j,q}(x) \end{aligned} \tag{34}$$

where

$$\phi_{j,q}(x) = 2^{j/2}\phi(2^j x - q) \tag{35}$$

$$\varphi_{j,q}(x) = 2^{j/2}\varphi(2^j x - q) \tag{36}$$

Here, $W_\phi(j_0,q)$ are called the approximation or scaling coefficients; $W_\varphi(j,q)$ are called the detail or wavelet coefficients; $\phi(x)$ is called scaling function and $\varphi(x)$ is called wavelet function; q determines the position of $\phi(x)$ along the x-axis; j determines $\phi(x)$'s width - how broad or narrow it is along the x-axis. The Inverter Discrete Wavelet Transform is defined as Eq. (34).

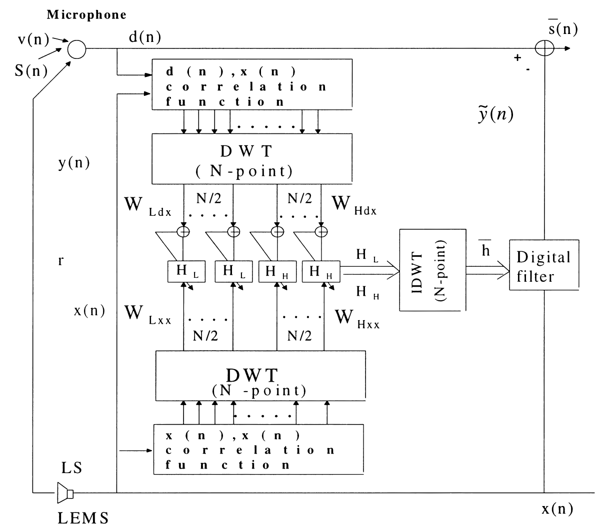

Fig. 4 Echo canceller by using WECLMS algorithm.

In Fig. 4, the structure of the WECLMS algorithm is shown. As shown in Fig. 4, first, we take the N-point Discrete Wavelet Transform (DWT) of the cross-correlation function and the autocorrelation function by Eqs. (32) and (33), respectively. The coefficients vector (N/2-point) can be written as:

$$DWT[R_{dx}(n,k)] = [W_{Ldx}(n), W_{Hdx}(n)] \quad (37)$$

$$DWT[R_{xx}(n,k)] = [W_{Lxx}(n), W_{Hxx}(n)] \quad (38)$$

where

$W_{Ldx}(n) = [W_{Ldx}(n,0), W_{Ldx}(n,1), \cdots, W_{Ldx}(n,N/2-1)]$

$W_{Hdx}(n) = [W_{Hdx}(n,0), W_{Hdx}(n,1), \cdots, W_{Hdx}(n,N/2-1)]$

$W_{Lxx}(n) = [W_{Lxx}(n,0), W_{Lxx}(n,1), \cdots, W_{Lxx}(n,N/2-1)]$

$W_{Hxx}(n) = [W_{Hxx}(n,0), W_{Hxx}(n,1), \cdots, W_{Hxx}(n,N/2-1)]$

$W_{Ldx}(n)$ is an approximation of the cross-correlation function. $W_{Hdx}(n)$ is a detail part of cross-correlation function. $W_{Lxx}(n)$ is an approximation of the autocorrelation function. $W_{Hxx}(n)$ is a detail part of autocorrelation function.

As like as the ECLMS algorithm the error signal is shown by:

$$\mathbf{e}_L(n) = \mathbf{W}^T_{Ldx}(n) - G_{Lxx}(n) * \mathbf{H}_L(n) \quad (39)$$

$$\mathbf{e}_H(n) = \mathbf{W}^T_{Hdx}(n) - G_{Hxx}(n) * \mathbf{H}_H(n) \quad (40)$$

where The "*" means the convolution operator. The $\mathbf{e}_L(n)$, $\mathbf{e}_H(n)$ are vertical vector errors for estimation of the approximation and detail of the cross-correlation function, respectively. $\mathbf{H}_L(n)$, $\mathbf{H}_H(n)$ are the estimation of the room impulse response in wavelet domain for the low-pass band and high-pass band, respectively. And

$$G_{Lxx}(n) = \begin{bmatrix} W_{Lxx}(n,0) & W_{Lxx}(n,1) & \cdots & W_{Lxx}(n,N/2-1) \\ W_{Lxx}(n,1) & W_{Lxx}(n,0) & \cdots & W_{Lxx}(n,N/2-2) \\ \vdots & \vdots & \ddots & \vdots \\ W_{Lxx}(n,N/2-1) & W_{Lxx}(n,N/2-2) & \cdots & W_{Lxx}(n,0) \end{bmatrix}$$

$$G_{Hxx}(n) = \begin{bmatrix} W_{Hxx}(n,0) & W_{Hxx}(n,1) & \cdots & W_{Hxx}(n,N/2-1) \\ W_{Hxx}(n,1) & W_{Hxx}(n,0) & \cdots & W_{Hxx}(n,N/2-2) \\ \vdots & \vdots & \ddots & \vdots \\ W_{Hxx}(n,N/2-1) & W_{Hxx}(n,N/2-2) & \cdots & W_{Hxx}(n,0) \end{bmatrix}$$

$G_{Lxx}(n)$ and $G_{Hxx}(n)$ are both Toeplitz matrix. We can update the tap coefficients as:

$$\mathbf{H}_L(n+1) = \mathbf{H}_L(n) + \frac{2\mu_L G_{Lxx}(n)\mathbf{e}_L(n)}{1 + tr[G_{Lxx}(n)G_{Lxx}(n)]} \quad (41)$$

$$\mathbf{H}_H(n+1) = \mathbf{H}_H(n) + \frac{2\mu_H G_{Hxx}(n)\mathbf{e}_H(n)}{1 + tr[G_{Hxx}(n)G_{Hxx}(n)]} \quad (42)$$

where μ_0 is the step size for tap coefficients adaptation and $tr[\cdot]$ means the trace operator.

Then we use the H_L and H_H to do the Inverse Discrete Wavelet Transform (IDWT) by Eq. (34).

$$IDWT(\mathbf{H}_L, \mathbf{H}_H) = \widetilde{h} \quad (43)$$

Finally, we copied $\widetilde{h}$ from correlation filter into the tap coefficients of the digital filter (DF in Fig. 4), to cancel the echo signal.

In the WECLMS algorithm, the correlation functions are decomposed by the high-pass and low-pass filters and down sample by 2. Therefore, we can adapt the estimation impulse response by using the different step-sizes in two bands, simultaneously. So that, the convergence speed is improved. Also the computational complexity is reduced, because of the downsampling process.

2.6 The Computational Complexity

In this section, we briefly discuss the computational complexity of the correlation-based algorithms. Consider first the standard LMS algorithm with N tap weight operating on real data. In this case, N multiplications are performed to compute the output and a further N multiplications are performed to update the tap weights, making for a total of $2N$ multiplications per iteration. For all kind of correlation-based algorithms, first we need extra $2N$ multiplications to compute the correlation functions $R_{xx}(n,k)$ and $R_{dx}(n,k)$ in Eqs. (22), (23) respectively. Then, as shown in Fig. 2, we also need extra N multiplications to compute the output of Digital Filter $\widetilde{y}(n)$. So for all kind of correlation-based algorithms, totally we need extra $3N$ multiplications, compared with the LMS algorithm. In CLMS algorithm we need N multiplications to estimate the cross-correlation function $\widetilde{R}_{dx}(n,k)$ in Eqs. (9) and N multiplications to update the tap coefficients in Eqs. (13), totally we need 5N multiplications. In the ECLMS algorithm considering tap coefficients adaptation computations as well as calculation for the estimation of the cross-correlation $\widetilde{R}_{dx}(n,k)$ between $d(n)$ and $x(n)$, we need $2N^2$ multiplications, in totally we need $2N^2+3N$ multiplications. On the other hands in the FECLMS algorithm we need three N-point FFTs and only $2N$ multiplications to estimate the cross-correlation function and the tap coefficients adaptation. So that in total the number of multiplication for the FECLMS algorithm is as follows:

$$3 \times \frac{N}{2}\log_2 N + 2N + 3N \quad (44)$$

In the WECLMS algorithm we need three N-point DWT

Table 1 The ratios of the computational loads.

N	$\frac{CLMS}{ECLMS}$	$\frac{FECLMS}{ECLMS}$	$\frac{WECLMS}{ECLMS}$
32	0.031	0.148	0.781
128	0.008	0.049	0.572
256	0.004	0.027	0.535
512	0.002	0.015	0.517

process. As we know the computation of the wavelet decomposition can be write as $2(L+1)N$, where L is the number of nonzero values of the scaling function (For Haar, $L = 2$, and for the Daubechies2, $L = 4$). We only need N^2 multiplications to estimate the cross-correlation function and the tap coefficients adaptation, because of the down sampling process. So that in total the number of multiplication for the WECLMS algorithm is as follows:

$$3 \times 2 \times (L+1) \times N + N^2 + 3N \tag{45}$$

As we can see, the CLMS algorithm is just 2.5 times complexity than the LMS algorithm. The ECLMS algorithm is very complex compare with the LMS algorithm, however the FECLMS algorithm is proposed to reduce the computational complexity.

In Table 1, the ratios of the computational loads for the CLMS, FECLMS and WECLMS to the ECLMS algorithms are given with respect to various number of tap coefficients N. For this comparison we need only to compare the computational loads in the different parts of the proposed algorithms.

So, for instance in $N = 512$ the WECLMS algorithm requires 51.7% of the computational loads for the ECLMS algorithm. The computational complexity is reduced. The FECLMS algorithm requires only 1.5% of computational loads for the ECLMS algorithm. This makes the hardware implementation of the FECLMS algorithm a realistic matter using a fewer chips of DSP or in considering of the mass production, it requires less LSI area.

3. Simulation Results

The acoustic echo impulse response, r_i, of the room is assumed to have exponential decaying shape that decreases to -60 dB after N samples as follows:

$$r_i = Randn[\exp(-8i/N)] \tag{46}$$

To measure the performance of the convergence of the algorithm, we use the ratio of distance of weight and impulse response, DW(n), which is defined as follows:

$$DW(n) = 10\log_{10}\left[\sum_{i=0}^{N-1}\left\|r_i - \widetilde{h}_i(n)\right\|^2 / \sum_{i=0}^{N-1}\|r_i\|^2\right] \tag{47}$$

In order to show the capability and robustness of the proposed new class of algorithm, we have performed several computer simulations by using the real speech data. Here, we use two independent speech signals, one is in English and another is in Japanese. The far-end signal $x(n)$ is the voice of a woman at her 20's and pronounced as "Good morning and welcome to IF I ONLY KNEW. . . " in English. The double-talk signal $s(n)$ is the voice of a woman at her 30's and pronounced as "KA RE GA IZEN KA RA, KAGA KU GIJYUTSU. . . " in Japanese. The sampling frequency is 8 kHz for both. The waveforms of the speech signals are shown in Fig. 5 and Fig. 6. The noise $\upsilon(n)$ is a Gaussian noise signal with zero mean. In the single-talk condition the signal to noise ratio (SNR) is 30 dB (compared with the far-end signal $x(n)$ and noise signal $\upsilon(n)$) and in the double-talk condition SNR is 10 dB (compared with the double-talk signal $s(n)$ and noise signal $\upsilon(n)$). In LMS, CLMS, ECLMS and FECLMS algorithm we set the step size equal to 0.01. In WECLMS algorithm, we use two different step sizes ($\mu_L = 0.001$, $\mu_H = 0.01$) to estimate the room impulse response in the two bands.

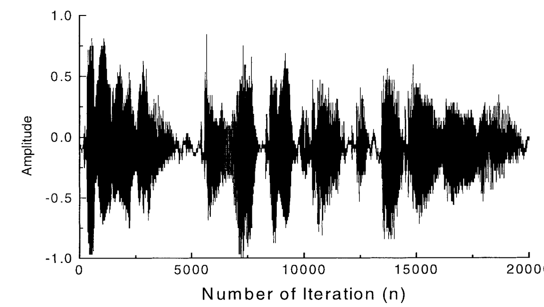

Fig. 5 The input speech signal from far-end.

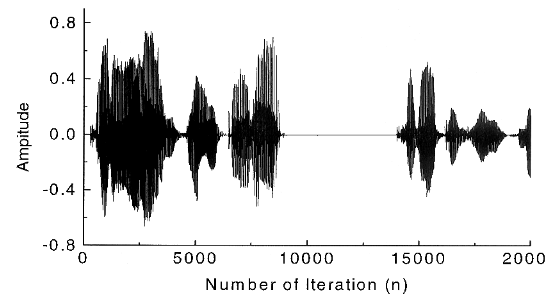

Fig. 6 The double-talk speech signal from near-end.

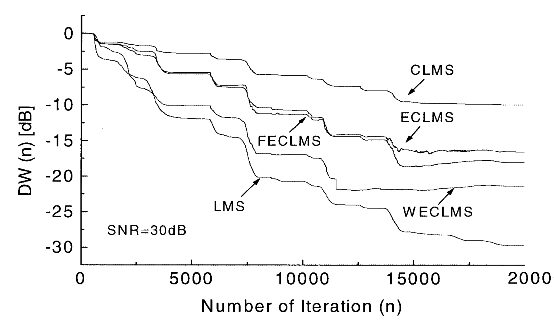

Fig. 7 Comparison between LMS, CLMS, ECLMS, FECLMS and WECLMS in noisy single-talk condition.

In Fig. 7, the convergence characteristics for LMS and proposed algorithms in noisy single-talk condition have been shown. The CLMS algorithm converges to -8 dB, the ECLMS algorithm reaches to -16 dB, the FECLMS al-

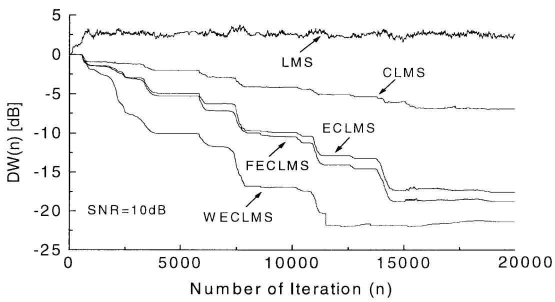

Fig. 8 Comparison between LMS, CLMS, ECLMS, FECLMS and WECLMS in noisy double-talk condition.

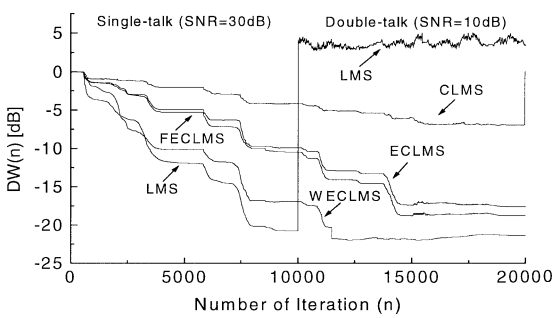

Fig. 9 Switching from single to double talk with the same echo path.

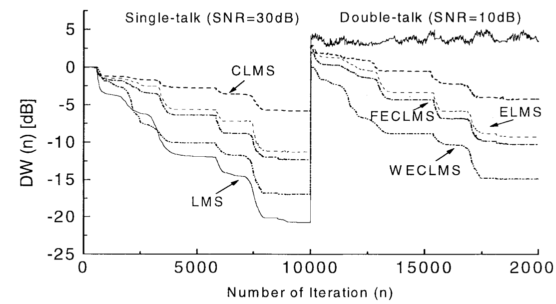

Fig. 10 Switching from single to double talk with the echo path changed.

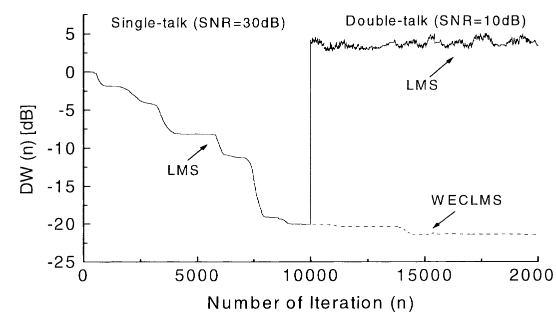

Fig. 11 Switching from single to double talk with the same echo path by using the different algorithm.

gorithm converges to −18 dB, the LMS algorithm reaches to −31 dB. The WECLMS algorithm is better than the CLMS, ECLMS and FECLMS algorithm and it converges to −22 dB. In Fig. 8, the proposed algorithms are compared with LMS algorithm in the noisy double-talk condition. As it shown in Fig. 8, the LMS algorithm hardly converges and totally blown up in the double-talk situation. The CLMS gives a better convergence than the LMS algorithm, and it converges to about −6 dB. The ECLMS algorithm reaches to −16 dB, the FECLMS algorithm converges to −18 dB. The WECLMS algorithm, which is the best among all algorithms, shows a steady convergence under noisy double-talk condition and it converges to −22 dB. Here, the convergence speed was also improved. Then, we note that the new class of algorithms is robust in the double-talk condition. In the next simulation in Fig. 9, we started with the noisy single-talk condition, then, at 10000-th iteration, we changed to double-talk condition, but the acoustic echo impulse response has not been changed here. We can see the robustness of proposed algorithm. These algorithms can continue the adaptation even after changed the single-talk to the double-talk condition. In Fig. 10, we started with the single-talk condition. Then, at 10000-th iteration, we changed the echo path impulse response and imposed the double-talk condition at the same time. As it shown in Fig. 10, the WECLMS algorithm has superior convergence characteristics comparing with the LMS, CLMS, ECLMS and FECLMS algorithm. In Fig. 11, we started with the noisy single-talk condition by using the LMS algorithm, then, at 10000-th iteration, we changed to double-talk condition by using the WECLMS algorithm, but the acoustic echo impulse response has not been changed here. We can see that there is no performance degradation being caused in this procedure.

4. Conclusion

In this paper, a new class of acoustic echo cancelling by using the correlation LMS algorithm for double-talk condition is presented. These algorithms are robust for tap adaptation in an echo canceller in the presence of the noisy double-talk condition. In the proposed CLMS algorithm the gradient search algorithm is simply obtained by the correlation function of the input signal. In the extended CLMS (ECLMS) algorithm, the cost function for tap adaptation is the sum of the lagged MSE of the correlation function. Then, an implementation of the ECLMS algorithm in the frequency domain is proposed. The frequency domain implementation is obtained using the fast Fourier transform where in its kernel the lag time of the correlation function is used as the time variable. In order to improve the convergence speed, an new implementation of the ECLMS algorithm in the wavelet domain is also proposed. The results of the computer simulation have shown superiority of the performance of the proposed algorithms over the conventional algorithm (LMS algorithm) in the double-talk condition. Also, a 22 dB convergence has been obtained as compared with other conventional algorithm, which totally does not converge in double-talk.

References

[1] S. Haykin, Adaptive Filter Theory, Third ed., Prentice Hall, 1996.

[2] P. Heitkamper, "An adaptation control for acoustic echo cancellers," IEEE Signal Process. Lett., vol.4, no.6, pp.170–172, June 1997.
[3] T. Gansler, "A double-talk resistant sub-band echo canceller," Signal Process., vol.65, pp.89–101, April 1997.
[4] T. Gansler, S.L. Gay, M.M. Sondhi, and J. Benesty, "Double-talk robust fast converging algorithm for network echo cancellation," IEEE Tarns. Speech Audio Process., vol.8, no.6, pp.656–663, Nov. 2000.
[5] M.R. Asharif, T. Hayashi, and K. Yamashita, "Correlation LMS algorithm and its application to double-talk echo cancelling," Electron. Lett., vol.35, no.3, pp.194–195, Feb. 1999.
[6] M.R. Asharif, A. Shimabukuro, T. Hayashi, and K. Yamashita, "Expanded CLMS algorithm for double-talk echo cancelling," Proc. IEEE, SMC'99, vol.1, pp.998–1002, Oct. 1999.
[7] M.R. Asharif, M. Shiroma, and K. Yamashita, "Design of echo canceler with frequency domain ECLMS algorithm," GlobeCom 2000, San Francisco, USA, Nov. 2000.
[8] R. Chen and M.R. Asharif, "Acoustic echo canceling in the double-talk condition," International Workshop on Acoustic Echo and Noise Control (IWAENC 2003), pp.39–42, Kyoto, Japan, Sept. 2003.
[9] A. Boggess and F.J. Narcowich, A First Course in Wavelet with Fourier Analysis, Prentice Hall, 2001.
[10] R.C. Gonzalez and R.E. Woods, Digital Image Processing, Second ed., Prentice Hall, 2002.

Rui Chen was born in Hunan, China, on March 10, 1974. He received the B.E. in information engineering from the south-central university for nationalities, China, M.E. degree in electrical engineering from the university of Ryukyu, Japan, in 1998, 2003, respectively. Now He is a Ph.D. student in the university of Ryukyu. His research interests are in the field of digital signal processing, acoustic echo cancelling, adaptive digital filtering.

Mohammad Reza Asharif was born in Tehran, Iran, on December 15, 1951. He received the B.Sc. and M.Sc. degrees in electrical engineering from the university of Tehran, Tehran, 1973, 1974, respectively and the Ph.D. degree in electrical engineering from the university of Tokyo, Tokyo in 1981. He was Head of Technical Department of T.V. broadcasting college (IRIB), Tehran, Iran from 1981 to 1985. Then, he was a senior researcher at Fujitsu Labs. Co. Kawasaki, Japan from 1985 to 1992. Then, he was an assistant professor in Department of Electrical and Computer Engineering, University of Tehran, Tehran, Iran from 1992 to 1997. Dr. Asharif is now a full professor at Department of Information Engineering, University of the Ryukyus, Okinawa, Japan since 1997. He has developed an algorithm and implemented its hardware for real time T.V. Ghost cancelling. He introduced a new algorithm for Acoustic Echo Canceller and he released it on VSP chips. Professor Asharif has contributed many publications to journals and conference proceedings. His research interests are in the field of digital signal processing, acoustic echo cancelling, active noise control, adaptive digital filtering, image and speech processing. Professor Asharif is a senior member of IEEE since 1998.

Iman Tabatabaei Ardekani was born in Tehran, Iran, on September 9, 1978. He received the M.Sc. degree in electrical engineering from the university of Tehran, Tehran, 2003.

Katsumi Yamashita was born in Hyogo, Japan in 1952. He received the B.E. degree in electrical engineering from Kansai University, M.E. degree in electrical engineering from Osaka Prefecture University, the Dr.Eng. degree in electrical engineering from Osaka University in 1974, 1976 and 1985, respectively. In 1982, he became an assistant professor in the Department of Electronics and Information Engineering at the University of the Ryukyus, where he was a professor in 1991. Now he is a professor in Graduate School of Engineering, Osaka Prefecture University. His current interests are in digital communication, digital signal processing. Prof. Yamashita is a member of the IEEE, IEE Japan.