Acoustic Echo and Noise Cancellation System for Hand-Free Telecommunication using Variable Step Size Algorithms

V. K. GUPTA[1], Mahesh CHANDRA[2], S. N. SHARAN[3]

[1]Department of ECE, IMS Engineering College, Ghaziabad, UP, India-201009
[2]Department of ECE, Birla Institute of Technology, Mesra, Ranchi, Jharkhand, India-835215
[3]Director, Greater Noida Institute of Technology, Greater Noida, UP, India

guptavk76@gmail.com, shrotriya69@rediffmail.com, snathsharan@yahoo.com

Abstract. *In this paper, acoustic echo cancellation with doubletalk detection system for a hand-free telecommunication system is implemented using Matlab. Here adaptive noise canceller with blind source separation (ANC-BSS) system is proposed to remove both background noise and far-end speaker echo signal in presence of double-talk. During the absence of double-talk, far-end speaker echo signal is cancelled by adaptive echo canceller. Both adaptive noise canceller and adaptive echo canceller are implemented using LMS, NLMS, VSLMS and VSNLMS algorithms. The normalized cross-correlation method is used for double-talk detection. VSNLMS has shown its superiority over all other algorithms both for double-talk and in absence of double-talk. During the absence of double-talk it shows its superiority in terms of increment in ERLE and decrement in misalignment. In presence of double-talk, it shows improvement in SNR of near-end speaker signal.*

Keywords

Adaptive filters, double-talk detection, ANC-BSS, ERLE, misalignment, SNR.

1. Introduction

In hands-free telephony and teleconferencing systems, acoustic echo reduces conversation quality, even at very small echo delays. In early telephony, microphone and loudspeaker are not enclosed in a single unit but they are separated and no sound could propagate between the speaker and the microphone. Therefore no echo would be transmitted back. Using a hands free loudspeaker telephone, the microphone and loudspeaker are enclosed in a single unit. The sound from the loudspeaker will be picked up by the microphone and transmitted back to the sender which is recognized as an echo. Acoustic echo canceller (AEC) suppresses the unwanted echo signal. For echo cancellation, initially echo path impulse response with an adaptive finite impulse response (FIR) filter is modeled. Subtraction of the filter output from the actual echo signal results in echo cancellation [1]-[3]. The adaptive filter converges to a good estimate of the echo path response and successfully cancels the echo in absence of near-end talk. However, the double-talk [4] situation occurs when both near-end speaker and far-end speakers speak at the same time. The near-end speech acts as uncorrelated noise to the adaptive algorithm and the filter may diverge, causing annoying audible echo to pass through to the far-end speaker. The best way to remove this problem is to halt filter adaptation in presence of near-end speech. This is the important role of the doubletalk detector (DTD). During double-talk detection, near-end speaker signal is corrupted by environment noise and far-end speaker echo signal. A combined adaptive noise canceller and blind source separation (ANC-BSS) system is proposed to remove noise and far-end echo signal from near-end signal. In ANC-BSS system, first ANC removes the environment noise then BSS is used to separate the near-end speaker signal from far-end echo signal. The block diagram of combined acoustic echo and noise canceller for hand-free telecommunication is shown in Fig. 1.

In this work, Matlab simulations are done inside an experimental room. For echo path impulse response, the test signal is generated by laptop loudspeaker and echo is recorded by the laptop microphone. The acoustic echo of far-end speaker may transmit back to the far-end speaker therefore it must be cancelled to improve the conversation between far-end speaker and near-end speaker. Both adaptive noise canceller and adaptive echo canceller are implemented using Least Mean Squares (LMS), Normalized Least Mean Squares (NLMS), Variable Step size Least Mean Squares (VSLMS) and Variable Step size Normalized Least Mean Squares (VSNLMS) algorithms [5]-[11].

2. Acoustic Echo Canceller Structure

In a loudspeaker-enclosure-microphone (LEM) system, both microphone and loudspeaker are either directly connected by an acoustic path or by a large number of reflections at the boundaries of the enclosure [3]. The impulse responses of LEM systems are highly sensitive to any changes such as the movement of a person carrying LEM system and material of the surfaces through which reflection take place. Thus, the impulse response of an LEM system is time-variant. The impulse response and frequency response of LEM system enclosed in an experimental room at any instant is shown in Fig. 2. The nonlinear processor (NLP) removes all signals below a preset threshold which is used to further remove the residual echo signal. During double talk, NLP remains in its idle position and it does not block the near-end speaker signal. Comfort-noise generator is used in absence of double-talk. The signal to be send to far-end speaker during the absence of double-talk situation is only residual echo. When residual echo becomes zero for some duration, the far-end speaker hears no sound. He might suspect that telephone line has gone dead. In this case, far-end speaker hears the comfort noise sound as shown in Fig. 3.The comfort-noise used here is white-Gaussian noise colored with the residual echo signal before NLP.

3. Acoustic Double-Talk Detection

In full duplex communication, double talk condition occurs when both near-end speaker and far-end speaker speak at the same time. When double talk occurs, the algorithms have difficulties to decide in between echo and near-end speech. If the canceller detects a double talk condition, the near-end speech will diverse the adaptive filter [4]. Therefore the role of double-talk detector is to stop the adaption through adaptive filter. In double talk detection process, a detection statistics is computed and compared with the preset threshold. Detection statistics is computed by different methods such as Geigel algorithm, cross- correlation method and normalized cross-correlation method. The normalized cross-correlation method is used for double-talk detection in this paper.

3.1 The Normalized Cross-Correlation Method

The microphone signal $d[n]$ can be expressed as a sum of the far-end speaker echo signal $x_E[n]$and the near-end speaker signal $x_N[n]$(neglecting noise influence first) i.e.

$$d[n] = x_E[n] + x_N[n]. \tag{1}$$

If impulse response of echo path of the room is h, then echo signal is:

$$x_E[n] = h^T x_F[n]$$

where $x_F[n]$ is far-end speaker signal. Therefore equation (1) can be written as

$$d[n] = h^T x_F[n] + x_N[n]. \tag{2}$$

From (2) (taking cross-correlation between $x_F[n]$and $x_N[n]$ equal to zero)

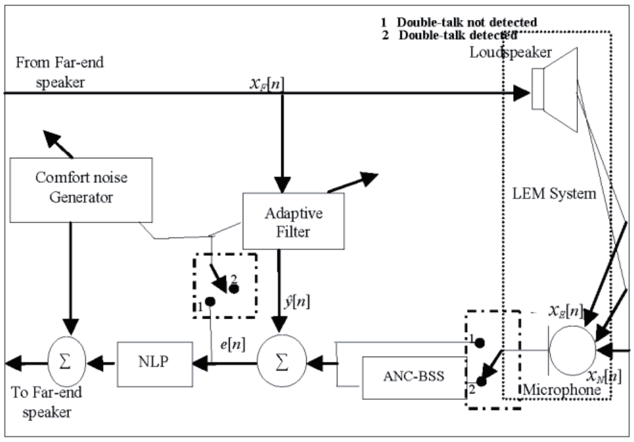


Fig. 1. A combined acoustic echo and noise canceller.

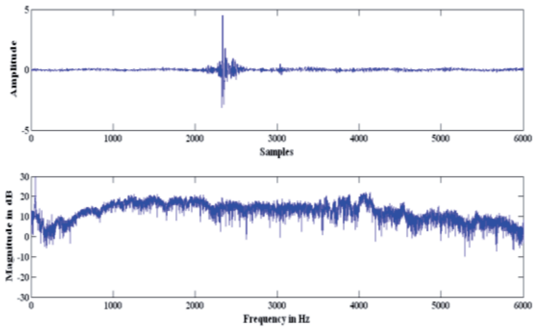


Fig. 2. Impulse response and frequency response of an experimental room.

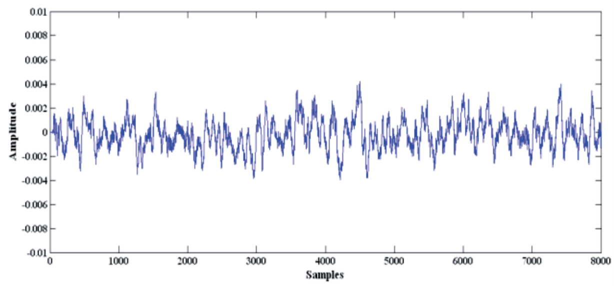


Fig. 3. Output of comfort noise generator.

$$\sigma_d^2 = h^T \chi_{FF} h + \sigma_N^2. \tag{3}$$

Now, the cross-correlation sequence of the far-end speaker and microphone signals can be expressed according to definition,

$$\begin{aligned}\chi_{FE} &= E[x_F[n]x_E[n]] \\ &= E[x_F[n]\{hx_F^T[n]\}] \\ &= h\chi_{FF}\end{aligned}.$$

Therefore

$$h = \chi_{FE}\chi_{FF}^{-1}. \tag{4}$$

Therefore, (3) can be written as

$$\sigma_d^2 = \chi_{FE}^T \chi^{-1}{}_{FF} \chi_{FE} + \sigma_N^2.$$

In absence of near-end speaker, i.e. $x_N[n] = 0$, then $d[n] = x_E[n]$. Therefore,

$$\sigma_d^2 = \chi_{Fd}^T \chi^{-1}{}_{FF} \chi_{Fd}$$

where

$$\chi_{Fd} = E[x_F[n] d[n]].$$

The detection statistics is suggested as:

$$\eta = \left[\frac{\chi_{Fd}^T \chi^{-1}{}_{FF} \chi_{Fd}}{\sigma_d^2} \right]. \tag{5}$$

The nominator is the power of the measured signal if no near-end speech is present, whereas the denominator is the actual power of the measured signal.

$\eta \approx 1$: there is no near-end speech signal present,

$\eta < 1$: otherwise.

4. ANC-BSS System

Combined adaptive noise canceller and blind source separation (ANC-BSS) system is used to separate near-end speaker signal from mixed noisy near-end speaker and far-end speaker echo signal. Fig. 4 shows the block diagram for source separation process. An adaptive noise canceller is used to remove the noise from corrupted signal. The mixed signal composed of near-end signal and far-end speaker echo signal is decomposed using Discrete Wavelet Transform (DWT). Further independent component analysis (ICA) is applied on the decomposed signal. Now the separated signals are reconstructed using Inverse Discrete Wavelet Transform (IDWT).

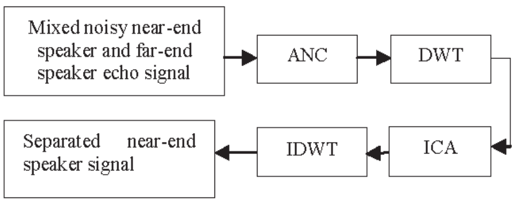


Fig. 4. Block diagram for source separation.

4.1 Adaptive Noise Canceller (ANC)

Adaptive noise cancellation system is shown in Fig. 5.The reference noise $\tilde{N}[n]$ is input to the transversal filter. The output of the transversal filter is $y[n]$ which is convolution of reference noise $\tilde{N}[n]$ and filter tap weight $w[n]$.The noisy signal $d[n]$ which consists of an information bearing signal $s[n]$ corrupted by noise $N[n]$. The $d[n]$ and $y[n]$ are compared to give the error signal $e[n]$.The adaptive filter coefficients are changed iteratively according to the error signal $e[n]$.The filter weights are adjusted continuously to minimize the error between $d[n]$ and $y[n]$, so that the output $e[n]$ is a close approximation of the signal $s[n]$. Both noise signals $N[n]$ and $\tilde{N}[n]$ are uncorrelated with the signal $s[n]$ while correlated with each other. The error $e[n]$ gives the estimated clean signal at the output.

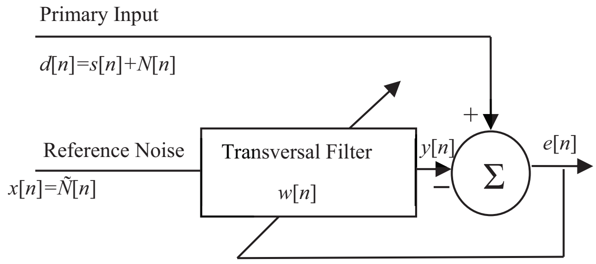


Fig. 5. Adaptive noise cancellation system.

The adaptive noise canceller system output is given by (6),

$$\begin{aligned} e[n] &= d[n] - y[n] \\ &= s[n] + N[n] - y[n] \end{aligned}. \tag{6}$$

Minimization of the estimate of mean square error is given by

$$\begin{aligned} &\min\left[E\left(e[n]\right)^2 \right] = \min\left[E\left[\left(s[n] + N[n] - y[n] \right)^2 \right] \right] \\ &= E[(s[n])^2] + \min\left[E[\left(N[n] - y[n] \right)^2] \right] \end{aligned} \tag{7}$$

Noise present in the output $e[n]$ is $[N[n]\text{-}y[n]]$. Since signal power $s[n]$ is uncorrelated with both $N[n]$ and $\tilde{N}[n]$, minimization of estimate of mean square error will minimize the noise power present in output $e[n]$ and output will be an exact replica of signal $s[n]$. Due to reduction in noise power, signal-to-noise ratio increases at the output.

4.2 Discrete Wavelet Transform (DWT)

The input signal $e[n]$ is decomposed into two sets of coefficients called approximation coefficients (denoted by c_a) and detail coefficients (denoted by c_d). These coefficients are obtained by convolving the input signal with a low-pass filter (for c_a) or a high-pass filter (for c_d) and then down-sampling the convolution result by 2. The size of c_a and c_d is half of the size of the input signal. The filters are determined by the chosen wavelet. Fig. 6(a) shows single-level DWT decomposition.

IDWT is the inverse process of wavelet decomposition. In contrast to decomposition, the reconstruction process is comprised of up-sampling and then filtering. The filters are determined by the type of the wavelet. Fig. 6(b) shows single-level IDWT reconstruction.

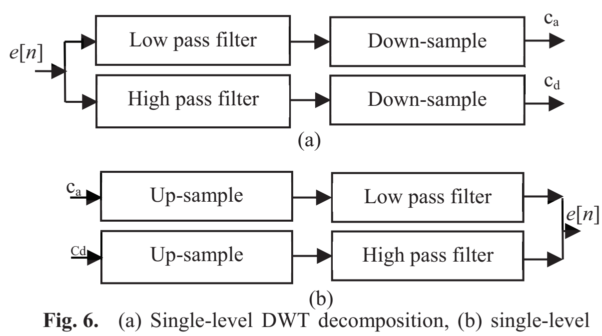


Fig. 6. (a) Single-level DWT decomposition, (b) single-level IDWT reconstruction.

4.3 Independent Component Analysis (ICA)

ICA is a method of signal processing and data analysis developed for Blind Source Separation (BSS). By the approach of ICA [12]-[15], even without any information of the source signals and the coefficients of transmission channel, source signals can be extracted only from the observations according to the stochastic properties of the input signals. ICA analysis is applied on a speech signal which is combination of two speech signals. ICA exploits the non-Gaussianity of the sources in the mixtures. For signal separation the non-Gaussian nature of signals are increased by preprocessing using wavelet packet decomposition.

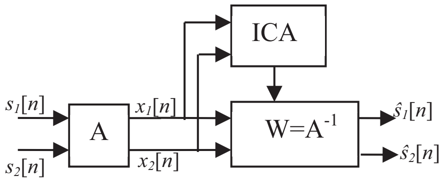


Fig. 7. Basic principle of ICA.

In the instantaneous mixture case, the sources are not observed directly but as a linear combination such that:

$$x_i[n] = \sum_{j=1}^{N} a_{ij} s_j[n] \tag{8}$$

where s are source signals, x are observed signals and $\mathbf{A} = [a_{ij}]$ is an unknown full rank mixing matrix. In practice, the goal of ICA is to find the inverse of $\mathbf{A}$, which is the unmixing matrix $\mathbf{W} = \mathbf{A}^{-1}$.The preprocessing transforms the observed signals to find an adequate representation where the signals distributions are non-Gaussian. For this, the wavelet transform is used to emphasize the non-Gaussian nature of the observed signals. Once the inverse matrix $\mathbf{W}$ is found with the wavelet packets based ICA, then the separation is performed using IDWT. For wavelet packet based ICA, only one wavelet coefficient node is selected and other coefficients are made zero before IDWT. The selection of this node is done as follows:

(i) Decompose the observed signal into wavelet packets.

(ii) Compute the Shannon entropy value at each node.

(iii) Select the node that has the lowest entropy.

The Shannon entropy is defined for each node (j, k) as:

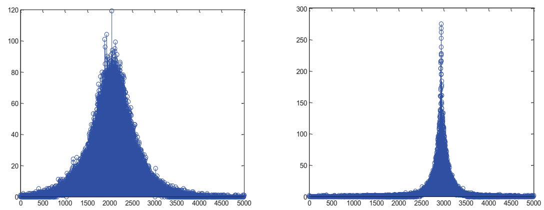
$$H(j,k) = -\sum_i p_i \log(p_i) \tag{9}$$

with

$$p_i = \frac{\left\| C_{j,k}[i] \right\|^2}{\left\| x \right\|^2}$$

where $C_{j,k}$ are wavelet coefficients and x is observed signal.

Fig. 8(a) and (b) shows the histogram of mixed signal and selected wavelet coefficient. It is observed that mixed signal distribution is more Gaussian than that of the selected coefficient. This means that non-Gaussian nature of signals is increased by preprocessing with DWT.



(a) (b)

Fig. 8. Histogram of (a) mixed signal, (b) selected wavelet coefficient.

5. Adaptive Algorithms

5.1 LMS Algorithm

It is well known and widely used algorithm due to its computational simplicity.The desired signal $d[n]$ is tracked by adjusting the filter coefficients $w[n]$. The input reference signal $x[n]$ is a known signal that is fed to the FIR filter. The difference between $d[n]$ and $y[n]$ is the error signal $e[n]$ as shown in Fig. 5. The error signal $e[n]$ is then fed to the LMS algorithm to compute the updated filter coefficients $w[n+1]$ to iteratively minimize the error. The weight update of LMS algorithm is done as per (10).

$$w[n+1] = w[n] + \mu e[n] x[n]\,. \tag{10}$$

The convergence time of the LMS algorithm depends on the step size μ. If μ is small, then it may take a long time to converge and this may defeat the purpose of using an LMS filter. However if μ is too large, the algorithm may never converge. The value of μ should be scientifically computed based on the environmental effects on $d[n]$.

5.2 NLMS Algorithm

The Fixed step size is the primary disadvantage of the LMS algorithm during all iterations. For fixed step size, statistics of the input signal is required before commencing the adaptive filtering operation. This is practically not desirable for online applications. The Normalised Least Mean Squares algorithm (NLMS) is a modified version of LMS algorithm where a variable step size value $\mu[n]$ is selected for each iteration of the algorithm. This step size is proportional to the inverse of the total expected energy of the instantaneous values of the coefficients of the input vector $x[n]$. The recursion formula for NLMS algorithm is given by the equation

$$w[n+1] = w[n] + \mu[n] e[n] x[n]\,. \tag{11}$$

where

$$\mu[n] = \frac{1}{x^T[n] x[n] + \varphi}\,.$$

Here, φ is a small positive constant in order to avoid division by zero when the values of the input vector are zero.

5.3 VSLMS Algorithm

The drawback of NLMS algorithm is that it has a fixed step size value for every tap weight in each iteration. In Variable Step size Least Mean Squares (VSLMS) algorithm, the step size $\mu[n]$ is different for each element of the filter tap weight vector $w[n]$ in one iteration. The step size and filter tap weight vectors are updated using (12) in each iteration.

$$\begin{aligned} & g_i[n] = e[n]x[n-i] \\ & \mu_i[n] = \mu_i[n-1] + \rho g_i[n] g_i[n-1] \\ & \mu_i[n] > \mu_{\max}, then \mu_i[n] = \mu_{\max} \\ & \mu_i[n] < \mu_{\min}, then \mu_i[n] = \mu_{\min} \\ & w_i[n+1] = w_i[n] + 2\mu_i[n] g_i[n] \end{aligned} \quad (12)$$

where $i = 0, 1, 2, \ldots, (N-1)$ for filter order N. where the parameter ρ is a small positive constant that controls the adaptive behavior of the step-size sequence $\mu_i(n)$. μ_{min}, μ_{max} are chosen to satisfy the convergence requirements of the algorithm.

5.4 VSNLMS Algorithm

The statistical knowledge of the input signal is required prior to the commencement of LMS algorithms and its variants NLMS and VSLMS in order to guarantee the stability of the algorithm. The major benefit of the NLMS algorithm is that it is designed to avoid this requirement by calculating an appropriate step size based upon the instantaneous energy of the input signal vector. If it is incorporated in the step size calculation into the variable step size algorithm, the stability for the filter increases without prior knowledge of the input signal statistics. The step size and filter tap weight vectors in one iteration are updated using (12), except $\mu_i[n]$ is given by (13)

$$\mu_i(n) = \mu_i(n-1) + \frac{\rho g_i(n) g_i(n-1)}{\|x(n-i)\|^2}. \quad (13)$$

6. Experimental Setup and Results

For Matlab simulation of echo canceller system, appropriate speech database is required. One clean sentence "DHOWBIN JAB SO KAR UTHTHI TO DEKHTI KI CHAWKA SAAF PADAA HAI AUR BARTAN MANJEY HUEYN HAIN" from Hindi Speech Database [16] has been taken as test sample for far-end signal. The noisy version of clean sentence "YAHA SAI LAGHBAG PANCH MEAL DAKSHIN PASCHIM MAI KATGHAR GAON HAI" from Hindi Speech Database [16] has been taken as test sample for near-end signal. The noisy version of this sentence was prepared by adding car noise from NOISEX-92 database [17] to this clean sentence at 0 dB, -5 dB and 5 dB SNR levels. The experimental setup for echo canceller is shown in Fig. 1. Experiments are performed with LMS, NLMS, VSLMS and VSNLMS algorithms. Step-size μ is taken as 0.14 in LMS and NLMS algorithms. $\mu_{\max}$ is taken as 0.1 in VSLMS algorithm and $\mu_{\min}$ is taken as 0.00001 in both VSLMS and VSNLMS algorithms. Adaptive filter order is taken as 45 in all algorithms.

6.1 During No Double-talk Detection

The complete flow chart for the implemented system is shown in Fig. 9.There is no double-talk situation when far-end speaker is speaking and near-end speaker is not speaking. Then adaptive filter operates to generate an estimated echo path impulse response to cancel the echo of far-end speaker signal such that the far-end speaker does not hears his own voice as echo. The performance of the echo canceller is measured in terms of Echo Return Loss Enhancement (ERLE) and Misalignment.

6.1.1 Echo Return Loss Enhancement (ERLE)

The Echo Return Loss Enhancement (ERLE) is a measure of the amount of echo suppressed by the acoustic echo canceller. It is defined as the ratio of power of original echo over the power of the residual echo signal after cancellation in dB,

$$ERLE = 10\log_{10} \frac{\text{power of the microphone signal}}{\text{power of the residual echo signal}}. \quad (14)$$

Measurement of ERLE is performed in the portion where there is no near-end signal but only the echo. The higher the ERLE, the better the AEC works.

6.1.2 Misalignment

Misalignment is a measure of closeness between the estimated impulse response $\hat{h}[n]$ and true impulse response $h[n]$ of the echo path. It is defined as the logarithmic normalized Euclidian distance between the true and estimated impulse response at each time instant.

$$Misalignment = 20\log_{10}\left\{\frac{\left\|\hat{h}[n] - h[n]\right\|}{\left\|h[n]\right\|}\right\}. \quad (15)$$

The lower the misalignment, the better will be the convergence of the adaptive filter.

Fig. 10 and Fig. 11 show the variations of ERLE and misalignment in dB for LMS, NLMS, VSLMS and VSNLMS algorithms. Tab. 1 shows the comparative performance of all algorithms in terms of maximum value of ERLE and average ERLE value in dB. VSNLMS algorithm shows its superiority among all in terms of maximum ERLE and average ERLE. Higher ERLE indicates that acoustic echo cancellation system removes the residual echo more efficiently. It is observed from Fig. 11(a) to Fig. 11(d) that VSNLMS has lowest mis-alignment among all algorithms. It is also observed that misalignment value is tending towards zero in all

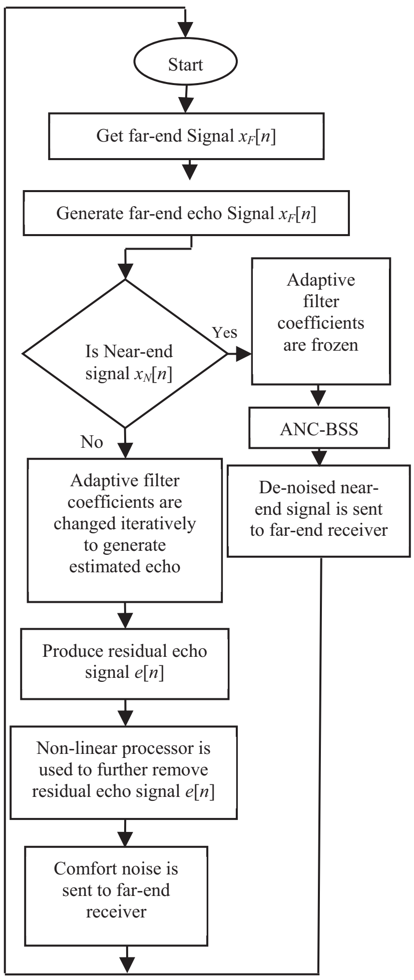


Fig. 9. Flowchart of combined acoustic echo and noise canceller system.

algorithms. However, VSNLMS takes a smaller number of iterations to converge than other algorithms. Lower misalignment means that the acoustic echo cancellation system estimates the echo path more accurately.

Algorithm	Maximum ERLE(dB)	Average ERLE(dB)
LMS	52.1	26.03
NLMS	56.1	27.45
VSLMS	56.6	28.12
VSNLMS	57.2	29.04

Tab. 1. Comparative ERLE for echo canceller of all algorithms.

6.2 During Double-talk Detection

During double-talk, double-talk detector stops the adaption through adaptive filter. The ANC-BSS system is activated. The separation of far-end speaker echo and near-end speaker was done using wavelets of the mixed signal. Two level of decomposition were done on the wavelets. Then ICA algorithm was applied on those wavelets for separation of source signal from the mixed signal. Noisy near-end signal corrupted by car noise at different SNR levels is given to the input of ANC-BSS system. Fig. 12

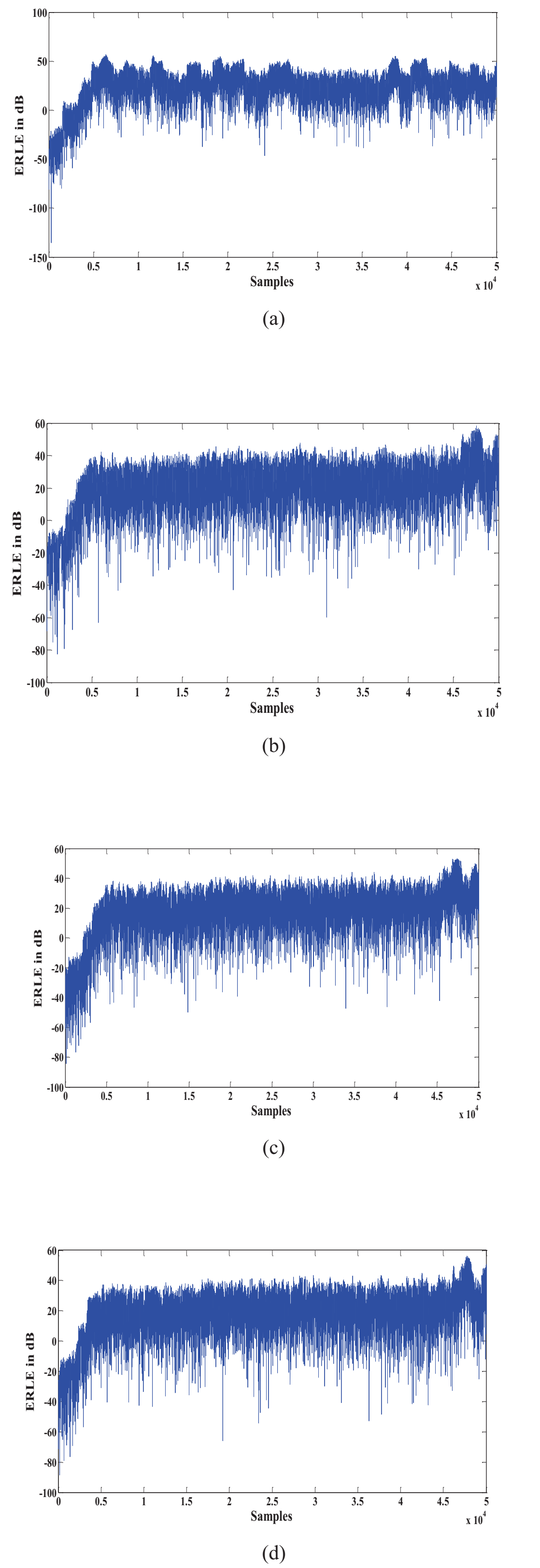
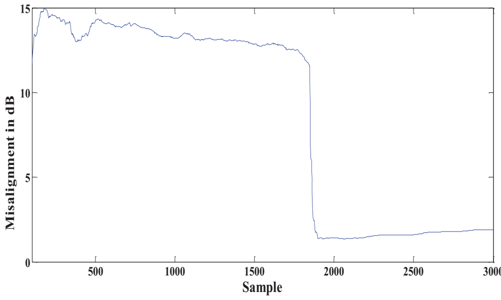
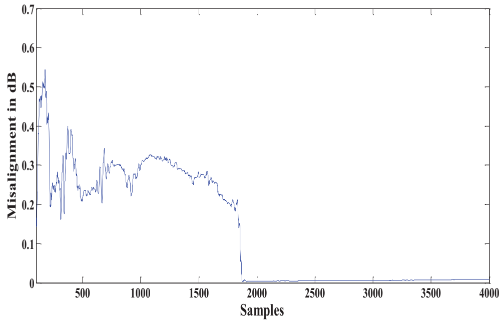


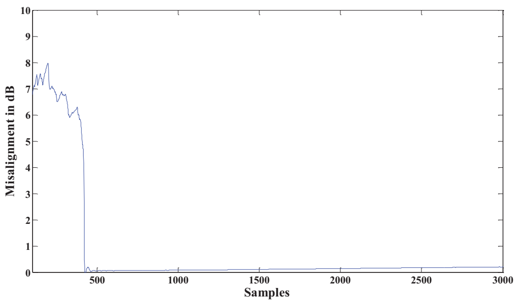
Fig. 10. Comparative performance of ERLE variations: ERLE variations for (a) LMS, (b) NLMS, (c) VSLMS, and (d) VSNLMS.



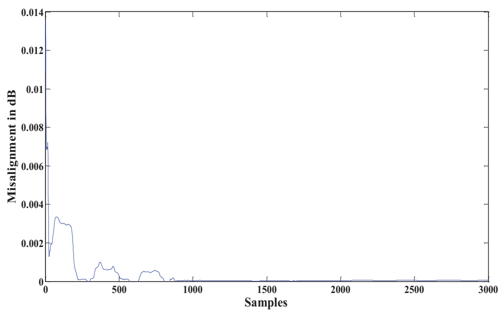
(a)



(b)



(c)



(d)

Fig. 11. Comparative performance of misalignment variations: misalignment variations for (a) LMS, (b) NLMS, (c) VSLMS, and (d) VSNLMS.

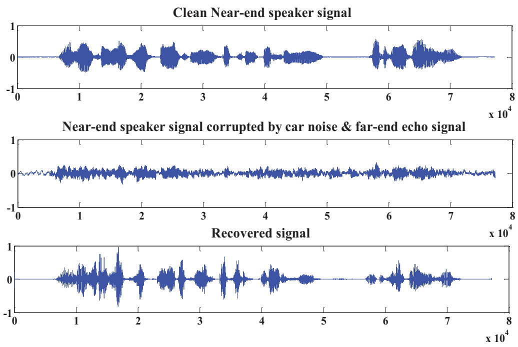


Fig. 12. Recovered signal during double-talk for near-end speaker signal corrupted by car noise at 0 dB SNR level and far-end echo signal.

shows the recovered near-end signal from corrupted signal using ANC-BSS. Tab. 2.shows the performance of ANC-BSS system using LMS, NLMS, VSLMS and VSNLMS adaptive algorithms at all input SNR levels in terms of improvement in SNR. It is observed that ANC-BSS system using VSNLMS algorithm has shown its superiority over other algorithms at all input SNR levels. Tab. 3 shows the average execution time of all three input SNR levels of ANC-BSS system using LMS, NLMS, VSLMS and VSNLMS adaptive algorithms. It is observed that ANC-BSS system using VSNLMS algorithm shows its superiority at the cost of increased computational complexity.

SNR of Near-end Noisy Signal in dB	SNR of Recovered Near-end Signal in dB			
	VSNLMS	VSLMS	NLMS	LMS
-5dB	5.09	4.22	2.24	2.20
0dB	20.9	17.23	13.22	12.46
5dB	23.4	20.11	15.45	14.78

Tab. 2. Performance of ANC-BSS system using different adaptive algorithms.

VSNLMS	VSLMS	NLMS	LMS
115.34	104.81	100.68	69.90

Tab. 3. Average execution time of all three input SNR levels of ANC-BSS system in seconds.

7. Conclusion

In this paper, a combined acoustic echo and noise canceller system based on normalized cross-correlation method for double-talk detection has been implemented. The performance of the system is evaluated both in presence and absence of double talk. Simulation results revealed that a combined acoustic echo and noise canceller system with VSNLMS algorithm is superior to other algorithms at all SNR levels.

References

[1] SONDHI, M. M. An adaptive echo canceller. *Bell Systems Technical Journal*, 1967, vol. 46, p. 497 - 510.

[2] ÅHGREN, P. Acoustic echo cancellation and doubletalk detection using estimated loudspeaker impulse responses. *IEEE Transactions on Speech and Audio Processing*, 2005, vol. 13, no. 6, p. 1231 - 1237.

[3] HANSLER, E., SCHMIDT, G. *Acoustic Echo and Noise Control: A Practical Approach*. Hoboken (NJ, USA): Wiley, 2004.

[4] CHO, J. H., MORGAN, D. R., BENESTY, J. An objective technique for evaluating doubletalk detectors in acoustic echo cancellers. *IEEE Transactions on Speech and Audio Processing*, 1999, vol. 7, no. 6, p. 718 - 724.

[5] HWANG, S. S., KIM, S. C., LEE, C. D. Real time acoustic echo cancellation technique based on normalized least mean square method. In *2nd International Conference on Computer Engineering and Technology*. Chengdu (China), 2010, vol. 6, p. 12 - 16.

[6] SHAO, T., ZHENG, Y. R., BENESTY, J. A variable step-size normalized sign algorithm for acoustic echo cancellation. *In IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*. Dallas (TX, USA), 2010, p. 333 - 336.

[7] HOMANA, I., TOPA, M. D., KIREI B. S., CONTAN, C. Adaptive algorithms for double-talk echo cancelling. In 9th International Symposium on *Electronics and Telecommunications.* Timişoara (Romania), 2010, p. 349 - 352.

[8] HAYKIN, S. *Adaptive Filter Theory*, 4th ed. Upper Saddle River (NJ, USA): Prentice Hall, 2002.

[9] SAYED, A. H., *Fundamentals of Adaptive Filtering*. Hoboken (NJ, USA): Wiley, 2003.

[10] ABOULNASR, T., MAYYAS, K. A robust variable step size LMS-Type algorithm: Analysis and simulations. In *International Conference on Acoustics, Speech, and Signal Processing*. Detroit (MI, USA), 1995, vol. 2, p. 1408 - 1411.

[11] GREENBERG, J. E. Modified LMS algorithms for speech processing with an adaptive noise canceller. *IEEE Transactions on Speech and Audio Processing*, 1998, vol. 6, no. 4, p. 338 - 358.

[12] WADA, T. S., JUANG, B.-H. H. F. Enhancement of residual echo for robust acoustic echo cancellation. *IEEE Transactions on Audio, Speech, and Language Processing*, 2012, vol. 20, no.1, p. 175 - 189.

[13] MIRARAB, M. R., SOBHANI, M. A, NASIRI, A. A. A new wavelet based blind audio source separation using kurtosis. In *3rd International Conference on Advanced Computer Theory and Engineering*. Chengdu (China), 2010, p. 550 - 553.

[14] MOUSSAOUI, R., ROUAT, J., LEFEBVRE, R. Wavelet based independent component analysis for multi-channel source separation. In *International Conference on Acoustics, Speech and Signal Processing*. Toulouse (France), 2006, vol. 5, p. 645 - 648.

[15] HIRAI, N., MATSUMOTO, H., FURUKAWA, T., FURUYA, K. A consideration of blind source separation using wavelet transform. In *International Symposium on Intelligent Signal Processing and Communication Systems*. 2004, p. 141 - 146.

[16] SAMUDRAVIJAYA, K., RAO, P. V. S., AGRAWAL, S. S. Hindi speech database. In *Sixth International Conference on Spoken Language Processing (ICSLP)*. Beijing (China), 2000, CDROM 00192.pdf.

[17] VARGA, A., STEENEKEN, H. J. M., JONES, D., The noisex-92 study on the effect of additive noise on automatic speech recognition system, *Reports of NATO Research Study Group* (RSG.10). 1992.

About Authors...

Vijay Kumar GUPTA was born in 1976. He has received the B.Tech degree from MJP Rohilkhand University, Bareilly in 2001 and M.Tech degree in Instrumentation from DAVV, Indore in 2003. Currently he is pursuing Ph.D. from Birla Institute of Technology, Mesra, Ranchi, India. His research interest includes signal and speech processing. He is working as an associate professor at the Department of Electronics and Communication Engineering in IMS Engineering College, Ghaziabad (UP), India. He has published more than 4 research papers in the area of speech and signal processing at national and international levels. He is a life member of ISTE, senior member of ICEIT and member of IEEE.

Mahesh CHANDRA was born in 1969. He received his B.Sc. from Agra University, Agra, India in 1990 and A.M.I.E. from I.E.I., Kolkatta, India in 1994. He received M.Tech. from J.N.T.U., Hyderabad-India in 2000 and Ph.D. from AMU, Aligarh, India in 2008. His areas of interest are speech, signal and image processing. He has worked as Reader and HOD in the Department of Electronics and Communication Engineering at S.R.M.S. College of Engineering and Technology, Bareilly, India from January 2000 to June 2005. Presently, he is working as an associate professor in the Department of Electronics and Communication Engineering, B.I.T., Mesra, Ranchi (Jharkhand)-India. He has published more than 54 research papers in the area of speech, signal and image processing at national and international levels. He is a Life Member of ISTE, Member of IEEE and IEI, Fellow member of IETE.

S. N. SHARAN has received his Ph.D. from IIT, Delhi, India. His areas of interest are Energy, Signal and Image Processing. He has worked as a professor at BITS, Pilani, India. Presently he is working as a director at GNIT, Gr. Noida, UP, India. He has published more than 30 research papers in the area of Speech, Signal and Image Processing at national and international levels.