

Multi-channel Acoustic Echo Cancellation using Adaptive Fuzzy Fusion Algorithm

Lekshmi .T

*Department of Electronics and communication
SCT College Pappanamcode*

Smitha P .S

*Asst. Professor
Department of Electronics and communication
SCT College Pappanamcode*

Abstract—In multi-channel acoustic echo cancellation(MAEC) system, the non-uniqueness problem occurs due to the correlation between the reference signals, the filter coefficients does not converge to the true echo path impulse response .It could affect convergence performance of the adaptive filtering. In this paper, adaptive fuzzy fusion algorithm is proposed. The new procedure expands on the idea of resampling in the frequency domain that efficiently and effectively alleviates the non-uniqueness problem for a multi-channel acoustic echo cancellation system while introducing minimal distortion to the signal. The amount of decorrelation in each sub band measured in terms of coherence can be controlled arbitrary by varying the parameters. The adaptive fuzzy fusion algorithm improvises, update and check operators obtain optimal solution for defined objective function. To obtain better solution the control parameters are adjusted. It achieves a superior performance in the echo reduction gain and offers the possibility of frequency selective decorrelation to further preserve the sound quality of the system. Simulation result for the proposed algorithm has shown a significant improvement in convergence rate compared with existing system.

Key words—Multi channel AEC, Sub band resampling, non-uniqueness problem, Misalignment problem, coherence.

I. INTRODUCTION

The speech from the far end caller is broadcast by the speaker phone or hand free cellular phone and the repeats itself by bouncing off the inside surface of the room. This repetition of sound is called an echo. Echoes are picked up by the near end microphone, creating a feedback loop where the far end caller hears an echo of his or her own sound. To solve this problem, we use the digital signal processing technique of acoustic echo cancellation(AEC) to stop the feedback and allow full duplex communication.

The effect of an echo mostly depend on the time delay between the initial and reflected sound waves (or sound signals), and the strength of the reflected sounds. In the case of acoustic echo, if the time delay is not long, then the echo can be perceived as soft reverberation, which adds artistic quality. Besides using wideband audio, the application of stereo phonic and multichannel devices adds to the sense of the sound realism in the teleconferencing system. The addition of an audio transmission channel allows users in the near end room to have the ability to distinguish between the speakers in the remote room. As in all acoustic hands free communication systems, acoustic echo cancellation is a requirement. Multichannel acoustic echo cancellation provides an additional set of signal processing challenges as compared to monophonic echo cancellation.

The echo canceller in the local room not only has to track changes in the near end echo paths, but also the echo path of the remote room. Therefore, the echo canceller has to retrain when the far end speaker moves or if a different speaker starts talking. Beside the increased number of echo paths to be modeled, another problem of multichannel acoustic echo cancellation is the non-uniqueness problem. Since the multiple channels will typically have linearly correlated signals the system of equations will have more unknown than known. Therefore there is no unique solution. There are several approaches to handle these problems. One approach is to introduce some distortion to the incoming far-end signal to decorrelate the channels, but obviously one cannot distort too much as it will cause a degradation in the signal quality. Introducing a simple non-linear transformation, such as a half wave rectifier is an effective method of decorrelating the signals without causing much perceptual loss of sound quality.

A number of preprocessing methods have been proposed to decorrelate stereo audio signal in order to enable an unambiguous identification of each echo path and to thus ensure robustness to changing sound source location. Early a two channel preprocessing approach in [1] is used to decorrelate the signals, it has high performance but has low convergence speed. Later propose a new solution based on [2] non-linear transformation to overcome these problems. It improves both the condition number of the covariance matrix and the misalignment not effective. SAEC algorithm [3], [8] establish the relationship between inter channel coherence of the two input signals and condition of the corresponding covariance matrix for stereo acoustic echo cancellation application. The residual echo enhancement (REE) procedure proposed in [4] is able to improve the acoustic echo cancellation performance in a very noisy acoustic mixing environment by utilizing the natural learning ability of the least mean square algorithm without precise estimation of the signal statistics. It demonstrates in this paper that the technique can also be applied effectively in a MCAEC setting to indirectly assist in the recovery of lost AEC performance due to the non-uniqueness problem.

The [5] presents a novel decorrelation procedure by frequency domain resampling in sub band. The new procedure expands on the idea of resampling in the frequency domain that efficiently and effectively alleviates the non-uniqueness problem for a multi-channel acoustic echo cancellation system while introducing minimal distortion to the signal.

The objectives of this work is to extend these contributions, provide a deep analysis of the performance bounds of the resampling procedure, and devise a proper overall scheme for Sub band resampling (SBR). Specifically, we analyze the link between the proposed resampling, the level of decorrelation, and achievable steady-state misalignment performance. We then derive new closed form expressions to demonstrate how resampling affects the misalignment for different types of reference signals, i.e., a white Gaussian noise or a speech signal, with or without the far end room impulse response. Following these analyses, we provide a novel, theoretically justifiable and perceptually motivated SBR strategy for achieving fast AEC convergence with a minimal signal distortion. In this paper we show that the proposed SBR scheme outperforms other decorrelation methods in terms of the convergence rate and the proposed speech quality.

The paper is organized as follows. Section II gives characteristics of multi-channel acoustic echo cancellation. Section III introduces the problems associated with MAEC. Section IV gives methodology and V is the discussions of the proposed method with other existing methods. Conclusions are given in the final section.

II. CHARACTERISTIC ANALYSIS

For the simplicity of the MAEC the echo paths that are linked with one of the microphones are taken. The microphone signal is given by [6]

$$y[n] = z_n[n] + v[n] + \sum_{p=1}^P d_p[n]$$

Where $y[n]$ is the microphone signal, $z_n[n]$ is the near end speech, $v[n]$ is the near end noise. P is the number of loudspeakers. $d_p[n] = \sum_{k=0}^{K-1} h_{p,k} x_p[n-k]$ is the acoustic echo generated by the p th loudspeaker.

For the P echo path, P adaptive filters are required to cancel the echo signal, and the error signal is given by [6],

$$e[n] = y[n] - \sum_{p=1}^P w_{p,c}^T x_{p,c}[n]$$

Where $w_{p,c}^T L[n] = [w_{p,c}^T 0[n] \dots w_{p,c}^T L[n]]^T$ is the adaptive filter coefficient vector for p th echo path. Here the actual length of near end room impulse response (K) is different from that of adaptive filter coefficient vector (L). The minimum mean square error can be computed by solving the normal equation in [6],

$$R_{xx,c}[n] w_{p,c}[n] = r_{yp,c}[n]$$

Where $R_{X_{FC}[n]} = E\{X_{FC}[n]X_{FC}^T[n]\}$ and $r_{yX_{FC}[n]} = E\{y[n]X_{FC}[n]\}$ with E is the expectation operator. The Figure 1 shows the stereo acoustic echo cancellation system with P=2.

III. PROBLEM ANALYSIS

There are mainly two problems associated with multi-channel AEC system, non-uniqueness problem and misalignment problem. The non-uniqueness problem and misalignment problem arises during the multi-channel AEC due to the correlation between the reference signals that degrades the performance of adaptive filtering algorithm[8]

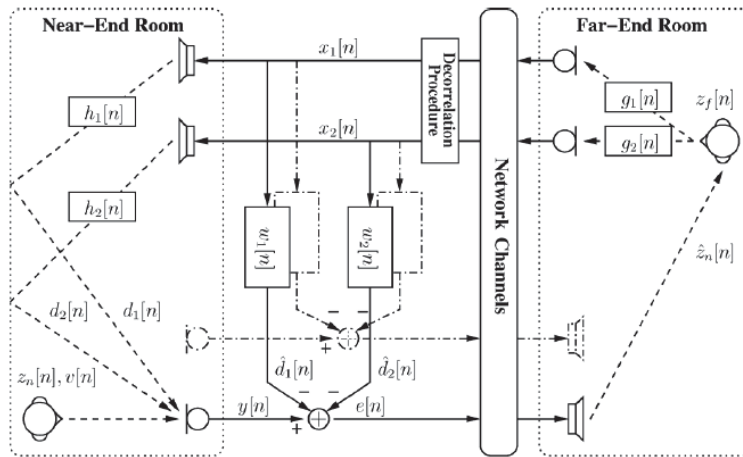


Figure 1. Block diagram of a stereo acoustic echo cancellation system [6]

A. Non-Uniqueness problem

Non uniqueness problem raises if the audio stream originate from the same source. This means that the normal equation to be solved by the adaptive filter is singular. So the echo canceller cannot provide a unique echo path solution. The filter coefficient does not converge to the true echo path impulse response. The fundamental problem that the two audio channels may carry linearly related signals which in turn may cause normal equation to be solved by the adaptive algorithm singular. ie, there is no unique solution to the equation but an infinite number of solution and all these solution depends on the transmission room.

Let $g_p:K$ be the p th far end room impulse response vector, $z_f[n]$ is the far end source signal and $z_n[n] = v[n] = 0$. The p th reference signal is given by [6]

$$x_p[n] = g_p^T z_f[n]$$

That is the solution for Multi-channel AEC depends not only on the near end room impulse response but also the far end room impulse response. The adaptive solution must vary when there is a change in far end room impulse response. This happens due to the multiple users in the room. Whenever the user turns to talk, the room impulse response is changing which causing degradation in the system performance

B. Misalignment problem

The mismatch between the adaptive filter and the actual room impulse response is termed as misalignment which is defined as in [6],

$$\epsilon = \frac{\|h_{FC} - \hat{w}_{FC}\|}{h_{FC}}$$

The length of the adaptive filter can be much lesser than the length of the near end room impulse response.

The coherence gives the amount of correlation between the two reference signals in the frequency domain. The power spectral density (PSD) and the cross spectral density (CSD) is given by [6],

$$S_{X_i X_j}[k] = \sum_{n=-\infty}^{\infty} r_{X_i X_j}[n] \exp^{-jkn} / L$$

Where $r_{X_i X_j}$ be the auto correlation and cross correlation for $i = j$ and $i \neq j$. (7) represent PSD when $i = j$ otherwise represent CSD. The coherence is defined by [6],

$$C_{X_i X_j}[k] = \frac{|S_{X_i X_j}[k]|^2}{S_{X_i X_i}[k] S_{X_j X_j}[k]}$$

C. Resampling technique

A decorrelation process is needed to decrease the coherence between the two signals. Resampling approach is used to address the non-uniqueness problem. The idea decorrelation by resampling is used to exploit the effect of sampling rate mismatch found out in [9]. A slight sampling rate mismatch between the audio channels is enough to decrease the correlation for a sufficient MAEC system. i.e., Decorrelation can be decreased by introducing a sampling rate mismatch between the correlated reference signals.

For discrete time signals, decorrelation is implemented as in [6] by up/down sampling a signal to a different sampling rate $f''s$ and playing back the resampled signal at the original rate fs . The resampling ratio is $R = f''s / fs$. A variable delay can be introduced through different values of resampling rate.

IV. METHODOLOGY

A decorrelation process is needed before the near end playback to improve the signal quality by reducing the non-uniqueness problem. The resampling process is related to time scaling process. $R < 1$ is the time expanding a signal $x(t)$ to $x(t/R)$. The time expanding process is built up a delay over the time between the original signal and the expanded signal. Through this delay, coherence between $x(t)$ and $x(t/R)$ is decrease.

Frequency domain resampling is to interpolate across frequency rather than across time [9], with appropriate expansion or compression, to reduce the computation through Fast Fourier Transform (FFT).

The algorithm used is

- Partitioned the input stereo audio signal to a limited range.
- Echo is created using a chebyshev filter
- Right audio signal is combined with this echo. Then the total signal is left audio signal, right echoed signal and the near end noise.
- An adaptive filter with LMS algorithm is used to remove the echo.
- Finally the step size of the adaptive filter made to adaptive.

Because of the adaptive nature of the adaptive filter step size the better result with maximum Echo Returns Loss Enhancement (ERLE) is obtained.

The Normalized Least Mean Square (NLMS) Algorithm is one of the adaptive algorithm usually used in echo cancellation system. By using NLMS algorithm output of adaptive filter can be computed as

$$y[n] = W^T(n)X(n)$$

Where $X(n) = [x(n) \ x(n-1) \ \dots \ x(n-N+1)]$ is the input vector and $W(n) = [w_0(n) \ \dots \ w_{N-1}(n)]$ is the weight vector. The error signal is $e(n) = d(n) - y(n)$

Next filter coefficient is given by,

$$W(n+1) = W(n) + \mu(n)e(n)x(n)/x^T(n)x(n)$$

Where $\mu(n)$ is the step size factor.

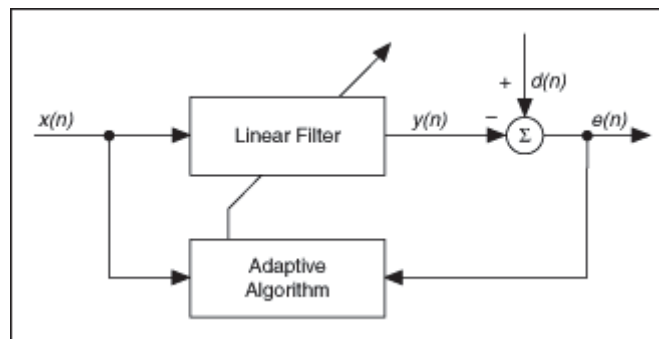


Figure 2. Adaptive filter

A. Adaptive Fuzzy Fusion Algorithm

Adaptive fuzzy fusion algorithm is an effective approach to handle non-linear signals such as stereo audio signals. The proposed procedure expands on the idea of fuzzy based adaptive resampling in the frequency domain, to directly assist the frequency domain adaptive filtering algorithm [10]. Figure shows the Fuzzy Logic Control (FLC) system block diagram. The value of step size is chosen such that adaptive filtering algorithm converges to optimum filter tap weights. If the value of step size is chosen to be large, the filter coefficient may change by a large value and at the second instant the coefficient may keep oscillating with a large variance about the optimal coefficient weights. If step size is chosen to be too small, time required to converge to the optimal coefficient values will be too large. The upper bound of step size is,

$$0 < \mu < 1/\gamma_{max}$$

Where μ is convergence factor and γ_{max} largest eigenvalue of the auto correlation matrix $R = E\{X(n)X^T(n)}$.

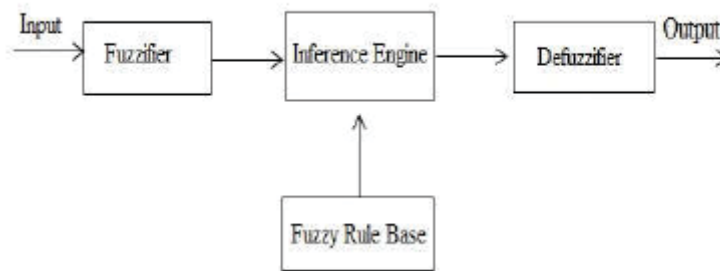


Figure 3. Block diagram of FLC [7]

V. SIMULATION RESULT

The multi-channel acoustic echo cancellation performance is evaluated by using adaptive fuzzy fusion algorithm which gives the optimum result compared with the existing system. Usually echo is measured in terms of ERLE. A higher value of ERLE means minimum amount of echo and smaller value of ERLE means maximum amount of echo. Figure 4 shows the ERLE graph of the proposed method which gives the optimum value compared with the existing system.

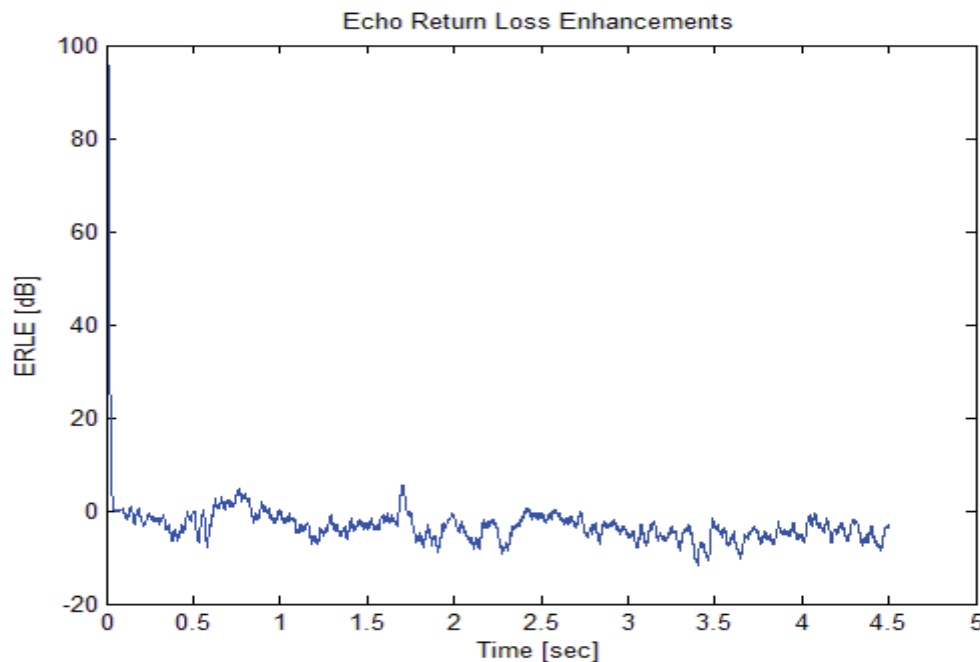


Figure 4. Echo return loss enhancement graph

VI. CONCLUSION

A new approach to Multi channel acoustic echo cancellation is developed using the adaptive fuzzy fusion algorithm. Adaptive fuzzy fusion algorithm is a variable step size adaptation algorithm. From the experimental results the echo removal using the proposed method was found that more quality than other existing algorithms. This is due to the maximum ERLE with minimum error rate.

REFERENCES

- [1] J. Herre, H. Buchner, W. Kellermann, "Acoustic Echo Cancellation For Surrounded Sound Using Perceptually Motivated Convergence Enhancement," IEEE ICASSP.,2007,pp.17-20.
- [2] T.S.Wada and B.H. Juang , "Multi channel acoustic echo cancellation based on residual echo enhancement with effective channel decorrelation via resampling," inProc.IWAENC,2010.
- [3] A. W. H. Khong, J.Benesty,andP.A. Naylor,"Stereophonic Acoustic Echo Cancellation: Analysis of the Misalignment in the Frequency Domain," IEEE Signal process.Lett., vol. 13,no. 1, pp 33-36. Jan.2006.
- [4] J.Wung, T.S.Wada,andB.H.Juang, "Decorrelation by resampling in frequency domain for multi channel acoustic echo cancellation based on residual echo enhancement ,"inProc.IEEE WASPAA,2011,PP.289-292.
- [5] J.Wung,T.S.Wada,and B.H.Juang,"Inter-Channel Decorrelation By Sub- Band Resampling In Frequency Domain," inProc.IEEE ICASSP , 2012, pp.29-32.
- [6] J.Wung,T.S.Wada, and B.H.Juang,"Inter channel decorrelation by sub band resampling for multi channel acoustic echo cancellation," IEEE Trans. on signal processing ,vol. 62,no. 8,April 15,2014
- [7] A.M.Geethu,K.S.Smitha, and D.Chengzhi,"A Fuzzy Logic Based Acoustic Echo Cancellation System,"International Journal of Engineering and Advanced Technology ,vol. 4, no. 6,August 2015, ISSN.2249-8958.
- [8] M.M.Sondhi,D.R.Morgan,andJ.L.Hall,"Stereo phonic acoustic echo cancellation -An overview of the fundamental problem,"IEEE Signal processing. Lett,vol. 2,no. 8,pp.148-151,Aug.1995.
- [9] M.M.Sondhi,D.R.Morgan,andJ.L.Hall,"On dealing with sampling rate mismatches in blind source separation and acoustic echo cancellation ,"inProc. IEEE WASPAA.,2007,pp.34-7