Echo Cancellation In Long Distance Telephone Circuits Using VSS-APA

Rajashree.s(ME) and Mahendran.n(HOD)
Dept of electronics and communication engineering
Ck College of Engineering and Technology
Cuddalore

ABSTRACT: In telephone circuits, Acoustic Echo Cancellation (AEC) provides one of the best solutions to the control of acoustic echoes generated by hands-free audio terminals. In this type of application, an adaptive filter identifies the acoustic echo path between the terminal's loudspeaker and microphone, i.e., the room impulse response. The filter output, which provides an electronic replica of the acoustic echo, is subtracted from the microphone signal to cancel the echo. Nevertheless, there are several specific and challenging problems associated with AEC applications. Hence, the residual echo caused by the part of the system that cannot be modeled acts like an additional noise and disturbs the overall performance. Second, the background noise that corrupts the microphone signal can be strong and highly non stationary. In my project, the Variable Step-Size Affine Projection Algorithm (APA) and some of its versions were found to be very attractive choices for AEC applications. The main advantage of this family of algorithms over the well known Normalized Least-Mean-Square (NLMS) algorithm consists of a superior convergence rate, especially for speech signals. This algorithm does not require any a priori information about the acoustic environment, so that it is easy to control in practice. The simulation results indicate the good performance of the proposed algorithm as compared to other members of the APA family.

I. INTRODUCTION

Acoustic Echo Cancellation (AEC) provides one of the best solutions to the control of acoustic echoes generated by hands-free audio terminals –. In this type of application, an adaptive filter identifies the acoustic echo path between the terminal's loudspeaker and microphone, i.e., the room impulse response. The filter output, which provides an electronic replica of the acoustic echo, is subtracted from the microphone signal to cancel the echo. Nevertheless, there are several specific and challenging problems associated with AEC applications. First, the echo path is extremely long (on the order of hundreds of milliseconds) and it may rapidly change at any time during the connection. The excessive length of the acoustic echo path in time is mainly due to the slow speed of sound through air; moreover, multiple reflections of walls and objects in the room increase this length. In addition, the impulse response of the room is not static overtime, since it varies with the ambient temperature, pressure, and humidity; also, movement of objects and human bodies can rapidly modify the acoustic impulse response. As a consequence of these aspects related to the acoustic echo path characteristics, the adaptive filter works most likely in an under-modeling situation, i.e., its length is smaller than the length of the acoustic impulse response. Hence, the residual echo caused by the part of the system that can not be modeled acts like an additional noise and disturbs the overall performance. Second, the background noise that corrupts the microphone signal can be strong and highly non stationary.

II. PROPOSED SYSTEM

In this system, the affine projection algorithm (APA) and different versions of it are very attractive choices for aec. however, an APA with a constant step-size parameter has to compromise between the performance criteria. therefore, a variable step-size APA (VSS-APA) represents a more reliable solution. in this paper, we propose a VSS-APA derived in the context of aec. most of the apas aim to cancel (i.e., projection order) previous a posteriori errors at every step of the algorithm. the proposed VSS-APA aims to recover the near-end signal within the error signal of the adaptive filter. consequently, it is robust against near-end signal variations (including double-talk). this algorithm does not require any a priori information about the acoustic environment, so that it is easy to control in practice.

III. SYSTEM DESIGN

A general AEC configuration is depicted in Fig. 1. The goal of this scheme is to identify an unknown system (i.e., acoustic echo path) using an adaptive filter. Both systems have finite impulse responses, defined by the real-valued vectors and , where superscript denotes transposition and is the time index; is the length of the echo path, while is the length of the adaptive filter. The signal is the far-end speech which goes through the acoustic impulse response , resulting the echo signal, . This signal is picked up by the microphone together with the near-end signal , resulting the microphone signal . The near-end signal can contain both the background noise, , and the near-end speech.

The output of the adaptive filter provides a replica of the echo, which will be subtracted from the microphone signal. The DTD block controls the algorithm behavior during double-talk; nevertheless, the proposed algorithm will be derived without involving the DTD decision.

The following relations define the classical APA

$$e(n)=d(n)-XT(n)\hat{h}(n-1) \qquad (1)$$

$$\hat{h}(n)=\hat{h}(n-1)+\mu X(n)[XT(n)X(n)]-1e(n) \qquad (2)$$

Where d is the desired signal vector of length , with denoting the projection order. The matrix is the input signal matrix, where (with) are the input signal vectors. The constant denotes the step-size parameter of the algorithm.

Let us rewrite (2) in a different form

$$\hat{h}(n)=\hat{h}(n-1)+X(n)[XT(n)X(n)]-1\mu(n)e(n) \qquad (3)$$

Where,

$$\mu(n)=diag\{\mu 0(n),\mu 1(n),\ldots,\mu p-1(n)\} \qquad (4)$$

Using the adaptive filter coefficients at time , the a posteriori error vector can be defined as

$$\varepsilon(n)=d(n)-XT(n)\hat{h}(n) \qquad (5)$$

It can be noticed that the vector from (1) plays the role of the a priori error vector. Replacing (3) in (5) and taking (1) into account, it results that

$$\varepsilon(n)=[Ip-\mu(n)]e(n) \qquad (6)$$

where denotes a identity matrix. In consistence with the basic idea of the APA, it can be imposed to cancel a posteriori errors, i.e., , where denotes a column vector with all its elements equal to zeros. Assuming that , it results from (6) that . This corresponds to the classical APA update (2), with the step-size . In the absence of the near-end signal, i.e., , the scheme from Fig. 1 is reduced to an ideal “system identification” configuration. In this case, the value of the step-size makes sense, because it leads to the best performance Nevertheless, it can be noticed that the AEC scheme from Fig. 1 can be interpreted as a combination between two classes of adaptive system configurations (according to adaptive filter theory –).

First, it represents a “system identification” configuration, because the goal is to identify an unknown system (i.e., the acoustic echo path) with its output corrupted by an apparently “undesired” signal (i.e., the near-end signal). However, it also can be viewed as an “interference cancelling” configuration, aiming to recover an “useful” signal (i.e., the near-end signal) corrupted by an undesired perturbation (i.e., the acoustic echo); consequently, the “useful” signal should be recovered in the error signal of the adaptive filter. Therefore, since the existence of the near-end signal cannot be omitted in AEC, a more reasonable condition is , where the column vector represents the near-end signal vector of length . Taking (6) into account, it results that

$$\varepsilon l+1(n)=[1-\mu l(n)]el+1(n)=v(n-1) \qquad (7)$$

where the variables and denote the th elements of the vectors and , with . The goal is to find an expression for the step-size parameter such that

$$E\{\varepsilon 2l+1(n)\}=E\{v2(n-1)\} \qquad (8)$$

where denotes mathematical expectation. Squaring and taking the expectations results in

$$[1-\mu l(n)]2\ E\{e2l+1(n)\}=E\{v2(n-1)\} \qquad (9)$$

By solving the quadratic (9), two solutions are obtained, i.e.,

$$\mu_l(n)=1\pm \frac{E\{v^2(n-1)\}}{\sqrt{E\{e^2_{l+1}(n)\}}} \qquad (10)$$

Following the analysis from , which states that a value of the step-size between 0 and 1 is preferable over the one between 1 and 2 (even if both solutions are stable, but the former has less steady-state mean square error with the same convergence speed), it is reasonable to choose

$$\mu_l(n)=1- \quad \underline{E\{v^2(n-1)\}} \qquad (11)$$

$$\sqrt{\ E\{e^2_{l+1}(n)\}}$$

From a practical point of view, has to be evaluated in terms of power estimates as

$$\mu_l(n) = 1 - \frac{\hat{\sigma}_v(n-l)}{\hat{\sigma}_{e_{l+1}}(n)}. \quad (12)$$

The variable in the denominator can be computed in a recursive manner, i.e.,

$$\hat{\sigma}^2_{e_{l+1}}(n) = \lambda\hat{\sigma}^2_{e_{l+1}}(n-1) + (1-\lambda)e^2_{l+1}(n) \quad (13)$$

where is a weighting factor chosen as , with ; the initial value is . The estimation of is not straightforward in real world applications like AEC. In this case, is the near-end signal (i.e., background noise or/and near-end speech) which is combined together with the acoustic echo, resulting the microphone signal (the only signal that is practically available). Several scenarios could be considered, as follows.

1) Single-Talk Scenario: In the single-talk case, the near-end signal consists only of the background noise, . Its power could be estimated during silences (and it can be assumed constant), so that becomes

$$\mu_l(n) = 1 - \frac{\hat{\sigma}_w}{\hat{\sigma}_{e_{l+1}}(n)}. \quad (14)$$

$$d(n) = y(n) + v(n) \quad (15)$$

$$E\{d^2(n)\} - E\{y^2(n)\} = E\{v^2(n)\}. \quad (16)$$

Also, let us assume that the adaptive filter coefficients have converged to a certain degree, so that

$$E\left\{y_L^2(n)\right\} \cong E\{\hat{y}^2(n)\}. \quad (17)$$

$$E\{v^2(n)\} + E\{q^2(n)\} = E\{d^2(n)\} - E\{\hat{y}(n)\}. \quad (18)$$

III. VSS-APA ALGORITHM

Initialization: $\hat{\mathbf{h}}(0)=\mathbf{0}_{L\times 1}$, $\hat{\sigma}_d^2(0)=0$, $\hat{\sigma}_{\hat{y}}^2(0)=0$

for $l=0$ to $p-1$

$\hat{\sigma}_{e_{l+1}}^2(0)=0$

For time index $n=1,2,\ldots$

$$\mathbf{e}(n)=\mathbf{d}(n)-\mathbf{X}^T(n)\hat{\mathbf{h}}(n-1)$$

$$\hat{y}(n)=\mathbf{x}^T(n)\hat{\mathbf{h}}(n-1)$$

$$\hat{\sigma}_d^2(n)=\lambda\hat{\sigma}_d^2(n-1)+(1-\lambda)d^2(n)$$

$$\hat{\sigma}_{\hat{y}}^2(n)=\lambda\hat{\sigma}_{\hat{y}}^2(n-1)+(1-\lambda)\hat{y}^2(n)$$

for $l=0$ to $p-1$

$$\hat{\sigma}_{e_{l+1}}^2(n)=\lambda\hat{\sigma}_{e_{l+1}}^2(n-1)+(1-\lambda)e_{l+1}^2(n)$$

$$\mu_l(n)=\left|1-\frac{\sqrt{\left|\hat{\sigma}_d^2(n-l)-\hat{\sigma}_{\hat{y}}^2(n-l)\right|}}{\xi+\hat{\sigma}_{e_{l+1}}(n)}\right|$$

$$\boldsymbol{\mu}(n)=\mathrm{diag}\{\mu_0(n),\mu_1(n),\ldots,\mu_{p-1}(n)\}$$

$$\hat{\mathbf{h}}(n)=\hat{\mathbf{h}}(n-1)+\mathbf{X}(n)\left[\delta\mathbf{I}_p+\mathbf{X}^T(n)\mathbf{X}(n)\right]^{-1}\boldsymbol{\mu}(n)\mathbf{e}(n)$$

Summarizing, the proposed VSS-APA is listed in Table I. A block diagram of this algorithm is presented in Fig., where the blocks with gray background indicate the specific operations of the proposed algorithm. As compared to the classical APA, the additional computational amount of the VSS-APA consists of multiplication operations, divisions, additions, and square-root operations. Taking into account the fact that the value of the projection order in AEC applications is usually smaller than 10 (e.g., common values could be , and 8), and the length of the adaptive filter is large (e.g., hundred of coefficients), it can be concluded that the computational complexity of the proposed VSS-APA is moderate and comparable with the classical APA or to any fast versions of this algorithm.

A. Block Diagram of VSS-APA

Define abbreviations and acronyms the first time they are used in the text, even after they have been defined in the abstract. Abbreviations such as IEEE, SI, MKS, CGS, sc, dc, and rms do not have to be defined. Do not use abbreviations in the title or heads unless they are unavoidable.

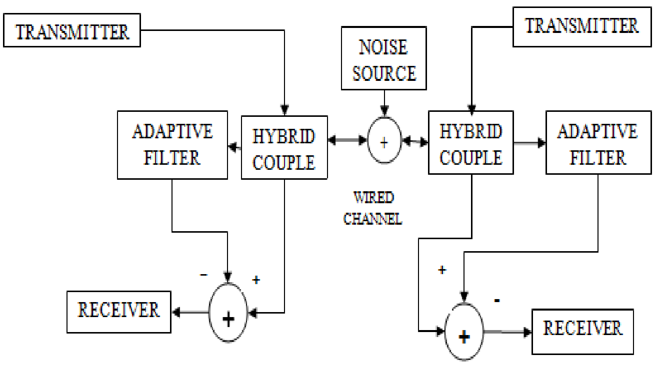

Fig 1: Block diagram of echo cancellation

It is interesting to notice that the step-size of the proposed VSS-APA does not depend explicitly on the near-end signal, even if it was developed taking into account its presence; consequently, a robust behavior under near-end signal variations (e.g., background noise variations and double-talk) is expected. Moreover, since only the parameters available from the adaptive filter are required and there is no need for a priori information about the acoustic environment, the proposed algorithm is easy to control in practice.

IV. SOFTWARE SPECIFICATION

MATLAB (matrix laboratory) is a multi-paradigm numerical computing environment. A proprietary programming language developed by Math Works, MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, creation of user interfaces, and interfacing with programs written in other languages, including C, C++, C#, Java, Fortran and Python. Although MATLAB is intended primarily for numerical computing, an optional toolbox uses the MuPAD symbolic engine, allowing access to symbolic computing abilities.

An additional package, Simulink, adds graphical multi-domain simulation and model-based design for dynamic and embedded systems. MATLAB is a high-performance language for technical computing. It integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical notation. Typical uses include:

- Math and computation
- Algorithm development
- Modeling, simulation, and prototyping
- Data analysis, exploration, and visualization
- Scientific and engineering graphics

A. The Matlab Language

This is a high-level matrix/array language with control flow statements, functions, data structures, input/output, and object-oriented programming features. It allows both "programming in the small" to rapidly create quick and dirty throw-away programs, and "programming in the large" to create complete large and complex application programs.

B. The Matlab Working Environment

This is the set of tools and facilities that you work with as the MATLAB user or programmer. It includes facilities for managing the variables in your workspace and importing and exporting data. It also includes tools for developing, managing, debugging, and profiling M-files, MATLAB's applications.

C. Handle graphics

This is the MATLAB graphics system. It includes high-level commands for two-dimensional and three-dimensional data visualization, image processing, animation, and presentation graphics. It also includes low-level commands that allow you to fully customize the appearance of graphics as well as to build complete Graphical User Interfaces on your MATLAB applications.

V. SOFTWARE SPECIFICATION

```
clear all
close all
hold off
%channel system order
sysorder = 5 ;
% Number of system points
N=1000;
inp = randn(N,1);
n = randn(N,1);
[b,a] = butter(2,0.25);
Gz = tf(b,a,-1);
h=  [0.0976;
    0.2873;
```

```
    0.3360;
    0.2210;
    0.0964;];
y = lsim(Gz,inp);
%adding noise
n = n * std(y)/(10*std(n));
d = y + n;
totallength=size(d,1);
N=60 ;
%VSSAP algorithm
w = zeros ( sysorder  , 1 ) ;
for n = sysorder : N
    u = inp(n:-1:n-sysorder+1) ;
    y(n)= w' * u;
    e(n) = d(n) - y(n) ;
    if n < 20
        mu=0.32;
    else
        mu=0.15;
    end
    w = w + mu * u * e(n) ;
end
%checking results
for n =  N+1 : totallength
    u = inp(n:-1:n-sysorder+1) ;
    y(n) = w' * u ;
    e(n) = d(n) - y(n) ;
end
nvar  = 0.1;
ip = randn(1000,1)*nvar;
hold on
subplot(3,1,3);
plot(d)
plot(y,'r');
title('System output') ;
xlabel('Samples')
ylabel('True and estimated output')
subplot(3,1,2);
semilogy((abs(e))) ;
title('Error curve') ;
xlabel('Samples')
ylabel('Error value')
subplot(3,1,1);
plot(0:999,ip);
title('Input Signal');
xlabel('Samples')
ylabel('Input')
grid; axis([0 1000 -4 4]);
```

VI. RESULTS AND DISCUSSION

It would be very interesting to compare the proposed VSS-APA with most of the VSS algorithms from the APA family and also with most of the double-talk robust algorithms. However, it is beyond the scope of this paper. Consequently, we choose to limit the framework only to the APA family by comparing the proposed solution with the regular APA and also with two other versions of it, as follows.

For the single-talk scenarios, the variable regularized APA (VR-APA) recently proposed inis included in our comparisons. The motivation behind this choice is twofold. First, due to its nature, this algorithm can be also considered a VSS-APA. Second, the experimental results presented inshow that it outperforms other VSS-APAs. The VR-APA update is

$$\hat{h}(n)=\hat{h}(n-1)+X(n)[\delta Ip+XT(n)X(n)]-e(n)$$

Its variable regularization factor is given by

$$\delta(n) = \min\left\{\frac{L\sigma_x^2}{\zeta}, \frac{p\sigma_w^2\sigma_x^2 L}{\hat{\sigma}_e^2(n) - p\sigma_w^2}\right\}$$

where is a positive design parameter. The power of the background noise and the power of the input signal have to be known within the algorithm, while the term is evaluated as

$$\hat{\sigma}_e^2(n) = \lambda\hat{\sigma}_e^2(n-1) + (1-\lambda)\|\mathbf{e}(n)\|^2$$

For the scenarios where there are near-end signal variations (i.e., background noise variations or double-talk), the robust proportionate APA (R-PAPA) proposed inis considered for comparison. Developed in the framework of robust statistics, this algorithm was initially applied for network echo cancellation, but this idea was found to be efficient even for less sparse echo path like in AEC. The R-PAPA is not included in the other single-talk scenarios since it is outperformed by the regular APA in these cases. The simulations were performed in an AEC context, as shown in Fig.

The length of the adaptive filter is set to 512 coefficients. The measured impulse response of the acoustic echo path is plotted in Fig. 3(a) (the sampling rate is 8 kHz); its entire length has 1024 coefficients. This length is truncated to the first 512 coefficients for a first set of experiments performed in an exact modeling case. Then, the entire length of the acoustic impulse response is used for a second set of experiments performed in the under-modeling case. The far-end signal is either an AR(1) process generated by filtering a white Gaussian noise through a first-order system , or a speech sequence . For the double-talk scenarios, the near-end speech is plotted in Fig. 3(c). An independent white Gaussian noise signal is added to the echo signal, with 20-dB signal-to-noise ratio (SNR) for most of the experiments. It is assumed that the power of the background noise is known for the VR-APA. The weighting factor (for VR-APA and VSS-APA) is computed.

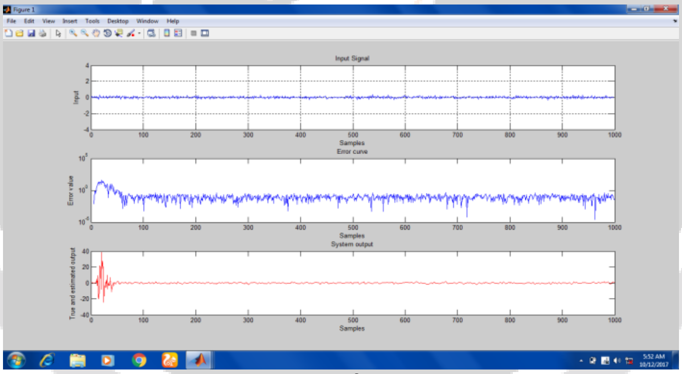

Fig 2: Output Waveform

VII. CONCLUSION

Variable Step-Size Affine Projection Algorithm (APA) and some of its versions were found to be very attractive choices for AEC applications. The main advantage of this family of algorithms over the well known normalized least-mean-square (NLMS) algorithm consists of a superior convergence rate, especially for speech signals. This algorithm does not require any a priori information about the acoustic environment, so that it is easy to control in practice. It have the good performance of the proposed algorithm as compared to other members of the APA family.

References

[1] Martin Schneider and Walter Kellermann," Multichannel Acoustic Echo Cancellation in the Wave Domain With Increased Robustness to Nonuniqueness", IEEE/ACM transactions on audio, speech, and language processing, vol. 24, no. 3, pp 518-529, march 2016.

[2] Sheng Zhang, Jiashu Zhang, and Hongyu Han," Roboust variable step size decorrelation normalized mean square algorithm and its application to acoustic echo cancellation", IEEE/ACM Transactions on Audio, Speech, and Language Processing,2016.

[3] Prof.S.Veerabhadraiah and Dr. M.Purnachandra Rao,"Application of Adaptive Filters in Desired signal Extraction", International conference on Signal Processing, Communication, Power and Embedded System (SCOPES)-2016.

[4] Silviu Ciochin, Constantin Paleologu, Jacob Benest, and Steven L. Grant," An Optimized NLMS Algorithm for Acoustic Echo Cancellation", IEEE/ACM Transactions on Audio, Speech, and Language Processing,2015.

[5] Pratik Shah, Ian Lewis, Steven Grant, Sylvain Angrignon,"Nonlinear Acoustic Echo Cancellation using Voltage and Current Feedback", IEEE/ACM Transactions on Audio, Speech, and Language Processing,2015.

[6] Keunsang Lee, Yonghyun Baek, and Youngcheol Park,"Nonlinear Acoustic Echo Cancellation Using a Nonlinear Postprocessor With a Linearly Constrained Affine Projection Algorithm", IEEE transactions on circuits and systems—ii: express briefs, vol. 62, no. 9,pp 881-886, september 2015.

[7] Haiquan Zhao, Yi Yu, Shibin Gao, Xiangping Zeng, and Zhengyou He," Memory Proportionate APA with Individual ctivation Factors for Acoustic Echo Cancellation," IEEE/ACM transactions on audio, speech, and language processing, vol. 22, no. 6,pp 1047-1056 june 2014.

[8] Jose M. Gil-Cacho, Toon van Waterschoot, Marc Moonen, and Søren Holdt Jensen,"A frequency-domain adaptive filter (FDAF) prediction error method (PEM) framework for double-talk-robust acoustic echo cancellation", IEEE/ACM Transactions on Audio, Speech, and Language Processing,2014.

[9] Sethaphak Sukhumalchayaphong and Chawalit Benjangkaprasert," Variable Forgetting Factor RLS Algorithm for Adaptive Echo Cancellation", 2014 14th International Conference on Control, Automation and Systems (ICCAS 2014),korea.

[10] Danilo Comminiello, Michele Scarpiniti, Luis A. Azpicueta-Ruiz, Jerónimo Arenas-García, and Aurelio Uncini, ," Nonlinear Acoustic Echo Cancellation Based on Sparse Functional Link Representations", IEEE/ACM transactions on audio, speech, and language processing, vol. 22, no. 7,pp 1172-1183, july 2014.

[11] Inseon Jang, Hyun Woo Kim, Do Young Kim," Acoustic Echo Cancellation for the Telepresence System", IEEE/ACM Transactions on Audio, Speech, and Language Processing,2013.

[12] Li Xu and Chu-lin Hou," An Improved Affine Projection Algorithm and Its Application in Acoustic Echo Cancellation,"3rd International Congress on Image and Signal Processing (CISP2010), china.

[13] Constantin Paleologu, Jacob Benesty, and Silviu Ciochina," A Variable Step-Size Affine Projection Algorithm Designed for Acoustic Echo Cancellation", IEEE transactions on audio, speech, and language processing, vol. 16, no. 8,pp 1466-1479, november 2008.

[14] W. Tingchan, V. Chutchavong and C. Benjangkaprasert," Performance of A Robust Variable Step-Size LMS Adaptive Algorithm for Multiple Echo Cancellation in Telephone Network", SICE-ICASE International Joint Conference 2006.

[15] Sen M. Kuo1 and Woon S. Gan2," Analysis of Nonlinear Residual Echo Suppressors for Telecommunications", IEEE/ACM Transactions on Audio, Speech, and Language Processing,pp 5994-5998,2005.

[16] Jianfeng Liu," Efficient and Robust Cancellation of Echoes With Long Echo Path Delay", IEEE transactions on communications, vol. 52, no. 8,pp 1288-1291, august 2004.

[17] Leonard G. Neumeyer, Vassilios V. Digalakis, and Mitchel Weintraub," Training Issues and Channel Equalization Techniques for the Construction of Telephone Acoustic Models Using a High-Quality Speech Corpus", IEEE transactions on speech and audio processing. vol. 2, no. 4,pp 590-598, october 1994.

[18] Philip C. Yip and Delores M. Etter," An Adaptive Technique for Multiple Echo Cancelation in Telephone Networks," IEEE/ACM Transactions on Audio, Speech, and Language Processing,1990

[19] .19.B. Widrow and S.D. Stearns, "Adaptive signal processing",1985.

[20] Stephen B. Weinstein,"Echo cancellation in telephone network", IEEE communications society magazine,1977.Ali Adid Amb, Sheikh Afrin , F.M Fahad and Hasan U. Zaman, "Acost Effective Way to Build a Web Controlled Search and CO Detector Rover" , IEEE Transcation (2017).