A Survey on Adaptive Algorithms for Noise Cancellation in Speech Signals

1G.Amjad Khan, 2Dr.K.E.Sreenivasa Murthy
1Assistant Professor, ECE Dept., 2Principal
1G.Pulla Reddy Engg. College (autonomous), Kurnool A.P.
2Sri Sai Insititue of Technology, Raichoti, Kadapa, A.P.
Gak80rec@gmail.com
kesmurthy@rediffmail.com

***Abstract*--The contamination of a signal of interest by other undesired signals (noise) is a problem encountered in many applications. The conventional linear digital filters with fixed coefficients exhibit a satisfactory performance in extracting the desired signal when the signal and noise occupy fixed and separate frequency bands. However, in most applications, the desired signal has changing characteristics which requires an update in the filter coefficients for a good performance in the signal extraction. Since the conventional digital filters with fixed coefficients do not have the ability to update their coefficients, adaptive digital filters are used to cancel the noise. The parameters of adaptive filters are mean square error (MSE) ,signal to noise ratio(SNR) ,convergence and computation time ,Misadjustment are the various techniques which gives the performance of the adaptive filters.**

***Keywords*—Adaptive filtering, Convergence time, LMS NLMS, CS-NLMS.**

I. INTRODUCTION

The objective of this study is to understanding the adaptive filter (AF) theory. This work will show the theory behind the adaptive filters and it will give examples of some applications. The idea of the study is that after consolidating the knowledge of this high level control technique (the adaptive filters) the researcher will have a huge range of application in most diverse areas. There will be presented possible algorithm's solutions and their performance results for some applications. Moreover, the work focuses on one class of application which is the main goal of the research. It is the interference cancelling (IC) also known as noise cancelling (NC). This paper begins with a succinct overview of adaptive filters. Furthermore, it introduces the final objective of the research: the adaptive noise cancellation (ANC) problem David S. Leeds has described an application of adaptive noise cancellation in the form of an adaptive multiple notch filter, to remove transmitter burst envelope noise that may be induced in high gain audio circuits of a wireless handset device. Since samples of the transmitter burst envelope are correlated to the additive noise that corrupts the sampled speech, therefore use this as a reference input to the ANC and also obtain the reference input by delaying the primary input to de-correlate the speech. material presented below can be found, for example, in [2].

Speech is the most fundamental way of communication used by human being. Speech signal is one dimensional signal and has comparatively low bandwidth of 8 kHz. This is very useful in various applications with different evolving technologies. Noise is an undesirable signal. When it is mixed with speech signal, it is very difficult to deliver actual information from one place to another. To differentiate original information containing signal from unwanted noise; filtering is a basic and common method. Filters extract useful information from its input depending upon its configuration [1].

II. ADAPTIVE FILTERING

As their own name suggests, adaptive filters are filters with the ability of adaptation to an unknown environment. This family of filters has been widely applied because of its versatility (capable of operating in an unknown system) and low cost (hardware cost of implementation, compared with the non-adaptive filters, acting in the same system). The ability of operating in an unknown environment added to the capability of tracking time variations of input statistics makes the adaptive filter a powerful device for signal processing and control applications [1]. Indeed, adaptive filters can be used in numerous applications and they have been successfully utilized over the years. As it was before mentioned, the applications of adaptive filters are numerous. For that reason, applications are separated in four basic classes: identification, inverse modelling, prediction and interference cancelling. These classes will be detailed in the next chapter. All the applications above mentioned, have a common characteristic: an input signal is received for the adaptive filter and compared with a desired response, generating an error. That error is then used to modify the adjustable coefficients of the filter, generally called weight, in order to minimize the error and, in some optimal sense, to make that error being optimized, in some cases tending to zero, and in another tending to a desired signal.

.

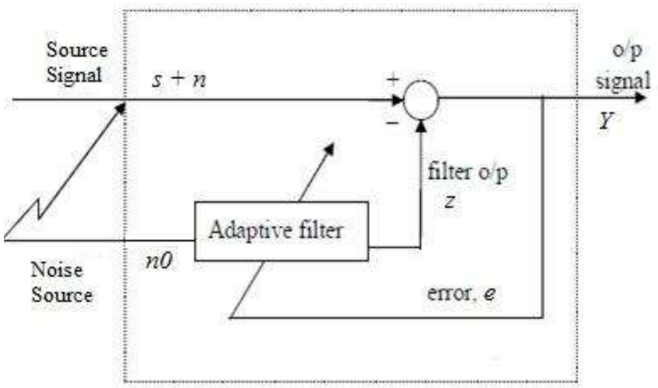

Fig. 1 Basic Adaptive Noise canceller

Adaptive noise canceller receives two inputs namely primary input and reference input. Primary input is combination of source signal *s* and noise signal *n*, uncorrelated with each other. While the reference input is another noise signal *n0*, correlated in some extent with noise *n* only. Reference input *n0* goes through adaptive filter producing *z* which is near estimate of primary input. This filter output is subtracted from primary input resulting in some residue i.e. error. Main purpose of using adaptive noise canceller here is to have output which is best fit in the least squares sense to the signal *s*. To achieve output, the error is fed back to adaptive filter that adjusts the filter using adaptive algorithm [4]. Thus minimizes total system output power.

Linear adaptive filters work on recursive algorithm which considers following important factors [1]

A.Rate of convergence: This is defined as the number of iterations required for the algorithm, in response to stationary inputs, to converge " close enough" to the optimum Wiener solution in the mean-square error sense. A fast rate of convergence allows the algorithm to adapt rapidly to a stationary environment of unknown statistics.

B.Mis-adjustment: For an algorithm of interest, this parameter provides a quantitative measure of the amount by which the final value of the mean square error, averaged over an ensemble of adaptive filters, deviates from the minimum mean square error produced by the Wiener filter.

C.Tracking: When an adaptive filtering algorithm operates in a non-stationary environment, the algorithm is required to track statistical variations in the environment. The tracking performance of the algorithm, however, is influenced by two contradictory features:

D.Robustness: For an adaptive filter to be a robust, small disturbances can only result in small estimation errors .The disturbances may arise from a variety of factors, internal or external to the filter.

E.Computational requirements: Here the issues of concern include (a) the number of operations required to make one complete iteration of the algorithm, (b) the size of memory locations required to store the data and the program, and (c) the investment required to program the algorithm on a computer.

F.Structure: This refer to the structure of information flow in the algorithm, determining the manner in which it is implemented in hardware form. For example, an algorithm whose structure exhibits high modularity, parallelism, or concurrency is well suited for implementation using very large scale integration

G.Numerical properties: When an algorithm is implemented numerically, inaccuracies are produced due to quantization errors, which in turn are due to analog to digital conversions of the input data and digital representation of internal calculations. It is latter source of quantization errors that poses a serious design problem. There are two basic issues of concern: numerical stability and numerical accuracy. Numerical stability is an inherent characteristic of an adaptive filtering algorithm. Numerical accuracy, on other hand, is determined by number of bits used in the numerical representation of the data samples and filter coefficients.

III. ADAPTIVE ALGORITHMS

As explained above, adaptive filter uses adaptive algorithm for adjustment of internal parameters to have correct output. Some basic adaptive algorithms along with approaches on which they are based are explained below:

A. Stochastic Gradient Approach

To develop a recursive algorithm for updating the tap weight of a transversal filter .we use an iterative procedure to solve the wiener solution. The iterative procedure is based on the method of steepest descent which is a well-known technique in optimization theory. Adaptive filter precisely uses second order function of tap weights. Updating of tap weights for recursive algorithm is carried out using Wiener-Hopf equations [1]. Basic algorithms under this approach are as follows

1) Least Mean Square (LMS): This algorithm was introduced by Widrow and Hoff in 1959[5]. Every tap weight in filter is updated in the direction of the gradient of the squared amplitude of an error signal with respect to that tap weight [2].This technique doesn't require prior knowledge of the detailed properties of the noise signal. In case of speech signal this technique takes advantage of the quasi-periodic nature of the speech [6]. The weight update equation for the least mean square algorithm is given coefficient update equation is given by,

$$w(n+1) = w(n) + \mu e^*(n)x(n) \qquad (3.1)$$

where $x(n)$ is the input Signal, $y(n)$ is the adaptive filter output and μ is the step size (or) convergence Parameter

$$y(n) = W^T(n)x(n) \qquad (3.2)$$

The error signal is given by the

$$e(n) = d(n) - y(n) \qquad (3.3)$$

Draw backs of LMS algorithm:

(i) LMS algorithm suffers from slower convergence rate.

(ii)The major disadvantage of LMS algorithm is its excess mean squared error, or misadjustment, which increases linearly with the desired signal power.

The main drawback of the LMS algorithm is that it is sensitive to the scaling of its input x(n). This makes it very hard to choose a step size μ that guarantees stability of the algorithm.

Convergence and stability in the mean of LMS: As the LMS algorithm does not use the exact values of the expectations, the weights would never reach the optimal weights in the absolute sense, but a convergence is possible in mean. That is even-though, the weights may change by small amounts, it changes about the optimal weights. However, if the variance, with which the weights change, is large, convergence in mean would be misleading. This problem may occur, if the value of step size μ is not chosen properly. Thus, an upper bound on μ is needed which is given as $0 < \mu < 2\ \lambda max$ Where λmax is an autocorrelation matrix, its Eigen vales are non-negative. If this condition is not fulfilled, the algorithm becomes unstable. The convergence of the algorithm [4] is inversely proportional to the Eigen value spread of the correlation matrix R. When the Eigen values of R are widespread, convergence may be slow. The Eigen value spread of the correlation matrix is estimated by computing the ratio of the largest Eigen value to the smallest Eigen value of the matrix. If μ is chosen to be very small then the algorithm converges very slowly. A large value of μ may lead to a faster convergence but may be less stable around the minimum value. Maximum convergence speed [4] is achieved when μ = 2 λ max +λmin Where λmin is the smallest Eigen value of R. Given that μ is less than or equal to this optimum, the convergence speed is determined by λmin, with a larger value yielding faster convergence. This means that faster convergence can be achieved when λmax is close to λmin, that is, the maximum achievable convergence speed depends on the Eigen value spread of R.

2. NORMALISED LEAST MEAN SQUARE (NLMS) ALGORITHM: The main drawback of the "pure" LMS algorithm is that it is sensitive to the scaling of its input. This makes it very hard to choose a learning rate μ that guarantees stability of the algorithm. The Normalised least mean squares (NLMS) filter [6], [7] is a variant of the LMS algorithm [1] that solves this problem by normalising with the power of the input.

To overcome this difficulty,we may use the "Normalized LMS filter "The adjustment applied vector at n+1 iteration is normalised with respect to the squared Euclidean norm of the tap input vector u(n) at n Hence the term"Normalised".

$$\hat{w}(n+1) = \hat{w}(n) + \frac{\widetilde{\mu(n)}}{\|x(n)\|^2} x(n)e^*(n) \qquad (3.4)$$

The product Vector $x(n)e^*(n)$is normalised with respect to the Squared Euclidean Norm of the tap input vector$x(n)$.

Advantages:

The Normalized LMS filter algorithm exhibits a rate of convergence that is faster than that of the standard LMS Algorithm for both uncorrelated and correlated input data.

. Moreover, NLMS provides faster Rate of Convergence, rapid tracking and low misalignment.

Disadvantage:

However, the NLMS algorithm has a problem; when the input vector $x(n)$ is small, then a rise of numerical difficulties may occur because by then we have to divide by a small value for tap-input Power $\|x(n)\|^2$.

3. Constrained Stability LMS Algorithm: A common major drawback of adaptive noise canceller based on LMS and NLMS algorithms is the large value of excess mean square error which results in signal distortion in the noise cancelled signal. Analysis reveals that if the LMS is chosen for de-noising, a larger step-size should be chosen, and the filter length should be kept small. For NLMS, step-size should be smaller and the filter length should also be small. But on choosing a larger step-size, the filter length should be increased to yield better performance In the CS LMS algorithm the time varying step size that is inversely proportional to the squared norm of the difference between two consecutive input vectors rather than the input data vector as in the NLMS.This algorithm provides significant improvements in decreasing mean squared error and consequently minimizing signal distortion.

The NLMS algorithm may be viewed as the solution to a constrained optimization problem .The problem of interest may be be stated as follows: given the tap input vector $w(n+1)$ so as to minimize the squared Euclidean norm of the Change $\delta w(n+1) = w(n+1) - w(n)$

In the tap weight vector $w(n+1)$ with respect to its old value $w(n)$,subject to the Constraint

$w^H(n+1)x(n) = d(n)$,where H denotes the Hermitian transpose .This constraint means that the a posteriori error sequence vanishes .In order to solve this optimization problem ,the method of Lagrange multipliers is used with the Lagrangian

$$L(w(n+1) = \|\delta w(n+1)\|^2 + Re[\lambda^* e^{(n+1)}(n)] \quad (3.4)$$

Where λ^* is the Lagrange multiplier ,thus obtaining the well known adaptation rule with the normalized step ize given by By $\mu = \frac{\tilde{\mu}}{\|x(n)\|^2}$ The latter constraint is restrictive in real applications ;The constrained optimization problem that provides the following cost function is Function

$$L(w(n+1) = \|\delta w(n+1)\|^2 + Re[\lambda^* \delta e^{(n+1)}(n)]$$

Where $\delta e^{[n+1]}(n) = e^{[n+1]}(n) - e^{[n+1]}(n-1)$.This equilibrium constraint ensures stability in the sequence of a posteriori errors.The Lagrange multiplier can be expressed

as $$\lambda^* = \frac{2\delta x^H(n)\delta w(n+1)}{\|\delta x(n)\|^2} = -\frac{2\left(\delta e^{[n+1]}(n) - \delta e^{[n]}(n)\right)^*}{\|\delta x(n)\|^2} \quad (3.5)$$

The weight update equation of constrained stability LMs algorithm is given by

$$w(n+1) = w(n) + \frac{\delta x(n)\delta e^*(n)}{\|\delta x(n)\|^2} \quad (3.6)$$

The weight adaptation rule can be made more robust by introducing a small positive constant ε into the denominator to prevent numerical instabilities in case of a vanishingly small squared norm $\|\delta x(n)\|^2$ and by multiplying the weight increment by a constant step size μ to control the speed of the adaptation. Note that the equilibrium condition enforces the convergence of the algorithm if $\|\delta x(n)\|^2 \neq 0$.

Several learning algorithms, where the learning relies on the concurrent change of processing variables, have been proposed in the past for de correlation, blind source separation, or de convolution applications. Stochastic information gradient (SIG) algorithms maximize (or minimize) the Shannon's entropy of the sequence of errors using an estimator based on an instantaneous value of the probability density function (pdf) and Parzen windowing. In this way, the CS-LMS algorithm can be considered as a generalization of the single sample-based SIG algorithm using variable kernel density estimators

IV.THEORETICAL REMARKS ON THE CS-LMS ADAPTATION

Once the CS-LMS method has been derived, a comparison is established with the NLMS algorithm. This section shows that, under some conditions:

1) CS-LMS and NLMS algorithms converge to the optimal Wiener solution ω_0 , and 2) for any fixed step size μ , the proposed CS-LMS exhibits improvements in excess minimum squared error (EMSE) and missadjustment (M) when compared to the NLMS algorithm.

Convergence Analysis of CS-LMS

Convergence Equivalence: Let $x(n)$ be the tap inputs to a transversal filter and $\omega(n)$ the corresponding tap weights. The estimation error $e(n)$ is obtained by comparing the estimate $y(n)$ provided by the filter with the desired response $d(n)$, that is $e(n) = d(n) - y(n)$. On the other hand, if the desired signal $d(n)$ is generated by the multiple linear regression model, i.e., $d(n) = w_o^H x(n) + e_o(n)$, where $e_o(n)$ is an uncorrelated white-noise process that is statistically independent of the input vector $x(n)$, then the CS-LMS adaptation converges to the Wiener solution $w_o(n)$ under stationary environment.

V.SIMULATION ANALYSIS

The tabulated values are calculated up to a maximum extent of 72000 iterations. The arrangement shows the deviation in the mean square error reduction factor.

Algorithm Features	**LMS**	**NLMS**	**CS-LMS**
Mean Square Error	0.0031	0.0012	0.0009
Time Elapsed [in Secs]	34.7	27.6	22.3

Table: Comparing algorithms at numerous iterations

VI.CONCLUSION

Suitable comparisons of presented algorithms imply the reduction in Mean Square Error to a maximum extent. The inference made from LMS is, it has less complexity compared to other two algorithms. However CS LMS proves more efficient by a reduced time lapse, though finding increase in complexity. By decreasing the signal noise ratio to more extent could reduce the cost of implementation which in turn could stand in reducing the complexity. In further the key points of each algorithm shall be abstracted to improve with a new phase. By studying the principles of adaptive noise cancellation (ANC), this article uses the regularization NLMS algorithm to improve its application in noise cancellation system and compares the simulation results of some performance with traditional NLMS algorithm. In speech signal processing, both methods can suppress interference signals to extract useful signals from the high background noise. However, the regularization CS LMS algorithm can exhibit a faster convergence, a greater stability and a better ability to suppress interference.

References

[1] S.Haykin, Adaptive Filter Theory, 4th edition. Englewood Cliffs, NJ: Prentice-Hall, 2002

[2] S.Hykin and A.Steinhardt, Eds., Adaptive Radar Detection and Estimation. New York: Wiley, 1992.

[3] S.L.Marple Jr., Digital Spectral Analysis with Applications. Englewood Cliffs, NJ:
Prentice-Hall, 1998.

[4] Leodardo S. Resende,Joao Marcos T. Ramano, Maurice G. Bellanger, “Split Wiener Filtering With Application in Adapive Systems”, IEEE Trans.Signal Processing, vol.52, No.3,March 2004.

[5] F. Beaufays, “Transform-domain adaptive filters: An analytical approach,” IEEE Trans. Signal Processing, vol. 43, pp. 422–431 Feb. 1995.

[6] J.M.Gorriz, Javier Ramirez, S.Cruces-Alvarez, Carlos G. Puntonet, Elmar W.Lang ,and Deniz Erdogmus, Senor Member IEEE," A Novel LMS Algorithm Applied To Adaptivel Noise Cancellation", IEEE Signal Processing Letters, Vol.16,No.1, pp.34-37, january 2009

[7] ALI H.SAYED, Adaptive Filters, John Willey & sons, 2003

[8] Paulo Sergio. Ramirez Diniz, Adaptive Filtering Algorithms and Practical Implementations

[9] SimonHaykin , B.Widrow, LeastMeanSquare Adaptive Filters, Johnwilley & sons, 2003

[10] Alexonder D.Poularikas, Zayed M.Ramadan, Adaptive Filtering Primer With MATLAB, Taylor & Francis Publishers, 2006.

[11] Allam Mousa, Marwa Quados, Sherin Bader, "Adaptive Noise Cancellation Algorithms Sensitivity to Parameters", IEEE Conference Publications on ICMCS (Multimedia And Computing Systems) pp. 1-5, 2011.

[12] Md Zia Ur Rahman," Non Stationary Noise Cancellation in Speech Signals Using Efficient Variable Step Size Higher order Filter", IJRRC, vol.2,No.2, pp. 414- 422, April 2011

[13] Georgi Iliev and Nikola Kasabov, "*Adaptive Filtering with Averaging in Noise Cancellation for Voice and Speech Recognition", Department of Information Science, University of Otago*.

[14] N.Sonbolestan, S.A.hadei,"A Fast Affine Projection Algorith Based on Matching Pursuit in Adaptive noise Cancellation Speech Enhancement",IEEE Conference Publications on ISMS 27-29 jan 2010.

[15] Thamer M. Jamel, Haider Abd Al-Latif mohamed, "Performance Improvements of Adaptive noise Canceller using New Adajusted Step Size LMS Algorithm", 3rd IEEE Conference on ICSPS august 27-28 2011

[16] E.CIfeachor and B.W.Jervis, Digital Signal Processing, A Practical Approach, Prentice Hall, 2002.

