Non Stationary Noise Removal from Speech Signals using Variable Step Size Strategy

K. Prameela, M. Ajay Kumar, Mohammad Zia-Ur-Rahman and Dr B V Rama Mohana Rao
Dept. of E.C.E., Narasaraopeta Engg. College, Narasaraopeta-522 601, India
E-mail: mdzr_5@ieee.org

Abstract—The aim of this paper is to implement various adaptive noise cancellers (ANC) for speech enhancement based on gradient descent approach, namely the least-mean square (LMS) algorithm and then enhanced to variable step size strategy. In practical application of the LMS algorithm, a key parameter is the step size. As is well known, if the step size is large, the convergence rate of the LMS algorithm will be rapid, but the steady-state mean square error (MSE) will increase. On the other hand, if the step size is small, the steady state MSE will be small, but the convergence rate will be slow. Thus, the step size provides a trade-off between the convergence rate and the steady-state MSE of the LMS algorithm.

An intuitive way to improve the performance of the LMS algorithm is to make the step size variable rather than fixed, that is, choose large step size values during the initial convergence of the LMS algorithm, and use small step size values when the system is close to its steady state, which results in Variable Step Size LMS (VSSLMS) algorithms. By utilizing such an approach, both a fast convergence rate and a small steady-state MSE can be obtained. By using this approach various forms of VSSLMS algorithms are implemented. These are robust variable step-size LMS (RVSSLMS) algorithm providing fast convergence at early stages of adaptation and modified robust variable step-size LMS (MRVSSLMS) algorithm. The performance of these algorithms is compared with conventional LMS and Kowngs VSSLMS algorithm. Finally we applied these algorithms on speech enhancement application. Simulation results confirm that the implemented RVSSLMS and MRVSSLMS are superior than conventional algorithms in terms of convergence rate and signal to noise ratio improvement (SNRI).

Keywords—Adaptive filtering, LMS algorithm, Noise Cancellation, Speech Processing, Variable Step Size.

I. INTRODUCTION

In real time environment speech signals are corrupted by several forms of noise such as such as competing speakers, background noise, car noise, and also they are subject to distortion caused by communication channels; examples are room reverberation, low-quality microphones, etc. In all such situations extraction of high resolution signals is a key task. In this aspect filtering come in to the picture. Basically filtering techniques are broadly classified as non-adaptive and adaptive filtering techniques. In practical cases the statistical nature of all speech signals is non-stationary; as a result non-adaptive filtering may not be suitable. Speech enhancement improves the signal quality by suppression of noise and reduction of distortion. Speech enhancement has many applications; for example, mobile communications, robust speech recognition, low-quality audio devices, and hearing aids.

Many approaches have been reported in the literature to address speech enhancement. In recent years, adaptive filtering has become one of the effective and popular approaches for the speech enhancement. Adaptive filters permit to detect time varying potentials and to track the dynamic variations of the signals. Besides, they modify their behavior according to the input signal. Therefore, they can detect shape variations in the ensemble and thus they can obtain a better signal estimation. The first adaptive noise cancelling system at Stanford University was designed and built in 1965 by two students. Their work was undertaken as part of a term paper project for a course in adaptive systems given by the Electrical Engineering Department. Since 1965, adaptive noise cancelling has been successfully applied to a number of applications. Several methods have been reported so far in the literature to enhance the performance of speech processing systems; some of the most important ones are: Wiener filtering, LMS filtering [1], spectral subtraction [2]-[3], thresholding [4]-[5]. On the other side, LMS-based adaptive filters have been widely used for speech enhancement [6]-[8].

In a recent study, however, a steady state convergence analysis for the LMS algorithm with deterministic reference inputs showed that the steady-state weight vector is biased, and thus, the adaptive estimate does not approach the Wiener solution. To handle this drawback another strategy was considered for estimating the coefficients of the linear expansion, namely, the block LMS (BLMS) algorithm [9], in which the coefficient vector is updated only once every occurrence based on a block gradient estimation. A major advantage of the block, or the transform domain LMS algorithm is that the input signals are approximately uncorrelated. Recently Jamal Ghasemi et.al [10] proposed a new approach for speech enhancement based on eigenvalue spectral subtraction, in [11] authors describes usefulness of speech coding in voice banking, a new method for voicing detection and pitch estimation. This method is based on the spectral analysis of the speech multi-scale product [12].

In practice, LMS is replaced with its Normalized version, NLMS. In practical applications of LMS filtering, a key parameter is the step size. If the step size is large, the convergence rate of the LMS algorithm will be rapid, but the steady-state mean square error (MSE) will increase. On the other hand, if the step size is small, the steady state MSE will be small, but the convergence rate will be slow. Thus, the step size provides a tradeoff between the convergence rate and the steady-state MSE of the LMS algorithm. The performance of the LMS algorithm may be improved by making the step size variable rather than fixed. The resultant approach with variable step size is known as variable step size LMS (VSSLMS) algorithm [13]. By utilizing such an approach, both a fast convergence rate and a small steady-state MSE can be obtained. Many VSSLMS algorithms are proposed during.
recent years [14]-[17]. In this paper, we considered the problem of noise cancellation in speech signals by effectively modifying and extending the framework of [1], using VSSLMS algorithms mentioned in [14]-[17]. For that, we carried out simulations on various real time speech signals contaminated with real noise. The simulation results show that the performances of the VSSLMS based algorithms are comparable with LMS counterpart to eliminate the noise from speech signals. Recently in [18] Karthik et.al demonstrated comparable with LMS counterpart to eliminate the noise from speech signals. Recently in [18] Karthik et.al demonstrated comparable with LMS counterpart to eliminate the noise from speech signals.

II. ADAPTIVE ALGORITHMS

A. Basic Adaptive Filter Structure

Figure 1 shows an adaptive filter with a primary input that is noisy speech signal $s_1$ with additive noise $n_1$. While the reference input is noise $n_2$, which is correlated in some way with $n_1$. If the filter output is $y$ and the filter error $e(n) = (s_1 + n_1) - y$, then

$$E[e^2] = E[(n_1 - y)^2] + E[s_1^2]$$

Minimizing the MSE results in a filter error output that is the best least-squares estimate of the signal $s_1$. The adaptive filter extracts the signal, or eliminates the noise, by iteratively minimizing the MSE between the primary and the reference inputs. The error estimation $e(n)$ is

$$e(n) = d(n) - w(n) \Phi(n)$$

Where $\Phi(n)$ is input data sequence. Coefficient updating equation is

$$w(n+1) = w(n) + \mu \Phi(n) e(n)$$

Where $\mu$ is an appropriate step size to be chosen as $0 < \mu < \frac{1}{\text{tr} R}$ for the convergence of the algorithm.

C. Kwong’s VSSLMS algorithm

The LMS type adaptive algorithm is a gradient search algorithm which computes a set of weights $w_k$ that seeks to minimize $E(d_k - X_k^T W_k)$The algorithm is of the form

$$W_{k+1} = W_k + \mu_k X_k e_k$$

Where $e_k = d_k + X_k^T W_k^* - w_k$ and $\mu_k$ is the step size. In the standard LMS algorithm $\mu_k$ is a constant. In this $\mu_k$ is time varying with its value determined by the number of sign changes of an error surface gradient estimate. Here the new variable step size or VSS algorithm, for adjusting the step size $\mu_k$ yields :

$$\mu_{k+1} = \alpha \mu_k + y \xi_k$$

and

$$\mu_{k+1} = \begin{cases} \mu_{\text{max}} & \text{if } \mu_{k+1} > \mu_{\text{max}} \\ \mu_{\text{min}} & \text{if } \mu_{k+1} < \mu_{\text{min}} \\ \mu_{k+1} & \text{otherwise} \end{cases}$$

where $0 < \mu_{\text{min}} < \mu_{\text{max}}$. The initial step size $\mu_0$ is usually taken to be $\mu_{\text{max}}$, although the algorithm is not sensitive to the choice. The step size $\mu_k$, is always positive and is controlled by the size of the prediction error and the parameters $\alpha$ and $\gamma$. Intuitively speaking, a large prediction error increases the step size to provide faster tracking. If the prediction error decreases, the step size will be decreased to reduce the misadjustment. The constant $\mu_{\text{max}}$ is chosen to ensure that the mean-square error (MSE) of the algorithm remains bounded. A sufficient condition for $\mu_{\text{max}}$

$$\mu_{\text{max}} \leq \frac{2}{\text{tr} R}$$

$\mu_{\text{min}}$ is chosen to provide a minimum level of tracking ability. Usually, $\mu_{\text{min}}$ will be near the value of $\mu$ that would be chosen for the fixed step size (FSS) algorithm $\alpha$ must be chosen in the range $(0, 1)$ to provide exponential forgetting.

D. Robust Variable Step-Size LMS (RVSSLMS) algorithm
A number of time-varying step-size algorithms have been proposed to enhance the performance of the conventional LMS algorithm. Simulation results comparing the proposed algorithm to current variable step-size algorithms clearly indicate its superior performance for cases of stationary environments. For non-stationary environments, our algorithm performs as well as other variable step-size algorithms in providing performance equivalent to that of the regular LMS algorithm [17].

The adaptation step size is adjusted using the energy of the instantaneous error. The weight update recursion is given by

\[ w^{(n+1)} = w(n) + \mu(n)e(n)X(n) \]

And updated step-size equation is

\[ \mu(n+1) = \alpha \mu(n) + \gamma \epsilon^2(n) \]

where \(0 < \alpha < 1, \gamma > 0\), and \(\mu(n+1)\) is set to or when it falls below or above these lower and upper bounds, respectively.

The constant \(\mu_{\text{max}}\) is normally selected near the point of instability of the conventional LMS to provide the maximum possible convergence speed. The value of \(\mu_{\text{max}}\) is chosen as a compromise between the desired level of steady state misadjustment and the required tracking capabilities of the algorithm. The parameter \(\gamma\) controls the convergence time as well as the level of misadjustment of the algorithm. At early stages of adaptation, the error is large, causing the step size to increase, thus providing faster convergence speed. When the error decreases, the step size decreases, thus yielding smaller misadjustment near the optimum. However, using the instantaneous error energy as a measure to sense the state of the adaptation process does not perform as well as expected in the presence of measurement noise. The output error of the identification system is

\[ e(n) = d(n) - X^T(n)W(n) \]

where \(d(n)\) is the desired signal is given by

\[ d(n) = X^T(n)W^*(n) + \xi(n) \]

\(\xi(n)\) is a zero-mean independent disturbance, and \(W^*(n)\) is the time-varying optimal weight vector. Substituting (8) and (9) in the step-size recursion, we get

\[ \mu(n+1) = \alpha \mu(n) + \gamma \epsilon^2(n) - 2\gamma \epsilon(n)X^T(n)X(n) \]

Where \(V(n) = W(n) - W^*(n)\) is the weight error vector. The input signal autocorrelation matrix, which is defined as \(R = E\{X(n)X^T(n)\}\), can be expressed as \(R = Q\Lambda Q^T\), where \(\Lambda\) is the matrix of eigen values, and \(Q\) is the model matrix of \(R\). Using \(V(n) = Q^T\epsilon(n)\) and \(X(n) = Q^T X(n)\), then the statistical behavior of \(\mu(n+1)\) is determined.

\[ E\{\mu(n+1)\} = \alpha E\{\mu(n)\} + \gamma E\{\epsilon^2(n)\} + E\{V^T(n)\Lambda V(n)\} \]

where we have made use of the common independence assumption of \(V(n)\) and \(X(n)\). Clearly, the term \(E\{V^T(n)\Lambda V(n)\}\) influences the proximity of the adaptive system to the optimal solution, and \(\mu(n+1)\) is adjusted accordingly. However, due to the presence of \(E\{\xi^2(n)\}\), the step-size update is not an accurate reflection of the state of adaptation before or after convergence. This reduces the efficiency of the algorithm significantly. More specifically, close to the optimum, \(\mu(n)\) will still be large due to the presence of the noise term \(E\{\xi^2(n)\}\).

The step size can be rewritten as

\[ \mu(n+1) = \alpha \mu(n) + \gamma E\{V^T(n)X(n)X^T(n)\} \]

It is also clear from above discussion that the update of \(\mu(n)\) is dependent on how far we are from the optimum and is not affected by independent disturbance noise. Finally, the considered algorithm involves two additional update equations compared with the standard LMS algorithm. Therefore, the added complexity is six multiplications per iteration. These multiplications can be reduced to shifts if the parameters \(\alpha, \beta, \gamma\), are chosen as powers of 2.

E. Modified Robust Variable Step-Size LMS (MRVSSLMS) algorithm

From the frame work of step size parameter of LMS algorithm, Kwongs and RVSSLMS algorithms the step size of MRVSS is given:

\[ \mu(n+1) = \begin{cases} \mu_{\text{max}} & \text{if } \mu(n+1) > \mu_{\text{max}} \\ \mu_{\text{min}} & \text{if } \mu(n+1) < \mu_{\text{min}} \\ \alpha \mu(n) + \gamma \rho^2(n) & \text{otherwise} \end{cases} \]

\[ \rho(n) = (1 - \beta(n)) \rho(n) + \beta(n) \epsilon(n) \epsilon(n-1) \]

\[ \beta(n+1) = \begin{cases} \beta_{\text{max}} & \text{if } \beta(n+1) > \beta_{\text{max}} \\ \beta_{\text{min}} & \text{if } \beta(n+1) < \beta_{\text{min}} \\ \eta \beta(n) + \lambda \epsilon^2(n) \end{cases} \]

where the parameters \(0 < \alpha, \eta, \gamma, \lambda > 0\). The \(\rho(n)\) is the time average of the error signal correlation at iteration time \(n\) and \(n+1\), and the \(\beta(n)\) is the time average of the square error signal, which is used to control the sensitivity of \(\rho(n)\) to the instantaneous error correlation. \(\min 0 < \mu_{\text{min}} < \mu_{\text{max}} ; 0 < \beta_{\text{min}} < \beta_{\text{max}} < 1\). The upper bound of step size \(\mu_{\text{max}}\) satisfied the mean square stability condition. The lower bound of the step size \(\mu_{\text{min}}\) is used to guarantee the excess MSE under the tolerant level. The parameter \(\beta\) should be less than 1 and larger than zero.

That is to say, when the algorithm is convergent, the instantaneous error power is very small and the error signal correlation is not sensitive to instantaneous error, and the accuracy of error signal correlation is enhanced. If the system is suddenly changed, the instantaneous error signal power is increased, which result to the enlargement of the correlation function of the error signal and the instantaneous error signal correlation, therefore the algorithm has a good tracking ability.
In one word, the MRVSS have good tracking ability and good anti-noise ability, which are the advantages of algorithm proposed in reference [15][17]. Using these strategies different adaptive noise cancellers are implemented to remove diverse form of noises from speech signals.

III. SIMULATION RESULTS

To show that RVSSLMS and MRVSSLMS algorithms are appropriate for speech enhancement we have used real speech signals with noise. In the figure number of samples is taken on x-axis and amplitude is taken on y-axis. In order to test the convergence performance we have simulated a sudden noise spike at 4000th sample. From the figure it is clear that the performance of the implemented RVSSLMS and MRVSSLMS algorithms is better than the conventional LMS and Kwongs VSSLMS algorithm. To prove the concept of filtering we have considered five speech samples contaminated with various real noises. These noises are high voltage murmuring, crane noise. For comparison purpose we also considered random noise removal. Generally the noise added to the speech signal when it is transmitted through free space is random in nature. The noisy speech signal is given as input to the adaptive filter structure shown in Figure 1, signal somewhat correlated with noise is given as reference signal. As the number of iterations increases error decreases and clean signal can be extracted from the output of the filter. These simulation results are shown in Figures 3, 4. To evaluate the performance of the algorithms SNRI is measured and tabulated in Tables I, II, III.

IV CONCLUSION

In this paper the problem of noise removal from speech signals using Variable Step Size based adaptive filtering is presented. For this, the same formats for representing the data as well as the filter coefficients as used for the LMS algorithm were chosen. As a result, the steps related to the filtering remains unchanged. The proposed treatment, however exploits the modifications in the weight update formula for all categories to its advantage and thus pushes up the speed over the respective LMS-based realizations. Our simulations, however, confirm that the ability of MRVSSLMS and RVSSLMS algorithms is better than conventional LMS and Kwongs VSSLMS algorithms in terms of SNR improvement and convergence rate. Hence these algorithm is acceptable for all practical purposes.
REFERENCES


Available online @ www.ijcscn.com

95