A Modified Adaptive Line Enhancer for Noisy Speech Signals

Maha Sharkas, M. Essam Khedr, and Amr Nasser

Abstract—Adaptive filter is a method for enhancing noisy speech signals. In this work, a refinement for the adaptive line enhancer proposed by Widrow and Sambur is introduced. The new algorithm is based on both the exact value of gradient and variable step-size. These values are computed and used for coefficient updating. The algorithm is applied on different cases of noisy speech signals as well as GSM signals.

The new algorithm gave good results even when noise is greater than signal which is considered a novel achievement in the field. It avoids the drawbacks of LMS algorithm introduced by widrow and others. The adaptive line enhancer (ALE) design method updates the FIR filter coefficients in terms of the exact value of gradient and a variable step size.

Keywords—Least mean squares, adaptive filters, adaptive noise canceller.

I. INTRODUCTION

ADAPTIVE filter is a kind of technology which is widely used in the field of the modern signal processing. It can detect and extract useful signal from the strong interference of noise environment. The adaptive noise canceller (ANC) whose objective is noise interference is the typical application of adaptive filters. Also, it restrains and attenuates the noise to improve SNR quality of the transmitting and receiving signal. Adaptive filter includes a digital filter and an adaptive algorithm. It uses least-mean-squares algorithm as its standard, then adjusts the filter coefficients to achieve the best filter characteristics using an adaptive algorithm. The least-mean-squares (LMS) algorithm is the most commonly method as it has a simple structure and a basic architecture but with a disadvantage in stability and convergence speed. Therefore, the adaptive noise canceller is designed adopting an improved LMS algorithm [1].

The ANC Based on Least Mean Square Algorithm has been extensively used in many researches [1-6]. The purpose of ANC is to remove the noise from a signal adaptively to improve the SNR. Because of its simplicity and ease of implementation, the LMS algorithm is the most popular adaptive algorithm. However, the LMS algorithm suffers from slow and data dependent convergence behavior.

One of the main disadvantages of the LMS algorithm is that it cannot be used when noise is greater than signal [7]. The algorithm can be modified to overcome this problem by using (Overlap and add - Overlap and save) methods as suggested in [7].

The objective of this paper is to implement a novel algorithm to extract signal from noise especially when noise is greater than signal. Fig.1 gives a simplified block diagram of the system.

Fig.1 A simplified block diagram of the system

The paper is organized as follows: section 2 reviews the Overlap and add - Overlap and save methods. The algorithm design steps are introduced in section 3. Results are presented in section 4 and the paper is concluded in section 5.

II. THE OVERLAP AND ADD - OVERLAP AND SAVE METHODS

In practice, the length of an input signal may be very large. To process such long sequences, the input signal can be sectioned into blocks of a length small enough to be processed by a given computer, and then add the output resulting from all the input blocks. This procedure can be used because the operation is linear. Sectioning the input allows the output to have a smaller processing delay.

There are two well-known techniques for performing a linear convolution using the FFT algorithm and these are referred to as the overlap-save and overlap-add sectioning methods. By overlapping elements of the data sequences and retaining only a subset of the final DFT product, a linear convolution between a finite-length sequence and an infinite-length sequence is readily obtained [8 and 9].

III. ALGORITHM DESIGN STEPS

The design steps include the following computations [9]:

1. The delay time p.
2. The maximum number of filter coefficients.
3. The exact value of gradient.
4. The variable step-size.

Maha Sharkas Author is with Arab academy for Science & Technology, Alexandria, Egypt; (e-mail:msharkas@aast.edu).
M. Essam khedr Author, is with Arab academy for Science & Technology, Alexandria, Egypt; (e-mail: khedr@vt.edu).
Amr Nasser. Author is with Arab academy for Science & Technology, Alexandria, Egypt; (e-mail: amrnaser2@yahoo.com).

http://dx.doi.org/10.15242/IIE.E0315003
A. Computing the delay time (P):

The adaptive line canceller ALE is a special case of the adaptive noise canceller ANC, where there is only one signal $x(n)$ available, the primary input, which is contaminated by noise. In such a case the reference signal $u(n)$ is the delayed replica of $x(n)$ such that $u(n) = x(n - p)$. Suppose the signal $x(n)$ consists of two components: a narrow band component $s(n)$, which has long range correlation, and a broadband component $v(n)$, which will tend to have short range correlation. Also, $L_{NB}$ and $L_{BB}$ are effectively the self-correlation lengths of the Narrowband (NB) and broadband (BB) respectively, as shown in figure 3. Beyond these lags, the respective correlations die out quickly. The delay $P$ is selected so that:

$$L_{NB} \leq P \leq L_{BB}$$

If $P$ is longer than the effective correlation length of BB, then the delayed replica $v(n - p)$, will entirely be uncorrelated with $v(n)$ which is a part of the primary signal. The adaptive filter will not be able to respond to this component. On the other hand, when $P$ is shorter than the correlation length of the NB component, the delayed replica of $s(n - p)$ that appears in the reference input will still correlate with $s(n)$ which is a part of the primary signal. In this case the filter will respond to cancel $v(n)$.

If $p$ is selected to be longer than both correlation lengths, then the reference input will become uncorrelated with the primary input and the adaptive filter will be turned off. In the opposite case, when the delay is selected to be less than both the correlation lengths, then both components of the reference signal will be correlated with the primary signal, and therefore the adaptive filter will respond to cancel the primary signal $x(n)$ completely.

The structure of the proposed ALE is illustrated in figure (3).

B. Computation of Maximum of filter (taps) coefficients

In the following, the maximum number of filter taps, $M_{\text{max}}$ is estimated, based on the cross correlation of the primary and the reference input samples. During $M$ intervals of length $T$, $M$ is the number of samples per observation window. The cross correlation for an assumed delay $\tau$ samples will

$$R(\tau) = \sum_{k=1}^{\infty} x(k)u(k-\tau)$$

$$= \sum_{k=1}^{\infty} x(k)x(k - p - \tau)$$

The estimated NB component, or filter output $s(n)$ is obtained from the $M$ forgoing samples $u(n), u(n-1), ..., u(n-M+1)$. The maximum correlation length $\tau_{\text{max}}$ is computed as the value of the variable $\tau$ at which the correlation between $x(n)$ and $u(n-\tau)$ is negligible with respect to the correlation between $x(n)$ and $u(n)$. Mathematically, $x(n)$ and $u(n-\tau)$ are completely uncorrelated if:

$$\frac{R(\tau)}{R(0)} < \varepsilon$$

Where $\varepsilon$ a small positive values, typical value for is $\varepsilon$ is $0.01$. The maximum number of filter taps is given by:

$$L = M_{\text{max}} = \tau_{\text{max}}$$

C. Computation of the exact value of the gradient

In this subsection, an exact value for the gradient is obtained. Under stationary conditions for the signal $u(n)$, the best estimate of $v(n)$ would be:

$$\hat{V}(n) = -\frac{1}{n} \sum_{i=1}^{n} e(i)u(i)$$

$$= \frac{-1}{n} \left[ (n-1)\hat{V}(n-1) + e(n)u(n) \right]$$

But, in the adaptive situation $u(n)$ is non-stationary, therefore it can be seen that the previous equation would not be a good estimate of $\hat{V}(n)$ because of its infinite memory. This estimate would become insensitive to changes for large values of $n$. To provide this effect, the previous equation becomes:

$$\hat{V}(n) = -\frac{1}{n} \left[ \sum_{i=1}^{n} \lambda^{n-i} e(i)U(i) \right]$$

$$= -\frac{1}{n} \left[ \sum_{i=1}^{n} \lambda^{n-i} e(i)u(i) + e(n)u(n) \right]$$

$$= \frac{-1}{n} \left[ (n-1)\hat{V}(n-1) + e(n)u(n) \right]$$

The choice of the forgetting factor $\lambda$ is often very important:

a) Theoretically one must have $\lambda = 1$ (for stationary input) to get convergence.

b) For the non-stationary case, the algorithm becomes more sensitive and the parameter estimates change very
quickly. For that reason, it is often an advantage to allow the forgetting factor to vary with time. Therefore, substitute \( \lambda \) by \( \hat{\lambda}(n) \). A typical choice is to let \( \lambda \) tend exponentially to 1. This can be written as

\[
\hat{\lambda}(n) = \lambda_0 \lambda(n-1) + (1 - \lambda_0)
\]

Typical value for \( \lambda_0 \) and \( \hat{\lambda}(0) \) are

\[
\lambda_0 = 0.99 \quad \text{and} \quad \hat{\lambda}(0) = 0.95
\]

D. Computation of a variable step size

An adaptive step-size must satisfy: 1. The speed of convergence should be fast. 2. When operating in stationary environments \([10]\), the steady state miss-adjustment values should be very small. 3. When operating in the non-stationary environments, the algorithm should be able to sense the rate at which the optimal coefficients are changing. The above goals are achieved when the step-size \( \mu(n) \) is adjusted as:

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

With \( 0 < \alpha < 1 \), \( \gamma > 0 \)

\[
\begin{aligned}
\mu_{\max} &> \mu_{\max} \\
\mu(n + 1) &< \mu_{\min} \\
\mu'(n + 1) &\text{Otherwise}
\end{aligned}
\]

The initial step-size \( \mu(0) \) is usually taken to be \( \mu_{\max} \). The choice of \( \mu_{\max} \) is based on stability considerations and it must be chosen in the range \( \frac{1}{\nu[R]} > \mu_{\max} > 0 \), where

\[
\nu[R] = L P_x
\]

\( P_x \) is the power of the input signal, which is approximated by

\[
P_x \approx \frac{1}{L} R(0)
\]

On the other hand, \( \mu_{\min} \) is chosen to provide the desired misadjustment, after the filter has converged. The misadjustment, ratio of the average excess mean-square error to the minimum mean-square error, is taken to be 5 percent, and \( \mu_{\min} \) is

\[
\mu_{\min} = 0.1 \mu_{\max}
\]

The value of \( \alpha \) appears to be a good choice for all experiments \( (\alpha = 0.97) \), while the value of \( \gamma \) is chosen arbitrary. Typical values of \( \gamma \) in stationary and non-stationary are \( 4.8 \times 10^{-4} \) and \( 7.65 \times 10^{-14} \) respectively.

IV. EXPERIMENTAL RESULTS

The speech is normally low-pass filtered at frequency of about 1 KHz, which is well above the maximum anticipated frequency range for pitch (500 Hz for female speech). Filtering helps to reduce the effect of higher formants and any high-frequency noise \([11]\). Two low pass filters are used:

The first one keeps all frequencies below 1000 Hz and eliminates all frequencies higher than 2000 Hz.

The second filter preserves the frequencies less than 500 Hz, and attenuates the frequencies higher than 1000 Hz. The corresponding cut-off digital frequencies in radians for the first filter are as follows.

Pass band frequency:

\[
\omega_p = \frac{f_p}{F_s/2} = \frac{1000}{8000/2} = 0.25
\]

Stop band frequency:

\[
\omega_s = \frac{f_s}{F_s/2} = \frac{2000}{8000/2} = 0.5
\]

Similarly, for the second filter we have:

\[
\omega_p = 0.125, \omega_s = 0.25
\]

The sampling frequency \( \frac{f_s}{2} \) is chosen to be fairly close to the upper limit of the ear’s sensitivity

Frame blocking is applied as follows:

(i) If we assume the signal is piecewise stationary, we can analyze the signal using a sliding window. The two used key parameters are:

(a) Frame duration \( T_f \).

(b) Window Duration \( T_w \).

Typical values are: \( T_f = 5 \text{ ms} \) , \( T_w = 10 \text{ ms} \)

(ii) The window duration controls the amount of averaging (or smoothing) used in power calculation.

- Each frame will contain 40 samples.
- Each window will contain 80 samples.

(iii) The amount of the overlap value is given by:

\[
(\text{overlap}) 100\% = \left( \frac{T_w - T_f}{T_w} \right) 100\% = 50\%
\]

Frame blocking is depicted in figure 4.
The proposed system referred to in figure 1 is simulated, using Matlab. The experiment is carried out in three main steps [10]:
Step 1: Preliminary adjustment.
Step 2: Determination of the optimum filter length.

a) It was necessary to record speech alone and noise alone. The noise speech is obtained by adding the noise to the clean signal according to the required SNR[12]

\[ SNR_{in}[dB] = 10 \log_{10}\left(\frac{signal\text{power}}{noise\text{power}}\right) \]

b) The speech signals are uttered by male speaker for the following three cases:
(i) Case (1): Voiced vowel [a] “Aah”
(ii) Case (2): For word [Ahmed]
(iii) Case (3): For complete sentence “How old are you”
(iv) A GSM signal
c) The noise signal is Gaussian type.
d) The sampling frequency used is 10024 Hz.
e) The duration of vowel to be synthesized is 105 msec sectioned into 7msec duration at which \( T_f = 5\text{msec} \) and \( T_w = 7\text{msec} \).

The results presented shows for all the above cases a clean signal, signal with added noise then the extracted signal using the modified technique. The SNR between the obtained output and the input is then given. The modified technique succeeded in retaining the original signal with good SNR as depicted in figures 5, 6 and 7.

Fig. 5: a: the clean signal for the vowel “a”, b: the signal with added noise, c: the extracted signal and d: the SNR between the output and the input.
The technique is also applied on GSM signals namely the GMSK signal as shown in figure 8. The extracted output signal is shown in figures 8 c & d with and without using the modified technique. The SNR without and with using the modified algorithm are compared in figure 8.e indicating the modified performance of the suggested technique.
An adaptive noise enhancement based on a modified LMS algorithm is implemented and tested. This modified adaptive line enhancer (ALE) can be used to detect and estimate weak signals in noise. Two modifications are introduced to the ALE using the FIR structure which are:

a) The exact value of the gradient is computed. The exact value emphasized the most recent values of the gradients and ignore the older one according to the value of the forgetting factor $\lambda$ ($\lambda$ lies between 0 and 1)

b) The variable step-size that depends on the square of predication error allows the adaptive filter to track changes in the system as well as produce small steady-state error.

This work demonstrates the ability of ALE to reduce additive periodic or stationary random noise in both periodic and random signals. The suggested modification allowed enhancing week signals and provided good SNR. It is also applied on GSM signals and succeeded in providing improved results.

REFERENCES


http://dx.doi.org/10.1002/9780470740156


http://dx.doi.org/10.15242/IIE.E0315003