SELECTIVE COEFFICIENT UPDATE OF GRADIENT-BASED ADAPTIVE ALGORITHMS

Tyseer Aboulnasr
Dept. of Elect. and Comp. Engg.
University of Ottawa
Ottawa, Canada K1N 6N5
aboulnas@eng.uottawa.ca

Khaled Mayyas
Dept. of Electrical Engineering
Jordan University of Science and Technology
Irbid, Jordan, 221 10

ABSTRACT

One common approach to reducing the computational overhead of the normalized LMS (NLMS) algorithm is to update a subset of the adaptive filter coefficients. It is known that the mean square error (MSE) is not equally sensitive to the variations of the coefficients. Accordingly, the choice of the coefficients to be updated becomes crucial. On this basis, we propose an algorithm that belongs to the same family but selects at each iteration a specific subset of the coefficients that will result in the largest reduction in the performance error. The proposed algorithm reduces the complexity of the NLMS algorithm, as do the current algorithms from the same family, while maintaining a performance close to the full update NLMS algorithm specifically for correlated inputs.

1. INTRODUCTION

Acoustic echo cancellation is a common application of adaptive filtering. The adaptive acoustic echo canceller requires several hundred taps in order to achieve satisfactory echo suppression. Even with the use of a NLMS-based echo canceller, the huge processing power required to implement such an echo canceller is beyond the capabilities of current DSP chips. Therefore, designing algorithms with reduced complexity can be extremely useful in such applications.

Several algorithms were proposed to reduce the computational cost of the NLMS algorithm. Such algorithms include the periodic NLMS algorithm [1], and the partial update algorithms [2, 3] where only a predetermined subset of the coefficients is updated every iteration. Inevitably, the penalty incurred by using these algorithms is a lower performance, in terms of convergence speed, than the regular NLMS algorithm where all coefficients are updated. The decrease in convergence speed is proportional to the reduction in complexity and can be sometimes a major drawback in their implementation in the case of long impulse responses.

The algorithm proposed here attempts to reduce the complexity of the NLMS algorithm while preserving a performance close to the regular NLMS algorithm. The algorithm is a member of the family of adaptive algorithms that updates a portion of their coefficients at each iteration, but it selects those coefficients adaptively to achieve the most reduction in the performance error. The algorithms in [1, 2, 3] choose those coefficients each iteration (or block of iterations) in a pre-specified fixed way. The proposed algorithm adds a maximum of $2\log_2(N)+2$; (N being the adaptive filter length) comparison operations over the computational overhead of the algorithms in [2, 3].

2. CURRENT ADAPTIVE ALGORITHMS WITH PARTIAL COEFFICIENTS UPDATES

Here, we consider two algorithms where only a subset of the adaptive filter coefficients is updated at each iteration. The first algorithm, namely the sequential NLMS (SNLMS) algorithm [4], updates M coefficients out of N coefficients at each iteration. The update procedure is given by

$$w_i(n+1) = \begin{cases} w_i(n) + \frac{\mu}{\mathbf{X^T}(n)\mathbf{X}(n)} e(n)x(n-i+1) \ , \\ \quad \text{if}(n-i+1) \ mod \ (N/M) = 0 \\ w_i(n) \quad , \ otherwise \end{cases} \tag{1}$$

where $\mathbf{W}(n) = [w_1(n), \ldots, \ w_N(n)]^T$ is the coefficient vector, $\mathbf{X}(n) = [x(n)\ x(n-1) \ldots x(n-N+1)]^T$ is the input data vector, and μ is the adaptation step size. This approach was used in [2, p.432] in updating the taps of the adaptive echo canceller. The complexity of the algorithm, excluding the overhead of calculating $\mathbf{X^T}(n)\mathbf{X}(n)$, is $N+M+1$ multiplications, $N+M$ additions, and a single division[1]. On the other hand,

[1]Note that $\mathbf{X^T}(n)\mathbf{X}(n)$ can be approximated using the simple recursion estimate of the power $p(n) = \lambda p(n-1) + (1-\lambda)x^2(n)$, $0 < \lambda < 1$. This method requires 3 multiplications, one addition, and one memory location. Another method to calculate $p(n) = \mathbf{X^T}(n)\mathbf{X}(n)$ is using $p(n) = p(n-1) + x^2(n) - x^2(n-N)$. This method requires one multiplication, 2 additions, and 2 memory locations.

the NLMS algorithm requires $2N+1$ multiplications, $2N$ additions, and one division.

The second algorithm divides the adaptive filter length into N/M successive blocks, each block has M coefficients [3]. At each iteration, one block is updated and blocks are updated in a sequential manner. The algorithm is referred to here as the sequential block NLMS (SBNLMS) algorithm. The complexity of this algorithm in terms of multiplications and additions is equivalent to the that of the SNLMS algorithm.

3. PROPOSED GRADIENT-BASED PARTIAL UPDATE ALGORITHM

For the above two algorithms, the reduction in the number of arithmetic operations per iteration has been achieved at the expense of the convergence time of the algorithms which is lowered to roughly N/M that of the standard NLMS algorithm. For LMS-type algorithms, it is generally noted that when updating all coefficients of the adaptive filter, the contribution due to the error in some coefficients is small while other coefficients have larger error contributions. In other words, the error function is not equally sensitive to variations in all coefficients. This is more evident for correlated inputs. Accordingly, even if "less important" coefficients are not unupdated at a given iteration, the algorithm performance will be marginally affected. The sensitivity of the performance error to the individual coefficient at each iteration depends on two factors: the shape of the mean-square error (MSE) surface and the location of that coefficient at that instance relative to the bottom of the MSE surface. The degree of the sensitivity of the error function to coefficients variations increases with the degree of "colour" in the input signal. This can be easily seen by noting that coloured inputs result in performance surfaces that have elliptical contours. In this case, the mean square error is extremely sensitive to the choice of the adaptive filter coefficients to be updated. This sensitivity is reflected in the steepness of the gradient vector components. Those coefficients with larger gradient components on the error surface result in considerably larger contributions to the reduction of the overall mean square error.

Consequently, a simple and direct criterion for the selection of coefficients to be updated is based on the magnitude of the corresponding gradient estimate in the direction of every individual coefficient at a given iteration. A larger gradient magnitude implies steeper descent in the direction of that coefficient, and consequently a larger reduction in error when that coefficient is updated. Moreover, when the coefficient is near its optimum value, the gradient along the direction of this coefficient will be very small. Therefore, not much reduction in error is attained when that coefficient is updated. The gradient estimate in the direction of the *ith* coefficient is $-2e(n)x(n-i+1)$, where $1 \leq i \leq N$. Clearly, since all gradient components involve the quantity $-2e(n)$, then the proposed criterion is to select coefficients associated with the M largest value of $|x(n-i+1)|$ for updating. Note that in low SNR environments, the error $e(n)$ can be very noisy leading to a noisy gradient estimate that can provide false information. However, it is clear that the determination of the subset to be updated does not depend on $e(n)$ ensuring that the algorithm is not sensitive to the effect of noise disturbances. The proposed algorithm is thus stated as follows. At each iteration, M coefficients out of N are updated. Those M coefficients are the ones associated with the M largest $|x(n-i+1)|$, $i=1,..,N$, at that iteration. The algorithm update equation can be written as

$$w_i(n+1) = \begin{cases} w_i(n) + \frac{\mu}{\mathbf{X^T}(n)\mathbf{X}(n)} e(n)x(n-i+1)\,, \\ \quad \text{if } i \text{ corresponds to one of the first} \\ \quad M \text{ maxima of } |x(n-i+1)|,\ i=1,..,N \\ w_i(n) \quad \textit{otherwise} \end{cases} \tag{2}$$

In terms of multiplications/additions, the proposed algorithm has the same complexity overhead of the two algorithms described in the previous section.

It should be noted that, at each iteration, a running sorting procedure in a descending order of $|x(n-i+1)|$, $i=1,..,N$ is required. The M coefficients that belong to the first M elements of the sorted vector are updated. In [5], a fast algorithm for running sorting of a sliding window of arbitrary N elements is proposed. The algorithm, named SORTLINE, uses the fact that at each iteration one new sample enters the window and one old sample is discarded . The algorithm requires, at most, $2\log_2(N)+2$ comparison operations per sample time. Consequently, the proposed partial update algorithm needs an extra $2\log_2(N)+2$ comparisons in addition to multiplications and additions required by the SNLMS or SBNLMS algorithm. For large N, the savings of $N-M$ multiplications and $N-M$ additions will exceed the additional complexity of $2\log_2(N)$ comparison operations.

4. SIMULATIONS

We examine here the acoustic echo cancellation application. The echo path is that of an anechoic room of 200 taps, measured at 8KHz sampling rate. The proposed algorithm is compared with the NLMS algorithm with the update of all N coefficients at every iteration, the SNLMS algorithm, and the SBNLMS algorithm. For the proposed algorithm, SNLMS, and SBNLMS, $M=25$ coefficients are updated every iteration. Both the SNLMS and SBNLMS are described in section 2. The algorithms are tested under different input signal environments: stationary white, stationary correlated, and nonstationary using a real speech signal.

In the first two examples, perfect modeling of the echo path is assumed, i.e., N=200. A white noise of 0.0001 variance is added to the desired signal. Results are obtained by averaging over 100 independent runs. In the last example, where a real speech is used, the adaptive filter is chosen to be of length $N = 150$ (undermodeling of the echo path).

In the first example, a zero-mean white Gaussian signal of unity variance is used. The full update NLMS, SBNLMS, SNLMS, and proposed algorithms are used with the same step size $\mu = 0.5$. It is clear from Fig.1 that the proposed algorithm outperforms the SNLMS and SBNLMS algorithms though the three algorithms update the same number of taps (M=25) at each iteration. Also, note that the loss in convergence speed of the proposed algorithm compared to the NLMS algorithm is marginal.

In the second example, the input signal is a highly correlated one generated by passing a zero-mean white Gaussian signal with unity variance through the filter $H(z) = \frac{1}{1-1.58z^{-1}+0.81z^{-2}}$. The step size value of the NLMS algorithm is $\mu = 0.5$. The proposed algorithm, SBNLMS, and SNLMS algorithms are used with the step size $\mu = 0.4$, which is chosen to achieve the same steady state MSE of the NLMS algorithm. The superiority of the performance of the proposed algorithm compared to the SBNLMS and SNLMS is obvious from Fig.2. Also, the performance of the proposed algorithm is still comparable to the NLMS algorithm. The slow convergence speed of the SBNLMS and SNLMS algorithm relative to the NLMS algorithm in Fig.1 and Fig.2 is not surprising, since the convergence speed of the SBNLMS and SNLMS algorithms is expected to be 8 times slower than the NLMS algorithm. On the other hand, the proposed algorithm strategy of updating the portion of the coefficients that leads to the greatest reduction in the error at each iteration has minimized the loss in performance compared to the case when all coefficients are updated. It can be seen that the improvement of the proposed algorithm in this case over SBLMS and SNLMS algorithms is significantly more than for the white input. This is due to the shape of the error surface for coloured inputs where the slope may vary dramatically from one coefficient to another.

In the last example, a speech signal of a male voice is used as an input signal. The adaptive filter is used with 150 taps to achieve a satisfactory echo return loss enhancement (ERLE) level. The ERLE is defined as

$$ERLE = 10 \log_{10}\left(\frac{E\{d^2(n)\}}{E\{e^2(n)\}} \right) \qquad (3)$$

All algorithms employ the same step size value $\mu = 1$. Fig.3 illustrates the ability of the proposed algorithm to operate as well as the NLMS algorithm, and better than the SBNLMS and SNLMS algorithms in this practical example. It is to be noted that all coefficients are incremented at each iteration when the NLMS algorithm is used. For the SBNLMS and SNLMS algorithms, each coefficient receives one update every fixed number of samples (150/25 = 6, for this example). However, the proposed algorithm picks the M (M=25) most "important" coefficients at each iteration. Thus, coefficient update is only performed when it is deemed to provide "reasonable" reduction in the error.

5. CONCLUSION

In this paper, we presented an adaptive algorithm that belongs to the family of algorithms that update only a subset of the adaptive filter coefficients at each iteration. The proposed algorithm selects the "important" coefficients to achieve the most reduction in error. The idea is based on the observation that some coefficients have steeper directions on the error surface than others and, therefore, make larger contribution in reducing the error. This case is more evident when the input signal is correlated. Simulation examples indicate that the proposed algorithm outperforms the SBNLMS and SNLMS algorithms and retains close performance to the full update NLMS algorithm. In terms of multiplications and additions, the proposed algorithm entails the same complexity as the SBNLMS or SNLMS algorithms. However, it involves, in the worst case, an additional $2 \log_2(N) + 2$ comparison operations.

6. REFERENCES

[1] W. A. Gardner, "Learning characteristics of stochastic-gradient-descent algorithms:A general study, analysis, and critique," *Signal Processing*, Vol. 6, No. 2, pp. 113-133, April, 1989.

[2] Digital Signal Processing Applications with the TMS3C20 Family, Texas Instruments, 1986.

[3] S. M. Kuo, and J. Chen, "Multiple Microphone Acoustic Echo Cancellation System with the Partial Adaptive Process," *Digital Signal Processing*, Vol. 3, pp. 54-63, 1993.

[4] S. C. Douglas, "Simplified Stochastic Gradient Adaptive Filters Using Partial Updating", *Proc. Sixth IEEE Digital Signal Processing Workshop*, Yosemite, CA, pp. 265-268, Oct., 1994.

[5] I. Pitas, "Fast Algorithms for Running Ordering and Max/Min Calculation," *IEEE Trans Circuits and Systems*, Vol. 36, No. 6, pp. 795-804, June, 1989.

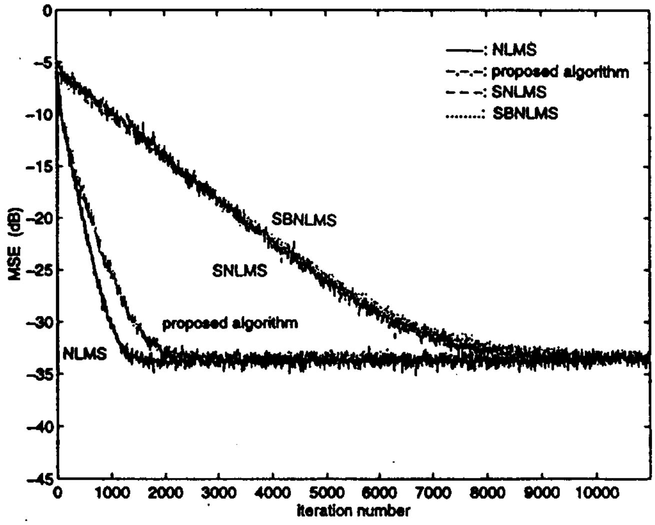

Fig. 1 Comparison of MSE between the full update NLMS, SBNLMS, SNLMS and proposed algorithm for white input case.

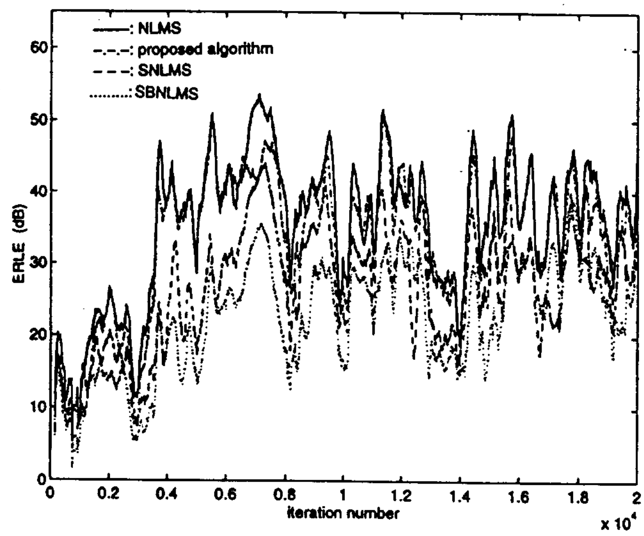

Fig.3 Comparison of MSE between the full update NLMS, SBNLMS, SNLMS and proposed algorithm for speech input case.

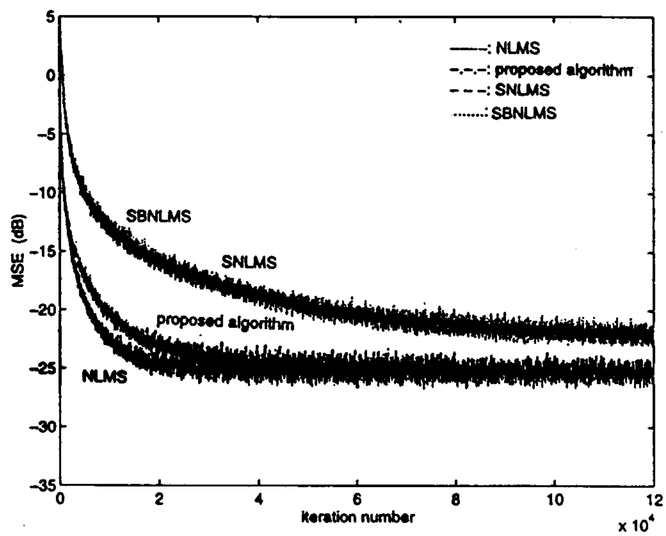

Fig.2 Comparison of MSE between the full update NLMS, SBNLMS, SNLMS and proposed algorithm for correlated input case.