

# FILTERED GRADIENT ALGORITHMS APPLIED TO A SUBBAND ADAPTIVE FILTER STRUCTURE

*J. A. Apolinário Jr.<sup>1</sup>, R. G. Alves<sup>2,3</sup>, P. S. R. Diniz<sup>4</sup>, and M. N. Swamy<sup>2</sup>*

<sup>1</sup>Instituto Militar de Engenharia  
Depto. Engenharia Elétrica  
Praça Gal. Tibúrcio, 80  
Rio de Janeiro, RJ  
22.290-270, Brazil  
apolin@ieee.org

<sup>2</sup>Concordia University  
Dept. Elec. and Comp. Eng.  
1455 de Maisonneuve Blvd. West  
Montreal, Quebec  
H3G 1M8 Canada  
swamy@ece.concordia.ca

<sup>3</sup>Clarity  
3290 W. Big Beaver Rd.  
Suite 220  
Troy, Michigan  
48084 US  
alves@clarityco.com

<sup>4</sup>Universidade Fed. do Rio de Janeiro  
COPPE/Prog. Eng. Elétrica  
P.O. Box 68504  
Rio de Janeiro, RJ  
21.945-970 Brazil  
diniz@lps.ufrj.br

## ABSTRACT

Adaptive filtering techniques in subbands have been recently developed for a number of applications including acoustic echo cancellation and wideband active noise control. In such applications, hundreds of taps are required resulting in high computational complexity and low convergence rate when using LMS based algorithms. For fullband systems, new algorithms which try to overcome these drawbacks have been investigated. A class of these algorithms employing variants of the filtered gradient adaptive (FGA) algorithm has been successfully developed. In this paper, we apply these techniques to a recently proposed subband adaptive filter structure in order to improve the convergence rate and the computational load. Computer simulations show the benefits obtained with these proposed algorithms.

## 1. INTRODUCTION

Adaptive filtering techniques, particularly using FIR filters in view of their stability and unimodal performance properties, are used in many applications. However, in some applications such as acoustic echo cancellation and wideband active noise control, the order of the adaptive filters is very high, resulting in a large number of operations for their implementation and hence presenting a slow convergence rate when using LMS based algorithms.

As an attempt to solve the above problem, subband processing techniques have been proposed for adaptive filters [1]-[3]. The advantages expected from subband processing are: (a) the computational complexity is approximately reduced in proportion to the number of subbands, because both the number of taps and weight update rate can be decimated in each subband; and (b) the convergence rate is

improved because the spectral dynamic range is greatly reduced in each subband.

For the fullband case, an algorithm known as the filtered gradient adaptive (FGA) algorithm and its variations have been developed in order to improve the convergence rate of applications where the input signal is highly correlated. The main idea of this paper is to employ these algorithms in the adaptive subband structure proposed in [3].

The paper is organized as follows. In Section 2, the subband adaptive structure proposed in [3] is described. Section 3 introduces the FGA algorithm for this afore-mentioned subband adaptive structure. Section 4 proposes the orthogonal projection FGA algorithms for the same structure. Simulation results are presented in Section 5, and Section 6 contains some concluding remarks.

## 2. THE SUBBAND ADAPTIVE FILTER STRUCTURE

The adaptive subband structure presented in [3] was derived from the filter bank structure with sparse adaptive subfilters of Fig. 1. In a system identification application, such a structure models exactly any FIR system if the sparse adaptive filters  $G_k(z)$  satisfy the following equation:

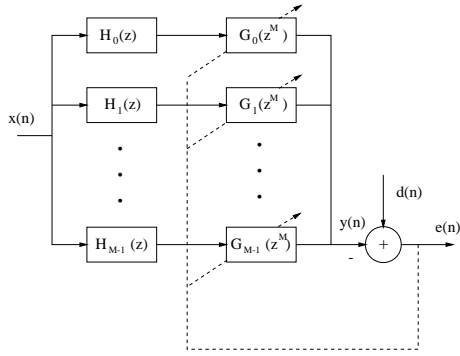
$$\begin{bmatrix} G_0(z) & G_1(z) & \cdots & G_{M-1}(z) \end{bmatrix} = \begin{bmatrix} P_0(z) & P_1(z) & \cdots & P_{M-1}(z) \end{bmatrix} \mathbf{F}_p(z) \quad (1)$$

where  $P_i(z)$  are the polyphase components of the unknown system transfer function  $P(z)$ , and  $\mathbf{F}_p(z)$  is the type-2 polyphase matrix of the synthesis bank which results in perfect reconstruction when associated with the analysis filters  $H_k(z)$  of Fig. 2.

By including maximally decimated perfect reconstruction analysis and synthesis banks following each sparse subfilter in Fig. 2, moving the sparse subfilters  $G_k(z^M)$  to the right of the decimators, and assuming that non-adjacent filters of the analysis bank have frequency responses which

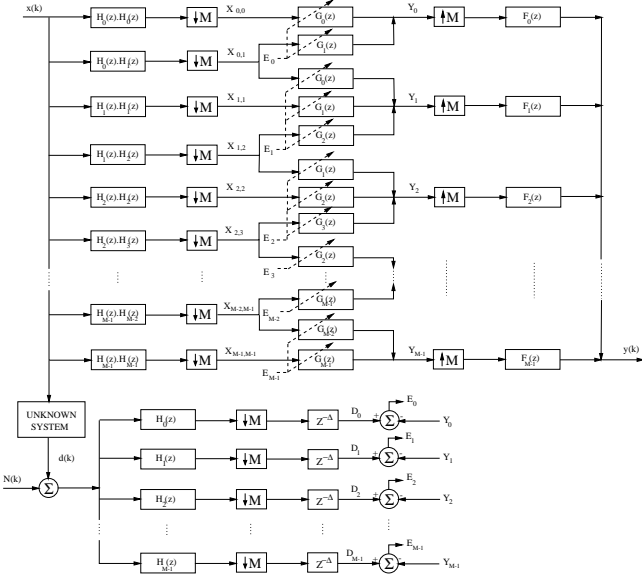
---

This work was partially supported by The Natural Sciences and Engineering Research Council of Canada, MICRONET - National Network of Centers of Excellence, and by FCAR, Quebec.



**Fig. 1.** Adaptive structure with an analysis filter bank and sparse subfilters.

do not overlap, the structure of Fig. 2 has been obtained [3]. Observe that, in the resulting structure, the subfilters  $G_k(z)$  operate at a rate which is  $1/M$ -th of the input rate, and that from (1), their lengths should be  $K = ((N + N_f)/M) - 1$ , where  $N$  is the length of the unknown system to be identified and  $N_f$  is the length of each synthesis filter  $F_k(z)$ .



**Fig. 2.** Adaptive subband structure with critical sampling applied to system identification problem.

For each subband of the general structure (corresponding to an arbitrary number ( $M$ ) of subbands) shown in Fig. 2, we define the *a priori* error signal given by

$$E_i(k) = D_i(k) - Y_i(k) \quad (2)$$

where  $Y_i(k)$  is given by

$$Y_i(k) = \begin{cases} \mathbf{X}_{0,0}^T(k) \mathbf{G}_0(k) + \mathbf{X}_{0,1}^T(k) \mathbf{G}_1(k), & i = 0 \\ \mathbf{X}_{i-1,i}^T(k) \mathbf{G}_{i-1}(k) + \mathbf{X}_{i,i}^T(k) \mathbf{G}_i(k) \\ \quad + \mathbf{X}_{i,i+1}^T(k) \mathbf{G}_{i+1}(k), & 0 < i < M-1 \\ \mathbf{X}_{M-2,M-1}^T(k) \mathbf{G}_{M-2}(k) \\ \quad + \mathbf{X}_{M-1,M-1}^T(k) \mathbf{G}_{M-1}(k), & i = M-1 \end{cases} \quad (3)$$

with  $\mathbf{X}_{i,j}(k) = [X_{i,j}(k) \ X_{i,j}(k-1) \ \dots \ X_{i,j}(k-N)]^T$ .

In the next two sections, two new algorithms will be derived for this structure using the above notation.

### 3. THE FGA ALGORITHM IN SUBBANDS

We start the derivation of the Filtered Gradient Adaptive (FGA) algorithm in subbands, for the particular structure presented in the former section, by defining the following objective function which is similar to what was done in [4] for the full band case.

$$J(k) = \sum_{i=0}^{M-1} \frac{1}{2} \sum_{j=0}^k \lambda^{k-j} E_i^2(j) \quad (4)$$

It is easy to obtain a recursive form for the above expression which is given by

$$J(k) = \lambda J(k-1) + \frac{1}{2} \sum_{i=0}^{M-1} E_i^2(k) \quad (5)$$

The updating equation for each  $i$ th subband coefficient vector can be carried out as follows.

$$\mathbf{G}_i(k+1) = \mathbf{G}_i(k) + \mu \mathbf{p}_i(k) \quad (6)$$

where the updating process is performed along the direction vector  $\mathbf{p}_i(k)$  regulated by step-size  $\mu$ .

The direction vector is chosen to be the negative gradient of the objective function being minimized or

$$\mathbf{p}_i(k) = -\nabla_{\mathbf{G}_i(k)} J(k) = \lambda \mathbf{p}_i(k-1) + \mathbf{g}_i(k) \quad (7)$$

where

$$\mathbf{g}_i(k) = \mathbf{X}_{i-1,i}(k) E_{i-1}(k) + \mathbf{X}_{i,i}(k) E_i(k) \\ + \mathbf{X}_{i,i+1}(k) E_{i+1}(k), \quad 0 < i < M-1 \quad (8)$$

Note that the above equation also stands for other values of  $i$  for which the first term is not present if  $i = 0$  and the last term is not present if  $i = M-1$ .

The FGA algorithm in subbands represented by (2), (6), and (7), in the case of full band or  $M = 1$ , is equivalent to the Momentum LMS algorithm [5] and its name (Filtered Gradient Adaptive) comes from the fact that the direction

vector may be considered the output of a filter with a single-pole  $\lambda$  with  $\mathbf{g}_i(k)$  being the input. Positive values of  $\lambda$  corresponds to a low-pass filter and encompasses a smoother convergence to this algorithm although no significant gain in convergence speed over the conventional LMS algorithm can be expected. [5] also links the MLMS algorithm with the Conjugate Gradient algorithm and gives the following stability conditions:  $|\lambda| < 1$  and  $0 < \mu < \frac{1+\lambda}{\lambda_N}$ , where  $\lambda_N$  is the highest eigenvalue of the input-signal autocorrelation matrix. Table 1 presents the subband FGA algorithm.

**Table 1.** The FGA algorithm in subbands.

For each subband $i$ :	
Compute $Y_i(k)$ as in	(3)
$E_i(k) = D_i(k) - Y_i(k)$	(2)
Compute $\mathbf{g}_i(k)$ as in	(8)
$\mathbf{p}_i(k) = \lambda \mathbf{p}_i(k-1) + \mathbf{g}_i(k)$	(7)
$\mathbf{G}_i(k+1) = \mathbf{G}_i(k) + \mu \mathbf{p}_i(k)$	(6)

#### 4. ORTHOGONAL PROJECTION FGA ALGORITHMS IN SUBBANDS

If we consider a slightly different definition for the objective function, we can accommodate the possibility of a time varying forgetting factor  $\lambda_k$  as in

$$\mathbf{p}_i(k) = \lambda_i(k) \mathbf{p}_i(k-1) + \mathbf{g}_i(k) \quad (9)$$

In order to improve the convergence speed of this class of algorithms, a procedure was introduced in [6] to determine  $\lambda_i(k)$  such that  $\mathbf{p}_i(k)$  and  $\mathbf{p}_i(k-1)$  are orthogonal or, equivalently,  $\mathbf{p}_i^T(k) \mathbf{p}_i(k-1) = 0$ . From (9) we make  $\lambda_i(k) \mathbf{p}_i^T(k-1) \mathbf{p}_i(k-1) + \mathbf{p}_i^T(k-1) \mathbf{g}_i(k) = 0$  such that the time-varying forgetting factor is given by

$$\lambda_i(k) = -\frac{\mathbf{p}_i^T(k-1) \mathbf{g}_i(k)}{\mathbf{p}_i^T(k-1) \mathbf{p}_i(k-1)} \quad (10)$$

From  $\mathbf{p}_i(k) = \lambda_i(k) \mathbf{p}_i(k-1) + \mathbf{g}_i(k) = \mathbf{g}_i(k) - \frac{\mathbf{p}_i^T(k-1) \mathbf{g}_i(k)}{\mathbf{p}_i^T(k-1) \mathbf{p}_i(k-1)} \mathbf{p}_i(k-1)$ , we see that  $\mathbf{p}_i(k)$  corresponds to  $\mathbf{g}_i(k)$  minus the projection of  $\mathbf{g}_i(k)$  onto  $\mathbf{p}_i(k-1)$ . This algorithm was named Orthogonal Gradient Adaptive (OGA) in [6] and its subband version (OGASB) is shown in Table 2.

A normalized version of the OGA algorithm is also available in [6]. Following a similar approach, we propose a normalized version for the subband case given by

$$\mathbf{g}_i(k) = \frac{\mathbf{X}_{i-1,i}(k) \mathbf{E}_{i-1}(k)}{\|\mathbf{X}_{i-1,i}(k)\|} + \frac{\mathbf{X}_{i,i}(k) \mathbf{E}_i(k)}{\|\mathbf{X}_{i,i}(k)\|} + \frac{\mathbf{X}_{i,i+1}(k) \mathbf{E}_{i+1}(k)}{\|\mathbf{X}_{i,i+1}(k)\|}, 0 < i < M-1 \quad (11)$$

The following table also presents this normalized version named NOGASB (from Normalized Orthogonal Gradient Adaptive) algorithm.

**Table 2.** The OGA and NOGA algorithm in subbands.

For each subband $i$ :	
Compute $Y_i(k)$ as in	(3)
$E_i(k) = D_i(k) - Y_i(k)$	(2)
Compute $\mathbf{g}_i(k)$ as in	(OGASB) (8)
	(NOGASB) (11)
$\lambda_i(k) = -\frac{\mathbf{p}_i^T(k-1) \mathbf{g}_i(k)}{\mathbf{p}_i^T(k-1) \mathbf{p}_i(k-1)}$	(10)
$\mathbf{p}_i(k) = \lambda_i(k) \mathbf{p}_i(k-1) + \mathbf{g}_i(k)$	(9)
$\mathbf{G}_i(k+1) = \mathbf{G}_i(k) + \mu \mathbf{p}_i(k)$	(6)

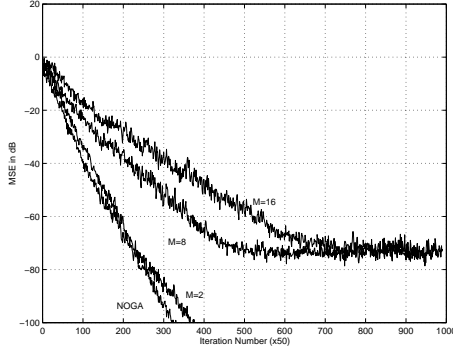
Although the (fullband) NOGA algorithm presents the *a posteriori* error equal to zero if  $\mathbf{g}(k) \perp \mathbf{p}(k-1)$ , the same does not apply to the NOGASB algorithm due to the typical coupling among channels for the subband structure used.

#### 5. SIMULATION RESULTS

The identification of a length  $N = 256$  FIR system is considered. In all the simulations, we have employed the value of the step-size which resulted in the best convergence rate for each algorithm. Experiments were performed with the subband structure of Fig. 2 with  $M = 2, 8$  and  $16$  subbands, and with perfect reconstruction analysis and synthesis cosine modulated filter banks with prototype filters of lengths  $N_h = 32, 128$  and  $256$ , respectively. We present here only the results for the NOGA algorithm, since the results for the FGA and OGA algorithms are similar to those obtained by the NOGA algorithm, but the convergence rate for these are worse than that for the NOGA algorithm.

Figure 3 presents the MSE evolution of the NOGA algorithm considering the subband and fullband systems, where the subband implementation uses the algorithm described in Table 2. The input signal is a colored noise sequence generated by passing a white noise sequence of unit variance through a first-order IIR filter with the pole located at  $z = 0.9$ . The step-size used are  $\mu = 0.1, 0.2$  and  $0.3$ , for  $M = 2, 8$  and  $16$ , respectively. The fullband uses  $\mu = 1$ . The subband structure converges to an MSE of the order of the stopband attenuation of the analysis filter (which is around  $-75$  db for  $M = 8$  and  $M = 16$ ), due the assumption of non-overlapping between non-adjacent analysis filters [3].

The results obtained in Fig. 3 show that increasing the number of subbands ( $M$ ), reduces the convergence rate. This result is opposite to that obtained using LMS based algorithms. Here, we can say that the smaller spectral dynamic range in each subband, due to the subband decomposition,



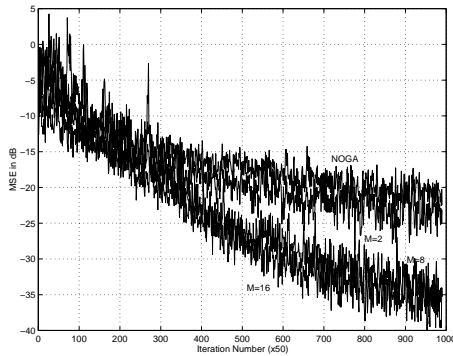
**Fig. 3.** MSE curves to NOGA considering fullband and subband system.

is not enough to dominate the fact that we are updating the coefficients of the adaptive subfilters with a rate  $M$  times smaller than the fullband system. Also, in a subband system, the total number of coefficients need to update is higher than that need in the case of the fullband system (see Eq. 1). However, we expect that by using input signals with higher spectral dynamic range, the subband algorithm will have a better performance in view of the convergence rate of the MSE.

Figure 4 presents the MSE evolution of the NOGA where the input signal  $x(k)$  is a colored noise sequence given by:

$$u(n) = w(n) - 0.99u(n-1) - 0.99u(n-2) - 0.99u(n-3) - 0.99u(n-4) \quad (12)$$

where  $w(n)$  is a white noise sequence of unit variance. The step-size used were  $\mu = 0.05, 0.15$  and  $0.3$ , for  $M=2, 8$  and  $16$ , respectively. The fullband uses  $\mu = 1$ .



**Fig. 4.** MSE curves to NOGA considering fullband and subband systems.

In Fig. 4 we confirm our expectancy, i.e., for a highly correlated input signal, the subband algorithm has a better convergence rate than the fullband system, and higher the number of subbands ( $M$ ), better will be the convergence rate of the adaptive algorithm.

Considering now the computational complexity (number of multiplications), for the subband FGA algorithm we obtain:

$$\frac{5N+1}{M} + \frac{2(3N_h-5)}{M} + 4\log_2 M \quad (13)$$

Comparing this with the computational complexity of the fullband FGA algorithm ( $3N+1$ ), we observe that for ( $N \gg N_h$ ), the subband algorithm is  $3M/5$  less complex than the fullband algorithm. This result can be extended to OGA and NOGA algorithms, where we can obtain similar conclusions.

## 6. CONCLUSIONS

We have derived new critical sampling subbands adaptation algorithms based on the FGA algorithm and its variants. Experimental results have shown that the convergence rate is improved when the input signals have a high spectral dynamic range. The convergence rate of the MSE may further be increased by using a higher number of subbands; also the computational complexity of the proposed algorithms is smaller than those of the fullband ones. Thus, we can state that for an application where the performance of the FGA-like algorithms are necessary (for instance, when the error bursting is a problem) and the number of coefficients is very high, these new subband implementations are good choices.

## 7. REFERENCES

- [1] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: analysis, experiments, and application to acoustic echo cancellation," *IEEE Trans. Signal Processing*, pp. 1862-1875, vol. 40, no. 8, Aug. 1992.
- [2] M. R. Petraglia and S. K. Mitra, "Performance Analysis of Adaptive Filter Structures Based on Subband Decomposition," in *Proc. IEEE Int. Symp. Circ. Syst.*, pp. 60-63, Chicago, IL, May 1993.
- [3] M. R. Petraglia, R. G. Alves and P. S. R. Diniz, "New Structures for Adaptive Filtering in Subbands with Critical Sampling," *To appear on IEEE Transactions on Signal Processing*, Dec. 2000.
- [4] J. S. Lim, "Fast adaptive filtering algorithm based on exponentially weighted least-square errors," *Electronic Letters*, vol. 35, no. 22, pp. 1913-1915, Oct. 1999.
- [5] S. Roy and J. J. Shynk, "Analysis of the momentum LMS algorithm," *IEEE Trans. on Acoust., Speech, and Signal Processing*, vol. 38, no. 12, pp. 2088-2098, Dec. 1990.
- [6] J. S. Lim, "New adaptive filtering algorithms based on an orthogonal projection of gradient vectors," *IEEE Signal Processing Letters*, vol. 7, Issue 11, pp. 314-316, Nov. 2000.