PAPER Special Section on Digital Signal Processing

LMS-Based Algorithms with Multi-Band Decomposition of the Estimation Error Applied to System Identification

Fernando Gil V. RESENDE, Jr.[†], Paulo S.R. DINIZ[††], *Nonmembers*, Keiichi TOKUDA[†††], Mineo KANEKO[††††], *and* Akinori NISHIHARA[†††††], *Members*

SUMMARY A new cost function based on multi-band decomposition of the estimation error and application of a different step-size for each band is used in connection with the least-mean-square criterion to improve the fidelity of estimates as compared to those obtained with conventional least-mean-square adaptive algorithms. The basic new idea is to trade off time and frequency resolutions of the adaptive algorithm along the frequency domain by using different step-sizes in the analysis of distinct frequencies in accordance with the frequency-localized statistical behavior of the input signal. The mathematical background for a stochastic approach to the multi-band decomposition-based scheme is presented and algorithms with fixed and variable step-sizes are derived. Computer experiments compare the performance of multi-band and conventional least-mean-square methods when applied to system identification.

key words: *adaptive signal processing, fixed and variable step-size LMS algorithms, filter banks, system identification*

1. Introduction

Least-squares (LS) and least-mean-square (LMS) optimization criteria are widely used in adaptive filtering systems. Improvement of factors such as computational complexity and fidelity, has been stimulating the design of a large quantity of algorithms that search for the most appropriate adjustable set of parameters which minimizes a cost function [1]–[3]. Consequently, the choice of the cost function crucially influences the performance of adaptive algorithms.

Recently, a new cost function based on the LS criterion with decomposition of the linear prediction error through a power complementary filter bank and analysis of each band with a different window has been shown to give rise to estimated parameters with more fidelity than conventional LS methods [4]–[6]. The basic advantage of the so-called multi-band method is that time and frequency resolutions of the adaptive algorithm can be traded off along the frequency domain giving rise to better estimates than those obtained with conventional methods, particularly when the statistical properties of the analyzed signal vary along the frequency spectrum.

In this paper, the basic mathematical relations that govern the application of the multi-band scheme with the LMS criterion are derived. An algorithm in which the estimation error is decomposed into subbands and a particular fixed step-size (FSS) is used for each band is first presented [7], [8]. In the following, an algorithm based on the multi-band structure with variable step-sizes (VSSs) is derived to allow for the trade off of time and frequency resolutions not only in the frequency domain but also in the time domain, which is suitable for nonstationary analysis. Plots of the running spectrum of the filter coefficients in a system identification problem for a nonstationary unknown system are used to compare the performances of the multi-band and conventional LMS methods.

This paper is organized as follows. Section 2 introduces a new cost function, and formulates the corresponding error-performance surface and Wiener-Hopf equations. Section 3 describes a multi-band LMS algorithm with FSSs based on the results of the preceding section. Section 4 presents a multi-band LMS algorithm with VSSs. Section 5 shows computer experiments comparing the performance of multi-band and conventional LMS algorithms. Section 6 presents our conclusions. The nomenclature used in this paper follows the standard contained in [2] with modifications that naturally arise from the multi-band characteristic of the presented structure.

2. Wiener Filtering with Multi-Band Decomposition of the Estimation Error

In this section, the Wiener filter theory is formulated for the multi-band decomposition of the estimation error structure. In Sect. 2.1, the equation related to the error-performance surface is derived and, in Sect. 2.2, the Wiener-Hopf equations are obtained when a linear transversal filter is used to estimate the desired response

Manuscript received Nov 27, 1996.
Manuscript revised March 4, 1997.

[†]The author is with the Department of Physical Electronics, Tokyo Institute of Technology, Tokyo, 152 Japan.
[††]The author is with the Prog. de Engenharia Elétrica e Depto. de Eletrônica, COPPE/EE/Federal University of Rio de Janeiro, Caixa Postal 68504, Rio de Janeiro, RJ, CEP 21945-970, Brazil.
[†††]The author is with the Department of Computer Science, Nagoya Institute of Technology, Nagoya-shi, 466 Japan.
[††††]The author is with the School of Information Science, Japan Advanced Institute of Science and Technology, Ishikawa-ken, 923–12 Japan.
[†††††]The author is with the Center for Research and Development of Educational Technology, Tokyo Institute of Technology, Tokyo, 152 Japan.

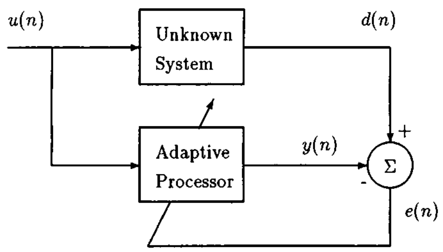


Fig. 1 Block diagram for system identification (modeling).

$d(n)$.

2.1 Error-Performance Surface

In adaptive filtering theory, several algorithms have been developed to solve the problem of updating the set of adjustable parameters in such a way that the mean-square value of the estimation error is minimized [1]–[3]. The conventional definition of the cost function as the mean-squared error is

$$J(n) = E\left[e(i,n)e^*(i,n)\right], \tag{1}$$

where $E[.]$ represents the expectation operator, the superscript symbol "$*$" means complex conjugation, and

$$e(i,n) = d(i) - y(i,n) \tag{2}$$

$$y(i,n) = \sum_{k=0}^{M-1} w_k^*(n)u(i-k), \tag{3}$$

where $e(i,n)$ is the estimation error, $d(i)$ is the desired response, $y(i,n)$ is the output of the linear transversal filter with coefficients $w_k(n)$, $k = 0,\ldots,M-1$, at time-index n, and $u(i)$ is the signal to be analyzed. A block diagram representation of the system described by the equations above is shown in Fig. 1.

In this paper, we are proposing a new cost function, which is defined as follows

$$J_{MB}(n) = \frac{1}{B}\sum_{j=1}^{B} J_{MB}^{(j)}(n) \tag{4}$$

$$= \frac{1}{B}\sum_{j=1}^{B} E\left[e^{(j)}(i,n)\left(e^{(j)}(i,n)\right)^*\right], \tag{5}$$

where the filtered estimation error for the jth-band, $e^{(j)}(i,n)$, is given as

$$e^{(j)}(i,n) = h^{(j)}(i) * e(i,n) \tag{6}$$

$$= d^{(j)}(i) - \sum_{k=0}^{M-1} w_k^*(n)u^{(j)}(i-k), \tag{7}$$

where

$$d^{(j)}(i) = h^{(j)}(i) * d(i) \tag{8}$$

$$u^{(j)}(i) = h^{(j)}(i) * u(i). \tag{9}$$

The symbol "$*$" stands for the convolution operation, and the analysis filter bank is represented by $h^{(j)}(i)$, $j = 1,\ldots,B$. If the analysis filter bank satisfies the power complementary property, then $J_{MB}(n) = cJ(n)$ (refer to the Appendix for a proof).

Substitution of (7) into (5) gives

$$\begin{aligned} J_{MB}(n) = \frac{1}{B}\sum_{j=1}^{B}\Bigg(& E\left[\left|d^{(j)}(i)\right|^2\right] \\ & - \sum_{k=0}^{M-1} w_k^*(n)E\left[u^{(j)}(i-k)\left(d^{(j)}(i)\right)^*\right] \\ & - \sum_{k=0}^{M-1} w_k(n)E\left[\left(u^{(j)}(i-k)\right)^* d^{(j)}(i)\right] \\ & + \sum_{k=0}^{M-1}\sum_{l=0}^{M-1} w_k^*(n)w_l(n) \\ & \cdot E\left[u^{(j)}(i-k)\left(u^{(j)}(i-l)\right)^*\right]\Bigg), \end{aligned} \tag{10}$$

which can be rewritten as

$$\begin{aligned} J_{MB}(n) = \frac{1}{B}\sum_{j=1}^{B}\Bigg(& \sigma_{d^{(j)}}^2 - \sum_{k=0}^{M-1} w_k^*(n)p^{(j)}(-k) \\ & - \sum_{k=0}^{M-1} w_k(n)\left(p^{(j)}(-k)\right)^* \\ & + \sum_{k=0}^{M-1}\sum_{l=0}^{M-1} w_k^*(n)w_l(n)r^{(j)}(l-k)\Bigg), \end{aligned} \tag{11}$$

where $\sigma_{d^{(j)}}^2$ is the variance of the filtered desired response $d^{(j)}(i)$, assumed to be of zero mean, and

$$p^{(j)}(-k) = E\left[u^{(j)}(i-k)\left(d^{(j)}(i)\right)^*\right] \tag{12}$$

and

$$r^{(j)}(l-k) = E\left[u^{(j)}(i-k)\left(u^{(j)}(i-l)\right)^*\right], \tag{13}$$

for $0 \leq l,k \leq M-1$, are respectively the jth-band cross-correlation between the filtered input signal $u^{(j)}(i)$ and the filtered desired response $d^{(j)}(i)$ for lag k; and the jth-band autocorrelation of the signal $u^{(j)}(i)$ with lag difference $(l-k)$. The cost function described in (11) is a second-order function of the tap weights $w_k(n)$ and this dependence can be visualized as a bowl-shaped $(M+1)$-dimensional surface with a unique minimum. This surface is referred to as the *error-performance surface* of the adaptive filter.

In the sequence, the cost function described in (11) is minimized giving rise to the Wiener-Hopf equations.

2.2 Wiener-Hopf Equations

Once the cost function $J_{MB}(n)$ described in (11) is a quadratic function of the adaptive filter coefficients $w_k(n)$, its minimization can be obtained by equaling its first derivative to zero. From (11), the derivative of $J_{MB}(n)$ with respect to the filter coefficients $w_k(n)$ is

$$\nabla_{w_k(n)} J_{MB}(n) = \frac{1}{B}\sum_{j=1}^{B}\left(-2p^{(j)}(-k) + 2\sum_{l=0}^{M-1} w_l(n) r^{(j)}(k-l)\right), \tag{14}$$

which is minimized when

$$\frac{1}{B}\sum_{j=1}^{B}\sum_{l=0}^{M-1} w_{o,l}(n) r^{(j)}(k-l) = \frac{1}{B}\sum_{j=1}^{B} p^{(j)}(-k), \tag{15}$$

for $k = 0, 1, \ldots, M-1$, so-called Wiener-Hopf equations, which can be rewritten in terms of the correlation matrix $\mathbf{R}_M$, the optimal filter coefficient vector $\mathbf{w}_{o,M}$, and the cross-correlation vector $\mathbf{p}_M$ as

$$\mathbf{R}_M \mathbf{w}_{o,M} = \mathbf{p}_M, \tag{16}$$

where

$$\mathbf{R}_M = \frac{1}{B}\sum_{j=1}^{B} \mathbf{R}_M^{(j)} \tag{17}$$

$$= \frac{1}{B}\sum_{j=1}^{B} E\left[\mathbf{u}_M^{(j)}(n)\left(\mathbf{u}_M^{(j)}(n)\right)^H\right] \tag{18}$$

$$\mathbf{u}_M^{(j)}(n) = \begin{bmatrix} u^{(j)}(n) \\ u^{(j)}(n-1) \\ \vdots \\ u^{(j)}(n-M+1) \end{bmatrix} \tag{19}$$

$$\mathbf{p}_M = \frac{1}{B}\sum_{j=1}^{B} \mathbf{p}_M^{(j)} \tag{20}$$

$$= \frac{1}{B}\sum_{j=1}^{B} E\left[\mathbf{u}_M^{(j)}(n)\left(d^{(j)}(n)\right)^*\right] \tag{21}$$

$$\mathbf{w}_{o,M} = \left[w_{o,0}, w_{o,1}, \ldots, w_{o,M-1}\right]^T, \tag{22}$$

where the superscripts "T" and "H" represent transposition and conjugate transposition, respectively.

In the next two sections, LMS algorithms which solve (16) for $\mathbf{w}_{o,M}$ are derived.

3. A Multi-Band LMS Algorithm with Fixed Step-Sizes

In Sect. 3.1, an LMS algorithm which uses a particular FSS for each band in the evaluation of an estimate of $\mathbf{w}_{o,M}$ is derived. In Sect. 3.2, convergence and stability conditions of the presented algorithm are analyzed.

3.1 The Algorithm

In particular, the solution of (16) involves two problems: the computation of the correlation matrix as shown in (18) is often unpractical since it relies on unknown quantities; even when an estimate of the correlation matrix is obtained, solution of (16) by inverting the estimated $\mathbf{R}_M$ at every time-index n is not computationally efficient.

As in conventional methods, a solution for the first problem mentioned in the above paragraph can be obtained by taking *instantaneous estimates.* With this approach, the correlation matrix and the cross-correlation vector can be respectively estimated as follows

$$\hat{\mathbf{R}}_M(n) = \frac{1}{B}\sum_{j=1}^{B} \hat{\mathbf{R}}_M^{(j)}(n) \tag{23}$$

$$= \frac{1}{B}\sum_{j=1}^{B} \mathbf{u}_M^{(j)}(n)\left(\mathbf{u}_M^{(j)}(n)\right)^H \tag{24}$$

$$\hat{\mathbf{p}}_M(n) = \frac{1}{B}\sum_{j=1}^{B} \hat{\mathbf{p}}_M^{(j)}(n) \tag{25}$$

$$= \frac{1}{B}\sum_{j=1}^{B} \mathbf{u}_M^{(j)}(n)\left(d^{(j)}(n)\right)^*. \tag{26}$$

One of the most popular computationally-efficient algorithms used in the solution of the Wiener-Hopf equations is based on the method of steepest descent [1]–[3]. For the multi-band case, a similar recursive equation for the coefficient vector can be written as

$$\hat{\mathbf{w}}_M(n+1) = \hat{\mathbf{w}}_M(n) + \frac{1}{2B}\sum_{j=1}^{B} \mu^{(j)}\left(-\hat{\nabla}_{w_k(n)} J_{MB}^{(j)}(n)\right) \tag{27}$$

$$= \hat{\mathbf{w}}_M(n) + \frac{1}{B}\sum_{j=1}^{B} \mu^{(j)}\left(\hat{\mathbf{p}}_M^{(j)}(n) - \hat{\mathbf{R}}_M^{(j)}(n)\hat{\mathbf{w}}_M(n)\right) \tag{28}$$

$$= \hat{\mathbf{w}}_M(n) + \frac{1}{B}\sum_{j=1}^{B} \mu^{(j)} \mathbf{u}_M^{(j)}(n)\left(e^{(j)}(n)\right)^*, \tag{29}$$

where $\mu^{(j)}$ is the step-size related to the jth-band and $\hat{\mathbf{w}}_M(n)$ represents the estimate of $\mathbf{w}_{o,M}$ at time-index n. Equations (7), (8), (9), and (29) form a multi-band LMS algorithm with FSSs.

3.2 Convergence and Stability Conditions

In this section, the value of the weight vector after convergence is first derived, and then a sufficient condition on the range of the step-sizes for each band which guarantees stability of the algorithm presented in Sect. 3.1 is obtained.

Equation (29) can be rewritten as

$$\hat{\mathbf{w}}(n+1) = \hat{\mathbf{w}}(n) + \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}\mathbf{u}^{(j)}(n)\Big(d^{(j)}(n) - \hat{\mathbf{w}}^{H}(n)\mathbf{u}^{(j)}(n)\Big) \tag{30}$$

Applying the expectation operator to both sides of the equation above, and assuming that $\mathbf{u}(n)$ and $\hat{\mathbf{w}}(n)$ are statistically independent, we have

$$E[\hat{\mathbf{w}}(n+1)] = E[\hat{\mathbf{w}}(n)] + \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}E[d^{(j)}(n)\mathbf{u}^{(j)}(n)] - \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}E[\mathbf{u}^{(j)}(n)\left(\mathbf{u}^{(j)}(n)\right)^{H}\hat{\mathbf{w}}(n)] \tag{31}$$

$$= E[\hat{\mathbf{w}}(n)] + \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}\mathbf{p}^{(j)} - \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}\mathbf{R}^{(j)}E[\hat{\mathbf{w}}(n)] \tag{32}$$

$$= \left[\mathbf{I} - \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}\mathbf{R}^{(j)}\right]E[\hat{\mathbf{w}}(n)] + \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}\mathbf{p}^{(j)} \tag{33}$$

$$= \left[\mathbf{I} - \tilde{\mathbf{R}}\right]E[\hat{\mathbf{w}}(n)] + \tilde{\mathbf{p}}, \tag{34}$$

where

$$\tilde{\mathbf{R}} = \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}\mathbf{R}^{(j)}, \tag{35}$$

$$\tilde{\mathbf{p}} = \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}\mathbf{p}^{(j)}. \tag{36}$$

In the same manner of the discussion on the convergence of the conventional LMS algorithm [1], [3], $E[\hat{\mathbf{w}}(n)]$ converges to $\tilde{\mathbf{w}}$ defined by

$$\tilde{\mathbf{R}}\tilde{\mathbf{w}} = \tilde{\mathbf{p}}, \tag{37}$$

with the condition

$$0 < \lambda_{max} < 2 \tag{38}$$

where λ_{max} is the largest eigenvalue of $\tilde{\mathbf{R}}$. It should be noted that $\tilde{\mathbf{w}}$ is obtained by minimizing

$$\tilde{J}_{MB}(n) = \frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}E\left[e^{(j)}(n)\left(e^{(j)}(n)\right)^{H}\right]. \tag{39}$$

Also, calling $P_e(\omega)$ and $P_u(\omega)$ the power spectrum of $e(n)$ and $u(n)$, respectively, we have

$$\tilde{J}_{MB}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} P_e(\omega)\frac{1}{B}\sum_{j=1}^{B}\mu^{(j)}\left|H^{(j)}(e^{j\omega})\right|^2 d\omega \tag{40}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_e(\omega)F(\omega)d\omega \tag{41}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_u(\omega)\left|W_o(z) - W(z)\right|^2 F(\omega)d\omega \tag{42}$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_u(\omega)\left|\sum_{k=0}^{M} w_{o,k}(n)z^{-k} - \sum_{k=0}^{N} w_k(n)z^{-k}\right|^2 F(\omega)d\omega, \tag{43}$$

where $w_{o,k}(n)$ and $w_k(n)$ represent the unknown system coefficients and the adaptive filter coefficients at time-index n, respectively. As a result, $\tilde{J}_{MB}(n)$ defines a frequency weighted mean square error, where the weighting function is $F(\omega)$. For applications such as system identification, it can be easily seen from (43) that if the unknown system and the adaptive filter have the same order, in the average, the adaptive filter coefficients $\hat{\mathbf{w}}(n)$ converge to the desired response $\tilde{\mathbf{w}}$, i.e., the unknown system coefficients $\mathbf{w}_{o,M}$.

Furthermore, as

$$\lambda_{max} < \text{trace}[\tilde{\mathbf{R}}], \tag{44}$$

and since the trace of a sum of matrices is equal to the sum of the traces of its submatrices, from (38), we can conclude that a sufficient condition for the range of the step-sizes of each band to ensure stability of the presented algorithm is

$$0 < \mu^{(j)} < \frac{2}{\text{trace}[\mathbf{R}^{(j)}]}. \tag{45}$$

4. A Multi-Band LMS Algorithm with Variable Step-Sizes

The multi-band LMS algorithm with VSSs presented in this section is directly obtained from the algorithm derived in the previous section in connection with the step-size adaptation technique described in [2], [9]–[11].

The basic idea is to independently update the step-sizes for every band j as follows

$$\mu^{(j)}(n+1) = \left[\mu^{(j)}(n) - \alpha^{(j)}\hat{\nabla}_{\mu^{(j)}}J^{(j)}_{MB}(n)\right]^{\mu^{(j)}_{max}}_{\mu^{(j)}_{min}} \tag{46}$$

where the bracket with $\mu_{max}^{(j)}$ and $\mu_{min}^{(j)}$ indicates that $\mu^{(j)}(n)$ is upper and lower bounded, $0 < \alpha^{(j)} \ll \mu_{min}^{(j)}$, and $\hat{\nabla}_{\mu^{(j)}} J_{MB}^{(j)}(n)$ is an estimate of

$$\nabla_{\mu^{(j)}} J_{MB}^{(j)}(n) = \frac{\delta J_{MB}^{(j)}(n)}{\delta \mu^{(j)}}. \tag{47}$$

Through a reasoning similar to that exposed in [2], [11], it can be shown that a recursive solution, obtained through an *instantaneous estimate* of $\nabla_{\mu^{(j)}} J_{MB}^{(j)}(n)$, for the update of $\mu^{(j)}(n)$ is given as

$$\mu^{(j)}(n+1) = \Bigg[\mu^{(j)}(n) + \alpha^{(j)} \mathrm{Re}\left[\left(\hat{\mathbf{\Psi}}_M^{(j)}(n)\right)^H \mathbf{u}_M^{(j)}(n) \left(e^{(j)}(n)\right)^* \right] \Bigg]_{\mu_{min}^{(j)}}^{\mu_{max}^{(j)}} \tag{48}$$

$$\hat{\mathbf{\Psi}}_M^{(j)}(n+1) = \left[\mathbf{I} - \mu^{(j)}(n) \mathbf{u}_M^{(j)}(n) \left(\mathbf{u}_M^{(j)}(n)\right)^H \right] \hat{\mathbf{\Psi}}_M^{(j)}(n) + \mathbf{u}_M^{(j)}(n) \left(e^{(j)}(n)\right)^*, \tag{49}$$

where

$$\hat{\mathbf{\Psi}}_M^{(j)}(n) = \frac{\delta \hat{\mathbf{w}}_M(n)}{\delta \mu^{(j)}}. \tag{50}$$

A complete multi-band VSSs algorithm is composed of Eqs (7)–(9), (29), (48), and (49), where $\mu^{(j)}$ in (29) is substituted by $\mu^{(j)}(n)$.

Whenever there is some knowledge of the statistics of the input signal, suitable choices of $\mu_{max}^{(j)}$ and $\mu_{min}^{(j)}$ will give rise to better responses. Special care must be taken when choosing $\mu_{max}^{(j)}$ to avoid instability of the system. A detailed analysis of the influence of the choice of α on the performance of the system as well as a study of merits and limitations of the step-size adaptation technique used in this section were analyzed in some detail in [9]–[11]. Even though various techniques can be used to update the step-sizes in connection with the presented multi-band scheme, the method described in [9] has been chosen because it has been supported by practical applications [10] as well as by theoretical results [11].

Since the algorithms presented in Sect. 3 and in Sect. 4 only differ with respect to the automatic tuning of the step-sizes, a similar analysis of convergence and stability conditions can be undertaken. Specific properties of the adaptive step-size algorithm used in Sect. 4 are analyzed in some detail in [9]–[11].

In the following, the presented multi-band LMS algorithms are applied to a system identification problem.

5. Computer Experiments

Computer experiments are shown in Fig. 2. The algorithms described in this paper were applied in the computation of filter coefficients estimates for the conventional (with $B = 1$) and multi-band LMS methods. The analysis bank used to decompose the estimation error is a 4-band linear-phase near-perfect-reconstruction filter bank designed in [12]. The spectrum of the unknown system begins with a high-frequency component and a low-frequency component. The high-frequency component stops when two spatially-close low-frequency components start. Fig. 2 (a) depicts the true running spectrum. In Fig. 2 (b), the multi-band LMS algorithm with FSSs values of [0.0001 0.0005 0.001 0.01], from lowest to highest bands, was used. In Fig. 2 (c), the multi-band LMS algorithm with VSSs and $\mu_{max}^{(j)}$ = [0.01 0.01 0.01 0.01] $\mu_{min}^{(j)}$ = [0.0001 0.0001 0.0001 0.0001] and $\alpha^{(j)}$ = [0.00001 0.00001 0.00001 0.00001], was applied. Fig. 2 (d) depicts the results obtained with the conventional LMS method for different step-sizes, from left to right: FSS with $\mu = 0.0001$, FSS with $\mu = 0.01$, and VSS with $\mu_{max} = 0.01$, $\mu_{min} = 0.0001$, and $\alpha = 0.00001$. All adaptive filters have 50 coefficients. Noise was added to the desired response. For the experiments shown in this paper, $\hat{\Psi}_M^{(j)}(n)$ is initialized with the null vector.

The experiment results shown in Fig. 2 exemplifies the advantages that multi-band LMS methods can offer as compared to the performance of conventional LMS methods whenever the statistical properties of the analyzed signal vary along the frequency domain. In this particular example, it can be observed that the multi-band method has the following characteristics:

- for the analysis of high-frequency components, as larger step-size values are used, time-resolution is improved at the expense of reducing frequency-resolution;
- for the analysis of low-frequency components, as smaller step-size values are used, frequency resolution is improved at the expense of reducing time-resolution.

Several filter banks have been experimented in the analysis of the error signal. Even though power complementariness is a desired property, designs that approximately satisfy this property have been successfully tested. With respect to the choice of the analysis bank, we may note the following points. As the analysis filters responses overlaps, the trade-off between time and frequency resolutions along the frequency domain varies smoothly. When filters with sharp responses are used, the different time-frequency resolution trade-off settings for different bands become more visible. Also, it must be noted that long filters increase the group delay, consequently, depending on the application, the designer must choose the appropriate bank by considering not only the response of the analysis filters but also the related group delay effect.

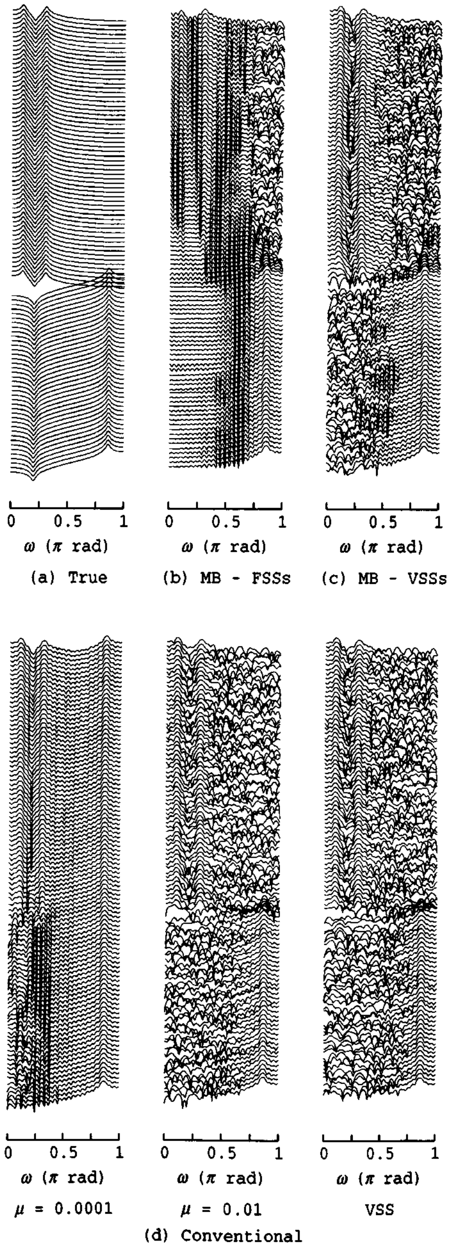


Fig. 2 Running spectrum of the filter coefficients estimates for the nonstationary unknown system shown in (a). Estimates obtained with the multi-band LMS method with FSSs [0.0001 0.0005 0.001 0.01], from lowest to highest bands, are depicted in (b), and estimates obtained with the multi-band LMS with VSSs are plotted in (c), while estimates obtained with the conventional LMS method with FSS of values 0.0001, 0.01, and with a VSS algorithm, from left to right, are displayed in (d). All adaptive filters have 50 coefficients.

6. Conclusions

The theory for a Wiener filter when the estimation error is decomposed into several bands and a particular step-size is correspondingly applied to each band was presented. LMS algorithms with FSSs and VSSs were derived for the multi-band structure. With the FSSs approach, time and frequency resolutions of the adaptive algorithm can be traded off along the frequency domain giving rise to better estimates as compared to those obtained with conventional LMS methods, particularly when the statistical properties of the signal to be analyzed vary along the frequency spectrum. With the VSSs algorithm, time and frequency resolutions are traded off not only in the frequency domain but also in the time domain, making it specially suitable for the analysis of nonstationary signals. Computer experiments comparing the performance of multi-band and conventional LMS methods with a nonstationary test signal are shown for the system identification problem. A study on novel techniques for automatic tuning of the convergence factor and its application to the multi-band scheme is subject of ongoing research [13].

Acknowledgment

The authors would like to express their appreciation to Profs. N. Fujii and S. Takagi of the Tokyo Institute of Technology for their input and encouragement. The authors are also grateful to the anonymous "reviewer 2" for his stimulating critiques. This work was partially supported by CNPQ/Ministry of Education of Brazil and Monbusho/Ministry of Education of Japan.

References

[1] B. Widrow and S. Stearns, "Adaptive signal processing," Englewood Cliffs, N.J., Prentice-Hall, 1985.
[2] S. Haykin, "Adaptive filter theory," Third ed., Englewood Cliffs, N.J., Prentice-Hall, 1996.
[3] P. Diniz, "Adaptive filtering: Algorithms and practical implementation," Boston, M.A., Kluwer Academic Publishers, 1997.
[4] F. Resende, K. Tokuda, and M. Kaneko, "AR spectrum estimation based on wavelet representation," Proc. of IEEE ISCAS, London, England, 1994.
[5] F. Resende, K. Tokuda, and M. Kaneko, "Adaptive AR spectral estimation based on wavelet decomposition of the linear prediction error," IEICE Trans. Fundamentals, vol.E79-A, pp.665–673, May 1996.
[6] F. Resende, K. Tokuda, M. Kaneko, and A. Nishihara, "Multi-band decomposition of the linear prediction error applied to adaptive AR spectral estimation," IEICE Trans. Fundamentals, vol.E80-A, pp.365–376, Feb. 1997.
[7] F. Resende, P. Diniz, K. Tokuda, M. Kaneko, and A. Nishihara, "LMS-based algorithms with multi-band decomposition of the linear prediction error," in Proc. of the 11th DSP Symposium, pp.379–384, Tottori, Japan, 1996.
[8] F. Resende, P. Diniz, K. Tokuda, M. Kaneko, and A. Nishihara, "Multi-band decomposition of the linear prediction error applied to the least-mean-square method with fixed and variable step-sizes," Proc. of IEEE ISCAS, pp.2176–2179, Hong Kong, 1997.
[9] A. Benveniste, M. Metivier, and P. Priouret, "Adaptive algorithms and stochastic approximation," New York, Berlin, Springer-Verlag, 1990.
[10] J. Brossier, "Egalisation adaptive et estimation de phase: Application aux communications sous-marines," Ph.D. thesis, del'Institut National Polytechnique de Grenoble,

1995.
[11] H. Kushner and J. Yang, "Analysis of adaptive step-size SA algorithms for parameter tracking," IEEE Trans. on Automatic Control, vol.40, pp.1403–1410, Aug. 1995.
[12] A. Soman, P. Vaidyanathan, and T. Nguyen, "Linear phase paraunitary filter banks: Theory, factorizations and designs," IEEE Trans. on Signal Processing, vol.41, Dec. 1993.
[13] F. Resende, P. Diniz, M. Kaneko, and A. Nishihara, "Adaptive AR spectral estimation based on multi-band decomposition of the linear prediction error with variable forgetting factors," Proc. of IEEE ICASSP, pp.2293–2296, Munich, Germany, 1997.

Appendix: Proof that Power Complementariness of the Analysis Bank Implies $J_{MB}(n) = cJ(n)$

In the appendix, it is proved that if the analysis bank satisfies the power complementary property, then $J_{MB}(n) = cJ(n)$, that is,

$$E\left[\frac{1}{B}\sum_{j=1}^{B} e^{(j)}(i)\left(e^{(j)}(i)\right)^*\right] = cE[e(i)e^*(i)], \tag{A·1}$$

where c is a positive constant.
Proof:
From the definition of the cost function $J_{MB}(n)$, given in (5), we have

$$J_{MB}(n) = \frac{1}{B}E\left[\sum_{j=1}^{B}\left|e^{(j)}(i)\right|^2\right] \tag{A·2}$$

$$= \frac{1}{B}\frac{1}{2\pi}\int_{-\pi}^{\pi} P_e(\omega)\sum_{j=1}^{B}\left|H^{(j)}(e^{j\omega})\right|^2 d\omega, \tag{A·3}$$

where $P_e(\omega)$ is the power spectrum of $e(n)$. If the analysis bank satisfies the power complementary property, that is,

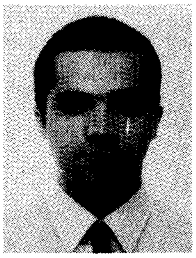
$$\sum_{j=1}^{B}\left|H^{(j)}(e^{j\omega})\right|^2 = c_1, \tag{A·4}$$

where c_1 is a positive constant, it can be seen from (A·3) that

$$J_{MB}(n) = \frac{c_1}{B}\frac{1}{2\pi}\int_{-\pi}^{\pi} P_e(\omega)d\omega \tag{A·5}$$

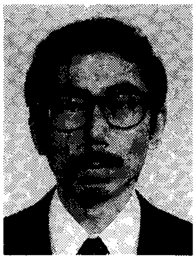
$$= cE[|e(n)|^2] = cJ(n), \tag{A·6}$$

where $c = \frac{c_1}{B}$.



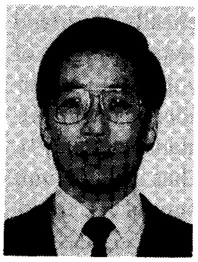
Fernando Gil V. Resende, Jr. received the B.Sc. degree from Instituto Militar de Engenharia, Brazil, in 1990, and the M.Sc. and Ph.D. degrees from Tokyo Institute of Technology, Japan, in 1994 and 1997, respectively, all in electrical and electronic engineering. His research interests include adaptive filtering theory, spectral estimation, and speech processing.

Paulo S.R. Diniz received the B.Sc. degree from Universidade Federal do Rio de Janeiro (UFRJ) in 1978, the M.Sc. degree from COPPE/UFRJ in 1981, and the Ph.D. degree from Concordia University, Montreal, PQ, Canada, in 1984, all in electrical engineering. Since 1979 he has been with the Department of Electronic Engineering, UFRJ. He has also been with the Program of Electrical Engineering, COPPE/UFRJ, since 1984, where he is presently Professor and had served as Chairman. From Jan. 1991 to July 1992 he was a visiting Research Associate in the Department of Electrical and Computer Engineering of University of Victoria, BC, Canada. His teaching and research interests are in analog and digital signal processing, electronic circuits and adaptive signal processing. He is a Senior Member of IEEE and is the Vice President of the IEEE-CAS Society.



Keiichi Tokuda received the B.E. degree in electrical and electronic engineering from the Nagoya Institute of Technology, Nagoya, Japan, the M.E. and D.E. degrees in information processing from the Tokyo Institute of Technology, Tokyo, Japan, in 1984, 1986, and 1989, respectively. From 1989 to 1996 he was a Research Associate at the Department of Electrical and Electronic Engineering, Tokyo Institute of Technology. Since 1996 he has been with the Department of Computer Science, Nagoya Institute of Technology as Associate Professor. His research interests include speech coding, speech synthesis and recognition, and signal processing. He is a member of IEEE and ASJ.

Mineo Kaneko received the B.E., M.E. and Dr.E. degrees in electrical and electronic engineering from Tokyo Institute of Technology in 1981, 1983 and 1986, respectively. In 1986 he joined the Department of Electrical and Electronic Engineering, Tokyo Institute of Technology as a Research Associate. From 1992 to 1996, he was an Associate Professor in the same Department of Tokyo Institute of Technology. Since 1996 he has been an Associate Professor in the School of Information Science, Japan Advanced Institute of Science and Technology, Hokuriku. His research interests include high-speed and fault tolerant VLSI signal processing, computer aided design of integrated circuits and wafer scale integration. He received best paper awards from IEEE APCCAS'92 and APCCAS'94 in 1992 and 1994, respectively. He is a member of IEEE.



Akinori Nishihara received the B.E., M.E. and Dr. Eng. degrees in electronics from Tokyo Institute of Technology in 1973, 1975 and 1978, respectively. Since 1978 he has been with Tokyo Institute of Technology, where he is now Professor of the Center for Research and Development of Educational Technology (CRADLE). His main research interests are in filter design, 1D and multi-D signal processing, and educational technology. From 1990 to 1994 he served as an Associate Editor of the IEICE Trans. Fundamentals, and is now serving as an Associate Editor of the IEEE Trans. Circuits & Systems II. He was 1995–96 Student Activities Committee Chair, IEEE Region 10 (Asia Pacific Region). Dr. Nishihara is a member of IEEE, EURASIP, ECS and JET.