A VARIABLE STEP-SIZE FREQUENCY-DOMAIN ADAPTIVE FILTERING ALGORITHM FOR STEREOPHONIC ACOUSTIC ECHO CANCELLATION

Laura Romoli, Stefano Squartini and Francesco Piazza

A3Lab - DIBET - Università Politecnica delle Marche
Via Brecce Bianche 1, 60131 Ancona Italy
l.romoli@univpm.it
web: www.a3lab.dibet.univpm.it

ABSTRACT

In this paper a gradient adaptive step-size algorithm for frequency domain adaptive filtering is proposed: the step-size adaptation is performed by differentiating the time-domain block error at the output of the block adaptive filter with respect to the step-size. By using the results obtained through a rigorous mathematical treatment in the single-channel situation and in the presence of a partitioned adaptive filter, an adaptation rule valid for two-channel adaptive algorithms, which may be easily extended to multi-channel adaptive algorithms, is derived. Experimental results concerning stereophonic acoustic echo cancellation are reported, in order to validate the effectiveness of the proposed algorithm.

1. INTRODUCTION

The least mean square (LMS) adaptive algorithm is used in a lot of adaptive signal processing applications, such as acoustic echo cancellation, because of its simplicity and ease of computation. In its standard form, it suffers from a slow convergence, which highly depends on the input vector, because the filter coefficients updating is directly proportional to it [1]. Therefore, its normalized version is introduced (NLMS): the normalization is performed by using the Euclidean norm of the input vector. Anyhow, the choice of the step-size is an open issue: there is a trade-off between faster convergence and lower steady-state performances. Stochastic gradient adaptive step-size approaches have been proposed to overcome this problem [2, 3], based on the idea of using for the step-size high values when the adaptive filter coefficients are far from the optimal solution and small values in the steady-state condition. Variations of these algorithms have been proposed, in order to lower the computational cost, especially as regards the Benveniste's approach [4].

The use of frequency-domain adaptive filtering is due to its interesting computational cost: as a matter of fact, the adaptation procedure is performed by taking advantage of the Fast Fourier Transform (FFT) efficiency (Fast-LMS) [5], allowing an improved convergence with low computational requirements [1, 6], even in the presence of long impulse responses (IRs). The main drawback of this algorithm is the input-output delay, because it is equal to the adaptive filter length: long IRs to be identified require long adaptive filters, that means a high delay. In [7], the generalization of frequency-domain adaptive filtering to the partitioning of the adaptive filter is proposed: in this case, the filter length is a positive integer multiple of the block size, which can be opportunely reduced in order to lower the input-output latency.

Recent approaches for varying the step-size in single-channel frequency-domain echo cancellation have been proposed [8, 9]. They are based on a step-size adaptation rule, derived from the computation, directly in the frequency-domain, of the derivative of the mean square error with respect to a parameter which gives an estimation of the misalignment. In [10] a coherence-based step-size control is proposed and applied to stereo echo cancellation. All these algorithms are focused on making the adaptive filtering robust in the presence of double-talk.

In this paper, the extension of the Mathews' stochastic gradient adaptive step-size approach to frequency-domain adaptive filtering is proposed, by differentiating the block error at the output of the block adaptive filter with respect to the step-size: this operation is performed by considering the time-domain error vector, differently from [8, 9], in which the adaptation rule is derived in the frequency-domain. The formula is initially derived for the Fast-LMS adaptive algorithm, even in the presence of a partitioned adaptive filter, and then extended to the stereophonic case, also showing the possibility of an easy generalization to the multi-channel case. As the proposed approach is focused on speeding up the convergence without taking into consideration double-talk situations, it can be applied to generic adaptive algorithms. A very recent approach concerning a bin-wise block varying step-size control method for frequency-domain LMS algorithm has been introduced in [11] for the same purpose, but only in the single-channel case.

The paper is organized as follows: first, a review of Mathews' algorithm will be presented, and then, starting from it, the proposed gradient adaptive step-size algorithm for frequency-domain adaptive filtering will be analytically derived for mono and stereo (multi-channel) case studies. Finally, some experimental results relative to stereophonic acoustic echo cancellation (SAEC) will be reported, in order to validate the proposed approach.

2. MATHEWS' GRADIENT ADAPTIVE STEP-SIZE ALGORITHM REVIEW

Considering a typical problem of impulse response identification, the instantaneous error signal at the output of the filter at the time instant n is defined as follows:

$$e_n = d_n - \mathbf{x}_n^T \mathbf{w}_n \tag{1}$$

where d_n is the desired signal, $\mathbf{x}_n$ is the input signal vector composed of the L most recent samples, $\mathbf{w}_n$ is the estimated impulse response vector of length L and $(\cdot)^T$ is the vector transpose operator. In [2] a variable step-size NLMS algorithm is proposed: the step-size is updated in order to obtain a variation proportional to the gradient of the squared error with respect to the convergence parameter μ, as described by the following equations:

$$\mathbf{w}_{n+1} = \mathbf{w}_n + \mu_n e_n \frac{\mathbf{x}_n}{\|\mathbf{x}_n\|^2} \tag{2}$$

$$\hat{\mu}_n = \mu_{n-1} - \frac{\rho}{2}\frac{\partial e_n^2}{\partial \mu_{n-1}} = \mu_{n-1} + \rho e_n e_{n-1} \frac{\mathbf{x}_n^T \mathbf{x}_{n-1}}{\|\mathbf{x}_{n-1}\|^2} \tag{3}$$

where ρ is a small positive constant, used to control the step-size update, and $\|\cdot\|^2$ is the Euclidean norm. As reported in [12], the convergence requirements limit the allowed values for μ_n during the adaptation. Therefore,

$$\mu_n = \begin{cases} \mu_{max} & \text{if } \hat{\mu}_n > \mu_{max} \\ \mu_{min} & \text{if } \hat{\mu}_n < \mu_{min} \\ \hat{\mu}_n & \text{otherwise} \end{cases} \tag{4}$$

with $0 < \mu_{min} < \mu_{max} < 2$.

3. THE PROPOSED ALGORITHM

Starting from the approach reported in section 2, it is useful to obtain an analogous step-size adaptation rule for frequency-domain adaptive filtering: in this case, the error $\mathbf{e}_m$ is a column vector of length M and eq. (3) can be rewritten in the time-domain as follows:

$$\hat{\mu}_m = \mu_{m-1} - \frac{\rho}{2}\frac{\partial \mathbf{e}_m^T \mathbf{e}_m}{\partial \mu_{m-1}} \tag{5}$$

where m is the block index and M is the length of the block in which the incoming sequence is partitioned. Assuming $\mathbf{X}_m$ and $\mathbf{w}_m$ as the input signal matrix ($M \times L$) and the filter coefficients vector ($L \times 1$) at index m respectively, the block adaptation rule is summarized as follows:

$$\begin{aligned} \mathbf{y}_m &= \mathbf{X}_m \mathbf{w}_{m-1} \\ \mathbf{e}_m &= \mathbf{d}_m - \mathbf{y}_m \\ \mathbf{w}_m &= \mathbf{w}_{m-1} + \mu_m \nabla_m \end{aligned} \tag{6}$$

where ∇_m is the block gradient estimate. In order to calculate eq. (5), it is necessary to introduce the differentiation rules for matrices and vectors [13]. In the following, only the differentiation rules used for the algorithm definition are reported:

$$\frac{\partial \mathbf{AF}}{\partial \mathbf{B}} = \frac{\partial \mathbf{A}}{\partial \mathbf{B}}(\mathbf{I}_t \otimes \mathbf{F}) + (\mathbf{I}_s \otimes \mathbf{A})\frac{\partial \mathbf{F}}{\partial \mathbf{B}} \tag{7}$$

$$\frac{\partial \mathbf{A}[\mathbf{C}(\mathbf{B})]}{\partial \mathbf{B}} = \left(\frac{\partial \operatorname{vec}(\mathbf{C})^T}{\partial \mathbf{B}} \otimes \mathbf{I}_p\right)\left(\mathbf{I}_t \otimes \frac{\partial \mathbf{A}}{\operatorname{vec}(\mathbf{C})}\right) \tag{8}$$

$$\frac{\partial \mathbf{y}}{\partial \mathbf{y}} = \operatorname{vec}(\mathbf{I}_q) \tag{9}$$

$$\frac{\partial \mathbf{y}^T}{\partial \mathbf{y}} = \mathbf{I}_q \tag{10}$$

$$\operatorname{vec}(\mathbf{AD}) = (\mathbf{I}_s \otimes \mathbf{A})\operatorname{vec}(\mathbf{D}) = (\mathbf{D}^T \otimes \mathbf{I}_p)\operatorname{vec}(\mathbf{A}) \tag{11}$$

where $\mathbf{A}$ is a $(p \times q)$ matrix, $\mathbf{F}$ is a $(q \times u)$ matrix, $\mathbf{B}$ is a $(s \times t)$ matrix, $\mathbf{C}$ is a $(r \times l)$ matrix, $\mathbf{y}$ is a $(q \times 1)$ vector, $\mathbf{D}$ is a $(q \times s)$ matrix and $\otimes$ is the Kronecker product.

Therefore, eq. (7) allows to rewrite eq. (5) as follows:

$$\hat{\mu}_m = \mu_{m-1} - \frac{\rho}{2}\left(\frac{\partial \mathbf{e}_m^T}{\partial \mu_{m-1}}\mathbf{e}_m + \mathbf{e}_m^T \frac{\partial \mathbf{e}_m}{\partial \mu_{m-1}}\right) \tag{12}$$

where $\frac{\partial \mathbf{e}_m^T}{\partial \mu_{m-1}}$ can be calculated by using the following equation:

$$\frac{\partial \mathbf{e}_m^T}{\partial \mu_{m-1}} = \frac{\partial (\mathbf{d}_m - \mathbf{y}_m)^T}{\partial \mu_{m-1}} = -\frac{\partial \mathbf{y}_m^T}{\partial \mu_{m-1}} \tag{13}$$

As described in eq. (6), $\mathbf{y}_m$ is a function of $\mathbf{w}_{m-1}$ which is in turn a function of μ_{m-1}. Therefore, by using (8), it results:

$$\begin{aligned} \frac{\partial \mathbf{e}_m^T}{\partial \mu_{m-1}} &= -\left(\frac{\partial \operatorname{vec}(\mathbf{w}_{m-1})^T}{\partial \mu_{m-1}} \otimes \mathbf{I}_1\right)\left(\mathbf{I}_1 \otimes \frac{\partial \mathbf{y}_m^T}{\partial \operatorname{vec}(\mathbf{w}_{m-1})}\right) = \\ &= -\nabla_{m-1}^T \mathbf{X}_m^T \end{aligned} \tag{14}$$

An analogous equation is valid for $\frac{\partial \mathbf{e}_m}{\partial \mu_{m-1}}$:

$$\frac{\partial \mathbf{e}_m}{\partial \mu_{m-1}} = -\mathbf{X}_m \nabla_{m-1} \tag{15}$$

At this stage, by using the results described by eq. (14) and (15), the adaptation rule for the step-size in the mono-channel case for a block adaptive algorithm can be derived as:

$$\begin{aligned} \hat{\mu}_m &= \mu_{m-1} - \frac{\rho}{2}(-\nabla_{m-1}^T \mathbf{X}_m^T \mathbf{e}_m - \mathbf{e}_m^T \mathbf{X}_m \nabla_{m-1}) = \\ &= \mu_{m-1} + \rho \mathbf{e}_m^T \mathbf{X}_m \nabla_{m-1} \end{aligned} \tag{16}$$

The step-size μ_m can be calculated from $\hat{\mu}_m$ by using eq. (4) and $\mathbf{X}_m \nabla_{m-1}$ can be efficiently calculated by using FFT.

3.1 IR partitioning case study

In acoustic echo cancellation applications, it is often useful to partition the impulse response in order to overcome huge latency problems. The equation (6) has to be rewritten by considering the K sections in which the adaptive filter is partitioned. Considering $\mathbf{X}_{k,m}$, $\mathbf{w}_{k,m-1}$ and $\nabla_{k,m}$ as the input signal, the filter coefficients and the block gradient estimate for each section k at block index m, it follows:

$$\begin{aligned} \mathbf{y}_m &= \sum_{k=1}^{K} \mathbf{X}_{k,m} \mathbf{w}_{k,m-1} \\ \mathbf{e}_m &= \mathbf{d}_m - \mathbf{y}_m \\ \mathbf{w}_{k,m} &= \mathbf{w}_{k,m-1} + \mu_m \nabla_{k,m} \text{ for } k = 1,2,...,K \end{aligned} \tag{17}$$

In eq. (17), all the filter sections are updated by using the same step-size, thus eq. (12) is still valid and the two derivatives $\frac{\partial \mathbf{e}_m^T}{\partial \mu_{m-1}}$ and $\frac{\partial \mathbf{e}_m}{\partial \mu_{m-1}}$ can be calculated by mean of the results obtained in section 3. It results:

$$\begin{aligned} \frac{\partial \mathbf{e}_m^T}{\partial \mu_{m-1}} &= \frac{\partial (\mathbf{d}_m - \sum_{k=1}^{K} \mathbf{y}_{k,m})^T}{\partial \mu_{m-1}} = -\sum_{k=1}^{K} \frac{\partial \mathbf{y}_{k,m}^T}{\partial \mu_{m-1}} = \\ &= -\sum_{k=1}^{K} (\nabla_{k,m-1}^T \mathbf{X}_{k,m}^T) \end{aligned} \tag{18}$$

$$\frac{\partial \mathbf{e}_m}{\partial \mu_{m-1}} = -\sum_{k=1}^{K} (\mathbf{X}_{k,m} \nabla_{k,m-1}) \tag{19}$$

The equations (18) and (19) allow to derive the adaptation rule for the step-size in the mono-channel case for a partitioned block adaptive algorithm. It follows:

$$\begin{aligned} \hat{\mu}_m &= \mu_{m-1} + \frac{\rho}{2}\left[\sum_{k=1}^{K} (\nabla_{k,m-1}^T \mathbf{X}_{k,m}^T)\right]\mathbf{e}_m + \\ &+ \frac{\rho}{2}\mathbf{e}_m^T \sum_{k=1}^{K} (\mathbf{X}_{k,m} \nabla_{k,m-1}) = \mu_{m-1} + \rho \mathbf{e}_m^T \sum_{k=1}^{K} (\mathbf{X}_{k,m} \nabla_{k,m-1}) \end{aligned} \tag{20}$$

As already noted in section 3, the step-size μ_m is calculated from $\hat{\mu}_m$ by using eq. (4) and $\sum_{k=1}^{K} (\mathbf{X}_{k,m} \nabla_{k,m-1})$ can be efficiently calculated in the frequency-domain for all the k sections. These results are consistent to those obtained in [9] for the frequency-domain, except for the differentiation with respect to the misalignment estimation, which has not been taken into consideration. In the following, it is described as the proposed approach can be easily extended to the stereophonic (multi-channel) case.

3.2 Extension to the stereophonic case

The extension of the proposed approach to the stereophonic case can be easily obtained by considering the previous results. Assuming a two-channel situation, where $\mathbf{X}_m^{(p)}$, $\mathbf{w}_{m-1}^{(p)}$ and $\nabla_m^{(p)}$ are the input signal, the filter coefficients and the block gradient estimate, respectively, for each channel p ($p = 1,2$), eq. (6) has to be modified as follows

$$\begin{aligned} \mathbf{y}_m &= \mathbf{X}_m^{(1)} \mathbf{w}_{m-1}^{(1)} + \mathbf{X}_m^{(2)} \mathbf{w}_{m-1}^{(2)} \\ \mathbf{e}_m &= \mathbf{d}_m - \mathbf{y}_m \\ \mathbf{w}_m^{(1)} &= \mathbf{w}_{m-1}^{(1)} + \mu_m \nabla_m^{(1)} \\ \mathbf{w}_m^{(2)} &= \mathbf{w}_{m-1}^{(2)} + \mu_m \nabla_m^{(2)} \end{aligned} \tag{21}$$

This is a simple extension of the mono-channel Fast-LMS to the two-channel algorithm, in which the step-size normalization factor in the adaptation rule involves only the auto power spectra of the input channels.

If the two filters $\mathbf{w}_m^{(1)}$ and $\mathbf{w}_m^{(2)}$ are updated by using the same step-size, eq. (12) is still valid and the two derivatives $\frac{\partial \mathbf{e}_m^T}{\partial \mu_{m-1}}$ and $\frac{\partial \mathbf{e}_m}{\partial \mu_{m-1}}$ can be calculated by mean of the previous results. It follows:

$$\begin{aligned}\frac{\partial \mathbf{e}_m^T}{\partial \mu_{m-1}} &= \frac{\partial\left(\mathbf{d}_m - \mathbf{y}_m^{(1)} - \mathbf{y}_m^{(2)}\right)^T}{\partial \mu_{m-1}} = -\frac{\partial \mathbf{y}_m^{(1)^T}}{\partial \mu_{m-1}} - \frac{\partial \mathbf{y}_m^{(2)^T}}{\partial \mu_{m-1}} = \\ &= -\nabla_{m-1}^{(1)^T}\mathbf{X}_m^{(1)^T} - \nabla_{m-1}^{(2)^T}\mathbf{X}_m^{(2)^T} \end{aligned} \tag{22}$$

$$\frac{\partial \mathbf{e}_m}{\partial \mu_{m-1}} = -\mathbf{X}_m^{(1)}\nabla_{m-1}^{(1)} - \mathbf{X}_m^{(2)}\nabla_{m-1}^{(2)} \tag{23}$$

At this stage, by using the results described by eq. (22) and (23), the adaptation rule for the step-size in the two-channel case is derived as:

$$\begin{aligned}\hat{\mu}_m &= \mu_{m-1} + \frac{\rho}{2}\left(\nabla_{m-1}^{(1)^T}\mathbf{X}_m^{(1)^T}\mathbf{e}_m + \nabla_{m-1}^{(2)^T}\mathbf{X}_m^{(2)^T}\mathbf{e}_m\right) + \\ &+ \frac{\rho}{2}\left(\mathbf{e}_m^T\mathbf{X}_m^{(1)}\nabla_{m-1}^{(1)} + \mathbf{e}_m^T\mathbf{X}_m^{(2)}\nabla_{m-1}^{(2)}\right) = \\ &= \mu_{m-1} + \rho\mathbf{e}_m^T\left(\mathbf{X}_m^{(1)}\nabla_{m-1}^{(1)} + \mathbf{X}_m^{(2)}\nabla_{m-1}^{(2)}\right)\end{aligned} \tag{24}$$

As already noted in previous sections, the step-size μ_m can be obtained from $\hat{\mu}_m$ by using eq. (4) and $\mathbf{X}_m^{(1)}\nabla_{m-1}^{(1)}$ and $\mathbf{X}_m^{(2)}\nabla_{m-1}^{(2)}$ can be efficiently calculated in the frequency-domain.

These results have to be matched with the results obtained in section 3.1 in order to derive a step-size adaptation rule for a stereophonic algorithm with partitioned IRs. The equation (21) has to be modified as follows:

$$\begin{aligned}\mathbf{y}_m &= \sum_{k=1}^{K}\left(\mathbf{X}_{k,m}^{(1)}\mathbf{w}_{k,m-1}^{(1)}\right) + \sum_{k=1}^{K}\left(\mathbf{X}_{k,m}^{(2)}\mathbf{w}_{k,m-1}^{(2)}\right) \\ \mathbf{e}_m &= \mathbf{d}_m - \mathbf{y}_m \\ \mathbf{w}_{k,m}^{(1)} &= \mathbf{w}_{k,m-1}^{(1)} + \mu_m\nabla_{k,m}^{(1)} \text{ for } k = 1,2,\ldots,K \\ \mathbf{w}_{k,m}^{(2)} &= \mathbf{w}_{k,m-1}^{(2)} + \mu_m\nabla_{k,m}^{(2)} \text{ for } k = 1,2,\ldots,K\end{aligned} \tag{25}$$

In eq. (25), the filters relative to each channel and each partition are updated by using the same step-size, thus eq. (12) is still valid. It results:

$$\begin{aligned}\frac{\partial \mathbf{e}_m^T}{\partial \mu_{m-1}} &= \frac{\partial\left(\mathbf{d}_m - \sum_{k=1}^{K}\mathbf{y}_{k,m}^{(1)} - \sum_{k=1}^{K}\mathbf{y}_{k,m}^{(2)}\right)^T}{\partial \mu_{m-1}} = \\ &= -\sum_{k=1}^{K}\left(\nabla_{k,m-1}^{(1)^T}\mathbf{X}_{k,m}^{(1)^T}\right) - \sum_{k=1}^{K}\left(\nabla_{k,m-1}^{(2)^T}\mathbf{X}_{k,m}^{(2)^T}\right)\end{aligned} \tag{26}$$

$$\frac{\partial \mathbf{e}_m}{\partial \mu_{m-1}} = -\sum_{k=1}^{K}\left(\mathbf{X}_{k,m}^{(1)}\nabla_{k,m-1}^{(1)}\right) - \sum_{k=1}^{K}\left(\mathbf{X}_{k,m}^{(2)}\nabla_{k,m-1}^{(2)}\right) \tag{27}$$

The adaptation rule for the step-size in the two-channel case with partitioning of the impulse response is described as follows:

$$\begin{aligned}\hat{\mu}_m &= \mu_{m-1} + \rho\mathbf{e}_m^T\sum_{k=1}^{K}\left(\mathbf{X}_{k,m}^{(1)}\nabla_{k,m-1}^{(1)}\right) + \rho\mathbf{e}_m^T\sum_{k=1}^{K}\left(\mathbf{X}_{k,m}^{(2)}\nabla_{k,m-1}^{(2)}\right) = \\ &= \mu_{m-1} + \rho\sum_{p=1}^{2}\left[\mathbf{e}_m^T\sum_{k=1}^{K}\left(\mathbf{X}_{k,m}^{(p)}\nabla_{k,m-1}^{(p)}\right)\right]\end{aligned} \tag{28}$$

As already noted in previous sections, the step-size μ_m can be obtained from $\hat{\mu}_m$ by using eq. (4) and the quantities of interest can be efficiently calculated in the frequency-domain.

The equation (28) can be easily generalized to the multi-channel case, by considering P channels.

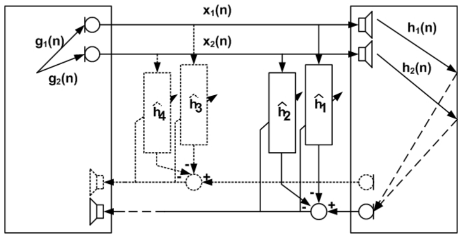
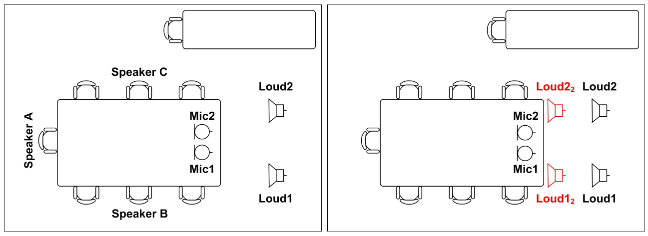


Figure 1: Scheme of a SAEC applied to teleconferencing.



(a) (b)

Figure 2: Scheme of the transmission (a) and the receiving (b) rooms.

4. EXPERIMENTAL RESULTS

Some experimental results concerning acoustic echo cancellation are described in this section, in order to validate the proposed method, focusing on the stereophonic case. Fig. 1 shows a schematic diagram of the application of stereophonic acoustic echo cancellation for teleconferencing. The adaptive algorithm used to identify the room IRs is described by eq. (25). $\rho = 0.0004$ and $\mu_0 = 0.001$ (as initial value) have been used for the variable step-size. It is worth underlying that ρ has to be a small positive constant, in order to prevent $\hat{\mu}_m$ from varying too quickly. Moreover, it has been noted that a small initial value for the step-size allows a better control at the beginning of the convergence in terms of filter stability. In the presence of a fixed step-size, it has been considered a vector $\mu = [\mu_0, \mu_1, \ldots, \mu_{K-1}]$, so that each filter section is characterized by a specific value for the step-size, that decreases according to an exponential decay. Indeed, the filter coefficients have to be weighted proportional to the expected variation of an IR: this variation gradually decreases according to an exponential behaviour that reflects the IR energy decay [14]. In the same manner, the filter sections have to be characterized by a step-size that exponentially decreases, as described by the following equations:

$$\mu_k = \begin{cases}\text{const} & \text{if } k = 0 \\ \mu_{k-1}\gamma & \text{if } k = 1,2,\ldots,K-1\end{cases} \tag{29}$$

where k is the section index and γ is a constant which describes the exponential decay.

In all tests, the adaptive filter has the same length of the room IRs (i.e. $N = L = 4096$ samples). It has been considered an input block size of 256 samples, thus the adaptive filters are partitioned into $K = 16$ sections. In order to overcome the non-uniqueness problem, which characterizes the multi-channel echo cancellation, a pre-processing of the two input channels has been applied, based on the approach proposed in [15]. Different test sessions have been carried out on a speech signal sampled at 16 kHz, in order to study the performances of the proposed approach in different operative situations:

1. a fixed speaker in the transmission room (*Test 1*);
2. two alternating speakers in the transmission room (*Test 2*);
3. a fixed speaker in the transmission room and an abrupt change of the IRs in the receiving room (*Test 3*).

		Misalignment [dB]	ERLE [dB]
Test 1	variable μ	-11.02	-25.12
	$\mu = 0.001$	-1.08	-2.21
	$\mu = 0.01$	-6.66	-13.93
	$\mu = 0.02$	-9.17	-19.32
	$\mu = 0.03$	-10.22	-22.09
	$\mu = 0.04$	-10.52	-23.51
Test 2	variable μ	-9.38	-22.16
	$\mu = 0.001$	-0.87	-2.30
	$\mu = 0.01$	-5.53	-13.54
	$\mu = 0.02$	-7.82	-18.46
	$\mu = 0.03$	-8.73	-20.33

Table 1: Misalignment and ERLE values after 10 s of simulation.

A Personal Computer running NU-Tech platform [16] has been used for evaluating performance, according to the following aspects:

- Echo Return Loss Enhancement (ERLE)

$$\mathrm{ERLE}_n = 10\log_{10}\frac{e_n^2}{d_n^2} \tag{30}$$

- Misalignment

$$\mathrm{Misalignment}_m = 20\log_{10}\frac{\|h_m - w_m\|^2}{\|h_m\|^2} \tag{31}$$

where h_m is the true echo path at block index m; the misalignment has been reported for just one of the two echo paths.

4.1 Transmission and receiving rooms setup

Tests have been carried out by using the IRs measured in a real room. Fig. 2(a) shows a scheme of the transmission room, describing the positions of speakers, microphones and loudspeakers. The acoustic paths from *SpeakerA* to the microphones *Mic*1 and *Mic*2 are used to derive the stereophonic signal sent to the receiving room. As regards *SpeakerB* and *SpeakerC*, the real recorded signal has been used for tests, thus the IRs matched to these positions have not been taken into consideration.

Fig. 2(b) shows a scheme of the receiving room. Two positions for each loudspeaker have been considered: the stereo echo canceller has initially to identify the acoustic paths from the loudspeakers *Loud*1 and *Loud*2 to *Mic*1 and then the acoustic paths from the loudspeakers *Loud*1_2 and *Loud*2_2 to *Mic*1, in order to test its behaviour in the presence of a variation in the receiving room.

4.2 Test 1

In this test session a fixed speaker in the transmission room has been considered (*SpeakerA* in Fig. 2(a)). Test results are shown in terms of misalignment (Fig. 3(a)) and ERLE (Fig. 3(b)): the step-size behaviour is reported in Fig. 3(c). It is evident that, in the presence of a variable step-size, the algorithm outperforms the performances obtained with a fixed step-size, whose values totally cover the range of values reached by the variable step-size. Indeed, too small and too high values force the adaptive filters with a fixed step-size to converge too slowly and too fastly, respectively, while a variable step-size is more flexible to the different convergence stage. In Tab. 1, misalignment and ERLE values (dB) after 10 s of simulation are reported.

4.3 Test 2

In this test session *SpeakerB* takes turn with *SpeakerC* (Fig. 2(a)). A real recorded speech signal has been applied to the stereophonic echo canceller. As in section 4.2, results are shown in terms of misalignment (Fig. 4(a)) and ERLE (Fig. 4(b)): they confirm the expected results, as shown even in Tab. 1 concerning misalignment and ERLE values (dB) after 10 s of simulation. Step-size behaviour has not been reported for the sake of brevity.

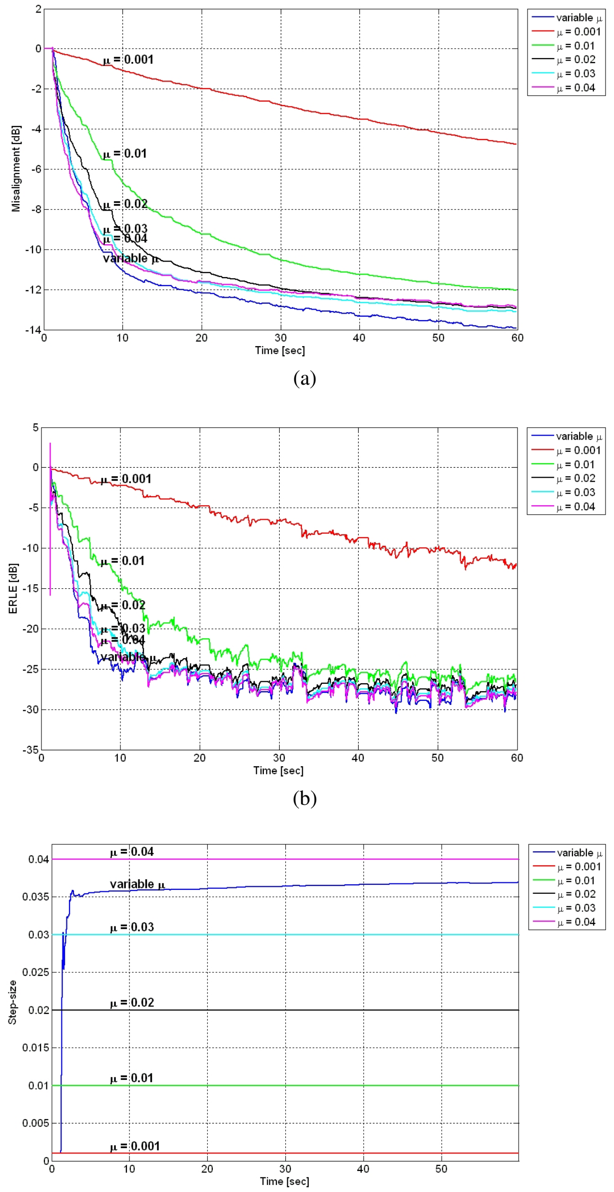


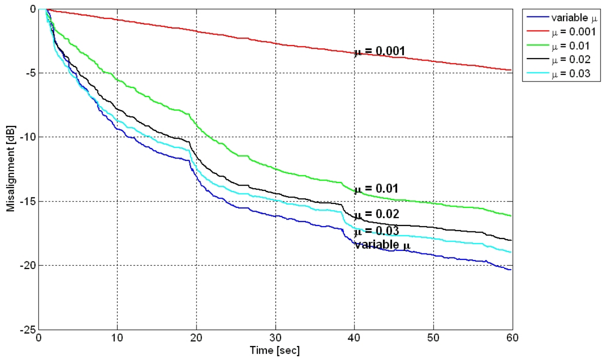
Figure 3: Misalignment (a), ERLE (b) and μ behaviour (c) for *Test 1*.

4.4 Test 3

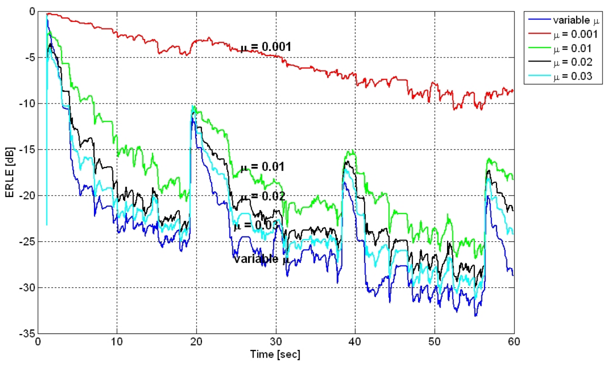
In this test session, a fixed speaker (*SpeakerA* in Fig. 2(a)) as in section 4.2 has been considered, but an abrupt change is applied in the receiving room after about 31 seconds (shift of the loudspeakers positions as described in Fig. 2(b)), in order to show the benefits of the variable step-size in the presence of an echo path variation. Results are shown in terms of ERLE (Fig. 5(a)), while Fig. 5(b) shows step-size behaviour, confirming again the advantages of using a variable step-size algorithm.

5. CONCLUSIONS

In this paper a frequency-domain adaptive algorithm with a time-varying step-size has been proposed, in order to improve convergence performances. First, starting from previous approaches related to NLMS adaptive algorithm, a step-size adaptation rule for Fast-LMS algorithm has been analytically derived, even consider-

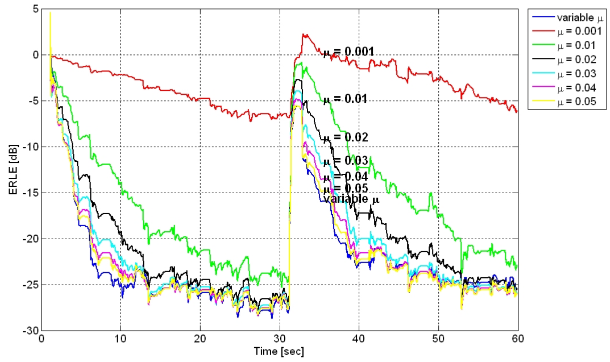


(a)

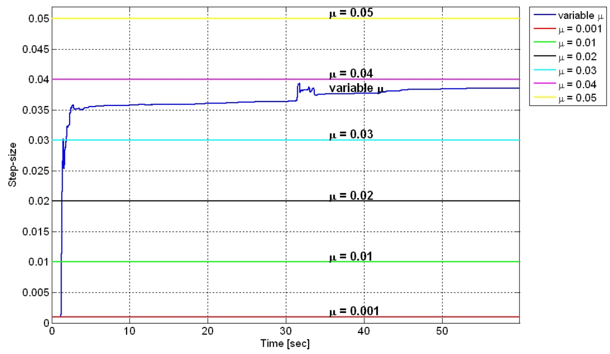


(b)

Figure 4: Misalignment (a) and ERLE (b) for *Test 2*.



(a)



(b)

Figure 5: ERLE (a) and μ behaviour (b) for *Test 3*.

ing a partitioned impulse response. Then, the obtained results have been used for the extension of this formula to the two-channel adaptive algorithm, showing the possibility of a simple generalization to the multi-channel case. Several test sessions concerning stereophonic acoustic echo cancellation have been carried out, in order to validate the effectiveness of the proposed algorithm in different operative situations.

REFERENCES

[1] S. Haykin, *Adaptive Filter Theory*. Prentice-Hall, Inc., 1996.

[2] V. Mathews and Z. Xie, "A stochastic gradient adaptive filter with gradient adaptive step size," *Signal Processing, IEEE Transactions on*, vol. 41, no. 6, pp. 2075–2087, Jun 1993.

[3] M. M. A. Benveniste and P. Priouret, *Adaptive Algorithms and Stochastic Approximations*. Springer-Verlag, 1990.

[4] W.-P. Ang and B. Farhang-Boroujeny, "A new class of gradient adaptive step-size LMS algorithms," *Signal Processing, IEEE Transactions on*, vol. 49, no. 4, pp. 805–810, Apr 2001.

[5] E. Ferrara, "Fast implementations of LMS adaptive filters," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 28, no. 4, pp. 474–475, Aug 1980.

[6] J. Shynk, "Frequency-domain and multirate adaptive filtering," *Signal Processing Magazine, IEEE*, vol. 9, no. 1, pp. 14–37, Jan 1992.

[7] J. Benesty and P. Duhamel, "A fast exact least mean square adaptive algorithm," *Signal Processing, IEEE Transactions on*, vol. 40, no. 12, pp. 2904–2920, Dec 1992.

[8] J.-M. Valin, "On adjusting the learning rate in frequency domain echo cancellation with double-talk," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 15, no. 3, pp. 1030–1034, March 2007.

[9] J. Valin and I. Collings, "A new robust frequency domain echo canceller with closed-loop learning rate adaptation," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 1, April 2007, pp. I–93–I–96.

[10] S. Emura and Y. Haneda, "A method of coherence-based step-size control for robust stereo echo cancellation," *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*, vol. 5, pp. 592–595, Apr 2003.

[11] K. Shi and X. Ma, "A frequency domain step-size control method for lms algorithms," *Signal Processing Letters, IEEE*, vol. 17, no. 2, pp. 125–128, Feb 2010.

[12] A. I. Sulyman and A. Zerguine, "Convergence and steady-state analysis of a variable step-size NLMS algorithm," *Signal Processing*, vol. 83, no. 6, pp. 1255–1273, 2003.

[13] J. Brewer, "Kronecker products and matrix calculus in system theory," *Circuits and Systems, IEEE Transactions on*, vol. 25, no. 9, pp. 772–781, Sep 1978.

[14] S. Makino, Y. Kaneda, and N. Koizumi, "Exponentially weighted stepsize NLMS adaptive filter based on the statistics of a room impulse response," *Speech and Audio Processing, IEEE Transactions on*, vol. 1, no. 1, pp. 101–108, Jan 1993.

[15] J. Herre, H. Buchner, and W. Kellermann, "Acoustic echo cancellation for surround sound using perceptually motivated convergence enhancement," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 1, April 2007, pp. I–17–I–20.

[16] A. Lattanzi, F. Bettarelli, and S. Cecchi, "NU-Tech: The Entry Tool of the hArtes Toolchain for Algorithms Design," *124th AES Convention, proceedings of*, May 2008.