ELectronics EXpress

IEICE Electronics Express, Vol.6, No.1, 20–26

Delay coefficients based variable step size algorithm for subband affine projection adaptive filters

Karthik Muralidhar[1a)], **Anoop Kumar Krishna**[2], **Kwok Hung Li**[3], **and Sapna George**[1b)]

[1] *STMicroelectronics Asia pacific Pte. Ltd., Singapore*

[2] *EADS Singapore Pte. Ltd.*

[3] *Nanyang Technological University, Singapore*

a) *mr.karthik@st.com*

b) *sapna.george@st.com*

Abstract: Subband adaptive filters are preferred in acoustic echo cancellation systems with long echo tail lengths due to speed of convergence and complexity savings. Recently, a new and novel subband affine projection (SAP) algorithm was reported based on the polyphase decomposition of the adaptive filter and noble identities. For good system performance it is important to have a good variable step size (VSS) algorithm as part of an adaptive filter. In this paper, based on the method of delay coefficients (DC), we propose[1] a VSS algorithm for the SAP adaptive filter, which is called as delay coefficients based variable step size subband affine projection algorithm (DC-VSS-SAP). We examine in detail the similarities and differences between DC method for the subband and fullband scenarios. Further, we show how the method of DC can be used to detect changes in echo paths and speed up convergence of the adaptive filter.

Keywords: subband affine projection, variable step size, delay coefficients

Classification: Science and engineering for electronics

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 Process.*, vol. 40, no. 8, pp. 1862–1875, Aug. 1992.

[2] H. Choi, S. H. Han, and H. D. Bae, "Subband adaptive filtering with maximal decimation using an affine projection algorithm," *IEICE Trans. Commun.*, vol. E89-B, no. 5, pp. 1477–1485, May 2006.

[1]The algorithms are protected and are part of a US(PSP)-61002897 patent filed by STMicroelectronics Pte. Ltd. in USPTO. This work was done while the second author was employed at STMicroelectronics.

DOI: 10.1587/elex.6.20
Received October 31, 2008
Accepted November 21, 2008
Published January 10, 2009

IEICE Electronics Express, Vol.6, No.1, 20–26

[3] H. Choi and H. D. Bae, "Subband affine projection algorithm for acoustic echo cancellation system," *EURASIP Journal on Advances in Signal Processing*, 2007.
[4] A. Mader, H. Puder, and G. Schmidt, "Step-size control for echo cancellation filters-An overview," *Signal Process.*, vol. 80, pp. 1697–1719, Sept. 2000.
[5] S. Haykin, "Normalized Least-Mean-Square Adaptive Filters," in *Adaptive Filter Theory*, fourth edition, pp. 320–344, Pearson Education Low Price Edition, Delhi, 2005.

1 Introduction

Subband adaptive filters are preferred in acoustic echo cancellation (AEC) systems with long echo tail lengths as it results in improved convergence and reduced complexity. Recently, affine projection adaptive filters (AF) are preferred as it results in RLS-like performance with LMS-like complexity. The use of subband adaptive filters generally require the presence of cross filters between adjacent subbands [1]. Recently, a new and novel subband affine projection algorithm (SAP) was published [2, 3]. The novelty of this method is that it avoids the use of cross filters by making use of the polyphase decomposition of the adaptive filter and noble identities.

Speech is a non-stationary signal with varying input powers and silence periods. In AEC applications the use of a variable step size (VSS) algorithm to control the step size of the AF is very important to achieve good system performance. More specifically, a high step size is needed during the transient phase of the AF or if there is a change in echo path to achieve quick convergence. Further, a small step size (close to 0) is needed when the far-end signal power is low (silence periods), near-end signal is present or the AF is in the steady state. We therefore develop a VSS algorithm for the SAP AF proposed in [2, 3]. The VSS algorithm is based on the method of delay coefficients (DC) [4, Section 4.2.1], [5, Section 6.3] and is called as delay coefficients-based variable step size subband affine projection algorithm (DC-VSS-SAP). We examine the similarities and differences between the proposed DC-VSS-SAP and its corresponding fullband method in [4, Eq. (19)], called as delay coefficients based affine projection (DC-AP) algorithm. We show that when the input signal is white in nature the subband method assumes a form which is very similar to that of the fullband case. Further, we show that the method of DC can be used to detect changes in echo paths and speed up the convergence of the AF in such scenarios.

This paper is organized as follows. In Section II, we describe the method of SAP [2, 3]. In Section III, we derive the DC-VSS-SAP algorithm based on the method of DC. In Section IV, we show how the DC method is used to detect changes in echo paths and speed up the reconvergence of the AF. Section V is about results and discussions and Section VI presents conclusions of the study in this paper.

DOI: 10.1587/elex.6.20
Received October 31, 2008
Accepted November 21, 2008
Published January 10, 2009

IEICE Electronics Express, Vol.6, No.1, 20–26

2 Subband Affine Projection Algorithm

A *novel subband affine projection* algorithm based on the polyphase decomposition of the AF was recently proposed in [2, 3]. In this section, we briefly introduce the SAP algorithm. The notations throughout the paper will be consistent with the ones used in [2]. SAP is described with respect to Fig. 3 in [2] which uses the polyphase decomposition of the adaptive filters and noble identities. The system corresponds to M subbands and it is desired to estimate an unknown echo path denoted by an $N \times 1$ vector $\mathbf{s}^*$. The projection order is denoted by P. The subband projection order is $P_s = P/M$. Through out this paper we assume the indices $i, j \in [0, \ldots, M-1]$. Using the analysis filters $\mathbf{h}_0, \ldots, \mathbf{h}_{M-1}$, the far-end signal, $x(n)$, and the near-end signal, $d(n)$, are partitioned into new sets of subband signals $x_{ij}(n)$ and $d_i(n)$, respectively. The M polyphase components of filter $\mathbf{s}^*$ are denoted by the $N/M \times 1$ vectors $\mathbf{s}_i^*$ and their z-transforms are related as $\mathbf{S}^*(z) = \sum_{i=0}^{M-1} z^{-i}\mathbf{S}_i^*(z^M)$. The polyphase weight vectors estimated by the AF at time instant n are denoted by the column vectors $\mathbf{s}_i(n)$ while the subband error signals are denoted by $e_i(n)$. Matrices $\mathbf{U}(n)$ and $\boldsymbol{\phi}(n) = \mathbf{U}^T(n)\mathbf{U}(n)$ comprise of far-end signal entries [2, Eq. (38), Eq. (39)] while $\mathbf{E}(n)$ is an error-signal vector [2, Eq. (40)]. Note that $\boldsymbol{\phi}(n)$ is dependent on the far-end signal power. Denote the step size by $\mu(n)$. Define $\mathbf{S}(n) = \left[\mathbf{s}_0^T(n), \ldots, \mathbf{s}_{M-1}^T(n)\right]^T$, $\tilde{\boldsymbol{\phi}}(n) = \boldsymbol{\phi}(n) + \delta\mathbf{I}_P$, where $\mathbf{I}_P$ is a $P \times P$ identity matrix and δ is a small regularization factor which prevents amplification of near-end noise when the far-end signal power is low [5, pp. 339]. The tap-update equation is given by

$$\mathbf{S}(n+1) = \mathbf{S}(n) + \mu(n)\mathbf{U}(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\mathbf{E}(n)\cdot \tag{1}$$

Define the *weight-error vector* $\tilde{\mathbf{S}}(n) = [\tilde{\mathbf{s}}_0^T(n), \ldots, \tilde{\mathbf{s}}_{M-1}^T(n)]^T$, where $\tilde{\mathbf{s}}_i(n) = \mathbf{s}_i^* - \mathbf{s}_i(n)$. The *undisturbed error vector* is defined as $\boldsymbol{\xi}(n) = \tilde{\mathbf{S}}^T(n)\mathbf{U}(n)$. Let $v_i(n)$ denote the ith subband component of near-end background noise $v(n)$. We have

$$\mathbf{E}(n) = \boldsymbol{\xi}^T(n) + \mathbf{v}(n) \tag{2}$$

where $\mathbf{v}(n) = \left[\mathbf{v}_0^T(n), \ldots, \mathbf{v}_{(M-1)}^T(n)\right]^T$ and $\mathbf{v}_i(n) = [v_i(n), \ldots, v_i(n-P_s+1)]^T$. The mean square deviation (MSD) is defined as $D(n) = E\left\{\tilde{\mathbf{S}}^T(n)\tilde{\mathbf{S}}(n)\right\}$. To prevent filter divergence, $\mu_{\text{opt}}(n)$ should be chosen such that $D(n+1) < D(n)$. The permissible range of $\mu(n)$ to achieve this is $0 < \mu(n) < 2\mu_{\text{opt}}(n)$ [2, Eq. (47)] and fastest convergence is achieved for

$$\mu_{\text{opt}}(n) = \frac{E\left\{\boldsymbol{\xi}(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\boldsymbol{\xi}^T(n)\right\}}{E\left\{\mathbf{E}^T(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\mathbf{E}(n)\right\}}. \tag{3}$$

3 Delay Coefficients based VSS-SAP (DC-VSS-SAP)

The AF is said to diverge if (1) results in large values of $\mathbf{S}(n+1)$. This happens due to what is known as near-end noise amplification [5, Sec. 6.5]. More specifically, during silence periods of the far-end signal, $\tilde{\boldsymbol{\phi}}(n)$ $\left(\tilde{\boldsymbol{\phi}}^{-1}(n)\right)$ is low (high), resulting in large values of $\mathbf{S}(n+1)$. The large value of $\mathbf{S}(n+1)$

DOI: 10.1587/elex.6.20
Received October 31, 2008
Accepted November 21, 2008
Published January 10, 2009

also occurs during the presence of a near-end signal ($\mathbf{E}(n)$ is high). In both the above scenarios, the VSS algorithm chooses very low values of $\mu(n)$ to offset the high values of $\mathbf{U}(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\mathbf{E}(n)$. It thereby ensures that $\mathbf{S}(n+1)$ does not attain large values, i.e., the AF does not diverge. In what follows we present the proposed DC-VSS-SAP algorithm.

Note that $\boldsymbol{\xi}(n)$ is not a measurable quantity in (3) as it is dependent on the unknown $\mathbf{s}^*$ and hence one cannot compute $\mu_{\text{opt}}(n)$ as per (3). To compute $\mu_{\text{opt}}(n)$, we need to use the method of delay coefficients [4]. In this section, we present DC-VSS-SAP. Let $\mathbf{K}(n) = E\{\tilde{\mathbf{S}}(n)\tilde{\mathbf{S}}^T(n)\}$ denote the *weight-error* correlation matrix. Define $\sigma_{\xi}^2(n) = E\left\{\boldsymbol{\xi}(n)\boldsymbol{\xi}^T(n)\right\}$. For a white input signal, it can be shown that [4, Eq. (18)], [5, Eq. (6.29)]

$$\sigma_{\xi}^2(n) \quad = \quad \frac{1}{N}D(n)\cdot \text{tr}\left\{\boldsymbol{\phi}(\text{n})\right\} \tag{4}$$

where tr denotes the trace of a matrix. Note that (4) holds good only for white input and is only approximately correct for speech signals (colored signal) [4, Eq. (18)]. The fluctuations in $\boldsymbol{\phi}^{-1}(n)$ from one iteration to the next are small enough compared to $\boldsymbol{\xi}(n)$ and $\mathbf{E}(n)$ [5, Eq. (6.22)]. Making use of the identity $\text{tr}\{\mathbf{MN}\}=\text{tr}\{\mathbf{NM}\}$ [5], we have

$$\begin{aligned} E\{\mathbf{E}^T(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\mathbf{E}(n)\} &= E\{\text{tr}\{\mathbf{E}^T(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\mathbf{E}(n)\}\} \\ &= E\{\text{tr}\{\mathbf{E}(n)\mathbf{E}^T(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\}\} \\ &\approx \text{tr}\{E\{\mathbf{E}(n)\mathbf{E}^T(n)\}\tilde{\boldsymbol{\phi}}^{-1}(n)\}, \\ E\{\boldsymbol{\xi}(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\boldsymbol{\xi}^T(n)\} &= E\{\text{tr}\{\mathbf{U}^T(n)\tilde{\mathbf{S}}(n)\tilde{\mathbf{S}}^T(n)\mathbf{U}(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\}\} \\ &\approx (1/N)\, D(n)\text{tr}\{\boldsymbol{\phi}(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\} \end{aligned} \tag{5}$$

where we have retained only the diagonal elements of $\mathbf{K}(n)$. Using these results in (3), the DC-VSS-SAP method calculates $\mu_{\text{opt}}(n)$ as

$$\mu_{\text{opt}}(n) = \frac{\hat{D}(n)\cdot \text{tr}\left(\boldsymbol{\phi}(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\right)}{N\cdot \text{tr}\left(R_{\mathbf{E}}(n)\tilde{\boldsymbol{\phi}}^{-1}(n)\right)} \tag{6}$$

where $R_{\mathbf{E}}(n)$ is computed as $R_{\mathbf{E}}(n) = \alpha R_{\mathbf{E}}(n-1) + (1-\alpha)\mathbf{E}(n)\mathbf{E}^T(n)$ and α is a smoothing constant. Let $[\mathbf{s}_i(n)]_p$ denote the pth element of the vector $\mathbf{s}_i(n)$. An estimate of MSD is given by [4, Eq. (23)], [5, Eq. (6.33)]

$$\hat{D}(n) = \frac{N}{b}\sum_{p=0}^{b-1}\left([\mathbf{s}^* - \mathbf{S}(n)]_p\right)^2 = \frac{N}{b}\sum_{i=0}^{M-1}\sum_{p=0}^{\frac{b}{M}-1}\left\{[\mathbf{s}_i(n)]_p\right\}^2 \tag{7}$$

where, as required by the method of DC, the first b values of $\mathbf{s}^*$ is 0. (This means that the first b/M values of $\mathbf{s}_i(n)$ are 0). This can be achieved by delaying the loudspeaker feed by b samples.

For colored inputs, $\boldsymbol{\phi}(n)$ is a block diagonal matrix [2, Eq. (39)], and for white inputs, it is a diagonal matrix. Retaining only the main diagonal of $\boldsymbol{\phi}(n)$ in (6) (the other diagonals are equal to 0), we arrive at DC based main diagonal approximation SAP (DC-MDA-SAP), which computes $\mu_{\text{opt}}(n)$ as

$$\mu_{\text{opt}}(n) = \frac{\hat{D}(n)}{\frac{1}{M}\sum_{i=0}^{M-1}\frac{E\{e_i^2(n)\}}{E\{x_i^2(n)\}}}. \tag{8}$$

DOI: 10.1587/elex.6.20
Received October 31, 2008
Accepted November 21, 2008
Published January 10, 2009

Note that (8) resembles the full band algorithm derived in [4, Eq. (19)], [5, Eq. (6.29)] except for the fact that the subband error powers are normalized by the corresponding subband input signal powers and averaged over all subbands. The matrix $\boldsymbol{\phi}(n)$ is a diagonal matrix for white inputs. Hence, for white inputs, the methods of DC-VSS-SAP and DC-MDA-SAP are considered equivalent. However, for colored signals, DC-MDA-SAP is only an approximation of DC-VSS-SAP. For large number of subbands, the signals $x_{ij}(n)$ can be considered to be white in nature and both the above methods should yield similar performance as is confirmed later by simulations.

4 Detection of changes in echo paths

Changes in echo paths happen quite frequently in an AEC system. The user could increase/decrease/mute the speaker volume and there could be people moving around in the room resulting in changing acoustic echo paths. A changing echo path increases the error signal $e_i(n)$ which results in very low values of $\mu(n)$ computed by the method of DC-VSS-SAP. The low values of the step size practically freezes the AF (no adaptation), resulting in very slow convergence. This observation, which is associated with the method of DC, was reported in [4, Section 4.2.1]. It therefore becomes necessary to quickly detect changes in the echo path and reset the AF. All step-size computations are started afresh and quick reconvergence is assured.

MSD increases during changes in echo path. Since $\hat{D}(n)$ is an estimate of MSD, it is expected to increase during echo path changes. Changes in echo paths can be detected using low pass filtering (LPF) or smoothing of $\hat{D}(n)$. The LPF or smoothed output is computed as $\bar{D}(n) = 0.9995\bar{D}(n-1)+0.0005\hat{D}(n)$. If $\bar{D}(.)$ increases monotonically for more than 20 continuous samples, we report a change in echo path. This way, the method of delay coefficients can be used as a change in echo path detector (CED). Once an echo change is detected the AF is reset to enable quick reconvergence to the new echo path. Simulation results are reported in Section V.

5 Results and discussions

All input signals are speech signals with a sampling rate of 16 kHz. The default values of $(M,P) = (2,4)$, SNR $= 20$ dB and $\alpha = 0.9$ are used. The other values used are $N = 632$, $b = 120$, $\mathbf{s}^{*T}\mathbf{s}^{*} = -0.42$ dB. We use misalignment (in dB), $10\log(\tilde{\mathbf{S}}^{T}(n)\tilde{\mathbf{S}}(n))$, as a performance measure.

Fig. 1 compares the performance between SAP [2] and the proposed DC-VSS-SAP. The SAP [2] algorithm uses a fixed step size of $\mu(n) = 0.6$, i.e., it uses no VSS algorithm. Recall that the proposed DC-VSS-SAP algorithm is the delay coefficients (DC) based variable step size (VSS) method used in conjunction with SAP [2]. It can be seen that the SAP [2] diverges during the silence periods of the far-end signal and during the presence of a near-end signal while the proposed DC-VSS-SAP ensures that the AF does not diverge. The observation is in line with the reasonings given in Section III.

DOI: 10.1587/elex.6.20
Received October 31, 2008
Accepted November 21, 2008
Published January 10, 2009

ELectronics EXpress

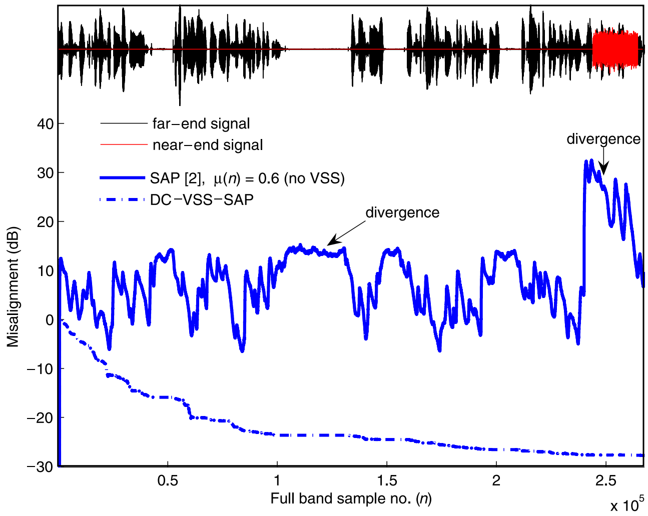

Fig. 1. Comparison between DC-VSS-SAP and SAP [2].

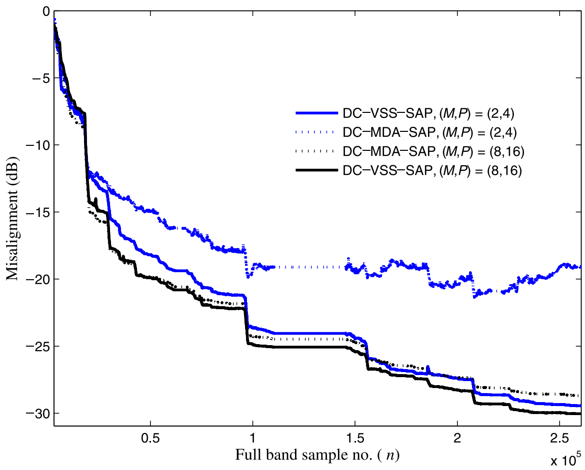

Fig. 2. Performance comparison between DC-VSS-SAP and DC-MDA-SAP algorithms. Input signals are speech signals with silence periods.

Fig. 2 is a comparison between DC-VSS-SAP and DC-MDA-SAP. It can be seen that for a large value of $M = 16$, both methods are equivalent and have the same performance, whereas, for $M = 2$, the DC-MDA-SAP has a loss in performance compared to DC-SAP. This is in accordance with the analysis in Section III.

Fig. 3 denotes results pertaining to the use of a change in echo path detector (CED). The echo path is abruptly changed at sample number 1e5. Use of a CED results in timely detection of echo path change and quick reconvergence compared to the case when CED is not used.

6 Conclusions

A delay coefficients-based variable step size algorithm was proposed for a recently published subband affine projection adaptive filter for use in AEC applications. The motivations for the use of a VSS algorithm was discussed.

DOI: 10.1587/elex.6.20
Received October 31, 2008
Accepted November 21, 2008
Published January 10, 2009

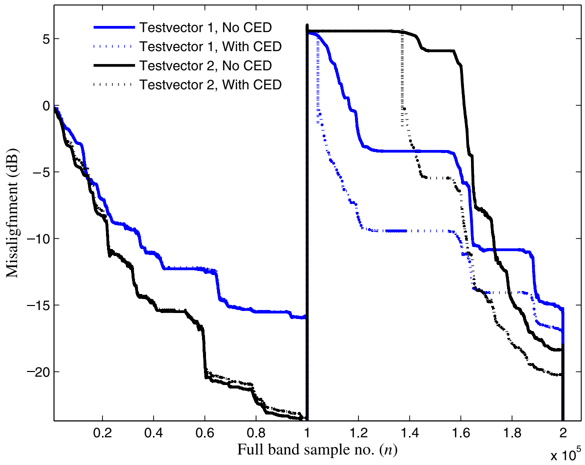

Fig. 3. Performance of DC-VSS-SAP in changing echo path conditions with and without CED. When CED is used, AF is reset on detection of an echo path change. The system is suddenly changed from $\mathbf{s}^*$ to $-\mathbf{s}^*$ at fullband sample number 1e5.

The VSS algorithm ensures the AF does not diverge by computing the optimum step size at each time instant. Differences and similarities between the VSS fullband and subband algorithms were mentioned. It was shown that for white inputs, the VSS subband algorithm assumes a form similar in nature to the fullband VSS algorithm. A change in echo path detector was proposed based on the method of DC which improves the convergence rate of the filter whenever there are changes in echo paths.

Acknowledgments

The first, second and fourth authors thank STMicroelectronics Asia Pacific Pte. Ltd. for permission to publish this work.

DOI: 10.1587/elex.6.20
Received October 31, 2008
Accepted November 21, 2008
Published January 10, 2009