A VARIABLE STEP SIZE EVOLUTIONARY AFFINE PROJECTION ALGORITHM

Felix Albu[1], Constantin Paleologu[1], and Jacob Benesty[2]

[1]Faculty of Electronics and Telecommunications, "Politehnica" University of Bucharest
1-3 Iuliu Maniu, Bucharest, Romania, email: {felix, pale}@comm.pub.ro
[2]INRS-EMT, University of Quebec
QC H5A 1K6, Montreal, Canada, email: benesty@emt.inrs.ca

ABSTRACT

It is well known that the affine projection algorithm (APA) offers a good tradeoff between convergence rate/tracking and computational complexity. Recently, the evolutionary APA (E-APA) with a variable projection order has been proposed. In this paper, we propose a variable step size (VSS) version of the E-APA, called VSS-E-APA. It is shown that the VSS-E-APA is robust to near-end signal variations. Also, it has both a fast convergence speed and a small steady-state error and a much reduced numerical complexity than the VSS-APA.

Index Terms— adaptive filters, affine projection algorithm, variable step size, echo cancellation.

1. INTRODUCTION

In echo cancellation systems, an adaptive filter is used to reduce the echo. The well-known normalized least-mean-square (NLMS) algorithm has been widely used in this context. Nevertheless, it converges slowly for acoustic echo cancellation (AEC) applications, where long length adaptive filters are used in order to model the acoustic echo paths. The affine projection algorithm (APA) [1] can be considered as a generalization of the NLMS algorithm that provides an improved convergence speed, especially for highly correlated signals, such as speech. In terms of convergence rate, the APA has a performance that rivals with the more complex recursive least-squares (RLS) algorithm in most practical situations. Many fast affine projection (FAP) algorithms have been proposed for acoustic echo cancellation systems (e.g., based on embedded fast RLS algorithm [2], Gauss-Seidel iterations [3], dichotomous coordinate descent (DCD) iterations [4], displacement structure theory [5], etc.). All previously mentioned FAP versions use a fixed projection order. It is known that if the projection order increases, the convergence speed is faster, but the steady-state error also increases. A variable projection order might lead to a lower steady-state error. In [6], an affine projection algorithm with an evolving projection order, called evolutionary APA (E-APA) was proposed. Based on findings from [7-8], the projection order was modified depending on the relationship between the output error and a threshold. However, the authors do not investigate practical implementations with reduced numerical complexity. Also, the E-APA uses a fixed step size. A variable step size (VSS) version could be a more reliable solution in case of near-end signal variations, including double-talk. Such a solution has been presented in [9] and the VSS-APA algorithm was proposed. It was shown that the VSS-APA is more robust in adverse conditions than the APA. Also, a computationally simpler version based on DCD, called VSS-DCD-AP was proposed in [10]. In this paper the same step size computational method is adapted to the less computationally demanding E-APA. A new algorithm, called variable step size evolutionary APA (VSS-E-APA) is proposed and compared with previously published E-APA and VSS-APA.

The outline of the paper is as follows. The proposed evolutionary APA using a VSS scheme (VSS-E-APA) is described in Section 2. In Section 3, the behavior of this algorithm for AEC systems is examined and compared to the VSS-APA and E-APA. Section 4 concludes this work.

2. VSS-E-APA

In the AEC configuration, the far-end signal, $x(n)$, goes through the echo path $\mathbf{h}$, providing the echo signal, $y(n)$. This signal is added to the near-end signal, $\nu(n)$ (which can contain both the background noise and the near-end speech), resulting the microphone signal, $d(n)$. The adaptive filter, defined by the vector $\hat{\mathbf{h}}(n)$, aims to produce at its output an estimate of the echo, $\hat{y}(n)$, while the error signal, $e(n)$, should contain an estimate of the near-end signal.

Through this paper, the following notation will be used: L is the length of the adaptive filter, K_n is the projection order of the E-APA at iteration n, $K_{\max}$ is the maximum projection order, δ is the regularization parameter, $\boldsymbol{\mu}_{K_n}(n) = diag\left[\mu_0(n), \mu_1(n), ..., \mu_{K_n-1}(n)\right]$ is the variable step-size diagonal matrix whose dimension varies with K_n, and its maximum size is $K_{\max}$, τ_0 a small constant to avoid division by zero, $\mathbf{I}_{K_n}$ denotes the $K_n \times K_n$ identity matrix, at time instant n, $\mathbf{d}(n)=\left[d(n)\ d(n-1)\ldots d(n-K_n+1)\right]^T$ is the desired vector, $\mathbf{y}(n)=\left[\hat{y}(n)\ \hat{y}(n-1)\ldots \hat{y}(n-K_n+1)\right]^T$ is the filter output vector, superscript T denotes transposition, and $\mathbf{x}(n)=\left[x(n), x(n-1),\ldots, x(n-L+1)\right]^T$ is the input signal vector. Also, we have the data matrix $\mathbf{X}(n)=\left[\mathbf{x}(n)\ \mathbf{x}(n-1)\ldots \mathbf{x}(n-K_n+1)\right]$, $\mathbf{R}(n)$ is the $K_n \times K_n$ auto-correlation matrix of the input signal, i.e., $\mathbf{R}(n)=\mathbf{X}^T(n)\mathbf{X}(n)+\delta\mathbf{I}_{K_n}$, the adaptive filter coefficients vector is $\hat{\mathbf{h}}(n)=\left[\hat{h}_0(n),\ldots,\hat{h}_{L-1}(n)\right]^T$, and $\xi(n)=\left[x(n), x(n-1),\ldots, x(n-K_{\max}+1)\right]^T$ is a $K_{\max}\times 1$ vector. Also, $\mathbf{r}(n)$ is a $K_{\max}\times 1$ autocorrelation vector, $\mathbf{p}(n)$ is a $K_n \times 1$ solution vector, and finally, $\mathbf{e}(n)$ is a $K_n \times 1$ vector.

The computation of the matrix $\mathbf{R}(n)$ can be made in an efficient way taking into account its symmetry. However, the size of this square matrix varies depending on the chosen projection order. There are two possible situations. In the first situation, if the projection order at time n is smaller than or equal to the projection order at time $n-1$, the matrix $\mathbf{R}(n)$ is updated by replacing the first row and column with the elements of $\mathbf{r}(n)$, while the bottom-right $(K_n-1)\times(K_n-1)$ sub-matrix is replaced with the top-left $(K_n-1)\times(K_n-1)$ sub-matrix of $\mathbf{R}(n-1)$. In the second situation, if $K_n > K_{n-1}$, the matrix $\mathbf{R}(n)$ is updated by replacing the first row and column with the elements of $\mathbf{r}(n)$, while the other elements are replaced with $\mathbf{R}(n-1)$. The update of $\mathbf{r}(n)$ is made on full $K_{\max}\times 1$ vector, although only $K_n \times 1$ vector is needed for updating $\mathbf{R}(n)$. The equations that define the proposed VSS-E-APA are summarized in the following table. The noise variance, σ_v^2, can be estimated online as in [9-11], or during the periods of silence, and the forgetting factor λ is evaluated as in [11]. The equations (11)—(14) from the Table 1 are introduced from the VSS-APA. However, the projection order can vary and therefore $\mathbf{e}(n)$ cannot be used to compute all $e_{l+1}^2(n)$ elements (especially for $l=K_n..K_{\max}-1$ if $K_n < K_{\max}$). An approximation is made by using previous $e_{l+1}^2(n-1)$ values [see Eq. (10) from Table 1].

Initialization: (1)

$\mathbf{x}(0)=\mathbf{0}_{L\times 1},\ \hat{\mathbf{h}}(0)=\mathbf{0}_{L\times 1},$

$K(0)=K_{\max}, \hat{\sigma}_d^2(0)=0, \hat{\sigma}_{\hat{y}}^2(0)=0,$

$C_1=\mu\sigma_v^2/(2-\mu), C_2=2\sigma_v^2/(2-\mu)$

$\mathbf{r}(0)=\mathbf{0}_{K_{\max}\times 1}, \xi(0)=\mathbf{0}_{K_{\max}\times 1}$

for $l=0$ to $K_{\max}-1$, $\hat{\sigma}_{e_{l+1}}^2(0)=0$

For time index $n=1,2,\ldots$

$$\mathbf{r}(n)=\mathbf{r}(n-1)+x(n)\xi(n)-x(n-L)\xi(n-L) \quad (2)$$

$$\mathbf{y}(n)=\mathbf{X}^T(n)\hat{\mathbf{h}}(n-1) \quad (3)$$

$$\mathbf{e}(n)=\mathbf{d}(n)-\mathbf{y}(n) \quad (4)$$

$$\eta_n=C_1K_{n-1}+C_2 \quad (5)$$

$$\theta_n=\eta_n-C_1 \quad (6)$$

IF $e^2(n)>\eta_n$

$$K_n=\min\{K_{n-1}+1, K_{\max}\} \quad (7)$$

ELSE IF $e^2(n)\le\theta_n$

$$K_n=\max\{K_{n-1}-1, 1\} \quad (8)$$

ELSE

$$K_n=K_{n-1} \quad (9)$$

IF $K_n<K_{\max}$

$$e_{K_n}(n)=e_{K_n-1}(n-1)\ldots \quad (10)$$

$$e_{K_{\max}-1}(n)=e_{K_{\max}-2}(n-1)$$

END

$$\hat{\sigma}_d^2(n)=\lambda\hat{\sigma}_d^2(n-1)+(1-\lambda)d^2(n) \quad (11)$$

$$\hat{\sigma}_{\hat{y}}^2(n)=\lambda\hat{\sigma}_{\hat{y}}^2(n-1)+(1-\lambda)\hat{y}^2(n) \quad (12)$$

For $l=0$ to $K_{\max}-1$

$$\hat{\sigma}_{e,l+1}^2(n)=\lambda\hat{\sigma}_{e,l+1}^2(n-1)+(1-\lambda)e_{l+1}^2(n) \quad (13)$$

For $l=0$ to K_n-1 (14)

$$\mu_l(n)=\left|1-\frac{\sqrt{\left|\hat{\sigma}_d^2(n-l)-\hat{\sigma}_{\hat{y}}^2(n-l)\right|}}{\tau_0+\hat{\sigma}_{e,l+1}(n)}\right|$$

Update $\mathbf{R}(n)$ using $\mathbf{r}(n)$ (15)

$$\mathbf{E}(n)=\boldsymbol{\mu}_{K_n}(n)\mathbf{e}(n) \tag{16}$$

$$\text{Solve } \mathbf{R}(n)\mathbf{p}(n)=\mathbf{E}(n) \tag{17}$$

$$\hat{\mathbf{h}}(n)=\hat{\mathbf{h}}(n-1)+\mathbf{X}(n)\mathbf{p}(n) \tag{18}$$

TABLE 1. VSS-E-APA Equations.

It will be shown in the next section that the effect of the approximation on convergence performance is not important.

The VSS-APA has $(2L+6)K_{\max}+6+O(K_{\max}^3)$ multiplications (the projection order is $K_{\max}$) [4] and the E-APA with the presented matrix update procedure has $(2L+3)K_n+O(K_n^3)$. In our implementation, for specific algorithm iteration where the chosen projection order is K_n, the VSS-E-APA has $(2L+1)K_n+3K_{\max}+8+O(K_n^3)$ multiplications, and the contribution of the VSS part is $3K_{\max}+6$ multiplications. It is obvious that the VSS part is only a very small fraction of the complexity of the VSS-APA and VSS-E-APA because $L \gg K_{\max}$ in typical AEC systems.

3. SIMULATIONS

Simulations were performed in an AEC context and VSS-APA, E-APA, and VSS-E-APA were compared. The measured impulse response of the acoustic echo path is plotted in Fig. 1a (the sampling rate is 8 kHz); its entire length has 1024 coefficients. This length is truncated to the first 512 coefficients (exact modeling case). An independent white Gaussian noise signal is added to the echo signal $y(n)$, with 30 dB signal-to-noise ratio (SNR) for most of the experiments and $\mu=0.2$. The maximum considered value of the projection order for all simulations is $K_{\max}=8$. In the following experiments, the speech sequence from Fig. 1b is used as the far-end signal and the signal from Fig. 1c is used as a near-end signal. The performance for the exact modeling scenario is evaluated in terms of the normalized misalignment (in dB), defined as $20\log_{10}\left[\left\|\mathbf{h}-\hat{\mathbf{h}}(n)\right\|/\left\|\mathbf{h}\right\|\right]$, where $\|\bullet\|$ denotes the l_2 norm.

Fig. 2a shows the misalignment curves in case of a single-talk scenario where a sudden change in the echo path after 21 seconds is simulated. In terms of the final misalignment, it can be seen that the variable projection order versions have a very close performance to that of VSS-APA. Initially, in the convergence phase, most of projection orders are chosen closer to the maximum allowed projection order, while in the steady state phase the projection orders are close to the minimum allowed projection order (Fig. 2b). The reduction in complexity of the evolutionary APAs is important (an average of 2200 multiplications for the E-APA and 2150 multiplications for the VSS-E-APA as opposed to 8400 for the VSS-APA). That corresponds to around 75% complexity reduction of VSS-E-APA in comparison with VSS-APA.

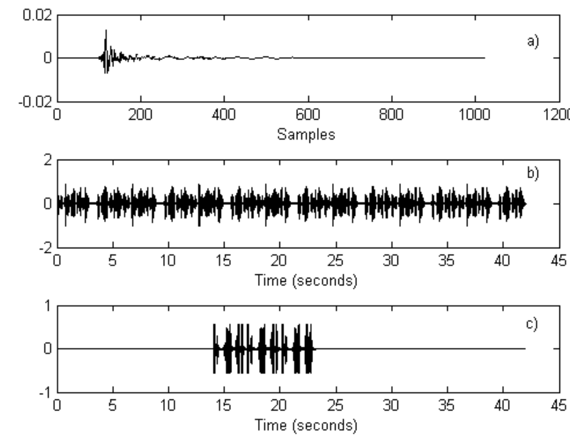

Fig. 1. a) Measured room acoustic impulse response. b) Far-end speech signal used in the experiments. c) Near-end speech signal used in the last experiment.

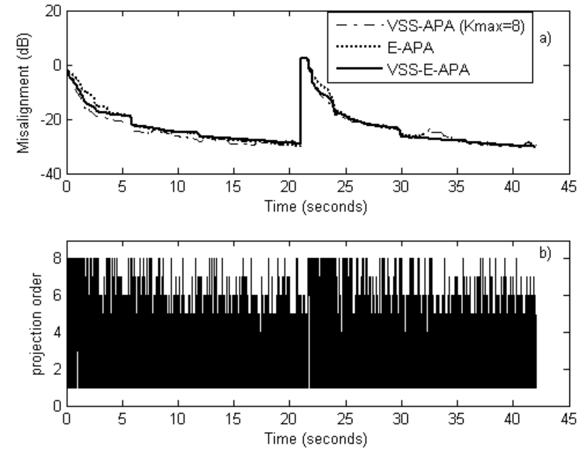

Fig. 2. a) Misalignment of the VSS-APA (p=8), E-APA, and VSS-E-APA. Single-talk case, L = 512, and SNR = 30dB. b) The projection orders chosen by the VSS-E-APA.

In Fig. 3, the investigated algorithms are used in a double-talk situation in conjunction with a Geigel DTD. Its settings were chosen as in [11]. It can be noticed from Fig. 4a that the robustness to double talk situations of the VSS-E-APA and VSS-APA is much better than that of E-APA. The E-APA is affected, even if it uses the Geigel DTD. It can be seen from Fig. 3b that the presence of near-end signals leads to the choice of many higher projection orders K_n in case of E-APA towards the end of the speech sequence. This fact doesn’t happen with VSS-E-APA that has mostly lower projection orders towards the end of the speech sequence (Fig. 3c).

Fig.4 shows the number of occurrences in the previously investigated cases. It can be noticed in Fig. 4a that most of K_n are 1 or 2 in the single-talk case.

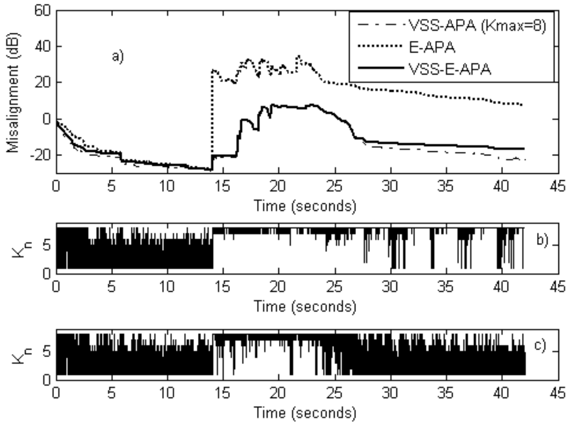

Fig. 3. a) Misalignment of the VSS-APA (p=8), E-APA, and VSS-E-APA. Double-talk case, exact-modelling case, L = 512, SNR = 30dB, and Geigel DTD is used. b) The projection orders chosen by the E-APA. c) The projection orders chosen by the VSS-E-APA.

The VSS-E-APA reduced complexity advantage over E-APA in the double-talk case is proved by the distribution of the chosen projection order shown in Fig. 4c and Fig. 4b respectively. The number of $K_n = 1$ and $K_n = 2$ in case of VSS-E-APA overcomes the number of occurrences of $K_n = 8$. In the E-APA case, there are twice $K_n = 8$ than the sum of all the other projection orders.

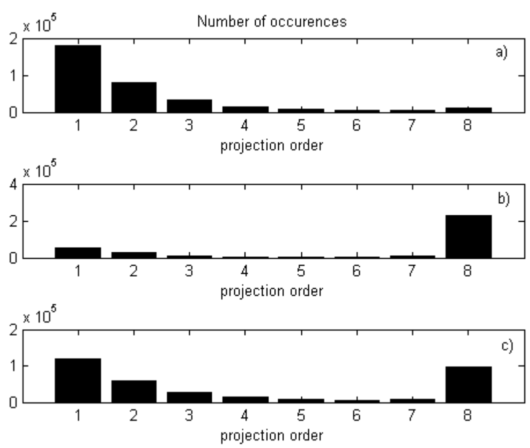

Fig. 4 The number of occurrences of each projection order between 1 and 8 for: a) VSS-E-APA in the single talk case. b) E-APA in the double-talk case c) VSS-E-APA in the double-talk case

The average number of multiplications is 6190 for the E-APA and only 3600 for the VSS-E-APA. Therefore, the complexity reduction over VSS-APA is about 26% for E-APA and more than twice for VSS-E-APA (67%). Also, VSS-E-APA is about 40% less complex than E-APA. Similar conclusions were obtained in the background noise variation and under-modeling case (not shown here due to the lack of space). In these cases, the computational savings of VSS-E-APA are significant and can reach 50% over E-APA and even 80% in comparison with VSS-APA.

4. CONCLUSIONS

The VSS-E-APA has been proposed for AEC. A variable step size was used in order to take into account the existence and the non-stationarity of the near-end signals. Simulation results have shown the superiority of the VSS-E-APA over the E-APA to near-end signal variations. The VSS-E-APA can be combined with a simple Geigel DTD in order to enhance its performance. The proposed algorithm leads to a more efficient implementation and significant computational savings in comparison with the VSS-APA and E-APA.

Acknowledgment: This work was supported by the UEFISCSU grant PN-II-RU-TE no. 7/5.08.2010

5. REFERENCES

[1] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Electronics and Communications in Japan*, vol. 67-A, no. 5, pp. 19—27, May 1984.

[2] S. L. Gay "A fast converging, low complexity adaptive filtering algorithm," in *Proc. IWAENC*, 1993, pp. 223—226.

[3] F. Albu, J. Kadlec, N. Coleman, and A. Fagan, "The Gauss-Seidel fast affine projection algorithm," in *Proc. IEEE SIPS*, 2002, pp. 109—114.

[4] Y. Zakharov, F. Albu, "Coordinate descent iterations in fast affine projection algorithm," *IEEE Signal Processing Letters*, Vol. 12, No. 5, pp. 353—356, May 2005.

[5] M. C. Tsakiris and P. A. Naylor, "Fast exact affine projection algorithm using displacement structure theory," in *Proc. DSP* 2009, pp. 69—74.

[6] S.-E. Kim, S.-J. Kong, and W.-J. Song, "An affine projection algorithm with evolving order," *IEEE Signal Processing Lett.*, vol. 16, no. 11, pp. 937—940, Nov. 2009.

[7] H.C. Shin and A. H. Sayed, "Mean-square performance of a family of affine projection algorithms," *IEEE Trans. Signal Processing*, vol. 52, no. 1, pp. 90—102, Jan. 2004.

[8] N. R. Yousef and A. H. Sayed, "A unified approach to the steady-state and tracking analysis of adaptive filters," *IEEE Signal Processing Lett.*, vol. 11, pp. 132—135, Feb. 2004.

[9] C. Paleologu, J. Benesty, and S. Ciochină, "A variable step-size affine projection algorithm designed for acoustic echo cancellation," *IEEE Trans. Audio, Speech, and Language Processing*, vol. 16, pp. 597—600, Nov. 2008.

[10] F. Albu, C. Paleologu, J. Benesty, and Y. V. Zakharov, "Variable step size dichotomous coordinate descent affine projection algorithm," in *Proc. IEEE EUROCON*, 2009, pp. 1366-1371.

[11] J. Benesty, H. Rey, L. R. Vega, and S. Tessens, "A nonparametric VSS NLMS algorithm," *IEEE Signal Process. Lett.*, vol. 13, no. 10, pp. 581—584, Oct. 2006.

The codes for the proposed algorithms can be obtained from
http://falbu.50webs.com/List_of_publications_aec.htm

The reference for the paper is:
F. Albu, C. Paleologu, J. Benesty, "A Variable Step Size Evolutionary Affine Projection Algorithm", ICASSP 2011, Prague, Czech Republic, May 2011, pp. 429-432