ACOUSTIC ECHO CANCELLATION IN THE PRESENCE OF DISTORTING LOUDSPEAKERS

Aurelio Uncini, Andrea Nalin and Raffaele Parisi

INFOCOM Dept. - University of Rome "La Sapienza"
Via Eudossiana 18, 00184 Rome – Italy
email: *aurel@ieee.org*

ABSTRACT

Acoustic echo cancellation is usually needed in hands-free telephone systems or in video conference applications. Conventional acoustic echo cancellers (AECs) basically rely on the assumption of a linear echo path. Their performance is then limited in the presence of nonlinearities, like those typically generated in low-quality loudspeakers.

The presence of nonlinearities require the adoption of nonlinear AECs. Truncated Volterra filters are usually applied to nonlinear echo cancellation, but the improvements obtained with respect to linear adaptive filters do not justify a computational load of at least $O(N^2)$ (like in the case of second-order Volterra filters) and simpler approaches are necessary.

In this paper the effect of loudspeaker nonlinearities is first investigated. Then a convenient new nonlinear architecture based on spline adaptive functions with computational cost of $O(N)$ is proposed as an effective alternative to computationally more expensive approaches.

1. INTRODUCTION

The problem of acoustic echo arises when a loudspeaker and a microphone operate in the same environment, such that the microphone picks up the signal radiated by the loudspeaker and its reflections on the borders of the enclosure.

Figure 1 shows the architecture of a typical acoustic echo canceller (AEC), which is based on an adaptive filter placed in parallel to the loudspeaker-enclosure-microphone system (LEM). The adaptive filter should match the impulse response of the LEM system in order to cancel out its effect from the transmitted signal. Typically, linear filters are employed.

However, the presence of nonlinearities in the LEM system can make the performance of linear filtering schemes inadequate [5]. A typical case is the presence of low-quality loudspeakers, that can introduce serious distortions in the acquired signal [5][7]. In these situations, nonlinear AECs constitute an attractive alternative to more conventional architectures [6]. At the same time they may require a higher computational cost that make them ineffective in practical applications.

Due to their local adaptation characteristics, splines demonstrated particularly effective in the design of non-linear adaptive systems [1][8][9]-[[11]. In particular, adaptive splines are able to guarantee the flexibility and generalization capabilities required, with a reduced computational overhead. For these reasons, they appear particularly suited for the design of nonlinear AECs.

The paper is organized as follows. In section 2 the reduction of performance of conventional linear algorithms in the presence of loudspeaker nonlinearities is illustrated with the use of real signals. In section 3 the proposed spline-based nonlinear filter is presented and in section 4 its computational complexity discussed. Finally, in section 5 experimental results are shown, in which the proposed structure is compared both to linear normalized LMS (NLMS) filter and to second-order Volterra non linear filter [2][3].

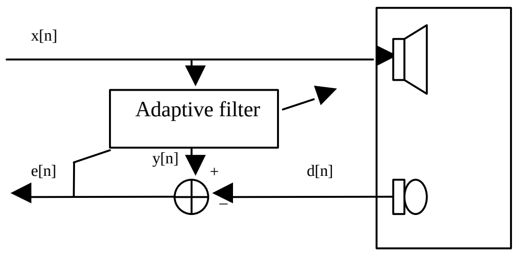

Fig 1. Architecture of a typical echo canceller

2. LOUDSPEAKER EFFECT ON THE PERFORMANCE OF LINEAR AECs

An index generally used to evaluate the performance of echo cancellers is the echo return loss enhancement (ERLE), which is expressed by:

$$ERLE = 10\log_{10}\left[\frac{E\{d^2[n]\}}{E\{e^2[n]\}}\right], \tag{1}$$

where notation of fig. 1 has been used.

In order to illustrate the reduction in the achievable ERLE of linear AECs due to the presence of nonlinearities, we have compared the experimental results of echo cancellation in a typical teleconference setup, using two different loudspeakers: a professional HI-FI loudspeaker and a common low-cost loudspeaker (like those used for PC workstations). Figure 2 shows the difference between the two speakers, in terms of introduced distorsions, for a 80 Hz sine.

The linear NLMS algorithm is the most commercially used optimization method in echo cancellation and is the baseline by which performance of alternative models is measured. Experimental results in two different cases are shown in fig. 3, referring respectively to white gaussian noise and male human voice. It is clear the difficulty of the linear approach (NLMS) in cancelling distorsions introduced by the inexpensive speaker.

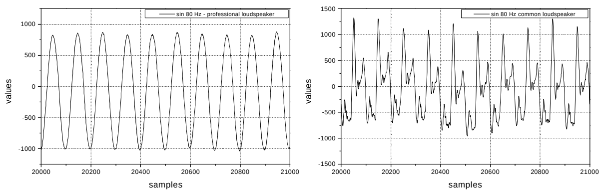

Fig. 2 80 Hz sine acquired using a professional loudspeaker (left) and a low-cost one (right)

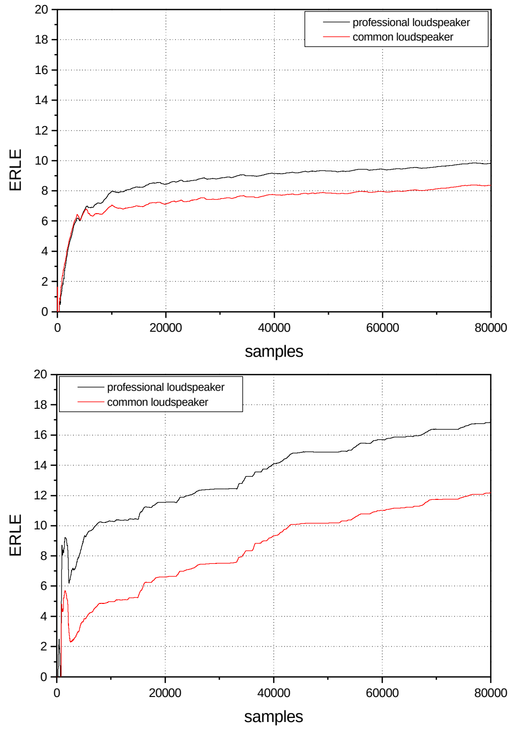

Fig. 3 Measured ERLE for a white gaussian noise and a male voice

3. SPLINE-BASED NON LINEAR ECHO CANCELLER

The architecture that we propose is composed by a FIR filter with a nonlinear function at its output. This structure can be viewed as a neural network with a single neuron. In order to increase the learning capabilities of this neuron, we consider an adaptive activation function, i.e. a function whose shape can be opportunely modified during the signal processing. In particular it has been showed [8][11] that adaptive spline activation functions are characterized by good generalization capabilities and learning speed. In addition, adaptive splines yield a reduction of the number of free parameters, still mantaining its adaptive capabilities.

The proposed nonlinearity is based on a LUT (Look Up Table), whose control points are adapted in order to optimize a specified cost function. The output is obtained by a proper interpolation scheme. More specifically, the type of interpolation adopted must guarantee a continuous first derivative in order to allow a gradient-based learning algorithm.

In order to take into account these requirements, the piece-wise polynomial spline interpolation scheme has been chosen [11]. In the following, we briefly recall the main features of adaptive splines and furnish a description of the learning strategy.

3.1 Adaptive splines

Let $h(x)$ be a general nonlinear function. Spline approximation consists in subdividing $h(x)$ in multiple tracts (*spans*), each one being locally approximated by a spline curve:

$$y = h(x) = \bar{h}(u,i) \tag{2}$$

Spline approximation requires a number of *control points* q_i and a local variable $u \in [0,1)$ for each span. In general, N+1 control points ($q_{x,0} < q_{x,1} < ... < q_{x,N}$) are considered, spaced by a uniform step size Δx. The output of the activation function is then obtained through sequential application of the following equations:

$$z = \frac{s}{\Delta x} + \frac{N-2}{2}, \tag{3}$$

$$z = \begin{cases} 1 & z < 1 \\ z & 1 \le z \le N-3 \\ N-3 & z > N-3 \end{cases}, \tag{4}$$

$$i = \lfloor z \rfloor, \tag{5}$$

$$u = z - i, \tag{6}$$

$$F_{yi}(u) = \sum_{j=0}^{3} q_{y,i+j} C_j(u), \tag{7}$$

where i is the index of the curve span of the function needed for the computation of the output, $C_j(u)$ are the spline basis functions and $F_{yi}(u)$ is the output.

Figure 4 shows two kinds of cubic splines (i.e. controlled by four control points), namely the Catmul-Rom and the B-spline, corresponding to two different interpolation schemes [11]. The B-spline will be used in the following.

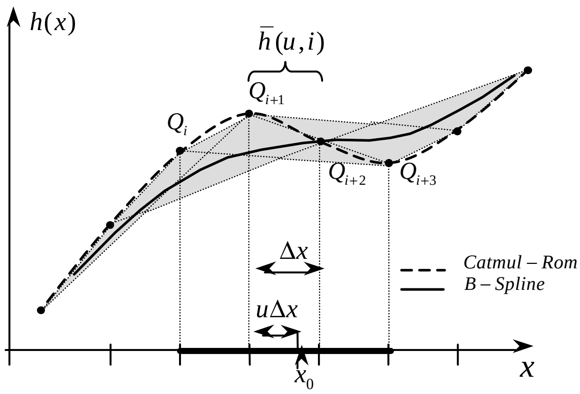

Fig. 4 Cubic-Spline interpolation of control points.

3.2 Learning algorithm

We have developed an LMS-type adaptation process. Considering the squared value of the istantaneous error $e[n]$ as a cost function, the filter coefficients are updated by the following formula:

$$\mathbf{w_{n+1}} = \mathbf{w_n} + m e[n]\frac{\partial f(s)}{\partial s}\mathbf{x_n}\ , \qquad (8)$$

where

$$\frac{\partial f(s)}{\partial s} = \frac{\partial f(u)}{\partial u}\frac{\partial u}{\partial s} = \frac{\partial f(u)}{\partial u}\frac{1}{\Delta x} = \frac{\partial F_{yi}(u)}{\partial u}\frac{1}{\Delta x} = (\sum_{j=0}^{3}\frac{\partial C_j(u)}{\partial u}q_{y,i+j})\frac{1}{\Delta x} \qquad (9)$$

Concerning the adaptation of the control points, we have:

$$q_{y,i+j}[n+1] = q_{y,i+j}[n] - m_q e[n]\frac{\partial e[n]}{\partial q_{y,i+j}[n]}\ , \qquad 0 \le j \le 3 \qquad (10)$$

where

$$\frac{\partial e[n]}{\partial q_{y,i+j}[n]} = -\frac{\partial F_{yi}(u)}{\partial q_{y,i+j}[n]} = -\frac{\partial}{\partial q_{y,i+j}[n]}\left(\sum_{k=0}^{3} q_{y,i+k}[n]C_k(u)\right) = -C_j(u)$$

$0 \le j \le 3$

Referring to [4], we also introduce the following normalization of the learning step-size of the coefficients:

$$m(n) = \frac{\tilde{m}}{\sum_{j=0}^{N}\left(\frac{\partial f(s)}{\partial s}x[n-j]\right)^2} = \frac{\tilde{m}}{\sum_{j=0}^{N}\left((\sum_{j=0}^{3}\frac{\partial C_j(u)}{\partial u}q_{y,i+j})\frac{1}{\Delta x}x[n-j]\right)^2} \qquad (11)$$

This choice for m allows to obtain for the non linear case the same improvements that NLMS yields with respect to the LMS algorithm in a linear framework.

4. COMPUTATIONAL COMPLEXITY

An interesting feature of the proposed approach is its computational cost in terms of MOPS (multiplications per sample). In fact, its order is $O(N)$ and it is then comparable with that of the linear NLMS filter.

The computational overhead required to compute the output of the adaptive non linear function and to adapt the control points is a fixed cost, which can be easily computed from equations (2) to (11) and depends on the chosen spline basis functions. In the case of the cubic B-Spline used in our simulations, this fixed cost has been estimated to be of 30 multiplications. As a consequence, the total number of multiplications required is $3N+30$ MOPS. However, this overhead with respect to NLMS filtering is substantially negligible for typical values of N.

5. EXPERIMENTAL RESULTS

The presented system was tested using a white Gaussian noise as excitation. A commercial loudspeaker (the same used in section 2) and a microphone were placed in an enclosure with low reverberation level, whose acoustic impulse response had a length estimated in 1024 coefficients. The echo acquired with the microphone was used for the simulations.

In fig. 5 it is shown the ERLE measured in the case of the proposed structure using the cubic B-Spline basis function, $\Delta x = 0.2$ and step-size values $\tilde{m} = 1$ e $\tilde{m}_q = 0.05$. For comparison purposes, the ERLE obtained with the linear NLMS filter and with non linear second order truncated Volterra filter [] with $N_2 = 10$ and $N_2 = 20$ are also shown. This experiment confirmed the better performance of the spline-based approach.

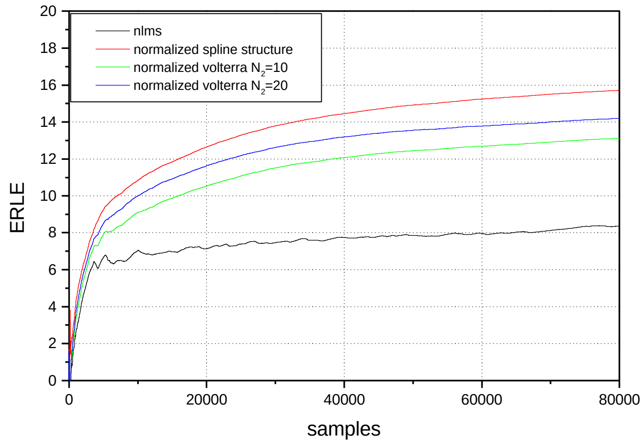

Fig.5 Measured ERLE for a white gaussian noise

REFERENCES

[1] V. J. Mathews, "Adaptive Polynomial Filters", *IEEE Signal Processing Magazine,* 8(3), pp. 10-26, July 1991.

[2] Stenger, L. Trautmann, R. Rabenstein, "Nonlinear Acoustic Echo Cancellation With 2nd Order Adaptive Volterra Filters", *ICASSP'99*, USA, 1999.

[3] J. Chen, J.Vandewalle, "Study of Adaptive Nonlinear Echo Canceller with Volterra Expansion", *ICASSP'89*, vol.2, pp. 1376 -1379.

[4] S. Kalluri, G.A. Arce, "A General Class of Nonlinear Normalized Adaptive Filtering Algorithms", *IEEE Transaction on Signal Processing*, vol. 47, NO. 8, August 1999.

[5] N. Birkett, R. A. Goubran, "Limitations of Handsfree Acoustic Echo Cancellers Due To Nonlinear Loudspeaker Distorsion and Enclosure Vibration Effects", *1995 IEEE ASSP Workshp on Appl. Of Signal Processing to Aud. And Acoustics*, New Paltz, New York, October 1995.

[6] N. Birkett, R. A. Goubran, "Acoustic Echo Cancellation using NLMS-Neural Network Structures", *Proceedings ICASSP95*, pp.3035-3038.

[7] N. Birkett, R. A. Goubran," Non-linear Loudspeaker Compensation for Hands-free Acoustic Echo Cancellation", *IEE Electronic Letters*, vol. 32, NO. 12, pp. 1063-1064, June 1996.

[8] C.T. Chang, W.D. Chang, "A Feedforward Neural Network with Adaptive Spline-based Activation Finctions", *Proc. of International Neural Network Society Annual Meeting WCNN95*, Washington D.C., USA, pp. 1695-1699.

[9] F. Piazza, A. Uncini, M. Zenobi, "Artificial Neural Networks with Adaptive Polynomial Activation Function", *Proc. of the IJCNN*, vol. 2, pp. 343-349, Bejing, China, 1992.

[10] L. Vecci, P. Campolucci, F. Piazza, A. Uncini, "Approximation Capabilities of Adaptive Spline Neural Networks", *Proceedings ICNN'97*, Houston, Texas USA, June 1997.

[11] S. Guarnieri, F. Piazza, A. Uncini, "Multilayer Feedfoward Networks with Adaptive Spline Activation Function", *IEEE Transaction on Neural Networks*, vol. 10, NO. 3, May 1999.