ACOUSTIC ECHO CANCELLATION USING NONLINEAR CASCADE FILTERS

J.-P. Costa, A. Lagrange and A. Arliaud

IUP GMI, University of Avignon
339 ch. des Meinajaries BP 1228 - 84911 Avignon FRANCE Cedex 9

ABSTRACT

The miniaturization of GSM handsets creates nonlinear acoustical echoes between microphones and loudspeakers when the signal level is high (hands-free communication). Several methods including nonlinear cascade filters and a bilinear filter are proposed to compensate these echoes. A bilinear filter is a restricted NARMAX (Nonlinear Autoregessive Moving Average with eXogenous inputs) filter. We will present an evaluation based on the standard ERLE (echo return loss enhancement) measure, between a simple linear adaptive FIR filter and various nonlinear filters. These experiments are carried out first on a simulated communication system, then on experimental signals.

1. INTRODUCTION

Nowadays hands-free telephones employ a linear adaptive filter in order to compensate acoustic echoes (echo canceller diagram, figure 1). Moreover, competitive audio consumer products require not only cheap signal processing hardware but also low-cost analog equipment and sound transducers. The nonlinear distortions produced by these electro-acoustic transmission systems can't be described and analyzed by standard methods based on linear system theory alone [1], [2],[3]. The commonly used linear acoustic echo canceller (AEC) can't compensate these kind of echoes. Therefore the far-end user may hear an annoying distorted echo of high volume speech portions, while echo of low volume portions is removed by the linear canceller (Adaptive FIR filter)[4].

Some methods have been studied for nonlinear echo cancellation. Volterra series based filters [5]. This filter method can represent a large class of nonlinear systems but implies a high computational complexity (section 3). Neural networks [6]; this cascade structure offer a new perspective but need an extra reference microphone [4]. NARMAX structure method; this is a general parametric model but need a pre-identification procedure [7].

More recently, cascade filter structures have been proposed [2],[4]. In [2] a non polynomial Wiener-Hammerstein model is presented with a saturation nonlinear memoryless function. This kind of function can't modeled a large class of nonlinearities. In [4] a polynomial hammerstein structure is proposed.

This work is funded in part by the company STEPMIND.

In section 2 cascade structures are proposed and described. Comments and limitations of these models are discussed in section 3. Results of echo cancellation on simulated and experimental signals are summarized in table 2 of section 5.

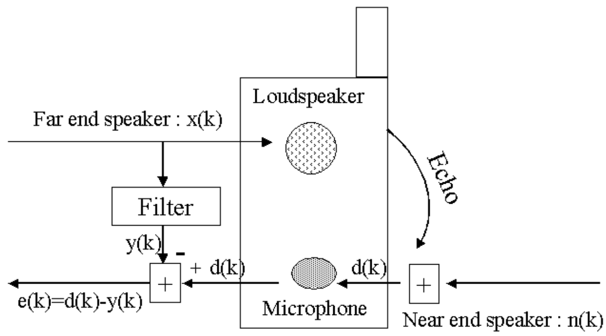

Fig. 1. Echo canceller diagram.

2. STRUCTURE OF FILTERS

Let y_k and x_k be respectively the observed output and input samples at k time.

2.1. FIR filter

The input/ouptut relationship is defined as follow :

$$y_k = \sum_{i=0}^{m_x-1} h_i x_{k-i} \,, \tag{1}$$

where m_x is the memory length of the filter. This linear filter is commonly used to compensate acoustic echoes.

2.2. Nonlinear structures

2.2.1. Volterra model

The discrete-time invariant Volterra filter with memory length m_x and order of nonlinearity D, with L samples, is defined by :

$$y_k = \sum_{i=1}^{D} \sum_{j_1,\cdots,j_i=0}^{m_x-1} \theta^i_{j_1,\cdots,j_n} x_{k-j_1} \cdots x_{k-j_i} \quad (2)$$

The Volterra filter is attractive because it's a straightforward generalization of the linear system description and the behaviour of many physical systems can be described with a Volterra filter [8].

2.2.2. Bilinear model

The Bilinear model is a parametric model, which contains cross terms [9]; it corresponds to a subclass of the NARMAX (Nonlinear Autoregessive Moving Average with eXogenous inputs) structure [7]:

$$y_k = \sum_{i=0}^{m_{x1}-1} \alpha_i x_{k-i} + \sum_{j=1}^{m_{y1}} \beta_j y_{k-j} + \sum_{i=0}^{m_{x2}-1} \sum_{j=1}^{m_{y2}} \gamma_{ij} x_{k-i} y_{k-j} \quad (3)$$

The inclusion of information from both lagged inputs and outputs provides a very concise representation for nonlinear systems

2.3. Nonlinear cascade structures

The main advantage of these structures is to introduce less parameters to be estimated.

2.3.1. Hammerstein model

It's a cascade of a memoryless polynomial filter and a FIR filter.

$$\begin{aligned} u_k &= \sum_{i=1}^{D} a_i x^i \\ y_k &= \sum_{i=0}^{m_x-1} h_i u_{k-i} \end{aligned} \quad (4)$$

2.3.2. Wiener model

It's a cascade of a FIR filter and a memoryless polynomial filter.

$$u_k = \sum_{i=0}^{m_x-1} h_i x_{k-i} \ , y_k = \sum_{i=1}^{D} a_i u_k^i \quad (5)$$

2.3.3. Wiener–Hammerstein model

It's a cascade of a FIR filter, a memoryless polynomial filter and a FIR filter.

$$\begin{aligned} u_k^1 &= \sum_{i=0}^{m_{x1}-1} h_i x_{k-i} \ , u_k^2 = \sum_{i=1}^{D} a_i (u_k^1)^i \\ y_k &= \sum_{i=0}^{m_{x2}-1} h_i u_{k-i}^2 \end{aligned} \quad (6)$$

3. COMMENTS AND LIMITATIONS

3.1. Volterra vs nonlinear cascade structures

Let y_k^V, y_k^W and y_k^H be respectively the output sequence of the Volterra, Wiener and Hammerstein models. Let's take $D = 2$ and $m_x = 2$:

$$\begin{aligned} y_k^V &= \underline{\underline{\theta_0^1}} x_k + \theta_1^1 x_{k-1} + \theta_0^2 x_k + \theta_1^2 x_{k-1} + \theta_{01}^2 \underline{x_k x_{k-1}} \quad (7) \\ y_k^W &= \underline{\underline{a_1 h_0}} x_k + a_1 h_1 x_{k-1} + a_2 h_0 x_k^2 + a_2 h_1 x_{k-1}^2 + \ldots \\ & a_2 h_0 h_1 \underline{x_k x_{k-1}} \quad (8) \\ y_k^H &= \underline{\underline{a_1 h_0}} x_k + a_1 h_1 x_{k-1} + a_2 h_0 x_k^2 + a_2 h_1 x_{k-1}^2 \quad (9) \end{aligned}$$

This basic example indicates that :

- The Volterra and Wiener models contain more terms than the Hammerstein model $\underline{x_k x_{k-1}}$ (all the cross terms are included).
- The number of parameters of the Volterra model is bigger than the Wiener and Hammerstein model (see table 1).
- In both the Wiener and Hammerstein models, fewer parameters are needed. A constraint is that the parameters of all terms are dependent ($\underline{\underline{a_1 h_0}}$ for the x_k term). This can leads to estimation errors.
- Note that the parameters of the Wiener–Hammerstein model are also dependent.
- All these cascade filter methods can be consider as a particular subclass of a Volterra series filter.

3.2. Number of parameters

The number of parameters to be estimated is crucial for the computational complexity and convergence of the filter algorithms. The number of parameters is evaluated for :

- EXAMPLE 1 : $D = 3$, $m_x = 10$, $m_{x1} = m_{x2} = 10$ and $m_{y1} = 0$ and $m_{y2} = 1$.
- EXAMPLE 2 :$D = 3$, $m_x = 20$, $m_{x1} = m_{x2} = 20$ and $m_{y1} = 0$ and $m_{y2} = 1$.

Model	Nb of parameters	Ex1	Ex2
FIR	m_x	10	20
Volterra	$(D+m_x)!/(D!m_x!)-1$	285	1770
Hammerstein	m_x+D	13	23
Wiener	m_x+D	13	23
Wiener-Hammerstein	$m_{x1}+m_{x2}+D$	23	43
Bilinear	$m_{x1}+m_{y1}+m_{x2}m_{y2}$	20	40

Table 1. Number of parameters.

Results are reported in table 1.

The number of parameters of the Volterra model quickly increases with D and m_x. As a consequence, large data sets are required in order to obtain an estimation of the model parameters with reasonable accuracy. For these reasons, we won't consider the Volterra model in the experiments of section 5.

As described in section 3.1 both cascade filters and Volterra filter can model nonlinear polynomial systems. However, for the cascade filters, if the amount of parameters needed is small, they are dependent from each other $(\mathbf{a}, \mathbf{h})$. Moreover, the polynomial order D is limited and fixed at $D = 3$ [6].

4. ADAPTATION ALGORITHMS

The parameters (excepted the a_i parameters of the cascade models) are updated using a stochastic gradient algorithm (NLMS) minimizing the expected value of the squared error signal e_k (with $e_k = d_k - \hat{d}_k$). The parameters of the polynomial filters are updated with a RLS algorithm, which allows a control of the adaptation speed which is an important issue in AEC [4]. Moreover, *a priori* and *a posteriori* errors are used for the convergence computation of the parameters. This leads to improve the quality of the echo cancellation process (see algorithm described in [4]).

5. EXPERIMENTS

5.1. Simulation results

Traditionally the nonlinear communication channel is modeled by saturation followed by a linear propagation. This representation does not take into account the nonlinear mechanical vibrations.

A commonly used function for modeling saturation is a sigmoid function (figure 4 for $\alpha = 1, 2$ and 5) well known in the neural networks community [10] and defined as follow :

$$\varphi(u) = (\frac{2}{1+\exp(-\alpha u)} - 1)\beta\,. \qquad (10)$$

The linear system is chosen as $H(z) = \frac{1}{1-\tau_1 z^{-1}}$, with $\tau_1 = 0.2$. The input sequence $\{x_k\}$ is i.i.d. $\mathcal{N}(0,1)$. For the

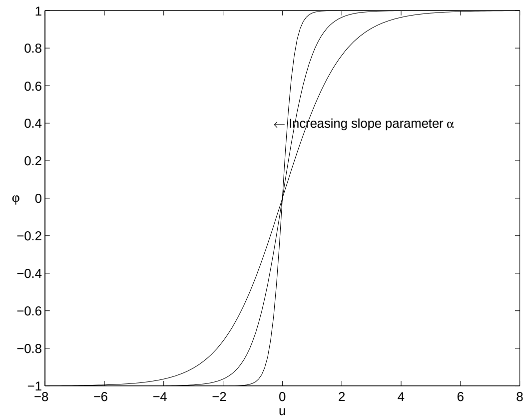

Fig. 2. Sigmoid function.

single talk context, the results are presented in terms of the Echo Return Loss Enhancement (ERLE), define by :

$$ERLE_{dB} = 10\log\left[\frac{\mathbf{d}^t\mathbf{d}}{\mathbf{e}^t\mathbf{e}}\right] \text{ (single talk context),}$$

and for the double talk context, we use the inverse of the normalized mean-squared error defined by :

$$Er_{dB} = 10\log\left[\frac{\mathbf{n}^t\mathbf{n}}{(\mathbf{e}-\mathbf{n})^t(\mathbf{e}-\mathbf{n})}\right] \text{ (double talk context),}$$

where $\mathbf{d}, \mathbf{e}, \mathbf{n}$ are respectively the observed, transmitted end local speech sequences (see figure 1). The polynomial order, and the memory length m_x are fixed : $D = 3$, $m_x = 10$, $m_{x1} = 10$. Results are obtained with $m_{y1} = 3, m_{x2} = 3, m_{y2} = 4$ for the Bilinear model and $m_{x2} = 10$ for the Wiener–Hammerstein model. The values given in table 2 are averaged over 100 realizations.

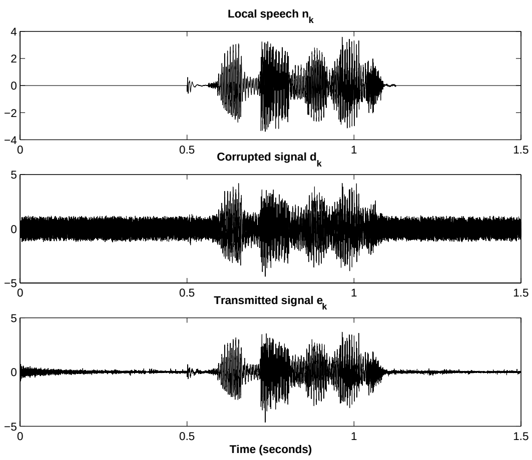

Fig. 3. Simulation results.

5.2. Experimental results

Using speech excitation and recordings from a real echo path, the nonlinear cascade filters are compared to the standard linear FIR filter. These sequences don't contain near-end speech ($n_k = 0$). The polynomial order, and the memory length m_x are fixed : $D = 3$, $m_x = 256$. Results are obtained with $m_{x1} = 100, m_{y1} = 70, m_{x2} = 2, m_{y2} = 4$ for the Bilinear model and $m_{x1} = 100, m_{x2} = 150$ for the Wiener–Hammerstein model.

MODEL	SIMULATION		EXPERIMENT
	$ERLE$	Er	$ERLE$
FIR	8.79	14.44	14.32
Hammerstein	**17.28**	**16.61**	**17.49**
Wiener	16.61	11.38	14.85
Wiener-Hammerstein	18.43	13.76	12.54
Bilinear	9.74	13.52	15.23

Table 2. Nonlinear echo cancellation results.

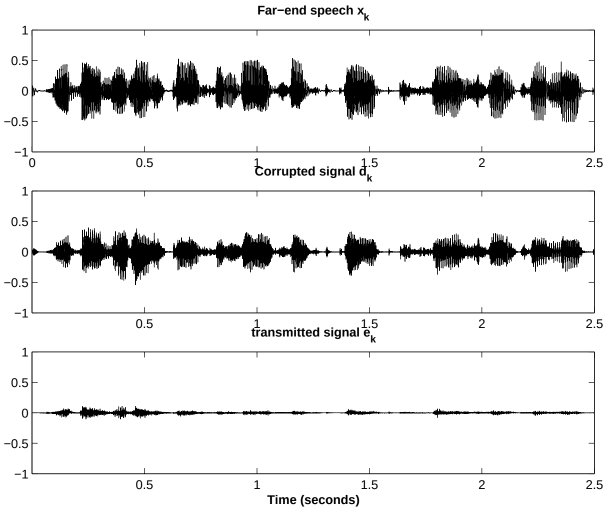

Fig. 4. Experimental results.

5.3. Comments

- The simulated example shows that the best results are obtained with the polynomial Wiener–Hammerstein model in terms of the ERLE measure, and the polynomial Hammerstein model for the double talk context.

- The experimental results indicate that the signals don't seem to have strong influence. Best results are obtained with the polynomial Hammerstein filter (fig. 4). For this reason the Wiener and the Wiener–Hammerstein filters don't have the expected results.

- The Bilinear filter, due to the complexity of its structure, has interesting preliminary results and need to be explored before being definitively rejected.

6. SUMMARY

New nonlinear structures of filters are proposed for acoustic echo cancellation : cascade structures and the Bilinear model. These structures are compared to the standard linear adaptive FIR filter method. For all the contexts tested, the Hammerstein filter has the best behaviour. Several issues require further study, for instance : so far we have ignored the presence of additive noise in the system. In practice, this assumption is unrealistic.

7. ACKNOWLEDGEMENTS

The authors are very grateful to F. Kuech and A. Stenger from the University of Erlangen-Nuremberg to provide experimental signals.

8. REFERENCES

[1] F. Heinle, R. Rabenstein, and A. Stenger, "A measurement method for the linear and nonlinear properties of electro-acoustic transmission systems.," *Signal Processing*, vol. 64, no. 1, pp. 49–60, 1998.

[2] B. S. Nollet and D. L. Jones, "Nonlinear echo cancellation for hands-free speakerphones," *NSIP*, September 1997.

[3] J-P. Costa, L. Pronzato, and E. Thierry, ""Nonlinear prediction by kriging, with application to noise cancellation"," *Signal Processing*, vol. 80, pp. 553–566, 2000.

[4] A. Stenger and W. Kellermann, "RLS-Adapted Polynomial for Nonlinear Acoustic Echo Cancelling," *Signal Processing*, vol. 80, pp. 1747–1760, Sep. 2000.

[5] A. Stenger, W. Kellermann, and R. Rabenstein, "Nonlinear acoustic echo cancellation with 2nd order adaptive volterra filters," *Icassp*, 1999.

[6] A. N. Birkett and R. A. Goubran, "Acoustic echo cancellation using nlms-neural network structures," *Icassp*, vol. 3, 1995.

[7] J-P. Costa, T. Pitarque, and E. Thierry, ""Using Orthogonal Least Square For Adaptive Nonlinear Filtering of GSM Signals"," *Icassp*, vol. 3, pp. 2397–2400, 1997.

[8] R. D. Nowak and B. D. Van Veen, "Random and Pseudorandom Inputs for Volterra Filter Identification," *IEEE trans. on Signal Processing*, vol. 42, no. 8, august 1994.

[9] F. Fnaiech and L. Ljung, ""Recursive identification of bilinear systems "," *Int. J. Control*, vol. 45, no. 2, pp. 453–470, 1987.

[10] S. Haykin, *"Adaptive Filter Theory"*, Third Edition, Prentice Hall, Inc, A Simon & Schuster Company, New Jersey, 1996.