USING ORTHOGONAL LEAST SQUARES IDENTIFICATION FOR ADAPTIVE NONLINEAR FILTERING OF GSM SIGNALS

Jean-Pierre COSTA, Thierry PITARQUE and Eric THIERRY

Laboratoire I3S, CNRS-UNSA, 250 Av Albert EINSTEIN
Sophia Antipolis, 06560 Valbonne FRANCE
e-mail : costa@alto.unice.fr

ABSTRACT

The miniaturization of GSM handsets creates nonlinear acoustical echoes between the microphone and the loudspeaker when signal level is high. Nonlinear adaptive filtering can tackle this problem but the computational complexity has to be reduced by restricting the number of coefficients introduced by nonlinear models. This paper compares performances of different nonlinear models. In a first training stage we use the OLS (Orthogonal Least Squares) identification method to find models using the fewest coefficients along with a good fitting accuracy. In a second filtering stage these parsimonious models are used to adaptively filter the GSM signals.

1. INTRODUCTION

The miniaturization of GSM handsets creates nonlinear acoustical echoes between the microphone and the loudspeaker when signal level is high. This nonlinear phenomena are introduced by mechanical propagation of vibrations along the handset. They are not to be confused with linear echoes (like in a car) which are processed bylinear adaptive filters such as NLMS (Normalised Least Mean square) [1].
The main items in the context of a nonlinear filtering are the choices of a suitable model and of the number of coefficients. You have to deal with the computation complexity (induced by the number of coefficients) and the maximum normalized estimation error required by the application constraints. Various authors [2] [3] have studied nonlinear models for identification of systems or for adaptive nonlinear filtering.
The goal of this paper is to compare performances of different nonlinear models using the following methodology :

- In a first training stage we use the OLS (Orthogonal Least Squares) identification method [4] [5] to find models using the fewest parameters along with a good fitting accuracy.
- In a second filtering stage these parsimonious models are used to adaptively filter the GSM signals (cf Fig 1) .

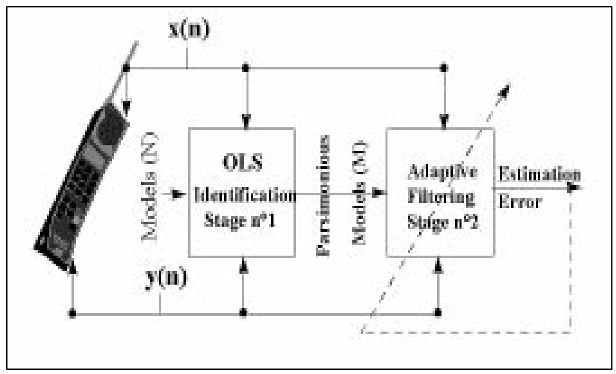

Figure 1: Identification and filtering methodology

Four classes of models are used. Input and output sequences are respectively denoted by $x(n)$ and $y(n)$, their memory lengths are n_x and n_y.
The first model is a linear ARMA model. It is used as a reference model.

$$y(n) = \sum_{i=0}^{n_x-1} a_i x(n-i) + \sum_{j=1}^{n_y} b_j y(n-j)$$

The second one is a Volterra model with a polynomial order D and a memory length n_x [6]. It is one of the most popular nonlinear models, but is also one of the most expensive because of the large number of coefficients used.

$$y(n) = \sum_{i=1}^{D} \sum_{j_1,\cdots,j_i=0}^{n_x-1} h_i(j_1,\cdots,j_n) x_{n-j_1}\cdots x_{n-j_n}$$

The following models are more recent, they generate cross-terms between the input signal x and the output signal y.
The third model is a Bilinear model [7] without exogenous input.

$$y(n) = \sum_{i=0}^{n_x-1} a_i x(n-i) + \sum_{j=1}^{n_y} b_j y(n-j) + \sum_{i=0}^{n_x-1}\sum_{j=1}^{n_y} c_{ij} x(n-i) y(n-j)$$

The last one is a NARMAX (Nonlinear Autoregessive Moving Average with eXogenous inputs) model [8].

$$y(n) = \sum_{i=0}^{n_x-1} a_i x(n-i) + \sum_{j=1}^{n_y} b_j y(n-j) + \sum_{i=0}^{n_x-1}\sum_{j=1}^{n_y} c_{ij} x(n-i) y(n-j) + \sum_{i=0}^{n_x-1}\sum_{j=0}^{n_{x-1}} d_{ij} x(n-i) x(n-j) + \sum_{i=1}^{n_y}\sum_{j=1}^{n_y} e_{ij} y(n-i) y(n-j)$$

Table 1 shows the different models and their number of coefficients for $n_x = 10$, $n_y = 10$ and $D = 3$ (order of nonlinearity for the Volterra model). The NARMAX model has been restricted to an order 2 of nonlinearity (no cubic terms), without exogenous input. It is clear from table 1, that excepted the ARMA model, it is necessary to decrease the number of coefficients.

$Model$	$Coefficients$	$Example$
$ARMA$	$n_x + n_y$	20
$Volterra$	$(D + n_x)!/(D!n_x!)$	285
$Bilinear$	$n_x + n_y + n_x n_y$	120
$NARMAX$	$n_x + n_y + n_x n_y + \frac{n_x^2+n_y^2+n_x+n_y}{2}$	230

Table 1: Model's coefficients

2. OLS IDENTIFICATION

The goal of the OLS identification is to reduce the number of coefficients using a criterion based on the energy of the desired output. Departing from a basis P such as

$$y = \sum_{i=1}^{N} a_i P_i + \eta = [y(0), \cdots, y(L-1)]^t \qquad (1)$$

where L is the observation length, N is the number of coefficients and η is a observation noise; we construct an orthogonal basis W such that :

$$\hat{y} = \sum_{i=1}^{M} g_i W_i \qquad (2)$$

$\hat{y}$ is the estimate of y. To avoid numerical problems, because N is typically large, OLS should select M orthogonal basis vectors with $N \gg M$. The energy of $\hat{y}$ is given by :

$$\epsilon_r = \sum_{i=1}^{M} g_i^2 W_i^2 \qquad (3)$$

At each step of the orthogonal decomposition of P, we select the vector W_k which maximizes the individual energy $g_i^2 W_i^2$. This iterative process is stopped either when $\epsilon_r < \epsilon$ (ϵ is an a priori fixed threshold), or $M = N$. After the OLS identification stage, the normalized estimation error of the model is estimated by :

$$E_r = \frac{(y - \tilde{y})^2}{y^t y} \qquad (4)$$

where $\tilde{y}$ is given for example for the ARMA model by :

$$\tilde{y} = \sum_{i=1}^{m_1} c_i x(n-\alpha_i) + \sum_{j=m_1+1}^{M} d_j \tilde{y}(n-\gamma_j) \qquad (5)$$

where m_1 is the reduced number of coefficients in x, and $M - m_1$ is the reduced number of coefficients in y with their respective α_i and γ_j lags.
As a validation of the method, we apply this approach to GSM handset signal. We can then compare model performances.

3. MODEL SELECTION

The "Priestley test" [9] is used to compare the performance of models. This statistical test for L samples is given by :

$$g = L \log(\frac{\sigma_1^2}{\sigma_2^2}) \qquad (6)$$

where σ^2, the residual variance, is defined by :

$$\sigma^2 = \frac{1}{L-1+d}\sum_{i=d+1}^{L} e_r(i)^2 \qquad (7)$$

with d the maximum length memory used by the model ($d = max[n_x, n_y]$), and $e_r = y - \hat{y}$. g is distributed on χ_q^2, where q is the difference in the number of parameters of the two compared models. The second model is accepted as a significant improvement over the first one if g exceeds its 95% confidence level. This test is used after the filtering stage.

4. IDENTIFICATION OF GSM SIGNALS

Experiments have been conducted with a GSM handset and a DSP (TMS 320C30) based acquisition board for different input signals (single frequency, white noise and speech). $x(n)$ represents the signal sent to the loudspeaker and $y(n)$ the signal received by the microphone (cf Fig 1). When the input signal is a sinusoid of frequency f_0, the output $y(n)$ has new frequencies like $3 * f_0$, $5 * f_0$ and more. Figure 2 shows the FFT of y(n) when f_0 is equal to 300 Hz.

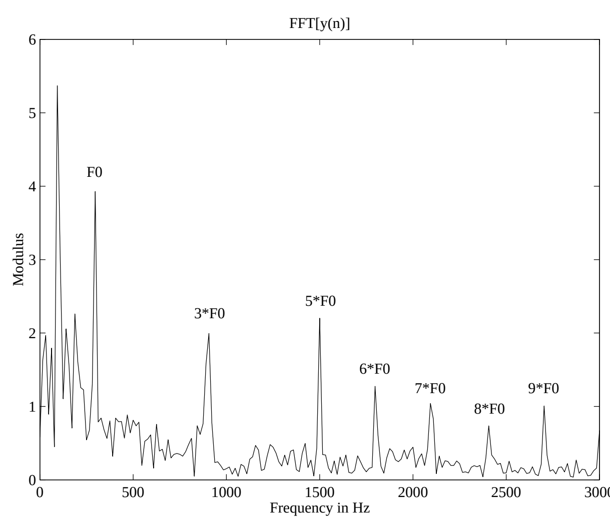

Figure 2: FFT[y(n)] for a single frequency input

Now x(n) is a white noise sequence and has 7000 samples ($L = 7000$), but for the identification step only the first 1000 samples are used, the filtering step use the rest of samples. OLS identification results are given in Table 2.

The comparison of the model should be done by considering the reduced number M of coefficients and the normalised estimation error ($\%E_r$). In this case the ARMA, Bilinear and NARMAX models give the best results. However the final comparison will be done after evaluation of the adaptive nonlinear filter performances.

$Model$	n_x	n_y	N	M	$\%\epsilon_r$	$\%E_r$
$ARMA$	40	10	50	26	95.01	12.08
$Volterra$	10	$\backslash$	285	285	50.19	47.79
$Bilinear$	25	10	285	42	95.02	13.17
$NARMAX$	25	10	665	36	95	11.56

Table 2: OLS results

5. NONLINEAR ADAPTIVE FILTERING OF GSM SIGNALS

To compare these models, we adaptively filter the second part of the GSM signal (i.e. the remaining 6000 samples) with a suitable forgetting factor, and initialize the filter's coefficients with the coefficients estimated by the OLS.
Figures 3 and 4 represent respectively the filtered signal for the ARMA, Bilinear and NARMAX model and the PSD of the outputs. The normalized estimation error ($\%E_r$) is :

- $123, 30\%$ for the ARMA model,
- 16.53% for the Bilinear model,
- 15.21% for the NARMAX model.

The PSD graphs confirm that the ARMA model does not have a good behavior in the filtering stage. Although the results of the OLS identification stage seem comparable for the three selected models, the adaptive filtering stage highlights the superiority of the Bilinear and NARMAX model in this case.
The Priestley test can be used now to compare the residual variance between the Bilinear and the NARMAX model. The NARMAX model is selected by this test, because the residual variance σ^2 is lower than the residual variance of the Bilinear model.

6. CONCLUSION

Given these results, we can draw several conclusions :

- the Volterra model has been rejected, because its coefficient number remains too high after the OLS identification,

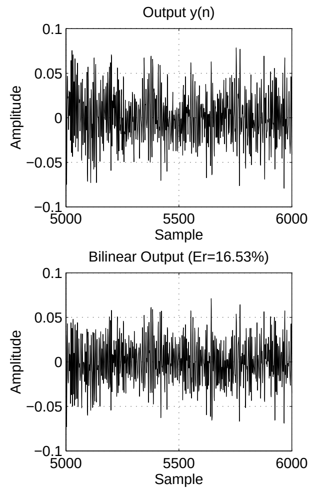

Figure 3: Filtering results for the different models

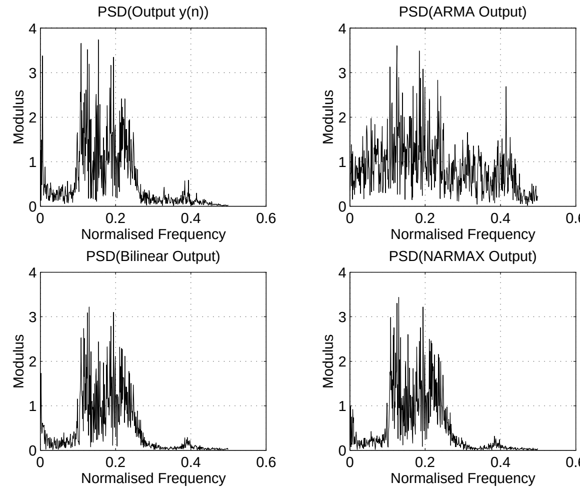

Figure 4: PSD of the outputs of various models

- the ARMA model has been rejected after the filtering stage, because its performance is not good enough,
- the Bilinear and the NARMAX models have the same behaviour in this application. They have a comparable normalized estimation error and after the filtering stage, their residual variance are similar, $\sigma^2_{Bilinear} = 0.8079$ and $\sigma^2_{NARMAX} = 0.7260$,
- the "Priestley test" has selected the NARMAX model, because this one has the best residual variance and the fewest coefficients,
- the OLS identification gives good reduction coefficient number. For the Bilinear and NARMAX model, this reduction takes the following value :
  - 85.26% for the Bilinear model,
  - 94.58% for the NARMAX model.

  with reasonable fitting and filtering accuracies.

This methodology, OLS identification followed by adaptive nonlinear filtering, presented in this paper has been successfully applied of GSM signals for a white noise input sequence. However the choice of the excitation input seems very important and we are currently working on this problem.

7. REFERENCES

[1] A. GILLOIRE and T. PETILLON. "A comparison of NLMS and RLS algorithms for identification of time varying systems with noisy output. application to accoustic echo cancellation". *Proc. EUSIPCO-90, Barcelona, Spain*, pages 417–420, 1990.

[2] J.C STAPLETON and S.C. BASS. "Adaptive Noise Cancellation for a Class of Nonlinear, Dynamic Reference Channels". *IEEE Transactions on Circuit and Systems*, cas-32(2):143–150, Feb. 1985.

[3] V.J. MATHEWS. "Adaptive Polynomial Filters". *IEEE Transactions on Signal Processing*, pages 10–26, 1991.

[4] M.J. KORENBERG. "Orthogonal Identification of Nonlinear difference Equation Models". *Proc. 28th Midwest Symp. Cir. Sys., Louisville, KY*, pages 90–95, Aug. 1985.

[5] S. CHEN, C.F. COWAN, and P.M. GRANT. "Orthogonal Least Squares Learning Algorithm for Radial Basis Function Networks ". *IEEE Transactions on Neural Networks*, 2(2):302–309, March 1991.

[6] T. KOH and E.J. POWERS. "Second Order Volterra Filtering and its Application to Nonlinear System Identification". *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-33(6):1445–1455, Dec. 1985.

[7] F. FNAIECH and L. LJUNG. "Recursive identification of bilinear systems ". *Int. J. Control*, 45(2):453–470, 1987.

[8] S. CHEN and S.A. BILLINGS. "Representation of nonlinear systems : the narmax model". *Int. J. Control*, 49(3):1013–1032, 1989.

[9] D. Christini, F. Bennett, K. Lutchen, H. Ahmed, J. Haussdorff, and N. Oriol. "Application of Linear and Nonlinear Time Series Modeling to Heart Rate Dynamics Analysis". *IEEE Transactions on Boimedical Engineering*, 42(4):411–415, 1995.