Parameter Identification of a Class of Nonlinear Systems for Robust Audio Watermarking Verification

Shay Mizrachi and David Malah
Department of Electrical Engineering
Technion, Israel Institute of Technology

1. Introduction

Audio watermarking systems for protecting ownership rights may need to withstand attacks that include passing the watermarked audio signal through nonlinear systems, with or without memory [1]. In the considered verification scheme, ownership verification requires the identification of the attacking system parameters. While linear filter parameters are relatively easy to identify by applying LMS type adaptive algorithms, the identification of the parameters of an attack system that includes a nonlinear component is usually difficult. A common approach for handling nonlinearities is to perform a Volterra series expansion [2]. The method can be considered as a generalization of linear filtering theory, by modeling the system as a linear system followed by a memoryless nonlinearity. However, the number of parameters involved in the series expansion, describing the input-output relationship, grows polynomially with the nonlinearity order. This may well be prohibitive in terms of computations, convergence time, and sometimes even in terms of the large amount of input/output data samples needed for the identification process (particularly when the system varies in time). [4] describes the FPET (Fixed Pole Expansion Technique) approach, which employs orthogonal basis function derived from fixed pole locations to expand the Volterra kernels and reduce the number of estimated parameters.

In this work we present an approach by which the memoryless nonlinearity is modeled by a piecewise approximation of the nonlinear function using a relatively small number of linear segments. It is shown that the adaptive LMS framework can be used to estimate the segments slopes. Furthermore, it is shown that the LMS-based identification processes of both the linear system coefficients (modeled as an FIR filter) and the memoryless nonlinearity parameters (slopes) can be combined into a simple and efficient single process. In addition to its ability to handle a cascade of linear (L) and nonlinear (N) components (both L-N and N-L cascade arrangements) this approach can be extended it in a modular way to handle other cascade arrangements like N-L-N, L-N-L, N-L-N-L, etc.

In the next section the Volterra series expansion is introduced and its limitations are discussed. Next, in section 3, we described the proposed solution, which is followed by conclusions in the last section.

2. The Volterra series expansion

A Volterra-based nonlinear adaptive filter consists of a linear filter that is followed by a memoryless non-linearity. The Volterra series provides a method for describing the input-output relationship for a nonlinear device with memory. Let x_n denote the input signal samples. Then, we may combine these input samples to get the following discrete Volterra kernels:

H_0 = zero-order (dc) term; $H_1[x_n]$= first-order (linear) term = $\sum_i h_i \cdot x_i$;

$H_2[x_n]$= second-order (quadratic) term = $\sum_i \sum_j h_{ij} \cdot x_i \cdot x_j$; $H_3[x_n]$= third-order (cubic) term = $\sum_i \sum_j \sum_k h_{ijk} \cdot x_i \cdot x_j \cdot x_k$,

and so on - for higher order terms. The filter parameters, i.e., the h coefficients, are determined analytically. Therefore, we can separate the nonlinear adaptive filter into two parts, as follows:

Part I: Expansion of the linear space into a nonlinear space according to the selected order of the Volterra term. This part uses $x_0, x_1, \ldots, x_n$ to built a larger set of signals: $u_1, u_2, \ldots, u_q$ where $q>n$. For example, the vector expansion for a (3,2) system, i.e., a system of order 2 and memory length 3, has the form: $\underline{u}= [1, x_0, x_1, x_2, x_0^2, x_0x_1, x_0x_2, x_1x_0, x_1^2, x_1x_2, x_2x_0, x_2x_1, x_2^2]^T$.

Part II: A linear FIR filter – this filter uses the elements of $\underline{u}$ as samples for creating the required signal, resulting in an adaptive filtering scheme as shown in Fig. 1, with $d(n)$ denoting the reference signal and $e(n)$ - the error signal.

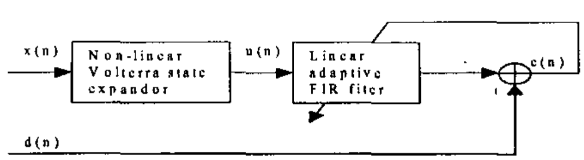

Fig. 1

It is seen that a Volterra series of order 2 requires $N+N^2$ coefficients, compared to N coefficients only for a first order expansion. Thus, the expansion complexity increases polynomially with the Volterra series order. Moreover, when dealing with attacks on a watermarking system, a filter of high order (N in the hundreds) followed by a soft nonlinearity may be used. Thus, a Volterra series order of two will result in an expansion having thousands of coefficients. One should note that besides complexity, a crucial issue could then be the prohibitive amount of data (input samples) needed for the estimation.

3. Proposed solution

3.1. Identification of a memoryless nonlinear system

Problem formulation: Given: $y_i = f(x_i)$; $\{x_i, y_i\}_{i=1}^{N}$, $x_i \in [0,1]$, where $f(x_i)$ is a non-linear function and $f(0)=0$.

Goal: Piecewise regression analysis for $y=f(x)$ using $\{x_i, y_i\}_{i=1}^{N}$.

Method: Find a piecewise approximation for $f(x)$ as described in Fig. 2.

Let L be the number of segments, of length ℓ, each.

Thus, with $x \in [0,1] \Rightarrow L \cdot \ell = 1$

Requirement: Continuity between segments.

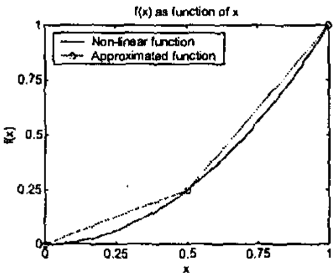

Fig. 2

For each segment, L_i, we define a corresponding slope, $\hat{a}_i$. The approximation, $\hat{f}(x)$, is therefore:

$$\hat{f}(x) = \begin{cases} \hat{f}_1(x) & 0 \le x \le \ell \\ \hat{f}_2(x) & \ell \le x \le 2\cdot\ell \\ \vdots & \\ \hat{f}_L(x) & (L-1)\cdot\ell \le x \le L\cdot\ell \end{cases} \quad \text{where} \quad \begin{aligned} &\hat{f}_1(x) = \hat{a}_1 \cdot x \\ &\hat{f}_2(x) = \hat{a}_2 \cdot (x-\ell) + \hat{a}_1 \cdot \ell \\ &\vdots \\ &\hat{f}_i(x) = \hat{a}_i \cdot (x-(i-1)\cdot\ell) + \ell \cdot \sum_{k=0}^{i-1} \hat{a}_k \qquad (\hat{a}_0 = 0) \end{aligned}$$

Now, we can redefine the problem: Using $\{x_i, y_i\}_{i=1}^{N}$, find $\{\hat{a}_i\}_{i=1}^{L}$ that minimize the approximation error according to a specific criterion.

Method A: Identification using an LS error criterion

In this method, the estimation of $\{\hat{a}_i\}_{i=1}^{L}$ is based on using all the input-output pairs in the *i*-th segment: $\{(x,y)\}_i \equiv \{(x,y) : (i-1)\cdot\ell \le x \le i\cdot\ell\}$. The idea is to calculate the estimated slopes $\{\hat{a}_i\}_{i=1}^{L}$ sequentially, starting with the first (i=1). Thus, the calculation of $\hat{a}_4$ uses the already computed $\hat{a}_3$, $\hat{a}_2$, and $\hat{a}_1$. The selected error criterion, for the *i-th* segment is $D_i = \sum_{(x,y)\in\{(x,y)\}_i} \left(\hat{a}_i \cdot (x-(i-1)\cdot\ell) + \ell\cdot\sum_{k=1}^{i-1}\hat{a}_k - y\right)^2$. Hence, by requiring:

$$\frac{\partial D_i}{\partial \hat{a}_i} = 2 \cdot \sum_{(x,y)\in\{(x,y)\}_i} \left(\hat{a}_i \cdot (x-(i-1)\cdot\ell) + \ell\cdot\sum_{k=1}^{i-1}\hat{a}_k - y\right)\cdot(x-(i-1)\cdot\ell) = 0,$$

$$\text{one obtains:} \quad \hat{a}_i = \frac{\sum_{(x,y)\in\{(x,y)\}_i} \left(y - \ell\cdot\sum_{k=1}^{i-1}\hat{a}_k\right)\cdot(x-(i-1)\cdot\ell)}{\sum_{(x,y)\in\{(x,y)\}_i} (x-(i-1)\cdot\ell)^2}$$

Method B: Identification using an adaptive system

In this method the signal samples are fed into an LMS (or a normalized LMS: NLMS) identification system, as shown in Fig. 3.

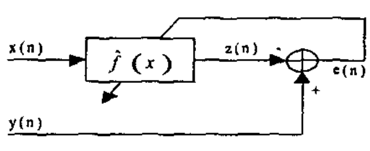

Fig. 3

The error criterion is: $\varepsilon^2 = E\{e(n)^2\} = E\{[y(n) - \hat{f}_{I(x)}(x(n))]^2\}$, where $I(x)$ denotes the segment number that contains incoming sample x, i.e., $I(x) = j \quad if \quad (j-1)\ell \le x \le j\cdot\ell$.

The resulting update equation for each incoming sample is [3]:

$$\hat{a}_j^{n+1} = \hat{a}_j^n + \Delta \cdot e(n) \cdot \begin{cases} \left(x(n) - \left(I(x(n)) - 1\right) \cdot \ell\right) & j = I(x(n)) \\ \ell & j < I(x(n)) \\ 0 & otherwise \end{cases}$$

where Δ is a constant factor (for LMS), or is varied to track the input signal level (for NLMS).

3.2 Identification of a nonlinear system with memory

To identify a nonlinear system with memory, constructed as a combination of a linear filter followed by a nonlinearity, we propose to combine the conventional LMS (or NLMS) adaptive filtering method with method B above, as shown in Fig. 4.

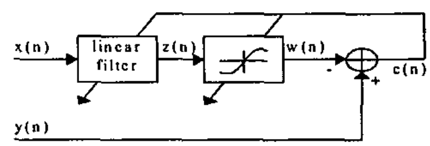

Fig. 4

The resulting update equations can be shown to be [3]:

1. For the linear filter coefficients:

$$h_k(n) = h_k(n-1) + \Delta \cdot e(n) \cdot x(n-k) \cdot \hat{a}_{I(z(n))};\ k = 0,1,\ldots,N\text{-}1$$

2. For the memoryless nonlinearity coefficients (slopes):

$$\hat{a}_j(n) = \hat{a}_j(n-1) + \Delta \cdot e(n) \cdot \begin{cases} z(n) - \left(I(z(n)) - 1\right) \cdot \ell & if \quad j = I(z(n)) \\ \ell & if \quad j < I(z(n)) \\ 0 & otherwise \end{cases}$$

3.3 Identification of other nonlinear systems with memory

The above approach can easily be extended to other nonlinear systems with memory that are composed of a cascade of a linear filter and a memoryless non-linearity. For example, an N-L combination, which means that the memoryless nonlinearity precedes the linear filter (while above we considered an L-N combination), as well as L-N-L and N-L-N, and so on. For example, the update equations for an N-L system are [3]:

1. For the linear filter coefficients:

$$\hat{h}_k^{n+1} = \hat{h}_k^n + \Delta \cdot e(n) \cdot z(n-j),\ \ k = 0,1,\ldots N-1$$

2. For the memoryless nonlinearity coefficients:

$$\hat{a}_j^{n+1} = \hat{a}_j^n + \Delta \cdot e(n) \cdot \sum_{k=0}^{N-1} \hat{h}_k(n) \cdot \begin{cases} x(n-k) - \left(I(x(n-k)) - 1\right) \cdot \ell & if \ j = I(x(n-k)) \\ \ell & if \ j < I(x(n-k)) \\ 0 & otherwise \end{cases}$$

4. Conclusions

The proposed identification method has been examined in [3] for an audio watermarking system by simulating a wide range of attacks of the above kind (as well as other attacks - like additive noise) and proved to dramatically increase the robustness of the verification system. Method B is preferable over method A (LS) since it is fully adaptive and can also handle nonlinear systems that vary slowly in time. The proposed method can be applied to a variety of combinations of filters and nolinearities, like N-L-N / N-L / L-N and can be useful also in modeling and estimating a variety of communication channels that contain a memoryless nonlinearity either at the transmitting end or receiving end or in both.

5. References

[1] M. D. Swanson, B. Zhu, A. H. Tewfik, and L. Bondey, "Robust audio watermarking using perceptual masking", Signal Processing 66 (1998), pp. 337-355.

[2] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, New Jersey, 1986.

[3] Shay Mizrachi, "Robust Detection Of Watermarks In Audio Signals", M.Sc. thesis, Technion, Sep 2000 [in Hebrew].

[4] Rifat Hacioglu, Geoffrey A. Williamson, "Reduced Complexity Volterra Models for Nonlinear System Identification"