ADAPTIVE AR MODEL IDENTIFICATION BASED ON THE FAEST FILTERS

S. D. Likothanassis [1,2], *E. N. Demiris* [1] *and D. G. Karelis* [3]
[1]Department of Computer Engineering and Informatics, Univ. of Patras, Patras 26500, GREECE
[2]Computer Technology Institute (C.T.I.), Patras 26100, GREECE
[3]Department of Electrical Engineering, T.E.I. of Patras, Patras 26334, GREECE.
Tel: +30 61 997755, Fax: +30 61 997706
E-mail: Likothan@cti.gr, Demiris@ceid.upatras.gr

ABSTRACT

A new adaptive approach for simultaneously selecting the order and identifying the parameters of an AutoRegressive model (AR) is presented. The proposed algorithm is based on the reformulation of the problem in the standard state space form and the subsequent implementation of a bank of fast a posteriori error sequential technique (**FAEST**) filters, each fitting a different order model. The problem is reduced then to selecting the true model, using the Multi-Model Partitioning (MMP) theory. Simulations illustrate that the proposed method is selecting the correct model order and identifies the model parameters, even in the case that the true model order does not belong to the bank of FAEST filters. The use of FAEST filters, reduce the computational effort, especially in the case of large order AR models. Finally, the algorithm is parallel by nature and thus suitable for VLSI implementation.

1 INTRODUCTION

Adaptive filtering has been a central issue in the field of signal processing for many years. Selecting the correct order and estimating the parameters of an AR model, is fundamental in many application areas and the related algorithms are part of many existing systems. The problem of fitting an AR model to a given time series has attracted much attention because it arises in a large variety of applications, such as adaptive control, speech analysis and synthesis, radar, sonar, seismology and biomedical engineering.

Several information theoretic criteria have been proposed for the model order selection task. The most well known of the proposed solutions for this problem include the Final Prediction Error (FPE), Akaike's Information Criterion (AIC) that was proposed by Akaike [1], [2], [3], the Minimum Description Length (MDL) Criterion that was proposed by Schwartz [16] and Rissanen [15] and a new approach based on the MDL criterion [13]. Most of the techniques that are performed by the above criteria are based on the assumption that the data are Gaussian and upon asymptotic results. Furthermore, they are two-pass methods; thus they can not be used in an on-line or adaptive fashion. In this paper, the method that is presented for simultaneous AR model order selection and identification is based on the adaptive MMP filters [9] - [11]. Performance bounds for the MMP filters have been given in [6]. The method is applicable to on-line/adaptive operation and is computationally efficient. Furthermore, it can be realized in a parallel processing fashion, a fact which make it amenable to VLSI implementation. A similar method for AR model order selection is found in [7], which is implemented using a bank of convetional or modified Kalman filters. The new method, by utilizing the computationally efficient FAEST filters is significally more efficient for large AR models [5].

The main points and the organization of this paper are the following: In section 2, the AR model order selection and identification problem is stated. In section 3 the problem is reformulated, the adaptive multi-model partitioning algorithm is briefly presented and it's application to the specific problem is discussed. In section 4, simulation examples and figures are presented that demonstrate the performance of the method. Finally section 5, summarizes the conclusions.

2 PROBLEM STATEMENT

An AR model can be represented as follows [4]:

$$A(q)y(t) = e(t) \tag{1}$$

or

$$y(t) = \sum_{i=1}^{n} a_i y(t-i) + e(t) \tag{2}$$

where N is the number of the noisy measurements of the discrete time process $y(t)$, $e(t)$ is a zero-mean white noise process, with variance R, not necessarily Gaussian, n is the order of the predictor and $a_i : i = 1, \dots, n$, are the predictor coefficients.

Let us further define the vector of coefficients as follows:

$$\vartheta(t) = [a_1(t), \dots, a_n(t)]^\top \tag{3}$$

where $0 \le t \le N$ (N denotes the number of samples) and coefficients a_i have been replaced by $a_i(t)$ to reflect the possibility that they are subject to random perturbations. This fact can be modeled by assuming that:

$$\vartheta(t+1) = \vartheta(t) + w(t), \quad t = 1, 2, \ldots, N \tag{4}$$

where $w(t) = [w_1(t), w_2(t), \ldots, w_n(t)]$ is a zero-mean white noise process with variance Q (we assume that $e(t)$ and $w(t)$ are independent). In order to complete the system description we assume the case that the prediction coefficients are constant in time, or slowly varying ($Q = 0$), so (4) is replaced by:

$$\vartheta(t+1) = \vartheta(t), \quad t = 1, 2, \ldots, N \tag{5}$$

The problem is now stated as follows: Given a set of observations $y(t)$ where $0 \leq t \leq N$, from an AR(n) process we have to determine the unknown parameter vector:

$$v = [n, \vartheta(t), R] \tag{6}$$

Clearly the problem is two-fold: one has both to select the order of the predictor and then to compute the predictor coefficients. Perhaps the most crucial part of the problem is the former.

Remarks:

1. We have to assign values to the variance R of the noise process $e(t)$. Assessing the value of R is not always an easy task. If R is not readily obtainable, it can be estimated using a technique described in [8]. The effect of estimating R via this technique is investigated in [14].

2. We assume that an a priori mean of the vector $\vartheta(0)$ can be set to zero when no knowledge about their values is available before any measurements are taken (we notice that is the most likely case). On the other hand the usual choice of the initial variance of the vector $\vartheta(0)$, denoted by $P(0/0)$ is $P(0/0) = mI$, where m is a large integer.

3. We assume that measurements of $y(t)$, are set to zero for $t < 0$; this technique is well known as "prewindowing".

3 REFORMULATION OF THE AR MODEL ORDER SELECTION AND IDENTIFICATION PROBLEM

Let us now assume that the order n is unknown and what we know is only that the true order satisfies the condition $n_0 \leq n \leq n_{MAX}$. This parameter is assumed to be a random variable with known or assumed a priori pdf $p(n/0) = p(n)$. It is clear that the true model is one of a family of models described by the relations (4) and (7) or (5) and (7), the true model being specified by the true value of the parameter n. The problem is then to select the correct model among various "candidate" models. In other words, we have to design an optimal estimator when uncertainly is incorporated in the signal model.

The solution to this problem for large order AR models which is described in the following, has been given by the FAEST algorithm [5] using the MMP theory [9] - [11]. The MMP algorithm operates on the following discrete model:

$$y(t) = h^\top(t/n)\vartheta(t) + e(t) \tag{7}$$

where

$$h^\top(t) = [y(t-1)y(t-2)\ldots y(t-n)] \tag{8}$$

The time updating of the LS Filter is:

$$\tilde{z}(t/n) = -y(t) + h^\top(t/n)\hat{\vartheta}(t-1/n)$$

$$\epsilon(t/n) = \frac{\tilde{z}(t/n)}{\alpha(t/n)}$$

$$\hat{\vartheta}(t/n) = \hat{\vartheta}(t-1/n) + w_m(t/n)\epsilon(t/n)$$

where $\alpha(t/n)$ is the otimal gain encountered in LS filters and $w_m(t/n)$ is the FAEST gain vector.

The optimal MMSE estimate of $\vartheta(t)$ is given by

$$\hat{\vartheta}(t) = \int_n \hat{\vartheta}(t/n)p(n/t)dn \tag{9}$$

The model-conditional pdf $p(n/t)$ is given by:

$$p(n/t) = \frac{L(t/t;n)}{\int_n L(t/t;n)p(n/t-1)dn}p(n/t-1) \tag{10}$$

where

$$L(t/t;n) = |P_z(t/t-1;n)|^{\frac{-1}{2}} exp\{-\frac{1}{2}\|\tilde{z}(t/n)\|$$

$$*P_z^{-1}(t/t-1;n)\}$$

$$P_z(t/t-1;n) = h^\top(t/n)P(t/t;n)h(t/n) + R$$

$$P(t/t;n) = P(t/t-1;n) - G(t/n)h^\top(t/n)P(t/t-1;n)$$

$$P(t+1/t;n) = P(t/t;n) + QI$$

$$G(t/n) = \frac{P(t/t-1;n)h(t/n)}{[h^\top(t/n)P(t/t-1;n)h(t/n) + R]} \tag{11}$$

Equations (9), (10) pertain to the case where n's pdf is continuous in n. When this is the case, one is faced with the need for a nondenumerable infinity of FAEST filters for the exact realization of the optimal estimator. The usual approximation performed to overcome this difficulty is to somehow approximate n's pdf by a finite sum [17]. Fortunately, in our case the sample space is naturally discrete, so that no approximations are necessary and the estimator is indeed optimal. Since this is the case, the integrals in (9), (10) must be replaced by summations running over all possible values of n. The important feature of the algorithm is that all the filters needed for its implementation can be independently realized. A wide range of applications require high-speed, real-time digital signal processing. Minimization of processing time is efficiently achieved in a parallel processing environment.

Figure 1, shows the new algorithm's implementation. This block diagram emphasizes the ability of implementing the algorithm in parallel, thus saving enormous computational time [12].

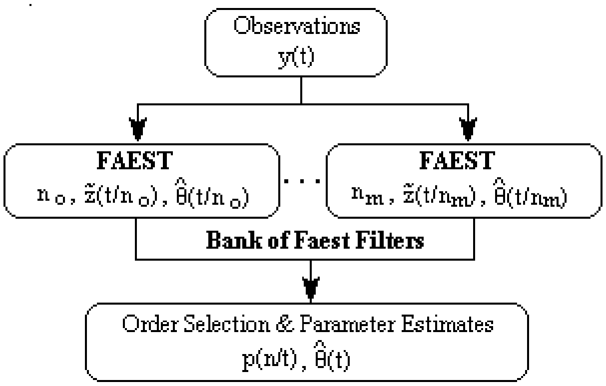

Figure 1: The structure of the proposed method

4 SIMULATION EXAMPLES

The proposed algorithm has been tested extensively on several simulation experiments. In this section two examples are discussed; first we present the case where the true order model belongs to the FAEST filters' bank and second the case where the correct order model satisfies the condition $n_0 \leq n \leq n_{MAX}$ but it does not exists in the filters' bank. To reduce realization dependency of the simulations we averaged over 100 Monte Carlo runs.

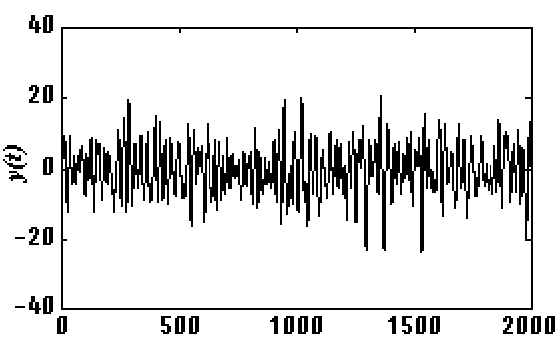

Figure 2: The AR process

The data record is of length 2000 and they are produced from an AR model of order 20 and the following coefficients' values:

$$[\ 1.32,\ 0.68,\ 1.04,\ -0.54,\ 0.25,\ -1.02,\ 0.39,$$

$$-0.22,\ 0.36,\ 0.1,\ 0.17,\ -0.29,\ 0.27,\ -0.35,$$

$$0.011,\ 0.062,\ 0.018,\ -0.047,\ 0.11,\ -0.067\ \ ]$$

Based in this model ($n = 20$), an AR system like that of (5) and (7) was simulated. Figure 2, shows the true AR process.

The added noise $e(t)$ is a zero-mean white noise process with variance R equal unity and the initial value for $P(0/0)$ is $1000I$.

Example 1: We have realized a bank of FAEST filters which contains eight filters of order (5, 10, 15, 20, 25, 30 , 35, 40) respectively.

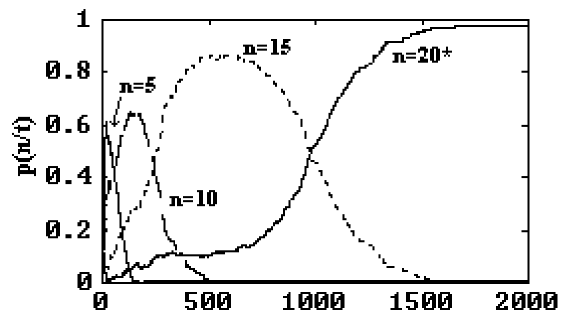

Figure 3: The a posteriori probabilities

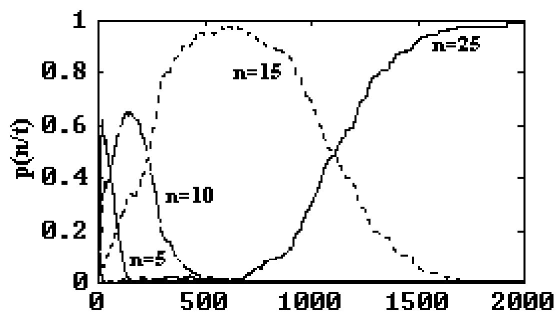

Figure 4: The a posteriori probabilities

Example 2: We have realized a bank of FAEST filters which contains eight filters of order (5, 10, 15, 25, 30, 35, 40, 45) respectively. Notice that the true order model does not belong to the bank of FAEST filters.

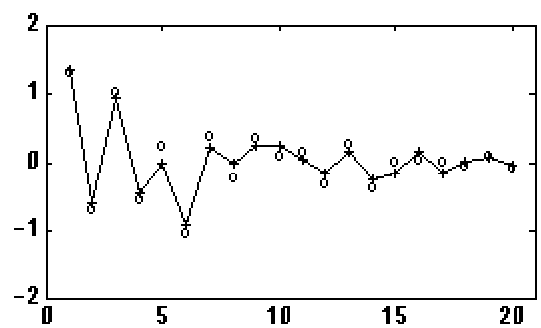

Figure 5: Parameter Estimates (+) - True Values (o)

Figure 3 and figure 4, depict the a posteriori probabilities associated with each value of n, for the examples 1

and 2 respectively.
Figure 5 for the example 1 and figure 6 for the example 2, shows how the true parameter estimates (denoted by $+$) tracks the true values of the parameters (denoted by o).

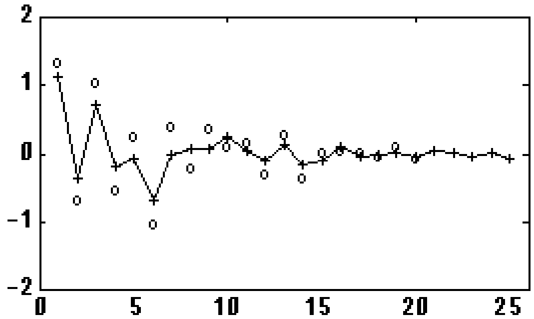

Figure 6: Parameter Estimates (+) - True Values (o)

5 CONCLUSIONS

The experiments, indicate that (a) in the case that the true model is one of the models included in the bank of the FAEST filters the proposed method is selecting the correct model order and identifies accurately the parameters and (b) in the case where the true model is not included to the filters bank, the algorithm converges to the closest model, in our experiment the model with order $n = 25$, by means of the Kullback information criterion minimization [6] and identifies successfuly the parameters as the last five parameters tends to zero.
As shown by the experiment results as the algorithm identifies the true model order, n_{true}, the $p(n_{true}/t)$ tends to unity and $p(n/t)$ tends to zero, for $n \neq n_{true}$. Thus, when the order changes, on line, the algorithm can sense the variation and is adapted to the true order model.
The superiority of this method is that it works in real time and even in the case that the system order is large enough, the algorithm identifies precisely the true order (by means of the MAP criterion) and the parameters in a sufficient number of iterations. Furthermore, the algorithm can be parallely implemented and also a VLSI implementation is feasible.

References

[1] H. Akaike. Fitting autoregressive models for prediction. *Ann. Inst. Of Stat Math*, 21:243-347, 1969.

[2] H. Akaike. Information theory and an extension of the maximum likelihood principle. *In Proc. 2nd Int. Symp. Inform. Theory. Budapest, Hungary: Akademia Kiado*, 267-281, 1973.

[3] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Automat. Conr.*, 26:1-18, 1977.

[4] B.D.O. Anderson and J.B. Moore. Optimal Filtering. *Englewood Cliffs, NJ, Prentice-Hall*, 1979.

[5] George Carayannis, Dimitris G. Manolakis and Nicholas Kalouptsidis. A Fast Sequential Algorithm for Least-Squares Filtering and Prediction. *IEEE Trans. Acoust., Speech, Signal Processing*, Vol ASSP-31, No 6, 1394-1402, 1983.

[6] Richard M. Hawkes and John B. Moore. Performance Bounds for Adaptive Estimation. *IEEE Proceedings*, Vol 64, No 8, 1976.

[7] S.K. Katsikas, S.D. Likothanassis and D.G. Lainiotis. AR model identification with unknown process order. *IEEE Trans. A.S.S.P.*, Vol 38, No 5:872-876, 1990.

[8] S.K. Katsikas. Optimal algorithms for geophysical signal processing. Ph. D. Thesis, University of Patras, Dept. of Computer Engineering and Informatics, Greece, 1986.

[9] D.G. Lainiotis. Optimal adaptive estimation: Structure and parameter adaptation and parameter adaptation. *IEEE Trans. Automat. Contr.*, Vol AC-16:160-170, 1971.

[10] D.G. Lainiotis. Partitioning: A unifing framework for adaptive systems I: Estimation. *Proc. IEEE*, 64:1126-1143, 1976.

[11] D.G. Lainiotis. Partitioning: A unifing framework for adaptive systems II: Estimation. *Proc. IEEE*, 64:1182-1198, 1976.

[12] D.G. Lainiotis, S.K. Katsikas and S.D. Likothanassis. Adaptive deconvolution of seismic signals: Performance, computational analysis, parallelism. *IEEE Trans. Acoust., Speech, Signal Processing*, Vol 36, No. 11:1715-1734, 1988.

[13] Gang Liang, D. Mitchell Wilkes and James A. Cadzow. ARMA Model Order Estimation Based on the Eigenvalues of the Covariance Matrix. *IEEE Trans. On Signal Processing*, Vol 41, No. 10, 1993.

[14] A.K. Mahalanabis and S. Prasad. On the application of the fast kalman algorithm to adaptive deconvolution of seismic data. *IEEE Trans. Geosci. Remote Sensing*, Vol. GE-21, No. 4, 1983.

[15] J. Rissanen. Modeling by shortest data description. *Automatica*, 14:465-471, 1978.

[16] G. Schwarz. Estimation of the dimension of a model. *Ann. Stat.*, 6:461-464, 1978.

[17] R.L. Shengbush and D.G. Lainiotis. Simplified parameter quantization procedure for adaptive estimation. *IEEE Trans. Automat. Contr.*, Vol AC-14:424-425, 1969.