EVOLUTIONARY MULTIMODEL PARTITIONING FILTERS FOR MULTIVARIABLE SYSTEMS*

G.N. Beligiannis[1,2] *K.G. Berketis*[3] *D.A. Fotakis*[1,2] *S.D. Likothanasis*[1,2]

[1] Department of Computer Engineering and Informatics
University of Patras, Rion, 265 00 Patras, GREECE
Tel: +30 61 997755; fax: +30 61 997706

[2] Computer Technology Institute
Kolokotroni 3, 262 21 Patras, GREECE

[3] Department of Mathematics
University of the Aegean, 832 00 Karlovassi, GREECE

e-mail: beligian@ceid.upatras.gr, {fotakis, likothan}@cti.gr

ABSTRACT

It is known that for the adaptive filtering problem, the Multi Model Adaptive Filter (MMAF) based to the Partitioning Theorem is the best solution. It is also known that Genetic Algorithms (GAs) are one of the best methods for searching and optimization. In this work a new method, concerning multivariable systems, which combines the effectiveness of MMAF and GAs' robustness has been developed. Specifically, the a-posteriori probability that a specific model, of the bank of the conditional models, is the true model can be used as fitness function for the GA. Although the parameters' coding is more complicated, simulation results show that the proposed algorithm succeeds better estimation of the unknown parameters compared to the conventional MMAF, even in the case where it is not included in the filters bank. Finally, a variety of defined crossover and mutation operators is investigated in order to accelerate algorithm's convergence.

1 INTRODUCTION

The adaptive filtering problem, with unknown time-invariant or time-varying parameters, has been a central issue in the field of signal processing. A variety of solutions to this problem has been proposed, that approximate the optimal solution, under certain conditions [1]-[3]. Usually, these methods are based on the assumption that the data are Gaussian and they are two-pass methods. Therefore, they cannot be used in an on-line or adaptive fashion. A different adaptive approach, based on the Partitioning Theorem, is the Multi Model Adaptive Filter (MMAF) [3], that operates on general, not necessarily Gaussian data pdf's. The MMAF converges to the optimal solution, if the model supporting the data is included to the filter's bank. Otherwise, it converges to the closer model by mean of the Kullback information criterion minimization. This is due to the fact that the number of the filters used in the MMAF bank is finite. It is clear that the evolution of the population of the filter's bank will improve the filter's performance.

Among the existing adaptive identification methods, we are particularly interested in the partitioned adaptive technique, since it is useful not only for identifying the noise statistics but also for estimating the unknown system parameters. Pioneer work in this area can be found in [3], due to Magill and Lainiotis. It is known that the linear filtering problem with unknown time-invariant or time-varying parameters, i.e. the adaptive filtering problem, reduces to a non-linear filtering problem, which has major difficulties in its realization. In particular, it is extremely difficult to access the effect of approximations made in the suboptimally realization of non-linear filters. However, partitioned adaptive filtering constitutes a partitioning of the original non-linear filters into a bank or set of much simpler linear elemental Kalman or Kalman-Busy filters. This realization is very simple to implement physically.

It is known that Genetic Algorithms (GAs) are one of the best methods for searching and optimization [4, 5]. They apply genetic operators (reproduction, crossover and mutation), in a population of individuals (sets of unknown parameters properly coded), in order to achieve the optimum value of the fitness function. By evolving the best individuals, in each generation, they converge to the (near) optimal solution. The main advantage of the GAs, is that they use the parameter's values, instead

*The authors' names are in alphabetical order.

of the parameters themselves. In this way, they search the whole parameter space.
In this work a new method which combines the effectiveness of MMAF and GAs' robustness has been developed. Specifically, the a-posteriori probability that a specific model, of the bank of the conditional models, is the true model can be used as fitness function for the GA. The solution to this problem was first introduced in [6, 7], where the case of scalar unknown parameters was considered. In this approach we face the problem for multivariable systems. Although the parameters' coding is more complicated, simulation results show that the proposed algorithm succeeds better estimation of the unknown parameters compared to the conventional MMAF. Furthermore, the algorithm identifies the true model in the case where it is not included in the filters bank. Finally, a variety of defined crossover and mutation operators is investigated in order to accelerate the algorithm's convergence.
The paper is organized as follows. In section 2 the problem of the discrete time adaptive estimation and the MMAF are stated. In sections 2.1 and 2.2 the proposed method and the GA structure are presented respectively. Section 3 contains the simulation results, while section 4 summarizes the conclusions.

2 ADAPTIVE ESTIMATION IN DISCRETE TIME

A multivariable discrete time linear system is described by the following vector difference equation:

$$x(k+1) = \Phi(k+1, k; \theta)\, x(k) + B(k, \theta)\, w(k) \tag{1}$$

$$z(k) = H(k, \theta)\, x(k) + v(k) \quad k = 1, 2, \ldots \tag{2}$$

where $x(k)$ is the n-dimensional state vector, $\Phi(\cdot)$ is the $n \times n$ state transition matrix, $B(\cdot)$ is an $n \times p$ matrix, and $\{w(k), k = 1, 2, \ldots\}$ is a p-vector Gaussian white noise sequence, $w(k) \sim N(0, Q(k, \theta))$. $z(k)$ is the m-vector observation, $H(\cdot)$ is an $m \times n$ matrix function, and $\{v(k) k = 1, 2, \ldots\}$ is an m-vector Gaussian white noise sequence, $v(k) \sim N(0, R(k, \theta))$. The distribution of $x(0)$ is Gaussian, $x(0) \sim N(x(0, \theta), P(0, \theta))$, and $x(0)$, $w(k)$ and $v(k)$ are assumed to be mutually independent. The unknown parameter vector θ has a known or assumed a-priori probability $p(\theta)$. The discrete time version of Lainiotis partition theorem is presented in the following theorem [2]:

Given the observation data $Z_k = \{z(l), l = 1, 2, \ldots, k\}$, the MMSE estimate $\hat{x}(k/k)$ of $x(k)$ is given by:

$$\hat{x}(k/k) = \int \hat{x}(k/k, \theta)\, p(\theta, k)\, d\theta \tag{3}$$

where $\hat{x}(k/k) = E[x(k)/Z_k]$ and $\hat{x}(k/k, \theta) \equiv E[x(k)/Z_k, \theta)]$. The a-posteriori probability of θ given Z_k, $p(\theta/Z_k) = p(\theta/k)$, is provided by the relation:

$$p(\theta/k) = \frac{L(k/\theta)\, p(\theta/k-1)}{\int L(k/\theta)\, p(\theta/k-1)\, d\theta} \tag{4}$$

where the $L(k/\theta)$s are the model conditional scalar likelihood. In the case that the values of the vector θ are discrete the integrals in relations (3) and (4) must be replaced by summations running over all possible values of the elements of the vector θ.
Multimodel partitioning algorithms have been shown to converge to the model closer to the true one - in the sense of the Kullback information measure minimization - when the true model is not a member of the space spanned by the filter model [3]. In this case, if a specific estimate of θ is desired, it can be found as a weighted sum of the discrete parameter sets:

$$\hat{\theta} = \sum_{i=1}^{M} \theta_i\, p(\theta_i/Z_k) \tag{5}$$

Thus, we can produce the MMSE estimate $\hat{x}(k/k)$ at the same time as identifying the unknown parameter vector θ.

2.1 Adaptive Estimation In Discrete Time Using MMAF-GAs

In the system model mentioned and described at the previous paragraphs there are two quantities which must be estimated; the state $x(k)$ and the unknown parameter vector θ. The estimation method is the Multi Model Adaptive Filter as it is obtained from Lainiotis partition theorem. Our goal is to achieve the optimal estimation for both estimated quantities and particularly for the unknown parameter vector θ. The only information we have for this vector is that it belongs to a set or a space (finite or infinite) or it has a specific probability distribution (follows specific probabilistic rules-distribution). It is obvious that if the unknown vector belongs to a finite-discrete set with small cardinality, the MMAF is the only appropriate and most effective method to estimate this unknown vector. It is also widely known that GAs perform better when the space which will be searched has a large number of elements. So, we do not need GAs when the unknown vector belongs to a finite-discrete set with small cardinality. GAs can be used when the unknown vector belongs to a space with large cardinality or belongs to an infinite space or follows a probability distribution. Then, we should optimize with GAs the a-posteriori probability function for the various values of the unknown vector θ. That means that we have to optimize the probability described from relation (4), for discrete θ, i.e. the following probability function:

$$p(\theta_i) = \frac{L(k/\theta_i)\, p(\theta_i/k-1)}{\sum_{j=1}^{M} L(k/\theta_j)\, p(\theta_j/k-1)} \tag{6}$$

which will be the fitness function for the GAs, for the several values of the unknown parameter vector θ underlying to the above constraints.

2.2 The GA Structure

The structure of GA who has been developed is described in the following: First we made an initial population of m matrixes each of them contains a possible value of the unknown parameter vector θ. Because the elements of θ belong to an infinite space (even when follow a probability distribution, the possible values are infinite), that consists a small discretesation of this infinite space, which is absolutely reasonable as GAs can search in a continuous space only after a discretesation of the space. For each of these matrixes we apply an MMAF and have as result the a-posteriori probability of the value of the matrix. This is the fitness of each matrix (as mentioned above the fitness function is the a-posteriori probability function). Since we have the fitness of each matrix we are able to perform the other genetic operators, i.e. reproduction, crossover and mutation. The reproduction operator will be the classic biased roulette wheel selection according to the fitness function value of each matrix. As far as crossover is concerned, we will use five crossover operators: Uniform Crossover, Even-Odd Crossover, One-Point Crossover, Arithmetic Crossover and Blend Crossover. Finally, we will use three mutation operators: Flip Mutator, Swap Mutator and Gaussian Mutator. This new generation of matrixes iterates the same process as the old one and all this process may be repeated as many generations as we desire or till the fitness function has value 1 (one) which is the maximum value it is able to have as a probability [6, 7].

3 EXPERIMENTAL RESULTS

The MMAF filters are implemented in MATLAB due to MATLAB's capability of manipulating matrices in a very easy and quickly way while the GA is implemented in an object-oriented enviroment (C++) using GAlib. GAlib is a very powerfull object-oriented programming tool for implementing GAs which has been developed in M.I.T.

The first experiment which has been used to confirm our last assumptions retains to a specific two dimensional system with unknown measurement matrix, which is the following:

$$x(k+1) = \begin{bmatrix} 0.93 & 0 \\ 0 & 0.93 \end{bmatrix} \cdot x(k) + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot w(k)$$

$$z(k) = H(k,\theta) \cdot x(k) + v(k)$$

where $k = 0, 1, \ldots, 30$ and the 2×2 matrix H is the unknown measurement matrix to be estimated, $H[i,j] \in [0,1]$.

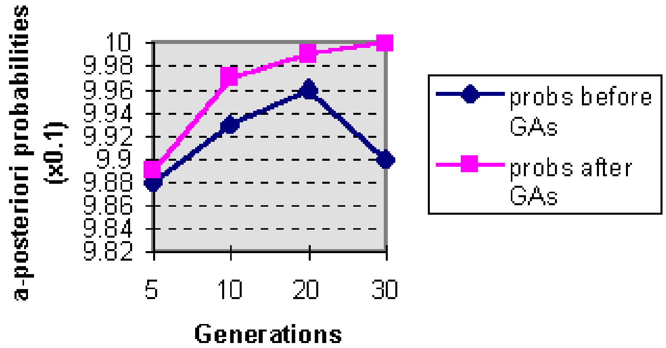

Figure 1: The evolution of the a-posteriori probabilities for the first experiment.

Also, $P(0) = \begin{bmatrix} 0.95 & 0 \\ 0 & 0.95 \end{bmatrix}$, $Q = \begin{bmatrix} 0.8 & 0 \\ 0 & 0.8 \end{bmatrix}$ and $R = \begin{bmatrix} 0.63 & 0 \\ 0 & 0.63 \end{bmatrix}$.

The first experiment's results compare the a-posteriori probability resulted from a single MMAF with the evolution and improvement which appears after some GA's generations. The presented results in Figure 1 show the difference between MAF before performing GAs and after performing GAs.

The second experiment which has been used to confirm our last assumptions retains to a specific two dimensional system with unknown measurement matrix, which is the following:

$$x(k+1) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -0.022 & 0.155 & 0.065 & -0.809 & 1.387 \end{bmatrix}$$

$$\cdot\, x(k) + \begin{bmatrix} 0 & -0.004 & 0 & 0 & 0 \\ 0 & -0.021 & 0 & 0 & 0 \\ 0 & -0.051 & 0 & 0 & 0 \\ 0.001 & -0.07 & 0 & 0 & 0 \\ 0.009 & -0.062 & 0 & 0 & 0 \end{bmatrix} \cdot w(k)$$

$$z(k) = H(k,\theta) \cdot x(k) + v(k)$$

where $k = 0, 1, \ldots, 30$ and the 3×5 matrix H is the unknown measurement matrix to be estimated, $H[i,j] \in [0, -30]$.

Also,

$$P(0) = \begin{bmatrix} 0.0001 & 0 & 0 & 0 & 0 \\ 0 & 0.0001 & 0 & 0 & 0 \\ 0 & 0 & 0.0001 & 0 & 0 \\ 0 & 0 & 0 & 0.0001 & 0 \\ 0 & 0 & 0 & 0 & 0.0001 \end{bmatrix},$$

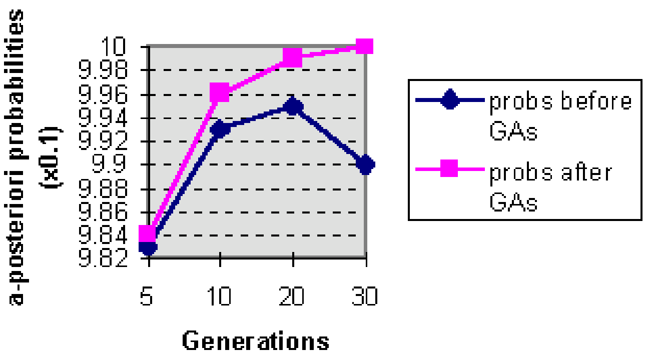

Figure 2: The evolution of the a-posteriori probabilities for the second experiment.

$$Q = \begin{bmatrix} 0.201 & 0 & 0 & 0 & 0 \\ 0 & 0.201 & 0 & 0 & 0 \\ 0 & 0 & 0.201 & 0 & 0 \\ 0 & 0 & 0 & 0.201 & 0 \\ 0 & 0 & 0 & 0 & 0.201 \end{bmatrix} \text{ and}$$

$$R = \begin{bmatrix} 1.026 & 0 & 0 \\ 0 & 6.908 & 0 \\ 0 & 0 & 158.658 \end{bmatrix}.$$

The second experiment's results compare the a-posteriori probability resulted from a single MMAF with the evolution and improvement which appears after some GA's generations. The presented results in Figure 2 show the difference between MMAF before performing GAs and after performing GAs.

Furthermore, according to both experiments' results the following conclusions were come to:

- As the population's size grows, the algorithm converges faster giving an a-posteriori probability higher than 0.999.
- As the crossover's probability grows, the algorithm converges faster giving an a-posteriori probability equivalent to or even higher than 0.999.
- As the mutation's probability grows, the algorithm converges faster giving an a-posteriori probability equivalent to or even higher than 0.999.
- As far as crossover operators are concerned, faster convergence is reached by using the Uniform Crossover and the Arithmetic Crossover operators.
- As far as mutation operators are concerned, faster convergence is reached by using the Flip Mutation and the Gaussian Mutation operators.

4 CONCLUSIONS

In this work a new evolutionary method for adaptive estimation of multivariable discrete time systems, with unknown parameters, has been proposed. The method combines the well known Adaptive Multi Model Partitioning theory with the effectiveness of the GAs. Simulation results show that the method performs significally better than the conventional MMAF. Although the parameter's coding is more complicated a variety of defined crossover and mutation operators was investigated, for the multivariable case, resulting in accelerations of the algorithm's convergence. Furthermore, the evolution of the initial population results to the identification of the true model, even in the case where it is not included in the in the initial population of the filter's bank. Finally, the method can be implemented in a parallel enviroment, since the MMAF as well as the GA are naturally parallel structured, thus increasing the computational speed.

References

[1] JERRY M. MENDEL, "Lessons in Digital Estimation Theory" Prentice-Hall, Englewood Cliffs, N.J., 1987.

[2] BRIAND.O.ANDERSON, JOHN B.MOORE, "Optimal Filtering" Prentice-Hall, Englewood Cliffs, N.J., 1979.

[3] KEIGO WATANABE "Adaptive Estimation and Control. Partitioning Approach" Prentice-Hall, U.K., 1992.

[4] DAVID E. GOLDBERG "Genetic Algorithms in Search, Optimization and Machine Learning" Addison-Wesley, Reading, Mass., 1989.

[5] ZBIGNIEW MICHALEWICZ "Genetic Algorithms + Data Structures = Evolution Programs", Springer Verlang, N.Y., 1992.

[6] K. G. Berketis, S. K. Katsikas and S. D. Likothanassis, "Multimodel Adaptive Filtering with Genetically Determined Varying Model Space", Int. Conference on Digital Signal Processing, DSP'97, 2-4 July, Greece, 1997.

[7] G. Berketis, S.K. Katsikas, and S. D. Likothanassis , "Multimodel Partitioning Filters and Genetic Algorithms", Journal of Non-Linear Analysis, vol.30-4, pp.2421-2428, Dec. 1997.