Blind Channel Identification Based on Higher-Order Cumulant Fitting Using Genetic Algorithms

S. Chen †, Y. Wu † and S. McLaughlin ‡

† Department of Electrical and Electronic Engineering, University of Portsmouth
Anglesea Building, Portsmouth PO1 3DJ, U.K.
Email: schen@ee.port.ac.uk

‡ Department of Electrical Engineering, University of Edinburgh
King's Buildings, Edinburgh EH9 3JL, U.K.

Abstract

A family of blind equalisation algorithms identifies a channel model based on a higher-order cumulant (HOC) fitting approach. Since HOC cost functions are multimodal, gradient search techniques require a good initial estimate to avoid converging to local minima. We present a blind identification scheme which uses genetic algorithms (GAs) to optimise a HOC cost function. Because GAs are efficient global optimal search strategies, the proposed method guarantees to find a global optimal channel estimate. A micro-GA implementation is adopted to further enhance computational efficiency. As is demonstrated in computer simulation, this GA based scheme is robust and accurate, and has a fast convergence performance.

1. Introduction

An important class of blind equalisation algorithms uses techniques based on HOCs [1]–[4]. A two-stage strategy is usually adopted, which first identifies a channel model using HOC fitting algorithms and then employs the estimated channel model to design an equaliser. The key step of this approach is its ability to obtain an accurate channel model. Once a channel model is available, a variety of existing equaliser design methods can be employed, ranging from simple linear inverse filter to sophisticated maximum likelihood sequence estimator, depending on a trade-off between performance and complexity. Therefore, we concentrate on blind channel identification using the HOC fitting approach in this paper.

HOC fitting cost functions are however multimodal, and conventional gradient techniques [3],[4] may converge to "wrong" solutions unless a good initial value for the channel parameters is provided, which is not always possible. To overcome the problem of local minima, simulated annealing has been implemented to optimize a HOC cost function [5]. We propose to use GAs [6]–[9] for blind channel identification based on HOC fitting. This GA based scheme is very robust and achieves a global optimal solution regardless of initial value of channel estimate. Furthermore, the number of parameters to be optimised in the problem of blind channel estimation is usually small, and GAs are particularly effective to solve this kind of optimisation problems. The micro-GA implementation [8] is adopted to further improve convergence rate.

The channel is modelled as a finite impulse response filter with an additive Gaussian white noise:

$$r(k) = \sum_{i=0}^{n_a} a_i s(k-i) + e(k) \quad (1)$$

Blind channel identification refers to the determination of the channel model $\mathbf{a} = [a_0\, a_1 \cdots a_{n_a}]^T$ using only the noisy received signal $\{r(k)\}$ and some knowledge of statistic properties of $s(k)$. We will assume a real-valued channel and a PAM symbol constellation. Extension to complex-valued channels and modulation schemes is straightforward.

0-8186-8005-9/97 $10.00 © 1997 IEEE

2. Higher-order cumulant fitting

Since $3rd$-order cumulants of $r(k)$ are zero, $4th$-order cumulants have to be used [10]. The $4th$-order cumulant fitting cost function adopted in our application is defined by:

$$J_4(\hat{\mathbf{a}}) =$$

$$\sum_{\tau=-\hat{n}_a}^{\hat{n}_a} \left(\hat{C}_{4,r}(\tau,\tau,\tau) - \gamma_{4,s} \sum_{i=\max\{0,-\tau\}}^{\min\{\hat{n}_a,\hat{n}_a-\tau\}} \hat{a}_i \hat{a}_{i+\tau}^3 \right)^2 \quad (2)$$

where $\hat{C}_{4,r}(\tau,\tau,\tau)$ is the diagonal slice of the time estimate $\hat{C}_{4,r}(\tau_1,\tau_2,\tau_3)$ of the $4th$-order cumulant, $\gamma_{4,s}$ the kurtosis of $s(k)$, $\hat{n}_a$ an estimated channel length and $\hat{\mathbf{a}}$ the channel estimate.

Most of the existing algorithms for HOC fitting employ gradient search techniques. HOC fitting cost functions are well-known to be multimodal, and gradient based optimisation methods may fail to work. Even with measures of providing good initial channel estimate, it has been observed that gradient algorithms sometimes converge to local minima [3]. Using GAs to optimize the cost function (2) has the advantage of guaranteeing to find a global optimal channel estimate.

A channel model (1) typically contains a few taps. Thus the number of parameters to be optimized in (2) is small, and GAs are efficient in solving this kind of "small-dimensional" optimisation problems. In reality, the channel order n_a is unknown. A simple method is to overfit with $\hat{n}_a \geq n_a$. Although this will complicate the cost function and may cause problems to gradient based methods, the GA based method is capable of identifying those nonexisting taps with (near) zero values. An inspection of the obtained channel estimate allows us to delete those insignificant taps.

3. Genetic algorithms

GAs belong to a problem solving approach based loosely on the evolution of species in nature. They differ from gradient optimisation techniques mainly in four aspects [7]. Firstly, GAs work with an encoding of the parameter set to be searched, not the parameters themselves. Secondly, unlike gradient techniques, which concentrate their efforts on a single potential solution in the search space, GAs search with a population of potential solutions. Thirdly, GAs use the value of the objective function (termed fitness), not derivatives, to evaluate potential solutions. Lastly, GAs use probabilistic transition rules. The seemingly undirected search is guided by the fitness value of each individual and how it compares with others.

A popular encoding scheme is the bit-string encoding [6], which is adopted in our application. A simple GA usually consists of three operations, namely selection, crossover and mutation [7], at each cycle or generation. An "elitist" strategy, which automatically copies a few of the best solutions in the population into the next genneration, is often incorporated. In crossover operation, we adopt multiple crossover points [7], and the number of crossover points in our application is 4.

Since our goal is to find a global optimum solution quickly, the micro-GA [8] offers certain advantages. The population size in a micro-GA is much smaller than those in "standard" GAs. This feature of the micro-GA not only makes it particularly useful for nonstationary optimisation problems but also improves convergence rate in general [8]. Simply adopting a very small population size and letting the search converge just once is not very useful apart from quickly allocating some local optimum. In a micro-GA, after the search has converged, the population is reinitialised with random values while the best individual found so far is copied to the newly generated population. The reinitialisation is repeated until no further improvement can be achieved.

In general, the more complex the search space is, the larger the population size should be. The population size in our micro-GA is 2 times of the number of parameters. The crossover rate is set to 1.0 to facilitate a high rate of information exchange while the mutation rate is set to 0.0 as the reinitialisation of the population will keep the diversity of potential solutions fairly well. Most of GAs adopt the proportional selection [7]. Due to small population size of the micro-GA, the tournament selection is used in choosing parents [8].

4. GA based cumulant fitting scheme

Our GA based scheme is summarized as follows:

Step 1. Given a set of $\{r(k)\}_{k=1}^{N}$, assume an over-length $\hat{n}_a$ and compute the time estimate of the required cumulant to form the cost function (2).

Step 2. With a set of initial channel parameter vectors $\{\hat{\mathbf{a}}_i\}_{i=1}^{n_p}$, where n_p is the population size, use the micro-GA to optimise the cost function (2).

The fitness function value f_i for $\hat{\mathbf{a}}_i$ is defined as

$$f_i = \frac{1}{J_4(\hat{\mathbf{a}}_i)} \tag{3}$$

A measure is used to ensure that each $\hat{\mathbf{a}}_i$ satisfies

$$\sum_{j=0}^{\hat{n}_a} \hat{a}_{i,j}^4 = \frac{\hat{C}_{4,r}(0,0,0)}{\gamma_{4,s}} \tag{4}$$

In the population initialisation, the taps of each $\hat{\mathbf{a}}_i$ first take values randomly from $(-1,\ 1)$, and each chosen $\hat{\mathbf{a}}_i$ is normalised using (4). When a new population is produced, each member is also normalised. This ensures that the population is inside the feasible set of channel models and has an effect of improving convergence rate.

It is well-known that sign and time-shift ambiguities exist in blind identification. Sign ambiguity is due to the fact that both $\mathbf{a}$ and $-\mathbf{a}$ are global optimal solutions of (2). Time-shift ambiguity can be illustrated as follows. Let the true channel be $[a_0 \cdots a_{n_a}]^T$, $a_0 \neq 0$ and $a_{n_a} \neq 0$. Let $\hat{n}_a = n_a + 2$. Then $[a_0 \cdots a_{n_a}\ 0\ 0]^T$, $[0\ a_0 \cdots a_{n_a}\ 0]^T$ and $[0\ 0\ a_0 \cdots a_{n_a}]^T$ are all global optimal solutions. A solution to time-shift ambiguity is to fix the first tap. We do not fix a tap but use the following measure. If the first tap of a member is zero (absolute value smaller than a threshold), a shifting is performed to ensure that the first tap is always nonzero. Complexity of the GA based scheme is determined by the number of cost-function evaluations. The micro-GA employed is specifically designed to minimize this complexity.

5. Simulation results

The two channels used in the simulation were:

Channel 1 $\mathbf{a} = [-0.21\ -0.50\ 0.72\ 0.36\ 0.21]^T$
Channel 2 $\mathbf{a} = [0.227\ 0.460\ 0.688\ 0.460\ 0.227]^T$

The performance of the algorithm was assessed through the best cost function value $J_4(\tilde{\mathbf{a}})$, where $\tilde{\mathbf{a}}$ was the best channel model in the population, and the mean tap error (MTE) defined by:

$$\text{MTE} = \| \pm \tilde{\mathbf{a}} - \mathbf{a} \|^2 = \sum_{i=0}^{\hat{n}_a} (\pm \tilde{a}_i - a_i)^2 \tag{5}$$

In the expression (5), $-\tilde{\mathbf{a}}$ is used if $\tilde{\mathbf{a}}$ converges to $-\mathbf{a}$. Otherwise $\tilde{\mathbf{a}}$ is used.

8-PAM data symbols were transmitted and 50000 noisy received data samples were used to compute the time estimate of the $4th$-order cumulant. All the results were averaged over 100 different runs. Figs. 1 and 2 depict evolutions of the cost function and the MTE for channel 1 with different signal-to-noise ratio (SNR) conditions and assumed channel lengths $\hat{n}_a$, respectively. Table 1 summarises the results (mean$\pm$variance) for channel 1 with a SNR of 20 dB. Results for channel 2 are similarly given in Figs. 3 and 4 and table 2.

Some observations can be drawn. The GA based scheme always finds a global optimal channel estimate and the optimisation process converges fast. Compared with other existing methods of HOC fitting, our method is more accurate and robust, as is demonstrated by very small variances of estimated channel taps. Our method is capable of identifying nonexisting channel taps with (near) zero values (at least an order smaller than values for existing taps). Thus, model order selection can simply be carried out by first assuming an overlength $\hat{n}_a$ and then inspecting the obtained channel estimate to delete those insignificant taps. We also performed a range of simulation with a data length of 25000 samples and 16-PAM symbols. The results, not shown here, also confirm the above observations.

6. Conclusions

A GA based scheme has been developed for blind channel identification based on HOC fitting. Apart from ensuring a global optimal channel estimate regardless of initial conditions, the proposed method is highly accurate and very robust. Small variances of channel estimates and insensitivity to noise achieved in our simulation was not reported previously in other existing methods. Our application also demonstrates advantages of using the micro-GA for fast global optimisation of multi-modal cost functions.

References

[1] K.S. Lii and M. Rosenblatt, "Deconvolution and estimation of transfer function phase and coefficients for non-Gaussian linear processes," *Ann. Statist.*, Vol.10, pp.1195–1208, 1982.

[2] G.B. Giannakis and J.M. Mendel, "Identification of nonminimum phase system using higher order statistics," *IEEE Trans. Acoustics, Speech and Signal Processing*, Vol.ASSP-37, pp.360–377, 1989.

[3] F.-C. Zheng, S. McLaughlin and B. Mulgrew, "Blind equalization of nonminimum phase channels: higher order cumulant based algorithm," *IEEE Trans. Signal Processing*, Vol.41, pp.681–691, 1993.

[4] J.K. Tugnait, "Blind equalization and estimation of digital communication FIR channels using cumulant matching," *IEEE Trans. Communications*, Vol.43, pp.1240–1245, 1995.

[5] J. Ilow, D. Hatzinakos and A.N. Venetsanopoulos, "Blind equalizers with simulated annealing optimization for digital communication systems," *Int. J. Adaptive Control and Signal Processing*, Vol.8, pp.501–522, 1994.

[6] J.H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI.: University of Michigan Press, 1975.

[7] D.E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA.: Addison-Wesley, 1989.

[8] K. Krishnakumar, "Micro-genetic algorithms for stationary and non-stationary function optimization," in *Proc. SPIE Intell. Cont. Adapt. Syst.*, 1989, Vol.1196, pp.289–296.

[9] L. Yao and W.A. Sethares, "Nonlinear parameter estimation via the genetic algorithm," *IEEE Trans. Signal Processing*, Vol.42, pp.927–935, 1994.

[10] J.M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proc. IEEE*, Vol.79, pp.278–305, 1991.

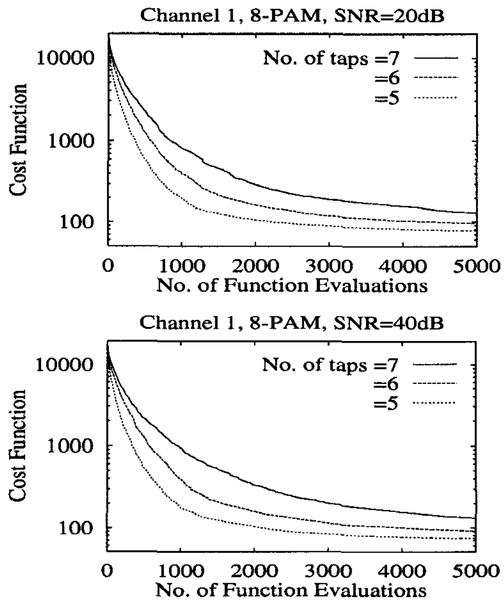

Figure 1. Cost function versus number of function evaluations for channel 1.

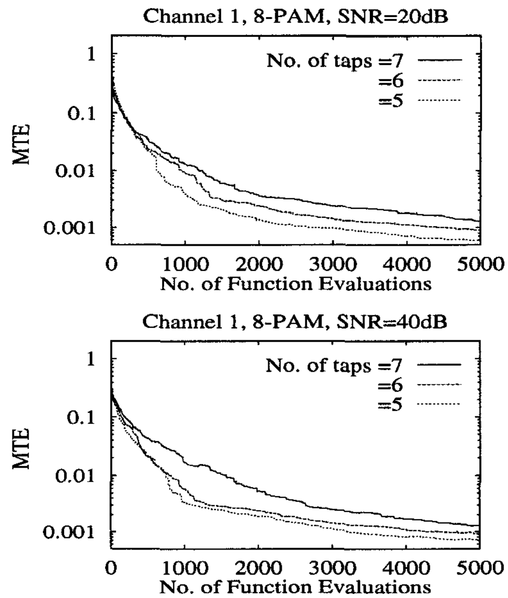

Figure 2. Mean tap error versus number of function evaluations for channel 1.

	true	estimate (mean±variance)		
	$n_a = 4$	$\hat{n}_a = 4$	$\hat{n}_a = 5$	$\hat{n}_a = 6$
a_0	-0.21	-0.20975±0.00014	-0.21124±0.00022	-0.21098±0.00020
a_1	-0.50	-0.49957±0.00003	-0.49975±0.00004	-0.49835±0.00005
a_2	0.72	0.72092±0.00002	0.71965±0.00002	0.72087±0.00003
a_3	0.36	0.35788±0.00006	0.35859±0.00010	0.35823±0.00011
a_4	0.21	0.21073±0.00016	0.21058±0.00013	0.20870±0.00017
a_5	—	—	-0.00041±0.00050	-0.00184±0.00050
a_6	—	—	—	-0.00173±0.00055

Table 1. Identification results for channel 1 with 8-PAM and SNR=20dB.

	true	estimate (mean±variance)		
	$n_a = 4$	$\hat{n}_a = 4$	$\hat{n}_a = 5$	$\hat{n}_a = 6$
a_0	0.227	0.22731±0.00080	0.22221±0.00099	0.21743±0.00119
a_1	0.460	0.45727±0.00097	0.45679±0.00127	0.45481±0.00242
a_2	0.688	0.68913±0.00070	0.68095±0.00066	0.67408±0.00181
a_3	0.460	0.45744±0.00052	0.46426±0.00077	0.46870±0.00262
a_4	0.227	0.22564±0.00079	0.22322±0.00074	0.22265±0.00217
a_5	—	—	0.01577±0.00781	0.01514±0.00653
a_6	—	—	—	0.00359±0.00485

Table 2. Identification results for channel 2 with 8-PAM and SNR=20dB.

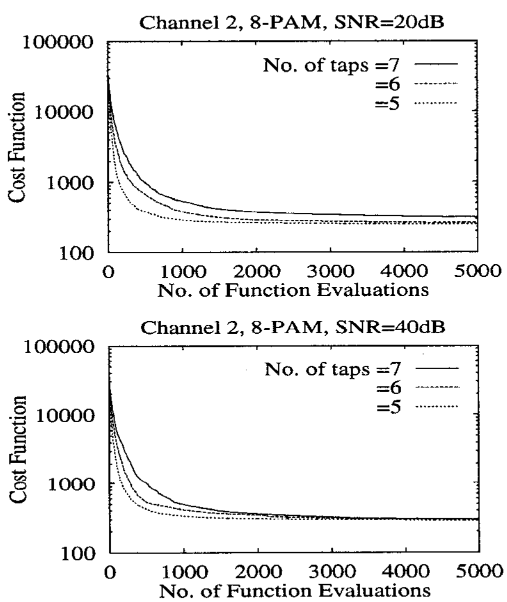

Figure 3. Cost function versus number of function evaluations for channel 2.

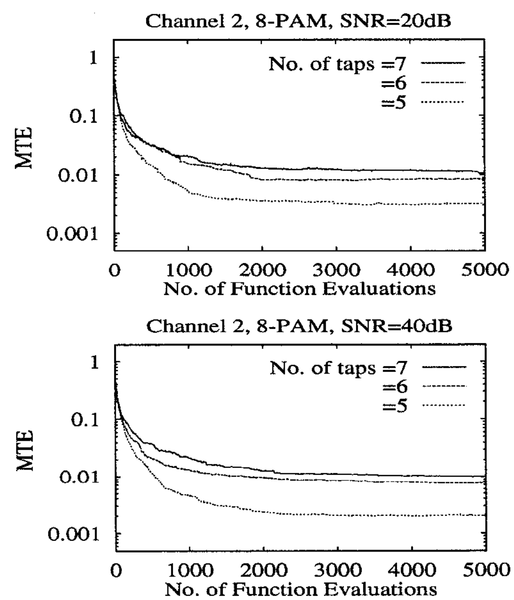

Figure 4. Mean tap error versus number of function evaluations for channel 2.