GENETIC ALGORITHM OPTIMISATION FOR MAXIMUM LIKELIHOOD JOINT CHANNEL AND DATA ESTIMATION

S. Chen and Y. Wu

Department of Electrical and Electronic Engineering
University of Portsmouth, Anglesea Building
Portsmouth PO1 3DJ, U.K.

ABSTRACT

A novel blind equalisation scheme is developed based on maximum likelihood (ML) joint channel and data estimation. In this scheme, the joint ML optimisation is decomposed into a two-level optimisation loop. An efficient version of genetic algorithms (GAs), known as a micro GA, is employed at the upper level to identify the unknown channel model and the Viterbi algorithm (VA) is used at the lower level to provide the maximum likelihood sequence estimation of the transmitted data sequence. The proposed GA based scheme is accurate and robust, and has a fast convergence rate, as is demonstrated in simulation.

1 INTRODUCTION

This paper considers blind equalisation based on the approach of ML joint channel and data estimation. When both the channel and transmitted data sequence are unknown, in theory, their optimal estimates can be obtained via the ML optimisation over channel and data jointly. The computational requirement of such a joint optimisation is, however, prohibitively large. In practice, approximations are adopted. A straightforward way is to employ a batch iterative process between data decoding and channel estimation [1]. Seshadri [2] presented a recursive algorithm for performing joint channel and data estimation. This algorithm may be viewed as an "enhanced" VA that retains several surviving sequences and associated channel estimates for each state of the trellis. The quantized channel algorithm [3] is a batch procedure that maintains a family of candidate channels with discrete parameters. Each channel model is used by the VA to decode data, and the algorithm selects the most likely quantized channel. The present study proposes a novel scheme for joint channel and data estimation using GAs [4]–[7].

We show that GAs are ideal for performing a ML joint channel and data estimation when combining with the VA. A two-layer strategy is suggested. At the top layer, a micro GA [6] searches the channel parameter space to optimise the ML criterion. The bottom layer consists of a number of VA units, one for each member of the channel population provided by the GA. Each VA unit decodes data based on the given channel model and feeds back the corresponding likelihood metric to the GA. Compared with other existing methods for joint channel and data estimation, the GA based scheme is more accurate. Simulation results also demonstrate that the GA based method is robust and has a fast convergence rate in terms of the total number of VA evaluations.

2 MAXIMUM LIKELIHOOD BLIND EQUALISATION

The channel is modelled as a finite impulse response filter with an additive noise source. Specifically, the received signal at sample k is given by

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

where n_a is the channel length, a_i are the channel taps, $e(k)$ is a Gaussian white noise with variance σ_e^2, and the symbol sequence $\{s(k)\}$ is independent. We will assume that the M-PAM scheme is used. The signal to noise ratio (SNR) of the system is defined as

$$\text{SNR} = \sigma_s^2 \left(\sum_{i=0}^{n_a-1} a_i^2 \right) / \sigma_e^2 \tag{2}$$

where σ_s^2 is the symbol variance.

When the channel is unknown and no training sequence is available, joint channel estimation and data detection can be performed based on the ML criterion. Let

$$\begin{aligned} \mathbf{r} &= [r(1)\; r(2) \cdots r(N)]^T \\ \mathbf{s} &= [s(-n_a+2) \cdots s(0)\; s(1) \cdots s(N)]^T \\ \mathbf{a} &= [a_0\; a_1 \cdots a_{n_a-1}]^T \end{aligned} \tag{3}$$

0-7803-4428-6/98 $10.00 © 1998 IEEE

be the vector of N received data samples, the transmitted data sequence and the channel tap vector, respectively. The probability density function of $\mathbf{r}$ conditioned on $\mathbf{a}$ and $\mathbf{s}$ is

$$p(\mathbf{r}|\mathbf{a},\mathbf{s}) = \frac{1}{(2\pi\sigma_e^2)^{N/2}} \times \exp\left(-\frac{1}{2\sigma_e^2}\sum_{k=1}^{N}\left(r(k) - \sum_{i=0}^{n_a-1} a_i s(k-i)\right)^2\right) \quad (4)$$

The joint ML estimate of $\mathbf{a}$ and $\mathbf{s}$ is obtained by maximizing $p(\mathbf{r}|\mathbf{a},\mathbf{s})$ over $\mathbf{a}$ and $\mathbf{s}$ jointly. Equivalently, the ML solution is the minimum of the cost function

$$J(\mathbf{a},\mathbf{s}) = \sum_{k=1}^{N}\left(r(k) - \sum_{i=0}^{n_a-1} a_i s(k-i)\right)^2, \quad (5)$$

that is,

$$(\mathbf{a}^*,\mathbf{s}^*) = \arg\left[\min_{\mathbf{a},\,\mathbf{s}} J(\mathbf{a},\mathbf{s})\right] \quad (6)$$

In theory, $(\mathbf{a}^*,\mathbf{s}^*)$ can be obtained. However, such an optimal solution is certainly too expensive to compute except for the simplest case. In practice, suboptimal solutions are adopted for computational efficiency.

The joint minimisation process (6) can be performed using an iterative loop first over the data sequences $\mathbf{s}$ and then over all the possible channels $\mathbf{a}$

$$(\mathbf{a}^*,\mathbf{s}^*) = \arg\left[\min_{\mathbf{a}}\left(\min_{\mathbf{s}} J(\mathbf{a},\mathbf{s})\right)\right] \quad (7)$$

The inner optimisation can be carried out using the VA. In order to obtain the true optimal solution, the outer optimisation must be performed over all the possible channels $\mathbf{a}$, the complexity of which is generally prohibitive. Usually, suboptimal solutions are sought by constraining the search to a finite set. For example, the quantized channel algorithm [3] uses a family of 2^{n_a} quantized channels. GAs are natural choices for performing the outer optimisation in (7).

3 GENETIC ALGORITHMS

The first step in applying GAs is to encode the parameters to be optimised. We use the popular binary encoding [4]. A simple GA usually consists of three operations, namely selection, crossover and mutation [5], at each cycle. An "elitist" strategy [7], which automatically copies a few of the best solutions in the population into the next generation, is often incorporated. In the crossover operation, we adopt multiple crossover points [5], and the number of crossover points in our application is equal to the number of the parameters.

The version of GA adopted is the micro GA [6]. The population size used in a micro GA is much smaller than those used in "standard" GAs. Simply adopting a very small population size and letting the search converge just once, however, is not very useful apart from quickly allocating some local optimum. Therefore, in a micro GA, after the search has converged, the population is reinitialised with random values while the best individual found so far is automatically copied to the newly generated population. The reinitialisation is repeated until no further improvement can be achieved.

A population size of 5 was suggested in [6] for the micro GA. Generally, however, the more complex the search space is, the larger the population size should be. In our application, the population size n_p is given by $n_p = 5 \times n_a$. This is still considerably smaller than a typical population size used by standard GAs. In our implementation, the crossover rate is set to 1.0, and the mutation rate is set to 0.0 (no mutation) as the reinitialisation of the population will keep the diversity of potential solutions fairly well. Due to the small population size of micro GA, the tournament selection [5],[6] is used in choosing parents for reproduction.

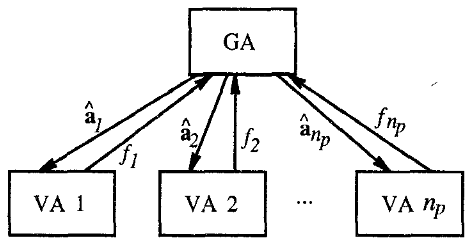

Figure 1: GA based scheme for joint channel and data estimation.

4 THE PROPOSED GA BASED SCHEME

The proposed scheme is depicted in Fig 1. The operations of the algorithm involves an initialisation phase and two loops. In the initialisation, a set of channel vectors $\{\hat{\mathbf{a}}_i\}_{i=1}^{n_p}$ is randomly chosen. The inner loop is summarized as follows:

Step 1. For $1 \le i \le n_p$, the i-th VA unit decodes data based on the given $\hat{\mathbf{a}}_i$, and feeds back the likelihood metric associated with the detected sequence, which is the fitness function value f_i corresponding to $\hat{\mathbf{a}}_i$.

Step 2. If the convergence test for the current population is satisfied, the inner loop is terminated. Otherwise, a new generation of $\{\hat{\mathbf{a}}_i\}_{i=1}^{n_p}$ is generated, and the algorithm goes back to step 1.

After the inner loop has converged, the population is reinitialized, and the inner loop restarts. If the best solutions found after two consecutive reinitializations remain unchanged, the outer loop is terminated. The channel length n_a is assumed to be known but the effects of incorrect channel length will be investigated. It is assumed that the channel is normalised. This assumption is realistic since the channel energy can always be estimated. The search range for each parameter is therefore $(-1,\ 1)$.

The quantized channel algorithm [3] has a similar form to the GA based method in the sense that it also employs a family of channel models. Our GA based method has a faster convergence rate in terms of the total number of VA evaluations. Seshadri's algorithm [2] is widely regarded as one of the best approaches for joint channel and data estimation. It is a recursive algorithm and has considerable computational advantages. Our GA method, however, is much more accurate, as will be demonstrated in the simulation study.

5 SIMULATION STUDY

Simulation was conducted to test the proposed GA scheme using two channels with the impulse response:

Channel 1 $\quad \mathbf{a} = [0.407\ 0.815\ 0.407]^T$

Channel 2 $\quad \mathbf{a} = [0.227\ 0.460\ 0.688\ 0.460\ 0.227]^T$

In practice, the performance of the algorithm can only be observed through the estimated mean square error (MSE). In simulation, the performnance can also be assessed by the estimated mean tap error (MTE):

$$\text{MTE} = ||\tilde{\mathbf{a}} - \mathbf{a}||^2 \tag{8}$$

where $\tilde{\mathbf{a}}$ is the best channel model in the population.

Figs. 2 to 5 depict the MTE performance versus the number of VA evaluations for 2-PAM and 8-PAM symbols, respectively. These results were obtained assuming correct n_a and were averaged over 100 different runs. Compared with the results of using the quantized channel algorithm given in [3], the GA based scheme required a smaller number of VA evaluations to achieve a same level of performance. The final results obtained by the GA based method were also more accurate.

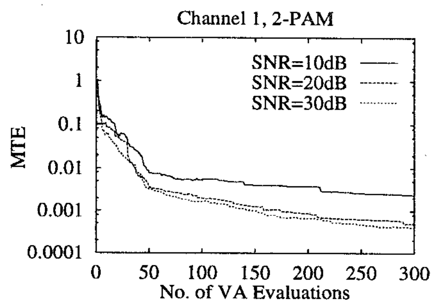

Figure 2: MTE as a function of VA evaluations averaged over 100 runs. 50 data samples used.

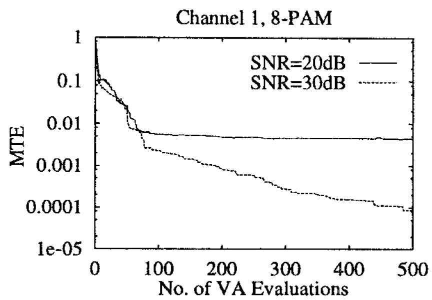

Figure 3: MTE as a function of VA evaluations averaged over 100 runs. 100 data samples used.

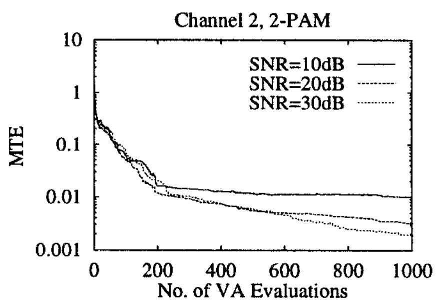

Figure 4: MTE as a function of VA evaluations averaged over 100 runs. 100 data samples used.

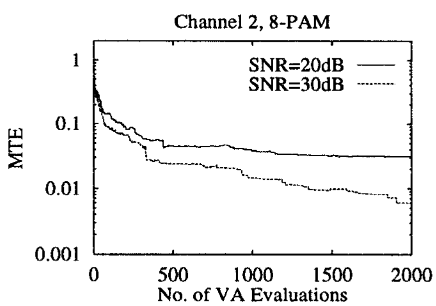

Figure 5: MTE as a function of VA evaluations averaged over 100 runs. 200 data samples used.

Table 1 shows the means and variances of the MTE over 100 runs for channel 1. It can be seen that convergence of the GA based scheme is consistent as is evident from the very small estimation variances. Tables 2 and 3 compare the MTEs and the numbers of received data samples used for the GA method and Seshadri's algorithm. The results of Seshadri's algorithm were estimated from the graphs in [2]. Our GA method is clearly much more accurate. This advantage is obtained at the cost of computational complexity.

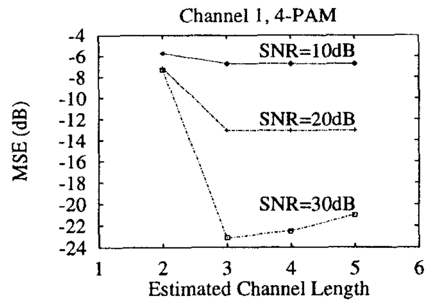

Figure 6: MSE as a function of estimated channel length averaged over 100runs.

In reality the channel length is unknown and has to be estimated. A simple solution is to run the GA based method with a set of different lengths. Fig. 6 illustrates the MSE performance versus the estimated channel length. As expected, when the estimated channel length is correct, the MSE curve achieves the minimum. In this way, the correct channel length can be identified.

6 CONCLUSIONS

A GA based method has been developed for blind equalisation based on ML joint channel and data estimation. Compared with the quantized channel approach, the GA based scheme is more accurate and computationally more efficient in terms of the total number of VA evaluations. As is demonstrated in the simulation study, the GA based scheme requires less received data samples and is more accurate, compared with the best recursive blind trellis search technique. This better performance is, however, obtained at the expense of computational complexity. Simulation results have also shown that the GA method converges consistently with very small estimation variances.

References

[1] M. Ghosh and C.L. Weber, "Maximum-likelihood blind equalization," in *Proc. SPIE* (San Diego), 1991, Vol.1565, pp.188–195.

[2] N. Seshadri, "Joint data and channel estimation using blind trellis search techniques," *IEEE Trans. Communications*, Vol.42, No.2/3/4, pp.1000–1011, 1994.

[3] E. Zervas, J. Proakis and V. Eyuboglu, "A quantized channel approach to blind equalization," in *Proc. ICC'92* (Chicago), 1992, Vol.3, pp.351.8.1–351.8.5.

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

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

[6] 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.

[7] L. Yao and W.A. Sethares, "Nonlinear parameter estimation via the genetic algorithm," *IEEE Trans. Signal Processing*, Vol.42, No.4, pp.927–935, 1994.

	SNR	MTE
2-PAM	10 dB	$1.70 \times 10^{-3} \pm 3.97 \times 10^{-6}$
300 VA calls	20 dB	$3.47 \times 10^{-4} \pm 1.99 \times 10^{-7}$
	30 dB	$1.05 \times 10^{-4} \pm 2.62 \times 10^{-8}$
8-PAM	20 dB	$2.08 \times 10^{-3} \pm 7.96 \times 10^{-6}$
500 VA calls	30 dB	$9.50 \times 10^{-5} \pm 4.16 \times 10^{-7}$

Table 1: Results (means±variances) for channel 1 averaged over 100 runs.

channel	Seshadri		GA scheme	
	MTE	N	MTE	N
1	0.02	100	0.003	50
2	0.08	100	0.01	100

Table 2: Performance comparison. 2-PAM and SNR=10 dB.

channel	Seshadri		GA scheme	
	MTE	N	MTE	N
1	0.03	800	0.0001	100
2	0.02	800	0.006	200

Table 3: Performance comparison. 8-PAM and SNR=30 dB.