International Telecommunications Symposium – ITS2002, Natal, Brazil

A New Method for Blind Channel Identification with Genetic Algorithms

Alysson Machado Costa [1], Romis Ribeiro de Faissol Attux [2], João Marcos Travassos Romano[2]

1 – DENSIS – FEEC - Unicamp
2 - DSPCom – DECOM – FEEC – Unicamp
alysson@densis.fee.unicamp.br, {romisri, romano}@decom.fee.unicamp.br

Abstract – **The first objective of this work is to analyze the suitability of a framework based on the use of genetic algorithms (GAs) to optimize blind equalization cost functions. This effort is motivated by advances in the field of GAs, as well as by recent breakthroughs in the study of blind techniques and their relations to the Wiener criterion. The second objective is to present and test a new proposal for blind channel identification. The method, whose success depends highly on the equalizer quality, was implemented through the established framework. The obtained results show that GAs are a valuable tool that can decisively aid filter adaptation and also demonstrate the efficiency of the identification proposal.**

I. INTRODUCTION

In the last decades, various works started to scrutinize the foundations of blind adaptive filtering theory. They can be generally divided in three classes: new proposals [1]; criteria / algorithms theoretical analysis [2, 3]; and investigations on the relation between different techniques [4].

Several conclusions were drawn from this great intellectual effort. Among them, we highlight here an important one: there is a strong correspondence between blind criteria minima and certain Wiener solutions [3, 4]. Since the Wiener filter formulation expresses in a straightforward manner the ideal of equalization, it provides a very natural basis for comparison.

So, if one can obtain good estimates of the best Wiener receivers *without the need of supervision*¸ then a powerful ground for equalizer adaptation will be established. However, this requires effective search methods for blind cost functions, since they possess different minimum points.

Following recent trends in the constant modulus (CM) criterion analysis, we consider that a good search method can provide good global convergence rates, with corresponding low value of MSE, without a supervision procedure. We chose a genetic algorithm (GA) to perform this task, inspired by its efficiency and by important advances and new conceptions in the field. We also propose a novel equalization-based method for blind channel identification, which is also robust to the unknown channel order.

The paper is divided as follows: **section 2** brings an overview on adaptive equalization to well pose the main concepts to be used in the sequel, including the classical CM criterion. In **section 3**, we propose the new channel identification method, to further implemented in a framework based on the use of GAs. Then **section 4** discusses the genetic algorithm principles and its implementation issues. **Section 5** brings the results while **section 6** states the conclusion and final remarks.

II. BACKGROUND ON ADAPTIVE EQUALIZATION

The goal of communications engineering is to assure proper message interchange between transmitter and receiver**,** by means of a given channel. By nature, such channel generally brings additive noise and distortion. To prevent unacceptable levels of distortion, one may employ an equalizer, a countermeasure device that filters the received signal, in order to recover the desired information. After these preliminary statements, a simple system model is represented by Fig. 1.

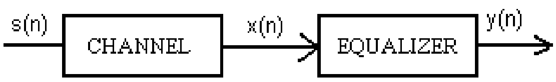

Fig. 1. A Simple Communication System Model

So, it becomes clear that the main objective of equalization is to provide an output signal mathematically represented by:

$$y(n) = K.s(n-d) \quad (1)$$

This establishes the so-called *zero-forcing (ZF) condition*, where K is a gain factor and *d* is the equalization delay. In most practical applications, linear filters play the equalizer role, so that:

$$y(n) = \mathbf{w}^T.\mathbf{x}(n) \quad (2)$$

where **w** is the equalizer parameter vector and **x**(n) = [x(n) x(n-1) ... x(n-L+1)]^T, where L is the equalizer length.

Then a crucial question remains: how to adjust the equalizer parameters **w** in order to obtain a condition as close as possible to (1)? If it is possible to count on *a priori* knowledge of the channel impulse response, the mathematical becomes rather simple. If not, a suitable optimization criterion must be employed for **w** determination. A very important one, that does not require channel information, is the *Wiener criterion* :

$$J_W = E\{[s(n-d) - y(n)]^2\} \qquad (3)$$

The major drawback is now the need of s(n) knowledge, since it is generally unavailable in communications applications.

The usual optimization procedure is to minimize (3) with respect to **w**, for a given value of *d*. So the criterion is unimodal and the well-known result is called *Wiener solution*. But it is worth noting that, for different values of *d*, we have distinct Wiener solution performances. Hence the Wiener criterion may be viewed as a multimodal one if the equalization delay is also considered as an optimization variable.

On the other hand, if s(n) is not available, we turn to the so-called *blind* or *non-supervised criteria*. Among them, the *constant modulus (CM) criterion* has received a considerable attention since its first proposition by Godard in 1980 [1]. Its cost function is expressed by:

$$J_{CM} = E[(|y(n)|^2 - R_2)^2] \qquad (4)$$

where R_2 is defined by:

$$R_2 = \frac{E[|s(n)|^4]}{E[|s(n)|^2]} \qquad (5)$$

Recent works have shown that the minima of (4) are strongly related with some of the Wiener solutions, for a certain set of delays. Inspired by [3, 4], we conjecture that this set may be the best possible in the MSE sense. This motivates the application of efficient search methods in (4), since succeeding in this task would allow the attainment of good Wiener solutions without supervision. Implementation particularities will be properly discussed in section 4.

III. A METHOD FOR BLIND CHANNEL IDENTIFICATION

Another important issue in communications is the identification of the channel impulse response. This procedure has many practical motivations, e.g., modeling and equalizers design; MLSE equalizer implementation; etc. Usually, identification is carried out in a supervised way, which cannot be practical, as previously mentioned.

So, blind techniques were also proposed in the literature. There are different possible approaches, based on cumulant analysis (e. g. [6]), or on signal recovery after equalization (e. g. [5]), to mention a few.

We propose a new *equalization-based approach* to blind identification, founded on the convolution commutative property. First of all, let us define h(n) and w(n) as channel and equalizer impulse responses, respectively. We may thus write:

$$y(n) = s(n)*h(n)*w(n) \qquad (6)$$

For the sake of illustration, we can imagine a two-coefficient channel and a 20-tap equalizer. It is quite reasonable to think that, by means of a proper adaptive algorithm, we will be close to the ZF condition. Now, if we switch channel and equalizer positions, the relation between s(n) and y(n) will not change. But now we have a 20-tap channel properly equalized by a two-tap (!) filter. The procedure is represented in Fig. 2.

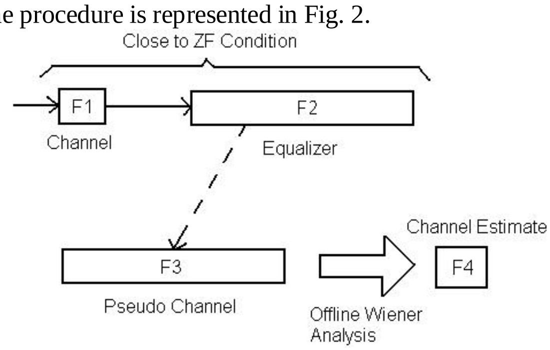

Fig. 2. Identification Process

So, the proposed procedure to identify a certain channel (F1) is the following:

1) Choose a “long-length” equalizer (F2) and a proper blind adaptive search method to adjust this filter, in order to ensure a condition rather close to the ZF.

2) Assume the obtained F2 is a “pseudo-channel” (F3) and perform an offline analysis to derive its best equalizer

(F4) in the Wiener sense, comparing every possible delay, given a channel order.

3) The best Wiener solution (F4) will be the best estimation of F1. The crucial point is that the F4 equalizer is rather close to the ZF condition, despite its short length, since it comes from a simple change of order in the (F1 + F2) chain. This concludes the procedure.

a certain channel (F1). We choose a large equalizer and a good blind adaptive search method to adjust this filter (F2), what will ensure a condition rather close to the ZF. Now, imagine that we face F2 as a “channel” (F3). We can easily perform an offline analysis that will give us the best equalizers in the Wiener sense, for every possible delay, given a certain order. From these, we pick the best (F4). If F1 + F2 gave us a good condition, it can also be attained by F3 + F4. As F3 is supposed to be larger than F4, the last will probably be rather close to F1, for the Wiener criterion has much to do with the idea of approximating the ZF condition (specially if the source is white). This concludes the estimation process.

The genetic CM equalizer has been chosen to obtain F2. For this reason, some background on this method will be posed in sequel.

IV. GENETIC ALGORITHMS

The basic ideas behind GA were first proposed by J. Holland [7], in 1975, and have been widely used in the most different applications since then.

The original GA works as follows: a set of solutions is generated and coded, usually in a string, named chromosome or individual. The set of individuals is said to be a population. For a given number of iterations – called *generations* – the individuals on the population go through *crossover* (new individuals are generated using information from two former ones), *mutation* (individuals are slightly changed), *evaluation* (to each individual is assigned a fitness function, supposed to measure its quality) and *selection* (individuals of the population are selected for the next generation, giving priority to the most adapted ones).

The analogy with nature comes clearly, since according to Darwin’s classical studies, individuals are generated via reproduction (crossover + mutation). The generated individuals compete amongst themselves and the best-adapted ones (evaluation) are more likely to reproduce and pass their genetic information to the next generation (selection).

The rationale behind GA is in the fact that nature was able to generate, via evolution, a wide range of different species, which have adapted to very different conditions. One expects that the individuals in the GA will also be able to evolve, adapting to the best solutions of the search space.

Mathematically, GA is a stochastic optimization method that does not need any information on the problem but one measure of quality of solutions. Though strongly relying on randomness, GA is quite more sophisticated than random search. Indeed, Holland argues that the key to understanding why GA works is in its implicit parallelism. According to Holland [7], when evaluating N individuals in a generation, a GA is implicitly working with a much higher number of solutions (N^3, according to his estimates) what makes the search much more effective.

Though very efficient in exploring the search space, GA’s are prone to premature convergence to one optimum. Actually, the maintenance of diversity in the population is one of the most important questions in GA studies. Mutation usually plays an important role, since it always introduce new information to the population.

Other methods that try to maintain diversity have also been proposed. For instance, some authors have proposed algorithms that introduce randomly generated individuals into the population. This can be done at every generation or when the diversity was considered to reach low levels. This procedure is known as *immigration*.

The rates of mutation and/or immigration must take into account that very high values can transform the algorithm into a simple random search, while very low rates can lead to a convergence to local optima, which are sometimes very distant from the global optimum.

In some cases, it may be very hard to match this compromise. In particular, search spaces with very strong multimodal characteristics are examples of rather difficult scenarios for classical GAs. For these cases, the literature has proposed alternative methods to obtain a wide exploration of the search space.

These strategies are named *Niching Methods* [8] and are based on the idea that, in an ecosystem, not all species compete amongst themselves. Some species may not interfere with others and some may even help the proliferation of others. Actually, each species has a niche and the individuals of that specie will compete only against themselves and against individuals of other species that share the same niche. Some of the most famous Niching methods are *Fitness Sharing* [9] and *Clearing* [10].

Fitness Sharing is probably the best-known niching method available in the literature. In this method, a measure of distance is made between each individual and all other individuals in the population. The fitness of each individual is divided by the number of individuals that are relatively close to it. The idea is to privilege individuals that explore new regions of the search space.

The main difficulty in Fitness Sharing is to determine what is “relatively” close in an unknown problem and the fact that very good individuals can be lost due to closeness to other individuals. It could be better to reduce the fitness of the other individuals but maintain the fitness of the best one. This is the idea of Clearing.

In the original Clearing Method [10], the best individuals in each niche are preserved and the others have the fitness reduced to zero. In this way, the presence of very close individuals is discouraged but no harm is done to the best individuals. Moreover, Petrowski argues that Clearing procedure has a lower complexity and is directly compatible with elitist strategies, where the best individual per niche is preserved from one generation to another. However, there is still the difficulty of determining the radius of the niche.

Proposed Algorithm

The GA to be proposed here works in the following way: a population is generated and goes through the classical operators of crossover and mutation. A Clearing Method Analysis precedes the selection phase.

The Clearing maintains the fitness of the best individual per niche and introduces random generated individuals (immigration) to substitute the other ones.

The *k* best individuals per niche are preserved in an elitist procedure. The difficulty of determining the niches is overcome by using an intrinsic characteristic of the problem: it is observed that the best equalizers for a channel usually contain a predominant coefficient. These coefficients are different for each one of the best minima. Therefore, the idea is to assume that solutions that have the predominant coefficient in the same position belong to the same niche.

The use of “the predominant coefficient” as the niche definition criterion fits very properly our application with a quite low computational effort. This would be not at all the case if we had to calculate every distance by using, for instance, the norms amongst all the individuals.

V. Results

We chose two channels to test the proposed equalization framework. One (C1) is a non-minimum phase channel with coefficients [1 0.4 0.9 1.4] and the other (C2) is also a non-minimum phase channel with coefficients
[1 1.2 –0.3 0.8].

Tab. 1 contains the GA settings for the simulations. The population size, mutation and crossover parameters are rather standard and have been used for all simulations. This is a very important information since it shows that the algorithm is robust and we don’t have to set the parameters for each new situation.

The Clearing Method parameters were also kept the same during the whole set of tests. We decided to preserve only the best individual of each niche (*k*=1) and the number of niches was, in each case, the number of parameters.

TABLE 1
G. A. SETTINGS

Parameter	**Value**
Standard GA	
Population Size, N_{ind}	30
Nb. of Crossovers per Gen.	10
Mutation probability, pm	0.1
Stopping Criterium	2000 Generations
Clearing Method	
Nb. of ind. preserved per niche, *k*	1
Number of niches, *q*	Equalizer order

The first test is carried out in a relatively small search space. The C1 channel and a five-parameter equalizer are considered. This filter order is not sufficient to achieve a good open-eye condition, but will fit nicely our requirements in showing the effectiveness of the GA.

Tab. 2 has been built from the outcomes of 50 simulations. It reveals that the method has *always* provided a solution rather close to the global Wiener minimum, what clearly indicates a very good performance. We also carried out a test with a 10 individuals population. In this case, global convergence was achieved in 82% of the cases and the mean residual error was 0.1905, close to the lowest possible cost. We conclude that a good performance can also be obtained with a lower computational burden.

TABLE 2
C1 + 5 COEF. EQUALIZER

Solution	**Residual MSE**	**Freq.**
[0.2183 -0.1873 -0.0596 -0.2804 0.5892]	0.1751	100%
Mean Residual MSE: 0.1751		

Let us now turn to a larger and more complex search space with the C1 channel and an eight-coefficient equalizer. In this case, we still have not a significant IIS reduction, since there are more minima and they have a higher degree of uniformity. These features account for a more difficult job.

Tab. 3 brings the corresponding results. We notice that convergence to good minima was predominant, with a remarkable global convergence rate. The inverse relation between MSE and frequency marks the coherence of the method and proves its distance from any kind of random search. The mean residual MSE indicates that it is reasonable to expect a "second minimum" performance in this case.

The last simulated scenario is formed by the combination of C2 channel with a seven-coefficient equalizer. This filter order is enough to provide a condition close to the ZF one. Tab. 4 brings the corresponding results. Again, we have a very good global convergence rate and a good proportion between residual MSE and frequency. This reveals, once more, the method efficiency. The mean residual MSE is lower than the MSE achieved by the second minimum. This provides a solid base for good performance expectation.

This concludes our investigations on equalization. The method revealed itself a powerful search tool, from which is acceptable to expect an average performance close to the best possible one. The proposed GA showed great efficiency, what is clear from Tabs. 2, 3 and 4, whose MSE / frequency relation is quite noticeable.

TABLE 3
C1 + 8-COEF. EQUALIZER

Solution	Residual MSE	Freq.
[0.1975 -0.1420 -0.1180 -0.2265 0.5200 0.1146 -0.1132 0.0848]	0.1293	48%
[-0.1191 0.0528 -0.0234 0.2825 -0.2372 -0.0465 -0.3045 0.6145]	0.1397	22%
[-0.0235 0.2256 -0.1475 -0.1057 -0.2763 0.5536 0.1190 -0.0666]	0.1445	12%
[-0.0237 -0.0247 0.2331 -0.1408 -0.0860 -0.3011 0.5466 0.0906]	0.1533	10%
[-0.1350 -0.1433 0.3712 0.1986 -0.1456 0.1708 -0.1339 0.0572]	0.1890	4%
[-0.0246 -0.1117 -0.1464 0.3829 0.1610 -0.1226 0.1730 -0.1008]	0.1951	4%
Mean Residual MSE: 0.1408		

Our attention will now be concentrated on the identification proposal (section 3). Its test is not necessarily related to the previously exposed equalization framework. However, to put things on that basis gives us a somewhat practical view on its capabilities.

TABLE 4
C2 + 7-COEF. EQUALIZER

Solution	Residual MSE	Freq.
[-0.0837 0.1543 -0.2303 0.4058 0.2894 -0.0522 -0.1528]	0.0312	48%
[0.1422 -0.2480 0.3825 0.3070 -0.0390 -0.1451 -0.0419]	0.0458	40%
[-0.2277 0.4120 0.3456 -0.0667 -0.1660 -0.0536 0.0615]	0.0917	8%
[0.0400 -0.0851 0.1579 -0.2249 0.3744 0.2651 -0.0714]	0.0918	2%
[-0.0373 0.0467 -0.0735 0.1732 -0.2449 0.3592 0.2547]	0.1022	2%
Mean Residual MSE: 0.0445		

We chose the last scenario (C2 + seven-coefficient equalizer) to test the method, since the corresponding minima are closer to the ZF solutions. First, we supposed that channel order was *a priori* known. Tab. 5 shows the estimates obtained from each minimum present in Tab. 4 and also brings an error index, defined by:

$$\text{Error} = ||\mathbf{h} - \mathbf{h}_{est}|| \quad (7)$$

which evaluates the difference between the estimate and the real channel vector.

TABLE 5
IDENTIFICATION RESULTS

Estimate	Error (Norm)	Freq.
[0.9486 1.2824 -0.3365 0.6844]	0.1553	48%
[0.8637 1.2860 -0.4421 0.7265]	0.2271	40%
[0.6690 1.4515 -0.5551 0.8458]	0.4899	8%
[1.1185 1.3836 -0.4100 0.3942]	0.4739	2%
[1.1842 1.2996 -0.6337 0.3420]	0.6041	2%

It is clear that the estimate quality tends to improve when good equalizers are employed. Notice that the minimum error is determined by structural limitations, and not by some limitation inherent to the method. By taking, for instance, the best Wiener solution for a 15-tap equalizer, the estimate would be $\mathbf{h}_{est}$ = [0.9944 1.2052 -0.3027 0.7933] and the error index, 0.0105. This proves the efficiency of the proposal.

The next step is to test an order estimation strategy, keeping the same scenario. This expedient is extremely desirable, for such knowledge is hardly at hand in a practical application.

Our idea is to improve F4 order while monitoring its residual MSE. We assume that a significant error reduction will take place in the transition to a good order estimate.

If we consider that channel order equals one, the residual MSE of the best Wiener solution is 0.5404. By augmenting this number, we obtain Tab. 6. There is a significant qualitative jump in the transition from order = 3 to order = 4. For higher orders, the system experiences a timid MSE reduction. We thus consider order = 4 a reasonable estimate. For the sake of illustrating the method robustness, we can observe that an overestimated order = 5 would lead to two "good estimates": $\mathbf{h}_{est}$ = [0.9447 1.2689 -0.3521 0.6726 -0.0921] and $\mathbf{h}_{est}$ = [0.0790 0.9588 1.2958 -0.3249 0.6878]. The first solution corresponds to addition of a zero *after* the significant coefficients and the second, to the addition of a zero *before* them.

TABLE 6
ORDER ESTIMATION PROCEDURE

Chosen Order	Residual MSE
1	0.5404
2	0.2430
3	0.1817
4	0.0221
5	0.0192
6	0.0158

VI. CONCLUSIONS

The objectives of this work were twofold: to propose and analyze an equalization framework based on genetic algorithms as well as a method for blind channel identification.

Results show that GAs can be successfully used in the adaptive equalization problem. They possess several desirable features which contribute to the verified good performance and justify its choice as the basis of the whole proposed approach.

One future task is, in our opinion, to look for memetic operators, i. e., evolutionary procedures based on the equalization problem peculiarities. This may help to reduce computational cost, which is still somewhat high, and can contribute to improve the search effectiveness. We also consider a very important task to look for new GA betterments in the future. Additional motivation comes from the concrete possibility of applying this framework to nonlinear supervised and blind equalization.

The identification method was successful. Its central idea was confirmed but a deeper study remains to be carried out to well acess all potentialities of the approach. It is also important to provide a comparative account on this and other blind identification methods and a careful test under noisy conditions. Noise influence certainly can be rather pernicious, for it can provoke a drift from "channel inversion" solutions. Clustering methods can be an effective countermeasure to this drawback.

ACKNOWLEDGEMENTS

We thank Fernando Von Zuben, Cristiano Panazio and Maurício de Castro, for their important contributions. Alysson Costa and Romis Attux thank, respectively, CAPES and FAPESP (process no. 01/00680-4) for their financial support.

REFERENCES

[1] D.N. Godard, "Self-recovering Equalization and Carrier Tracking in Two-dimensional Data Communication Systems", IEEE Trans. on Comm., Vol. COMM-28, No. 11, pp. 1867-1875, 1980.

[2] S. Haykin, *Adaptive Filter Theory*, 3rd Edition, Prentice Hall, 1996.

[3] C. R. Johnson, P. Schniter, T. J. Endres, J. D. Behm, D. R. Brown, R. A. Casas "Blind Equalization using the Constant Modulus Criterion: a Review", Proc. of the IEEE, Vol. 86, No. 10, pp. 1927-1950, 1998.

[4] H. Zeng, L. Tong, C. R. Johnson "An Analysis of Constant Modulus Receivers", IEEE Trans. on Sig. Proc., Vol. 47, No. 11, pp. 2990-2999, 1999.

[5] P. Schniter, R. A. Casas, P. Touzni, C. R. Johnson, "Performance Analysis of Godard-based channel identification", IEEE Trans. on Sig. Proc., Vol. 49, No. 8, 2001.

[6] Z. Ding, J. Liang, "A cumulant matrix subspace algorithm for blind single FIR channel identification", IEEE Trans. on Sig. Proc., Vol. 49, No. 2, pp. 325-333, 2001.

[7] J. H. Holland, *Adaptation in Natural and Artificial Systems*, Univ. of Michigan Press, 1975.

[8] B. Sareni, L. Krähenbül, "Fitness Sharing and Niching Methods Revisited", Trans. on Evol. Comput., Vol. 2, No. 3, pp. 97-106, 1998.

[9] D.E. Goldberg, J. Richardson, "Genetic Algorithms with Sharing for Multimodal Function Optimization", in Genetic Algorithms and their Applications: Proc. Of the 2nd Int. Conf. on Genetic Algorithms, pp. 41-49, 1987.

[10] A. Pétrowski, "A Clearing Procedure as a Niching Method for Genetic Algorithms", IEEE 3rd International Conference on Evolutionary Computation (ICEC'96), Nagoya, may 20-22, 1996.