NEURAL NETS FILTERS: INTEGRATED CODING AND SIGNALING IN COMMUNICATION SYSTEMS.

M.E. Santamaría, M.A. Lagunas, M. Cabrera

TSC Dept., ETSI Telecom., Apdo. 30002, 08080 Barcelona, Spain

SUMMARY

This paper presents the potential of the herein so-called neural net filters in communication systems. The use of integrated signaling and coding techniques in modern modulation schemes has not a right correspondence in the same digital signal processing integrated tools. The neural net filter seems to be the needed contribution. Introduced from de basic neural structure associated with coding or finite word length representations of samples, it is shown that a FIR filter with finite representation of its output could be viewed at a two layers neural net.

Further generalizations attending some currently interesting problems in communications are provided.

To asses some preliminary conclussions, a few experiments in equalization of non-linear with memory communications channels is reported. From these experiments is easy to conclude the potential of neural networks in integrated tools for signal processing and decoding.

1. INTRODUCTION

The success in the last decade of adaptive algorithms in tap-delay or lattice structures promoted many attempts to extend its philosophy to other architectures for digital signal processing. The use of infinite response filters or transform domain schemes represents somehow a little advance in the use of adaptive algorithms. Probably the first approach of adaptive algorithms was done, in an application solved environment, in the field of pattern recognition. The design of a pattern classifier from prototypes or just training it represents the classic problem for an adaptive algorithm, where a processor looks the system response and a given reference or prototype making the changes in consequence with the system performance. In fact, classical linear adaptive filters could be encompassed as a supervised learning problem where the reference plays the same role that the prototypes in supervised learning.

The basic neuron can be derived easily from a classic beamformer in antenna array processing as depicted in Fig. 1.

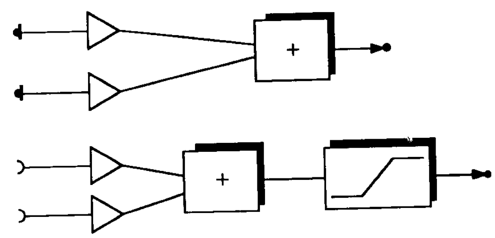

Fig. 1. Array beamformer (above) and neural model (below)

It is easy to see that a neuron evolves from a classic beamformer just introducing a non-linearity in producing the beamforming output. The use of a non-linear device in the beamformer could be based in expected non-linearities in the propagation channel playing the neuron, with respect the communication channel, the same role that the compander/expander in communication systems.

Other aspect that could deserve further attention is up to what degree convergence or tracking behavior could be improved by adaptively increase or decreasing the deviation from linearity of the non-linear device.

But, more important, an unexpected problem arises in neural nets in the same way that with linear array beamformers. Time varying systems interconected, even when using the same update policy, change their performance depending in the way they are located in a cascaded structure whenever the initial state is not the same or all the systems are not equal. Furthermore, adaptive lineal filters in paralell with different objectives to be minimized together with other systems in cascade provide nowdays an open field of research.

At this time it is important to remind that, from the point of view of the adaptive algorithm, the only difference between a linear beamformer and a neuron is just the objective to be minimized. In other words, a linear beamformer with a non-linear objective, lets say exp y_n being y_n the sample output, becomes equal from de adaptive point of view that a neuron with non-linearity equal to the objective adapted to the zero state.

In summary, adaptive linear systems in cascade or in paralell architectures with different objectives arise to a neural network with the involved unknowns that this field currently represents for designers.

This work will deal with the applications of neural networks in those fields associated with communications where time varying linear systems need to be used or just represent one step ahead in the use of digital signal processing in advanced receivers. The structure of neural net to be under study will be the multiple layer feed-forward network.

2. SIGNALING AND CODING

Many communications systems and concretly digital receivers, probably due to the apparence of adaptive equalizers, were achieving higher data rates than classical receivers by including fast adaptive equalizers with different structures FIIR or FIR, DFE, fractionally spaced, etc. In other sense coding techniques with increasing complexity supported by cheap technology shown new upper bounds in communications performance. Regardless the superiority of coding versus equalizers still the drawbacks derived from the difference between the objective in coding (i.e. the Hamming distance) was constraining the potential of both procedures.

When CPM and, mainly TCM, were reported, the gap between signalling and coding disappears and at the same time the interest of digital signal processing tools in digital communication receivers increases. A brief description of the receiver reduces to a Viterbi like decoder where branching represents the potential of coding and ranking each brunch at every sample input depends on a digital signal processing algorithm. The situation is ploted in Fig. 2.

Every branch in the trellis diagram of the decoder is quoted by the coding only certain connections are allowed; after some input samples that could be fixed or not depending on the delay required to the decoding process, the decision is done based on the most likely path. It is when this decision have to be done when ranking every branch the importance of the DSP takes place. To gain a little bit more inside, the channel code associated to a candidate branch is resignaled on the DSP which with a channel model try to reproduce the input $\underline{x}_n$; the euclidean distance or magnitude of the difference dictates the ranking of the transition. In order to harmonize both distances in branching, the Hamming for close symbols and the signals supporting those symbols in terms of euclidean distance, TCM includes signaling and coding at the transmiter as an integrated tool to achieve better performance.

To stablish the relationship between TCM like communication systems and neural networks, let us consider a soft decision decoder where a given vector of input samples X_n, being $X_n = (x_{n-1}, \ldots, x_{n-a})$ coded with a finite number of bits, have to produce the most likely bit pattern Y_n at the output.

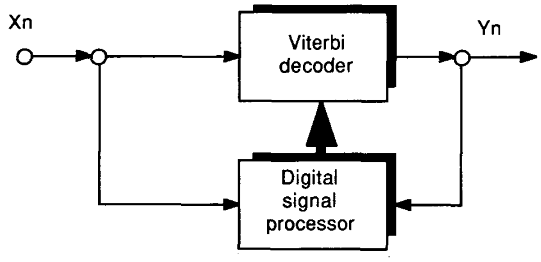

Fig. 2. Viterbi decoder with side information from a DSP.

Thus, at least from the point of view of the input/output relationship a Viterbi decoder or CPM and TCM decoders using soft-decision and side information could be envisaged as a potential application of neural nets. Note that in this problem the set of prototypes will be random in nature due to the communication chanel and the receiver noise including acquisition errors and synchronism missadjustments.

If it is possible to use neural nets in such a problem there is no longer any need to implement companders or expanders for non-linear effects due to the communication chanel. This allows to avoid the undesired coupling effects between companders and adaptive equalizers in cascade. Really the number of unknown problems to be faced by the inclussion of adaptive non-linear systems with memory, as concerns with signal and noise distribution both in probability or in their spectral density, previously to the decoder has not well suited reported solutions currently.

Furthermore, even in the case of binary detection again the adaptive equalizer together with the decision device could be viewed as a neuron in the sense that the samples at the different taps represents the neuron input and the binary decision as its output. Adaptive algorithms for this problem work also quite well in a single layer neural net.

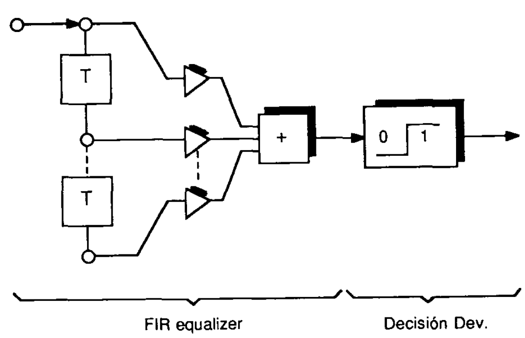

Fig. 3. FIR equalizer in a binary detection receiver and its similarity with a single neuron.

We will describe in the next section in which sense decoders regardless its complexity could be modelled as a neural net.

3. NEURAL NET FILTERS

The existence of a neuron in a digital signal processor it is most likely that initially could appear to us. In the previous section has been shown that any receiver, no matter the complexity of the modulation and coding, when including the decoder in cascade the filters or equalizers we have a neuron. But this kind of processor appears just in a classic FIR filter which output has to be converted to a finite precission arithmetic.

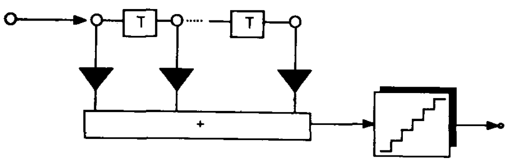

Fig. 4. A neuron as a FIR filter with A/D converter.

Taking this last approach, it is worthwhile to note that many of the powerfull tools in signal processing, involving concepts as Wiener filtering with reference or adaptive algorithms using the instantaneous gradient among others, have been used with a high degree of performance. The question to be answered is up to what degree the neural nets can be viewed as an step ahead to achieve a generalized filtering framework. We know, as mentioned before, that even without non-linearities time varying systems in cascade offer many opportunities because just the order they are interconnected changes the behavior of the system. This is a very important point in considering multilayer feed-forward neural networks where number of internal layers and number of hidden units play an important role.

To connect a generalized filter structure with a multilayer neural net and the preceeding examples of neurons in communication systems or in a single FIR system, let us remind that the A/D or the decoder can be modelled as a neural net and once we get this architecture for it, the time delay structure of the preceeding filter can also be integrated on it giving the corresponding neural net depicted in Fig. 5.

The reader could realize that the shape of the non-linearity is smoothed to a centered sigmoid instead of the thereshold device that corresponds to an A/D or decoder. This is not longer a problem since the parameter of the sigmoid could be modyfied in the adaptive process or, more easy than this, just adding a thereshold device at the output nodes.

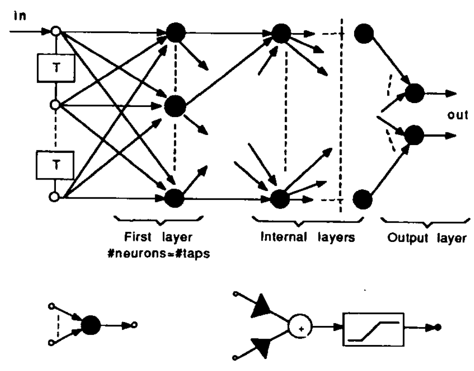

Fig. 5. Diagram of the generalized filter as a feed-forwars neural net.

That scheme can be named as a generalized or neural filter. Note that depending on the application the following input / output patterns can be used depending of the application we are dealing with.

input	output	application
digital	digital	soft-decision decoder
analog	digital	TCM decoder
analog	analog	generalized filter

Table I. Input-output choices for a neural filter.

The neural net architecture selected is a feed forward network and the update algorithm will be the back propagation rule with some modifications to be mentioned in the next section. Regardless a step ahead will be the use of more complex architectures, this paper as a firts attempt in reporting the potential of neural filter will focuss in the next section just the feed-forward scheme.

4. NEURAL WIENER FILTER

Among the possible choices or table I, seems to be that the case of analog input and output is the case of most interest to evaluate or to aprize the ideas reported herein. The filter, which behavior is going to be described, is as depicted in Fig. 5 with a simgle output unit providing an analog output.

The simulated environment is a digital source which symbols signaled with NRZ pulses are introduced to a communication channel with two sections: First section is an ARMA filter with coefficients $(1 - 0.2\,Z^{-1} + 0.1\,Z^{-2}) \,/\, (1 + 0.5\,Z^{-1})$; and, a second section formed by an order 3 non-linearity with coefficients $x + 0.2\,x^2 + 0.15\,x^3$. At the output of the channel, white gaussian noise is added at SNR dB. The reference for the Wiener filter is the signaling output (see Fig. 6). Both the output and the reference are applied to the adaptive neural filter.

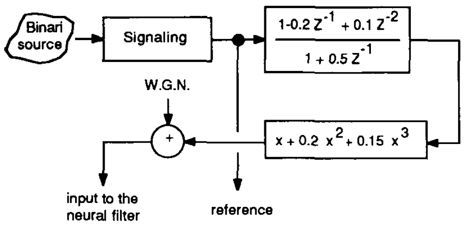

Fig. 6. Simulated communications environment to evaluate a neural filter with reference.

The adaptive algorithm was the backward propagation rule and in order to compare with a classic approach, in Figure 7 the learning curve of the neural net filter (top) is compared with a linear Wiener filter (bottom)

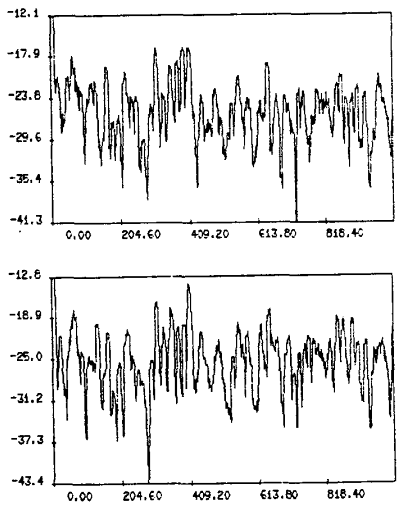

Fig. 7. Learning curves for neural filter (top) and linear Wiener filter (bottom).

From this experiment and others carried out by the authors it can be concluded that:

- The number of neurons in the first layer is related only with the memory of the communication channel. In our case above 8 taps not significant improvement was observed.

- Saturation in the non-linear devices only shows up in the first part of the learning curve. After acquisition and during tracking, the non-linear devices use no more than 80% of their dynamic range.

- The number of internal layers is not very critical in the experiments performed, but at least one internal layer was needed to obtain adequate quality. Seems to be that the number of layers is related with the order of the non-linearity in the communications channel. In this case using more than three layers increase complexity and relents convergence with no improvement in the missadjustment error.

- With respect the number of units in the hidden layers, no guidelines have been obtained up to now and further work with different environments are needed.

In conclusion it can be said that neural net filters represents the alternative to extent the potencial of linear adaptive equalizers for the non-linear equalization problem. With respect to the use of neural nets in fast equalization and decoding further work is needed, but the experiments using 8-PSK over the simulated channel and adding a decision device at the single output unit the performance looks very promissing. Next step is to use a number of output units equal to the source alfabet and evaluating the learning curve in terms of the global BER of the receiver. Any case, the intermediate step will be the use of digital codewords as reference in the learning period. A good performance in this last case will be in favour of the use of a global neural receiver for signaling / coding processing like in a TCM receiver.

5. REFERENCES

<1> J.G. Proakis. "Digital Communications". Mc. Graw-Hill. 1983.

<2> Mendel and Fu. "Adaptive, Learning, and Pattern Recognition Systems: Theory and Applications". Academic Press. Mathematics in Science and Engineering. Vol. 66. 1970.

<3> B. Widrow, R. Winter. "Neural nets for adaptive filtering and adaptive pattern recognition". IEEE comp., March 1988, pp. 25-39.

<4> D.E. Rumelhart, J.L. McClelland. "Paralell Distributed Processing", Vol. I and II, MIT Press, Cambridge, Mass., 1986.