QUALITY ASSESSMENT OF HYBRID NONLINEAR FILTERS

Mo Chen and Danilo P. Mandic

Imperial College London
Department of Electrical and Electronic Engineering
Exhibition Road, London, SW7 2BT, UK
E-mail: {mo.chen, d.mandic}@imperial.ac.uk

ABSTRACT

Traditionally, research on adaptive signal processing has been conducted with the aim of designing adaptive filters with high performance in terms of some prescribed performance measure. However, little is known about how such filters influence the nature of the processed signal. Based upon some recently introduced results on dealing with nonlinearity within a signal in hand, we provide a critical assessment of the qualitative performance of common linear and nonlinear filters and their combinations. An insight into the performance of so called hybrid filters is provided, which is achieved for combinations of standard nonlinear (neural) and linear filters. It is shown that depending on the application, it is important not only to look for best filter performance in terms of some quantitative measure of the error but also for a filter that will not change the character of a signal. Simulation results support the analysis.

1. INTRODUCTION

Much research has been conducted to devise an "optimal" filter according to a certain pre–defined criterion [1][2]. This has been achieved not only for linear (Wiener filter) but even more importantly for nonlinear filters [1][3]. The goal is usually to optimize (maximize) some performance criterion, a quantitative measure of the quality of performance. These optimization problems become more complex when we look for a sequential solution to a global optimization problems. In the adaptive filtering literature, for instance, the least mean square (LMS) and recursive least squares (RLS) are such first and second order algorithms [1], which are an approximate solution to a global "batch" optimal filtering problem. These solutions are sub–optimal, but computationally simple and widely employed in practice.

In the area of nonlinear filters, neural nonlinear filters are emerging in a variety of applications [3]. They are able to cope with the complexity of the problem and represent an alternative to linear adaptive filters for applications where linear filters cannot cope with the nonlinearity of a signal. Moreover, nonlinear filters might be a viable alternative for solving linear problems [4]. Since prediction is at the core of both nonlinear and linear adaptive signal processing and machine learning, we shall restrict ourselves, without the loss in generality, to only problems involving prediction. Notice that an input to an adaptive filter can be either linear or nonlinear[1]. A linear signal is the one which is generated by a linear stochastic model, while a nonlinear signal is the one which deviates from the linear one. Methods for quantifying nonlinear properties of a signal have been recently developed and include the so called the third-order autocovariance (C3) [7] and the asymmetry due to time reversal (REV) [7] as well as correlation dimension [8].

In some applications, such as medical ones, the nonlinear/linear nature of a signal conveys information about the health state of a patient. For instance, the nature of the heart rate variability signal changes from stochastic (linear) to chaotic (nonlinear) depending on whether the patient is healthy or not [9][10]. Therefore for such applications it is important not only to employ an adaptive filter which provides high performance but also the one which will not change the nature of the signal.

In this paper, therefore, we provide an initial investigation into the qualitative properties of linear and nonlinear neural adaptive filters and their combinations (hybrid filters) with respect to preserving the nature of a signal in hand. The analysis is supported by quantitative performance measures and illustrative examples highlighting the need to take into account the nature of processed signals when choosing adaptive filters and algorithms for online applications.

[1]It is important to differentiate between the linearity/nonlinearity properties of a system, which can be examined by superposition, from the linear/nonlinear properties of a signal, a theory which has only recently emerged in the physics literature [5] and its applications in signal processing are still in their infancy [3][6].

0-7803-8484-9/04/$20.00

2. NONLINEARITY DETECTION METHOD

Several methods for detecting the linear/nonlinear nature of a signal have been proposed, such as the "Deterministic Versus Stochastic" (DVS) plot [5], δ-ϵ Method [11], Correlation Exponent [8]. Among them, the recently proposed Delay Vector Variance (DVV) method is best suited for signal processing applications [3][12]. The algorithm is based upon examination of predictability of a signal in the phase space and is summarized below. For a given embedding dimension m:

- Generate delay vector (DV): $\mathbf{x}(k) = [x_{k-m}, \ldots, x_{k-1}]^T$ and corresponding target x_k,
- The mean, μ_d, and standard deviation, σ_d, are computed over all pairwise Euclidean distances between DVs, $\|\mathbf{x}(i) - \mathbf{x}(j)\| (i \neq j)$,
- The sets $\Omega_k(r_d)$ are generated such that $\Omega_k(r_d) = \{\mathbf{x}(i) | \|\mathbf{x}(k) - \mathbf{x}(i)\| \leq r_d\}$, *i.e.*, sets which consist of all DVs that lie closer to $\mathbf{x}(k)$ than a certain distance r_d, taken from the interval $[\max\{0, \mu_d - n_d\sigma_d\}; \mu_d + n_d\sigma_d]$, *e.g.*, N_{tv} uniformly spaced distances, where n_d is a parameter controlling the span over which to perform the DVV analysis,
- For every set $\Omega_k(r_d)$, the variance of the corresponding targets, $\sigma_k^2(r_d)$, is computed. The average over all sets $\Omega_k(r_d)$, normalized by the variance of the time series, σ_x^2, yields the 'target variance', $\sigma^{*2}(r_d)$:

  $$\sigma^{*2}(r_d) = \frac{\frac{1}{N}\sum_{k=1}^{N} \sigma_k^2(r_d)}{\sigma_x^2} \tag{1}$$

  We only consider a variance measurement *valid*, if the set $\Omega_k(r_d)$ contains at least $N_0 = 30$ DVs, since too few points for computing a sample variance yields unreliable estimates of the true variance. A sample of 30 data points for estimating a mean or variance is a general rule-of-thumb.

The presence of a strong deterministic component will lead to small target variances for small spans. The minimal target variance, $\sigma_{min}^{*2} = min_{r_d}[\sigma^{*2}(r_d)]$, is a measure for the amount of noise which is present in the time series.

In the following step, the linear or nonlinear nature of the time series is examined by performing the DVV test on both the original and a number of surrogate time series, using the optimal embedding dimension of the original time series. Due to the standardization of the distance, these plots can be conveniently combined in a *scatter diagram*, where the horizontal axis corresponds to the DVV plot of the original time series, and the vertical axis to that of the surrogate time series. If the surrogate time series yield DVV plots similar to that of the original time series, the 'DVV scatter diagram' coincides with the bisector line, and the original time series is judged to be linear. If not, then the original time series is thought to be nonlinear.

3. HYBRID VERSUS STANDARD FILTERS

It is an open question whether the use of hybrid filters, for instance a neural network followed by a linear FIR filter trained by LMS or RLS can improve the overall performance, as compared to the performance of single filters. In particular, it has been recently suggested that a cascaded combination of a recurrent neural network (RNN) and FIR filter can separately predict the nonlinear and linear component of a signal, respectively [1]. Although intuitively clear, there has been little practical evidence whether this is the case. Figure 1 shows a block diagram of a hybrid filter. To

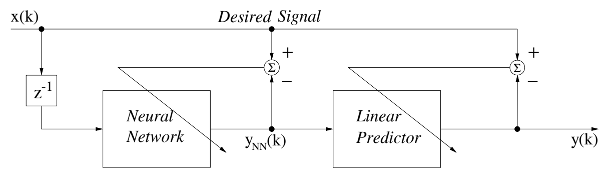

Fig. 1. A hybrid filter for prediction: A temporal neural network followed by an FIR filter.

shed further light on the performance of such filters and to highlight the need for a compromise between a quantitative measure of performance and preservation of the nature of a signal, we employ some recent results from the nonlinearity detection in signals and apply them to adaptive filtering problems.

The filters considered were both nonlinear and linear together with their combinations. The nonlinear neural filters were the dynamical perceptron (nonlinear FIR filter) trained by the nonlinear gradient descent algorithm (NGD) and a recurrent perceptron trained by the real time recursive learning (RTRL) algorithm. The linear filters considered were standard FIR filters trained by LMS and RLS.

The data we considered were a benchmark linear signal, given by [13]

$$\begin{aligned} x_k &= 0.8x_{k-1} + 0.15x_{k-2} + \nu_k + 0.3\nu_{k-1} \\ x_0 &= 1, x_1 = 0.7 \end{aligned} \tag{2}$$

and a benchmark nonlinear signal, given by [14]

$$\begin{aligned} z(k) &= \frac{z(k-1)}{1 + z(k-1)^2} + r^3(k) \\ r(k) &= 1.79r(k-1) - 1.85r(k-2) + 1.27r(k-3) \\ &\quad - 0.41r(k-4) + n(k) \end{aligned} \tag{3}$$

where $\{n(k)\}$ and $\{\nu_k\}$ are realizations of white Gaussian noise $\mathcal{N}(0,1)$. For these signals their DVV scatter diagrams are shown in Figure 2. The linear signal has its DVV scatter diagram on the bisector line, indicating its linear nature, whereas the DVV scatter diagram for the nonlinear signal (right diagram in Figure 2) deviates from the bisector line indicating nonlinearity present in the signal.

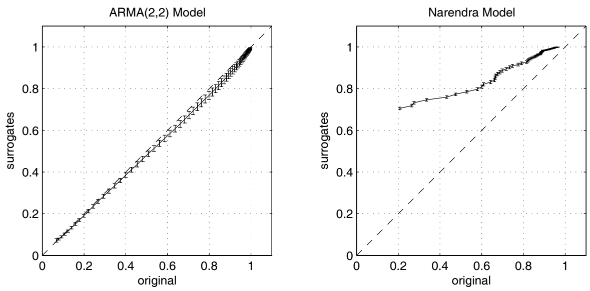

Fig. 2. Nonlinear nature of the considered signals. Left: linear signal. Right: nonlinear signal.

4. SIMULATIONS

We calculated the quantitative performance criterion, the prediction gain, given by $R_p = 10\log\frac{s^2}{e^2}$, that is the logarithmic ratio between the signal variance and prediction error variance for all the classes of filters considered. In addition we compared the DVV scatter diagrams to illustrate how the nature of a signal changes with the use of different classes of filters, which is the main purpose of this paper. Simulations were performed with averaging of 100 realizations of independent trials.

Figure 3 illustrates the quantitative prediction gains R_p and the qualitative DVV scatter diagrams, for the NGD and RTRL algorithms used to train the nonlinear neural dynamical perceptron, recurrent perceptron, and their combination (hybrid filter) with a linear FIR filter trained by LMS and RLS. The experiment was conducted for prediction of the linear benchmark signal (2). The DVV scatter diagrams show the nonlinearity information about the output of such filters. From the Figure, all of the filters and their combinations were able to preserve the linear nature of the filtered signal. In other words, in terms of preserving the nature of the signal (linear in this case), both the linear and nonlinear filters and their combinations performed well on a linear ARMA(2,2) signal, indicated by the fact that all the DVV scatter diagrams in Figure 3 are on the bisector line of the scatter diagram. In terms of the prediction gain, the NGD and RTRL performed similarly, and the hybrid filters performed better than single filters. The cascaded combination of a dynamical perceptron and FIR filter trained by RTRL and RLS gave the best performance.

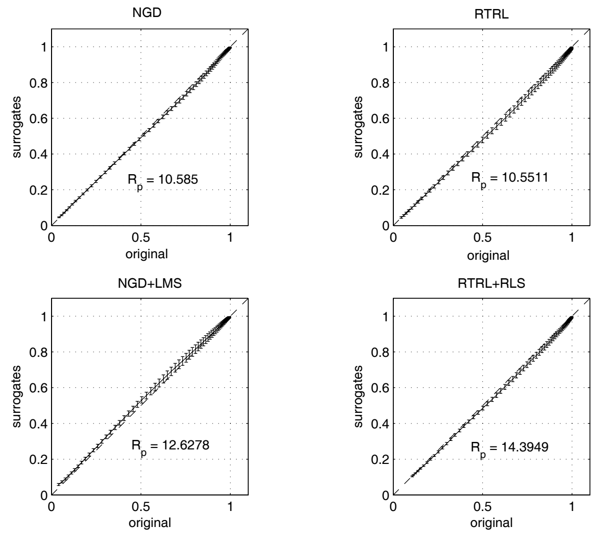

Fig. 3. Qualitative and quantitative comparison of the performance between nonlinear neural and linear filters for a linear benchmark signal (2).

Figure 4 illustrates a similar experiment performed on prediction of the benchmark nonlinear signal (3). The DVV scatter diagrams show the nonlinearity information about the output of such filters. From Figure 4, both nonlinear filters trained with NGD and RTRL performed poorly on their own in terms of the prediction gain. However, looking at the nature of the signal, from Figure 2, they preserved the nature of the benchmark nonlinear signal. The recurrent perceptron trained by the RTRL, showed worse quantitative performance but better qualitative performance. The hybrid filters, in the bottom row of Figure 4, performed better than the considered nonlinear filters. A hybrid filter consisting of a combination of a dynamical percpetron trained by NGD and an FIR filter trained by LMS, showed a considerable increase in prediction gain, however, the signal was considerably linearized and the DVV scatter diagram showed a significant change in the nature of the predicted signals. The bottom right diagram in Figure 4 shows the performance of a hybrid filter consisting of a recurrent perceptron trained by RTRL followed by a FIR filter trained by RLS algorithm. This case gave the best performance out of all combinations of hybrid filters considered in terms of preserving the nature of the signal in hand. The qualitative performance gain for this combination was the second best of all the combinations, whereas the nature of the signal was reasonably preserved.

It is natural to ask a question whether exchanging the order of filters within a hybrid filter will affect the performance.

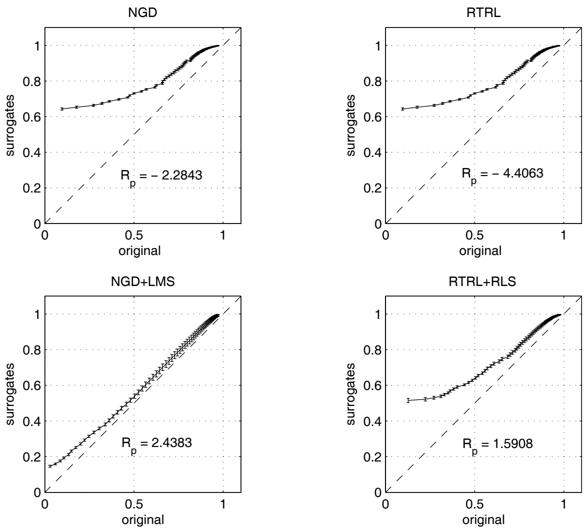

Fig. 4. Qualitative and quantitative comparison of the performance between nonlinear neural and linear filters for a nonlinear benchmark signal (3).

Given the highly nonlinear nature of the problem it is expected that the performances will be significantly different. To this purpose we re-run the experiments for the nonlinear benchmark signal. The results of the experiments are shown in Figure 5. Figure 5 confirms that exchanging the order of

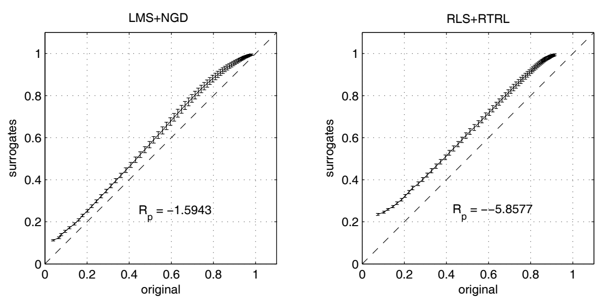

Fig. 5. Qualitative and quantitative comparison of the performance between hybrid filters for a linear and nonlinear benchmark signal. The filter order is interchanged from the one in previous experiments.

the filters within a hybrid filter does not provide the same performance as the original order of filters. In the experiment, both the quantitative performance was considerably worse and also the nature of the predicted signal changed significantly towards the linear one.

5. CONCLUSIONS

We have illustrated the need to consider not only the quantitative performance but also to preserve the nature of a signal when applying real time adaptive filters. To this purpose, we have used the Delay Vector Variance (DVV) method for characterizing the nonlinearities present in the original signal and have compared the nonlinear characteristics of the original signal with those of its predicted versions. This has been achieved for both standard neural and linear filters and hybrid filters consisting of a cascade of the two. It has been illustrated that in some cases, a high prediction gain yields linearized predictions, which may cause a problem in applications where the nature of a signal is of critical importance.

Acknowledgements: The authors would like to acknowledge Dr. Temujin Gautama from KU Leuven, Belgium for fruitful discussions and insightful comments.

6. REFERENCES

[1] S. Haykin, *Neural Networks, A Comprehensive Foundation*, Prentice Hall, 1994.

[2] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice Hall, 1985.

[3] D. P. Mandic and J. A. Chambers, *Recurrent Neural Networks for Prediction: Learning Algorithms, Architectures and Stability*, John Wiley & Sons, 2001.

[4] T. Adali, “Why a nonlinear solution for a linear problem ?,” *in Proc. IEEE Workshop on Neural Networks for Signal Processing (NNSP), Madison, WI*, pp. 157–165, 1999.

[5] A.S.Weigend and N.A.Gershenfeld, *Time Sereis Prediction: Forecasting the future and understanding the past*, Addison-Wesley, 1993.

[6] T. Gautama, D. P. Mandic, and M.M. Van Hulle, “Signal nonlinearity in fmri: A comparison between bold and mion,” *IEEE Trans. Med. Imaging*, vol. 22, no. 5, pp. 636–644, 2003.

[7] T. Schreiber and A. Schmitz, “On the discrimination power of measures for nonlinearity in a time series,” *Phys. Rev. E*, vol. 55, no. 5, pp. 5443–5447, 1997.

[8] P. Grassberer and I. Procaccia, “Measuring the stangeness of strange attractors,” *Physica D.*, vol. 9, no. 2, pp. 189–208, 1983.

[9] I. Radojicic, T. Gautama, and D. P. Mandic, “A comparison of two novel methods for characterization of heart rate variability series,” *in Proceedings of the 16th Biennial International EURASIP Conference BIOSIGNAL2002, Brno, Czech Republic*, vol. 7, pp. 84–87, 2002.

[10] T. Gautama, D. P. Mandic, and M.M. Van Hulle, “A novel method for determining the nature of time series,” *IEEE Transactions on Biomedical Engineering*, vol. tba, no. tba, pp. tba, 2003.

[11] D. Kaplan, “Exceptional events as evidence for determinism,” *Physica D.*, vol. 73, no. 1, pp. 38–48, 1994.

[12] T. Gautama, D. P. Mandic, and M.M. Van Hulle, “Indications of nonlinear structures in brain electrical activity,” *Phys. Rev. E*, vol. 67, pp. 046204–1–046204–5, 2003.

[13] W. Barnettt, A. Gallant, M. Hinich, J. Jungeilges, D. Kaplan, and M. Jensen, “A single-blind controlled competition among tests for nonlinearity and chaos,” *J.of Econometrics*, vol. 82, no. 1, pp. 297–302, 1998.

[14] K. S. Narendra and K. Parthasarathy, “Identification and control of dynamical systems using neural networks,” *IEEE Transactions on Neural Networks*, vol. 1, no. 1, pp. 4–27, 1990.