SIGNAL MODALITY CHARACTERISATION USING COLLABORATIVE ADAPTIVE FILTERS

Beth Jelfs and Danilo P. Mandic

Department of Electrical and Electronic Engineering,
Imperial College London, SW7 2BT, United Kingdom
{beth.jelfs,d.mandic}@imperial.ac.uk

ABSTRACT

A method for extracting information (or knowledge) about the nature of a signal is presented, this is achieved by tracking the dynamics of the mixing parameter within a hybrid filter rather than the actual filter performance. Implementations of the hybrid filter for tracking the nonlinearity and the sparsity of a signal are illustrated and simulations on benchmark synthetic data in a prediction configuration support the analysis. It is then shown that by combining the information obtained from both hybrid filters it is possible to use this method to gain a more complete understanding of the nature of signals and changes in signal modality.

***Index Terms*—** adaptive filters, collaborative signal processing, distributed signal processing, signal modality characterisation

1. INTRODUCTION

Hybrid filters have been previously introduced to improve the performance of adaptive filters, illustrating that by collaborative learning using a combination of subfilters of different characteristics it is possible to achieve better overall performance than that obtained from any of the individual subfilters [1]. One of the keys to designing the hybrid filter is the method in which the subfilters are combined, one simple but effective method is to combine the outputs of the subfilters in a convex manner. Convexity can be described as [2]

$$\lambda x + (1-\lambda)y \text{ where } \lambda \in [0,1] \tag{1}$$

For x and y being two points on a line, as shown in Fig. 1, their convex mixture (1) will lie on the same line between x and y. By using a convex combination of adaptive filters it

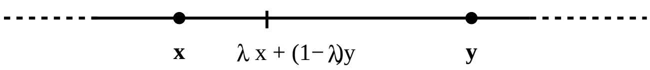

Fig. 1. Convex Combination

is possible to not only to obtain the best properties from the constituent subfilters but also to improve the overall stability of the filter, as should one subfilter fail to converge the hybrid filter tracks the output of the second subfilter [3, 4].

Figure 2 shows the block diagram of a hybrid filter consisting of two adaptive subfilters combined in a convex manner. At every time instant k, the output of the hybrid filter, $y(k)$, is an adaptive convex combination of the output of the first subfilter $y_1(k)$ and the output of the second subfilter $y_2(k)$, and is given by

$$y(k) = \lambda(k)y_1(k) + (1-\lambda(k))\, y_2(k). \tag{2}$$

The outputs of the two subfilters are dependent on the algorithms used to train the subfilters based on the common input vector $\mathbf{x}(k) = [x_1(k), \ldots, x_N(k)]^T$ for filters of length N. The outputs are given by $y_1(k) = \mathbf{x}^T(k)\mathbf{w}_1(k)$ and $y_2(k) = \mathbf{x}^T(k)\mathbf{w}_2(k)$ with corresponding weight vectors $\mathbf{w}_1(k) = [w_{1,1}(k), \ldots, w_{1,N}(k)]^T$ and $\mathbf{w}_2(k) = [w_{2,1}(k), \ldots, w_{2,N}(k)]^T$, where each subfilter is updated by its own error $e_1(k)$ and $e_2(k)$, using a common desired signal $d(k)$.

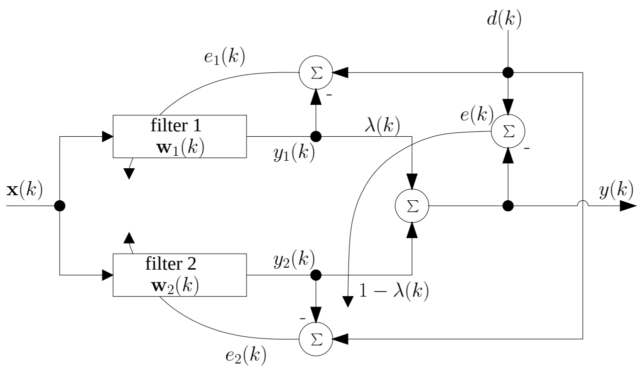

Fig. 2. Hybrid Filter Structure

Hybrid filters using convex combinations of two subfilters both being trained by the same algorithm have been shown to perform well in stationary environments and to always perform at least as well as the better of the two subfilters [3]. An alternative to this using a combination of two subfilters trained by two different algorithms has shown that by careful selection of the training algorithms it is possible to take the desired properties of both subfilters to give a better overall performance [4].

By understanding that the output of the hybrid filter will always be dominated by the better performing of the two subfilters, it is natural to assume that this information can be obtained by observing the adaptive mixing parameter λ. With this in mind we show that it is possible to design a hybrid filter which uses subfilters trained by algorithms with known different properties and that the behaviour of λ within such a combination will reveal not only which filter is currently giving the best response to the input signal, but also that with appropriately chosen subfilters the response of λ can then be used to reveal knowledge about the nature of the input signal.

2. SIGNAL CHARACTERISATION

Signal characterisation is a key topic of multidisciplinary research, but its applications in signal processing and machine learning have only recently become apparent. The range of signals spanned by just the characteristics of nonlinearity and stochasticity are shown in Fig. 3 (modified from [5]) and whilst there are some small areas which are well understood these tend to be extremes in nature, such as extremely nonlinear, deterministic signals (chaos), or linear and stochastic signals represented by autoregressive moving averages (ARMA). These extremes however do not cover the majority of real world signals, and the presence of factors such as noise or uncertainty leads to most real world signals being represented in the areas (a), (b), (c) or '?'. Knowing more about the nature of the signal being processed can provide valuable information in many areas such health or weather conditions, it can also provide prior knowledge for the selection of appropriate models as use of incorrect systems can lead to problems in their use or training and in some situations (such as the use of nonlinear model in absence of nonlinearity) can add extra unnecessary computational complexity. Many existing approaches to signal characterisation are based upon hypothesis testing, describing the signal in a statistical manner [6]. There is however a need for online approaches to signal characterisation which can not only identify the nature of a signal but also track any changes in signal modality. Some disadvantages of existing approaches are due to their tendency to rely on underlying models [7], making their application somewhat limited. To overcome these limitations we propose a much more flexible method based on collaborative adaptive filtering.

3. HYBRID FILTERS FOR SIGNAL CHARACTERISATION

Whilst previous implementations of hybrid filters based on convex combinations have focused mainly on the quantitive

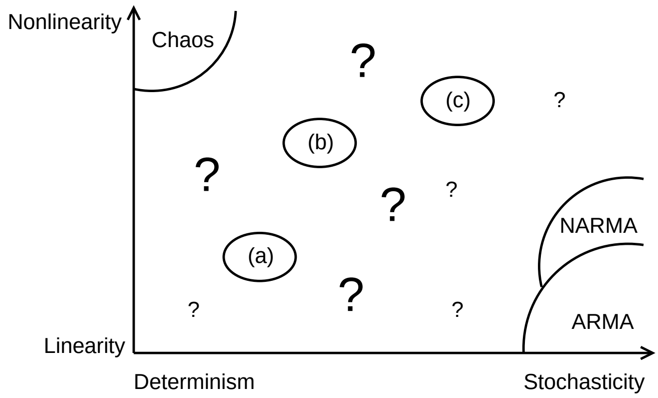

Fig. 3. Deterministic vs. stochastic nature and linear vs. nonlinear nature of real world signals

improvement in performance, our approach relies on observing the evolution of the mixing parameter λ. As such the hybrid filter is designed to have two constituent subfilters with sufficiently different characteristics so that when the mixing parameter λ is observed we gain an insight into the nature of the signals. In addition through the evolution of the mixing parameter, we can also track any changes in the modality of the signals [8].

Using the hybrid filter structure as illustrated in Fig. 2, the convex mixing parameter $\lambda(k)$ is updated based on minimisation of the quadratic cost function $E(k) = \frac{1}{2}e^2(k)$. Although in order to preserve the inherent characteristics of the constituent subfilters, which are the basis of our approach, and as previously stated the subfilters are updated based on the errors generated based on their individual algorithms, the parameter λ is updated based on the overall error $e(k)$. Using the following gradient adaptation

$$\lambda(k+1) = \lambda(k) - \mu_\lambda \nabla_\lambda E(k)_{|\lambda=\lambda(k)}, \tag{3}$$

where μ_λ is the adaptation step-size, from (2) and (3), an LMS type adaptation for the λ update can be obtained as

$$\begin{aligned} \lambda(k+1) &= \lambda(k) - \frac{\mu_\lambda}{2} \frac{\partial e^2(k)}{\partial \lambda(k)} \\ &= \lambda(k) + \mu_\lambda e(k)(y_1(k) - y_2(k)). \end{aligned} \tag{4}$$

In order to preserve the convexity of the update it is essential that the value of λ remains within the range $0 \leq \lambda(k) \leq 1$ and to this effect a hard bound on the values which λ can take was implemented. It is worth noting at this point that although alternative methods have been used to preserve the convexity of the function and have been proven to give good results (such as the sigmoid function used in [3]) this is not appropriate in our case as our primary interest is not in the filter output but in the behaviour of λ.

3.1. Convergence and Computational Complexity

Providing either one or both of the constituent subfilters converge due to the convex nature of the hybrid filter it too will also converge and will always perform at least as well as the better of the two subfilters [3]. The computational complexity of the hybrid filter is naturally a combination of the computational complexities of the algorithms used to update the constituent subfilters. The additional complexity required for the update of the mixing parameter λ is minimal requiring only an additional 4 multiplications and 5 additions and only becomes relevant if the overall weight update $\mathbf{w}(k) = \lambda(k)\mathbf{w}_1(k) + (1 - \lambda(k))\mathbf{w}_2(k)$ is also of interest.

4. TRACKING CHANGES IN SIGNAL MODALITY

In order to illustrate the the capability of the hybrid filter to track changes in signal modality example hybrid filters using combinations of subfilters suited to linear inputs and either nonlinear or sparse inputs have been designed.

4.1. Nonlinear Hybrid filter

For the nonlinear hybrid filter, the constituent finite impulse response (FIR) subfilters (one linear and the other saturation type nonlinear) were trained by the normalised least mean square (NLMS) algorithm [9] and the normalised nonlinear gradient descent (NNGD) algorithm [10]. These two algorithms were chosen to train the subfilters as the NLMS is widely used and known for its robustness and excellent steady state properties whereas the NNGD has faster convergence and better tracking capabilities making it more suited to nonlinear inputs than the NLMS. By exploiting these properties it is possible to show that the hybrid filter has excellent tracking capabilities for signals.

The output of the NLMS trained subfilter y_{NLMS} is generated from [9]

$$\begin{aligned} y_{NLMS}(k) &= \mathbf{x}^T(k)\mathbf{w}_{NLMS}(k) \\ e_{NLMS}(k) &= d(k) - y_{NLMS}(k) \\ \mathbf{w}_{NLMS}(k+1) &= \mathbf{w}_{NLMS}(k) + \\ &\quad \frac{\mu_{NLMS}}{\|\mathbf{x}(k)\|_2^2 + \varepsilon(k)} e_{NLMS}(k)\mathbf{x}(k) \end{aligned} \tag{5}$$

and y_{NNGD} is the corresponding output of the NNGD trained subfilter given by [10]

$$\begin{aligned} y_{NNGD}(k) &= \Phi\left(net(k)\right) \\ net(k) &= \mathbf{x}^T(k)\mathbf{w}_{NNGD}(k) \\ e_{NNGD}(k) &= d(k) - y_{NNGD}(k) \\ \mathbf{w}_{NNGD}(k+1) &= \mathbf{w}_{NNGD}(k) + \\ &\quad \eta(k)\Phi'\left(net(k)\right) e_{NNGD}(k)\mathbf{x}(k) \\ \eta(k) &= \frac{1}{C + \left[\Phi'\left(net(k)\right)\right] \|\mathbf{x}(k)\|_2^2} \end{aligned} \tag{6}$$

where the step-size parameter of the NLMS filters is μ_{NLMS} and ε is the regularisation term. In the case of the NNGD $\Phi(\cdot)$ represents a nonlinear activation function and C a constant representing the ignored higher terms, for simulation purposes these were $\tanh(\cdot)$ and unity respectively.

4.2. Sparse Hybrid Filter

To track the changes in the sparseness of a signal the subfilters of the hybrid filter were trained by the signed sparse LMS (SSLMS) [11] and the NLMS, the NLMS was selected for the nonsparse filter as it was found to be a better choice than the LMS due to its faster convergence speeds allowing it to adapt quickly to changes in the input signal (preventing the sparse filter from dominating). The output of the NLMS trained subfilter is given as above (5) and the output of the corresponding SSLMS trained subfilter $y_{SSLMS}(k)$ is given by

$$\begin{aligned} y_{SSLMS}(k) &= \mathbf{x}^T(k)\mathbf{w}_{SSLMS}(k) \\ e_{SSLMS}(k) &= d(k) - y_{SSLMS}(k) \\ \mathbf{w}_{SSLMS}(k+1) &= \mathbf{w}_{SSLMS}(k) + \\ &\quad \mu\left(|\mathbf{w}_{SSLMS}(k)| + \varepsilon\right) e_{SSLMS}(k)\mathbf{x}(k) \end{aligned} \tag{7}$$

4.3. Simulations

By evaluating the resultant hybrid filters in an adaptive one step ahead prediction setting with the length of the adaptive filters set to $N = 10$, it is possible to illustrate the ability of the hybrid filter to identify the modality of a signal of interest. The behaviour of λ has been investigated for benchmark synthetic linear, nonlinear and sparse inputs. Values of λ were averaged over a set of 100 independent simulation runs, for the inputs described by a stable linear AR(4) process:

$$\begin{aligned} x(k) = 1.79x(k-1) - 1.85x(k-2) + 1.27x(k-3) \\ - 0.41x(k-4) + n(k) \end{aligned} \tag{8}$$

a benchmark nonlinear signal [12]:

$$x(k+1) = \frac{x(k)}{1 + x^2(k)} + n^3(k) \tag{9}$$

and a benchmark sparse distribution [11], where $n(k)$ is a zero mean, unit variance white Gaussian process. The convex combinations were presented with an input signal which alternated from linear (8) to nonlinear (9) then linear to sparse. The input signal was alternated every 100 samples and the corresponding dynamics of the mixing parameters $\lambda(k)$ are shown in Fig. 4, where a value of $\lambda = 1$ corresponds to the output of the NNGD/SSLMS trained subfilters and a value of $\lambda = 0$ corresponds to the output of the NLMS trained subfilters. It is clear from Fig. 4 that the value of λ adapts to be dominated by the filter most suited to the current dynamics of the input signal. As expected (as sparsity can be considered a subset of nonlinearity) the nonlinear hybrid filter obtains similar results for both the nonlinear and sparse inputs, whereas the sparse hybrid filter shows a marked difference in levels of sparsity for the same inputs.

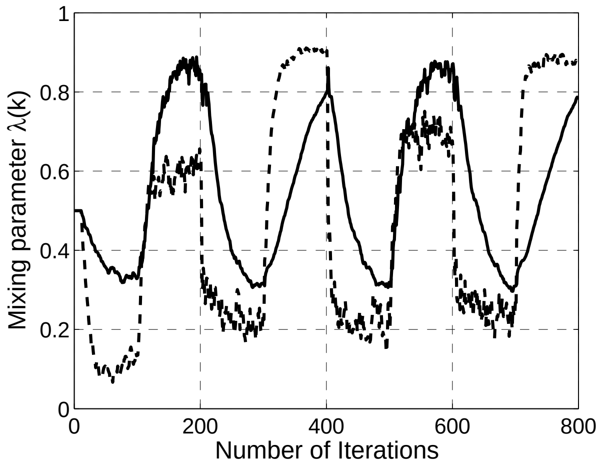

Fig. 4. Evolution of the mixing parameters λ for hybrid filters combining NNGD and NLMS subfilters (solid line) and SSLMS and NLMS (broken line) for an input signal alternating from linear to nonlinear and linear to sparse every 100 samples

5. TRACKING MULTIPLE CHANGES IN SIGNAL MODALITY

These algorithms have also been shown to give good results not only on synthetically generated data but also real world data in the form of EEG data from epileptic seizures [13] and speech data [14]. It is natural to consider whether these results can therefore be combined to allow us to track not only changes in nonlinearity but also at the same time changes in sparsity, this would prove of particular interest as changes in the sparseness of a signal can be considered to be a subset of the changes in the nonlinearity of a signal. Intuitively there should be a certain degree of correlation between the changes in the evolution of the mixing parameters λ of the sparse hybrid filter and that of the nonlinear hybrid filter.

Figure 5 shows the response of λ for the both the nonlinear hybrid filter and the sparse hybrid filter for the alternating input signal previously described, with the solid lines representing the linear sections, the broken lines the nonlinear sections and the dotted lines the sparse sections. For the linear sections although the evolution of the two λs do not follow the same path, there is an obvious correlation between them and the difference in responses can be attributed to the different learning rates of the subfilters of each hybrid filter. For the nonlinear and the sparse signals, however, the sparsity and saturation type nonlinearity are different phenomena and the sparse and nonlinear filter behaved differently. This representation is similar to the phase space representation in chaos theory, and allows for the signal modality characterisation to be considered within the framework of nonlinear dynamics. These results highlight the use of this technique in building a complete understanding of the nature of signals and has natural extensions both by using third dimensions (in this case combining the sparse and nonlinear filters in a hybrid filter) and also by using alternative filters to explore different signal characteristics.

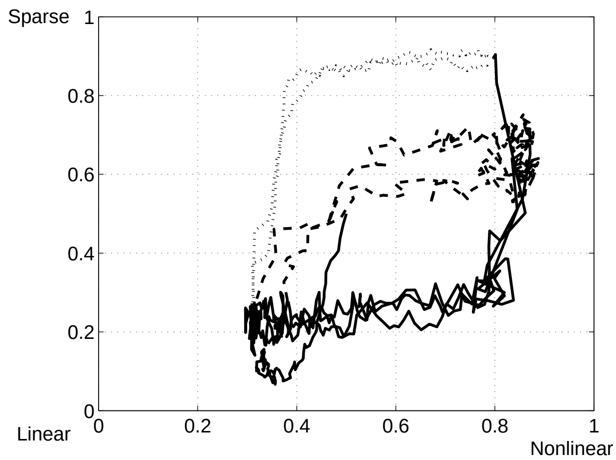

Fig. 5. Comparison of the evolution of the mixing parameters for linear/nonlinear and sparse/nonsparse for an input signal alternating every 100 samples. Solid line: nonlinear sections, broken line: linear sections, dotted line: sparse sections

To demonstrate the application of this method to real-world data, two sets of EEG data showing the onset of epileptic seizures were analysed. Figure 6 shows the EEG data, along with the corresponding evolution of the mixing parameters λ for both hybrid filters and the resultant changes in nonlinearity against sparsity. These results show that the proposed approach can not only effectively detect changes in the nature of the EEG signals which can be very difficult to achieve otherwise, but also identify which are changes in nonlinearity and which are also changes in sparsity.

6. CONCLUSIONS

We have highlighted that as well as offering improved performance it is also possible to use convex hybrid filters to gain information and track changes in signal modality. This is achieved through exploiting the different performance capabilities of key adaptive filtering algorithms and tracking the evolution of the adaptive convex mixing parameter λ within the hybrid filter structure. We have also presented a method by which it is possible to build on this information by combining the responses of several mixing parameters to obtain a

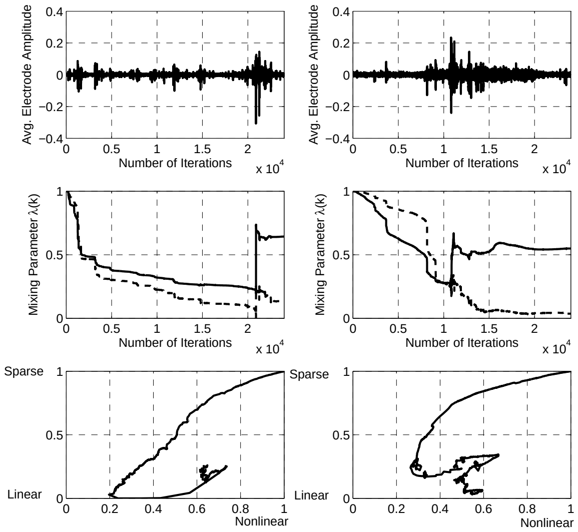

Fig. 6. Top panels: EEG epileptic seizure data. Middle panels: corresponding evolution of λ solid line: nonlinear hybrid filter, broken line: sparse hybrid filter. Bottom panels comparison of nonlinearity and sparsity, evolution over time starting from coordinates (1,1)

more complete understanding of the nature of signals.

7. REFERENCES

[1] S.S. Kozat and A.C. Singer, "Multi-stage adaptive signal processing algorithms," in *Proceedings of the IEEE Sensor Array and Multichannel Signal Processing Workshop*, 2000, pp. 380–384.

[2] A. Cichocki and R. Unbehauen, *Neural Networks for Optimisation and Signal Processing*, Wiley, 1993.

[3] J. Arenas-Garcia, A.R. Figueiras-Vidal, and A.H. Sayed, "Mean-square performance of a convex combination of two adaptive filters," *IEEE Transactions on Signal Processing*, vol. 54, no. 3, pp. 1078–1090, 2006.

[4] D. Mandic, P. Vayanos, C. Boukis, B. Jelfs, S.L. Goh, T. Gautama, and T. Rutkowski, "Collaborative adaptive learning using hybrid filters," in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2007*, 15-20 April 2007, vol. 3, pp. 921–924.

[5] T. Schreiber, "Interdisciplinary application of nonlinear time series methods," *Physics Reports*, vol. 308, no. 1, pp. 1–64, 1999.

[6] T. Schreiber and A. Schmitz, "Discrimination power of measures for nonlinearity in a time series," *Physical Review E*, vol. 55, no. 5, pp. 5443–5447, 1997.

[7] H. Mizuta, M. Jibu, and K. Yana, "Adaptive estimation of the degree of system nonlinearity," in *Proceedings IEEE Adaptive Systems for Signal Processing and Control Symposium (AS-SPCC)*, 2000, pp. 352–356.

[8] B. Jelfs, P. Vayanos, S. Javidi, V.S.L. Goh, and D.P. Mandic, "Collaborative adaptive filters for online knowledge extraction and information fusion," in *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, D. Mandic, M Golz, A. Kuh, D. Obradovic, and T. Tanaka, Eds. Springer, 2007.

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

[10] D.P. Mandic, "NNGD algorithm for neural adaptive filters," *Electronics Letters*, vol. 36, no. 9, pp. 845–846, April 2000.

[11] R.K. Martin, W.A. Sethares, R.C. Williamson, and C.R. Johnson, Jr., "Exploiting sparsity in adaptive filters," *IEEE Trans. Signal Proc.*, vol. 50, no. 8, pp. 1883–1894, 2002.

[12] 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.

[13] B. Jelfs, P. Vayanos, M. Chen, S. L. Goh, C. Boukis, T. Gautama, T. Rutkowski, T. Kuh, and D. Mandic, "An online method for detecting nonlinearity within a signal," in *Proceedings 10th International Conference on Knowledge-Based & Intelligent Information & Engineering Systems KES2006*, 2006, vol. 3, pp. 1216–1223.

[14] P. Vayanos, M. Chen, B. Jelfs, and D.P. Mandic, "Exploiting nonlinearity in adaptive signal processing," in *Proceedings ISCA Workshop on Non Linear Speech Processing*. 2007, Springer.