Signal Processing Techniques for Knowledge Extraction and Information Fusion

Danilo Mandic • Martin Golz • Anthony Kuh •
Dragan Obradovic • Toshihisa Tanaka
Editors

Signal Processing Techniques for Knowledge Extraction and Information Fusion

Springer

Editors

Danilo Mandic
Imperial College London
London
UK

Martin Golz
University of Schmalkalden
Schmalkalden
Germany

Anthony Kuh
University of Hawaii
Manoa, HI
USA

Dragan Obradovic
Siemens AG
Munich
Germany

Toshihisa Tanaka
Tokyo University of Agriculture
and Technology
Tokyo
Japan

ISBN: 978-0-387-74366-0 e-ISBN: 978-0-387-74367-7

Library of Congress Control Number: 2007941602

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

Preface

This book emanated from many discussions about collaborative research among the editors. The discussions have focussed on using signal processing methods for knowledge extraction and information fusion in a number of applications from telecommunications to renewable energy and biomedical engineering. They have led to several successful collaborative efforts in organizing special sessions for international conferences and special issues of international journals. With the growing interest from researchers in different disciplines and encouragement from Springer editors Alex Greene and Katie Stanne, we were spurred to produce this book.

Knowledge extraction and information fusion have long been studied in various areas of computer science and engineering, and the number of applications for this class of techniques has been steadily growing. Features and other parameters that describe a process under consideration may be extracted directly from the data, and so it is natural to ask whether we can exploit digital signal processing (DSP) techniques for this purpose. Problems where noise, uncertainty, and complexity play major roles are naturally matched to DSP. This synergy of knowledge extraction and DSP is still under-explored, but has tremendous potential. It is the underlying theme of this book, which brings together the latest research in DSP-based knowledge extraction and information fusion, and proposes new directions for future research and applications. It is fitting, then, that this book touches on globally important applications, including sustainability (renewable energy), health care (understanding and interpreting biomedical signals) and communications (extraction and fusing of information from sensor networks).

The use of signal processing in data and sensor fusion is a rapidly growing research area, and we believe it will benefit from a work such as this, in which both background material and novel applications are presented. Some of the chapters come from extended papers originally presented at the special sessions in ICANN 2005 and KES 2006. We also asked active researchers in signal processing with specializations in machine learning and multimodal signal processing to make contributions to augment the scope of the book.

This book is divided in four parts with four chapters each.

Collaborative Signal Processing Algorithms

Chapter 1 by Jelfs et al. addresses hybrid adaptive filtering for signal modality characterization of real-world processes. This is achieved within a collaborative signal processing framework which quantifies in real-time, the presence of linearity and nonlinearity within a signal, with applications to the analysis of EEG data. This approach is then extended to the complex domain and the degree of nonlinearity in real-world wind measurements is assessed.

In Chap. 2, Hirata et al. extend the wind modelling approaches to address the control of wind farms. They provide an analysis of the wind features which are most relevant to the local forecasting of the wind profile. These are used as prior knowledge to enhance the forecasting model, which is then applied to the yaw control of a wind turbine.

A collaborative signal processing framework by means of hierarchical adaptive filters for the detection of sparseness in a system identification setting is presented in Chap. 3, by Boukis and Constantinides. This is supported by a thorough analysis with an emphasis on unbiasedness. It is shown that the unbiased solution corresponds to existence of a sparse sub-channel, and applications of this property are highlighted.

Chapter 4 by Zhang and Chambers addresses the estimation of the reverberation time, a difficult and important problem in room acoustics. This is achieved by blind source separation and adaptive noise cancellation, which in combination with the maximum likelihood principle yields excellent results in a simulated high noise environment. Applications and further developments of this strategy are discussed.

Signal Processing for Source Localization

Kuh and Zhu address the problem of sensor network localization in Chap. 5. Kernel methods are used to store signal strength information, and complex least squares kernel regression methods are employed to train the parameters for the support vector machine (SVM). The SVM is then used to estimate locations of sensors, and to track positions of mobile sensors. The chapter concludes by discussing distributed kernel regression methods to perform localization while saving on communication and energy costs.

Chapter 6, by Lenz et al., considers adaptive localization in wireless networks. They introduce an adaptive approach for simultaneous localization and learning based on theoretical propagation models and self-organizing maps, to demonstrate that it is possible to realize a self-calibrating positioning system with high accuracies. Results on real-world DECT and WLAN groups support the approach.

In Chap. 7, Host-Madsen et al. address signal processing methods for Doppler radar heart rate monitoring. This provides unobtrusive and ubiquitous detection of heart and respiration activity from distance. By leveraging recent advances in signal processing and wireless communication technologies, the authors explore robust radar monitoring techniques through MIMO signal processing. The applications of this method include health monitoring and surveillance.

Obradovic et al. present the fusion of onboard sensors and GPS for real-world car navigation in Chap. 8. The system is based on the position estimate obtained by Kalman filtering and GPS, and is aided by corrections provided by candidate trajectories on a digital map. In addition, fuzzy logic is applied to enhance guidance. This system is in operation in a number of car manufacturers.

Information Fusion in Imaging

In Chap. 9, Chumerin and Van Hulle consider the detection of independently moving objects as a component of the obstacle detection problem. They show that the fusion of information obtained from multiple heterogeneous sensors has the potential to outperform the vision-only description of driving scenes. In addition, the authors provide a high-level sensor fusion model for detection, classification, and tracking in this context.

Aghajan, Wu, and Kleihorst address distributed vision networks for human pose analysis in Chap. 10. This is achieved by collaborative processing and data fusion mechanisms, and under a low bandwidth communication constraint. The authors employ a 3D human body model as the convergence point of the spatiotemporal and feature fusion. This model also allows the cameras to interact and helps the evaluation of the relative values of the derived features.

The application of information fusion in E-cosmetics is addressed by Tsumura et al. in Chap. 11. The authors develop a practical skin color analysis and synthesis (fusion) technique which builds upon both the physical background and physiological understanding. The appearance of the reproduced skin features is analysed with respect to a number of practical constraints, including the imaging devices, illuminants, and environments.

Calhoun and Adalı consider the fusion of brain imaging data in Chap. 12. They utilize multiple image types to take advantage of the cross information. Unlike the standard approaches, where cross information is not taken into account, this approach is capable of detecting changes in functional magnetic resonance imaging (fMRI) activation maps. The benefits of the information fusion strategy are illustrated by real-world examples from neurophysiology.

Knowledge Extraction in Brain Science

Chapter 13, by Mandic et al. considers the "data fusion via fission" approach realized by empirical mode decomposition (EMD). Extension to the complex domain also helps to extract knowledge from processes which are strongly dependent on synchronization and phase alignment. Applications in real-world brain computer interfaces, e.g., in brain prosthetics and EEG artifact removal, illustrate the usefulness of this approach.

In Chap. 14, Rutkowski et al. consider some perceptual aspects of the fusion of information from multichannel EEG recordings. Time–frequency EMD features, together with the use of music theory, allow for a convenient and unique audio feedback in brain computer and brain machine (BCI/BMI) interfaces. This helps to ease the understanding of the notoriously difficult to analyse EEG data.

Cao and Chen consider the usefulness of knowledge extraction in brain death monitoring applications in Chap. 15. They combine robust principal factor analysis with independent component analysis to evaluate the statistical significance of the differences in EEG responses between quasi-brain-death and coma patients. The knowledge extraction principles here help to make a binary decision on the state of the consciousness of the patients.

Chapter 16, by Golz and Sommer, addresses a multimodal approach to the detection of extreme fatigue in car drivers. The signal processing framework is based on the fusion of linear (power spectrum) and nonlinear (delay vector variance) features, and knowledge extraction is performed via automatic input variable relevance detection. The analysis is supported by results from comprehensive experiments with a range of subjects.

London,
October 2007

Danilo Mandic
Martin Golz
Anthony Kuh
Dragan Obradovic
Toshihisa Tanaka

Acknowledgement

On behalf of the editors, I thank the authors for their contributions and for meeting such tight deadlines, and the reviewers for their valuable input.

The idea for this book arose from numerous discussions in international meetings and during the visits of several authors to Imperial College London. The visit of A. Kuh was made possible with the support of the Fulbright Commission; the Royal Society supported visits of M. Van Hulle and T. Tanaka; the Japan Society for the Promotion of Science (JSPS) also supported T. Tanaka.

The potential of signal processing for knowledge extraction and sensor, data, and information fusion has become clear through our special sessions in international conferences, such as ICANN 2005 and KES 2006, and in our special issue of the International Journal of VLSI Signal Processing Systems (Springer 2007). Perhaps the first gentle nudge to edit a publication in this area came from S.Y. Kung, who encouraged us to organise a special issue of his journal dedicated to this field. Simon Haykin made me aware of the need for a book covering this area and has been inspirational throughout.

I also thank the members of the IEEE Signal Processing Society Technical Committee on Machine Learning for Signal Processing for their vision and stimulating discussions. In particular, Tülay Adalı, David Miller, Jan Larsen, and Marc Van Hulle have been extremely supportive. I am also grateful to the organisers of MLSP 2005, KES 2006, MLSP 2007, and ICASSP 2007 for giving me the opportunity to give tutorial and keynote speeches related to the theme of this book. The feedback from these lectures has been most valuable.

It is not possible to mention all the colleagues and friends who have helped towards this book. For more than a decade, Tony Constantinides has been reminding me of the importance of fixed point theory in this area, and Kazuyuki Aihara and Jonathon Chambers have helped to realise the potential of information fusion for heterogeneous measurements. Maria Petrou has been influential in promoting data fusion concepts at Imperial. Andrzej Cichocki and his team from RIKEN have provided invigorating discussions and continuing support.

A special thanks to my students who have been extremely supportive and helpful. Beth Jelfs took on the painstaking job of going through every chapter and ensuring the book compiles. A less dedicated and resolute person would have given up long before the end of this project. Soroush Javidi has created and maintained our book website, David Looney has undertaken a number of editing jobs, and Ling Li has always been around to help.

Henry Goldstein has helped to edit and make this book more readable. Finally, I express my appreciation to the signal processing tradition and vibrant research atmosphere at Imperial, which have made delving into this area so rewarding.

Imperial College London, *Danilo Mandic*
October 2007

Contents

Contributors

Tülay Adalı
University of Maryland Baltimore County, Baltimore
MD 21250, USA
adali@umbc.edu

Hamid Aghajan
Department of Electrical Engineering
Stanford University
CA, USA
hamid@wsnl.stanford.edu

Kazuyuki Aihara
Institute of Industrial Science
The University of Tokyo
153-8505, Japan
aihara@sat.t.u-tokyo.ac.jp

Joachim Bamberger
Siemens AG
Otto-Hahn-Ring 6
81730 Munich, Germany
joachim.bamberger@siemens.com

Bruno Betoni Parodi
Siemens AG
Otto-Hahn-Ring 6
81730 Munich, Germany
betoni@gmail.com

Olga Boric-Lubecke
Kai Sensors, Inc./University of Hawaii, Honolulu
HI 96822, USA
olga@ieee.org

Christos Boukis
Athens Information Technology
Peania/Athens
19002, Greece
cbou@ait.edu.gr

Vince Calhoun
The MIND Institute/University of New Mexico, 1101 Yale Boulevard
Albuquerque, NM 87131, USA
vcalhoun@unm.edu

Jianting Cao
Saitama Institute of Technology
Saitama
369-0293, Japan
cao@sit.ac.jp

Jonathon A. Chambers
Advanced Signal Processing Group
Loughborough University
Loughborough, UK
j.a.chambers@lboro.ac.uk

Zhe Chen
Massachusetts General Hospital
Harvard Medical School
Boston, MA 02114, USA
zhechen@neurostat.mgh.harvard.edu

Nikolay Chumerin
Katholieke Universiteit Leuven
Herestraat 49, bus 1021
B-3000 Leuven, Belgium
nikolay.chumerin@med.kuleuven.be

Andrzej Cichocki
Brain Science Institute
RIKEN, Saitama
351-0198, Japan
cia@brain.riken.jp

Anthony Constantinides
Imperial College London
Exhibition Road, London
SW7 2BT, UK
agc@imperial.ac.uk

Vanessa Su Lee Goh
Nederlandse Aardolie Maatschappij
B.V., PO Box 28000
9400 HH Assen, The Netherlands
Vanessa.Goh@shell.com

Martin Golz
University of Applied Sciences
Schmalkalden
Germany
m.golz@fh-sm.de

Uwe Hanebeck
Universitt Karlsruhe (TH)
Kaiserstr. 12
76131 Karlsruhe, Germany
uwe.hanebeck@ieee.org

Kai Heesche
Siemens AG
Otto-Hahn-Ring 6
81730 Munich, Germany
kai.heesche@siemens.com

Yoshito Hirata
Institute of Industrial Science
The University of Tokyo
153-8505, Japan
yoshito@sat.t.u-tokyo.ac.jp

Joachim Horn
Helmut-Schmidt-University/
University of the Federal Armed
Forces, 22043 Hamburg, Germany
joachim.horn@hsu-hh.de

Anders Høst-Madsen
Kai Sensors, Inc./University
of Hawaii, Honolulu
HI 96822, USA
ahm@hawaii.edu

Soroush Javidi
Imperial College London
Exhibition Road, London
SW7 2BT, UK
soroush.javidi@imperial.ac.uk

Beth Jelfs
Imperial College London
Exhibition Road, London
SW7 2BT, UK
beth.jelfs@imperial.ac.uk

Richard Kleihorst
NXP Semiconductor Research
Eindhoven
The Netherlands
Richard.Kleihorst@nxp.com

Anthony Kuh
University of Hawaii
Honolulu
HI 96822, USA
kuh@spectra.eng.hawaii.edu

Henning Lenz
Siemens AG
Oestliche Rheinbrueckenstr. 50
76187 Karlsruhe, Germany
henning.lenz@siemens.com

Wai Yie Leong
Agency for Science, Technology
and Research, (A*STAR) SIMTech
71 Nanyang Drive, Singapore 638075
waiyie@ieee.org

David Looney
Imperial College London
Exhibition Road, London
SW7 2BT, UK
david.looney06@imperial.ac.uk

Victor M. Lubecke
Kai Sensors, Inc./University of
Hawaii, Honolulu
HI 96822, USA
lubecke@ieee.org

Danilo P. Mandic
Imperial College London
Exhibition Road, London
SW7 2BT, UK
d.mandic@imperial.ac.uk

Yoichi Miyake
Graduate School of Advanced
Integration Science, Chiba University
263-8522, Japan
miyake@faculty.chiba-u.jp

Toshiya Nakaguchi
Graduate School of Advanced
Integration Science, Chiba University
263-8522, Japan
nakaguchi@faculty.chiba-u.jp

Dragan Obradovic
Siemens AG
Otto-Hahn-Ring 6
81730 Munich, Germany
dragan.obradovic@siemens.com

Nobutoshi Ojima
Kao Corporation
Japan
ojima.nobutashi@kao.co.jp

Byung-Kwon Park
University of Hawaii
Honolulu
HI 96822, USA
byungp@hawaii.edu

Nicolas Petrochilos
University of Hawaii
Honolulu
HI 96822, USA
petro@ahi.eng.hawaii.edu

Tomasz Rutkowski
Brain Science Institute
RIKEN, Saitama
351-0198, Japan
tomek@brain.riken.jp

Markus Schupfner
Harman/Becker Automotive Systems
GmbH, Moosacherstr. 48
80809 Munich, Germany
MSchupfner@harmanbecker.com

David Sommer
University of Applied Sciences
Schmalkalden
Germany
d.sommer@fh-sm.de

George Souretis
Imperial College London
Exhibition Road, London
SW7 2BT, UK
g.souretis@imperial.ac.uk

Hideyuki Suzuki
Institute of Industrial Science
The University of Tokyo
153-8505, Japan
hideyuki@sat.t.u-tokyo.ac.jp

Andrei Szabo
Siemens AG
Otto-Hahn-Ring 6
81730 Munich, Germany
andrei.szabo@siemens.com

Toshihisa Tanaka
Tokyo University of Agriculture and Technology
Japan
tanakat@cc.tuat.ac.jp

Norimichi Tsumura
Graduate School of Advanced Integration Science, Chiba University
263-8522, Japan
tsumura@faculty.chiba-u.jp

Marc M. Van Hulle
Katholieke Universiteit Leuven
Herestraat 49, bus 1021
B-3000 Leuven, Belgium
marc.vanhulle@med.kuleuven.be

Phebe Vayanos
Imperial College London
Exhibition Road, London
SW7 2BT, UK
foivi.vayanos@imperial.ac.uk

Hui Wang
Siemens AG
Otto-Hahn-Ring 6
81730 Munich, Germany
hui.wang.ext@siemens.com

Chen Wu
Department of Electrical Engineering
Stanford University
CA, USA
chenwu@stanford.edu

Yonggang Zhang
Advanced Signal Processing Group
Loughborough University
Loughborough, UK
Y.Zhang5@lboro.ac.uk

Qin Zhou
Broadcom Inc.
USA
lucy.qinzhou@gmail.com

Chaopin Zhu
Juniper Networks
Sunnyvale
CA 94089, USA
czhu@juniper.net

Part I

Collaborative Signal Processing Algorithms

1

Collaborative Adaptive Filters for Online Knowledge Extraction and Information Fusion

Beth Jelfs, Phebe Vayanos, Soroush Javidi, Vanessa Su Lee Goh,
and Danilo P. Mandic

We present a method for extracting information (or knowledge) about the nature of a signal. This is achieved by employing recent developments in signal characterisation for online analysis of the changes in signal modality. We show that it is possible to use the fusion of the outputs of adaptive filters to produce a single collaborative hybrid filter and that by tracking the dynamics of the mixing parameter of this filter rather than the actual filter performance, a clear indication as to the nature of the signal is given. Implementations of the proposed hybrid filter in both the real $\mathbb{R}$ and the complex $\mathbb{C}$ domains are analysed and the potential of such a scheme for tracking signal nonlinearity in both domains is highlighted. Simulations on linear and nonlinear signals in a prediction configuration support the analysis; real world applications of the approach have been illustrated on electroencephalogram (EEG), radar and wind data.

1.1 Introduction

Signal modality characterisation is becoming an increasingly important area of multidisciplinary research and large effort has been put into devising efficient algorithms for this purpose. Research in this area started in mid-1990s but its applications in machine learning and signal processing are only recently becoming apparent. Before discussing characterisation of signal modalities certain key properties for defining the nature of a signal should be outlined [8, 21]:

1. Linear (strict definition) – A linear signal is generated by a linear time-invariant system, driven by white Gaussian noise.
2. Linear (commonly adopted) – Definition 1 is relaxed somewhat by allowing the distribution of the signal to deviate from the Gaussian one, which can be interpreted as a linear signal from 1. measured by a static (possibly nonlinear) observation function.

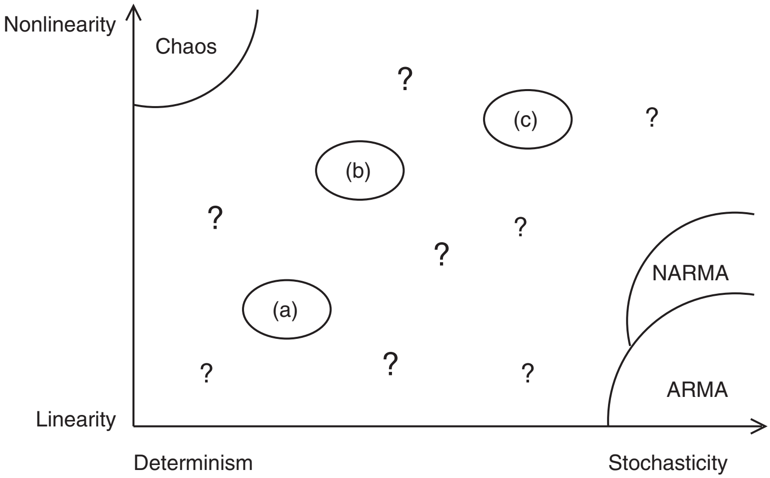

Fig. 1.1. Deterministic vs. stochastic nature or linear vs. nonlinear nature

3. Nonlinear – A signal that cannot be generated in the above way is considered nonlinear.
4. Deterministic (predictable) – A signal is considered deterministic if it can be precisely described by a set of equations.
5. Stochastic – A signal that is not deterministic.

Figure 1.1 (modified from [19]) illustrates the range of signals spanned by the characteristics of nonlinearity and stochasticity. While signals with certain characteristics are well defined, for instance chaotic signals (nonlinear and deterministic) or those produced by autoregressive moving average (ARMA) models (linear and stochastic signals), these represent only the extremes in signal nature and do not highlight the majority of signals which do not fit into such classifications. Due to the presence of such factors as noise or uncertainty, any real world signals are represented in the areas (a), (b), (c) or '?'; these are significant areas about which we know little or nothing. As changes in the signal nature between linear and nonlinear and deterministic and stochastic can reveal information (knowledge) which is critical in certain applications (e.g., health conditions) the accurate characterisation of the nature of signals is a key prerequisite prior to choosing a signal processing framework.

The existing algorithms in this area are based on hypothesis testing [6, 7, 20] and describe the signal changes in a statistical manner. However, there are very few online algorithms which are suitable for this purpose. The purpose of the approach described in this chapter is to introduce a class of online algorithms which can be used not only to identify, but also to track changes in the nature of the signal (signal modality detection).

One intuitive method to determine the nature of a signal has been to present the signal as input to two adaptive filters with different characteristics, one nonlinear and the other linear. By comparing the responses of each filter, this can be used to identify whether the input signal is linear or not. While this is a very useful simple test for signal nonlinearity, it does not provide an online solution. There are additional ambiguities due to the need to choose many parameters of the corresponding filters and this approach does not rely on the "synergy" between the filters considered.

1.1.1 Previous Online Approaches

In [17] an online approach is considered which successfully tracks the degree of nonlinearity of a signal using adaptive algorithms, but relies on a parametric model to effectively model the system to provide a true indication of the degree of nonlinearity. Figure 1.2 shows an implementation of this method using a third-order Volterra filter and the normalised least mean square (NLMS) algorithm with a step size $\mu = 0.008$ to update the system parameters. The system input and output can be described by

$$
\begin{aligned}
u[k] &= \sum_{i=0}^{I} a_i x[k-i] \text{ where } I = 2 \text{ and } a_0 = 0.5,\ a_1 = 0.25,\ a_2 = 0.125, \\
y[k] &= F(u[k];k) + \eta[k], \qquad (1.1)
\end{aligned}
$$

where $x[k]$ are i.i.d uniformly distributed over the range $[-0.5, 0.5]$ and $\eta[k] \sim \mathcal{N}(0, 0.0026)$. The function $F(u[k];k)$ varies with k

$$
F(u[k];k) = \begin{cases} u^3[k] \text{ for } 10{,}000 < k \leq 20{,}000, \\ u^2[k] \text{ for } 30{,}000 < k \leq 40{,}000, \\ u[k] \text{ at all other times.} \end{cases} \qquad (1.2)
$$

The output $y[k]$ can be seen in the first trace of Fig. 1.2, the second and third traces show the residual estimation errors of the optimal linear system and

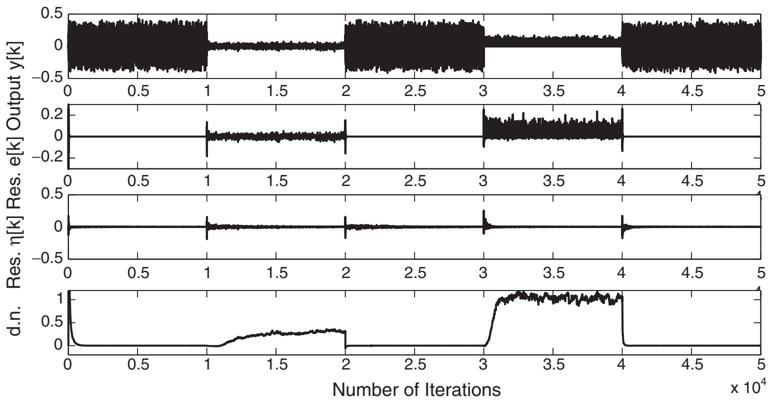

Fig. 1.2. Estimated degree of signal nonlinearity for an input alternating from linear to nonlinear

Volterra system, respectively, the final trace is the estimated degree of signal nonlinearity. While these results show that this approach can detect changes in nonlinearity and is not affected by the presence of noise, this may be largely due to the nature of the input signal in question being particularly suited to the Volterra model.

This type of method relies on the nature of the nonlinearity under observation being suited to the actual signal model; in real world situations it is not always possible to know the nonlinearity in advance, therefore their application is limited. To overcome these limitations, we propose a much more flexible method based on collaborative adaptive filtering.

1.1.2 Collaborative Adaptive Filters

Developing on the well-established tracking capabilities of adaptive filters using combinations of adaptive subfilters in a more natural way produces a single hybrid filter without the need for any knowledge of underlying signal generation models. Hybrid filters consist of multiple individual adaptive subfilters operating in parallel and all feeding into a mixing algorithm which produces the single output of the filter [4, 13]. The mixing algorithms are also adaptive and combine the outputs of each subfilter based on the estimate of their current performance on the input signal from their instantaneous output error.

Many previous applications of hybrid filters have focused mainly on the improved performance they can offer over the individual constituent filters. Our aim is to focus on one additional effect of the mixing algorithm, that is, to show whether it can give an indication of which filter is currently responding to the input signal most effectively. Therefore, intuitively by selecting algorithms which are particularly suited to one type of input signals, it is possible to cause the mixing algorithm to adapt according to fundamental properties of the input signal.

A simple form of mixing algorithm for two adaptive filters is a convex combination. Convexity can be described as [5]

$$\lambda x + (1 - \lambda) y \quad \text{where} \quad \lambda \in [0, 1]. \tag{1.3}$$

For x and y being two points on a line, as shown in Fig. 1.3, their convex mixture (1.3) will lie on the same line between x and y.

For convex mixing of the outputs of adaptive filters, it is intuitively clear that initially λ will adapt to favour the faster filter (that is the filter with faster learning rate) and following convergence it will favour the filter with

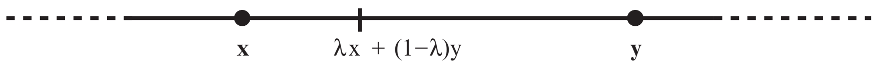

Fig. 1.3. Convexity

better steady-state properties[1]; should one of the subfilters fail to converge, the values of λ adapt such that the hybrid filter follows the stable subfilter [16]. The approach in this chapter focuses on observing the dynamics of mixing parameter λ, to allow conclusions to be drawn about the current nature of the input signal.

1.2 Derivation of The Hybrid Filter

Unlike the existing approaches to hybrid adaptive filters which focus on the quantitative performance of such filters, in this case the design of the hybrid filters is such that it should combine the characteristics of two distinctly different adaptive filters. Signal modality characterisation is achieved by making the value of the "mixing" parameter λ adapt according to the fundamental dynamics of the input signal. In this chapter we illustrate applications of this method for characterisation of nonlinearity and complexity on both synthetic and real world data, but this method can be equally well applied to any other signal characteristics. With that in mind we start from the general derivation of the convex hybrid filter before moving on to specific implementations.

Figure 1.4 shows the block diagram of a hybrid filter consisting of two adaptive filters 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

$$y(k) = \lambda(k)y_1(k) + (1 - \lambda(k))\, y_2(k), \tag{1.4}$$

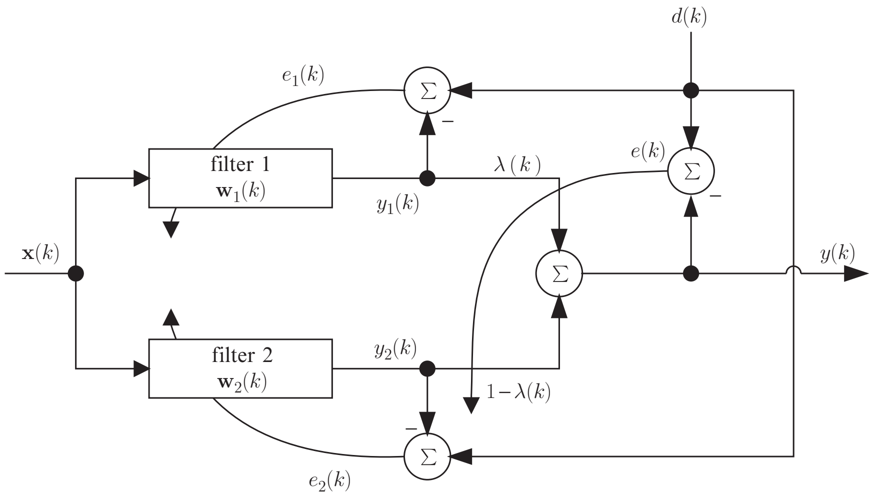

Fig. 1.4. Convex combination of adaptive filters (hybrid filter)

[1] Unlike traditional search then converge approaches this method allows for potentially nonstationary data.

where $y_1(k) = \mathbf{x}^{\mathrm{T}}(k)\mathbf{w}_1(k)$ and $y_2(k) = \mathbf{x}^{\mathrm{T}}(k)\mathbf{w}_2(k)$ are the outputs of the two subfilters with corresponding weight vectors $\mathbf{w}_1(k) = [w_{1,1}(k), \ldots, w_{1,N}(k)]^{\mathrm{T}}$ and $\mathbf{w}_2(k) = [w_{2,1}(k), \ldots, w_{2,N}(k)]^{\mathrm{T}}$ which are dependent on the algorithms used to train the subfilters based on the common input vector $\mathbf{x}(k) = [x_1(k), \ldots, x_N(k)]^{\mathrm{T}}$ for filters of length N.

To preserve the inherent characteristics of the subfilters, which are the basis of our approach, the constituent subfilters are updated by their own errors $e_1(k)$ and $e_2(k)$, using a common desired signal $d(k)$, whereas the parameter λ is updated based on the overall error $e(k)$. The convex mixing parameter $\lambda(k)$ is updated based on minimisation of the quadratic cost function $E(k) = \frac{1}{2}e^2(k)$ using the following gradient adaptation:

$$\lambda(k+1) = \lambda(k) - \mu_\lambda \nabla_\lambda E(k)_{|\lambda=\lambda(k)}, \tag{1.5}$$

where μ_λ is the adaptation step-size. From (1.4) and (1.5), using an LMS type adaptation, the λ update can be obtained as

$$\lambda(k+1) = \lambda(k) - \frac{\mu_\lambda}{2}\frac{\partial e^2(k)}{\partial \lambda(k)} = \lambda(k) + \mu_\lambda e(k)(y_1(k) - y_2(k)). \tag{1.6}$$

To ensure the combination of adaptive filters remains a convex function, it is critical that λ remains within the range $0 \le \lambda(k) \le 1$. In [4] the authors obtained this through the use of a sigmoid function as a post-nonlinearity to bound $\lambda(k)$. Since, to determine the changes in the modality of a signal, we are not interested in the overall performance of the filter but in the dynamics of parameter λ, the use of a sigmoid function would interfere with true values of $\lambda(k)$ and was therefore not appropriate. In this case a hard limit on the set of allowed values for $\lambda(k)$ was therefore implemented.

1.3 Detection of the Nature of Signals: Nonlinearity

Implementations of the hybrid filter described above using the LMS algorithm [23] to train one of the subfilters and the generalised normalised gradient descent (GNGD) algorithm [15] for the other, have been used to distinguish the linearity/nonlinearity of a signal [11]. The LMS algorithm was chosen as it is widely used, known for its robustness and excellent steady-state properties whereas the GNGD algorithm has a faster convergence speed and better tracking capabilities. By exploiting these properties it is possible to show that due to the synergy and simultaneous mode of operation, the hybrid filter has excellent tracking capabilities for signals with extrema in their inherent linearity and nonlinearity characteristics.

The output of the LMS trained subfilter y_{LMS} is generated from [23]

$$\begin{aligned} y_{\mathrm{LMS}}(k) &= \mathbf{x}^{\mathrm{T}}(k)\mathbf{w}_{\mathrm{LMS}}(k), \\ e_{\mathrm{LMS}}(k) &= d(k) - y_{\mathrm{LMS}}(k), \\ \mathbf{w}_{\mathrm{LMS}}(k+1) &= \mathbf{w}_{\mathrm{LMS}}(k) + \mu_{\mathrm{LMS}} e_{\mathrm{LMS}}(k)\mathbf{x}(k) \end{aligned} \tag{1.7}$$

and y_{GNGD} is the corresponding output of the GNGD trained subfilter given by [15]

$$\begin{aligned} y_{\mathrm{GNGD}}(k) &= \mathbf{x}^{\mathrm{T}}(k)\mathbf{w}_{\mathrm{GNGD}}(k) \\ e_{\mathrm{GNGD}}(k) &= d(k) - y_{\mathrm{GNGD}}(k) \\ \mathbf{w}_{\mathrm{GNGD}}(k+1) &= \mathbf{w}_{\mathrm{GNGD}}(k) + \frac{\mu_{\mathrm{GNGD}}}{\|\mathbf{x}(k)\|_2^2 + \varepsilon(k)} e_{\mathrm{GNGD}}(k)\mathbf{x}(k) \\ \varepsilon(k+1) &= \varepsilon(k) - \rho\mu_{\mathrm{GNGD}} \frac{e_{\mathrm{GNGD}}(k)e_{\mathrm{GNGD}}(k-1)\mathbf{x}^{\mathrm{T}}(k)\mathbf{x}(k-1)}{\left(\|\mathbf{x}(k-1)\|_2^2 + \varepsilon(k-1)\right)^2} \end{aligned} \tag{1.8}$$

where the step-size parameters of the filters are μ_{LMS} and μ_{GNGD}, and in the case of the GNGD ρ is the step-size adaptation parameter and ε the regularisation term.

By evaluating the resultant hybrid filter 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 and nonlinear inputs. Values of λ were averaged over a set of 1,000 independent simulation runs, for the inputs described by a stable linear AR(4) process:

$$x(k) = 1.79x(k-1) - 1.85x(k-2) + 1.27x(k-3) - 0.41x(k-4) + n(k) \tag{1.9}$$

and a benchmark nonlinear signal [18]:

$$x(k+1) = \frac{x(k)}{1 + x^2(k)} + n^3(k), \tag{1.10}$$

where $n(k)$ is a zero mean, unit variance white Gaussian process. The values of the step-sizes used were $\mu_{\mathrm{LMS}} = 0.01$ and $\mu_{\mathrm{GNGD}} = 0.6$. For the GNGD filter $\rho = 0.15$ and the initial value of the regularisation parameter was $\varepsilon(0) = 0.1$. Within the convex combination of the filters, filter 1 corresponds to the GNGD trained subfilter and filter 2 to the LMS trained subfilter, the step-size for the adaptation of $\lambda(k)$ was $\mu_{\lambda} = 0.05$ and the initial value[2] of $\lambda(0) = 1$.

From the curves shown in Fig. 1.5 it can be seen the value of $\lambda(k)$ for both inputs moves towards zero as the adaptation progresses. As expected, the output of the convex combination of adaptive filters approaches the output of the LMS filter y_{LMS} predominately. This is due to the better steady-state properties of the LMS filter when compared to the GNGD filter, which due to its constantly 'alert' state does not settle in the steady state as well as the LMS. In the early stages of adaptation, the nonlinear input (1.10) adapts to become dominated by the LMS filter much faster than the linear input and

[2] Since GNGD exhibits much faster convergence than LMS, it is natural to start the adaptation with $\lambda(0) = 1$. This way, we avoid possible artefacts that may arise due to the slow initial response to the changes in signal modality.

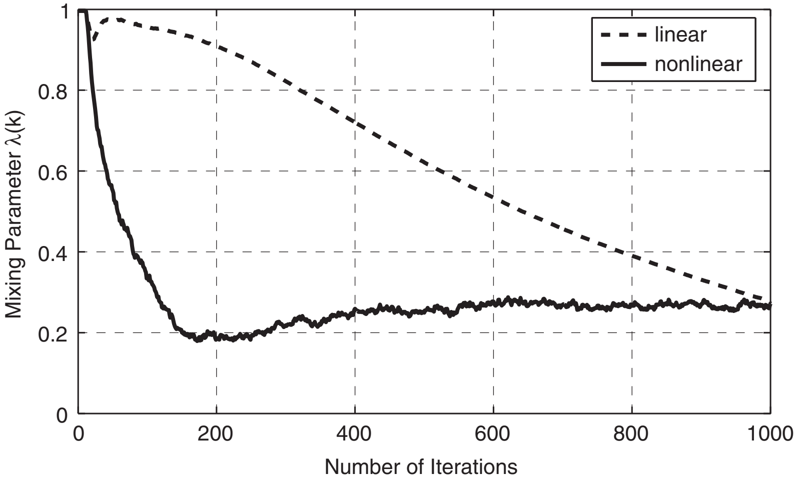

Fig. 1.5. Comparison of the mixing parameter λ for linear and nonlinear inputs

rapidly converges, whereas the linear input (1.9) changes much more gradually between the two filters.[3]

1.3.1 Tracking Changes in Nonlinearity of Signals

It is also possible to use changes in λ along the adaptation to track the changes in signal modality. Since the behaviour of λ as a response to the different inputs is clearly distinct, especially in the earliest stages of adaptation, the convex combination was presented with an input signal which alternated between linear (1.9) and nonlinear (1.10). The input signal was alternated every 200 samples and the corresponding dynamics of the mixing parameter $\lambda(k)$ are shown in Fig. 1.6. From Fig. 1.6 it is clear that the value of $\lambda(k)$ adapts in a way which ensures that the output of the convex combination is dominated by the filter most appropriate for the input signal characteristics.

To illustrate the discrimination ability of the proposed approach, the next set of simulations shows the results of the same experiment as in Fig. 1.6, but for a decreased number of samples between the alternating segments of data. Figure 1.7 shows the response of $\lambda(k)$ to the input signal alternating every 100 and 50 samples, respectively. There is a small anomaly in the values of λ immediately following the change in input signal from nonlinear to linear, which can be clearly seen in Fig. 1.7 around sample numbers $100i,\ i = 1, 2, \ldots,$ where the value of λ exhibits a small dip before it increases. This is due to the fact that the input to both the current AR process (1.9) and the tap inputs to both filters use previous nonlinear samples where we are in fact predicting the

[3] Both filters perform well on a linear input and are competing along the adaptation.

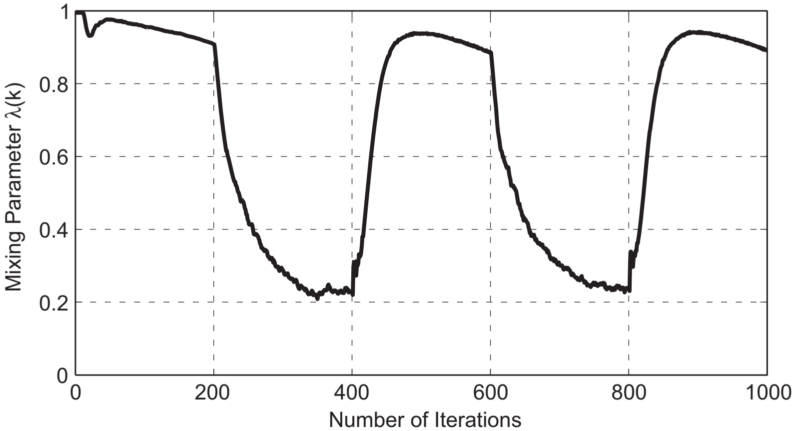

Fig. 1.6. Evolution of the mixing parameter λ for input nature alternating every 200 samples

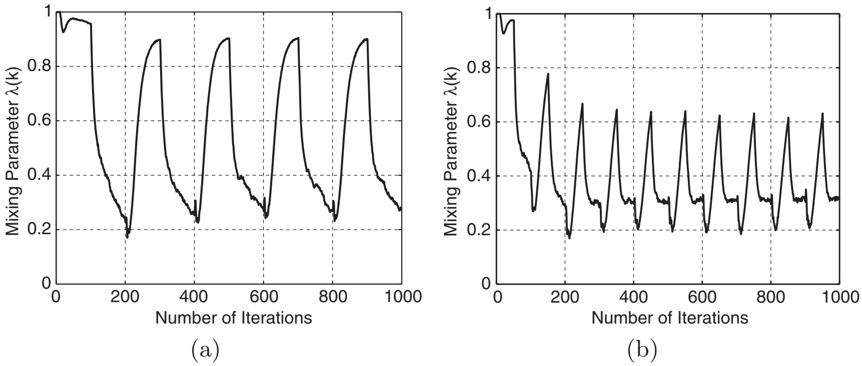

Fig. 1.7. Evolution of the mixing parameter λ for a signal with input nature alternating between linear to nonlinear. (**a**) Input signal nature alternating every 100 samples and (**b**) input signal nature alternating every 50 samples

first few "linear" samples. This does not become an issue when alternations between the input signals occur less regularly or if there is a more natural progression from "linear" to "nonlinear" in the the input signal.

Real World Applications

To examine the usefulness of this approach for the processing of real world signals, a set of EEG signals has been analysed. Following the standard practice, the EEG sensor signals were averaged across all the channels and any trends in the data were removed. Figure 1.8 shows the response of λ when applied to two different sets of EEG data from epileptic patients, both showing the