ELSEVIER
Neurocomputing 49 (2002) 439–443
NEUROCOMPUTING
www.elsevier.com/locate/neucom

Book review[1]

Adaptive blind signal and image processing: learning algorithms and applications
A. Cichocki, S. Amari, Wiley, New York, 2002, 586pp., ISBN 0471 60791 6

Blind signal and image processing (BSIP) is an exciting and emerging research topic in fields such as neural networks, advanced statistics, and biomedical signal/image processing, and, over the past decade, has established solid theoretical foundations and many real-world applications. The "blind" processing of signals, based on unsupervised learning in principal, do not assume any prior information in the form of desired training data, signal distributions, or parameter values of mixing systems. In fact, BSIP has recently become one of the hottest research topics in many application areas such as biomedical engineering, medical imaging, speech processing, remote sensing, and communication systems because the original sources and filtering/mixing models in these applications are usually unknown.

The main objective of the present book is to provide a comprehensive treatment of the problems, algorithms, and applications in BSIP, which are presented in three major areas: blind signal separation and extraction (BSS/BSE), independent component analysis (ICA), and multichannel blind deconvolution and equalization (MBD/MBE). Earlier, some of these terms are similarly, interchangeably, or sometimes confusingly treated in the literature. The present book gives clear definitions and identifies the principles and techniques relevant in each case. The term BSS is more general and refers to the broad problem of separating or estimating the waveforms of original sources from an array of observations or measurements, without knowing the characteristics of mixing or transmission channels. The ICA attempts to recover sources from such observations as independent as possible. On the other hand, BSS attempts to separate the sources even if they are not statistically independent. In MBD the sources, having unknown distributions and statistics, are presumed to be convolved by the transmission, or, more generally, mixing systems of unknown characteristics. The learning algorithms of BSS can be generalized or extended to MBD by using the relationship between convolution in time domain and multiplication in frequency domain. The blind identification problem emphasizes on identifying and characterizing the mixing or transmission channels by observing several sources simultaneously traveled through; the emphasis on blind equalization is the reconstruction of original interested signals transmitted through ambiguous or unknown channels. The blind extraction refers to the sequential approach of the BSS problem, in which a single source or set of sources with desired characteristics are extracted sequentially, one by one, from a set of their mixtures.

This book provides an extensive introduction to the techniques used in all areas of BSIP and offers a general overview of the basics of important ICA, BSS/E, and MBD/E solutions and algorithms as well as the fundamental mathematical background needed to understand and utilize them. The book delves into the techniques of second- and higher-order statistics, blind spatial and temporal decorrelation, robust whitening, blind filtering, matrix factorizations, robust principal component analysis (PCA), minor component analysis (MCA), sparse representations, automatic dimension reduction, features extraction in high dimensional data, and noise reduction and related problems. The algorithms are presented with

[1] Material to be included in this section can be submitted to: Prof. K.J. Cios, Computer Science and Engineering Dept., University of Colorado at Denver, Campus Box 109, 1200 Larimer Street, Denver, CO 80217-3364, USA. E-mail: kcois@carbon.cudenver.edu

0925-2312/02/$ - see front matter
PII: S0925-2312(02)00638-0

clear descriptions and sufficient mathematical detail and analysis; some interesting benchmarks are made available to compare performance of various unsupervised learning algorithms. Moreover, the book often provides illustrative examples to demonstrate the characteristics and performances of the described algorithms. The computer simulations are also presented for the validity and comparison of the derived learning algorithms.

The authors present the book as a textbook which can also be considered as a monograph of the authors since the results are new. The authors of the book, Andrzej Cichocki and Shun-Ichi Amari, are well known in the areas of ICA/BSS, PCA, and MBD/MBE. In this monograph, they detail all the relevant theory behind their pioneering work in areas like natural gradient approaches and adaptive blind algorithms for various applications. The text presents several recent results and ideas, mainly developed by the authors and their colleagues and collaborators, and further developments. With explanations and references to earlier techniques, the authors collate their findings in BSIP into a complete textbook. Furthermore, the book provides an in-depth coverage of various applications of BSIP in image processing, telecommunications, audio and biomedical signal processing, etc.

In the sequel, a review of the algorithms and applications described in the book is given briefly chapter by chapter.

Chapter 1 formulates fundamental problems of the blind signal processing (BSP) and illustrates several real-world problems as potential application areas. The chapter begins with giving important definitions and describing basic mathematical and physical models of BSP. The BSP problem is characterized by the absence of training data and a priori knowledge about the parameters of sources and mixing systems; its ramifications—BSS/BSE, ICA, MBD/MBE—are differentiated clearly with definitions. Possible extensions to handle noisy data and the non-linear case are highlighted; an illustration is given for MBD by using the state-space representation. The latter half of the chapter reviews several potential and promising applications: blind separation of electrocardiography (ECG) signals of fetus and mother, enhancement and decomposition of electromyography (EMG) signals, processing of electroencephalography (EEG) and magnetoencephalography (MEG) signals, cocktail party problem, and image restoration and understanding.

In Chapter 2, the authors deal with mathematical models of signals and images encountered in various applications, with an emphasis on solving the systems of linear algebraic equations arising from related blind processing. The chapter provides a tutorial to the problem of solving large overdetermined and underdetermined systems of linear equations especially when there is an uncertainty in the parameter values, and/or the systems are contaminated by noise. The overdetermined systems of linear equations are described as filtering, enhancement, deconvolution and identification problems while the underdetermined cases described as inverse and extrapolation problems. Least-squares (LS) problem and its recent extensions, such as weighted LS, extended total LS, data LS, and least absolute deviation, are first treated. A wide range of extrapolation, approximation, interpolation, regularization and inverse problems are then converted to minimum norm problems of underdetermined systems of linear equations. The chapter describes several novel on-line fast adaptive and iterative algorithms to deal with arbitrary noise statistics and special conditions. Sparse and overcomplete signal representations and solutions to associated problems are presented, demonstrating iterative LS solutions and FOCUSS algorithms for the minimum fuel problem (MFP), and the non-negative matrix factorization (NMF).

Chapter 3 introduces unsupervised learning algorithms to detect or extract useful features, regularities, and correlations of data or signals in order to detect, separate, or decorrelate underlying or representative signals with little or no knowledge. Normalized (constrained) Hebbian and anti-Hebbian learning rules are described as simple variants of basic unsupervised learning algorithms—in particular, learning algorithms for PCA, singular value decomposition (SVD), and MCA. The chapter provides an extensive overview of the most important adaptive algorithms for PCA, MCA, and SVD and their extensions such as robust and multistage PCA; the cost functions, derivations, and stability analyses of the learning rules are provided. Further, subspace methods based on AIC and MDL criteria and power method are described as robust techniques for dimensionality reduction.

Temporal, spatial, and spatio-temporal decorrelations, based only on second-order statistics (SOS), play vital roles in signal processing, particularly as preprocessing techniques. In Chapter 4, the authors present decorrelation and SOS techniques for robust blind identification as a basis for subspace methods in spectrum analysis and array processing. Variants of robust whitening transforms are discussed as preprocessing to improve performance of adaptive systems; after prewhitening, the BSS or ICA tasks usually become easier and better posed (less ill-conditioned) because the input to subsequent signal separation is described by a unitary matrix. The chapter discusses and analyzes a number of efficient and robust adaptive and batch algorithms for spatial whitening (eigenvalue decomposition (EVD)/SVD), orthogonalization (Gram–Schmidt orthogonalization), spatio-temporal and time-delayed blind decorrelation, and multistage approaches for BSS (multistage EVD/SVD). The chapter also describes several promising robust algorithms for blind identification and separation of non-stationary and/or colored sources, such as joint approximate diagonalization techniques including SOBI and JADE algorithms.

Chapter 5 describes the approaches for BSE, the sequential approach to BSS problem, in which a single source or set of sources is first extracted and the other sources are then extracted using a deflation process. The chapter introduces two different models and approaches for BSE. The first approach is based on higher-order statistics (HOS) by assuming that sources are mutually statistically independent and non-Gaussian (expect at most one) and using some measure of non-Gaussianity as the independence criteria; the algorithms based on Kurtosis (e.g., KuickNet) and fixed-point algorithms are discussed under this model. The second approach based on SOS assumes that source signals have some temporal structure, i.e., the sources are colored with different autocorrelation functions or equivalently different shape spectra; several on-line and batch algorithms for blind extraction of sources temporally correlated and/or corrupted by additive noise are developed and statistical analyses presented. Batch algorithms are implemented using bank of band pass filters.

The authors describe information theoretic approaches for learning in adaptive BSS and ICA, with an emphasis on natural gradient (NG) algorithms, in Chapter 6. Several measures of stochastic independence, such as mutual information (MI) and Kullback–Leibler divergence, are explored as promising contrast schemes; the corresponding NG algorithms are derived and analyzed in the general linear, orthogonal, and Stiefel manifolds. NG algorithms are extended for blind extraction of arbitrary group of sources and generalized for signals with wide class of distributions and unknown number of sources. The optimal choice of non-linear activation functions for various distributions, e.g., Gaussian, Laplacian, impulsive, and uniformly distributed signals, based on a generalized Gaussian-distributed model, is discussed. A family of efficient and flexible algorithms that exploit non-stationarity of signals is also derived; the dynamic properties and stability of the proposed NG algorithms are analyzed in detail.

Chapter 7 presents several associative or locally adaptive algorithms for the ICA, starting from a modified version of Jutten–Herault algorithm. These algorithms are comparatively efficient and simple and have straightforward electronic implementations; most are shown robust and perform well with noisy, badly scaled, or ill-conditioned signals. The proposed algorithms are further extended for the cases when the number and the distributions of sources are unknown (e.g., flexible ICA). In particular, locally adaptive Hebbian/anti-Hebbian learning algorithms and their implementations using multi-layer neural networks are presented; EASI algorithms are generalized. Finally, the optimal choice of non-linear activation function and general local stability conditions are discussed.

In Chapter 8, the authors propose several robust techniques for BSS and ICA to handle signals contaminated by large additive noise and extend the existing adaptive algorithms with equivariant properties to minimize the bias in the estimations of mixing or separating matrices, caused by measurement noise. Dynamical recurrent neural network approaches (e.g., Amari–Hopfield neural network) are presented for simultaneous estimation of the unknown mixing matrix and source signals and minimization of noise in the extracted output signals. The optimal choice of non-linear activation functions under various noise distributions, assuming a generalized Gaussian-distributed noise model, is also discussed.

In Chapter 9, the authors extend the adaptive NG algorithms for various MBD models and present relatively simple and efficient adaptive and batch algorithms for MBD/MBE of single-input/multiple-output (SIMO) and multiple-input/multiple-output (MIMO) minimum phase and non-minimum phase dynamical systems. Criteria for equalization and separation/deconvolution are discussed. By exploring the relationship between standard BSS/ICA and MBD, the algorithms derived in the previous chapters are extended for MBD/E; particularly, the NG approach for instantaneous mixtures is extended to convolutive dynamical models. A family of equivariant algorithms is derived; their stability and convergence properties analyzed. Furthermore, a Lie group and Riemannian metric are introduced on the manifold of FIR filters and using the isometry of the Riemannian metric, the NG on the FIR manifold is described. Based on the minimization of MI, an NG algorithm is presented for the causal minimum phase finite impulse response (FIR) multichannel filter. Next, using information back-propagation, an efficient implementation of the learning algorithm for the non-causal FIR filters is derived.

Chapter 10 introduces the concept of estimating functions, a unified framework to elucidate common structures and differences in most of the ICA/BSS and MBD algorithms. First, using the principles of information geometry, the estimating functions are defined in semi parametric statistical models which include the unknown functions as parameters of the model. The Newton method is automatically derived by the standardized estimating functions. Error and stability analyses are then presented, in terms of estimating functions, in order to facilitate choosing unknown parameters. Furthermore, the chapter extends the discussion of the convergence and efficiency of the batch estimator and NG learning for BSS/MBD, by using the semi parametric statistical model and the standardized estimating functions derived with efficient score functions. The geometrical properties of the manifold of the FIR filters based on the Lie group structure are presented. Finally, both batch learning and natural gradient learning are shown to be super efficient under given non-singular conditions.

The authors introduce the state-space description of dynamical systems as a powerful and flexible generalized model for BSS/MBD in Chapter 11. Some perspective and new insight into MIMO blind separation and filtering are initially given; NG learning algorithms are derived in the state-space framework. The main advantage of the state-space description is that not only it gives the internal description of the blind system, but also there exist various equivalent canonical types of state-space realizations, such as balanced realization and observable canonical forms, which are useful in parameterizing some specific classes of BSS/MBD models. The state-space model is used to provide much more general descriptions than the standard FIR convolutive filter models discussed in Chapter 9.

Chapter 12 extends and generalizes the state-space description of linear BSS/MBD models to non-linear dynamical models. The non-linear case is not only challenging but also intractable in the absence of any a priori knowledge about the non-linear mixing or filtering processes. Therefore, this chapter describes only some special non-linear models in order to simplify the problem for specific applications; the case of semi-blind separation and filtering is assumed, in which some information about the mixing and separating system and source signals are available. First, the invertibility and internal representation of the models are discussed. Then specific examples of models are considered in detail: the Wiener model, the Hammerstein model and non-linear autoregressive moving average (NARMA) models, and hyper radial basis function (RBF) neural network.

Appendix 13 presents mathematical preliminaries to help understanding of the material presented in the book; many useful definitions and formulas for matrix algebra and differentiation and of distance measures are given.

Each chapter of the book is presented in step-by-step and self-contained manner and easy to follow independently from other chapters. Many illustrative diagrams, graphs, and plots that accompany the presentation of basic concepts and algorithms offer valuable help for understanding the principles of BSS, ICA and MBD. Important algorithms are summarized and presented clearly in useful tables. Illustrative examples and computer implementation and simulations are provided for most algorithms. The accompanying CD-ROM includes an electronic interactive version of the book with hyperlinks, full-color figures, and text and a user friendly demo package for performing ICA and BSS/BSE.

Researchers, students, and practitioners from a variety of disciplines, who have interest in BSIP, will find this accessible volume both helpful and informative. This is a self-contained and complete volume on the subjects of BSS, PCA, ICA, MBD. The authors present many real-world applications in these areas and compare and discuss the pros and cons of various approaches and techniques. Compared with other ICA books, this book has much depth in mathematical and graphical interpretations in the subject matter and covers the broad subject of BSIP. Extensive, more than 1300, references on these areas are collected in a book for the first time. This book is a step toward making the subject of BSIP not only a common field of research but also a reference for those looking for new challenging topics and applications. I highly recommend the book to any reader interested in blind signal and image processing, especially independent component analysis.

Jagath C. Rajapakse
School of Computer Engineering, Nanyang Technological University,
Block N4 Nanyang Avenue, 639798
Singapore, Singapore
E-mail address: asjagath@ntu.edu.sg