KALMAN FILTER AND STATE-SPACE APPROACH TO BLIND DECONVOLUTION

L.-Q. Zhang, A. Cichocki and S. Amari
Brain-style Information Systems Research Group
RIEKN Brain Science Institute
Wako shi, Saitama 351-0198, JAPAN

Abstract. State-space model has been introduced as approach to blind deconvolution of dynamical systems. An efficient learning algorithm was developed for training the external parameters [20] and the Kalman filter was also applied to to compensate for the model bias and reduce the effect of noise [21] for linear systems. In this paper we generalize the Kalman filter to blind deconvolution of semi-nonlinear systems. First, we introduce a general framework of the state space approach for blind deconvolution and review the state of the art of state space approach for blind deconvolution. The adaptive natural gradient learning algorithm for updating external parameters is presented by minimizing a certain cost function, which is derived from mutual information of output signals. In order to compensate for the model bias and reduce the effect of noise, the extended Kalman filter is applied to the blind deconvolution setting. A new concept, called hidden innovation, is introduced so as to numerically implement the Kalman filter. A computer simulation is given to show the validity and effectiveness of the state-space approach.

INTRODUCTION

Blind separation/deconvolution of sources has attracted growing attention in the field of signal processing since it was introduced by Jutten and Herault [14]. Blind source separation or independent component analysis opens a new signal processing paradigm applicable to many fields in modern science and engineering. There are significant potential applications of blind separation/deconvolution in various fields, such as wireless telecommunication systems, sonar and radar systems, audio and acoustics, image enhancement and biomedical signal processing (EEG/MEG signals)[2, 4, 5, 6, 10, 12, 14]. The blind source separation/deconvolution problem is to recover independent sources from sensor outputs without assuming any a priori knowledge of original signals besides some statistic features [3, 4, 10].

The state-space description of systems [15] is a new generalized model for blind separation and deconvolution. There are several reasons why the state-space models are advantageous for blind deconvolution. Although transfer function models are equivalent to the state-space ones in the linear case, it is difficult to exploit any

0-7803-6278-0/00$10.00 (C) IEEE

common features that may be present in the real dynamic systems. The main advantage of the state space description for blind deconvolution is that it not only gives the internal description of a system, but there are various equivalent types of state-space realizations for a system, such as balanced realization and observable canonical forms. In particular, it is known how to parameterize some specific classes of models which are of interest in applications. In addition, it is easy to tackle the stability problem of state-space systems using the Kalman Filter. Moreover, the state-space model enables a much more general description than standard finite impulse response (FIR) convolutive filtering. All of the known filtering models, such as AR, MA, ARMA, ARMAX and Gamma filterings, could also be considered as special cases of flexible state-space models.

The state space formulation of blind source separation/deconvolution was discussed by Salam et al [17]-[18], Zhang et al [20]-[21] and Cichocki et al [8]-[9]. An efficient learning algorithm was developed by Zhang and Cichocki [20] to train the output matrices by minimizing the mutual information. In order to compensate for the model bias and reduce the effect of noise, a state estimator approach [21] was also proposed by using the Kalman filter. Cichocki et al extended the state space approach to nonlinear system [8], and an effective two-stage learning algorithm was presented [7] for training the parameters in demixing models.

In this paper we introduce a general framework of state space approach for multichannel blind deconvolution. First, we formulate blind deconvolution in the form of semi-nonlinear systems and discuss the recoverability of this model. The state space model enables us to divide the parameters in demixing models into two types: internal parameters and external parameters. They are trained in different ways. The internal parameters are independent of the individual signal separation problems; they are usually trained in an off-line manner, according to a set of signal separation problems. In contrast, the external parameters are trained individually for each separation problem. The natural gradient algorithm is presented for training the external parameters and stability analysis is also given. The extended Kalman filter is employed to compensate for the model bias and reduce the effect of noise. Finally, a computer simulation is given to show the validity and effectiveness of the state-space approach.

GENERAL FORMULATION

Assume that unknown source signals $\mathbf{s}(t) = (s_1(t), \cdots, s_n(t))^T$ are stationary zero-mean i.i.d processes and mutually statistically independent. Suppose that the unknown source signals $\mathbf{s}(k)$ are mixed by a stable nonlinear dynamic system, and $\mathbf{u}(k) \in \mathbf{R}^n$ is the vector of sensor signals, which are available to signal processing. In this paper, we present a semi-nonlinear dynamic system as a demixing model

$$\mathbf{x}(k+1) = \mathcal{F}(\mathbf{x}(k), \mathbf{u}(k), \Theta) \quad (1)$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \quad (2)$$

where $\mathbf{u}(k) \in \mathbf{R}^n$ is the vector of sensor signals, $\mathbf{x}(k) \in \mathbf{R}^M$ is the state vector of the system, $\mathbf{y}(k) \in \mathbf{R}^n$ is designated to recover source signals in certain sense, $\mathcal{F}_N$

is a nonlinear mapping, described by a general nonlinear capability neural network, Θ is the set of parameters (synaptic weights) of the neural network. The dimension M of the state vector is the order of the demixing system. In this demixing model, the output equation is assumed to be linear. The restriction is reasonable since in many practical problems, the measurement is a linear combination of certain variables. The simplification of the demixing model is simply for clarity of presentation and for easier derivation of learning algorithms. One can extend the results to the general case.

In blind deconvolution, we neither know the nonlinear mappings, nor the order of the mixing system. We need to estimate the order and approximate nonlinear mappings of the demixing system. There are several criteria for estimating the dimension of a system in system identification, such as AIC and FPE criteria. The order estimation problem in blind deconvolution is quite difficult, but interesting. It remains an open problem that is not discussed in this paper. It should be noted that if the dimension of state vector in the demixing model is overestimated, i.e., it is larger than that in the mixing model, the separated signals may contain some auxiliary time-delays, which are acceptable in blind deconvolution. If the mapping, $\mathcal{F}_N$ is linear, the nonlinear state space model will reduce to the standard multichannel blind deconvolution.

Internal Representation

The state space description [15]-[16] allows us to divide the variables into two types: the internal state variable $\mathbf{x}(k)$, which produces the dynamics of the system, and the external variables $\mathbf{u}(k)$ and $\mathbf{y}(k)$, which represent the input and output of the system, respectively. The vector $\mathbf{x}(k)$ is known as the state of the dynamic system, which summarizes all the information about the past behavior of the system that is needed to uniquely predict its future behavior, except for the purely external input $\mathbf{u}(k)$. The term state plays a critical role in mathematical formulation of a dynamical system. It allows us to realize the internal structure of the system and to define the controllability and observability of the system as well. In the state space framework, it becomes much easier to discuss the stability, controllability and observability of dynamical systems.

In the state space formulation, the parameters in the state equation of the demixture are referred to as internal representation parameters (or simply internal parameters), and the parameters in the output equation as external ones. Such a distinction enables us to train the demixing model in two stages: internal representation and output separation. In the internal representation stage, we will make the state space as sparse as possible such that the output signals can be represented as a linear combination of the state vector $\mathbf{x}(k)$ and input vector $\mathbf{u}(k)$. The internal parameters could be trained in off-line way or by using information backpropagation and Kalman filtering. In the second stage, we let the internal parameters fixed and train the external parameters by blind deconvolution algorithms.

Invertiability by State Space Model

Assume that the number of sensor signals equals the number of source signals, i.e. $m = n$. In the following discussion, we restrict the mixing model to the following form,

$$\mathbf{x}(k+1) = \overline{\mathcal{F}}(\mathbf{x}(k), \mathbf{s}(k)) \tag{3}$$

$$\mathbf{u}(k) = \overline{\mathbf{C}}\mathbf{x}(k) + \overline{\mathbf{D}}\mathbf{s}(k) \tag{4}$$

where the state equation is a nonlinear dynamic system, and the output equation is a linear one. From a theoretical point of view, we can easily find the inverse of the state space models in the same form, if the matrix $\overline{\mathbf{D}}$ is invertible. In fact, the inverse system is expressed by

$$\mathbf{x}(k+1) = \overline{\mathcal{F}}(\mathbf{x}(k), \overline{\mathbf{D}}^{-1}(\mathbf{y}(k) - \overline{\mathbf{C}}\mathbf{x}(k))), \tag{5}$$

$$\mathbf{s}(k) = \overline{\mathbf{D}}^{-1}(\mathbf{u}(k) - \overline{\mathbf{C}}\mathbf{x}(k)). \tag{6}$$

This means that if the mixing model is expressed by (3) and (4), we can recover the source signals using the inverse system (5) and (6). There is an advantage to the state space model in that we do not need to inverse any nonlinear functions explicitly.

LEARNING ALGORITHM

Stochastic gradient optimization methods for parameterized systems suffer from slow convergence due to the statistical correlation of the processes signals. While quasi-Newton and related methods can be used to improve convergence, they also suffer from heavy computation and numerical instability, as well as local convergence problems.

The natural gradient search scheme proposed by Amari [1, 2] is an efficient technique for solving iterative estimation problems. For a cost function $l(\mathbf{y}, \mathbf{W})$, the natural gradient $\tilde{\nabla} l(\mathbf{y}, \mathbf{W})$ defines the steepest ascent direction of $l(\mathbf{y}, \mathbf{W})$. In this paper we use the natural gradient approach to develop an efficient learning algorithm for blind deconvolution.

Our objective is to train the demixing model such that the output signals are as spatially mutually independent and temporarily i.i.d. as possible. In this paper, we employ the Kullback-Leibler Divergence as a cost function. The Kullback-Leibler Divergence between the pdf $p_{\mathbf{y}}(\mathbf{y})$ of $\mathbf{y}$ and $q_{\mathbf{y}}(\mathbf{y}) = \prod_{i=1}^{n} p_{y_i}(y_i)$ is given by

$$\mathcal{D}(p, q) = \int p_{\mathbf{y}}(y) \log\left(\frac{p_{\mathbf{y}}(\mathbf{y})}{\prod_{i=1}^{n} p_i(y_i)}\right) d\mathbf{y}, \tag{7}$$

where $p_{y_i}(y_i)$ is the marginal probability density function of y_i. The divergence $l(\mathbf{W})$ is a nonnegative functional, which measures the mutual independence of the output signals $y_i(k)$. The output signals $\mathbf{y}$ are mutually independent if and only if $l(\mathbf{W}) = 0$. In order to implement the statistical on-line learning, we reformulate the cost function as

$$l(\mathbf{y}, \mathbf{W}) = -\log|det(\mathbf{D})| - \sum_{i=1}^{n} \log q(y_i), \tag{8}$$

where $q(y_i)$ is an estimation of the true probability density function $p_i(y_i)$ of source signals.

In this section, we develop a learning algorithm to update the external parameters $\mathbf{W} = [\mathbf{C}, \mathbf{D}]$ in the demixing model. In order to improve learning performance, we define a new search direction, which is related to the natural gradient, developed by Amari [1, 2].

In fact, the relation between the natural gradient and the ordinary gradient can be defined by

$$\tilde{\nabla} l = \nabla l \begin{bmatrix} \mathbf{I} + \mathbf{C}^T\mathbf{C} & \mathbf{C}^T\mathbf{D} \\ \mathbf{D}^T\mathbf{C} & \mathbf{D}^T\mathbf{D} \end{bmatrix}. \tag{9}$$

where $\nabla l = \left[\frac{\partial l(\mathbf{y},\mathbf{W})}{\partial \mathbf{C}} \; \frac{\partial l(\mathbf{y},\mathbf{W})}{\partial \mathbf{D}}\right]$ [19, 22]. It is easy to see that the preconditioning matrix

$$\begin{bmatrix} \mathbf{I} + \mathbf{C}^T\mathbf{C} & \mathbf{C}^T\mathbf{D} \\ \mathbf{D}^T\mathbf{C} & \mathbf{D}^T\mathbf{D} \end{bmatrix}$$

is symmetric positive definite, and this expression is the extension of Amari's natural gradient to the state space model.

Hence, the natural gradient learning algorithm to update matrices $\mathbf{C}$ and $\mathbf{D}$ is described as

$$[\Delta\mathbf{C} \; \Delta\mathbf{D}] = -\eta(k)\tilde{\nabla} l(\mathbf{y}, \mathbf{W}), \tag{10}$$

or we rewrite it into the explicitly form

$$\begin{aligned} \Delta\mathbf{C}(k) &= \eta\left((\mathbf{I} - \varphi(\mathbf{y})\mathbf{y}^T)\mathbf{C} - \varphi(\mathbf{y})\mathbf{x}^T\right), & (11) \\ \Delta\mathbf{D}(k) &= \eta\left(\mathbf{I} - \varphi(\mathbf{y})\mathbf{y}^T\right)\mathbf{D}. & (12) \end{aligned}$$

Stability of Learning Algorithm

In this section we discuss the stability of the natural gradient algorithm (11) and (12). The equilibrium points of the dynamical system satisfy

$$E[\varphi(\mathbf{y}(k))\mathbf{x}^T(k)] = 0 \tag{13}$$

$$E\left[\mathbf{I} - \varphi(\mathbf{y}(k))\mathbf{y}^T(k)\right] = 0 \tag{14}$$

Clearly, the true solution $\mathbf{C}$ and $\mathbf{D}$ is the solution of (13) and (14). However, this does not guarantee that the $\mathbf{C}(k)$ and $\mathbf{D}(k)$ converges to the true solution even locally. This is because if the true solution is an unstable equilibrium point of (11) and (12), the learning sequence $\mathbf{C}(k)$ and $\mathbf{D}(k)$ will never converge to it.

Assume that $\mathbf{y}(k)$ is the recovered signal, which is spatially mutually independent and temporarily i.i.d. We will prove that $\mathbf{y}(k)$ is a locally stable equilibrium point of the learning algorithm (11) and (12) under certain conditions on the source signals.

The main idea to analyze the stability conditions is to consider the variational system of (11) and (12) at the equilibrium point. If all the eigenvalues of the variational matrix have negative real parts, the derived equilibrium point is asymptotically stable. We use the following notation

$$\sigma_i^2 = E[y_i^2],\ \kappa_i = E[\varphi_i'(y_i)],\ m_i = E[y_i^2 \varphi_i'(y_i)] \tag{15}$$

where $\varphi_i' = \frac{d\varphi_i}{dy_i}$. After tedious calculation, we obtain the stability conditions for (11)-(12) as follows

$$m_i + 1 > 0,\ \kappa_i > 0,\ \kappa_i \kappa_j \sigma_i^2 \sigma_j^2 > 1, \tag{16}$$

for $i, j = 1, \cdots, n$ and covariance matrix $E(\mathbf{x}\mathbf{x}^T)$ is positive definite.

If the mixing system is linear, the condition that covariance matrix $E(\mathbf{x}\mathbf{x}^T)$ is positive definite can be further simplified.

STATE ESTIMATOR – THE KALMAN FILTER

There is a drawback in training the parameters in demixing model using information backpropagation algorithm. It may suffer from instability of demixing systems. In order to overcome the problem, we employ the Kalman filter to estimate the state of demixing system. From output equation (2), it is observed that if we can accurately estimate the state vector $\mathbf{x}(k)$, then we can separate mixed signals using the learning algorithm (11) and (12).

The Kalman filter is a powerful approach for estimating the state vector in state-space models. Consider the demixing model with noise terms,

$$\mathbf{x}(k+1) = \mathcal{F}(\mathbf{x}(k), \mathbf{u}(k), \boldsymbol{\Theta}, \mathbf{w}(k)) \tag{17}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) + \mathbf{v}(k) \tag{18}$$

where the random variables $\mathbf{w}(k)$ and $\mathbf{v}(k)$ represent the process and measurement noise. To estimate a state vector of the nonlinear system, we begin with the variational equation

$$\delta\mathbf{x}(k+1) \approx +\mathbf{A}_k \delta\mathbf{x}(k) + \mathbf{W}_k \mathbf{w}(k), \tag{19}$$

where $\mathbf{A}_k$ is the Jacobian matrix of partial derivatives of $\mathcal{F}$ with respect to $\mathbf{x}$,

$$\mathbf{A}_{k,ij} = \frac{\partial \mathcal{F}_i}{\partial x_j}(\mathbf{x}(k), \mathbf{u}(k), \boldsymbol{\Theta}, \mathbf{0}), \tag{20}$$

and $\mathbf{W}_k$ is the Jacobian matrix of partial derivatives of $\mathcal{F}$ with respect to $\mathbf{w}$,

$$\mathbf{W}_{k,ij} = \frac{\partial \mathcal{F}_i}{\partial w_j}(\mathbf{x}(k), \mathbf{u}(k), \boldsymbol{\Theta}, \mathbf{0}). \tag{21}$$

Suppose that the random variables of $\mathbf{w}$ and $\mathbf{v}$ have the following probability density functions

$$p(w(k)) \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k),\ p(v(k)) \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k). \tag{22}$$

where $\mathbf{Q}_k$ and $\mathbf{R}_k$ are covariance matrices of $\mathbf{w}$ and $\mathbf{v}$. Given theses approximations, the Kalman filter equation used to the state $\hat{\mathbf{x}}$ is

$$\hat{\mathbf{x}} = \mathbf{x}(k) + \mathbf{K}_k \mathbf{r}(k), \tag{23}$$

where the matrix $\mathbf{K}_k$ is called the Kalman gain. $\mathbf{r}(k)$ is called the innovation or residual which measures the error between the measured(or expected) output $\mathbf{y}(k)$ and the predicted output $\mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k)$. There are a variety of algorithms with which to update the Kalman filter gain matrix $\mathbf{K}$ as well as the state $\mathbf{x}(k)$; refer to [13] and [11] for more details.

However, in the blind deconvolution problem there exists no explicit residual $\mathbf{r}(k)$ to estimate state vector $\mathbf{x}(k)$ because the expected output $\mathbf{y}(t)$ here means the source signals, and we cannot measure the source signals. In order to solve the problem, we present a new concept called hidden innovation in order to implement the Kalman filter in the blind deconvolution case. Since updating matrices $\mathbf{C}$ and $\mathbf{D}$ will produce an innovation in each learning step, we introduce a hidden innovation as follows

$$\mathbf{r}(k) = \Delta\mathbf{y}(k) = \Delta\mathbf{C}\mathbf{x}(k) + \Delta\mathbf{D}\mathbf{u}(k), \tag{24}$$

where $\Delta\mathbf{C} = \mathbf{C}(k+1) - \mathbf{C}(k)$ and $\Delta\mathbf{D} = \mathbf{D}(k+1) - \mathbf{D}(k)$. The hidden innovation presents the adjusting direction of the output of the demixing system and is used to generate an a posteriori state estimate. Once we define the hidden innovation, we can employ the commonly used Kalman filter to estimate the state vector $\mathbf{x}(k)$, as well as to update the Kalman gain matrix $\mathbf{K}$. The updating rule in this paper is described as follows:

(1) Compute the Kalman gain

$$\mathbf{K}_k = \mathbf{P}_k \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_k \mathbf{C}_k^T + \mathbf{R}_k)^{-1}.$$

(2) Update estimate with hidden innovation

$$\begin{aligned} \mathbf{r}(k) &= \Delta\mathbf{C}\mathbf{x}(k) + \Delta\mathbf{D}\mathbf{u}(k), \\ \hat{\mathbf{x}}(k) &= \mathbf{x}(k) + \mathbf{K}_k \mathbf{r}(k). \end{aligned}$$

(3) Update the error covariance

$$\hat{\mathbf{P}}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)\mathbf{P}_k.$$

(4) evaluate the state vector ahead

$$\mathbf{x}(k+1) = \mathcal{F}(\hat{\mathbf{x}}(k), \mathbf{u}(k), \Theta, \mathbf{0}).$$

(5) evaluate the error covariance ahead

$$\mathbf{P}_{k+1} = \mathbf{A}_k \hat{\mathbf{P}}_k \mathbf{A}_k^T + \mathbf{W}_k \mathbf{Q}_k \mathbf{W}_k^T.$$

where $\mathbf{Q}_k$ and $\mathbf{R}_k$ are the covariance matrices of the noise vector $\mathbf{w}$ and output measurement noise $\mathbf{v}_k$, respectively, and $\mathbf{A}_k$ and $\mathbf{W}_k$ are calculated by (20) and (21).

The theoretic problems such as convergence and stability remain to be analyzed. Simulation experiments show that the algorithm, based on the Kalman filter, can separate the convolved signals very well.

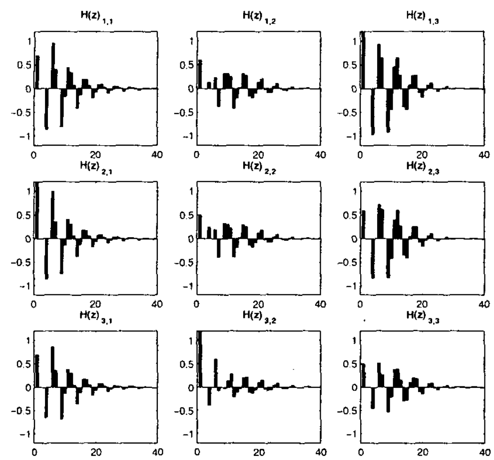

Figure 1: The coefficients of $\mathbf{H}(z)$ of the mixing system

COMPUTER SIMULATION

In this section we present computer simulations to demonstrate the validity and effectiveness of the natural gradient algorithm and the Kalman filter for blind deconvolution. Because the approximation of internal parameters will unavoidably produce a model bias, we employ the Kalman filter to compensate for the model bias and reduce the effect of noise. Several numerical simulations have been performed to demonstrate the performance of the Kalman filter. Here we give only one illustrative example.

Example: The coefficients of the transfer function of a linear system is plotted in Fig. 1. It is assumed to be unknown for the algorithm.

Assume that source signals are i.i.d quadrature amplitude modulated (QAM). The Gaussian noise represented by $\mathbf{v}$ was zero mean with a covariance matrix $0.1\mathbf{I}$. The initial values for matrices $\mathbf{A}$ and $\mathbf{B}$ in the state equation are chosen to be ones in canonical controller form and the initial value for matrix $\mathbf{C}$ is set to a zero matrix or given randomly in the range (-1,1), and $\mathbf{D} = \mathbf{I}_3 \in \mathbf{R}^{3\times 3}$.

We use the natural gradient algorithm (GD) to train the output matrices $\mathbf{C}$ and $\mathbf{D}$, and employ the Kalman filter (KF) to estimate the state vector $\mathbf{x}(k)$ of the system. Figure 2 illustrates the three output signals of demixing model for the time interval $1 \le k \le 300$, $1501 \le k \le 1800$ and $3001 \le k \le 3300$, respectively by using the natural gradient algorithm and the extended Kalman filter.

It is worth noting that the output signals may converge to the characteristic QAM constellation, up to an amplitude and phase rotation factors ambiguities.

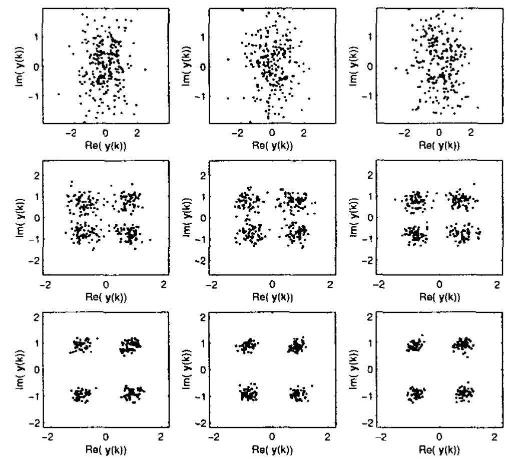

Figure 2: Output constellations

CONCLUSION

In this paper we have presented a general framework of state space approach for multichannel blind deconvolution. The state space model allows us to divide blind deconvolution in two steps: supervised learning for internal parameters and unsupervised learning for external parameters. Adaptive learning algorithms for updating external parameters are developed by minimizing a cost function, which is derived from mutual information of output signals. An state estimator based on the Kalman filter is also presented in order to amend the model bias and reduce the effect of noise. A computer simulation is given to demonstrate the validity and effectiveness of the state-space approach.

REFERENCES

[1] S. Amari, **Differential–geometrical methods in statistics, Lecture Notes in Statistics**, vol. 28, Berlin: Springer, 1985.

[2] S. Amari, "Natural gradient works efficiently in learning," **Neural Computation**, vol. 10, pp. 251–276, 1998.

[3] S. Amari and J.-F. Cardoso, "Blind source separation– Semiparametric statistical approach," **IEEE Trans. Signal Processing**, vol. 45, pp. 2692–2700, Nov. 1997.

[4] A. Bell and T. Sejnowski, "An information maximization approach to blind separation and blind deconvolution," **Neural Computation**, vol. 7, pp. 1129–1159, 1995.

[5] J.-F. Cardoso and B. Laheld, "Equivariant adaptive source separation," **IEEE Trans. Signal Processing**, vol. SP-43, pp. 3017–3029, Dec 1996.

[6] A. Cichocki and R. Unbehauen, "Robust neural networks with on-line learning for blind identification and blind separation of sources," **IEEE Trans Circuits and Systems I : Fundamentals Theory and Applications**, vol. 43, no. 11, pp. 894–906, 1996.

[7] A. Cichocki and L. Zhang, "Two-stage Blind Deconvolution Using State-Space Models

(Invited)," in **Proceedings of the Fifth International Conference on Neural Information Processing(ICONIP'98)**, Kitakyushu, Japan, Oct. 21-23 1998, pp. 729–732.

[8] A. Cichocki, L. Zhang and S. Amari, "Semi-blind and State-Space Approaches to Nonlinear Dynamic Independent Component Analysis," in **Proc. NOLTA'98**, 1998, pp. 291–294.

[9] A. Cichocki, L. Zhang and T. Rutkowski, "Blind separation and filtering using state space models," in **Proc of '99 Int. Symposium on Circuits and Systems, ISCAS'99**, Orlando, Florida, May 30-June 2 1999, pp. V–78–81.

[10] P. Comon, "Independent component analysis: a new concept?" **Signal Processing**, vol. 36, pp. 287–314, 1994.

[11] M. S. Grewal and A. P. Andrews, **Kalman Filtering: Theory and Practice**, Prentice Hall, 1993.

[12] Y. Hua, "Fast maximum likelihood for blind identification of multiple FIR channels," **IEEE Trans. Signal Processing**, vol. 44, pp. 661–672, 1996.

[13] O. L. R. Jacobs, **Introduction to Control Theory**, Oxford University Press, 1993.

[14] C. Jutten and J. Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," **Signal Processing**, vol. 24, pp. 1–10, 1991.

[15] T. Kailath, **Linear Systems**, Englewood Cliffs, NJ: Prentice Hall, 1980.

[16] R. J. Ober, "Balanced Canonical Forms," in S. Bittanti and G. Picci (eds.), **Identification, Adaptation, Learning**, Springer, NATO ASI Series, pp. 120–183, 1996.

[17] F. Salam, "An adaptive network for blind separation of independent signals," in **Proc. '93 IEEE International Symposium on Circuits and Systems (ISCAS)**, 1993, pp. 431–434.

[18] F. Salam and G. Erten, "The State space framework for blind dynamic signal extraction and recovery," in **Proc of '99 Int. Symposium on Circuits and Systems, ISCAS'99**, Orlando, Florida, May 30-June 2 1999, pp. V–66 –69.

[19] L. Zhang, S. Amari and A. Cichocki, "Natural gradient approach to blind separation of over- and undercomplete mixtures," in **Proceeding of Independent Component Analysis and Signal Separation(ICA'99)**, Aussois, France, 1999, pp. 455–460.

[20] L. Zhang and A. Cichocki, "Blind Deconvolution/Equalization Using State-Space Models," in **Proceeding of Int'l IEEE Workshop on Neural Networks for Signal Processing(NNSP'98)**, Cambridge, UK, August 31-September 2 1998, pp. 123–131.

[21] L. Zhang and A. Cichocki, "Blind Separation of Filtered Source Using State-Space Approach," in M. Kearns, S. Solla and D. Cohn (eds.), **Advances in Neural Information Processing Systems**, Cambridge, MA: MIT press, vol. 11, pp. 648–654, 1999.

[22] L. Zhang and A. Cichocki, "Natural gradient approach to blind deconvolution of dynamical systems," in **Proceeding of ICA'2000**, Helsinki, Finland, 2000, pp. 27–32.