Blind Source Separation via Principal Component Analysis

Seungjin Choi

Abstract - Various methods for blind source separation (BSS) are based on independent component analysis (ICA) which can be viewed as a nonlinear extension of principal component analysis (PCA). Most existing ICA methods require certain nonlinear functions (which leads to higher-order statistics) depending on the probability distributions of sources, whereas PCA is a linear learning method based on second-order statistics. In this paper we show that the PCA can be applied to the task of BSS, provided that sources are spatially uncorrelated but temporally correlated. Since the resulting method is based on only second-order statistics, it avoids the nonlinear function and is able to separate mixtures of several colored Gaussian sources, in contrast to the conventional ICA methods.

Keywords - Blind source separation, eigen-decomposition, independent component analysis, principal component analysis.

1. Introduction

Blind source separation (BSS) is a statistical method which aims at recovering unknown sources from their linear instantaneous mixtures without any prior knowledge of the mixing matrix nor sources. It has drawn lots of attractions in signal processing and artificial neural network communities since it is a fundamental problem that is encountered in many practical applications such as blind beamforming [1], multiuser communications [2], [3], speech processing [4], image processing [5], [6] and biomedical signal analysis [7], [8] where multiple sensors are involved.

Most existing methods for BSS are based on ICA, the task of which is to decompose multivariate data into a linear sum of statistically independent components. In Comon's seminar paper [9], it was shown that ICA could perform BSS when sources are statistically independent. ICA-based BSS methods include [9], [10], [11], [12], [13], [14], [15], [16], [17]. These methods exploit the higher-order statistical structure of the data either implicitly or explicitly.

Alternatively the task of BSS can be achieved by the nonlinear PCA [18] where certain nonlinear functions heuristically employed in the conventional PCA. Since these methods are based on higher-order statistics, they are involved with certain nonlinear functions that are related to the knowledge of probability distributions of sources which are unknown in advance. Some methods for selecting an appropriate nonlinear function in the task of BSS were developed [19], [21], [22], [17].

In this paper we show that the standard PCA can be applied to the task of BSS, provided that sources are spatially uncorrelated but temporally correlated. The resulting method is based on only second-order statistics, so it avoids nonlinear function. Moreover, the method is also able to separate the mixtures of several colored Gaussian sources, whereas the ICA method can not.

The rest of the paper is organized as follows. Next section describes the problem formulation of BSS and model assumptions. In Section Ⅲ we explain how the PCA can be applied to the task of BSS and present two methods of BSS based on the standard PCA. Numerical examples are given in Section Ⅳ to confirm the validity of the proposed methods. Finally conclusions are drawn in Section Ⅴ.

2. Problem Formulation

In the simplest form of BSS, it is assumed that an n-dimensional observation vector $\boldsymbol{x}(t) = [x_1(t) \cdots x_n(t)]^T$ is generated by

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t), \tag{1}$$

where $\boldsymbol{A} \in \boldsymbol{R}^{n\times m}$ is called the mixing matrix, $s(t)$ is the n-dimensional vector whose elements are called sources.

The task of BSS is to recover source vector $s(t)$ up to its possibly scaled and re-ordered version. That is, the estimate of source vector, $\hat{s}(t)$ is required to satisfy $\hat{s}(t) = \boldsymbol{P\Lambda s}(t)$ where $\boldsymbol{P}$ is some permutation matrix and $\boldsymbol{\Lambda}$ is some nonsingular diagonal matrix. The transformation by $\boldsymbol{P\Lambda}$ (generalized permutation matrix)

This work was supported in part by Korea Ministry of Science & Technology under Brain Science and Engineering Research Program and by KOSEF 2000-2-20500-009-5.

Manuscript received: Sep. 28, 2000 Accepted: Feb. 15, 2001.

Seungjin Choi is with Department of Computer Science & Engineering, POSTECH, San 31 Hyoja-dong, Nam-gu, Pohang Kyungbuk 790-784, Korea.

is referred to as a transparent transformation. In other words, the task is to build a linear transformation W (demixing matrix) such that $WA = P\Lambda$ is satisfied.

Throughout this paper the following assumptions are made:

AS1 The mixing matrix A is of full rank.

AS2 Sources are spatially uncorrelated stochastic signals with zero mean and unit variance, i.e.,

$$E\{ s_i(t)s_j(t-\tau)\} = \delta_{ij} \quad \forall \tau, \tag{2}$$

where δ_{ij} is the Kronecker delta equal to 1 for $i=j$, otherwise it is zero and $E\{\cdot\}$ denotes the statistical average operator.

AS3 Each source has non-zero temporal correlation, i.e.,

$$E\{s_i(t)s_i(t-\tau)\} \neq 0 \text{ for } \tau \in T \text{ and } i=1,\cdots,n, \tag{3}$$

where T is a set of time-lag for which sources have non-zero temporal correlations.

Remarks

· In general, the ICA requires the assumption of spatial independence among sources. Here we assume only spatial uncorrelatedness between sources.

· Many natural signals such as speeches and images have non-vanishing temporal correlations. Thus it is desirable to use the temporal structure of sources.

3. Proposed Methods

This section describes our method of BSS based on PCA with brief review of PCA and whitening.

A. PCA

The PCA is a classical multivariate data analysis method that is useful in linear feature extraction and data compression. It is essentially equivalent to Karhunen-Loève transformation and closely related to factor analysis. All these methods are based on second-order statistics of the data.

The PCA finds a linear transformation $y = Hx$ such that the retained variance is maximized. It can be also viewed as a linear transformation which minimizes the reconstruction error [23]. The row vectors of H correspond to the normalized orthogonal eigenvectors of the data covariance matrix.

One simple approach to PCA is to use singular value decomposition (SVD). Let us denote the data covariance matrix by $R_x(0) = E\{x(t)\ x^T(t)\}$. Then the SVD of $R_x(0)$ gives

$$R_x(0) = UDU^T, \tag{4}$$

where U is the eigenvector matrix (i.e., modal matrix) and D is the diagonal matrix whose diagonal elements correspond to the eigenvalues of $R_x(0)$. Then the linear transformation H in PCA is given by

$$H = U^{T.} \tag{5}$$

Adaptive algorithms for PCA are also available. These include Oja's subspace rule [24], GHA [25], and APEX [26], to name a few.

b. Whitening

The task of whitening is to find a linear transformation $y = Vx$ such that $E\{ yy^T\} = I$ where $I \in R^{n\times n}$ is the identity matrix. In other words, the whitening aims at eliminating cross-correlations as well as normalizing the variance to be unity. It follows from the decomposition in (4) that the transformation $V = D^{-\frac{1}{2}}U^T = R_x^{-\frac{1}{2}}(0)$ leads to $E\{ yy^T\} = I$.

Alternatively the data vector x can be whitened in adaptive fashion. For example, the global decorrelation rule [27] has the form

$$V(t+1) = V(t) + \eta\{I - y(t)y^T(t)\}\ V(t), \tag{6}$$

where $\eta > 0$ is a learning rate. The recurrent network can be also used for whitening [28].

C. PCA Based BSS

The data vector x is first whitened by a linear transformation $V = D^{-\frac{1}{2}}U^T$. Denote the whitened vector by $z = Vx$. Then the whitened data vector z has the form

$$z(t) = Bs(t), \tag{7}$$

where $B = VA$ is an orthogonal mixing matrix since $E\{ z(t)z^T(t)\} = I$.

Let us consider the signal $y(t)$ that is the sum of $z(t)$ and $z(t-1)$, i.e.,

$$y(t) = z(t) + z(t-1). \tag{8}$$

The correlation matrix of $y(t)$ is

$$\begin{aligned} R_y(0) &= E\{[z(t)+z(t-1)][z(t)+z(t-1)]^T\} \\ &= 2I + R_z(1) + R_z^T(1), \end{aligned} \tag{9}$$

where $R_z(1)$ is the time-delayed correlation matrix defined by

$$R_z(1) = E\{ z(t)z^T(t-1)\}. \tag{10}$$

Since the correlation matrix $R_y(0)$ is symmetric, it has the following eigen-decomposition

$$R_y(0) = U_y D_y U_y^T. \tag{11}$$

Note that from (7) and (8) we have

$$R_y(0) = B\{ 2I + R_s(1) + R_s^T(1)\} B^T, \tag{12}$$

where $2I + R_s(1) + R_s^T(1)$ is a diagonal matrix from the assumptions (AS2) and (AS3). Since B is an orthogonal matrix, Eq. (12) represents the eigen-decomposition of $R_y(0)$. Thus, it follows from (11) and (12) that the orthogonal mixing matrix B is equal to the eigenvector matrix U_y up to a permutation and sign of eigenvectors, provided that the diagonal elements of $2I + R_s(1) + R_s^T(1)$ are distinct. This leads to the estimate of the mixing matrix, $\widehat{A} = V^{-1} U_y$.

A simple PCA analysis of the signal $y(t)$ results in a solution to the task of BSS. The main advantage of this method is to exploit only second-order statistics, so it avoids any nonlinear function which is important in the conventional ICA methods. We can apply any adaptive PCA algorithm [25], [26] for calculating the matrix U_y in on-line fashion. The method is summarized below.

Algorithm Outline: BSSPCA-1

- We whiten the data x by a linear transformation $V = D^{-\frac{1}{2}} U^T$, i.e.,

$$z(t) = V x(t). \tag{13}$$

- We compute the signal $y(t) = z(t) + z(t-1)$.
- Apply the PCA to the signal $y(t)$ to obtain U_y where $R_y(0) = U_y D_y U_y^T$.
- The demixing matrix W is

$$W = U_y^T V. \tag{14}$$

In fact the method BSSPCA-1 can be viewed as the simultaneous diagonalization of two correlation matrices.

We define $v(t) = Wx(t) = U_y^T z(t)$. Then one can easily see that both $R_v(0)$ and $R_v(1) + R_v^T(1)$ are diagonalized by the transformation W. The following theorem explains why the simultaneous diagonalization of these two matrices gives the solution to the problem of BSS.

Theorem 1: Let $\Lambda_1, \Lambda_2, D_1, D_2 \in R^{n\times n}$ be diagonal matrices with nonzero diagonal entries. Suppose that $G \in R^{n\times n}$ satisfies the following decompositions:

$$D_1 = G\Lambda_1 G^T, \tag{15}$$

$$D_2 = G\Lambda_2 G^T. \tag{16}$$

Then the matrix G is the generalized permutation matrix, i.e.,
$G = P\Lambda$ if $D_1^{-1} D_2$ and $\Lambda_1^{-1}\Lambda_2$ have distinct diagonal entries.
See Appendix for proof.

Alternatively we can also find the demixing matrix W that simultaneously diagonalizes two correlation matrices. The method is described below.

Algorithm Outline: BSSPCA-2

- We whiten the data x by a linear transformation $V = D^{-\frac{1}{2}} U^T$, i.e.,

$$z(t) = V x(t). \tag{17}$$

- We compute the time-delayed correlation matrix $M_z(1) = \frac{1}{2}\{ R_z(1) + R_z^T(1)\}$.
- We compute the SVD of $M_z(1)$,

$$M_z(1) = U_z D_z U_z^T. \tag{18}$$

- The demixing matrix W is

$$W = U_z^T V. \tag{19}$$

Remarks

- We define $M_s(\tau) = \frac{1}{2}\{ R_s(\tau) + R_s^T(\tau)\}$. In similar manner we also define $M_x(\tau)$. Note that both $M_s(0)$ and $M_s(1)$ are diagonal matrices. The simultaneous diagonalization of $M_s(0)$ and $M_s(1)$ requires

$$W M_x(0) W^T = \Lambda_0,$$
$$W M_x(1) W^T = \Lambda_1, \tag{20}$$

where Λ_0 and Λ_1 are some nonsingular diagonal matrices. The demixing matrix W corresponds to the generalized eigenvector matrix that satisfies

$$M_x^{-1}(0) M_x(1) W^T = W^T \Lambda_0^{-1} \Lambda_1. \tag{21}$$

It is known that the generalized eigen-decomposition performs the simultaneous diagonalization [29], so the method BSSPCA-2 can also be implemented using the generalized eigen-decomposition. It was implemented using the recurrent network in [30]. The similar idea was also exploited in [31].

• In BSSPCA-1, our analysis holds for any $\tau \in T$. Instead of (8), we can consider

$$y(t) = z(t) + z(t-\tau), \tag{22}$$

as long as $R_s(\tau)$ is invertible diagonal matrix with distinct diagonal elements. Moreover we can also consider a linear sum

$$y(t) = \sum_{i=0}^{L} \alpha_i z(t-i). \tag{23}$$

• In BSSPCA-2, we can consider $M_z(\tau)$ for any $\tau \in T$. We can also consider a linear combination $\sum_{i=1}^{L} \beta_i M_z(i)$.

4. Numerical Experiment

We demonstrate the useful behavior of our methods (BSSPCA-1 and BSSPCA-2) through numerical experiment. We have chosen one speech and one music signals that were sampled at 8 kHz, and three independent colored-Gaussian sources. These five sources constituted the 5-dimensional source vector s. The 5-dimensional observation x were generated via the linear model (1) using the mixing matrix A that was chosen as

$$A = \begin{bmatrix} -2.095 & -0.539 & -0.579 & -0.566 & 0.774 \\ 0.041 & -0.686 & 1.627 & -0.755 & 0.059 \\ -0.971 & 0.132 & 0.959 & 0.103 & 1.079 \\ 0.913 & 0.196 & 0.826 & 0.142 & -0.617 \\ -0.472 & 0.866 & 1.537 & 0.785 & -0.351 \end{bmatrix}.$$

Total number of samples were 10000. The time-lag $\tau = 1$ was chosen in both BSSPCA-1 and BSSPCA-2.

In order to evaluate the performance of methods, we calculated the performance index (PI) defined by

$$PI = \frac{1}{2(n-1)} \sum_{i=1}^{n} \left\{ \left(\sum_{k=1}^{n} \frac{|g_{ik}|}{\max_j |g_{ij}|} - 1 \right) + \left(\sum_{k=1}^{n} \frac{|g_{ki}|}{\max_j |g_{ji}|} - 1 \right) \right\}, \tag{24}$$

where g_{ij} is the (i,j)-element of the global system matrix $G = WA$ and $\max_j g_{ij}$ represents the maximum value among the elements in the ith row vector of G, $\max_j g_{ji}$ does the maximum value among the elements in the ith column vector of G. The performance index defined in (24) tells us how far the global system matrix G is from a generalized permutation matrix (transparent transformation). The zero performance index means perfect separation, however in practice very small number (say, .001) is acceptable performance.

In this experiment, we compared the performance of our methods (BSSPCA-1, BSSPCA-2) to that of popular BSS algorithms such as JADE [1], flexible ICA method [17], extended infomax [22]. Since JADE, flexible ICA, and extended infomax are based on higher-order statistics, one expect that they have difficulty in separating out several Gaussian sources. Hinton diagrams of the global system matrix $G = WA$ are shown in Fig. 1. The extended infomax showed similar performance as the flexible ICA, so it was omitted here. In Hinton diagram, each square's area represents the magnitude of element of the matrix. Dark gray color represents for negative value and light gray color does for positive value. One can see that in BSSPCA-1 and BSSPCA-2, each row and column of the matrix bG has only one dominant non-zero value, in contrast to JADE and flexible ICA. Table I summarizes the value of performance index for each method.

Table 1 Performance comparison of various methods in term of PI

algorithm	performance index
BSSPCA-1	3.22×10^{-4}
BSSPCA-2	3.15×10^{-4}
JADE [1]	0.126
Flexible ICA [32]	0.293
Extended Infomax [21], [22]	0.332

5. Conclusions

In this paper, we have shown that the standard PCA could be applied to the task of BSS, provided that sources are spatially uncorrelated but temporally correlated. We have presented two methods (BSSPCA-1, BSSPCA-2) that are useful for BSS. The adaptive implementation of BSSPCA-1 and BSSPCA-2 is straightforward since any existing adaptive PCA algorithms can be employed. The main advantage of proposed methods is to exploit only second-order statistics, so the methods avoid any nonlinear function, in contrast to the conventional ICA methods. However the methods requires the careful selection of time-lag τ which guarantees that the corresponding time-delayed correlation matrix of source vector has distinct diagonal elements. We are currently working on a way of finding a linear combination (23) that guarantees that the corresponding time-delayed correlation matrix of source vector has distinct diagonal elements. We also working on the extension of the proposed methods to handle the noisy observation vector that was not considered in this paper.

Acknowledgment

This work was supported in part by Korea Ministry of Science & Technology under Brain Science and Engineering Research Program and by KOSEF 2000-2-20500-009-5.

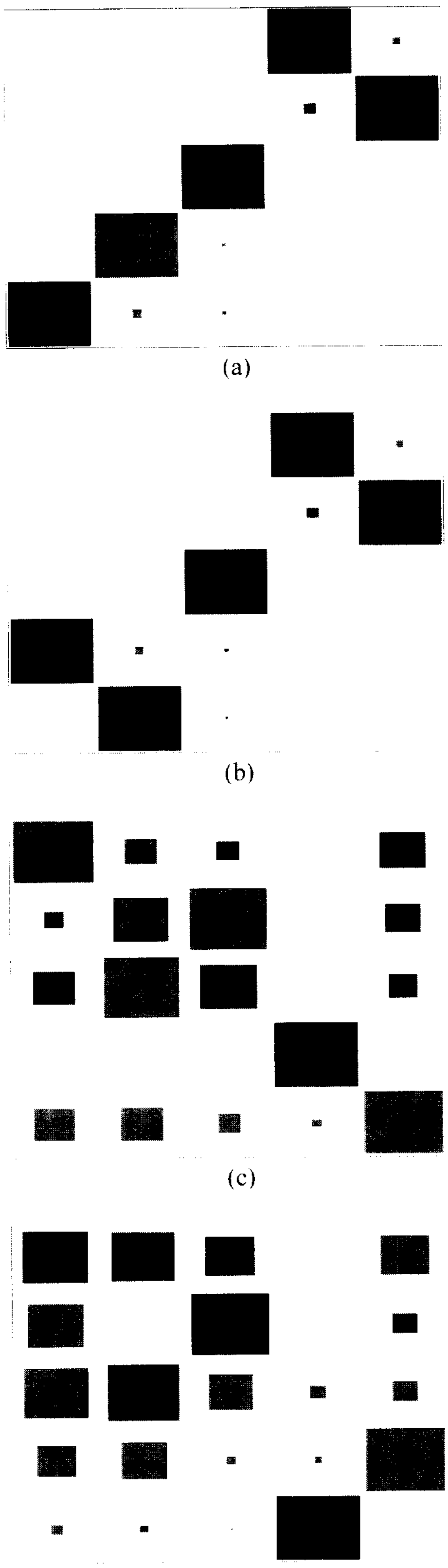

Fig. 1 Hinton diagrams of global system matrices G : (a) BSSPCA-1; (b) BSSPCA-2; (c) JADE; (d) Flexible ICA

Appendix

The proof of Theorem 1 is given here. From (15), there exists an orthogonal matrix $\boldsymbol{Q}$ such that

$$\left(G\Lambda_1^{\frac{1}{2}} \right) = \left(\boldsymbol{D}_1^{\frac{1}{2}} \right)\boldsymbol{Q}. \tag{25}$$

Hence,

$$G = \boldsymbol{D}_1^{\frac{1}{2}} \boldsymbol{Q}\Lambda_1^{-\frac{1}{2}}. \tag{26}$$

Substitute (26) into (16) to obtain

$$\boldsymbol{D}_1^{-1} \boldsymbol{D}_2 = \boldsymbol{Q} \Lambda_1^{-1} \Lambda_2 \boldsymbol{Q}^T. \tag{27}$$

Sine the right-hand side of (27) is the eigen -decomposition of the left-hand side of (27), the diagonal elements of $\boldsymbol{D}_1^{-1} \boldsymbol{D}_2$ and $\Lambda_1^{-1} \Lambda_2$ are the same. From the assumption that the diagonal elements of $\boldsymbol{D}_1^{-1} \boldsymbol{D}_2$ and $\Lambda_1^{-1} \Lambda_2$ are distinct, the orthogonal matrix $\boldsymbol{Q}$ must have the form $\boldsymbol{Q} = \boldsymbol{P\Psi}$, where $\boldsymbol{\Psi}$ is an diagonal matrix whose diagonal elements are either $+1$ or -1. Hence, we have

$$\begin{aligned} G &= \boldsymbol{D}_1^{\frac{1}{2}} \boldsymbol{P\Psi}\Lambda_1^{-\frac{1}{2}} \\ &= \boldsymbol{PP}^T \boldsymbol{D}_1^{\frac{1}{2}} \boldsymbol{P\Psi}\Lambda_1^{-\frac{1}{2}} \\ &= \boldsymbol{P}\Lambda, \end{aligned} \tag{28}$$

where

$$\Lambda = \boldsymbol{P}^T \boldsymbol{D}_1^{\frac{1}{2}} \boldsymbol{P\Psi}\Lambda_1^{-\frac{1}{2}} \qquad Q.E.D. \tag{29}$$

References

[1] J. F. Cardoso and A. Souloumiac, "Blind beam-forming for non-Gaussian signals," *IEE Proceedings-F*, vol. 140, no. 6, pp. 362-370, 1993.
[2] L. Castedo, C. J. Escudero, and A. Dapena, "A blind signal separation method for multiuser communications," *IEEE Trans. Signal Processing*, vol. 45, pp. 1343-1348, 1997.
[3] S. Choi, "Adaptive blind signal separation for multiuser communications: An information- theoretic approach," *Journal of Electrical Engineering and Information Science*, vol. 4, pp. 249-256, Apr. 1999.
[4] S. Choi, H. Hong, H. Glotin, and F. Berthommier, "Multichannel signal separation for cocktail party speech recognition: A dynamic recurrent network," *in Proc. ICSLP*, (Beijing, China), 2000.
[5] A. Bell and T. Sejnowski, "The independent

components of natural scenes are edge filters," *Vision Research*, vol. 37, no. 23, pp. 3327-3338, 1997.

[6] S. Choi and O. Lee, "Sparse independent component analysis: Efficient coding of natural scenes?," *Journal of Electrical Engineering and Information Science*, vol. 5, pp. 84-92, Apr. 2000.

[7] SS. Makeig, A. Bell, T. P. Jung, and T. Sejnowski, "Blind separation of auditory event-related brain responses into independent components," *Proc. of National Academy of Sciences*, vol. 94, pp. 10979-10984, 1997.

[8] S. Choi, A. Cichocki, and S. Amari, "Fetal electro-cardiogram data analysis using independent component analysis," *in Proc. Asia-Pacific Conf. Medical and Biological Engineering* (APCMBE '99), (Seoul, Korea), 1999.

[9] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, no. 3, pp. 287-314, 1994.

[10] A. Cichocki and R. Unbehauen, "Robust neural networks with on-line learning for blind identific-ation and blind separation of sources," *IEEE Trans. Circuits and Systems-I: Fundamental Theory and Applications*, vol. 43, pp. 894-906, 1996.

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

[12] S. Amari and A. Cichocki, "Adaptive blind signal processing neural network approaches," *Proc. of the IEEE, Special Issue on Blind Identification and Estimation*, vol. 86, no. 10, pp. 2026-2048, Oct. 1998.

[13] D. T. Pham, "Blind separation of instantaneous mixtures of sources via an independent component analysis," *IEEE Trans. Signal Processing*, vol. 44, no. 11, pp. 2768-2779, 1996.

[14] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, pp. 1483-1492, 1997.

[15] S. Choi and A. Cichocki, "A linear feedforward neural network with lateral feedback connections for blind source separation," *in Proc. IEEE Work- shop on Higher-order Statistics*, (Ban, Canada), pp. 349-353, 1997.

[16] S. Choi, R. Liu, and A. Cichocki, "A spurious equlibria-free learning algorithm for the blind separation of non-zero skewness signals," *Neural Processing Letters*, vol. 7, pp. 61-68, 1998.

[17] S. Choi, A. Cichocki, and S. Amari, "Flexible independent component analysis," *Journal of VLSI Signal Processing*, vol. 26, pp. 25-38, Aug. 2000.

[18] J. Karhunen, E. Oja, L. Wang, R. Vigario, and J. Joutsensalo, "A class of neural networks for independent component analysis," *IEEE Trans. Neural Networks*, vol. 8, pp. 486-504, 1997.

[19] S. C. Douglas, A. Cichocki, and S. Amari, "Multi-channel blind separation and deconvolution of sources with arbitrary distributions," *in Neural Networks for Signal Processing VII* (J. Principe, L. Gile, N. Morgan, and E. Wilson, eds.), pp. 436-445, 1997.

[20] A. Cichocki, I. Sabala, S. Choi, B. Orsier, and R. Szupiluk, "Self-adaptive independent component analysis for sub-Gaussian and super-Gaussian mixtures with unknown number of source signals," *in Proc. Int. Symp. Nonlinear Theory and Applications*, pp. 731-734, 1997.

[21] M. Girolami, "An alternative perspective on adaptive independent component analysis algori- thms," *Neural Computation*, vol. 10, no. 8, pp. 2103-2114, Nov. 1998.

[22] T. W. Lee, M. Girolami, and T. Sejnowski, "Independent component analysis using an extend-ed infomax algorithm for mixed sub-Gaussian and super-Gaussian sources," *Neural Computation*, vol. 11, no. 2, pp. 609-633, 1999.

[23] K. I. Diamantaras and S. Y. Kung, *Principal Component Neural Networks: Theory and App-lications*. John Wiely & Sons, INC, 1996.

[24] E. Oja, "Neural networks, principal component analysis, and subspaces," *International Journal of Neural Systems*, vol. 1, pp. 61-68, 1989.

[25] T. D. Sanger, "Optimal unsupervised learning in a single-layer linear feedforward neural network," *Neural Networks*, vol. 2, no. 6, pp. 459-473, 1989.

[26] S. Y. Kung and K. I. Diamantaras, "A neural network learning algorithm for adaptive principal component extraction(APEX)," *in Proc. ICASSP*, pp. 861-864, 1990.

[27] S. C. Douglas and A. Cichocki, "Neural networks for blind decorrelation of signals," *IEEE Trans. Signal Processing*, vol. 45, no. 11, pp. 2829-2842, 1997.

[28] S. Choi, S. Amari, and A. Cichocki, "Natural gradient learning for spatio-temporal decorrelation: Recurrent network," *IEICE Trans. Fundamentals*, vol. E83-A, pp. 2715-2722, Dec. 2000.

[29] K. Fukunaga, An *Introduction to Statistical Pattern Recognition*. New York, NY: Academic Press, 1990.

[30] L. Molgedey and H. G. Schuster, "Separation of a mixture of independent signals using time delayed correlations," *Physical Review Letters*, pp. 3634-3637, 1994.

[31] L. Tong, V. C. Soon, Y. F. Huang, and R. Liu, "AMUSE: a new blind identification alogrithm," *in Proc. ISCAS*, pp. 1784-1787, 1990.

[32] A. Belouchrani, K. Abed-Merain, J. F. Cardoso, and E. Moulines, "A blind source separation technique

using second order statistics," *IEEE Trans. Signal Processing*, vol. 45, pp. 434-444, Feb. 1997.

Seungjin Choi was born in Seoul, Korea, in 1964. He received the B.S. and M.S. degrees in electrical engineering from Seoul National University, Korea, in 1987 and 1989, respectively and the Ph.D degree in electrical engineering from the University of Notre Dame, Indiana, in 1996.

He was a Visiting Assistant Professor in the Department of Electrical Engineering at University of Notre Dame, Indiana during the Fall semester of 1996. He was with the Laboratory for Artificial Brain Systems, RIKEN, Japan in 1997 and was an Assistant Professor in the School of Electrical and Electronics Engineering, Chungbuk National University from 1997 to 2000. He is currently Assistant Professor of Computer Science and Engineering at Pohang University of Science and Technology, and is affiliated with the Intelligent Multimedia Lab as well as the Brain Research Center.

He has also been an adjunct senior researcher at Laboratory for Advanced Brain Signal Processing, BSI, Japan since 1998. His primary research interests include probabilistic/statistical learning, statistical (blind) signal processing, independent component analysis, data analysis, pattern recognition, and multiuser communications.

Dr. Choi is a Technical Committee member of IEEE Neural Networks for Signal Processing (NNSP).