ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 243-248

From Source Separation to Independent Component Analysis. An Introduction to the Special Session

Christian Jutten [1]

INPG-TIRF, 46 Av. Félix Viallet, F-38000 Grenoble
and GDR-PRC ISIS
E-mail: chris@tirf.inpg.fr

1 Introduction

First works in source separation, and the concept itself, have been introduced in early 80's by Ans, Hérault and Jutten [10], for modelling the biological problem of motion coding. Although these works had a weak impact in the neural network community, they provided a large interest in the signal processing community during the last 10 years, especially in France and Europa. Since a few years, sessions in signal processing conferences and more recently in neural networks conferences, have been devoted to the problem of source separation: for instance, GRETSI 93 and GRETSI 95 (Juan-Les-Pins, France), NOLTA 95 (Las Vegas, USA), ISCASS 96 (Atlanta, USA), EUSIPCO 96 (Trieste, Italia), NIPS 96 post-workshop (Denver, USA), and this special session.
The objectives of this paper is to give an idea of current researches in this domain, and to introduce the nine papers which have been accepted in the ESANN 97 special session. For a more complete state-of-the-art, one can refer to the invited paper of Karhunen in ESANN'96 [13].
The first section presents the basic models, assumptions and objectives of source separation. The second section gives a short overview of key results and key papers in the case of instantaneous mixtures. The third section recalls main results for convolutive mixtures. The section four addresses briefly the problem of source separation in non-linear mixtures. In section five, the more general concept of independent component analysis is introduced.

2 The problem

2.1 Mixture models

Let us consider the observations at the outputs of a set of n sensors: $x_j(t)$, $j = 1, \dots, n$. These observations are generally unknown superimpositions of p unknown sources $s_i(t)$, $i = 1, \dots, p$, assumed to be statistically independent and non-Gaussian :

$$x_j(t) = \mathcal{F}_j(s_1(t), \dots, s_p(t)),\ j = 1, \dots, n, \tag{1}$$

[1] Ch. Jutten is professor in the Institut des Sciences et Techniques of Université Joseph Fourier of Grenoble.

where the $\mathcal{F}_j$ are unknown functions. The problem of source separation consists in retrieving the sources $s_i(t)$, $i = 1, \ldots, p$, only from the observations $x_j(t)$, $j = 1, \ldots, n$.
The model (1) is too general, and one usually prefer to address simpler mixture models. In the simplest case, we assume that observations are linear instantaneous mixtures of the sources. Denoting the observation vector $\mathbf{x}(t)$ and the source vector $\mathbf{s}(t)$:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{2}$$

where $\mathbf{A}$ is an unknown $n \times p$ matrix, with scalar entries.
If we model the superimpositions by linear filters, the model becomes a convolutive model:

$$\mathbf{x}(t) = [\mathbf{A}(z)]\mathbf{s}(t), \tag{3}$$

where $\mathbf{A}(z)$ is an unknown $n \times p$ polynomial matrix, whose entries are linear filters.

2.2 Assumptions and indeterminacies

The main assumption of the problem is the statistical independence of the sources. However, we can point out that the choice of a particular model, for instance (2) or (3), is also a very strong assumption. In the case of instantaneous mixtures, the separation structure will be a matrix $\mathbf{B}$. In the case of convolutive mixtures, the separation structure will be a polynomial matrix $\mathbf{B}(z)$. Then, source separation methods consist in estimating a matrix $\mathbf{B}$ (or $\mathbf{B}(z)$), such that components of the output vector $\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t)$ (or $\mathbf{y}(t) = [\mathbf{B}(z)]\mathbf{x}(t)$) are statistically independent. Finally, we generally assume that the number of sensors is larger or equal to the source number.

One of the particularities of the source separation problem is made up by solution indeterminacies.
In the case of instantaneous mixtures, one can find the sources up to a scale and up to a permutation. It means that the separating matrix is a matrix satisfying:

$$\mathbf{B}\mathbf{A} = \mathbf{\Delta}\mathbf{P}, \tag{4}$$

where $\mathbf{\Delta}$ and $\mathbf{P}$ are a diagonal matrix and a permutation matrix, respectively.
In the case of convolutive mixtures, the sources are generally obtained up to a filter and up to a permutation. The separating polynomial matrix $\mathbf{B}$ satisfies:

$$[\mathbf{B}(z)][\mathbf{A}(z)] = [\mathbf{\Delta}(z)]\mathbf{P}, \tag{5}$$

where $\mathbf{\Delta}(z)$ is a diagonal polynomial matrix and $\mathbf{P}$ is a permutation matrix.

2.3 Principles of solutions

Methods for source separation, in both convolutive and instantaneous mixtures, are based on the statistical independence of the sources. Two random variables

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 243-248

u and v are independent if and only if:

$$p_{uv}(x, y) = p_u(x) p_v(y). \tag{6}$$

A first way of exploiting the independence definition is to use the Kullback-Leibler divergence:

$$I(p_u, p_v) = \int p_u(x) \log \frac{p_u(x)}{p_v(y)} dx dy. \tag{7}$$

This definition can be easily generalized for a familly of independent random variables. However, it is not very convenient to use, because the source pdf's are unknown. A first approach consists in using a Gram-Charlier or Edgeworth expansion of the source pdf's. A second idea is to write (6) using the second characteristic functions: thus, one can prove that the independence definition involves *all* the cumulants. This second definition is also difficult to use, because it requires an infinity of conditions.
In order to solve the problem of source separation in instantaneous mixtures, Comon introduced in 1991 [5], [6] the notion of contrast function, which is a function $\gamma(p_{\mathbf{y}})$ of the pdf of the random vector $\mathbf{y}$. Simple examples of constrast functions involve fourth-order cumulants [5], [6], [8], [17]. Generalization of contrast functions for convolutive mixtures have been more recently proposed by Comon [7].

3 Instantaneous linear mixtures

It is the simplest case, and has been intensively studied. Theoretical aspects of the initial Hérault, Jutten and Ans algorithm were published, and numerous other solutions based on various parametrizations of the mixtures, or based on new contrast functions, were designed [4], [2], [17], [8]. However, except in the case of two mixtures of two sources, it has not been possible to prove that minimization of contrast function always leads to a separation solution, even if practically spurious solution has never been observed. The strongest result is obtained by Delfosse and Loubaton [8]: all the minima of their contrast function corresponds exactly to a separation solution. Thus, the result of S. Choi and R.W. Liu in the first paper of the session is of high interest: it gives a necessary and sufficient condition, without spurious equilibrium, for separation in the case of non-zero skewness sources.
Another very important result is due to Cardoso and Laheld [15] who computed and compared performances of a few methods, and proposed a familly of equivariant algorithms, i.e. whose performances do not depend on the mixture matrix. These algorithms are particularly efficient for ill-conditionned mixtures.
Finally, many algorithms have been studied for special sources: binary or multi-valued sources, sources with bounded pdf, etc. [3]. This often leads to simpler or to more efficient algorithms. For instance, Pajunen's algorithm, in the second paper of the session, is able to separate more sources than sensors. In the case

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 243-248

where the source number is different of the sensor number, the extraction of sources introduced by Malouche and Macchi is an interesting approach.

4 Convolutive linear mixtures

For narrow band sources, it is easy to check that a convolutive mixture $\mathbf{A}(z)$ reduces to an instantaneous linear mixture $\mathbf{A}$ with complex entries. For wide band sources, the mixture must be modelled by filters. First algorithms for convolutive linear mixtures of wide band sources are due to Nguyen Thi and Jutten [12], [21], Yellin and Weinstein [23], [24], Van Gerven and Van Compernolle [22]. Recently, other approaches based on blind equalization algorithms and involving special parametrization (with Hankel or Sylvester matrices) have been studied [9], [16]. Such an approach is followed by X.-R. Cao and J. Zhu in their paper. For convolutive mixtures, Nguyen Thi and Jutten generalized the Hérault-Jutten algorithm used for instantaneous mixtures. Filter parameters $c_{ij}(k)$, where k is the tap number, are optimized with the following adaptation rule:

$$\Delta c_{ij}(k) = \mu f(y_i(t))g(y_j(t-k)). \tag{8}$$

In the case of instantaneous mixtures, Pham *et al.* [20] proved how to choose optimally the non-linear function f and g. Generalization of this idea for convolutive mixtures is done by Charkany and Deville in their paper. The optimal choice of the non-linear functions is also addressed by L. Xu, C.C. Cheung, J. Ruan and S.I. Amari in the more general framework of independent component analysis (ICA).

5 Non-linear mixtures

Let s_1 and s_2 be two independent random variables, then $f(s_1)$ and $g(s_1)$, where f and g are any inversible functions, are still independent random variables. Then, using only source independence, it seems difficult to separate independent sources in non-linear mixtures without distortion. In their contribution, one of the first for non-linear mixtures with [19], Taleb and Jutten study a special non-linear mixture, called post non-linear mixture, which allows separation without non-linear distortion, and they propose an efficient adaptive algorithm.

6 Independent component analysis

In the case of sources without time correlation, the source separation problem corresponds to a general data analysis method, called independent component analysis (ICA). First introduced and compared to principal component analysis (PCA) by Jutten and Hérault [11], ICA has been formalized by Comon [5], [6] who first emphasized on relations with entropy [1], [18]. Relations with non-linear PCA have been also studied by Karhunen *et al.* [14]. Contributions on

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 243-248

ICA will finish the session. Girolamy and Fyfe will develop relations between ICA and exploratory projection pursuit, while Nadal and Parga will focus on relation between entropy and ICA.

References

[1] A. Bell and T. Sejnowski. An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, 7(6), 1995.

[2] J.-F Cardoso. Source separation using higher order moments. In *Proceeding of ICASSP*, pages 2109–2212, Glasgow, Scotland, May 1989.

[3] J.-F. Cardoso and C. Jutten. Separation of sources: Really blind ? In *NOLTA 95*, Las Vegas, USA, 1995.

[4] P. Comon. Separation of sources using higher-order cumulants. In *SPIE Vol. 1152 Advanced Algorithms and Architectures for Signal Processing IV*, San Diego (CA), USA, August 8-10 1989.

[5] P. Comon. Independent component analysis. In J.-L. Lacoume, M. A. Lagunas, and C. L. Nikias, editors, *International Workshop on High Order Statistics*, pages 111–120, Chamrousse, France, July 1991.

[6] P Comon. Independent component analysis, a new concept ? *Signal Processing*, 36(3):287–314, 1994.

[7] P Comon. Contrasts for multichannel blind deconvolution. *IEEE Signal Processing Letters*, 3(7):209–211, 1996.

[8] N. Delfosse and Ph. Loubaton. Adaptive blind separation of independent sources: A deflation approach. *Signal Processing*, 45:59–83, 1995.

[9] A. Gorokov and Ph. Loubaton. Second-order blind identification of convolutive mixtures with temporally correlated sources: a subspace based approach. In *Proc. European Signal Processing Conf. EUSIPCO 96*, pages 2093–2096, Trieste, Italia, September 1996.

[10] J. Hérault, C. Jutten, and B. Ans. Détection de grandeurs primitives dans un message composite par une architecture de calcul neuromimétique en apprentissage non supervisé. In *Actes du Xeme colloque GRETSI*, pages 1017–1022, Nice, France, 20-24 Mai 1985.

[11] C. Jutten and J. Hérault. Independent components analysis versus principal components analysis. In *Proc. European Signal Processing Conf. EUSIPCO 88*, pages 643–646, Grenoble, France, September 1988.

[12] C. Jutten, L. Nguyen Thi, E. Dijkstra, E. Vittoz, and J. Caelen. Blind separation of sources : an algorithm for separation of convolutive mixtures. In *International Signal Processing Workshop on Higher Order Statistics*, pages 273–276, Chamrousse, France, July 1991.

ESANN'1997 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 16-17-18 April 1997, D-Facto public., ISBN 2-9600049-7-3, pp. 243-248

[13] J. Karhunen. Neural approaches to independent component analysis and source separation. In *ESANN'96, European Symposium on Artificial Neural Networks*, pages 249–266, Bruges, Belgium, April 1996.

[14] J. Karhunen and J. Joutsensalo. Generalizations of principal component analysis, optimization problems, and neural networks. *Neural Networks*, 8(4):549–562, 1995.

[15] B. Laheld and J.-F. Cardoso. Adaptative source separation without prewhitening. In *Proc. European Signal Processing Conf. EUSIPCO 94*, pages 183–186, Edinburgh, Scotland, September 1994.

[16] A. Mansour, C. Jutten, and Ph. Loubaton. Subspace method for blind separation of sources in convolutive mixtures. In *Proc. European Signal Processing Conf. EUSIPCO 96*, pages 2081–2084, Trieste, Italia, September 1996.

[17] E. Moreau and O. Macchi. New self-adaptive algorithms for source separation based on contrast functions. In *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics*, pages 215–219, South Lac Tahoe, USA (CA), June 1993.

[18] J.-P. Nadal and N. Parga. Nonlinear neurons in the low-noisy limit: a factorial code maximizes information transfer. *Network*, 5:565–581, 1994.

[19] P. Pajunen, A. Hyvarinen, and J. Karhunen. Non linear source separation by self-organizing maps. In *ICONIP 96*, Hong-Kong, September 1996.

[20] D. T. Pham, Ph. Garat, and C. Jutten. Separation of a mixture of independent sources through a maximum likelihood approach. In *Proc. European Signal Processing Conf. EUSIPCO 92*, pages 771–774, Brussels, Belgium, August 1992.

[21] H.L. Nguyen Thi and C. Jutten. Blind sources separation for convolutive mixtures. *Signal Processing*, 45:209–229, 1995.

[22] S. Van Gerven and D. Van Compernolle. Feedforward and feedback in a symmetric adaptive noise canceller: Stability analysis in a simplified case. In J. Vandewalle, R. Boite, M. Moonen, and A. Oosterlinck, editors, *Signal Processing VI, Theories and Applications*, pages 1081–1084, Brussels, Belgium, August 1992. Elsevier.

[23] D Yellin and E. Weinstein. Criteria for multichannel signal separation. *IEEE Trans. Signal Processing*, pages 2158–2168, August 1994.

[24] D Yellin and E. Weinstein. Multichannel signal separation: methods and analysis. *IEEE Trans. Signal Processing*, pages 106–118, January 1996.