

A Survey of Signal Separation Algorithms

Otman Chakkor

Dep. Telecommunication and
mechatronics
Abdelmalek Essaadi University

Carlos Garcia Puntonet

Architecture and Computer
Technology
University of Granada

Mohammed Essaadi

ENSIAS
Morocco

ABSTRACT

Blind separation of sources refers to the problem of recovering source signals from their mixtures using only the observed mixtures. The BSS algorithms have developed intensively for linear, nonlinear and convolutive mixtures, the purpose of this paper is to present a description of those algorithms for each situation.

Keywords

Blind signal separation, independent component analysis, maximum likelihood, information theory

1. INTRODUCTION

The Blind Signal separation or blind source separation is the method used to recuperate N unknown signals (sources) from an M observed mixtures through unknown propagation channels. Generally, a method for solving BSS problem is to find a linear transformation of the measured signals such that the resulting source signals are statistically independent from each other as possible. Herault and al [21, 22] have proposed the BSS for the first time as a mathematical adaptive solution for biological problem. At the same time Barnes and al [9] have proposed a similar solution in communication.

Many kinds of approaches have been presented concerning the blind separation of sources, with applications to real problems in areas such as communications, feature extraction, pattern recognition, data visualization, speech processing and biomedical signal analysis (EEG, MEG, fMRI, etc.)

There are some hypotheses that we have to take into account before to proceed to the signal separation:

- 1- The statistical independence between sources.
- 2- We can have at maximum one Gaussian source
- 3- The mixture can be modeled as: linear, convolutive and non-linear.
- 4- We can also consider that we have:
 - a. $N = M$ (easy)
 - b. $N > M$ (undetermined case)
 - c. $N < M$ (over determined case).

One of the important issues in signal processing is to find a set on statistically independent components from the observed data by linear or non-linear transformation.

In the next section, we describe in details the different signal separation algorithms for linear, non-linear and convolutive mixtures:

2. ALGORITHMS

In this section, we will describe the different signal separation algorithms in all possible situations:

2.1 Linear Instantaneous Mixture

Where the function A is considered as a real Matrix:

$$Y = A.S + N \quad (1)$$

Where: A : mixture matrix, S : sources, Y : observed signals and N : Noise

In this section, we will present the principal signal separation algorithms for linear instantaneous mixture:

2.1.1 Second order of statistics (SOS): Methods to discriminate using the correlation matrix of data at different instants, assuming that all sources have different power spectral density. Authors who have used these methods are Fety (diagonalizing matrices covariance), whose idea was improved by Tong giving rise to AMUSE algorithm, Comon, Amari, Pham-Garat, Belouchrani (SOBI), A.Tomé (Matrix Pencil).

2.1.2 High order of statistics (HOS): can be used to define contrast function (or objective function for source separation) which are simple approximations to those derived from Maximum likelihood approach. The cumulants represents the high-order information. Cumulants satisfies several useful properties. Some of them are shared by moments, but other not. For zero-mean random variables a, b, c, d , the second-order cumulants are:

$$Cum[a, b] \triangleq E\{ab\} \text{ Which is identical to second-order moments and fourth-order cumulants are:}$$

$$Cum[a, b, c, d] \triangleq E\{abcd\} - E\{ab\}E\{cd\} - E\{ac\}E\{bd\} - E\{ad\}E\{bc\} \quad (2)$$

The moments and cumulates enjoy the multi-linearity property:

$$Cum[\alpha X, Y, \dots, Z] \triangleq \alpha Cum[X, Y, \dots, Z] \quad (3)$$

and

$$Cum[X_1 + X_2, Y, \dots, Z] \triangleq Cum[X_1, Y, \dots, Z] + Cum[X_2, Y, \dots, Z]. \quad (4)$$

And invariances property by the permutation of indices:

$$\text{Cum}[X_1, X_2, \dots, X_r] = \text{Cum}[X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(r)}] \quad (5)$$

The algorithms based on high-order statistics were used by Herault-Jutten using the moments, generated implicitly by the application of nonlinear odd function to observations (adaptive algorithms). Methods that explicitly use the cumulants were proposed by Cruces, Lacoume, Garat, Mansour, Jutten, Moreau-Macchi, Martín-Clemente, Hyvarinen, Comon or Cardoso. Those methods optimize the contrast functions that depend on the cumulants (module, square, etc. ...) and/or based on matrix decompositions.

2.1.3 Geometrical method: used to calculate the independent components from space observation of the mixtures (for signals with unimodal probability density function. In principle, they work well with few numbers of signals, and for sources with large module kurtosis, i.e. they are away from Gaussianity or zero kurtosis (Prieto, Puntonet).

The basic idea behind the geometric separation method is that in the source space $\{s_1, \dots, s_\beta \subset R^n\}$, where s_i represent a fixed number of samples of the source vector S with zero mean, the clusters along the axes of the coordinate system are transformed by A into clusters along different lines through the origin. In order to estimate the sources, the detection of those n new axes allows us to determine a demixing matrix B which is equivalent to A^{-1} .

Recently, geometric ICA algorithms have received further attention due to their relative ease of implementation [1] [2]. They have been applied successfully to the analysis of real world biomedical data [3] [4] and have been extended to non-linear ICA problems [5] also.

2.1.4 Information theory: Lacoume have proved that the mutual information is a contrast function, Bell and Sejnowski proposed an information maximization method for separation and deconvolution (Infomax), while Pham proposed an algorithm using mutual information (MMI criterion). Laheld and Cardoso introduced an algorithm (EASI), leading to an adaptive rule having the important concept of relative gradient and also obtained by Amari defining "natural gradient".

Minimizing the Mutual Information is the simple way to measure the independence between random variables Y_1, \dots, Y_n is by the mutual information $I(Y_1, \dots, Y_n)$. The main characteristic of this measure is that the variable $Y_i, i=1, \dots, N$ are statistically independent only if $I(Y_1, \dots, Y_n) = 0$

The mutual information can be expressed also using the divergence of Kullback-leibler between $p_Y(y)$ and marginal

$$I(Y_1, \dots, Y_N) = D_{p_Y \parallel \prod_{i=1}^N p_{Y_i}} \quad (6)$$

The principle of the INFOMAX method is to maximize the transfer of the information between the input (source signals) and the output (observed signals) [Linsker 1988]. After that, Bell and Sejnowski [Bell 96] have presented the method of the maximum of the Entropy (ENTMAX) as an application of the principle of the INFOMAX theorem to the signal separation problem. Both methods try to maximize the transfer of the information between the input x and the output y :

$$I(Y; X) = h(Y) - h(Y|X) \quad (7)$$

2.1.5 Maximum likelihood: those methods have been used by Gaeta-Lacoume (with Gram-Charlier developments) by Pham (for iid sources, independent and identically distributed, and non-iid), Garat (QMV QMV-I and-II) and Belouchrani (for iid Gaussian noise, EM algorithm, MLS and SMLS).

The likelihood of source separation model is the probability of a data set as a function of the parameters of the model. The model $x = A \cdot s$ is parameterized by the pair (A, q) , with A is the mixing matrix and q is the density for the source vector.

The density q for the vector x for a given pair (A, q) is given by:

$$p(x; A, q) = |\det(A)|^{-1} \cdot q(A^{-1} \cdot x) \quad (8)$$

Since we have T samples $X_{1:T} = [x(1), \dots, x(T)]$ modeled as independent, then we have: $p(X_{1:T}) = p(X_1) \times \dots \times p(X_T)$ Thus the normalized log likelihood of $X_{1:T}$ for the parameter pair (A, q) is:

$$\begin{aligned} \frac{1}{T} &= \log p(X_{1:T}; A, q) \\ &= \frac{1}{T} \sum_{t=1}^T \log q(A^{-1} \cdot x(t)) - \log |\det(A)| \end{aligned} \quad (9)$$

2.1.6 Non-stationary signals: Have also been developed algorithms for separating non-stationary signals. Matsuoka, for example, was the first to develop a method assuming that the energy ratio of two signals is a non-constant function of time and using the covariances matrix and a stochastic gradient method.

The linear mixture model is assumed in most of the papers devoted to blind signal separation. A more realistic model for mixture should be non-linear:

2.2 Non-linear mixture

Such mixtures can be classified into two groups: 'Pure' non-linear and the post-nonlinear mixtures. For the first type, Burel and Krob have used HOS and adaptive methods, the neural networks algorithms were used by (Pajunen, Yang, Amari Cichocki), self-organizing systems "SOM" or and geometric based methods were used by Mansour, Puntonet, Rojas, Bauer, Lang and Theis. For post-nonlinear, the entropy was used by Taleb and Jutten, radial basis functions (Tan), pattern repulsion by Almeida or ensemble learning by Oja, Valpola.

2.3 Convolutive mixtures

In the literature, we can find the algorithms using second-order statistics (SOS): These procedures use causal FIR filters or subspace based methods (applied when the number of sources is less than number of sensors). Authors have used these techniques are Mansour, Delphosse and Loubaton. In Overall, this is to reduce the problem to a mixture Snapshot statistics of order two for, then use statistics of order four and obtain separation.

The first algorithm using high-order statistics to separate two signals was proposed by Jutten, where the channel was considered as a finite impulse response filter ("FIR") generalizing for convolutive mixtures what was developed for linear mixtures, ie separating the sources by imposing the annulation of cross-moments. Charkani improved method of searching for optimal moment-generating functions.

A convolutive mixture problem can be solved using approximations in frequency domain, it's transformed into a linear problem using the Fourier transform. However we still have the problem of indeterminacies permutation. After the Fourier transform and the use of narrowband filters, we use a post-separation process. This solution was proposed by Capdevielle, Trillion, Charkani, Herault, Castedo and Dapena, the latter using cross four order cumulants

3. INDEPENDENCE

The fundamental assumption in many signal separation algorithms is the statistical independence of the sources.

Theorem 2.1 Let $x = A.s$ be a n -dimension regular mixture of mutually independent random variables, with at most one Gaussian, $y = B.x$ has mutually independent components iff:

$B.A = P.D$, where P and D are permutation and diagonal matrices, respectively.

x_i and y_j are statistical independent if the density function can be factorized :

$$p(x_i, y_j) = p(x_i).p(y_j) \quad (10)$$

Independence must be distinguished from uncorrelatedness, which means that:

$$E\{y_i y_j\} - E\{y_i\}.E\{y_j\} = 0, \text{ for } i \neq j \quad (11)$$

Independence is in general a much stronger requirement than uncorrelatedness. If y_i are statistically independent, then we have:

$$E\{g_1(y_i).g_2(y_j)\} - E\{g_1(y_i)\}.E\{g_2(y_j)\} = 0, \text{ for } i \neq j \quad (12)$$

for any measurable function g_1 and g_2 .

The statistical independence of the signals means that the cross-cumulants of all orders must be zero, usually, but in the practice, we use the fourth order cumulants. The second order of statistics (SOS) is not sufficient to carry out the separation if the sources have similar covariance matrices

The third-order statistics is not valid for signals with symmetric function probability density function

4. SIMULATIONS

We used four signals (speech), mixed using a random mixing matrix)

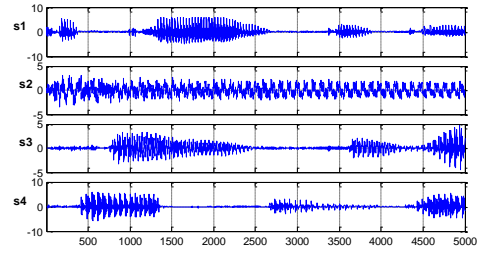


Figure 1 Sources (speech)

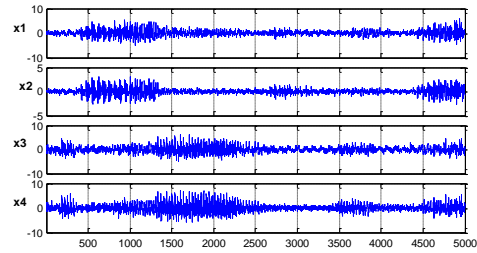


Figure 2 Mixed sources

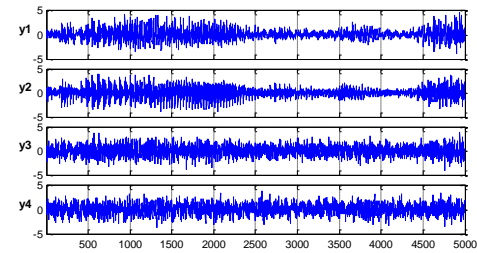


Figure 3 Estimated sources using AMUSE algorithm

The performance index = 0,484680 and the mean (SIR) for S = 2,5447.

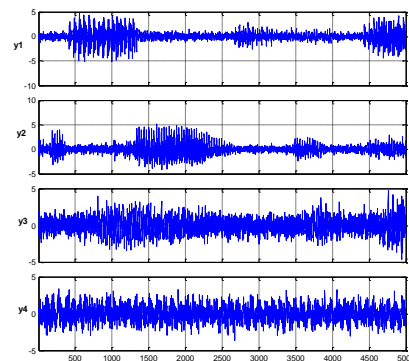


Figure 4 Estimated sources using FastICA (HOS)

The performance index = 0,130668 and the mean (SIR) for S = 5,9289.

5. CONCLUSION

The main objective of this work is to describe the in details different source separation algorithms. We discussed the signal separations methods for different situations, linear, nonlinear, instantaneous, as well as convolutive mixtures. Different interesting aspects of the behavior of the algorithms were presented.

REFERENCES

- [1] C. G. Puntonet and A. Prieto, "An adaptive geometrical procedure for blind separation of sources," *Neural Processing Letters*, vol. 2, 1995.
- [2] C. G. Puntonet and A. Prieto, "Neural net approach for blind separation of sources based on geometric properties," *Neurocomputing*, vol. 18, pp. 141–164, 1998.
- [3] Ch. Bauer, M. Habl, E.W. Lang, C.G. Puntonet, and M.R. Alvarez, "Probabilistic and geometric ICA applied to the separation of EEG signals," M.H.Hamza, ed., *Signal Processing and Communication (Proc.SPC'2000)*, IASTED/ACTA Press, Anaheim, USA, pp. 339 – 346, 2000.
- [4] C.G. Puntonet, Ch. Bauer, E.W. Lang, M.R. Alvarez, and B. Prieto, "Adaptive-geometric methods: application to the separation of EEG signals," P.Pajunen, J.Karhunen, eds., *Independent Component Analysis and Blind Signal Separation (Proc. ICA'2000)*, pp. 273–278, 2000.
- [5] C.G.Puntonet, M.R.Alvarez, A.Prieto, and B.Prieto, "Separation of speech signals for nonlinear mixtures," *Lecture Notes in Computer Science*, vol. 1607, pp. 665–673, 1999.
- [6] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955. (references)
- [7] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [8] I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [9] Y. Barness, J. Carlin and M. Steinberger: "Bootstrapping adaptive cross pol cancellers for satellite communications", *Int. Conf. On*
- [10] *Communication*, Philadelphia, Pennsylvania, 1982.
- [11] R. Nicole, "Title of paper with only first word capitalized," *J. Name Stand. Abbrev.* in press.
- [12] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
- [13] M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989.
- [14] Kevin R. Fall, W. Richard Stevens, *TCP/IP Illustrated, Volume 1: The Protocols*, 2nd ed., Addison-Wesley, USA, 2011.
- [15] Mayank Suhrid, Kiran B Ladhane, Mahendra Singh, Vishwas A Sawant, "Lateral Load Capacity of Rock Socketed Piers Using Finite Difference Approach", *Scientific & Academic Publishing, Journal of Civil Engineering Research*, vol.1, no.1, pp.1-8, 2011.
- [16] Mohamed Almorsy, John Grundy and Amani S. Ibrahim, "Collaboration-Based Cloud Computing Security Management Framework" , in *Proceedings of 2011 IEEE 4th International Conference on Cloud Computing*, pp. 364-371, 2011.
- [17] A. Karnik, "Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP", M. Eng. thesis, Indian Institute of Science, India, 1999.
- [18] J. Padhye, V. Firoiu, D. Towsley, "A stochastic model of TCP Reno congestion avoidance and control", *Univ. of Massachusetts, Tech. Rep. 99-02*, 1999. *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification*, IEEE Std. 802.11, 1997.
- [19] J. Herault, C. Jutten, B. Ans: "Detection de grandeurs primitives dans un message composite par une architecture de calcul neuromimetique en apprentissage non supervise", *X Colloque GRETSI, Nice, France*, pp. 1017-1022, May, 1985.
- [20] C. Jutten and J. Herault: "Independent Component Analysis versus Principal Component Analysis", *EUSIPCO-1988, Signal processing IV*, pp. 643-646, Grenoble, France, 1988.