FREQUENCY-DOMAIN INFOMAX FOR BLIND SEPARATION OF CONVOLUTIVE MIXTURES

Cristina Mejuto, Adriana Dapena, Luis Castedo

Departamento de Electrónica y Sistemas, Universidad de La Coruña
Campus de Elviña s/n, 15.071 La Coruña, SPAIN
Tel: 34 981 167150, Fax: 34 981 167160, e-mail: cris@des.fi.udc.es, adriana@des.fi.udc.es

ABSTRACT

Blind Source Separation (BSS) is a well-known problem that arises in a large number of signal processing applications. In this paper we present a new approach developed in the context of unsupervised learning of Neural Networks (NN) and based on the *Infomax Principle* for the separation of linear mixtures of sources with memory (convolutive mixtures). The problem is solved in the frequency domain, turning the convolution operation in the time domain into a multiplication in the frequency domain. The simulation results show the performance of the proposed algorithm for the separation of convolutive mixtures of white sources.

1. INTRODUCTION

Most of the approaches that have been developed to solve the Blind Source Separation (BSS) problem consider that the input of the separating system is a linear and instantaneous mixture of the sources. Unfortunately, this kind of mixture is seldom found in real world applications and it is more suitable to consider mixing systems that can be modeled with Finite Impulse Response (FIR) filters. In this case the separation can be achieved using criteria based on Higher Order Statistics (HOS) to invert the effect of the FIR filters in the mixture [1].

Most of the solutions to the convolutive problem are time domain extensions of adaptive algorithms for instantaneous mixtures under the hypothesis of statistically independence among the sources [2]. A different alternative is to work in the frequency domain [3, 4]. Applying the Discrete Fourier Transform (DFT) to a data window, i.e., the Short-Time Fourier Transform (STFT), it is possible to convert the general problem of convolutive mixtures into several problems of instantaneous mixtures, one for each frequency. In this way it is possible to apply the same algorithms used for the instantaneous case to each frequency band and, finally, recover the original sources in the time domain using the Inverse Short-Time Fourier Transform (ISTFT). In addition, since for each frequency the sources can be obtained with a different order, a permutation stage is needed before the ISTFT (see figure 1).

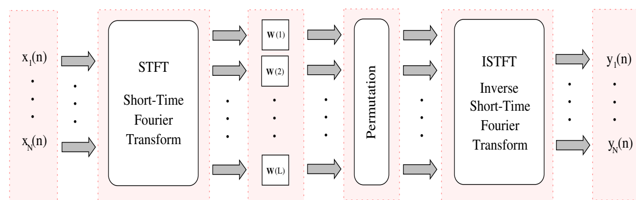

Figure 1: Separation in the frequency domain.

In this paper we present a new approach to solve the convolutive mixtures of white sources based on a previous work for instantaneous mixtures [5] that uses the *Infomax Principle*. Section 2 presents the problem statement. In section 3 an extension of the information transfer maximization in the frequency domain is showed and the learning rule is derived. In section 4 a method is proposed to solve the permutation problem. Section 5 presents some simulation experiments and section 6 contains the conclusions.

2. PROBLEM STATEMENT

Let us consider a vector of observations $\mathbf{x}(n) = [x_1(n), \cdots, x_N(n)]^T$ which is a convolutive mixture of a vector of unknown sources $\mathbf{s}(n) = [s_1(n), \cdots, s_N(n)]^T$, that is,

$$\mathbf{x}(n) = \sum_{k=-\infty}^{\infty} \mathbf{H}(k)\mathbf{s}(n-k) \tag{1}$$

where $\mathbf{H}(n)$ is an unknown $N \times N$ matrix of linear filters representing the mixing system. The element $h_{ij}(n)$ of this matrix denotes the impulse response of the filter over the j source in its way towards the i

sensor. We will assume that the sources are statistically independent, zero-mean, unit-variance and temporally white.

To recover the sources, $\mathbf{x}(n)$ is processed through a linear Multi-Input Multi-Output (MIMO) system which should not only recover the sources but also remove the effect due to the memory in the mixing system. The output $\mathbf{y}(n)$ after a linear filtering of the observations can be expressed as follows

$$\mathbf{y}(n) = \sum_{k=-\infty}^{\infty} \mathbf{W}^H(k)\mathbf{x}(n-k) \quad (2)$$

where $\mathbf{W}(n)$ is a $N \times N$ matrix of linear filters representing the separating system and $y_i(n)$ is a particular output $y_i(n) = \sum_{j=1}^{N} w_{ij}(n) * x_j(n)$, being $*$ the linear convolution operator.

Combining (1) and (2), we can express the output in the z-domain as follows

$$\mathbf{Y}(z) = \mathbf{W}^H(z)\mathbf{X}(z) = \mathbf{W}^H(z)\mathbf{H}(z)\mathbf{S}(z) = \mathbf{G}^H(z)\mathbf{S}(z) \quad (3)$$

The objective to recover the original sources is to select the filter matrix $\mathbf{W}(z)$ in order that each output corresponds to a single and different source. When this occurs, the matrix $\mathbf{G}(z)$ can be expressed as $\mathbf{G}_{(opt)}(z) = \mathbf{\Delta}(z)\mathbf{P}(z)$, where $\mathbf{\Delta}(z) = Diag(\alpha_{11}z^{-\tau_1}, \cdots, \alpha_{NN}z^{-\tau_N})$, being α_{ii} a complex constant, and $\mathbf{P}(z)$ a permutation matrix.

In practice we will assume that the filters in $\mathbf{H}(n)$ can be modeled as FIR causal filter of order M. Then, we can express the vector of observations $\mathbf{x}(n)$ given in (1) as follows

$$\mathbf{x}(n) = \sum_{k=0}^{M-1} \mathbf{H}(k)\mathbf{s}(n-k)$$

In the same way, assuming FIR filters of order F in the separating system, we can expressed the output vector $\mathbf{y}(n) = [y_1(n), \cdots, y_i(n), \cdots, y_N(n)]^T$ as

$$\mathbf{y}(n) = \sum_{k=0}^{F-1} \mathbf{W}^H(k)\mathbf{x}(n-k) \quad (4)$$

3. EXTENSION OF INFOMAX PRINCIPLE TO CONVOLUTIVE MIXTURES

In [5] a learning rule to update the coefficients of a single layer nonlinear Neural Network (NN) has been derived from a unsupervised learning paradigm called *Infomax Principle* [6]. Based on this principle, it is proposed as a criterion for the separation of instantaneous mixtures the maximization of the information transfer between the input and the output of the NN. This is equivalent to the maximization of the output entropy $H(\mathbf{u})$, being $\mathbf{u}$ the vector of the outputs after the nonlinear activation function of the NN, i.e., $u_i = g(y_i),\ i = 1, \cdots, N$. Choosing as nonlinearity the function $g(x) = \int_{-\infty}^{x} exp(-||t|^2 - 1|^2)dt$, it is demonstrated in [5] that the maximization of $H(\mathbf{u})$ is equivalent to the minimization of the following cost function

$$\phi(\mathbf{W}) = \sum_{i=1}^{N} E[||y_i|^2 - 1|^2] - \ln|det\ \mathbf{W}| \quad (5)$$

The previous function admits another interesting interpretation from the perspective of blind adaptive filtering. The first part of (5) is the extension to MIMO systems of the well-known Constant Modulus (CM) criterion [7] widely used in blind equalization. When this criterion is used to solve the BSS problem, it can lead to a situation where the same source is extracted at different outputs. This situation is avoided by the second term because, when it occurs, two columns of $\mathbf{W}$ are linearly dependent and the second part of (5) grows very large. In [8] it is also presented an analysis of the stationary points of (5), which demonstrates its ability to separate instantaneous mixtures of sources with negative kurtosis.

To extend the criterion (5) to the separation of convolutive mixtures in the frequency domain, we will consider a new instantaneous mixing and separating model for each frequency band [8]. Towards this aim, we split each particular observation $x_j(n)$, in R non overlapped segments $x_j^{(r)}(m)$ of length K (see figure 2), where $r = 0, \cdots, R-1$ and $m = 0, \cdots, K-1$, i.e.,

$$x_j(n) = \sum_{r=0}^{\infty} x_j^{(r)}(n - rK)$$

where each segment is obtained from the following expression

$$x_j^{(r)}(m) = \begin{cases} x_j(rK+m) & 0 \le m \le K-1 \\ 0 & rest \end{cases}$$

Taking into account this segmentation, we can consider a similar model to the one described in (4) for each window $x_j^{(r)}(m)$ and to define a vector $\mathbf{y}^{(r)}(m) = [y_1^{(r)}(m), \cdots, y_i^{(r)}(m), \cdots, y_N^{(r)}(m)]^T$ whose components are obtained as follows

$$y_i^{(r)}(m) = \sum_{j=i}^{N} w_{ij}(n) * x_j^{(r)}(m)$$

We define now the L points DFT of $Y_j^{(r)}(k)$ as

$$Y_j^{(r)}(k) = W_{ij}(k)X_j^{(r)}(k)$$

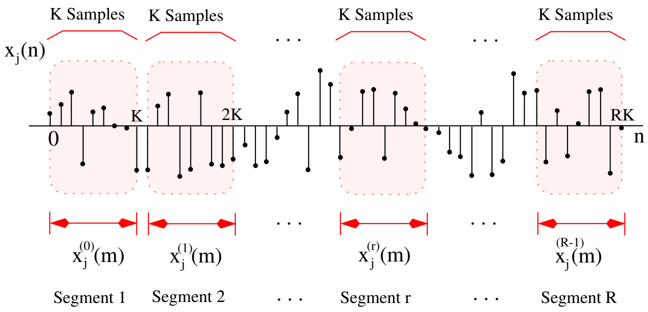

Figure 2: Segmentation of the observations.

where $W_{ij}(k)$ and $X_j^{(r)}(k)$ represent the L points DFTs of the sequences $w_{ij}(m)$ and $x_j^{(r)}(m)$ of length F and K respectively , being $L \geq F + K - 1$

$$W_{ij}(k) = \sum_{m=0}^{L-1} w_{ij}(m) e^{-j\frac{2\pi}{L}km} \quad k = 0, \cdots, L-1$$

$$X_j^{(r)}(k) = \sum_{m=0}^{L-1} x_j^{(r)}(m) e^{-j\frac{2\pi}{L}km} \quad k = 0, \cdots, L-1$$

Computing the IDFT of $Y_j^{(r)}(k)$, we obtain each component of the vector $\mathbf{y}^{(r)}(m)$

$$y_i^{(r)}(m) = \frac{1}{L} \sum_{k=0}^{L-1} Y_j^{(r)}(k) e^{j\frac{2\pi}{L}km}$$

To recover a particular output $y_i(n)$ we add the delayed sequences, $y_i^{(r)}(m)$, taking into account the overlapping of $F-1$ points due to the shift of K points of each input segment

$$y_i(n) = \sum_{r=0}^{\infty} y_i^{(r)}(n - rK) \quad (6)$$

This reconstruction procedure is called *overlap-add* [6].

Taking into account the notation derived above, the mixing model for each frequency band can be expressed now as follows

$$\mathbf{X}^{(r)}(k) = \mathbf{H}(k)\mathbf{S}^{(r)}(k)$$

where $\mathbf{S}^{(r)}(k) = [S_1^{(r)}(k), \cdots, S_N^{(r)}(k)]^T$, $\mathbf{X}^{(r)}(k) = [X_1^{(r)}(k), \cdots, X_N^{(r)}(k)]^T$ and $\mathbf{H}(k)$ is a $N \times N$ matrix of the form

$$\mathbf{H}(k) = \begin{bmatrix} H_{11}(k) & \cdots & H_{1j}(k) & \cdots & H_{1N}(k) \\ \vdots & & \vdots & & \vdots \\ H_{i1}(k) & \cdots & H_{ij}(k) & \cdots & H_{iN}(k) \\ \vdots & & \vdots & & \vdots \\ H_{N1}(k) & \cdots & H_{Nj}(k) & \cdots & H_{NN}(k) \end{bmatrix}$$

In the same way, the separating system can be defined as

$$\mathbf{Y}^{(r)}(k) = \mathbf{W}^H(k)\mathbf{X}^{(r)}(k)$$

where $\mathbf{W}(k)$ is a $N \times N$ matrix and $\mathbf{Y}^{(r)}(k) = [Y_1^{(r)}(k), \cdots, Y_N^{(r)}(k)]^T$.

Extending the previous notation to the proposed criterion in (5) and taking into account the model showed in figure 3, we obtain that the output entropy of the NN is

$$H(\mathbf{U}^{(r)}(k)) = ln\ |det(\mathbf{W}^H(k))| - \sum_{i=1}^{N} E[||Y_i^{(r)}(k)|^2 - 1|^2]$$

where $\mathbf{U}^{(r)}(k) = [U_1^{(r)}(k), \cdots, U_N^{(r)}(k)]^T$ and $U_i^{(r)}(k) = g(Y_i^{(r)}(k))$ being $g(.)$ the non linear activation function of the NN defined above.

The maximization of the output entropy, $H(\mathbf{U}^{(r)}(k))$, is equivalent to the minimization of the following cost function for each frequency k

$$\min_{\mathbf{W}(k)} \phi(\mathbf{W}(k)) \stackrel{\text{def}}{=} \sum_{i=1}^{N} E[||Y_i^{(r)}(k)|^2 - 1|^2] - ln\ |det(\mathbf{W}^H(k))| \quad (7)$$

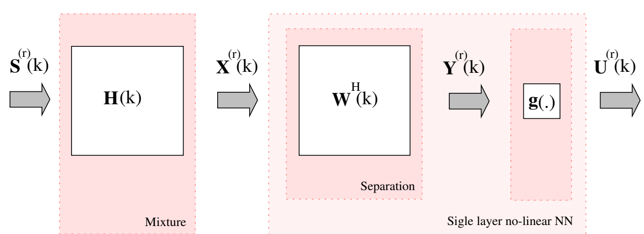

Figure 3: Mixing model and separating MIMO system for each frequency k.

The optimization problem (7) can be solved with a simple stochastic gradient adaptive algorithm, that allows us to recursively update the coefficients of the matrices $\mathbf{W}(k)$ in each frequency

$$\mathbf{W}_{n+1}(k) = \mathbf{W}_n(k) + \mu \nabla_{\mathbf{W}(k)} \phi(n)$$

where $\mathbf{W}_n(k)$ is the value of the matrix $\mathbf{W}(k)$ at time n, μ is a positive constant and $\nabla_{\mathbf{W}(k)}\phi(n)$ is the gradient of the function with respect to the matrix $\mathbf{W}(k)$ at time n. A similar procedure presented in [5] leads us to the following learning rule to minimize the cost function (7)

$$\begin{aligned} \mathbf{W}_{n+1}(k) = \mathbf{W}_n(k) &+ \mu\ (4\ E[\mathbf{X}_n^{(r)}(k)(\mathbf{Y}_n^{(r)}(k))^H] \\ &- 4\ E[\mathbf{X}_n^{(r)}(k)(\mathbf{Y}_n^{(r)}(k))^H \mathbf{D}(\mathbf{Y}_n^{(r)}(k))] \\ &+ [(\mathbf{W}_n(k))^H]^{-1}) \end{aligned} \quad (8)$$

where $(.)^H$ denotes the hermitic operator and $\mathbf{D}(\mathbf{Y}^{(r)}(k)) = Diag(|Y_1^{(r)}(k)|^2, \cdots, |Y_N^{(r)}(k)|^2)$.

The cost function (7) can also be minimized using a relative stochastic gradient algorithm [12, 13] as

$$\begin{aligned}
\mathbf{W}_{n+1}(k) &= \mathbf{W}_n(k) + \mu\ \mathbf{W}_n(k)(\mathbf{W}_n(k))^H \nabla_{\mathbf{W}(k)}\phi(n) \\
&= \mathbf{W}_n(k) + \mu\ \mathbf{W}_n(k)(4\ E[\mathbf{Y}_n^{(r)}(k)(\mathbf{Y}_n^{(r)}(k))^H] \\
&- 4\ E[\mathbf{Y}_n^{(r)}(k)(\mathbf{Y}_n^{(r)}(k))^H \mathbf{D}(\mathbf{Y}_n^{(r)}(k))] + \mathbf{I}) \quad (9)
\end{aligned}$$

where $\mathbf{I}$ denotes the identity matrix.

4. PERMUTATION PROBLEM

BSS algorithms for instantaneous mixtures of sources do not take into account the order in which the original sources are recovered. In this case the separation is achieved when $\mathbf{G}_{(opt)} = \mathbf{\Delta P}$ where $\mathbf{\Delta}$ is a diagonal matrix and $\mathbf{P}$ is a permutation matrix. As a consequence, when the BSS of convolutive mixtures is solved in the frequency domain the separation is achieved when $\mathbf{G}_{(opt)}(z) = \mathbf{\Delta}(z)\mathbf{P}(z)$ where $\mathbf{\Delta}(z)$ is a diagonal matrix and $\mathbf{P}(z)$ is a permutation matrix.

Since at each frequency we are applying an algorithm for instantaneous mixtures, it is possible that the permutation matrix $\mathbf{P}(z)$ be different for each one of these frequencies. When this occurs, we have a permutation problem that can lead us to a wrong reconstruction of the spectrum of the recovered sources at the output of the separating system.

Although our algorithm seems to converge to the same permutation matrix for all the frequencies, as it will be shown in the simulations, this is an important problem to take into account and for whose solution different approaches have been proposed [4, 10, 11]. A first approach consists of considering the statistically relationship among each one of the frequencies of the estimated sources. The simplest relation would be to determine the correlation between two frequency bands. However, the correlation between two bands of adjacent frequencies of the same source depends on the form of the spectral density [4]. This is because it is not always easy to define a threshold which allows us to decide if a particular frequency band belongs to one or another source. Moreover, when the sources are white, as in our case, the correlation between two frequency bands of the same source would be zero. We propose a different alternative, which consists of determining the fourth-order cross-cumulant of the outputs in all the frequencies taking into account the different possible combinations, i.e.,

$$\begin{aligned}
Cum(Y_m(i)), Y_l(j)) &= E[|Y_m(i)|^2 |Y_l(j)|^2] \quad (10) \\
&- E[|Y_m(i)|^2]\ E[|Y_l(j)|^2] \\
&- |E[Y_m(i)Y_l^*(j)]|^2 - |E[Y_m(i)Y_l(j)]|^2
\end{aligned}$$

where $m, l = 1, \cdots, N$ and $i, j = 1, \cdots, L$, being $i < j$. Since the sources are statistically independent, the cross-cumulants (10) will be non-zero when the outputs $Y_m(i)$ and $Y_l(j)$ extract the same source. In contrast, expression (10) will be zero when different sources are extracted.

5. SIMULATIONS

In this section we present the results of a computer simulation carried out to illustrate the behavior of the learning rule proposed in (9) with a step size $\mu = 3 \times 10^{-2}$. We have considered an environment with two white and statistically independent sources, a 4-QAM and a 16-QAM. As the mixing system we have considered FIR filters of order $M = 4$ obtained by truncating to four samples the transfer matrix

$$\mathbf{H}(z) = \begin{bmatrix} -0.8\frac{0.1+z^{-1}}{1+0.1z^{-1}} & 0.5\frac{0.2+z^{-1}}{1+0.2z^{-1}} \\ 0.5\frac{0.5+z^{-1}}{1+0.5z^{-1}} & -0.8\frac{0.1+z^{-1}}{1+0.1z^{-1}} \end{bmatrix}$$

The result of this mixture in the time domain is showed in the left column of figure 4. We have considered segments of length $K = 2$ and the expectations have been estimated with 512 samples (block approximation [8]). In the separating system we have used FIR filters of order $F = 11$. The right column of figure 4 shows the outputs obtained after applying the adaptive learning rule in each one of the $L = 12$ frequencies after 250 iterations. It is clearly seen that our approach is able to successfully recover the original sources.

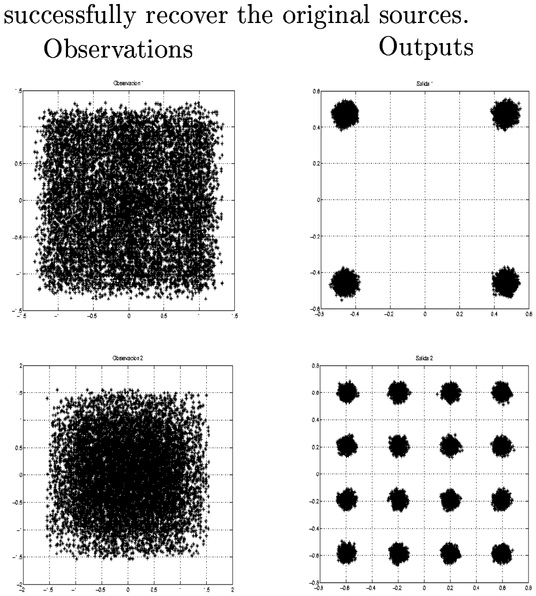

Figure 4: Separation from the convolutive mixture of two sources 16-QAM y 4-QAM.

In order to measure the performance of our algorithm at each frequency band we used the following performance index described in [9], which is zero when the matrix $\mathbf{G}(z) = \mathbf{W}^H(z)\mathbf{H}(z)$ corresponds to the separation of the sources, i.e., when $\mathbf{G}_{(opt)}(z) = \mathbf{\Delta}(z)\mathbf{P}(z)$,

$$\begin{aligned}\rho(\mathbf{G}(z)) &= \sum_{i=1}^{N}\left(\sum_{j=1}^{N}\frac{|g_{ij}(z)|^2}{\max_l(|g_{il}(z)|^2)} - 1\right) \\ &+ \sum_{j=1}^{N}\left(\sum_{i=1}^{N}\frac{|g_{ij}(z)|^2}{\max_l(|g_{lj}(z)|^2)} - 1\right)\end{aligned}$$

Figure 5 illustrates the evolution of this index for each one of the $L = 12$ frequencies. We can see that the algorithm has separated the sources in all the frequencies. Table 1 shows the cross-cumulants computed from (10) among the outputs at different frequencies. We can see that the fourth-order cross-cumulants $Cum(Y_1(i), Y_1(j))$ and $Cum(Y_2(i), Y_2(j))$ for all the frequencies are non-zero. In contrast, the fourth-order cross-cumulants $Cum(Y_1(i), Y_2(j))$ take values close to zero. This means that the same source has been extracted at the same output for all the frequencies.

In addition, it is interesting to compare the results in Table 1 with the theoretic values obtained from the stability analysis presented in [8]. Extending the results derived in [8] to the frequency domain, and assuming that the $i-th$ source is extracted at the $i-th$ output for all the frequencies, we obtain that at each frequency k

$$\mathbf{G}^2_{opt}(k) = Diag\left(\frac{1+\sqrt{1+c_1(k)}}{2c_1(k)}, \frac{1+\sqrt{1+c_2(k)}}{2c_2(k)}\right)$$

where $c_i(k) = E[|S_i(k)|^4]$ is the fourth-order moment of the $i-th$ source at frequency k. In our experiments, we have observed that the fourth-order moments of the sources in the frequency domain are $c_1(k) = 1.4881\ \forall k$ and $c_2(k) = 1.6664\ \forall k$, and that the kurtosis are $k_1(i) = -0.512\ \forall i$ and $k_2(i) = -0.338\ \forall i$, $i = 1, \cdots, L$. Then for these particular sources we obtain that the matrix $\mathbf{G}^2_{opt}(k) = Diag(0.866, 0.790)$. Now, using the properties of the cumulants, we obtain that the theoretic values for the fourth-order cross-cumulant of the outputs are $Cum(Y_1(i), Y_1(j)) = g^2_{11}(i)g^2_{11}(j)k_1(i) = -0.3840$ and $Cum(Y_2(i), Y_2(j)) = g^2_{22}(i)g^2_{22}(j)k_2(i) = -0.2084$. Note that these values are close to the experimental results shown in Table 1.

6. CONCLUSIONS

In this paper we address the blind source separation of convolutive mixtures of sources. This problem can be solved in the time or in the frequency domain. When the second option is chosen, the convolutive problem can be treated as L problems of instantaneous mixtures

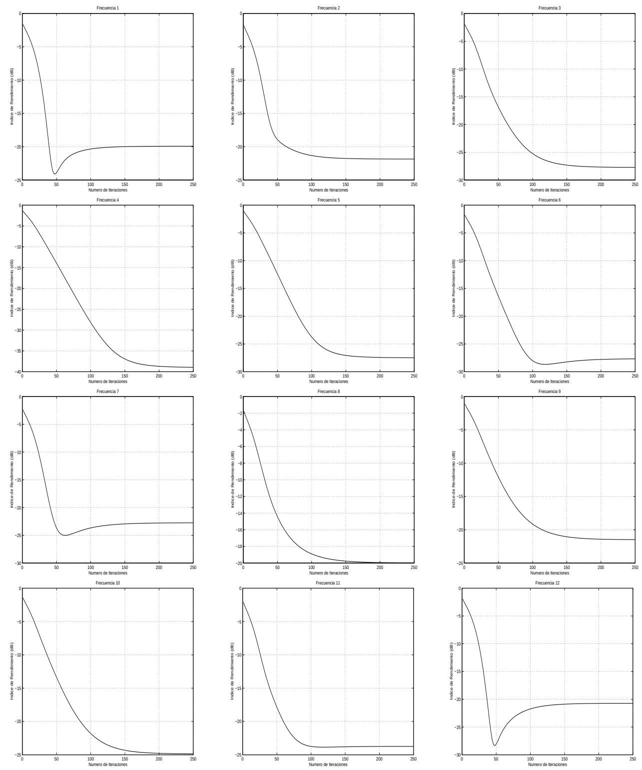

Figure 5: Performance index for each frequency.

$Cum(Y_1(i), Y_1(j))$

	$Y_1(1)$	$Y_1(2)$	$Y_1(3)$	$Y_1(4)$
$Y_1(1)$	-	-0.3670	-0.3720	-0.3735
$Y_1(2)$	-	-	-0.3751	-0.3764
$Y_1(3)$	-	-	-	-0.3813
$Y_1(4)$	-	-	-	-

$Cum(Y_1(i), Y_2(j))$

	$Y_2(1)$	$Y_2(2)$	$Y_2(3)$	$Y_2(4)$
$Y_1(1)$	-0.0016	-0.0024	-0.0019	-0.0018
$Y_1(2)$	-	-0.0063	-0.0072	-0.0075
$Y_1(3)$	-	-	-0.0105	-0.0106
$Y_1(4)$	-	-	-	-0.0106

$Cum(Y_2(i), Y_2(j))$

	$Y_2(1)$	$Y_2(2)$	$Y_2(3)$	$Y_2(4)$
$Y_2(1)$	-	-0.2970	-0.1962	-0.1974
$Y_2(2)$	-	-	-0.1979	-0.1988
$Y_2(3)$	-	-	-	-0.2000
$Y_2(4)$	-	-	-	-

Table 1: Fourth-order cross-cumulant values.

for each one of the frequency bands obtained after applying a DFT to a data segment of the observations. In this way, any algorithm for the BSS of complex instantaneous mixtures of sources can be applied to each one of the L frequencies.

As separation criterion for each frequency, we have proposed an extension of the *Infomax Principle* used in the unsupervised learning of NN [5]. The optimization problem has been solved with a simple stochastic adaptive gradient algorithm in two versions, conventional and natural or relative. The original sources in the time domain can be recovered after the reconstruction of their spectrum and an inversion process of the DFTs of each segment. The simulation results show the behavior of the relative algorithm with a block approximation for the BSS of convolutive mixtures of white sources.

7. ACKNOWLEDGMENTS

The work was supported by program FEDER (1FD97-0082).

8. REFERENCES

[1] D. Yellin and E. Weinstein, "Criteria for Multichannel Signal Separation", in *IEEE Trans. Signal Process*, vol. 42 no. 8, pp. 2158-2168, August, 1994.

[2] T. Nguyen and C. Jutten, "Blind Source Separation for Convolutive Mixtures", in *Signal Processing*, vol. 45 no. 23, pp. 209-229, August, 1995 .

[3] T. Lee and A. Bell and R. Lambert, "Blind Separation of Multiple Speakers in a Multipath Environment", in *Proceedings of NIPS97*, 1997.

[4] V. Capdevielle and C. Serviere and J.L. Lacoume, "Blind Separation of Wide-band Sources in the Frequency Domain", in *Proceedings of ICASSP95*, pp. 2080-2083, 1995.

[5] C. Mejuto and L. Castedo, "A Neural Network Approach to Blind Source Separation", in *Proceedings of NNSP97*, pp. 486-495, Florida, USA, September, 1997.

[6] S. Haykin, "Neural Networks, A Comprehensive Foundation", Macmillan College Publishing Company, Inc., 1996.

[7] D. N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems", *IEEE Trans. on Communications*, vol. COM-28, pp. 1867-1875, November 1980.

[8] C. Mejuto, "Nuevas Aproximaciones al Problema de Separación Ciega de Fuentes", PhD Thesis, Universidad de La Coruña, SPAIN, September, 1999.

[9] Z. Malouche and O. Macchi, "Adaptive Unsupervised Extraction of One Component of a Linear Mixture with a Single Neuron", *submitted to IEEE Trans. Neural Networks*, 1996.

[10] L. Parra and C. Spence, "Convolutive Blind Source Separation based on Multiple Decorrelation", in *Proceedings IEEE Workshop Neural Networks for Signal Processing NNSP98*, Cambridge, UK, September, 1998.

[11] H-C Wu and J.C. Principe "Simultaneous Diagonalization in the Frequency Domain (SDIF) for Source Separation", in *Proceedings ICA99*, pp. 245-250, Aussois, FRANCE, January, 1999.

[12] S. Amari, A. Cichocki and H. Yang, "A New Learning Algorithm for Blind Source Separation", *Advances in Neural Information Processing Systems 8* , 1996.

[13] J. F. Cardoso and B. H. Laheld, "Equivariant Adaptive Source Separation", *IEEE Trans.Signal Processing* , vol. SP-44, pp. 3017-3030, Dec. 1996.