BLIND WIDEBAND SOURCE SEPARATION*

Montse Nájar, Miguel A. Lagunas, Ignasi Bonet

Department of Signal Theory and Communications
Universitat Politècnica de Catalunya
08071 BARCELONA, SPAIN

ABSTRACT

This paper deals with the general problem of separating two independent wideband sources when they are mixed by unknown filters. In order to solve this problem, a backward framework is proposed which is composed of two different stages. The first one consists of two linear predictors devoted to improve the source separation, whitening the input signals. Their coefficients are calculated applying the LMS algorithm, minimizing the mean squared errors between the predicted signals and the output of the separation network. The second stage is formed by decoupling filters that have to be blindly estimated imposing an independence criterion to the outputs.

1. INTRODUCTION

The problem of source separation is addressed in this paper. The attention paid to this subject has recently increased because of its numerous possible applications. For instance in array processing, when several sources from different directions impinge on an aperture of sensors simultaneously. Another example is the problem of enhancing a desired speech signal in the presence of noise or an interference, or separating several speakers using multiple microphones.

Following the pioneer task in the source separation problem by Common, Herault and Jutten [1] and [2], and Bar-Ness [3], among others, some schemes have been developed for multitone tracking [4] and DOA estimation [5] and [6]. In all these architectures the sources are considered narrowband signals, therefore, the unknown coupling filters can be modeled as simple constant gains. However, when wideband signals are involved in the mixture, this approach can not be applied. Previous solutions to the separation of wideband sources have been proposed by Bar-Ness [7], using Bootstrapped algorithms, by Jutten [8], approximating statistical independence applying non-linear functions, and by Yellin [9], employing the cross-bispectra of the reconstructed signals.

This paper describes a new framework composed of two different stages. In the first one, the parameters of two linear predictors have to be estimated. In the second one, two linear decoupling filters should be adjusted imposing statistical independence to the outputs. Both stages interact each other in order to improve the learning. It is important to remark the close relationship of this system with the Estimate-Maximize (EM) algorithm refered in [10], which is applied to the noise cancellation problem. The main advantage of our system in front of the EM algorithm is the suppression of the spectral parameter estimation of the signal which is needed in that algorithm.

First of all we will formulate the problem describing the mixture model and finding the theoretical solutions, next, an adaptive solution is proposed to solve this problem: finally, some simulations results and conclusions will be shown.

2. MIXTURE MODEL

The mixture of two signals can be characterized by two coupling filters, usually considered as MA (Moving Average) models. Furthermore, it can be assumed that the signals are given as a linear combination of their past values and some input taken as white noise. In other words, it is possible to model the signals, before the mixture, as the output of a white noise through an AR (autoregressive) filter. Consequently, the general model considered in this paper is represented in Figure 1 where the input signals (s_1, s_2) are independent white noise sources, and (e_1, e_2) are the resulting signals after the mixture of the corresponding impulse responses H_{ij}. H_{11} and H_{22} are the direct transfer functions considered as AR models, while H_{12} and H_{21}, the coupling transfer functions, are taken as ARMA models therefore, they are the results of the convolution between the direct AR filters and the mixture MA filters.
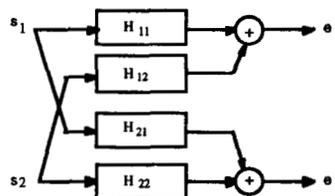

Figure 1. Mixture Model

The frequency domain mixing signals can be described by the matrix formulation:

$$\begin{bmatrix} E_1(f) \\ E_2(f) \end{bmatrix} = \begin{bmatrix} H_{11}(f) & H_{12}(f) \\ H_{21}(f) & H_{22}(f) \end{bmatrix} \begin{bmatrix} S_1(f) \\ S_2(f) \end{bmatrix} \qquad (1)$$

3. THEORETICAL SOURCE SEPARATION

The goal of source separation is to recover the original sources from the observation of the data vector $\underline{e}$. There exist two basic architectures to solve this problem: the Feed-Forward and the Feed-Backward.

The Feed-Forward scheme is derived from the inverse of the mixture matrix:

$$\begin{bmatrix} S_1(f) \\ S_2(f) \end{bmatrix} = \frac{1}{\Delta(f)} \begin{bmatrix} H_{22}(f) & -H_{12}(f) \\ -H_{21}(f) & H_{11}(f) \end{bmatrix} \begin{bmatrix} E_1(f) \\ E_2(f) \end{bmatrix} \qquad (2)$$

This structure depicted in Figure 2 can be easily implemented for wideband applications since no loops are involved on it, but it shows an important drawback: the difficult learning or estimation of the transfer functions caused by the determinant $\Delta(f)$. The formula dependence on the determinant makes quite complex any attempt to compute the gradient with respect to any coefficient of the parametric representation, assuming the model of the mentioned transfer functions.
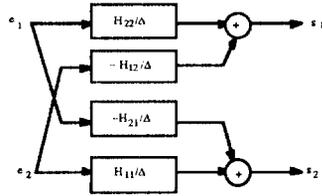

Figure 2. The Feed-Forward separation model

The learning problems which show up in the Feed-Forward architecture can be avoided using the Feed-Backward architecture. In order to find the theoretical solution of the source separation problem by the Feed-Backward model, we assume that the architecture proposed in [8] can be applied:

$$\begin{bmatrix} X_1(f) \\ X_2(f) \end{bmatrix} = \begin{bmatrix} E_1(f) \\ E_2(f) \end{bmatrix} \cdot \begin{bmatrix} 0 & G_{12}(f) \\ G_{21}(f) & 0 \end{bmatrix} \begin{bmatrix} X_1(f) \\ X_2(f) \end{bmatrix} \qquad (3)$$
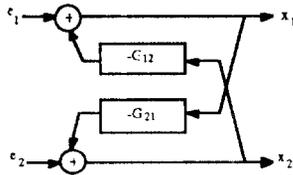

Figure 3. Previous Feed-Backward separation model

Substituting equation (1) in (3), and imposing that each of the obtained outputs x_1 and x_2 must be proportional to only one of the original sources s_1 and s_2, it is possible to obtain the required relations between the transfer functions G_{ij} and H_{ij}:

$$G_{12}(f) = \frac{H_{21}(f)}{H_{11}(f)} \qquad G_{21}(f) = \frac{H_{12}(f)}{H_{22}(f)} \qquad (4)$$

It is important to remark that these transfer functions correspond to MA models from the mixture model assumption.

The resulting signals at the output of these filters are:

$$X_1(f) = H_{11}(f)\, S_1(f) \qquad X_2(f) = H_{22}(f)\, S_2(f) \qquad (5)$$

So, we can conclude that the scheme represented in Figure 3 is not sufficient for an accurate recovery of the sources. It is necessary to add a second stage to our system, that consist of dividing the outputs (x_1, x_2) by the direct filters $(H_{11}\ H_{22})$ respectively. Thus, the whole architecture is obtained:
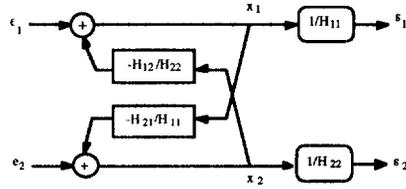

Figure 4. The Feed-Backward separation model

Likewise, another Feed-Backward architecture can be obtained changing the order of these stages, which means to divide the inputs (e_1, e_2) by the respective direct filters, and afterwards, apply the corresponding decoupling filters

4. ADAPTIVE SOLUTION

In most of the interesting cases the coupling systems are unknown, for this reason it is necessary to implement a blind adaptive solution to estimate the desired signals. Two architectures have been developed, corresponding to the two possible Feed-Backward solutions treated in the previous section, the only difference between them is the order of the two stages that conform the solution. In spite of obtaining similar results with the two schemes, the behavior of the system is slightly better placing the stage devoted to cancel the effect of the direct filters (H_{11}, H_{22}) in front of the decoupling stage. Therefore, we will focus on this case, which implementation is depicted in the Figure 5.

The transfer functions that appear in both stages of the separation network can be approximated by FIR filters. The first one can be designed with linear predictors applying the LMS adaptive algorithm, minimizing the mean squared error between the predicted signals and the outputs of the second

stage (s_1, s_2). Since the goal of the system is to separate independent signals, the second stage is estimated imposing an independence criterion on the outputs (s_1, s_2):

$$\phi = |E\{s_1 s_2\}|^2|_{min} \qquad (6)$$

The learning rule is obtained from the derivative of this expression with respect to each filter coefficient. Taking the instantaneous gradient, the coefficient vectors at instant n are adapted as:

$$\underline{h}_1(n) = \underline{h}_1(n) + \mu \ s_2(n)^2 s_1(n) \ \underline{s}_2 \qquad (7)$$

$$\underline{h}_2(n) = \underline{h}_2(n) + \mu \ s_1(n)^2 s_2(n) \ \underline{s}_1 \qquad (8)$$

where $s_i = [s_i(n-1), s_i(n-2), ...s_i(n-Q)]$, and Q is the number of coefficients of the i-filter, and μ is normalized by the power of the two signals.

It is important to notice the higher order character of the correction term. This kind of learning has been used previously in [4] and [6].

Although the original signals are considered white noises the proposed scheme is extensible to any kind of signals. Let us suppose that the desired signals are the responses of the AR filters to independent white noise. First we will recover the noise signals, because the behavior of the separation stage is better for signals with flat spectra. In order to achieve that, the input signals are whitened in the first stage of the system and, simultaneously, the inverse of the AR filters (predictors) are estimated. So, the desired sources can be recovered, provided that the estimated AR filters will be placed behind the separation network.
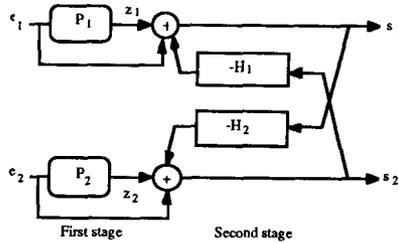

First stage   Second stage

Figure 5. System implementation

5. SIMULATION RESULTS

The first simulation we present is a mixture of two white noise signals where the direct AR filters were of second order and , of course, identical to the AR parts of the coupling filters (given the model), while the MA parts of these filters were of order 6. The imposed order for the estimated predictors was 10, and for the decoupling FIR filters was 6. Anyway, the results obtained in other simulations for different orders were very similar. The parameter μ used for the filter coefficients estimation was 0.001 for the two stages. In Figure 6 the mean squared errors between the estimated signals and the original ones are shown up to 5000 samples (the coefficient adaptation is done sample by sample).

This result can be compared with the one depicted in Figure 7, these graphics correspond to the mean squared errors obtained for the same mixture model, when the separation is made without the first stage, that is, without prewhitening the mixing signals. Certainly, the first stage added to the separation network increases considerably the convergence of the system.

Another interesting simulation result is shown in the two last figures. In this case the original sources were an speech signal and white noise. The mixture was done with two MA filters (order 6), the AR filters were not included because they have only sense as signal generators from noise. In Figure 8 the original speech signal which corresponds to the sentence 'Queso tierno', and the noisy signal after the mixture with a SNR of 0 dB are represented. In Figure 9 the estimated speech signals are depicted (the estimated noise signals lack of interest). The first one was obtained applying the two stages architecture with predictors and decoupling filters of order six, and inverting the estimated predictors at the output in order to recover the original signal after the separation. The second one was computed only with the decoupling stage. The parameter μ for the coefficients adaptation was in the two cases 0.001.

Again, the convergence of the signal separation is fast prewhitening the input signals. Nevertheless, it can be observed from the figures that in the last samples the estimation without predictors is more accurate than with them. This problem may be solved improving the prediction algorithm and consequently the recovering of the original signal after the separation.

Some simulations with two speech signals have been done and anew, the system with the predictors stage provides better performance. The really significant effect can be heard but not seen in the pictures. It is for this reason and for the lack of space that the results of these simulations have not been included in this paper.
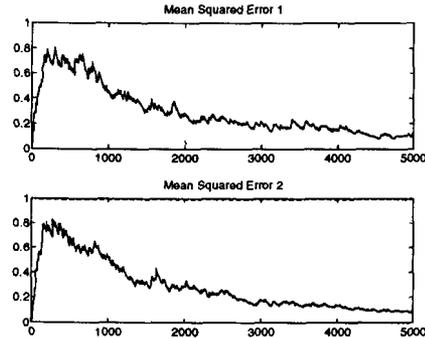

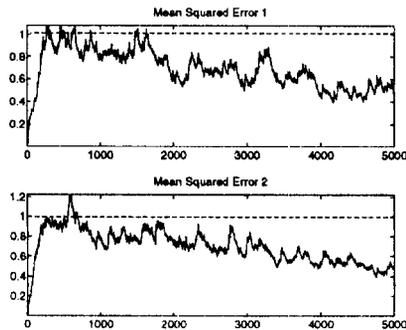*Figure 6. Mean Squared Errors obtained with the two stages network*

Mean Squared Error 1

Mean Squared Error 2

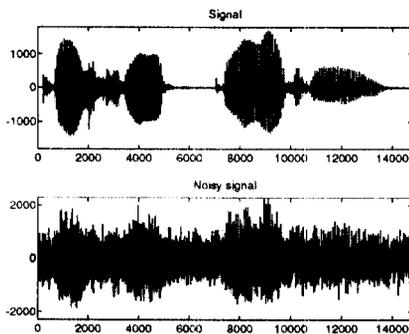Figure 7. Mean Squared Errors obtained without predictors

Signal

Noisy signal

Figure 8. Speech signal and noisy signal

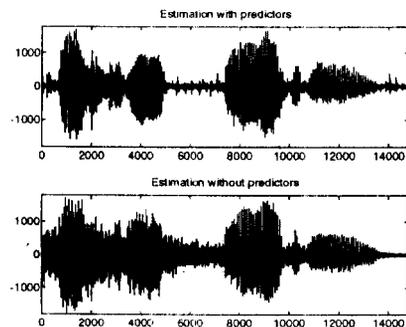Estimation with predictors

Estimation without predictors

Figure 9. Speech estimations

6. CONCLUSIONS

A new adaptive architecture composed of two different stages has been proposed to solve the problem of wideband source separation. In the first stage the mixing signals are whitened by two linear predictors. In the second one the separation in the strict sense is done imposing an independence criterion to the outputs. The simulation results prove that the separation network has a better convergence adding the first stage.

A possible improvement to the method could be achieved by applying better algorithms for the estimation of the predictors. For instance, using the lattice architecture, in such a way that the whitening of the signals will be larger and the recovering of the sources after the separation by the predictors inversion will be also better.

REFERENCES

[1]   Jutten C. ,"Calcul neuromimetique et traitement du signal: Analyse en composants independents", Ph. Tesis, in french, INPG-USMG Grenoble, 1987.

[2]   Common P., Jutten C., Herault J., "Blind separation of sources Part II: Problem statement", Signal Processing, vol 24, n° 1, pp. 11-21, July 1991.

[3]   Bar-Ness Y. ,Carlin J. W. , Steinberger M. L. , "Bootstrapping Adaptive Interference Cancelers: Some Practical Limitations", Globcom 1982, pp F3.7.1-F3.7.5.

[4]   Lagunas M. A. ,Pagés A. , "Multitone tracking with coupled EKFs and high order learning". Proc. ICASSP-92, pp. V153-V156, San Francisco, USA.

[5]   Nájar M. , Lagunas M. A. , Pérez-Neira A. , "Source Separation based on coupled single DOA Estimation Processors", Proc. ICASSP-93, pp. IV336-IV339, Minneapolis, April 1993.

[6]   Nájar M, Lagunas M. A. , "Simultaneous DOA Estimation based on Kolmogorov's Theorem", Proc. IEEE Signal Processing Workshop on Higher-Order Statistics, pp.280-284, June 1993.

[7]   Bar-Ness Y. , "Broadband interference cancellation using a Bootstrapped approach", Proc. ICASSP 93, Minneapolis, April 1993.

[8]   Jutten C. ,Nguyen Thi L. , Dijkstra E. , Vittoz E. , Caelen J. , "Blind Separation of Sources: an Algorithm for Separation of Convolutive Mixtures", Proc. Signal Processing Workshop on High Order Statistics, pp 275-278, Chanrrouse, Francia, July 1991.

[9]   Yellin D. ,Weinstein E. ,"Multi-Channel Signal Separation based on Cross-Bispectra", IEEE Signal Processing Workshop on Higher-Order Statistics, pp.270-274, June 1993.

[10]   Feder M. ,Oppenheim A. , "Methods for Noise Cancellation based on the EM Algorithm", Proc. ICASSP-87, pp. 6.11.1- 6.11.4.

[11]   Oliva X, Lacoume, "Blind Separation of sources", Report Athos, Grenoble, April 1993.