A FLEXIBLE ICA APPROACH TO A NOVEL BSS CONVOLUTIVE NONLINEAR PROBLEM: PRELIMINARY RESULTS

Daniele Vigliano, Raffaele Parisi and Aurelio Uncini
Dipartimento INFOCOM, Università di Roma "La Sapienza" – Italy Via Eudossiana, 18, 00184 Roma – Italy, daniele.vigliano@poste.it; parisi@infocom.uniroma1.it; aurel@ieee.org;

Abstract. This paper introduces a Flexible ICA approach to a novel blind sources separation problem. The proposed on line algorithm performs the separation after the convolutive mixing of post nonlinear convolutive mixtures. The Flexibility of the algorithm is given by the on line estimation of the score function performed by Spline Neurons. Experimental results are described to show the effectiveness of the proposed technique.

Key Words: Blind Source Separation, Flexible ICA, Spline Adaptive function, Mutual Information.

1. Introduction

The first studies about Independent Component Analysis aimed at resolving the famous Cocktail party problem first in static, then in reverberant environments. A critical issue is that linear mixing models are too unrealistic and "poor" in a lot of real situations. The approach to nonlinear convolutive problems are not too widely diffused until now.

Important theoretical results in nonlinear static ICA are in [Hyvarinen et al., 1999]. Several papers considers Post Nonlinear Mixing problem (PNL) in static [Taleb, 2002] and in convolutive [Milani et al., 2002][Zade et al., 2002] environment but only few of them (see [Taleb et al., 1999][Hyvarinen et al., 1999]) explore the existence and uniqueness of the solution. Recent advances in BSS of nonlinear mixing models have been reviewed in [Jutten et al., 2003]. A growing interest is also in the so called Flexible ICA since it improves the quality of separation introducing a better pdf matching and allows a faster learning.

Actually recent studies try to improve the severity of mixing models moving from single block nonlinear structures (convolutive or at least static) to multi block structures. In [Solazzi et al., 2004] sources are recovered from a PNL mixing followed by an instantaneous mixing; in [Vigliano et al., 2004][Vigliano et al., 2004] the mixing environment is composed by a PNL mixing block followed by a convolutive one. This paper explores the solution of the BSS problem in a novel, more severe, convolutive nonlinear mixing environment: the convolutive mixing follows a PNL convolutive mixing block.

2. The nonlinear issue

This section introduces BSS problem in nonlinear environment. Considering an N vector of independent sources $\mathbf{s}[n]$ and a vector of signals $\mathbf{x}[n]$ received by a N-sensor array. The general formulation of the hidden mixing model is:

$$\mathbf{x}[n] = \mathcal{F}\{\mathbf{s}[n],...,\mathbf{s}[n-L]\} \tag{1}$$

in which $\mathcal{F}\{\cdot\}$ is a dynamic nonlinear distorting function. The solution of the BSS problem can be expressed as: $\mathbf{y}[n] = \mathcal{H}\{\mathbf{s}(n)\} = \mathcal{G} \circ \mathcal{F}\{\mathbf{s}(n)\}$. In instantaneous environments ICA recovers the original sources up to some trivial acceptable non-uniqueness: outputs can be scaled and delayed version of flipped inputs. Convolutive mixing environments add a stronger non-uniqueness: the filtering indeterminacies. Convolutive mixtures are separable but applying channel-by-channel filters to the independent recovered signals, outputs are still independent.

This indeterminacy may be unacceptable since it can strongly distort the sources. In any case after separation it is possible to equalize the outputs producing acceptable results. According to these reasons filtering indeterminacy will no more considered in the rest of this paper.

In the more general convolutive nonlinear case (1), the issue of separating mixture with the only constraint of independent output signals and no other a priori assumption is affected by a strong non uniqueness [Jutten et al., 2003]. Several well known examples show that some maps, given independent inputs, produce independent outputs even with non diagonal Jacobian matrix. Independence constraint alone is not strong enough to recover original sources from generic nonlinear mixing environments [Taleb, 2002].

The main issue for generic nonlinear problems is to ensure the presence of conditions (in term of sources, mixing environment, recovering structure) granting at least theoretically the possibility to achieve the desired solution. In [Hyvarinen et al., 1999] authors proposed a constructive way (a Gram-Schmidt like method) to obtain solutions of the separation problem in a static nonlinear mixing environment; in order to grant the uniqueness of the solutions some constraints have been applied to the mixing environment.

The idea introduced is general: adding some “soft” constraint to the problem (like a priori “trivial” assumptions) can produce the uniqueness of the solution. In this paper the a priori knowledge of the mixing model is exploited to design the recovery network: the so called “mirror” demixing model is used.

3. The mixing-demixing structure

This section explores the recovery of separated sources from nonlinear convolutive mixing; the a priori knowledge of the mixing model has been used to design the recovering network. The mixing environment modelled in this paper is represented in figure 1. In which $\mathbf{A}[k]$ and $\mathbf{B}[k]$ are $N \times N$ FIR matrices with respectively L_a and L_b filter taps and $\mathbf{F}\left[\mathbf{p}[n]\right] = \left[f_1\left[p_1[n]\right], \; f_N\left[p_N[n]\right]\right]^T$ is the $N \times 1$ vector of nonlinear distorting functions.

The closed form for mixing model is: $\mathbf{x}[n] = \mathcal{F}[\mathbf{s}] = \mathbf{B}[n] * \mathbf{F}\left[\mathbf{A}[n] * \mathbf{s}[n]\right]$; it enlarges the set of mixing environments from which it is possible to recover separated signals. According to the uniqueness requirements expressed in the previous section the recovering structure mirrors the mixing model. The closed form for recovered outputs is:

$$\mathbf{y}[n] = \mathcal{G}[\mathbf{x}] = \mathbf{Z}[n] * \mathbf{G}\left[\mathbf{W}[n] * \mathbf{x}[n]\right] = \sum_{h=0}^{K_Z - 1} \mathbf{Z}[h]\mathbf{G}\left[\sum_{k=0}^{K_W - 1} \mathbf{W}[k]\mathbf{x}[n-k-h]\right] \quad (2)$$

In which $\mathbf{G}[.]$ is the $N \times 1$ vector of nonlinear compensating functions, one for each channel; $\mathbf{W}[k]$ and $\mathbf{Z}[k]$ are $N \times N$ FIR matrices with K_w and K_z filter taps.

Introducing the knowledge about the particular kind of mixing model is the key to avoid the strict non uniqueness of the solution; such assumption limits the weakness of the output independence condition reducing the cardinality of all possible independent output solutions; with this constraint the problem of recovery the original sources is not ill posed any more.

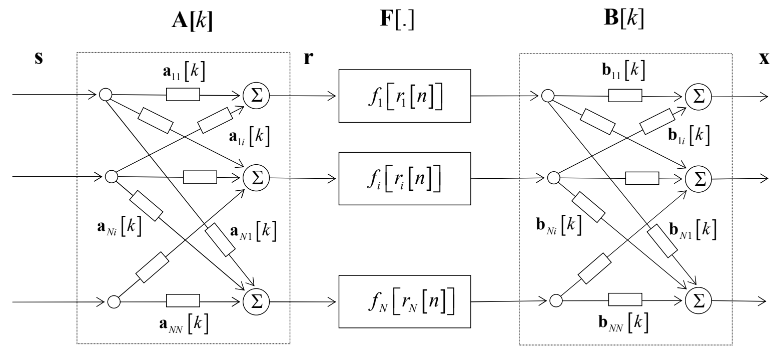


Figure 1. The Block diagram of the convolutive nonlinear mixing model

The use of FIR filter blocks grants the stability of the whole demixing structure.

4. The blind demixing algorithm and the network model

This section explores the blind demixing algorithm, the adaptive network and the network used to compensate the nonlinear distortion. The blind algorithm performs an on-line adaptive learning of the network parameters $\mathbf{\Phi}$ on the base of the output independence estimation. The learning is realized minimizing the Mutual Information $I\{\mathbf{\Phi},\mathbf{y}\}$ between outputs, with a steepest descent algorithm: $\mathbf{\Phi}(k+1)=\mathbf{\Phi}(k)-\eta_{\mathbf{\Phi}}\left[\partial I\{\mathbf{\Phi},\mathbf{y}\}/\partial\mathbf{\Phi}\right]$. The choice of a gradient based minimization procedure lead to terms like:

$$\frac{\partial}{\partial\mathbf{\Phi}}\log\left[p_{y_i}(y_i)\right]=\frac{\partial p_{y_i}(y_i)/\partial y_i}{p_{y_i}(y_i)}\frac{\partial y_i}{\partial\mathbf{\Phi}}=\psi_i(y_i)\frac{\partial y_i}{\partial\mathbf{\Phi}} \tag{3}$$

in which $\psi_i(y_i)$ are the so called Score Functions (SF). In this paper, the Spline Neurons are used to perform the on-line estimation of both Score Functions and nonlinear compensating functions (for a detail about Spline Neurons see [Solazzi et al., 2004][Uncini et al., 2004]). The most attractive property of Spline Neurons, as function estimator, is local learning: for each learning step only the four control points nearest to the training input are considered; no matter how many control points the Spline curve has.

The direct estimation of SF has been performed MSE approach ([Taleb, 2002] for details) but learning rules result still blind:

$$\frac{\partial \varepsilon}{\partial \mathbf{Q}_i^{\psi_j}} = \left[\frac{1}{4} \mathbf{T}_u \mathbf{M} \mathbf{T}_u \mathbf{M} \mathbf{Q}_i^{\psi_j} + \frac{1}{\Delta} \dot{\mathbf{T}}_u \mathbf{M} \right] \tag{4}$$

in which **M** is a matrix of coefficients, **T** is the vector local abscissa and Δ is the distance between the abscissas of adjacent control points.

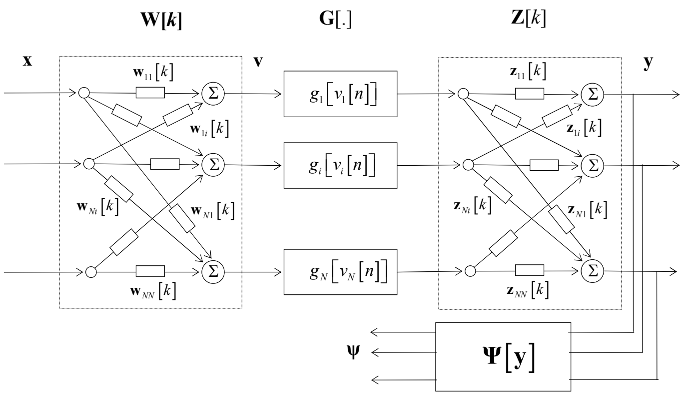


Figure 2. Feed Forward network used for the nonlinear blind deconvolution and separation.

Figure 2 shows the network used to perform the separation, it is a cascade of blocks well described in literature and previously used to resolve more simple problems. Deriving the cost function $I\{\mathbf{\Phi},\mathbf{y}\}$ with respect the learning parameter $\mathbf{\Phi}$ results:

$$\frac{\partial I\{\mathbf{\Phi},\mathbf{y}[n]\}}{\partial \mathbf{\Phi}} \simeq \frac{\partial \Im\{\mathbf{\Phi},\mathbf{y}[n]\}}{\partial \mathbf{\Phi}} = \\ = -\frac{\partial}{\partial \mathbf{\Phi}} \sum_{n=0}^{M} \left[\log\left|\det \mathbf{Z}(0)\right| + \log \prod_{i=1}^{N} g_i\left[v_i[n]\right] + \log\left|\mathbf{W}(0)\right| + \sum_{i=1}^{N} \log p_{y_i}(y_i) \right] \tag{5}$$

In (5) the expected value of the signals has been replaced by the instantaneous value. The learning rules for the elements of the FIR matrices **Z**[*k*] and **W**[*k*], and for the control points $\mathbf{Q}^g$ of the Spline neurons that compensate the nonlinear distorting functions are:

$$\partial \Im / \partial \mathbf{Z}[k] = -\mathbf{Z}[k]^{-T} \delta_k - \mathbf{\Psi}_y{}^T \mathbf{v}[n-k] \tag{6}$$

$$\partial\Im/\partial\mathbf{Q}_i^{g_j} = -\left[\dot{\mathbf{T}}_u\mathbf{M}/\dot{\mathbf{T}}_u\mathbf{M}\mathbf{Q}_i^{g_j} + \mathbf{\Psi}_y\left(\mathbf{Z}[0]\right)_j \mathbf{T}_u\mathbf{M}\right] \tag{7}$$

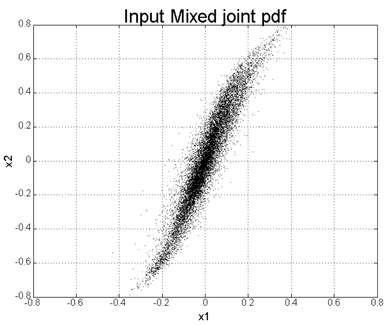
$$\partial\Im/\partial\mathbf{W}[k] = -\mathbf{Z}[0]^{-T}\delta_k - \left[\ddot{g}_1(r_1)/\dot{g}_1(r_1)\cdots \quad \ddot{g}_N(r_N)/\dot{g}_N(r_N)\right]^T \mathbf{x}[n-k] + -\sum_p\left(\mathbf{Z}[p]^T\mathbf{\Psi}\right)\mathbf{v}^T[n-p]\mathbf{x}[n-p-k] \tag{8}$$

in which **M** and **T** have the same sense as in (4).

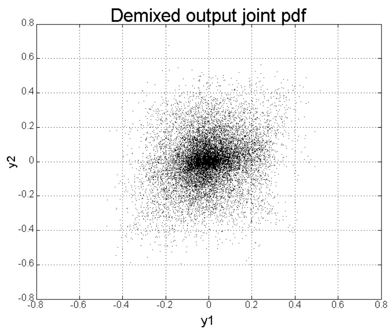
One of the main problem using FIR is the length of filters: real convolutive problems or simply non trivial ones require a large number of filter taps; must be noted that learning time grows in an exponential way with the FIR length.

5. Experimental results

This section collects the experimental result of the proposed architectures. The algorithm is able to perform the separation of *N*-channel mixtures but in order to make it possible the proper visualization of results only a pair of sources are considered: a male and a female voice speaking respectively "*Le donne i cavalier l'arme*" and "*Riperdo una seconda volta quegli esigui beni*".



a)



b)

Figure 3. a) Joint pdf of input mixture; b) Joint pdf of output demixed signals.

Figure 3 a) shows the pdf of mixed signal (the typical plot of the joint pdf of nonlinearly mixed sources) and figure 3 b) the ones of resulting signals after a 1200 epochs training: the typical plot of the joint pdf of separated signals. The recovering network has 103 Spline control points and a 15 taps FIR matrixes. The nonlinear distortions applied in this test are: $\mathbf{F}\left[f_1(p_1), f_2(p_2)\right] = \left[p_1 + 2p_1^3, 0.5p_2 + \tanh(7p_2)\right]$.

The mixing environment applied are invertible mixing MIMO channels; with respect to figure 1: $\mathbf{A} = \begin{bmatrix} 0.8 - 0.3z^{-1} + 0.3z^{-2} & 0.5 + 0.2z^{-1} - 0.2z^{-2} \\ -0.5 + 0.6z^{-1} + 0.2z^{-2} & 0.3 + 0.2z^{-1} - 0.1z^{-2} \end{bmatrix}$,

$\mathbf{B} = \begin{bmatrix} 0.7 + 0.1z^{-1} + 0.4z^{-2} & 0.4 - 0.3z^{-1} + 0.1z^{-2} \\ 0.6 + 0.5z^{-1} - 0.1z^{-2} & 0.8 + 0.2z^{-1} + 0.3z^{-2} \end{bmatrix}$.

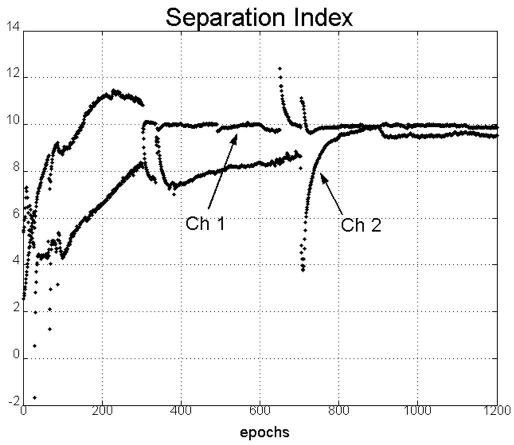


Figure 4. Separation index ratio index during the training.

The Separation index S_j (dB) introduced in [Shobben et al., 1999] measures the separation of the channel *j*-th.

$$S_j = 10\log\left[E\left\{\left(y_{\sigma(j),j}\right)^2\right\} \Big/ E\left\{\sum_{k\neq j}\left(y_{\sigma(j),k}\right)^2\right\}\right] \qquad (9)$$

In (9) $y_{i,j}$ is the *i*-th output signal when only the *j*-th input signal is present while $\sigma(j)$ is the output channel corresponding to the *j*-input. The trend of this index (Figure 4) confirms the growing of separation during the training. Figure 4 shows that, after a first period, the algorithm performs the separation of the output signals. The reason of the starting transient has been the number of blocks each of one separately have to converge to the optimum values.

6. Conclusion

This paper explores a novel mixing environment for which the BSS performed by ICA is granted. Preliminary result on separation assures a quite good sources recovery after the convolutive mixing of a PNL convolutive mixtures. Although a good separation level has been reached, we are carrying researches on improving it and on granting better output quality. The FIR recovering network performs the on line estimation of the score function by the Spline Neurons. Spline Neurons perform also the nonlinear compensating function estimation.

References

Jutten, C., Karhunen, J., (2003), “Advances in Nonlinear Blind Sources Separation”, 4th International Symposium on ICA and BSS (ICA2003), April 2003, Nara, Japan.

Taleb, A., (2002), “A Generic Framework for Blind Sources Separation in Structured Nonlinear Models”, In IEEE Trans. on signal processing, vol. 50. no 8 August 2002.

Taleb, A., Jutten, C., (1999), “Sources Separation in post nonlinear mixtures”, In IEEE Trans. on signal processing, vol. 47. no 10 August 1999.

Hyvarinen, A., Pajunen, P., (1999), “Nonlinear Independent Component Analysis: Existence and Uniqueness Results”, Neural Networks 12(2): 429-439, 1999.

Solazzi, M., Uncini, A., (2004), “Spline Neural Networks for Blind Separation of Post-Nonlinear-Linear Mixtures”, In IEEE Trans. on Circuits and Systems I Fundamental Theory and Applications, Vol. 51 , No. 4, pp 817 – 829, April 2004.

Vigliano, D., Parisi, R., Uncini, A., (2004), “A novel recurrent network for independent component analysis of Post Nonlienar convolutive mixtures”, Proc. of IEEE ICASSP’04, Montreal, Canada, May 17-21, 2004.

Uncini, A., Vecci, L., Piazza, F., (1998), “Learning and approximation capabilities of adaptive Spline activation function neural network”, In NN, Vol. 11, no. 2, pag. 259-270 March 1998.

Vigliano, D., Parisi, R., Uncini, A., (2004), “Nonlinear ICA solution for convolutive mixing of PNL mixture”, Proc. of IEEE ISCAS’04, Vancouver, Canada, May 23-26, 2004.

Milani, F., Solazzi, M., Uncini, A., (2002), “Blind Source Separation of convolutive nonlinear mixtures by flexible spline nonlinear functions”, Proc. of IEEE ICASSP’02, Orlando, USA, May, 2002.

Zade, M. B., Jutten, C., Najeby, K., (2001), “Blind Separating, Convolutive Post nonlinear Mixture”, ICA 2001 In Proc. of the 3rd Workshop on Independent Component Analysis and Signal Separation (ICA2001), San Diego (California, USA), 2001, pp. 138–143.

Shobben, D., Torkkola, K., Smaragdis, P., (1999), “Evaluation of blind signal separation methods”, In Proc. of ICA and BSS, Aussois, France, January 11-15, 1999.