Self-organizing Blind MIMO Deconvolution

Inbar Fijalkow, Philippe Gaussier
Equipe de Traitement des Images et du Signal, UPRES-A CNRS 8051
ETIS / ENSEA - Univ. de Cergy-Pontoise, 6 av du Ponceau
95014 Cergy-Pontoise Cdx, France
`fijalkow, gaussier@ensea.fr`

Abstract

We address the dynamical architecture design of linear filters for the blind adaptive restoration of several sources from convolutive mixtures. In [2], we presented a self-organizing architecture based on the dynamical stability properties of neural network lateral inhibition rules. In this paper, we improve the proposed algorithm to suit more than two sources and study its robustness to change of sources.
Keywords: **Blind MIMO deconvolution**, *Self-organization, Neural competition, Adaptive deconvolution.*

1. Introduction

The problem of blind source deconvolution is to recover the independent source signals from linear and convolutive mixtures without the knowledge of the mixing matricial transfer function. This problem is motivated by many potential applications such as biological analysis (brain activity, EEG), voices separation (cocktail party, visio-conference), or multi-user detection in telecommunications. In all these problems, several emitters sharing the same bandwidth are received by several sensors after a spatial and temporal mixing due to multipath propagation. The received signals are then the sum of convolutive mixtures of all sources where the transfer function from one source to one sensor depends on the paths for propagation. It is therefore desirable to restore each signal adaptively and blindly since the speakers can be slowly moving and the associated transfer function can not be estimated by training.

The problem of convolutive mixtures can not be reduced to the simpler instantaneous problem for which a lot of recent research has been performed. However, some criteria, such as the Constant Modulus (CM) criterion [3], can be considered for both cases when each source to be restored is an independent and identically distributed sequence (i.i.d.) and the sources independent from each other. In [5], a glob-ally convergent approach was proposed based on a hierarchical architecture to force each filter to restore a source that the other filters are not selecting. Although the algorithm in [5] gives the desired asymptotical results, it is not always suited to the filters initialization for a given (unknown system). In a recent contribution [2], we have proposed a self-organizing filters structure using dynamical stability properties exhibited by lateral inhibition mechanisms used in neural networks (see the dynamic field theory for the control of motor behavior for instance [6]). The proposed structure seems to be quite promising but was suited only in the case of 2 sources. In this proposal, we extend our proposal by explaining in depth the role of the lateral inhibitions and show its interest in the case of changing sources.

2. Problem and simulation setting

2.1. Problem setting

Let us consider P source symbols $(s_k(n))_{k=1,..,P}$ emitted from different locations and observed at the output of L spatially distributed sensors. At each instant, these observations are collected in a L-variate vector $\mathbf{y}(n)$ viewed as the sum of the P sources contributions as in $\mathbf{y}(n) = \sum_{k=1}^{P} \sum_{m=0}^{Q} \mathbf{h}_k(m) s_k(n-m)$. In the absence of noise, the FIR linear deconvolution problem consists in finding P transfer filters $(\mathbf{g}_p)_{p=1,...,P}$ such that $x(n) = \sum_{k=0}^{N-1} \mathbf{g}_p^{k\top} \mathbf{y}(n-k) = s_p(n-\nu_p)$ restores one source up to some source permutation, delay ν_p and scaling.

The blind extraction of one arbitrary source (with an arbitrary delay) using the CM algorithm (CMA) is possible if the system invertibility conditions are satisfied (see for instance [5]). The CMA updating expression is

$$\mathbf{g}_k^{(n+1)} = \mathbf{g}_k^{(n)} - \mu(|x(n)|^2 - 1)x(n)\mathbf{Y}^*(n)$$

$\mathbf{Y}(n) = (\mathbf{y}^\top(n), ...\mathbf{y}^\top(n-N))^\top$ The use of such an approach to extract several sources requires to make sure that the same source is not selected twice, even with different delays. This can be achieved using decorrelation constraints

1

as in the hierarchical criterion proposed in [5] and in a more refined way in the proposed approach.

2.2. Decorrelation approaches

All outputs decorrelation based algorithms (see [2] for a more complete list of references) can be written as:

$$\mathbf{g}_k^{(n+1)} = CMA_k(n) - \mu_2 \sum_{l \neq k} \theta_{k,l}(n) \mathbf{C}_{k,l}(n) \qquad (1)$$

where $\mathbf{C}_{k,l}(n) = \sum_m (g_k^T \hat{R}(m) g_l^* \hat{R}(m) g_l^*)$ is the gradient of the correlation between the outputs k and l, with $\hat{R}(m)$ estimate of $R(m) = [\mathbf{Y}(n)\mathbf{Y}^{*T}(n-m)]$ from the data flow.

In the hierarchical approach [5], $\theta_{k,l}$ equals 0 if $l \geq k$ and a given constant otherwise.

2.3. Simulations

In the following we explain our approach through simulation examples so that we need to set the simulations setting at once. In the simulations, the filters outputs and the global impulse responses are displayed[1].

Table I : Impulse response of 2-sources / 3-sensors system.

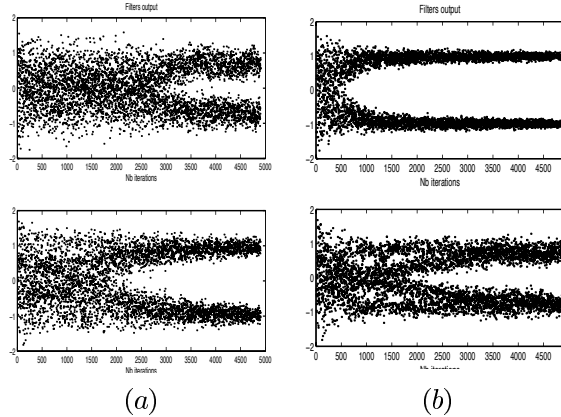| $\mathbf{h}_1(0)$	$\mathbf{h}_1(1)$	$\mathbf{h}_1(2)$	$\mathbf{h}_2(0)$	$\mathbf{h}_2(1)$	$\mathbf{h}_2(2)$
0.0400	0.9359	0.3686	0.7575	0.4756	0.6541
0.7591	0.0907	0.2895	0.8320	0.9996	0.9651
0.5619	0.2927	0.7802	0.7853	0.1786	0.1635

(a) $\qquad\qquad$ (b)

Figure 1. Hierarchical filters outputs with the initialization: (a) $g_1(z) = z^{-3}$, $g_2(z) = z^{-10}$, **(b) swapped.**

In Figure 1, we can see the hierarchical algorithm outputs when the channel is given by Table I when the filters g_1

---

[1]The quadrant i, j represents the impulse response between the source i and the output j.

ans g_2 are initialized by center-spike vectors corresponding to basins of attraction of the same source (with the non-zero component at the respective positions 3 and 10 for initialization (a) and swapped for initialization (b)). The corresponding global impulse responses are plotted in Figure 2.
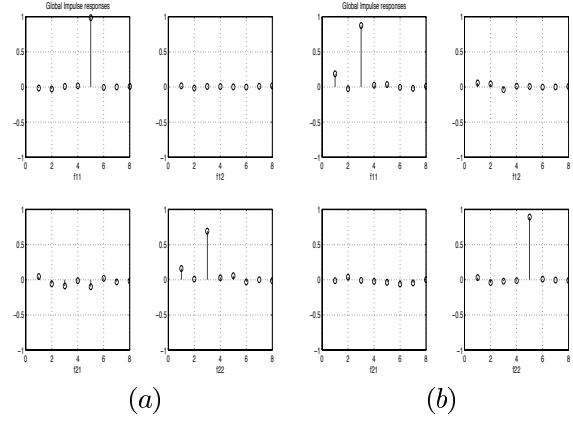

(a) $\qquad\qquad$ (b)

Figure 2. Hierarchical filters global impulse responses with the initialization: (a) $g_1(z) = z^{-3}$, $g_2(z) = z^{-10}$, **(b) swapped.**

3. Self-organizing hierarchy

3.1. Principle

In the self-organized approach [2], the scalar $\theta_{k,l}(n)$ are adapted so to match the optimal hierarchy and to suppress the decorrelation terms when no longer required. In the case of more than two sources, we propose to use,

$$\theta_{k_1,k_2}(n) = (1 - \bar{\phi}_{k_1}(n))\bar{\phi}_{k_2}(n), \quad k_1 \neq k_2 \qquad (2)$$

where $\bar{\phi}_{k_1}(n) \in [0, 1]$ measures the success of the k_1th filter to restore a source (not restored by others).

One may think of averaging in $\phi_k^{moy}(n)$ the values of $\phi_k(n) = f(1 - (|x_k(n)|^2 - 1)^2)/e_{max}$. Here e_{max} is the maximal error allowed to measure success and f is a decision device, it can be a sigmoïd. In that case, *if* $\phi_{k_1}^{moy}(n) > \phi_{k_2}^{moy}(n)$, *then k_1 is the winner*, so that we set $\bar{\phi}_{k_1}(n) = 1$ and $\bar{\phi}_{k_2}(n) = 0$. A soft-decision device can be used with $\bar{\phi}_{k_1}(n) = f(\phi_{k_1}^{moy}(n) - \phi_{k_2}^{moy}(n))$. When one CMA error is larger than e_{max}, $\phi_k(n) = 0$ so that the corresponding $\bar{\phi}_k(n)$ will decrease and will be less likely to win. This is not a problem since we assume that such an error level is too large for a successful source deconvolution. In Figure 3, we display the averaged successes (over 100 samples) $\phi_k^{moy}(n)$ in the case of the simulation setting and with initialization (a). We can see that averaging followed by a

decision rule is not sufficient to decide of a winner before $n = 1500$ which is the time required by the hierarchical approach to ensure being in different basins of attraction in the same setting.
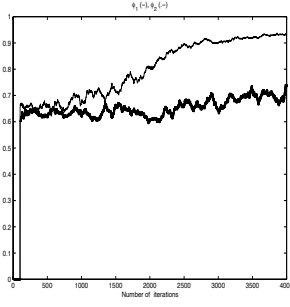

Figure 3. Evolution of $\phi_1^{moy}(n)$ and $\phi_2^{moy}(n)$

Therefore, we propose to stabilize the decision procedure in incorporating in the update of $\bar{\phi}_k(n)$ a feedback signal from the different neurons outputs at the previous iteration:

$$\bar{\phi}_{k_1}(n+1) = f((1-\lambda)\bar{\phi}_{k_1}(n)+\alpha_1\phi_{k_1}(n)-\sum_{k_2\neq k_1} w_{k_1,k_2}\bar{\phi}_{k_2}(n)) \quad (3)$$

λ and α_1 are respectively the forgetting and acquiring factors of the integration of the success measure. This competition rule (3) so called Winner Take All, is based on biological brain modeling and is represented in Figure 4, [4]. $\phi_k(n)$ is called the neuron input and $\bar{\phi}_k(n)$ the neuron output.
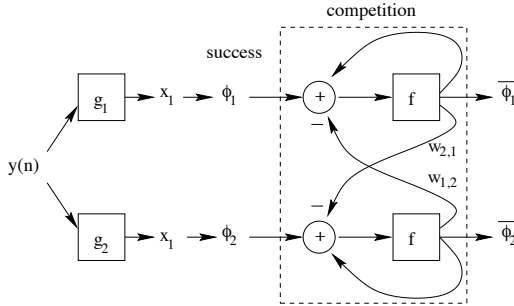

Figure 4. Competition scheme

3.2. Lateral-Inhibition rule

For more than 2 sources, we need to adapt the w (otherwise only one winner is allowed) by taking the decorrelation into account. Then, several filters win when they restore different sources. We follow the work of [1] to adapt $w_{k_1,k_2}(n + 1)$ so to increase inhibition when the two filters outputs are correlated and not successful. On the oppo-

site, the inhibition from a given source should stop if the filters outputs are uncorrelated (or sufficiently decorrelated).
$$w_{k_1,k_2}(n + 1) =$$

$$(1-\lambda_2)w_{k_1,k_2}(n)+\alpha_2\phi_{k_1}(n)\phi_{k_2}(n)f_s(c\bar{o}rr_{k_1,k_2}(n)) \quad (4)$$

for $k_1 \neq k_2$. λ_2 is a forgetting factor. f_s is a binary function defined by $f_s(c) = 0$ if c is smaller than a given threshold t_s, 1 otherwise. The last term of (4), so-called anti-hebbian term in neural networks, is meant to reinforce the decorrelation when the two outputs are correlated and to decrease it otherwise. The integrated correlation is given by:

$$c\bar{o}rr_{k_1,k_2}(n) = (1-\lambda_3)c\bar{o}rr_{k_1,k_2}(n)+\lambda_3\frac{corr_{k_1,k_2}(n)}{\|\mathbf{g}_{k_1}(n)\|\|\mathbf{g}_{k_2}(n)\|}$$

$corr_{k_1,k_2}(n) = \mathbf{g}_{k_1}^{\top}(n)\mathbf{C}_{k_1,k_2}(n)$. λ_3 is the forgetting factor.

Simulations

We compare the filters outputs and the global impulse responses[2] for the hierarchical algorithm and for the new algorithm.
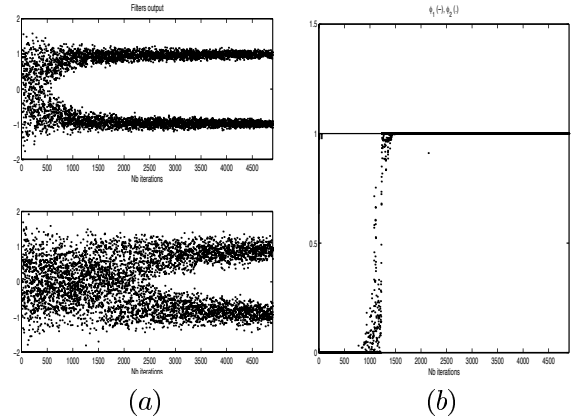

(a) (b)

Figure 5. (a) Self-organized filters output, (b) measures of success $\bar{\phi}_1$ full line, $\bar{\phi}_2$ dotted line

The simulation setting is $w_{k_1,k_2}(0) = 1$, $\lambda_2 = 0.001$, $\alpha_2 = 1$, $t_s = 0.2$ $\lambda_3 = 0.01$. First note that the self-organized algorithm always behaves similarly to the hierarchical best case. Moreover, with the self-organized hierarchy, the decorrelation term is no longer effective, after 2000 iterations, both filters are updated by CMA only. Obviously the two different sources have been restored almost perfectly in mean. Nevertheless, the second filter output (Figure 5 (a)) shows a quite large variance. To overcome this phenomenon, one may think of introducing a forgetting

---

[2]The quadrant i, j represents the impulse response between the source i and the output j.

factor (or leakage), or of switching to a decision directed mode.

Finally in Figure 5 (b), the neuron outputs are displayed. $\bar{\phi}_1$ wins almost instantaneously, and remains at 1. $\bar{\phi}_2$ is first set to 0, being therefore subject the decorrelation constraint from the first neuron. After only 1000 iterations, the two filters outputs are decorrelated enough meaning that the second filter succeeded in being in a new basin of attraction. This induces a decrease of the inhibition weight, so that the second neuron is also allowed to win. The filters adaptations are then completely decoupled (CMA adaptation only). This shows the relevance of adapting the inhibition weight.

4. Robustness to sources changes

In this section we want to show on simulations the interest of the new approach when the simulation setting changes. This can happen when a source disappears or when a new source appears. We consider the simulation setting presented in Section 2 where one source disappears at iteration 3200 and a new one appears at iteration 5300. The new source corresponds to a binary (± 1) source which contribution to the 3 sensors is described by:

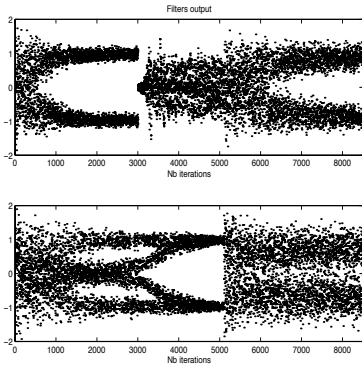| $\mathbf{h}_3(0)$	$\mathbf{h}_3(1)$	$\mathbf{h}_3(2)$
0.421	0.762	0.463
0.544	0.235	0.936
0.291	0.722	0.535

Figure 6. Self-organized filters outputs.

In Figure 6, we observe that the new approach allows to keep the extracted sources independently of the filter that leads a hierarchy. If we call $s2$ the source that remains all along, it is extracted by the second filter independently of the disappearence of the first source and the apparition of the new source. In the case of the hierarchical approach, a different scenario may take place: when the first source disappears, the leading filter tries to extract the remaining s_2

and pushes therefore the other filter to 0. The time for convergence to the new setting is much increased and the hierarchical approach can not compete with the new approach that has already converged. Moreover, when a new source appears, the filter extracting s_2 is not disturbed too much (only its variance increases) even if the new source is more powerful than s_2. Again, this nice behaviour is not guaranteed with the hierarchical approach. In Figure 7, we can see that the outputs of the hierarchical approach, with the two initial settings. In particular, in Figure 7 (b), the hierarchy is totally corrupted by the disappearance of one source.
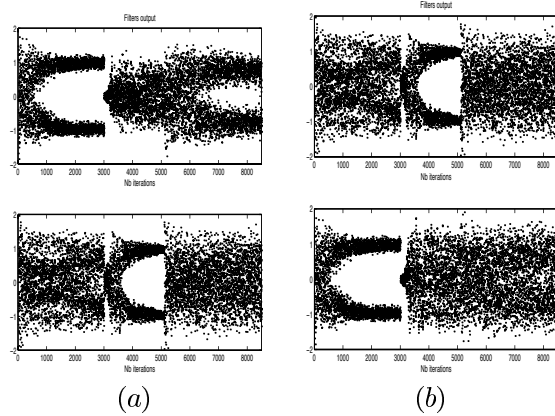

(a) (b)

Figure 7. Hierarchical filters outputs with the initialization: (a) $g_1(z) = z^{-3}$**,** $g_2(z) = z^{-10}$**, (b) swapped.**

We can see, on Figure 8, that the measures of success are not disturbed much by the different changes. On Figure 9,
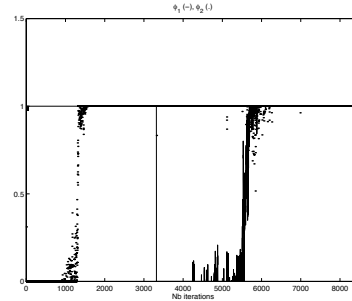

Figure 8. Self-organized approach: measures of success $\bar{\phi}_1$ full line, $\bar{\phi}_2$ dotted line.

the control of the lateral inhibition term $c\bar{o}r_{1,2}(n)$ is displayed with the instantaneous correlation showing how this term is able to enlight the important changes and decreases very fast when the filters reach basins of attraction linked to different sources. When s_2 is the only source to be active,

the high level of the inhibition pushes the first filter towards 0, which is a good starting point for the arrival of a new source.
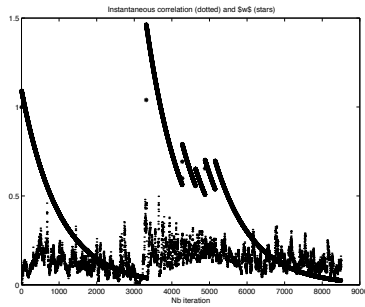

Figure 9. Compeion parameters: '.' instantaneous correlation, '*' $c\bar{o}r r_{1,2}(n)$.

5. Conclusion

We have motivated and proposed a new self-organinzing approach for the blind deconvolution of spatio-temporal MIMO mixtures. The need for lateral inhibition rules for the dynamical organization of the filters hierarchy was shown. The analysis of this approach is very complicated because of the combined adaptations. Nevertheless, we can observe, from the simulations, the robustness of the proposed approach to abrupt changes in the mixtures composition.

References

[1] P. Földiák, *Adaptive network for optimal linear feature extraction*, in Int. Joint Conf. on Neural Net., pp.401-405, 1989.

[2] I. Fijalkow, P. Gaussier, *Self-organizing Blind MIMO Deconvolution using Lateral-Inhibition*, in Proc. ICA-99, France, January 1999.

[3] D. Godard, *Self-recovering equalization and carrier tracking in two dimensional data communication systems*, IEEE Trans. on Com., 28:1867-1875, 1980.

[4] D.E. Rumhelhart, D. Zipser, *Feature discovery by competitive learning*, Cognitive science, 9:75-112, 1985.

[5] A. Touzni, I. Fijalkow, J.R. Treichler, M.G. Larimore, *A globally convergent approach for blind MIMO adaptive deconvolution*, submitted to IEEE Tr. on Signal Proc., May 1998.

[6] G. Schöner, M. Dose, C. Engels, *Dynamics of behavior: theory and applications for autonomous robot architectures*, in Robotics and Autonomous Sys., 16:(2-4)213–245, Dec., 1995.