JOINT DIAGONALIZATION OF COMPLEX SPATIAL-WAVELET MATRICES FOR BLIND SOURCES SEPARATION OF NON STATIONARY SOURCES

El Mostafa Fadaili[1], *Vincent Vigneron*[1,2]

[1]IBISC, CNRS FRE 2873, 91020 Evry-Courcouronnes, France
email: {el-mostafa.fadaili,vincent.vigneron}@ibisc.univ-evry.fr
[2]équipe SAMOS-MATISSE, CES CNRS UMR 8174, 90 rue de Tolbiac, 75634 Paris cedex 13, France
email: vigneron@univ-paris1.fr

ABSTRACT

This communication concerns the problem of blind sources separation (BSS) of non stationary sources in the instantaneous mixture case. We consider an approach based on joint-diagonalization of some hermitian matrices constructed by a family of spatial complex cross-wavelet transform. We show that mixing matrix can be estimated by using some algebraic properties of considered matrices. Computer simulations are provided to demonstrate the effectiveness of the proposed method, which is compared with other approach based on the joint-diagonalization.

***keywords** : Blind sources separation, wavelet transform, cross-wavelet, joint-diagonalisation.*

1. INTRODUCTION

We consider the blind separation of a *instantaneous mixture* of signals called *sources*. The problem has found numerous solutions in the past twenty years.
However, more recently, interest on solutions based on the use of joint-diagonalization of matrices from spatial time-frequency representations was growing [1, 3, 4, 5, 6].
In fact, such an approach, by taking advantage of the non-stationarity of the sources, makes it possible to consider a wider class of source signals than the classical "statistically independent random sources". The selection of time-frequency $(t-f)$ points that correspond only to sources auto-term with a view to build a set of matrices to be joint-diagonalized, is the main point with this kind of methods.
The purpose of this communication is to provide a original manner to build the set of Hermitian matrices to be joint-diagonalized. The approach is not based on spatial time-frequency representations, but on spatial time-scale $(t-s)$ ones – so-called *wavelets transform* – who have recently emerged as a strong mathematical tool for processing different types of signals [7, 9].
After introducing cross-wavelet and spatial-wavelet definitions in section 3, we will show that mixing matrix can be estimated by exploiting some algebraic properties of spatial-wavelet transform of sources (section 4). Section 5 will introduce a criterion for the automatic selection of matrices to be joint-diagonalized.

2. MATHEMATICAL MODELING

Suppose that N signals are received on M sensors ($M \geq N$). In matrix and vector notations, the input/output relationship of the mixing model reads:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \tag{1}$$

with $\mathbf{A}$ the (M,N) real mixing matrix which is assumed full rank, $\mathbf{x}(t) = [x_1(t),\ldots,x_M(t)]^T$ the observations vector and $\mathbf{s}(t) = [s_1(t),\ldots,s_N(t)]^T$ the deterministic sources vector. ^{T} denotes the transposition operator.
Recall that the BSS problem consists in estimating a "*separating*" matrix, say $\mathbf{B}$, when applied to the observation as

$$\mathbf{y}(t) = \mathbf{B}\mathbf{x}(t), \tag{2}$$

that yields an estimation of the source signals.
Let's define $\mathbf{G} = \mathbf{B}\mathbf{A}$ as the resulting matrix of the global system. The source separation problem is solved when one has found a matrix $\mathbf{B}$ in such a way that

$$\mathbf{G} = \mathbf{D}\mathbf{P} \tag{3}$$

where $\mathbf{D}$ is an invertible diagonal matrix which corresponds to arbitrary attenuations for the restored sources and $\mathbf{P}$ is a permutation matrix which corresponds to an arbitrary order of restitution of source signals.
A discriminating property for source signals is always required in order to perform separation, *e.g.* statistical independence for random signals, decorrelation, stationarity, etc. In the following, we consider only nonstationary deterministic signals whose *time-scale representation do not overlap completely*, *i.e.* the signatures of sources in the time-scale plane are localized in "sparse" areas. Under such assumption, we are capable to find time-scale points in each area corresponding to a different source signals (see for an illustration Fig. 2).

3. SPATIAL-WAVELET TRANSFORM

The proposed approach is based on the use of spatial-wavelets transform (SWT) and theirs properties. We briefly recall in this section outstanding points related to their utilization. More details about wavelets transform can be found in [7, 9].

3.1 Wavelet transform

A wavelet family is defined by its scale and shift parameters a and b resp. as

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}}\psi(\frac{t-b}{a}), \tag{4}$$

where so-called mother wavelet $\psi(t)$ is a real or complex function. The wavelet transform of signal $s(t)$ is the inner product:

$$W_s(a,b) = \int_{-\infty}^{+\infty} \psi_{a,b}(t) s^*(t) dt, \quad (5)$$

where $a \in \mathbf{R}^+$, $b \in \mathbf{R}$ and * denotes the conjugaison operator.
The most important properties of wavelets are the *admissibility* and the *regularity* conditions. These are the properties which gave wavelets their name. It can be shown that square integrable functions $\psi(t)$ satisfying the admissibility condition :

$$\int_{-\infty}^{+\infty} \frac{|\Psi(w)|^2}{|w|} dw < +\infty, \quad (6)$$

where $\Psi(w)$ stands for the Fourier transform of $\psi(t)$, can be used to first analyze and then reconstruct a signal without loss of information.

3.2 Cross-wavelet and spatial-wavelet transform

The cross-wavelet transform (CWT) between two signals $s_1(t)$ and $s_2(t)$ is defined, when the same mother wavelet $\psi(t)$ is considered, as :

$$W_{s_1 s_2}(a,b) = W_{s_1}(a,b).W^*_{s_2}(a,b). \quad (7)$$

From (7), one can define the spatial-wavelet transform (SWT) matrix of vector $\mathbf{s}(t) = [s_1(t), \ldots, s_N(t)]$ as the collection of the cross-wavelet components :

$$\{\mathbf{W_s}(a,b)\}_{i,j} = W_{s_i s_j}(a,b). \quad (8)$$

Under the mixture model (1), $\forall (i,j) \in \{1,\ldots,M\}$:

$$\begin{aligned} W_{x_i x_j}(a,b) &= W_{x_i}(a,b).W^*_{x_j}(a,b) \\ &= \sum_{l,m=1}^{N} a_{il} a_{jm} W_{s_l s_m}(a,b). \quad (9) \end{aligned}$$

Then, the SWT of vector $\mathbf{x}(t)$ can be written as :

$$\mathbf{W_x}(a,b) = \mathbf{A}\mathbf{W_s}(a,b)\mathbf{A}^T. \quad (10)$$

If the mother wavelet $\psi(t)$ is complex[1], then matrix $\mathbf{W_s}(a,b)$ is $N \times N$ square Hermitian. When matrices $\mathbf{W_s}(a,b)$ are diagonal for some (a,b), $\mathbf{A}$ can be estimated using simultaneous joint-diagonalization algorithm of matrices $\{\mathbf{W_x}, (a,b)\}$.

4. NON ORTHOGONAL JOINT DIAGONALIZATION

Let us now describe briefly the problem of non-orthogonal joint-diagonalization. We consider the set $\mathscr{M}$ of K matrices $\mathbf{M}_i$, $i \in \{1,\ldots,K\}$ which all admit the following decomposition : there exists a matrix $\mathbf{A}$ and K diagonal matrices $\mathbf{D}_i$, $i \in \{1,\ldots,K\}$ such that

$$\mathbf{M}_i = \mathbf{A}\mathbf{D}_i\mathbf{A}^T \ , \quad \forall i \in \{1,\ldots,K\} \ .$$

The problem is to estimate the matrix $\mathbf{A}$ and the diagonal matrices $\mathbf{D}_i$, $i \in \{1,\ldots,K\}$ from the matrices set $\mathscr{D}$.

[1] This condition will be considered in all next sections.

When $\mathbf{A}$ is orthogonal, the above problem has been reported in [1] where some solution can be found. For the *non-orthogonal* case, we propose to consider the following objective function:

$$\mathscr{C}(\mathbf{B}) = \sum_{i=1}^{K} \|\mathsf{OffDiag}\{\mathbf{B}^T\mathbf{M}_i\mathbf{B}\}\|^2, \quad (11)$$

where the operator $\mathsf{OffDiag}\{\cdot\}$ is defined as the zero-diagonal matrix built from the off-diagonal components of the matrix argument. In fact we are looking for matrix $\mathbf{B}$ that minimizes the criterion $\mathscr{C}(\mathbf{B})$. In that case, this optimal matrix argument plays directly the role of a separating matrix.

In [3], we have proposed an algorithm for the optimization of $\mathscr{C}(\mathbf{B})$ without the unitary constraint. Other methods without constraint can be found in [10].

5. TIME-SCALE POINT SELECTION

In order to build the matrice set $\mathscr{M} = \{\mathbf{W_x}(a,b)/(a,b)\}$ to be joint-diagonalized, it is necessary to find, only from observation data, a set of points (a,b) in time-scale plane for which the matrices $\mathbf{W_s}(a,b)$ are diagonal.
Let us discuss here about different algebraic structure of matrices $\mathbf{W_s}(a,b), (a,b) \in \mathbf{R}^+ \times \mathbf{R}$. We can distinguish three different cases :

- Matrix $\mathbf{W_s}(a,b)$ is zero, then $\|\mathbf{W_x}(a,b)\| = 0$ and the point (a,b) is not considered.
- Matrix $\mathbf{W_s}(a,b)$ has no particular algebraic structure, then, $\mathbf{W_x}(a,b)$ is not "interesting" for the matrice set $\mathscr{M}$.
- Finally, matrix $\mathbf{W_s}(a,b)$ is diagonal, so $\mathbf{W_x}(a,b)$ can be added to $\mathscr{M}$.

In the last case, *i.e.* when matrix $\mathbf{W_s}(a,b)$ is diagonal, according to hermitian symmetry and because the matrix $\mathbf{A}$ is real, the matrice $\mathbf{W_x}(a,b)$ is real. This last property can be exploited to select point (a,b) in the time-scale plane where $\mathbf{W_s}(a,b)$ is diagonal :

$$\|\mathbf{W_x}(a,b)\| > \varepsilon_1 \quad (12)$$
$$\|\Im\{\mathbf{W_x}(a,b)\}\| < \varepsilon_2 \quad (13)$$

where $\Im\{.\}$ and $\|.\|$ respectively stands for the imaginary part and the Euclidian norm, ε_1 and ε_2 being fixed thresholds.
We can also generalize time-frequency points selection procedure proposed in [6] to time-scale plane. The idea of proposed criterion is to keep points with a sufficient energy and then to use the rank one property to detect single auto-terms. Indeed, if matrix $\mathbf{W_s}(a,b)$ is diagonal, then it is rank one, because, if we suppose that $W_{s_i}(a,b) \neq 0$ and $W_{s_j}(a,b) \neq 0$ for some $i \neq j$, then, from (7) $W_{s_i s_j}(a,b) \neq 0$ and matrix $\mathbf{W_s}(a,b)$ will not be diagonal.
Notice that other points selection procedure in time-frequency plane requiring spatial whitening step can be found in [1, 5]. We recall that spatial whitening stage limits the performance of separation due to error estimation of whitening matrix. For this reason, it is not considered here both for the joint-diagonalization and the time-scale point selection procedure.

6. COMPUTER SIMULATIONS

In this simulation, we consider $N = 3$ source signals (Figure 1) of length $T = 2048$ and $M = 3$ mixtures. The sources are defined by : $s_1(t) = sin(250\pi t^2)$, $s_2(t) = sin(150\pi t^3)$, $s_3(t) = sin(350\pi t)$, $t \in [0,1]$.
The mixing matrix is :

$$\mathbf{A} = \begin{pmatrix} 1 & 0.7 & 0.4 \\ -0.3 & 1 & -0.7 \\ 0.9 & 1.5 & 1 \end{pmatrix} \qquad (14)$$

The following performance index is used as a measure of the separation quality [8] :

$$\begin{aligned} I(\mathbf{G}) &= \frac{1}{N(N-1)} \sum_{i=1}^{N} \left(\sum_{j=1}^{N} \left(\frac{|(\mathbf{G})_{i,j}|^2}{\max_{\ell} |(\mathbf{G})_{i,\ell}|^2} - 1 \right) \right) \\ &+ \frac{1}{N(N-1)} \sum_{j=1}^{N} \left(\sum_{i=1}^{N} \left(\frac{|(\mathbf{G})_{i,j}|^2}{\max_{\ell} |(\mathbf{G})_{\ell,j}|^2} - 1 \right) \right), \end{aligned}$$

with $N = 3$ the dimension of $\mathbf{G}$. This index, given in dB, is defined by $I(.)\ dB = 10\log(I(.))$. It measures the distance between the global matrix and the product of diagonal matrix and permutation matrix.
We consider the complex Morlet mother wavelet defined as : $\psi(t) = \frac{1}{\sqrt{\pi f_b}} e^{2i\pi f_c t} e^{-\frac{t^2}{f_b}}$. Time-scale values (a,b) are calculated on $[1,\ldots,32] \times [1,\ldots,T]$.
The sources are represented in time in Figure 1 and in the time-scale plane in Figure 2, Figure 3 shows mixtures in the same plane. In Figure 4 are selected $K = 3944$ points by respect to criteria (12) and (13) with $\varepsilon_2 = 10^{-2}$ and ε_1 equal to the mean of $\|\mathbf{W}_\mathbf{x}(a,b)\|$ from all points (a,b). As shown in this figure, every selected point must corresponds to only one source and all other points are rejected. In other words, the points where there is “interference” between two time-scale signatures (Figure 3) are not selected.
After joint-diagonalizing the set matrices $\mathscr{M}$ built from selected points, matrix $\mathbf{B}$ is estimated and the value of performance index is $I(\mathbf{BA})\ dB = -69.7677$.
Finally, in Figure 6, the performance index of the proposed approach (WBSS) is compared to JADE[2] algorithm [2] in noisy mixture case for SNR varying between $5dB$ and $50dB$. The noisy mixture is achieved by adding random signal with uniform distribution to each components of the vector $\mathbf{x}(t)$. The SNR is defined by $10\log(\frac{1}{\sigma^2})$, where σ^2 denotes the variance of the noise. The result shows the superiority of WBSS as compared to JADE at higher SNR.

7. CONCLUSION

We emphasize in this paper that the separation of non stationary deterministic sources having different time-scale representations can be realized. Proposed solution is based on the joint-diagonalization of matrices set built from spatial time-scale transform, using some automatic selection criterion in time-scale plane. We have illustrated the effectiveness of the proposed method thanks to computer simulations.

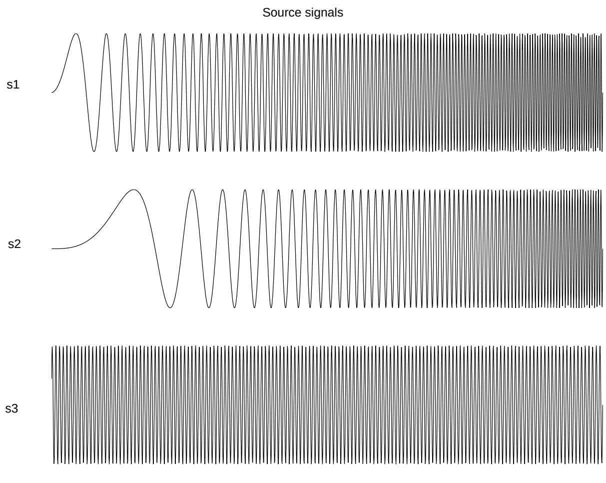

Figure 1: Source signals $s_1(t)$, $s_2(t)$ and $s_3(t)$ before mixing.

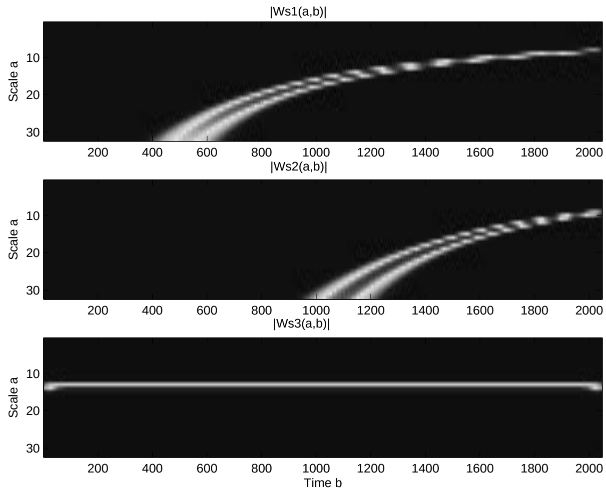

Figure 2: Time-scale transform of sources using complex Morlet wavelet.

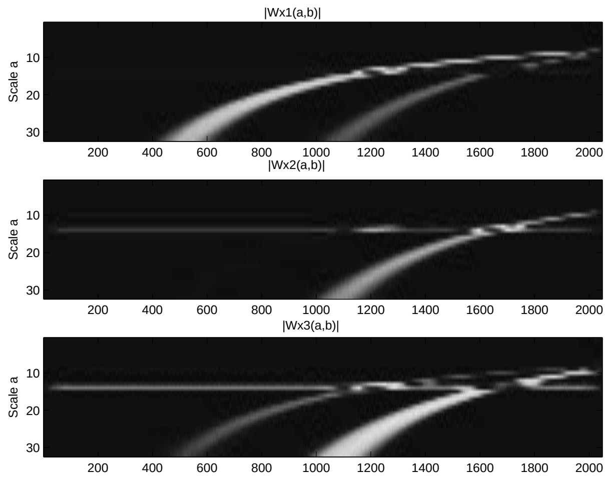

Figure 3: Time-scale transform of mixtures using complex Morlet wavelet.

[2]This algorithm is based on use of joint diagonalization of cumulant matrices.

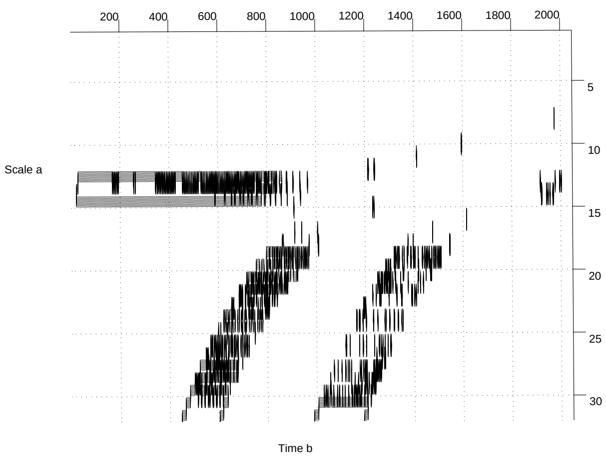

Figure 4: Time-scale points (a,b) selected to build matrices to be diagonalized. Each "energetic" point corresponds to only one source.

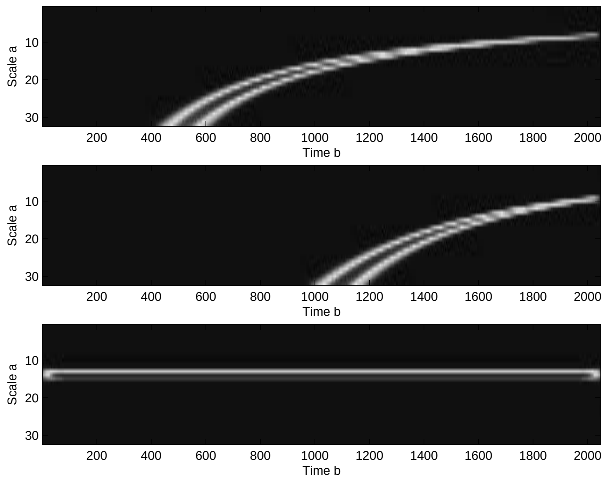

Figure 5: Time-scale representations of the three reconstructed sources.

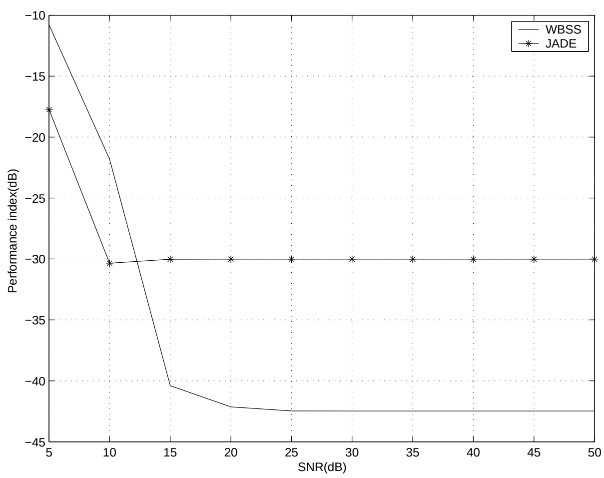

Figure 6: Performance index (dB) for proposed approach (WBSS) and JADE.

A deep study of noisy mixture can be interesting for future work since the wavelet transform can be used both for noise cancellation and source separation. A more challenging case will be also when the mixing system is a convolutive one, i.e., when the sources are mixed through a linear filtering operation.

REFERENCES

[1] A. Belouchrani, M. G. Amin, "Blind source separation based on time-frequency signal representations," *IEEE Trans. Signal Processing*, Vol. 46, No. 11, pp. 2888-2897, Nov. 1998.

[2] J.-F. Cardoso and A. Souloumiac, "Blind Beamforming for non Gaussian Signals", *IEEE Proceedings-F*, Vol. 40, pp. 362-370, 1993.

[3] E.M. Fadaili, N. Thirion-Moreau, E. Moreau, "Combined Non-Orthogonal Joint Zero-diagonalization and Joint Diagonalization for Sources Separation", *In Proc. IEEE Workshop on Statistical Signal Processing*, Bordeaux, France, July 2005.

[4] E.M. Fadaili, N. Thirion-Moreau, E. Moreau, "Non orthogonal joint diagonalization /zero diagonalization for sources separation based on time-frequency distributions", *IEEE Transactions on Signal Processing*, vol. 55, No. 4, April 2007.

[5] C. Févotte and C. Doncarli, "Two contributions to blind source separation using time-frequency distributions", *IEEE Signal Processing Letters*, Vol. 11, No. 3, pp. 386-389, March 2004.

[6] L. Giulieri, H. Ghennioui, N. Thirion-Moreau and Eric Moreau, "Non-Orthogonal Joint-Diagonalization of Spatial Quadratic Time Frequency Matrices for Source Separation", *IEEE Signal Processing Letters*, Vol. 12, No. 5, pp. 415-418, May 2005.

[7] S. Mallat, "A thiry for multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pattern Anal. and Mach. Intelligence*, 11(7), 1989, p. 674-693.

[8] E. Moreau, "A generalization of joint-diagonalization criteria for source separation", *IEEE Trans. Signal Processing*, Vol. 49, Nř 3, pp. 530–541, March 2001.

[9] M.B. Ruskai et al, "Wavelets and their applications", *Jones and Bartlett Publ*, Boston ,1992.

[10] A.Yeredor, "Non-Orthogonal Joint Diagonalization in the Least Square Sense With Application in Blind Source Separation". *IEEE Transactions on signal processing*, vol. 50, No. 7, July 2002.