A WAVELET APPROACH TO CONVOLUTIVE BLIND SEPARATION OF NON-STATIONARY SOUND SOURCES

Kiyotaka Takada * **Kazushi Nakano** *
Hirokazu Watai *

* *Dept. of Electronic Engineering, The University of Electro-Communications, 1-5-1 Chofu-ga-oka, Chofu, Tokyo 182-8585, Japan*

Abstract: This paper considers a wavelet-based approach to the blind separation of signals (BSS) including reflections such as sound. The BSS for the instantateous mixture case is achieved by maximizing the mutual information between input and output signals or by zeroing the cross-correlations of separated signals. For making these methods applicable to the convolutive mixture case, it is important to transform signals in the time domain into those in the time-frequency domain by the wavelet transforms with the Gabor function. Simulation results demonstrate the effectiveness of our wavelet-based approach.

Keywords: blind source separation, convolutive mixing process, wavelet transform, Gabor function, InfoMax method, cross-correlation method.

1. INTRODUCTION

The blind source separation (BSS) in the independent component analysis (ICA) attracts much of our attention to its broad applications: crosstalk removal in multichannel communications, improvement over beamforming microphones for audio, and discovery of independent sources in various biological signals such as EEG, MEG and so on. The BSS is a signal processing technique for reconstructing source signals, which are statistically independent of each other, only by using the measurements given through a linear mixing process. Since the mixing process is unknown, the BSS could be achieved if the estimates of source signals become independent of each other.

One of the major approaches to the BSS is to use higher-order statistics. For example, Bell and Sejnowski (1995) presented the InfoMax principle, in which application of the InfoMax principle to source separation maximizes an output entropy. This algorithm needs iterative calculation and requires non-linear optimization. On the other hand, Molgedey and Schuster (1994) presented the algorithm using second-order statistics in the form of correlation function. Their algorithm needs neither higher-order statistics nor iterative calculations.

When applying a BSS system to a real acoustic environment such as a number of people talking in a room, the BSS system performs poorly because it removes mainly the sounds from the jammer direction. This is the reason for the difficulty of BSS in reverberant environments. Such a mixture is called a "convolutive mixture," not "instantaneous mixture."

This paper presents a wavelet-based algorithm for the BSS in convolutive mixing process. First of all, we transform signals in the time domain into those in the time-frequency domain by the wavelet transform (WT) (Mallat, 1998), and make a set of time series fixed at each frequency. Next, we apply the BSS algorithms to the time series at a fixed frequency. Gathering and rearranging the time series for all the frequencies, and transforming from the time-frequency domain to the time domain, we can obtain a solution of this type of BSS problem. However, this is not always true because of the so-called scaling and permutation issues. Thus, each signal is independently replaced at each scale a, and also its magnitude remains indefinite. For these issues, we try to divide the spectrogram using an independent signal at each scale a. Furthermore, we try to solve the permutation issue by using the correlation between envelopes. Lastly, we demonstrate the superiority of our wavelet-based approach to the InfoMax and cross-correlation methods via computer simulation.

2. REVIEW OF BSS PROBLEM

2.1 Original Problem Statement

The source signals are denoted by

$$\mathsf{s}(t) = [s_1(t), \ldots, s_n(t)]^T \tag{1}$$

In the BSS, there is the assumption that each component of $\mathsf{s}(t)$ is independent of each other. For defining the statistical independence of $\mathsf{s}(t)$, we use the formula of

$$p(s) = \prod_{i=1}^{n} p_i(s_i) \tag{2}$$

This means that, if any different signals are independent of each other, their joint distribution is represented as product of marginal distributions. The mixed signals (recorded signals) are denoted by

$$\mathsf{x}(t) = [x_1(t), \ldots, x_n(t)]^T \tag{3}$$

In the basic BSS problem, we assume that the mixed signals are only linear mixtures of sources. That is, the following formula represents non-delayed linear mixing.

$$\mathsf{x}(t) = A\mathsf{s}(t) \tag{4}$$

where A is an unknown constant matrix denoting linear channels. Another case is a matrix of FIR filters, which is used as a model of recording in a real environment. In this paper, we focus on the later case.

The task of BSS is to estimate matrix B lead to independent signals without knowing the information on the channels A and the probability distribution of source signals $\mathsf{s}(t)$. The separated signals are represented by

$$\mathsf{y}(t) = B\mathsf{x}(t) \tag{5}$$

Ideally, we have to find B to be the inverse of the channels A. We however cannot obtain the information on the amplitude and the order of the source signals, where there remains the indefiniteness of amplitude and permutation. This fact is represented by the following relation:

$$BA = PD \tag{6}$$

where P is a permutation matrix, and D is a diagonal matrix.

Major approaches to the BSS are based on the following approaches:

- Equalizing the joint distribution function and marginal distribution functions of the separated signals.
- Zeroing the cross-correlation functions of the separated signals as much as possible.

We will give an account of these principles.

2.2 InfoMax Method

There are some approaches based on probability distributions. We review "InfoMax method" in this section.

The InfoMax method was suggested by Bell and Sejnowski (1995). The algorithm makes independent signals by maximizing the mutual information between the mixed and separated signals.

Assuming that the input signals x and the output ones y of network, the mutual information $I(\mathsf{y},\mathsf{x})$ between x and y is defined by

$$I(\mathsf{y},\mathsf{x}) = H(\mathsf{y}) - H(\mathsf{y}|\mathsf{x}) \tag{7}$$

where $H(\mathsf{y})$ is the entropy of outputs y, and $H(\mathsf{y}|\mathsf{x})$ is the conditional entropy of outputs y with inputs x.

A learning rule is given by computing the gradient with respects to weight of network B. Here, considering that $H(\mathsf{y}|\mathsf{x})$ does not depend on weight B, the gradient is written as

$$\frac{\partial}{\partial B} I(\mathsf{y},\mathsf{x}) = \frac{\partial}{\partial B} H(\mathsf{y}) \tag{8}$$

Maximizing the mutual information between inputs and outputs corresponds with maximizing the entropy of outputs. Also it leads to minimizing the mutual information among each component of outputs, and the output signals become independent signals. Assuming that the probability density functions of inputs and outputs are denoted as $p(\mathsf{x})$ and $p(\mathsf{y})$ respectively, their relationship is given by the formula:

$$p(\mathsf{y}) = \frac{p(\mathsf{x})}{|\ J\ |}, \qquad J = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_n}{\partial x_1} & \cdots & \frac{\partial y_n}{\partial x_n} \end{bmatrix} \tag{9}$$

By use of this relationship, the entropy of outputs $H(\mathsf{y})$ and eq.(8) are rewritten espectively by

$$H(\mathsf{y}) = -\int_{-\infty}^{\infty} p(\mathsf{y}) \log p(\mathsf{y}) d\mathsf{y} = -E[\log p(\mathsf{y})]$$
$$= E[\log | J |] - E[\log p(\mathsf{x})] \quad (10)$$

$$\frac{\partial}{\partial B} I(\mathsf{y}, \mathsf{x}) = \frac{\partial}{\partial B} H(\mathsf{y})$$
$$= \frac{\partial}{\partial B} E[\log | J |] - \frac{\partial}{\partial B} E[\log p(\mathsf{x})] \quad (11)$$

The second term does not depend on the weight B, so a learning rule of the weight B is lastly represented as

$$\Delta B \propto \frac{\partial}{\partial B} H(\mathsf{y}) = \frac{\partial}{\partial B} E[\log | J |] = J^{-1} \cdot \frac{\partial J}{\partial B} \quad (12)$$

In this paper, we set $\mathsf{y} = \text{sgm}(\mathsf{u})$ and $\mathsf{u} = B\mathsf{x}$, where $\text{sgm}(\mathsf{u}) = 1/(1 + e^{-\mathsf{u}})$. Thus, eq.(12) is rewritten by

$$\frac{\partial}{\partial B} H(y) = (B^T)^{-1} + (\mathbf{1} - 2\mathsf{y})\mathsf{x}^T \quad (13)$$

where $\mathbf{1}$ is a vector whose components are one.

We can obtain the separating matrix B_{t+1} by updating B_t with respect to each measured data.

$$B_{t+1} = B_t + \eta \Delta B \quad (\eta : \text{learning ratio}) \quad (14)$$

2.3 Cross-correlation Method

There are some algorithms for zeroing cross-correlations (Murata, et al., 1998). We use the algorithm that consists of two procedures, "sphering" and "rotation".

(1) Sphering
The sphering is an operation for orthogonalizing the source signals in the measuring coordinate. Let us define a covariance matrix of measurements

$$R = E\left[\mathsf{x}(t)\mathsf{x}(t)^T\right] \quad (15)$$

and define its square root inverse

$$\sqrt{R^{-1}} = \sqrt{\Lambda^{-1}} Q^T \quad (16)$$

where Q and Λ are an orthogonal matrix and a diagonal matrix, respectively, which satisfy

$$R = Q\Lambda Q^T \quad (17)$$

and $\sqrt{\Lambda^{-1}}$ denotes a diagonal matrix, and each of whose elements is given by the square root of Λ^{-1}'s elements. By transforming the measurement vector such that

$$\mathsf{z}(t) = \sqrt{R^{-1}} \mathsf{x}(t) \quad (18)$$

the covariance matrix of the new vector $\mathsf{z}(t)$ is orthogonalized, i.e.

$$E\left[\mathsf{z}(t)\mathsf{z}(t)^T\right] = \sqrt{R^{-1}} R \sqrt{R^{-1}}^T = I \quad (19)$$

where I is the identity matrix.

(2) Rotation
Even after sphering measured signals, there still remains an ambiguity of rotation. The correct rotation is determined by removing the off-diagonal elements of the correlation matrix at several time delay. A possible implementation is to find an orthogonal matrix C which minimizes

$$\sum_{k=1}^{r} \sum_{i \neq j} \left|(CD_kC^T)_{ij}\right|^2 \quad (20)$$

where $(CD_kC^T)_{ij}$ denotes the (i, j)-element of matrix CD_kC^T and

$$D_k = E\left[\mathsf{z}(t)\mathsf{z}(t + \tau_k)^T\right] \quad (k = 1, \ldots, r) \quad (21)$$

To solve this approximate simultaneous diagonalization problem, Cardoso and Souloumiac (1996) proposed a Jacobi-like algorithm. We use their method in our implementation.

With these two operations, the separation matrix B is given by

$$B = C\sqrt{R^{-1}} \quad (22)$$

3. WAVELET-BASED BSS FOR CONVOLUTIVE MIXTURE

3.1 Wavelet Transform

The wavelet transform is considered as an extension of the Fourier transform (FT) so as to overcome the difficulty in analyzing any local property of signals. A mother wavelet $\psi(t)$ is a function of zero average, which is dilated with a scale a, and translated by a position b such that

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad (23)$$

The WT of $f(t)$ at the scale a and position b is then defined by correlating $f(t)$ with a wavelet atom of eq.(23):

$$F(a, b) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \psi^*\left(\frac{t-b}{a}\right) dt \quad (24)$$

where "ψ^*" denotes the complex conjugate of "ψ". Like the windowed FT, the WT can measure the time-frequency variations of spectral components, but it has a different time-frequency resolution.

3.2 Filter Characteristic of Wavelet Transform

By applying the Parseval formula, the WT can be rewritten in the frequency-integral form.

$$F(a,b) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{f}(\omega)\hat{\psi}^{*}_{a,b}(\omega)d\omega \tag{25}$$

where "$\hat{\psi}$" denotes the FT of "ψ". The wavelet coefficient $F(a,b)$ thus depends on the value of $f(t)$ (and $\hat{f}(\omega)$) in the time-frequency domain where the energy of $\psi_{a,b}(t)$ (and $\hat{\psi}_{a,b}(\omega)$) is concenrated. Time varying harmonics are detected from the scale and position of high amplitude wavelet coefficients.

According to eq.(25), the WT has a filter characteristic. Correctly speaking, if $\hat{\psi}_{a,b}(\omega)$ is localized in the frequency domain, the WT corresponds to band-pass filter. For example, when using the Gabor function which has the minimum uncertainty in the time-frequency domain, a mother wavelet $\psi(t)$ is written by

$$\psi(t) = \pi^{-\frac{1}{4}}\sqrt{\frac{\omega_p}{\gamma}}\exp(-\frac{\omega_p^2}{2\gamma^2}t^2)\exp(j\omega_p t) \tag{26}$$

where γ is a positive constant relating to the admissibility condition (zero-average condition), and ω_p is the center frequency. Its FT and the scale-shift transform are written respectively by

$$\hat{\psi} = \pi^{\frac{1}{4}}\sqrt{\frac{2\gamma}{\omega_p}}\exp(-\frac{\gamma^2}{2\omega_p^2}(\omega-\omega_p)^2) \tag{27}$$

$$\hat{\psi}_{a,b} = \exp(-jb\omega)\sqrt{a}\hat{\phi}(a\omega) = \exp(-jb\omega)\sqrt{a} \cdot\pi^{\frac{1}{4}}\sqrt{\frac{2\gamma}{\omega_p}}\exp(-\frac{\gamma^2}{2\omega_p^2}(a\omega-\omega_p)^2) \tag{28}$$

Fig. 1 shows an example of the frequency characteristics of eq.(28). As shown in Fig. 1, $\hat{\psi}_{a,b}(\omega)$ is a band-pass filter showing localized frequency characteristics with a peak at $\omega = \omega_p/a$. Hereafter, we will use this Gabor function as $\psi(t)$ (Takada, et al., 2004).

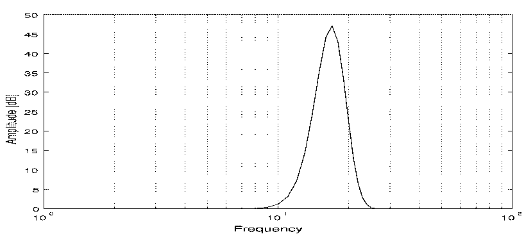

Fig. 1. Spectrogram of $\hat{\psi}_{a,b}$ ($\gamma = 2\pi,\ \omega_p = 0.1$)

3.3 Wavelet-based BSS for Convolutive Mixture

The measured signals based on convolutive mixture are defined by

$$\mathsf{x}(t) = A(t) * \mathsf{s}(t) \tag{29}$$

where $*$ denotes the convolution integral. We now take the wavelet transform of $\mathsf{x}(t)$ with $\psi(t)$.

$$X(a,b) = \int_{-\infty}^{\infty} \mathsf{x}(t)\frac{1}{\sqrt{a}}\psi^{*}(\frac{t-b}{a})dt \tag{30}$$

Eq.(30) is computed actually by discretizing in the following manner:

$$\omega = \omega_p/a = \omega_0 \times 2^{-\alpha} \quad (\alpha = 0, \Delta\alpha, 2\Delta\alpha, \ldots)$$
$$b = 0, \Delta T, 2\Delta T, \ldots \tag{31}$$

where ω_0 is the maximum frequency to be considered.

When using the Gabor function as $\psi(t)$, the relationship between the measured and source signals is

$$X(\omega,b) \simeq \hat{A}(\omega)S(\omega,b) \tag{32}$$

where $\hat{A}(\omega)$ is the FT of $A(t)$, and $X(\omega,b)$ and $S(\omega,b)$ are the WTs of $\mathsf{x}(t)$ and $\mathsf{s}(t)$, respectively (Tabaru, et al., 2005).

By fixing the frequency ω, $X(\omega,b)$ is represented as

$$X_{\omega}(b) = X(\omega,b) \tag{33}$$

Since this equation is only time series of b, we can apply any non-delay BSS algorithm to convolutive cases.
After applying the algorithm, we can get the time series of estimates whose components are mutually independent for each frequency ω,

$$U_{\omega}(b) = B_{\omega}X_{\omega}(b) \tag{34}$$

Gathering and rearranging the time series for all the frequencies, and transforming from the time-frequency domain to the time domain, we can obtain a solution of this type of blind separation. However, this is not always true because of the so-called scaling and permutation issues. Thus, each signal is independently replaced at each a, and also its magnitude remains indefinite.

First, to solve the scaling problem, we try to divide the spectrogram using an independent signal at each scale a.

$$V_{\omega}(b;i) = B_{\omega}^{-1}\begin{pmatrix} 0 \\ U_{i,\omega}(b) \\ 0 \end{pmatrix} \tag{35}$$

where index i denotes the dependence of the spectrograms at ω on the i-th independent component of $U_\omega(b)$. In order to obtain $V_\omega(b;i)$, we utilize B_ω and B_ω^{-1}. Then, $V_\omega(b;i)$ does not have an ambiguity of magnitude.

The second problem is permutation. Based on the non-stationarity of the source signals, if the split band-passed signals $V_\omega(b;i)$ originate from the same source signals, it is natural to assume that they are under the influence of a similar modulation in amplitude. We define an operator ε such that

$$\varepsilon V_\omega(b;i) = \frac{1}{2M}\sum_{t_s^0=b-M}^{b+M}\sum_{j=1}^{n}|V_{j,\omega}(t_s';i)| \quad (36)$$

where M is a positive constant, and $V_{j,\omega}(b;i)$ denotes the j-th element of $V_\omega(b;i)$. We define also its inner product and norm such that

$$\varepsilon V_\omega(i)\cdot\varepsilon V_{\omega^0}(j) = \sum_b \varepsilon V_\omega(b;i)\cdot\varepsilon V_{\omega^0}(b;j) \quad (37)$$

$$\|\varepsilon V_\omega(i)\| = \sqrt{\varepsilon V_\omega(i)\cdot\varepsilon V_\omega(i)} \quad (38)$$

We solve the permutation by sorting them. Sorting is determined with the correlation between the envelopes of band-passed signals.

- Sort ω in order of the weakness of correlation between independent components in ω. This is done by sorting in increasing order of

$$\mathrm{SIM}(\omega) = \sum_{i\neq j}\frac{\varepsilon V_\omega(i)\cdot\varepsilon V_\omega(j)}{\|\varepsilon V_\omega(i)\|\cdot\|\varepsilon V_\omega(j)\|} \quad (39)$$

$$\mathrm{SIM}(\omega_1)\le\mathrm{SIM}(\omega_2)\le\ldots\le\mathrm{SIM}(\omega_N) \quad (40)$$

- For ω_1, assign V_{ω_1} to Y_{ω_1} as it is:

$$Y_{\omega_1}(b;i) = V_{\omega_1}(b;i) \quad (i=1,\ldots,n) \quad (41)$$

- For ω_k, find the permutation $\sigma(i)$ which maximizes the correlation between the envelope of ω_k and the aggregated envelope from ω_1 through ω_{k-1}. This is achieved by maximizing

$$\sum_{i=1}^{n}\frac{\varepsilon V_{\omega_k}(\sigma(i))\cdot\left(\sum_{j=1}^{k-1}\varepsilon Y_{\omega_j}(i)\right)}{\|\varepsilon V_{\omega_k}(\sigma(i))\|\cdot\|\varepsilon\left(\sum_{j=1}^{k-1}\varepsilon Y_{\omega_j}(i)\right)\|} \quad (42)$$

  within all the possible permutation σ of i.
- Assign the appropriate permutation to Y_{ω_k}:

$$Y_{\omega_k}(b;i) = V_{\omega_k}(b;\sigma(i)) \quad (43)$$

As a result, we can solve the permutation ambiguity and obtain the separated spectrograms:

$$Y(\omega,b;i) = Y_\omega(b;i) \quad (44)$$

4. SIMULATION RESULTS

4.1 Preliminaries

We used a set of data mixed on a computer, and applied the BSS algorithm to it. Fig. 2 shows the source signals recorded on the computer. The measured signals mixed by the following matrix are shown in Fig. 3: (see the last reference):

$$\mathsf{x}(t) = A(t)*\mathsf{s}(t) = \begin{pmatrix}\sum_j a_{1j}(t)*s_j(t)\\ \sum_j a_{2j}*s_j(t)\end{pmatrix}$$

$$a_{ij}*s_j(t) = \sum_{\tau=0}^{\infty}a_{ij}(\tau)s_j(t-\tau) \quad (45)$$

where we used actually the following values (http://www.murata.elec.waseda.ac.jp/murata/lecture/ice/note/):

$$A(t) = \begin{pmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{pmatrix}$$

$$\begin{cases}a_{11} = 0.469\\ a_{12} = 0, 0.137, 0.103, 0.0824, 0.0549, 0.0347\\ a_{21} = 0, 0.177, 0.134, 0.106, 0.0718, 0.0442\\ a_{22} = 0.588\end{cases}$$

We evaluated the performances of the BSS algorithms using the ESR (Error to Signal Ratio) defined by

$$\mathrm{ESR}_i = 10\log_{10}\frac{\sum_t e_i(t)^2}{\sum_t s_i(t)^2}$$

$$e_i(t) = \hat{s}_i(t) - s_i(t) \quad (46)$$

where index i denotes the order of component.

4.2 Results by Two Methods

Fig. 4 shows the results by the InfoMax method, and Fig. 5 does the results by the method based on zeroing of the cross-correlations.

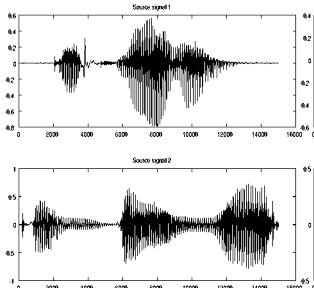

Fig. 2. Source signals

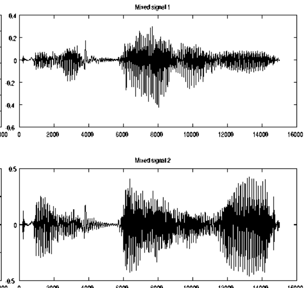

Fig. 3. Mixed signals

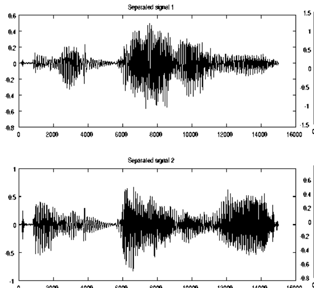

Fig. 4. InfoMax method

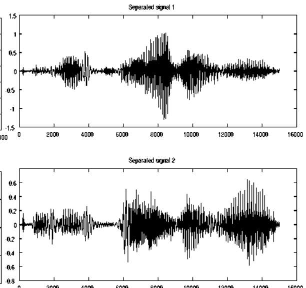

Fig. 5. Method for zeroing of cross-correlations

As compared with the instantaneous mixture case, the ESR for the convolutive mixture case deteriorated only by 2[dB] for the InfoMax method, while it deteriorated seriously for the cross-correlation method. Next, we considered the measurement noise case:

$$\mathsf{y}(t) = A(t) * \mathsf{s}(t) + n(t) \tag{47}$$

where $n(t)$ denotes the white Gaussian with zero-mean. Then, the ESR when applying the wavelet-based approach deteriorated very little in low NSR (Noise to Signal Ratio). This is due to denoising by a set of band-pass filters based on Gabor wavelet (Takada, et al., 2004). In addition, the ESR when applying the FT with do-nothing window instead of the WT deteriorated seriously even in the noise-free case.

Table 1. ESR comparison of two methods

	InfoMax	Cross-Correlations
ESR_1[dB]	-10.90	-8.95
ESR_2[dB]	-11.20	-7.34

5. CONCLUSIONS

This paper considered an application of the WT to the BSS problem for the convolutive mixture case. Generally speaking, the advantage of the WT is the resolution in the high-frequency range. However, the proposed algorithm may not necessarily yields remarkable results in practice.

This is because that

- There is a small number of sound components in the high-frequency range. Power of sound signals is generally concentrated in the low-frequency range, while it is not in the high-frequency range. Since the BSS algorithm is performed at each frequency, it might be difficult to separate the signals with small components existing in the high-frequency range.
- There might be a local-minimum problem due to the simplest optimization technique as in eq.(14).

When using the WT, we can set the frequency ranges according to the scale a. Therefore, we can apply the algorithm to a wide class of objects with every frequency characteristic. If not knowing any frequency characteristics of signals to be separated in advance, we can use the WT only for prefiltering as in 3.2. After examining the signals by the WT and obtaining the frequency information on the signals, we could perform the BSS algorithm only for a specified frequency band.

There still remain the following subjects:

- Hesse and James (2004) have recently proposed a method of BSS based on the discrete wavelet transform (DWT) and wavelet packet (WPT). We should compare our method with the above method from the viewpoint of separation accuracy.
- In connection with the above, the continuous wavelet transform (CWT) using the Gabor function is computationally intensive rahter than the DWT. Thus, the pseudo-CWT approach could be efficient for a compact wavelet representation of signals (Shinohara, 2004).

REFERENCES

Bell, A. and T. Sejnowski (1995). An Information Maximization Approach to Blind Separation and Blind Deconvolution, Neural Comput., **7**, 1129-1159.

Cardoso, J. F. and A. Souloumiac (1996). Jacobi Angles for Simultaneous Diagonalization, SIAM J. Mat. Anal. Appl., **17**, 1, 161-164.

Cardoso, J. F. (1997). Infomax and Maximum Likelifood for Blind Source Separation, IEEE Lectures on Signal Processing, **4**, 4, 112-114.

Mallat, S. (1998). *A Wavelet Tour of Signal Processing*, chapter.1, Academic Press.

Hesse, C. W., C. J. James (2004). An Efficient Time-Frequency Approah to Blind Source Separation Based on Wavelets, 5th Int. Conf. on ICA and BSS - ICA2004, 22-24.

Molgedey, L. and H. G. Schuster (1994). Separation of a Mixture of Independent Signals Using Time Delayed Correlations, Phs. Rev. Lett., **72**, 23, 3634-3637.

Murata, N. and S. Ikeda (1998). An Approach to Blind Source Separation Based on Temporal Structure of Speech Signals, Technical Report, RIKEN BSI.

Parra, L. and C. Spence (2000). Convolutive Blind Separation of Non-Stationary Sources, IEEE Trans. on Speach and Audio Processing, **8**, 3, 320-327.

Shinohara, W. (2002). A Study on Abnormality Detection eith Pseudo-CWT Approach, SICE Annual Conference 2004, 2541-2543.

Tabaru, T., K. Nakano, S. Shin and T. Matsuo (2005). Plant Data Analysis by Gabor Wavelet, SICE Annual Conference 2005 (to appear).

Takada, K., M. Maeda, H. Watai and K. Nakano (2004). On Wavelet-based Convolutive Blind Separation of Non-stationary Sound Sources, SICE Annual Conference 2004, 1362-1367.