Fetal-Maternal ECG Signal Separation from Two Channels based on the Continuous Wavelet Transform and the JADE Algorithm

Said Ziani, Atman Jbari, Larbi Belarbi

Abstract— this paper presents a new approach based on the continuous wavelet transform CWT and JADE algorithm for the blind source separation. The JADE algorithm has been widely used to separate the fECG and mECG signals from 8 recordings or channels, while not all of this number is needed. The present work will show that it isn't the number of channels that counts but rather the quality of the channel used in terms of the energy or the information that it carries. That's why before introducing the 8 channels to the JADE algorithm we will make a selection by the mathematical microscope Continuous Wavelet Transform CWT and we will show that the number of channels will be reduced to 5 or 3 or even to 2 channels. This algorithm has been validated on several real data.

Keywords*—*blind source separation BSS. continuous* wavelet transform. Independent Component Analysis ICA. *JADE.

I. INTRODUCTION

Fetal electrocardiogram signal, non -invasively taken from the abdominal of a pregnant woman is an efficient diagnostic tool for evaluating the health status of fetus. Heart defects are being the most common birth defects and the leading cause of sudden prenatal death [**2**].The cardiac defect may be very slight so that the baby appears healthy and normal for many years after birth, but suddenly becomes so severe due to that its life is in immediate danger. Since heart defects originate in the early weeks of pregnancy when the heart is forming [**3**], the regular monitoring of the fetal heart and the early detection of cardiac abnormalities may help obstetrics and pediatric cardiologist to prescribe proper medications in time, or to consider the necessary precautions during delivery. The fECG can be measured by placing electrodes on the mother's abdomen. However, this signal has very low power and is mixed with several sources of noise and interference. These include fetal brain activity, and power line interference. Moreover, its variability is increased by factors related to gestational age, position of the electrodes, skin impedance, etc. Nevertheless, the main contamination is the maternal ECG [**3**], since its amplitude is much higher. The basic problem is to extract the fECG from the mixture of mECG and fECG signals, where the interfering mECG is a much stronger signal. Since 1960, many different methods [9, 10, 11, 12, 13] have been developed for detecting the fetal ECG. Most of these methods focus on multi-channel mixtures of signals [12]. One direct method subtracts a thoracic maternal ECG from the abdominal composite ECG. A more recent approach [13] employs Independent Component Analysis (ICA), which extracts the sources from their mixtures by assuming the sources are statistically independent. Relatively few works address the problem separating ECG signals recorded on a single-channel. Kanjilal et. al. [4,9,10,14] developed a method for single-channel signals by first detecting both the maternal and fetal heart beats. Next the signal is "cut" into pieces. These pieces are aligned (to form a matrix) and SVD is then performed to obtain the ECG complex. In this paper we will first give a detailed overview of the theory of independent component analysis ICA, JADE algorithm and the continuous wavelet transform CWT, then we will present our new method and finally a simulation Matlab for the validation of the methods.

The authors are with the Research laboratory of Electrical Engineering, Electronic Systems Sensors and Nanobiotechnology, ENSET Mohammed V University in Rabat, Marocco

II. PRELIMINARY RESULTS

A-*Blind Sources Separation*

1-Description of the problem

BSS **[1, 8, 9, 10, 29],** consist of recovering unobserved sources from several observed mixture. Typically the observations are obtained at the output of a set of sensors, where each sensor receives a different combination of the sources signals. The adjective "blind" stresses the fact that:
-The sources signal are not observed

- No information is available about the mixture

This is a sound approach when modeling the transfer from the sources to the sensors is too difficult; it is unavoidable when no prior information is available about the transfer. The lack of prior knowledge about the mixture is compensated by a statistically strong but often physically plausible assumption of independence between the source signals. The simplest BSS model assumes the existence of **n** independent signals $s_1(t),..,s_n(t)$ and the observation of as many mixtures $x_1(t),..,x_m$(t), these mixtures being linear and instantaneous, i.e., $x_i = \sum_{j=1}^{m} a_{ij}\, s_j$ for each i=1……… n

This is represented compactly by the mixing equation

$$\boldsymbol{x} = \boldsymbol{As} + \mathbf{n} \qquad (1)$$

Where:

-**x** = $[x_1,..x_m]^T$ is an m×1 column vector collecting the observed signals.

- **s** = $[s_1,..s_n]^T$ is an n×1 column vector collecting the source signals.

- **n** = $[\mathrm{n}_1,..n_m]^T$ is an m×1 column vector collecting the noise signals.

-**A** is the square n×n mixing matrix

It is assumed that:

-m ≥ n

-rang (A) = n

– **s** is the realization of a centered vector random sequence of non-zero components.

- **n** is the realization of a centered vector random sequence of covariance $\boldsymbol{\sigma}^2\, \boldsymbol{I_n}$ independent of **s**

2-Fundamental assumptions:

↘ The sources are sequences **i**dentically and **i**ndependently **d**istributed **iid**

$\forall$ i ϵ $[1, n]$, $\quad p(s_i) = \prod_{t=0}^{T-1} p(s_i[t])$ and

$\forall$ (t, t'),

$$p(s_i[t]) = p(s_i[t'])$$

↘ Sources are mutually independent non-Gaussian stochastic processes with zero mean.

$$p(s_i) = \prod_{i=1}^{n} p(s_i)$$

↘ One source at most is Gaussian

3-Fundamental theorem

Let **x** be a vector with independent components**[1,8,9,10,29]** , of which at most one is Gaussian, and whose densities are not reduced to a point-like mass. Let **B** be an orthogonal **n×n** matrix and y the vector **y=Bx**. Then the following three properties are equivalent:

i) The components $\boldsymbol{y_i}$ are pairwise independent.

ii) The components $\boldsymbol{y_i}$ are mutually independent.

*iii) **B=DP**, **D** diagonal, **P** permutation.*

In short,

We will therefore seek the separation matrix **B** such that the components of **y** are the most independent possible, in the sense of a certain measure of independence between the components of y, and the problem of the BSS becomes then that of the Independents Components Analysis **ICA**. We will assume **n** = **0** and **m=n**

The problem can be summarized in the following figure:

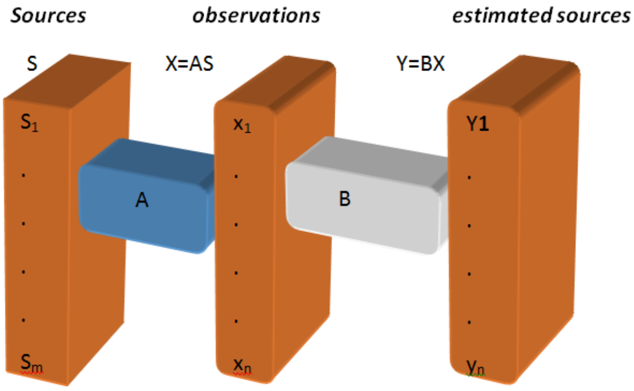


Fig.1. Mixing and separating sources

B-Independent Component Analysis ICA

ICA is a method of Blind Source Separation (BSS) [4].The assumption underlying ICA is that each row of the data matrix is weighted sum of pure source signals, the weights being proportional to the contribution of the corresponding pure signals to that particular mixture. The original source signals and their proportions in the analyzed mixtures are unknown. In ICA, **x** is not seen as a collection of points in a multidimensional space, but rather as a collection of signals with a certain number of common sources. ICA aims to extract these pure sources, underlying the observed signals, as well as their concentration in each mixture. ICA finds the independent component by maximizing the statistical independence of the estimated components. The two broadest definitions of independence for ICA are minimization of mutual information and maximization of non-Gaussianity. Typical algorithms for ICA use the centering and whitening as preprocessing steps in order to simplify and reduce the complexity of the problem.

The theory of ICA is in several steps:

1-Contrast function

A contrast function $\mathbf{C}\{\boldsymbol{v}\}$ is a function of probability density of random sequence $\boldsymbol{v} \in \mathbb{R}^{\boldsymbol{n}\times\mathbf{T}}$

$$\begin{cases} \forall \mathrm{M}\in\mathbb{R}^{\boldsymbol{n}\times\mathbf{n}}, \quad \mathbf{C}\{\boldsymbol{Ms}\} \geq \mathbf{C}\{\boldsymbol{s}\}, \\ \mathbf{C}\{\boldsymbol{Ms}\} = \mathbf{C}\{\boldsymbol{s}\} \quad \textit{if and only if} \quad \boldsymbol{M \sim I_n} \end{cases} \qquad \mathbf{(2)}$$

If **B** is a matrix minimizes $\mathbf{C}\{\boldsymbol{Bx}\}$= $\mathbf{C}\{\boldsymbol{BAs}\}$

So $BA \sim I_n$

And consequently **B=A**$^{-1}$
Let us denote that **y=Bx**
It is important to remember that $C\{y\}$ is a function of probability density following the assumptions in (A-2)

2- Orthonormal contrast functions

To the extent that the exact independence can not be recovered the mutual decorrelation of the components of **y** is desirable, so it must explicitly be taken into account in the minimization of the contrast function. We shall therefore speak of minimization of $C\{y\}$ under the whiteness constraint. This could correspond to two options:

↘Or a prior spatial whitening of the observations has been carried out so that the estimate of **B** has been reduced to the estimate of an orthonormal separation matrix **V** such that:

$$C\{y\} = C\{Vz\}$$

↘Or the minimization of a contrast $C\{y\} = C\{Bx\}$ is carried out under the constraint**:** $R_{yy} = I_n$
Where R_{yy} is a Correlation matrix

3- Maximum likelihood ML approach

Either q any hypothetical distribution of any sample s [.] of the sources. The log-likelihood is thus written:

$$L(B,q) = log|\det(B)| + \frac{1}{T}\sum_{t=0}^{T-1} \log q(Bx[t]) \quad (3)$$

Asymptotically if $T \to +\infty$ the criterion of ML Can be formulated as a KULLBACK DIVERGENCE between the two probability densities functions, $p_v(u)$ and $p_w(u)$, such that:

$$K\{v\|w\} \stackrel{\text{def}}{=} \int p_v(u) ln \frac{p_v(u)}{p_w(u)}\, du \quad (4)$$

An important property of the Kullback divergence is that it is zero if and only if v and w have the same distribution.
It is thus possible to associate the ML with the following contrast:

$$C_{ML}\{y\} = K\{Bx\|s'\} \quad (5)$$

In this respect, the ml estimator offers a new interpretation:
It consists in finding the separation matrix B which minimizes the kullback divergence between the density of **y=Bx** and the density of the sources.

4- ICA in practice :Approximation of the kulback divergence by cumulant

Let a, b, c and d be four real centered random variables. The fourth-order cumulant of these random variables is defined by:

$$c_{abcd} \stackrel{\text{def}}{=} E\{abcd\} - E\{ab\}E\{cd\} - E\{ac\}E\{bd\} - E\{ad\}E\{bc\} \quad (6)$$

If the random variables **a, b, c and d** can be separated into two sets of mutually independent random variables then the four-order cumulant is canceled **[1,8,9,10,29].**
The fourth-order cumulant of a real centered random variable is given by:

$$k_a \stackrel{\text{def}}{=} c_{aaaa} \quad (7)$$

k_a is called Kurtosis.

By using the Edgeworth expansion the Kulback divergence of two random vectors v and w is given by:

$$K\{v\|w\} \approx \frac{1}{4}\sum_{ij}(r_{v_iv_j} - r_{w_iw_j})^2 + \frac{1}{48}\sum_{ijkl}(c_{v_iv_jv_kv_l} - c_{w_iw_jw_kw_l})^2 \quad (8)$$

Where $v = [v_1, \dots, v_n]^T$ and $w = [w_1, \dots, w_n]^T$

Having this approximation it is possible to approximate the contrast function defined by **(5)**
So**:**

$$C_{ML}\{y\} \approx C_{24}\{y\} \quad (9)$$

Minimize $C_{24}\{y\}$ under constrained whiteness allows to write:

$$C^\circ_{ML}\{y\} \approx C^\circ_4\{y\} \quad (10)$$

With:

$$C^\circ_4\{y\} \stackrel{\text{def}}{=} \frac{1}{48}\sum_{ijkl}(c_{y_iy_jy_ky_l} - k_{s'_i}\delta_{ijkl})^2 \quad (11)$$

It can still be shown that:

$$C^\circ_{kurt}\{y\} \stackrel{\text{def}}{=} \frac{1}{48}\sum_{ijkl \neq iiii} {c_{y_iy_jy_ky_l}}^2 \quad (12)$$

And if $\boldsymbol{E\{yy^T\} = I_n}$

We will have

$$\boldsymbol{C^\circ}_{\boldsymbol{kurt}}\{\boldsymbol{y}\} = -\frac{\mathbf{1}}{\mathbf{48}}\sum_{i=1}^{n} k_{y_i}^2$$

(13)

Since the kurtosis of a Gaussian random variable is zero, maximizing the sum of the kurtosis to the square of y returns once again to exclude as much as possible the densities of the components of **y** from a Gaussian density. The maximization of the sum of kurtosis squared amounts to canceling the crossed cumulants of **y.** The **JADE** method that we will now propose proposes to minimize only certain cumulant slices, this leading to an efficient minimization procedure by simultaneous diagonalization.

C-JADE algorithm

The jade algorithm **[1,8]** designed by Cardoso & Souloumiac consists in minimizing the following contrast function:

$$\boldsymbol{C^\circ}_{\boldsymbol{JADE}}\{\boldsymbol{y}\} \stackrel{\text{def}}{=} \sum_{ijk} {c_{y_i y_j y_k y_l}}^2$$

(14)

And this minimization will be reduced to the simultaneous diagonalization of a matrix set of cumulants.

Definition 1

Let v= $[v_1 , \ldots , v_n]^T$ be a random vector of $\mathbb{R}^{n\times 1}$ and M a matrix $\epsilon\mathbb{R}^{n\times n}$. We define the matrix of cumulants by:

$$\forall (i, j) \in [\![1, n]\!]^2, \quad |\boldsymbol{Q_v(M)}|_{ij} = \sum_{kl} c_{v_i v_j v_k v_l}\, m_{kl}$$

(15)

The matrices of cumulants thus defined will allow us to reduce some slices of the tensor structure of cumulants of order 4 to a simpler matrix structure.

Definition 2

Let M a matrix $\epsilon\mathbb{R}^{n\times n}$, We define the diagonal measure by:

$$\text{Off (M)} = \sum_{i\neq j} {m_{ij}}^2$$

(16)

Result:

Let $\mathcal{N}$ be a base of the matrix set $\epsilon\mathbb{R}^{n\times n}$

Then:

$$\boldsymbol{C^\circ}_{\boldsymbol{JADE}}\{\boldsymbol{Vz}\} \stackrel{\text{def}}{=} \sum_{M\in\mathbb{N}} \boldsymbol{off(V^T Q_v(M) V)}$$

(17)

It can thus be seen that the minimization of the contrast function $\boldsymbol{C^\circ}_{\boldsymbol{JADE}}\{\boldsymbol{Vz}\}$ with respect to the orthonormal matrix **V** can be achieved by the simltaneous diagonalization of the matrix assembly

$$\{\boldsymbol{Q_z(M_i)}, \boldsymbol{M_i}\in\mathbb{N}\}$$

Remark

The simultaneous diagonalization algorithm used in JADE is inspired by the JACOBI method for the calculation of the eigenvalues of a symmetric real matrix. The algorithm proposed by Cardoso and Souloumiac makes it possible to carry out an approximate simultaneous diagonalization because in practice the cumulant matrices are not rigorously diagonalisable simultaneously

III. THE CONTINUOUS WAVELETS TRANSFORM: CWT

A –Introduction

This part covers the basic theory of the continuous wavelet transform. We will first determine what constitutes a wavelet, how it is used in the transformation of a signal and what is can tell us about the signal. We will look at the energy-preserving features of the wavelet transform and how it may be used to produce scalogram.

B –The wavelet

The wavelet transform is a method of converting a signal into another form which either makes certain features of the original signal more amenable to study or enables the original data set to e describe more succinctly. To perform a wavelet transform we need a wavelet which is localized waveform. In fact, a wavelet is a function $\varphi(t)$ which satisfies certain mathematical criteria. A selection of wavelets commonly used in practice is shown in figure (2).

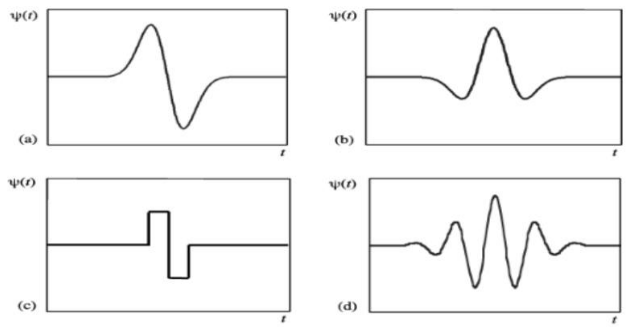


Fig.2. *(a) Gaussian wave. (b)Mexican hat. (c) Haar. (d) Morlet* [17, 18]

C –Requirements for the wavelet

In order to be classified as a wavelet, a function must satisfy certain mathematical criteria. These are:

1.Wavelet must have finite energy:

$$E = \int_{-\infty}^{+\infty} |\psi(t)|^{2} \; dt < \infty$$
(18)

Where E is the energy of a function equal to the integral of its Squared magnitude and the vertical brackets | | represent the modulus operator which gives the magnitude of ψ(t). If (t) is a complex function the magnitude must be found using both its real and its complex parts.

2.If $\widehat{\psi(f)}$ is the Fourier transform of ψ(t),i.e.

$$\widehat{\Psi(\mathrm{f})} = \int_{-\infty}^{+\infty} \psi(t) e^{-i2\pi f t} \, dt$$
(19)

Then the following condition must hold:

$$\mathrm{C_g} = \int_{0}^{+\infty} \frac{|\widehat{\Psi(\mathrm{f})}|^2}{f} \, df$$
(20)

This implies that the wavelet has no zero frequency components, $\widehat{\psi(0)} = 0$ or, to put it another way, the wavelet (t) must have a zero mean. Equation (3) is known as the *admissibility condition* and $\mathrm{C_g}$ is called the *admissibility constant. The value of* $\mathrm{C_g}$ depends on the chosen wavelet and is equal to π for the Mexican hat wavelet.

C-The energy spectrum of the wavelet

Wavelets satisfying the admissibility condition are in fact *bandpass filters.* This means in simple terms that they let through only those signals components within a finite range of frequencies (the pass band) and in proportions characterized by the energy spectrum of the wavelet. A plot of the squared magnitude of the Fourier transform against frequency for the wavelet gives its *energy spectrum.*

D-The wavelet transforms

1-Preliminary

Now we have chosen a mother wavelet, how do we put it to good use in a signal analysis capacity? First we require our wavelet to be more flexible than that defined earlier, i.e. (t). We can perform two basic manipulations to make or wavelet more flexible: we can stretch and squeeze it (dilation) or we can move it (translation). Figure III-2(a) shows the Mexican wavelet stretched and squeezed to respectively double and half its original width on the time axis. This dilation and contraction of the wavelet is governed by the dilation parameter a which, for the Mexican hat wavelet, is the distance between the centre of the wavelet and its crossing of the time axis. The movement of the wavelet is governed by the translation parameter b. Figure III-2(b) shows the movement of a wavelet along the time axis from b_1 via b_2 to b_3.These shifted and dilated versions of the mother wavelet are denoted $\psi(\frac{t-b}{a})$

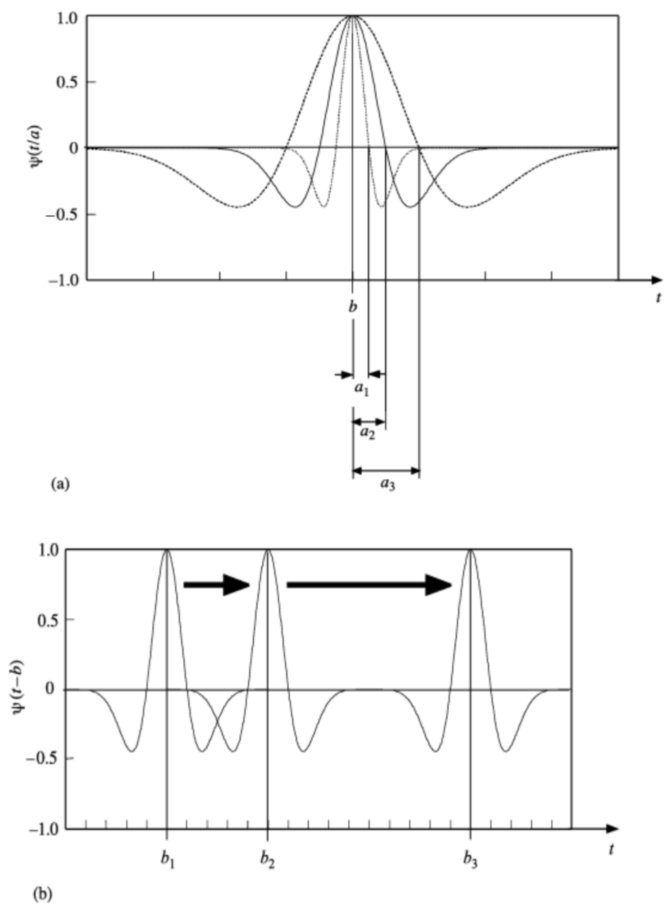


Fig.3 .Dilation (a) and(b) translation of a wavelet **[17, 18]**

2-Definition

The wavelet transform of a continuous signal with respect to the wavelet function is defined as

$$\mathrm{T\ (a,\ b) = w\ (a)} \int_{-\infty}^{+\infty} x(t) \psi^{*}(\frac{t-b}{a}) \, dt \qquad (21)$$

Where w (a) is a weighting function. The asterisk indicates that the complex conjugate of the wavelet function is used in the transform. The wavelet transform can be thought of as the cross-correlation of a signal with a set of wavelets of various 'width'. Typically w (a) is set to $\frac{1}{\sqrt{a}}$ for reasons of energy conservation. Thus the wavelet transform is written

$$T(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} x(t)\psi^*(\frac{t-b}{a})\, dt \qquad (22)$$

This is the continuous wavelet transform CWT. Take a closer Look at this equation. It contains both dilated and translated wavelet ($\frac{t-b}{a}$) and the x(t) signal, where x(t) could be a beating heart. In mathematical terms this is called a convolution. The normalized wavelet is often written more compactly as

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

We can express the wavelet transform in even more compact form as an inner product:

$$T(a, b) = < x, \psi_{a,b} > \qquad (24)$$

3-The inverse wavelet transforms

As with its Fourier counterpart, there is an *inverse wavelet transform, defined as*

$$x(t) = \frac{1}{Cg} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} T(a, b)\psi_{a,b}(t) \frac{da db}{a^2}$$
(25)

This allows the original signal to be recovered from its wavelet transform by integrating over all scales and location, a and b. Note that for the inverse transform, the original wavelet function is used, rather than is conjugate which is used in the forward transformation. If we limit the integration over a range of a scale rather than all a scales, we can perform a basic filtering of the original signal.

4-The signal energy: wavelet-based energy

The total energy contained in a signal, x(t), is defined as its integrated squared magnitude

$$E = \int_{-\infty}^{+\infty} |x(t)|^2 \, dt = \|x(t)\|^2$$
(26)

For this equation to be useful the signal must contain finite energy. The relative contribution of the signal energy contained at a specific a scale and b location is given by the two-dimensional wavelet energy density function:

$$E(a, b) = |T(a, b)|^2 \qquad (27)$$

A plot of $E(a, b)$ is known as a ***scalogram.*** In practice, all functions which differ from $|T(a, b)|^2$ by only a constant multiplicative factor are also called scalogram. The scalogram can be integrated across a and b to recover the total energy in

the signal using the admissibility constant, C_g, as follows:

$$E = \frac{1}{Cg} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} |T(a, b)|^2 \psi_{a,b}(t) \frac{da db}{a^2}$$

(28)

The relative contribution to the total energy contained within the signal at a specific a scale is given by the scale dependent energy distribution:

$$E = \frac{1}{Cg} \int_{-\infty}^{+\infty} |T(a, b)|^2 db$$

(29)

Peaks in E (a) highlight dominant energetic scales within the signal.

5-The wavelet transform in terms of the Fourier transform

As we saw in equation (7), the wavelet transform is the convolution of the signal with wavelet function. Hence we can employ the convolution theorem to express the wavelet transform in terms of products of the Fourier transforms of the signal, $\widehat{x(f)}$, and wavelet, $\widehat{\psi_{a,b}}(f)$, as follows:

$$T(a, b) = \int_{-\infty}^{+\infty} \widehat{x(f)} \widehat{\psi_{a,b}}(f)^* \, df$$

(30)

Where we note that the conjugate of the wavelet transform is used. The Fourier transform of the dilated and translated wavelet is

$$\widehat{\psi_{a,b}}(f) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} \psi(\frac{t-b}{a})\, e^{-2\pi i f t}\, dt$$

(31)

The equivalence between the time convolution and Fourier integrals for determining T (a, b) is depicted in figure III-3. This is a particularly useful result when using discretized approximations of the continuous wavelet transform in practice with large large signal data sets, as the fast Fourier transform (FFT) algorithm may be employed to facilitate rapid calculation of the wavelet transform and its inverse. In addition, the Fourier transform of the wavelet function, $\widehat{\psi_{a,b}}(f)$, is usually known in analytic form and hence need not be computed using an FFT.

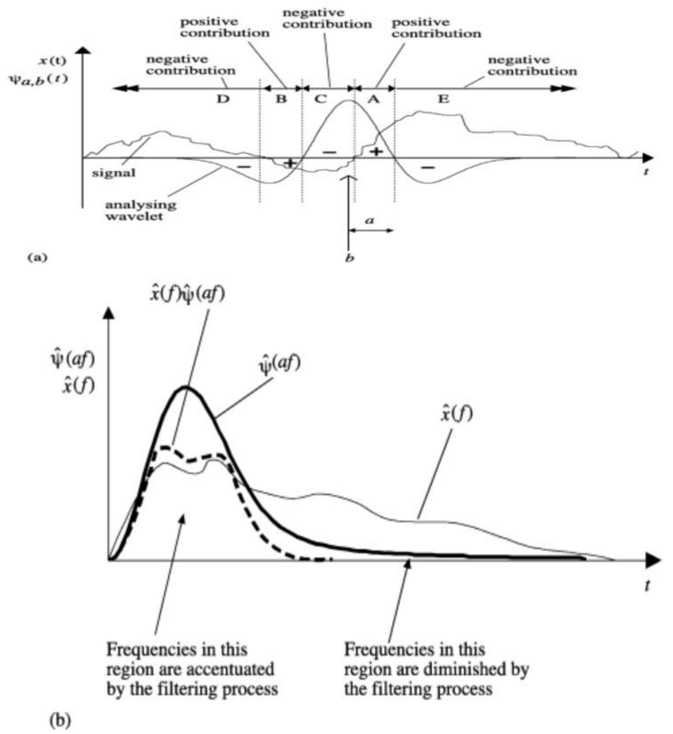


Fig.4. (a and b) Schematic representation of the wavelet transform in its time and frequency representations**[17, 18].**

6-Heisenberg boxes

This mechanism represents the richness of time-scale analysis. It allows a small scale to carry out a precise temporal examination of the signal corresponding to the high frequencies. On the contrary, the low frequencies are examined with a very high frequency resolution:

$$\Delta t \Delta f \geq 1/4\pi \qquad \text{(III-15)}$$

This is illustrated in figure III-4:

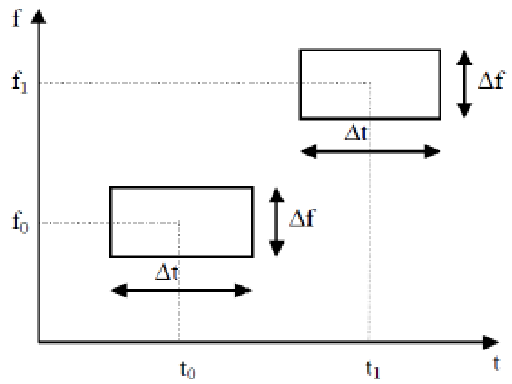
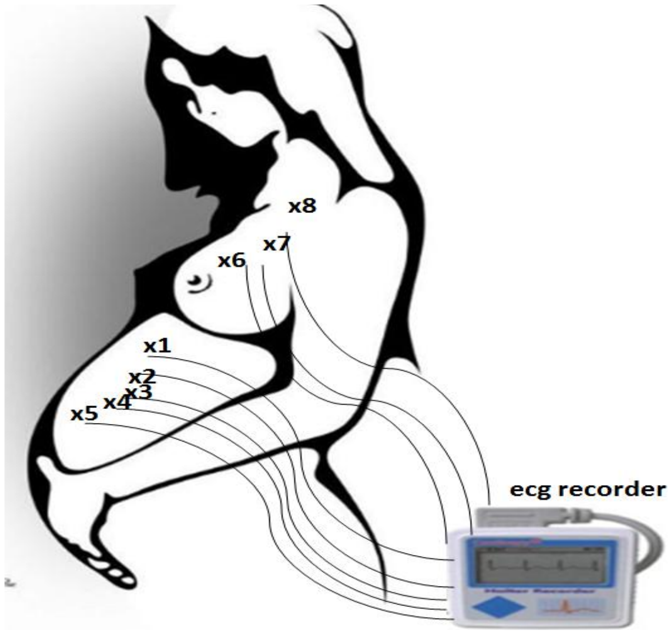


Fig.5. Heisenberg boxes in the time –frequency plane

IV. METHODS AND RESULTS

A-ECG recordings



The data is extracted from the DaISy database (Database for the Identification of Systems). And corresponding to ECG records collected simultaneously by eight electrodes Figure IV-2. The sampling frequency is 250 Hz. Duration T=10s
We used the MATLAB 2015a on Windows 7.

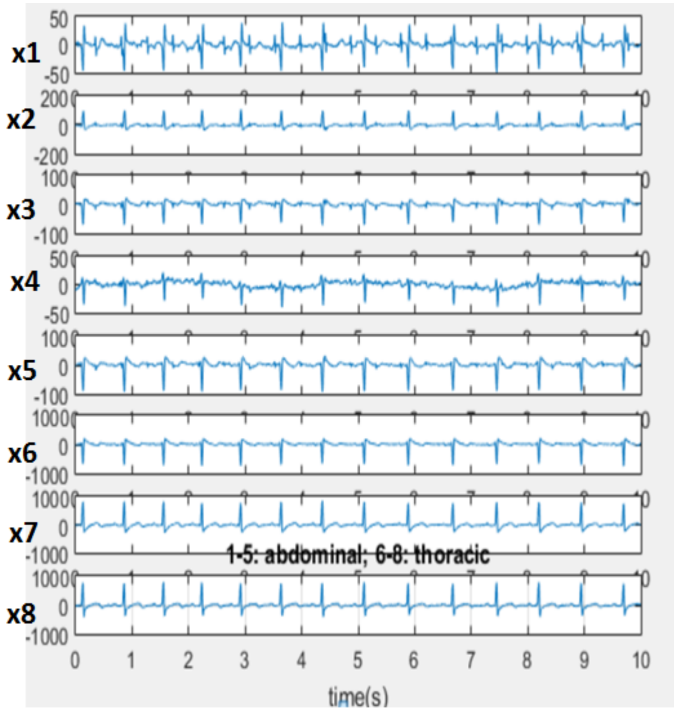


Fig.6. 8 Channel set of cutaneous data recordings.

B-Method

1-Algorithm

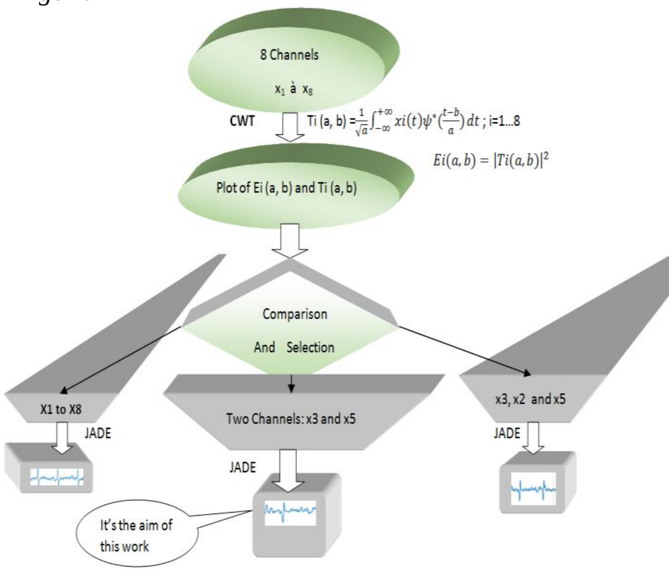
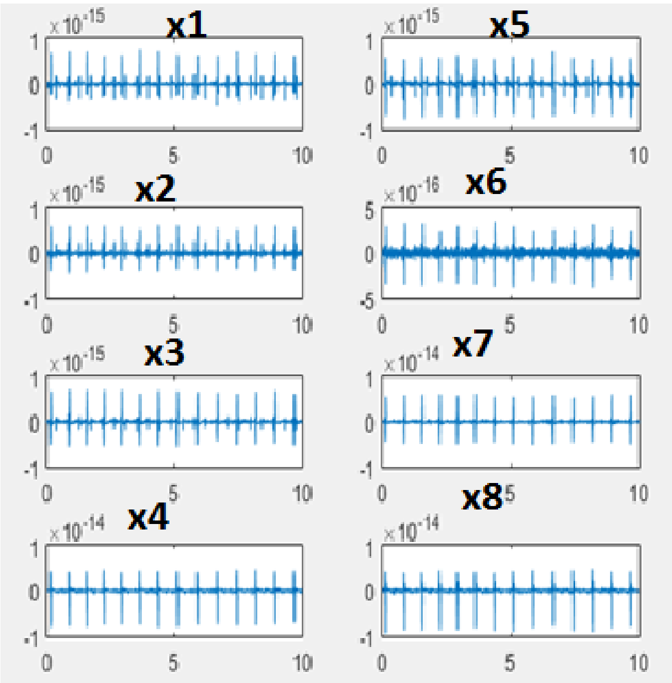


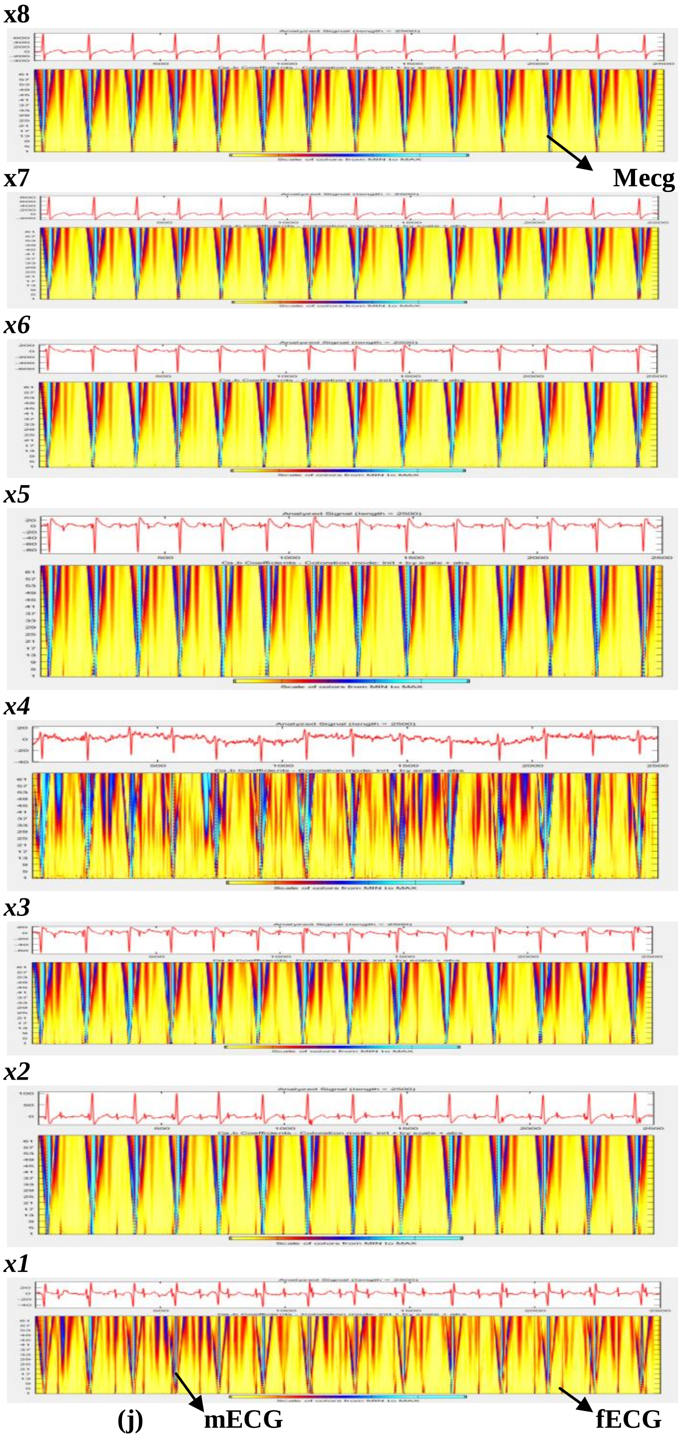
Fig.8. Algorithm

2-Continuous Wavelet Transform CWT

In order to make a good selection of the useful signals. Our approach consists to compute at a first time the continuous wavelet transform CWT of x1 to x8 signals and in second time to plot Ti (a, b) as a function of time and Ei (a,b) in the time-scale space. Finally we will calculate the mean energy for a low scale a, in order to be able to make the comparison and the selection of the useful signals for JADE algorithm **Fig (4).**Simulations are made by the Morley wavelet. The results found are:



(i)



(k)

Signal	X1	X2	X3	X4	X5
energy	E1=1.13 10^{-32}	E2=2.48 10^{-33}	E3=5.51 10^{-33}	E4=2.99 10^{-33}	E5=8.87 10^{-33}

Fig.9.
(i) Representation of the temporal evolution of Ti (a, b)
(j) Plotting of Ei (a, b) in the time-scale space (Scalogram)
(k) Mean energy for the scale a=0.1

C- Results:

1- *Results 1:*

We see that the average energy of x1 is large, but unfortunately it is richer in information on the cardiac activity of the mother than that of the fetus so it is not a good candidate for JADE. Because it mustn't be forgotten that the real problem is the fetal electrocardiogram extraction fECG.

2- Results 2:

The electrical activity of the fetus is absolutely absent in the signals x6 to x8: Are useless for JADE.

3- *Results 3:*

The energy classification of the four remaining signals

$E(x5) \geq E(x3) \geq E(x2) \geq E(x4)$

The first finding is that it is possible to reduce the number of channels from 8 to 4.

4- *Results 4:*

The two channels 2 and 4 are a comparable energy, therefore only one of the two is considered in the simulation.

5- *Results 5:*

There are three possibilities (cases):

a-Four channels: x2 to x5

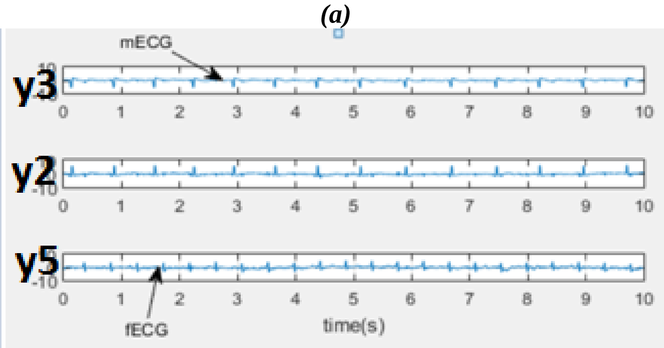
b-Three channels: x3, x2 and x5;

c-Two channels x3 and x5;

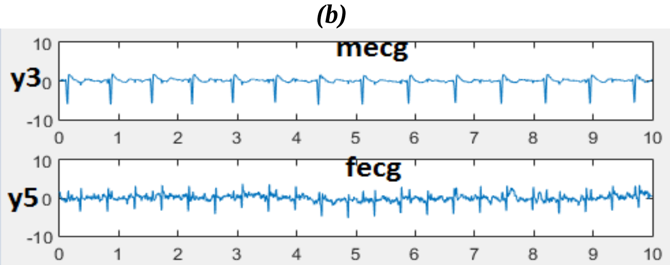
6- Results 6:

The application of the JADE function under Matlab in the three cases gives the following results:

(a)



(b)



(c)

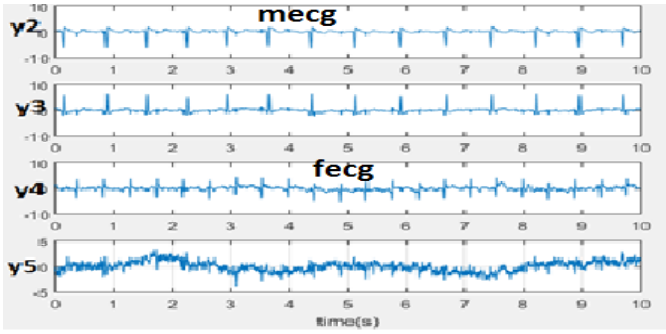


Fig.10.*Sources estimated via JADE Algorithm:*
(a) Tree channels: x3, x2 and x5
(b) Two channels: x3 and x5
(c)Four channels x2 to x5

V. DISCUSSION AND CONCLUSION

A-Discussion

The blind source separation was bieb successful in the three cases. More, the case of the ***two channels*** gave a surprising result because the fECG signal is extracted with well distinguished complex QRS.

In order to measure the quality of separation we are computing the fourth-order cumulant ***table1*** defined by:

$$CUM = C_{yiyiyjyk}$$

Cases	Two channels	Three channels	Four channels
CUM	0.0019	0.0025	0.0063

Table1: Values of fourth order cumulant

** These values indicate a very great independence between the extracted components and therefore a high quality of separation.

B- Conclusion:

Indeed the use of JADE as a blind source separation algorithm is not a new approach, but what is new in this article is the fact of combining the JADE algorithm and the continuous wavelets transform CWT in order to reduce the number of channels from eight to two .More, this combination led us to three very important and originals results:

1- JADE is not limited and doesn't require multiple recordings as reported in several articles [12].
2- The energy criterion is more efficient than the statistical criterion of independence.
3- A careful modification of JADE in the formation of cumulant matrices can lead to important results and even to solve the case of a single channel.

REFERENCES

[1] C.Févotte.Approche temps- fréquence pour la séparation de sources non-stationnaire.Phd thesis ,université de Nantes , octobre 2003..

[2] Congenital Heart Defects. March of Dimes. 2005.[online].Available: http://www.marchofdimes.com/professionals/14332_1212.asp.

[3] M.E.Pierpont,C.T.Basson,B.D.Gelb,T.M.Gilglia,E.Goldmuntz,G.McGee ,C.A.Sable,D.Srivastava,and C.L. Webb, ‘ Genetic basis for congenital heart defects: Current knowledge a scientific statement from the americain heart association congenital cardiac defects committee,council on cardiovascular disease in the young:Endorsed by the americain academy of pediatrics,’Circulation, vol. 115, no. 23, pp. 3015-3038, 2007.

[4] Janjarasjitt, S. A new QRS detection and ECG signal extraction technique for fetal monitoring.PhD thesis, Case Western Reserve University. (2006).

[5] J.Pan and W.L.Tompkins,’’Areal time QRS Detection Algorithm,’’in IEEE Trans.Biomed.Eng.,vol 32,1985,pp.230-236.

[6] Mallat S. Wavelet Tour of Signal Processing, 2nd ed.San Diego: Academic Press.1998:162-219,249-254.

[7] M. Niknazar,B. Rivet, and C. Jutten, and G.D. Clifford, ‘’A Nonlinear Bayesian Filtering Framework for ECG Denoising,’’ Biomed.

[8] Cardoso JF. Infomax and maximum likelihood for blind source separation. IEEE Signal Proc Lett 1997;4(4):112–4.

[9] Cardoso JF. Blind signal separation: statistical principles. Proc IEEE 1998;9(10):2009–25.

[10] Cardoso JF, Souloumiac A. Jacobi angles for simultaneous diagonalization. SIAM J Matrix Anal Appl 1996;17(1):161–4

[11] De Lathauwer L, Comon P, De Moor B, Vandewalle J. « ICA algorithms for 3 sources and 2 sensors ». In: HOS 99, IEEE Signal Processing Workshop on Higher-Order Statistics, Caesarea, Israel, June 14– 16. 1999. p. 116–20

[12] Albera L, Ferreol A, Chevalier P, Comon P. « ICAR, a tool for blind source separation using fourth order statistics only » IEEE Transactions On Signal Processing. 2005.

[13] Comon and C. Jutten. Handbook of Blind Source Separation. Independent Component Analysis and Applications. Academic Press, Amsterdam, 2010.

[14] P. Comon, “Independent component analysis, a new concept?,” Signal Processing (Special Issue Higher Order Statistics), vol. 36, no. 3, pp. 287–314, Apr. 1994.

[15] J.-F. Cardoso, C. Jutten, and P. Loubaton, Eds., in Proc. 1st Int. Workshop Independent Component Analysis and Blind Signal Separation (ICA’99), Aussois, France, Jan. 1999.

[16] J. F. Cardoso and A. Souloumiac. Blind beamforming for non gaussian signals. IEE Proceedings-F, 140(6) :362–370, December 1993.

[17] Paul S Adison.Napier University,Edinburgh,UK.The illustrad wavelet transform Hanbook .2002.

[18] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. 1996

[19] M.E.Pierpont,C.T.Basson,B.D.Gelb,T.M.Gilglia,E.Goldmuntz,G.McGee ,C.A.Sable,D.Srivastava,and C.L. Webb, ‘ Genetic basis for congenital heart defects: Current knowledge a scientific statement from the americain heart association congenital cardiac defects committee,council on cardiovascular disease in the young:Endorsed by the americain academy of pediatrics,’Circulation, vol. 115, no. 23, pp. 3015-3038, 2007.

[20] J.Pan and W.L.Tompkins,’’Areal time QRS Detection Algorithm,’’in IEEE Trans.Biomed.Eng.,vol 32,1985.

[21] Mallat S. Wavelet Tour of Signal Processing, 2nd ed.San Diego: Academic Press.1998:.

[22] M. Nguyen. Analyse Multi-dimensionnelle et analyse par ondelettes des signaux sismiques. PhD thesis, Institut National Polytechnique de Grenoble, 1999.

[23] A. Roueff, B. Nicolas, J. Chanussot, and J. Mars. Unsupervised seismic wave separation in the time/scale plane. In 63st Meeting of European Association of Geophysicists and Engineers, Amsterdam, Expended Abstracts, 2001

[24] A. Hyvarinen, E. Oja, "A Fast Fixed-Point’ Algorithm for Independent Component Analysis", *Neural Computation*, vol. 9, no. 7, pp. 1483-1492, 1997.

[25] P.P. Kanjilal, G. Saba, "Fetal ECG Extraction from Single-Channel Maternal ECG using Singular Value Decomposition", *IEEE trans. on biomedical Engineering*, vol. 44, no. 1, Jan 1997.

[26] T. W. Lee, Independent Component Analysis-Theory and Applications, Kluwer, 1998.

[27] V. Zarzoso, J. Millet-Roig, A.K. Nandi, "Fetal ECG Extraction from Maternal Skin Electrodes Using blind Source Separation and Adaptive Noise Cancellation Techniques", *Computers in Cardiology*, pp. 431-434, 2000.

[28] P.comon. independent component analysis a new concept signal processing vol 36 no 3 pp 287-314 april 1994

[29] S.D. Yanowitz, A.M. Bruckstein, A new method for image segmentation, CVGIP 46 (1989) 82}95

[30] W. Niblack, An Introduction to Image Processing, Prentice-Hall, Englewood Cli!s, NJ, 1986.

[31] L. Eikvil, T. Taxt, K. Moen, A fast adaptive method for binarization of document images, International Conference on Document Analysis and Recognition, ICDAR '91, France, 1991,.

[32] J. Parker, Gray level thresholding on badly illuminated images, IEEE Trans. Pattern Anal. Mach. Intell. 13 (8) (1991) .

[33]] A. Perez, R.C. Gonzalez, An iterative thresholding algorithm for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-9 (6) (1987) .

[34]] J. Sauvola, M. PietikaKinen, Page segmentation and classi- "cation using fast feature extraction and connectivity analysis, International Conference on Document Analysis and Recognition, ICDAR '95, Montreal, Canada, 1995,.

[35] s. aboud g barkai s.mashiash d.sadeh quantifica tion of the fetal electrocardiogram using averaging technique comput biod .med 20 pp 147-155,1990.

[36] i. daubechies .ten lectures on wavelets, society for industrial and applied mathematics 1992.

[37] a. parashiv – lonescu, c jutten , k.aminian b. najafi and ph.robert.source separation in strong noisy mixtures : a study of wavelet de noising pre – processing proceed of ICASSP 2002 Orlando (USA) may, 2002.

[38] d.t pham and j.f Cardoso blind separation of instantaneous mixture of nonstationary sources, IEEE trans on sp.49 (9) pp 1837-1848, 2001.

[39] Hui Zhang, Jason E.Fritts, Sally A. Goldman,” Image Segmentation Evaluation: A Survey UnsupervisedMethods”,2008.

[40] Nir Milstein, “Image Segmentation by Adaptive Thresholding”,Spring 1998.

[41] Sang uk lee, seok yoon chung and Rae hong park, “A Comparative Performance Study of Several Global Thresholding Techniques for Segmentation”, Computer Vision Graphics And Image Processing52, 171-190, 1990

[42] Graham Leedham, Chen Yan, Kalyan Takru, Joie Hadi Nata Tan and Li Mian, “Comparison of Some Thresholding Algorithms for Text/Background Segmentation in Difficult Document Images”,Proceedings of the Seventh International Conference on Document Analysis and Recognition , 2003.

[43] Senthilkumaran N1 and Vaithegi s2 image segmentation by using thresholding techniques for medical imges Computer Science & Engineering: An International Journal (CSEIJ), Vol.6, No.1, February 2016.

[44] J. Sauvola and M. Pietikainen, “Adaptive document image binarization,” Pattern Recognition 33(2), pp. 225–236, 2000.

[45] Graham Leedham, Chen Yan, Kalyan Takru, Joie Hadi Nata Tan and Li Mian “Thresholding Algorithms for Text/Background Segmentation in Difficult Document Images” Seventh International Conference on Document Analysis and Recognition (ICDAR 2010).

[46] S. Beucher. Mathematical morphology. Ecoles des Mines, Spring School, Fontainebleau, 1980.

[47]] P. Kumar and E. Foufoula-Georgiou. Wavelet analysis for geophysical application. Reviews of Geophysics, 1997, pp.385-412.

[48] J. I. Qing, R. Xinbo, Y. E. Zhihong, "Worst conducted EMI spectrum of the critical conduction mode boost PFC converter", *IEEE Trans. Power Electron.*, vol. 30, no. 3, pp. 1230-1241, Mar. 2015..

[49] J. S. Calderon-Piedras, A. D. Orjuela-Canon, D. Sanabria-Quiroga, "Blind source separation from single channel audio recording using ICA algorithms", *Proc. 2014 XIX Symp. Image Signal Process. Artif. Vis.*, pp. 1-5, 2014.sameni extraction of fetal cardiac signals from an array of maternal abdominal recording ph.d dissertation sharif university opf thecnology (2008)

[50] hassanpour, a. parsaie fetal ECG extraction using wavelet transform . international conference on waeb technologies and internet commerce (2006) pp 179-179.

[51] l.o.s alvares y.nf oprdonez , a.g salvator j.m roig noninvasive FECG for edtimating the fetal heart rate image signal processing and artificial vision (stsiva) . 2013 .

[52] b.azzerboni F.L foresta n.mammone al a new approach based on wavelket ICA algorithms for fetal electrocardiogram extraction proceeding of European symposium on artificial neural networks (2005).

[53] F.L foresta N. mammone anf f.c morabito , insependent component and wavelet analysis for FECG extraction : the staveform evaluation, IEEE international conference on comuta ional intelligence for measurement systems and applications (2005) .

[54] L.jiang p. cai j yang the source numer estimation based on the beam eigenvalue method, IEEE international conference on industrial electronics and applic tions (2007) , 2727-2731.

[55] A. M. Reddy and B. Raj, "Soft Mask Methods for singlechannel speaker separation," IEEE Transactions on Audio, Speech, and Language Processing, vol. 15, Aug. 2007.

[56] A. M. Reddy and B. Raj, "A Minimum Mean squared error estimator for single channel speaker separation," in International Conference on Spoken Language Processing (Inter-Speech), 2004.

[57] T. Kristjansson, J. Hershey, and H. Attias, "Single microphone source separation using high resolution signal reconstruction," in IEEE International Conference Acoustics, Speech and Signal Processing (ICASSP), 2004.

[58] M. H. Radfar, W. Wong, R. M. Dansereau, and W. Y. Chan, "Scaled factorial Hidden Markov Models: A new technique for compensating gain differences in model-based single channel speech separation," in IEEE International Conference Acoustics, Speech and Signal Processing (ICASSP), 2010.

[59] E. M. Grais and H. Erdogan, "Regularized nonnegative matrix factorization using gaussian mixture priors for supervised single channel source separation," Computer Speech and Language, vol. 27, no. 3, pp. 746-762, May 2013.

[60] E. M. Grais and H. Erdogan, "Gaussian mixture gain priors for regularized nonnegative matrix factorization in single-channel source separation," in Annual Conference of the International Speech Communication Association (InterSpeech), 2012.

[61] E. M. Grais and H. Erdogan, "Hidden Markov Models as priors for regularized nonnegative matrix factorization in singlechannel source separation," in Annual Conference of the International Speech Communication Association (InterSpeech), 2012.

[62] E. M. Grais and H. Erdogan, "Spectro-temporal postenhancement using MMSE estimation in NMF based singlechannel source separation," in Annual Conference of the International Speech Communication Association (InterSpeech), 2013.

[63] E. M. Grais and H. Erdogan, "Single channel speech music separation using nonnegative matrix factorization and spectral masks," in International Conference on Digital Signal Processing, 2011.

[64] Yan-Bo Lin1*, Yuan-Shan Lee1*, Tuan Pham1 , Tzu-Chiang Tai2 , and Jia-Ching Wang1. A Novel Approach for Single Channel Source Separation 2016 International Conference on Consumer Electronics-Taiwan.

[65] Ephraim, Y., Van Trees, H.L., A signal subspace approach for speech enhancement. ICASSP, vol. 2, pp. 355-358.

[66] K. Kinoshita, M. Delcroix, S. Gannot, E. A. P. Habets, R. Haeb-Umbach, W. Kellermann, V. Leutnant, R. Maas, T. Nakatani, B. Raj, A. Sehr, T. Yoshioka, "A Summary of the REVERB Challenge: State-of-the-art and Remaining Challenges in Reverberant Speech Processing Research", *EURASIP Journal on Advances in Signal Processing*, vol. 2016, no. 7, pp. 1-19, 2016.

[67] X.-L. Zhang and D. Wang, "Deep ensemble learning for monaural speech separation," IEEE/ACM Trans. Audio, Speech, Lang. Process., vol. 24, no. 5, pp. 967–977, May 2016.

[68] E. M. Grais , G. Roma, A. J. R. Simpson, and M. D. Plumbley, "Combining mask estimates for single channel audio source separation using deep neural networks," in Proc. InterSpeech, 2016.

[69] Emad M. Grais, Gerard Roma, Andrew J. R. Simpson, and Mark D. Plumbley, Fellow, Two Stage Single Channel Audio Source Separation using Deep Neural NetworksIEEE **2017**. IEEE/ACM Transactions on Audio, Speech, and Language Processing **2017**.

[70] Ivica Kopriva, Senior Member IEEE, Wei Ju, Bin Zhang, Fei Shi, Dehui Xiang, Kai Yu, Ximing Wang, Ulas Bagci and Xinjian Chen*, Senior Member IEEEI. Single-channel Sparse Nonnegative Blind Source Separation Method for Automatic 3D Delineation of Lung Tumor in PET Images. IEEE Journal of Biomedical and Health Informatics. 2016.

[71] [58] Q. He, S. Su, and R. Du, "Separating mixed multi-component signal with an application in mechanical watch movement," Dig. Sig. Proc., vol. 18, pp. 1013-1028, 2008.

[72] [59] D. Gunawan, and D. Sen, "Iterative Phase Estimation for the Synthesis of Separated Sources From Single-Channel Mixtures," IEEE Sig. Proc. Let., vol. 17, no. 5, pp. 421-424, May 2010.

[73] [60] R. M. Parry, and I. Essa, "Phase-Aware Non-negative Spectrogram Factorization," Lect. Notes Comp. Sci., vol. 4666, pp. 536-543, 2007.

[74] [61] B. Gao, W. L. Woo, and B. W.-K. Ling, "Machine Learning Source Separation Using Maximum A Posteriori Nonnegative Matrix Factorization," IEEE Trans. on Cybernetics, vol. 44, no. 7, pp. 1169-1179, July 2014.

[75]] J. I. Qing, R. Xinbo, and Y. E. Zhihong, "Worst conducted EMI spectrum of the critical conduction mode boost PFC converter," IEEE Trans. Power Electron., vol. 30, no. 3, pp. 1230–1241, Mar. 2015.