EEG energy analysis based on MEMD with ICA pre-processing

Yunchao Yin*, Jianting Cao*,[‡], Toshihisa Tanaka^{†,‡}

* Saitama Institute of Technology, 1690 Fusaiji Fukaya-shi 369-0293 Saitama Japan

E-mail: {n0503hpg, cao}@sit.ac.jp, Tel: +81-48-585-6854

† Tokyo University of Agriculture and Technology, 2-24-16 Nakacho Koganei-shi Tokyo 184-8588 Japan

‡ Brain Science Institute, RIKEN, 2-1 Hirosawa Wako-shi Saitama 351-0198 Japan

Abstract—Analysis of EEG energy is a useful technique in the brain signal processing. This paper presents a data analysis method based on multivariate empirical mode decomposition (MEMD) with ICA pre-processing to calculate and evaluate the energy of EEG recorded from the quasi brain deaths. The main advantage of introducing ICA pre-processing is that we can reduce the noise and other unexpected components. The simulation results illustrate the effectiveness and performance of the proposed method in brain death determination.

I. INTRODUCTION

EEG energy analysis is important and useful in the brain signal processing. In the determination of brain death, EEG energy analysis is used to evaluate the brain activity. Several methods of EEG energy analysis such as empirical mode decomposition (EMD) [1] and multivariate empirical mode decomposition (MEMD) [2] have been proposed to evaluate the brain activity [3]. The MEMD is a fully data-driven time-frequency technique which adaptively decomposes a set of signals into a finite set of amplitude-frequency modulated components, namely intrinsic mode functions (IMFs). Since the influence of environmental noise such as additive noise or/and the power supply interference, so that the accuracy of EEG energy analysis by using MEMD is not high.

In this paper, we focus on a novel data analysis method based on MEMD with ICA pre-processing to calculate and evaluate the energy of EEG recorded from patients. In the pre-processing stage, the joint approximate diagonalization of eigenmatrices (JADE) [4] algorithm associated with the developed factor analysis (FA) method [5] is applied to reduce the power of additive noise and power supply interference. In the MEMD stage, each channel is decompose into a small number of IMFs with a specific frequency. The desirable components can be extracted from the decomposed IMF based on the frequency band. In the post processing stage, the energy of EEG recorded from patients is calculated by fast Fourier transformation (FFT). The computer simulations are given to demonstrate the effectiveness of the proposed method not only in the artificial data but also in the quasi brain death EEG data.

The paper is organized as follows. In Section II, we first introduce ICA associated our developed factor analysis algorithms, and the recently developed MEMD method; in Section III, we first conduct the computer simulation to validate our

method by using artificial data and we compare MEMD to MEMD with ICA results, then we use real-world EEG data collected from the quasi brain death patient. Section IV includes the conclusion.

II. METHOD OF EEG DATA ANALYSIS

A. ICA with FA

In this subsection, we present a high level additive noise reduction technique based on the factor analysis (FA). Combining this technique with one of the standard ICA algorithms (for example, using JADE algorithm [4]), we can reduce the power of additive noise, and extract source signal efficiently.

The model based on the practical EEG examination can be formulated by

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{e}(t), \quad t = 1, 2, \cdots,$$
 (1)

where $\mathbf{x}(t) = [x_1(t), \cdots, x_m(t)]^T$ represent the observed m signals observed from sensor at time t. Each sensor signal $x_i(t)$ contains n common components (e.g. brain activities, interference components, etc.) represented by the vector $\mathbf{s}(t) = [s_1(t), \cdots, s_n(t)]^T$ and a unique component (called additive noise) which is an element in the vector $\mathbf{e}(t) = [e_1(t), \cdots, e_m(t)]^T$. Since the source components are overlapped, and transferred rapidly to the sensors, an element of the numerical matrix $\mathbf{A} \in \mathbf{R}^{m \times n} = (a_{ij})$ can be consider as a quantity related to the physical distance between i-th sensor and j-th source. Base on this definition of Eq. (1), we note that a source component s_i at least contributes to more than two sensors, and a noise component e_i contributes at most only one sensor. In the model, our goal is to estimate the unknown independent sources \mathbf{s} .

There are two kinds of noise components that have to be reduced or discarded in the EEG data analysis. The first kind of noise is called additive noise which is generated from each sensor. The standard ICA is usually failed to reduce such kind of noise. Therefore, we apply the FA technique in the pre-processing step to reduce the power of additive noises at first. The second kind of noise is a common component such as environmental interference. This kind of noise can be discarded after the independent source decomposition.

Let us rewrite Eq. (1) set in a data matrix form as

$$\mathbf{X}_{(m \times N)} = \mathbf{A}_{(m \times n)} \mathbf{S}_{(n \times N)} + \mathbf{E}_{(m \times N)}, \tag{2}$$

where N denotes data samples. When the sample size N is sufficiently large, the covariance matrix of the data can be written as $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^T + \mathbf{\Psi}$, where $\mathbf{\Sigma} = \mathbf{X}\mathbf{X}^T/N$, and the covariance of additive noise components \mathbf{E} represented by $\mathbf{\Psi} = \mathbf{E}\mathbf{E}^T/N$ is a diagonal matrix. For convenience, we assume that \mathbf{X} has been divided by \sqrt{N} so that the covariance matrix can be given by $\mathbf{C} = \mathbf{X}\mathbf{X}^T$.

To estimate both matrix ${\bf A}$ and the diagonal elements of ${\bf \Psi}$ from the data, we employ a cost function as

$$L(\mathbf{A}, \mathbf{\Psi}) = tr \left[\mathbf{A} \mathbf{A}^T - (\mathbf{C} - \mathbf{\Psi}) \right] \left[\mathbf{A} \mathbf{A}^T - (\mathbf{C} - \mathbf{\Psi}) \right]^T.$$
(3)

Minimizing the cost function, we obtain an estimate $\hat{\Psi}$ such as $\hat{\Psi} = diag(\mathbf{C} - \hat{\mathbf{A}}\hat{\mathbf{A}}^T)$. The estimate for $\hat{\mathbf{A}}$ can be obtained from $\partial L(\mathbf{A}, \Psi)/\partial \mathbf{A} = 0$. Here, we employ eigenvalue decomposition $\hat{\mathbf{A}} = \mathbf{U}_n \mathbf{\Lambda}_n^{\frac{1}{2}}$, where $\mathbf{\Lambda}_n$ is a diagonal matrix whose elements are the n largest eigenvalues of \mathbf{C} . The columns of \mathbf{U}_n are the corresponding eigenvectors.

The FA method plays the same role in source decorrelation as the standard principal component analysis (PCA) method, however the noise variance Ψ is taken into account. The difference between the two methods is that the PCA approach is used to fit both the diagonal and off-diagonal elements of \mathbf{C} , whereas the FA approach is used to only fit the off-diagonals elements of \mathbf{C} . Based on this property, the FA approach enables us to reduce high level additive noise which is very important in EEG energy analysis.

Once the estimates for $\hat{\mathbf{A}}$ and $\hat{\mathbf{\Psi}}$ converge to stable values, we can finally compute the score matrix using

$$\mathbf{Q} = \left[\hat{\mathbf{A}}^T \hat{\boldsymbol{\Psi}}^{-1} \hat{\mathbf{A}}\right]^{-1} \hat{\mathbf{A}}^T \hat{\boldsymbol{\Psi}}^{-1}.$$
 (4)

From the above result, the new transformation data can be obtained by employing $\mathbf{z} = \mathbf{Q}\mathbf{x}$. Note that the covariance matrix is $E\left[\mathbf{z}\mathbf{z}^{T}\right] = \mathbf{\Lambda}_{n} + \mathbf{C}\mathbf{\Psi}\mathbf{C}^{T}$, which implies that the source signals in a subspace are decorrelated.

The rotation procedure in JADE uses matrices $\mathbf{F}(\mathbf{M})$ formulated by a fourth-order cumulant tensor of the outputs with an arbitrary matrix \mathbf{M} as

$$\mathbf{F}(\mathbf{M}) = \sum_{k=1}^{K} \sum_{l=1}^{L} \operatorname{Cum}(z_i, z_j, z_k, z_l) m_{lk},$$
 (5)

where $\operatorname{Cum}(\cdot)$ denotes a standard cumulant and m_{lk} is the (l,k)-th element of matrix \mathbf{M} . The correct rotation matrix \mathbf{W} can be obtained by diagonalizing the matrix $\mathbf{F}(\mathbf{M})$; namely, $\mathbf{WF}(\mathbf{M})\mathbf{W}^T$ approaches to a diagonal matrix. After the FA and ICA approaches, the decomposed independent components $\mathbf{y} \in \mathbf{R}^n$ can be obtained from a linear transformation as

$$\mathbf{y}(t) = \mathbf{W}\mathbf{z}(t),\tag{6}$$

where $\mathbf{W} \in \mathbf{R}^{n \times n}$ is also termed as the demixing matrix.

The estimated sensor signal \hat{x} are obtained as

$$\hat{\mathbf{x}}(t) = \mathbf{W}^{-1}\mathbf{Q}^{-1}\mathbf{y}(t),\tag{7}$$

B. Multivariate Empirical Mode Decomposition (MEMD)

1) Existing EMD Algorithm: EMD decomposes the original signal into a finite set of amplitude- and/or frequency-modulated components, termed intrinsic mode functions (IMFs), which represent its inherent oscillatory modes [1]. More specifically, for a real-valued signal x(k), the standard EMD finds a set of N IMFs $\{c_i(k)\}_{i=1}^N$, and a monotonic residue signal r(k), so that

$$x(k) = \sum_{i=1}^{n} c_i(k) + r(k).$$
 (8)

IMFs $c_i(k)$ are defined so as to have symmetric upper and lower envelopes, with the number of zero crossings and the number of extrema differing at most by one. The process to obtain the IMFs is called sifting algorithm.

The first complex extension of EMD was proposed in [8]. An extension of EMD to analyze complex/bivariate data which operates fully in the complex domain was first proposed in [9], termed rotation-invariant EMD (RI-EMD). An algorithm which gives more accurate values of the local mean is the bivariate EMD (BEMD) [10], where the envelopes corresponding to multiple directions in the complex plane are generated, and then averaged to obtain the local mean. An extension of EMD to trivariate signals has been recently proposed in [11]; the estimation of the local mean and envelopes of a trivariate signal is performed by taking projections along multiple directions in three-dimensional spaces.

2) The Proposed n-Variate EMD Algorithm[2]: For multivariate signals, the local maxima and minima may not be defined directly because the fields of complex numbers and quaternions are not ordered [11]. Moreover, the notion of 'oscillatory modes' defining an IMF is rather confusing for multivariate signals. To deal with these problems, the multiple real-valued projections of the signal is proposed in [2]. The extrema of such projected signals are then interpolated componentwise to yield the desired multidimensional envelopes of the signal. In MEMD, we choose a suitable set of direction vectors in n-dimensional spaces by using: (i) uniform angular coordinates and (ii) low-discrepancy pointsets.

The problem of finding a suitable set of direction vectors that the calculation of the local mean in an n-dimensional space depends on can be treated as that of finding a uniform sampling scheme on an n sphere. For the generation of a pointset on an (n-1) sphere, consider the n sphere with centre point C and radius R, given by

$$R = \sum_{j=1}^{n+1} (x_j - C_j)^2.$$
 (9)

A coordinate system in an n-dimensional Euclidean space can then be defined to serve as a pointset on an (n-1) sphere. Let $\{\theta_1, \theta_2, \cdots, \theta_{n-1}\}$ be the (n-1) angular coordinates, then

an *n*-dimensional coordinate system having $\{x_i\}_{i=1}^n$ as the *n* coordinates on a unit (n-1) sphere is given by

$$x_n = sin(\theta_1) \times \cdots \times sin(\theta_{n-2}) \times sin(\theta_{n-1}).$$
 (10)

Discrepancy can be regarded as a quantitative measure for the irregularity (non-uniformity) of a distribution, and may be used for the generation of the so-called 'low discrepancy pointset', leading to a more uniform distribution on the nsphere. A convenient method for generating multidimensional 'low-discrepancy' sequences involves the family of Halton and Hammersley sequences. Let x_1, x_2, \dots, x_n be the first n prime numbers, then the ith sample of a one-dimensional Halton sequence, denoted by r_i^x is given by

$$r_i^x = \frac{a_0}{x} + \frac{a_1^2}{x^2} + \frac{a_2^3}{x^3} + \dots + \frac{a_s^{s+1}}{x^s},$$
 (11)

where base-x representation of i is given by

$$i = a_0 + a_1 \times x + a_2 \times x^2 + \dots + a_s \times x^s.$$
 (12)

Starting from i = 0, the *i*th sample of the Halton sequence then becomes

$$(r_i^{x_1}, r_i^{x_2}, r_i^{x_3}, \cdots, r_i^{x_n}).$$
 (13)

Consider a sequence of *n*-dimensional vectors $\{\mathbf{v}(t)\}_{t=1}^T =$ $\{v_1(t), v_2(t), \cdots, v_n(t)\}$ which represents a multivariate signal with n-components, and $\mathbf{x}^{\theta_k} = \{x_1^k, x_2^k, \cdots, x_n^k\}$ denoting a set of direction vectors along the directions given by angles $\theta_k = \{\theta_1^k, \theta_2^k, \cdots, \theta_{n-1}^k\}$ on an (n-1) sphere. Then, the proposed MEMD suitable for operating on n-variate time series is summarized in the following.

- 1) Choose a suitable pointset for sampling on an (n-1)
- Calculate a projection, denoted by $p^{\theta_k}(t)\}_{t=1}^T$, of the input signal $\{\mathbf{v}(t)\}_{t=1}^T$ along the direction vector \mathbf{x}^{θ_k} , for all k (the whole set of direction vectors), giving
- pθk(t) k=1 as the set of projections.
 Find the time instants {tike time instants time instants {tike time instants time instants
- 5) For a set of K direction vectors, the mean $\mathbf{m}(t)$ of the envelope curves is calculated as

$$\mathbf{m}(t) = \frac{1}{K} \sum_{k=1}^{K} \mathbf{e}^{\theta_k}(t). \tag{14}$$

6) Extract the 'detail' d(t) using d(t) = x(t) - m(t). If the 'detail' d(t) fulfills the stoppage criterion for a multivariate IMF, apply the above procedure to x(t) - d(t), otherwise apply it to d(t).

The stoppage criterion for multivariate IMFs is similar to the standard one in EMD, which requires IMFs to be designed in such a way that the number of extrema and the zero crossings differ at most by one for S consecutive iterations of the sifting algorithm. The optimal empirical value of S has been observed to be in the range of 2-3 [12]. In the MEMD, we apply this criterion to all projections of the input signal and stop

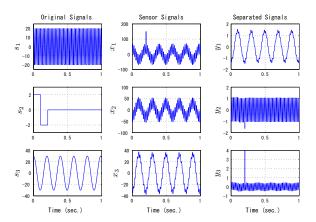


Fig. 1: ICA decomposed results

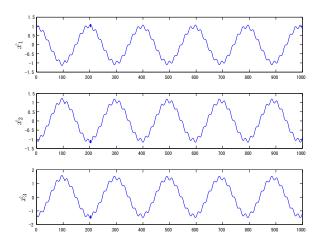


Fig. 2: Estimated sensor signals x used y_1 back-propagation

the sifting process once the stopping condition is met for all projections.

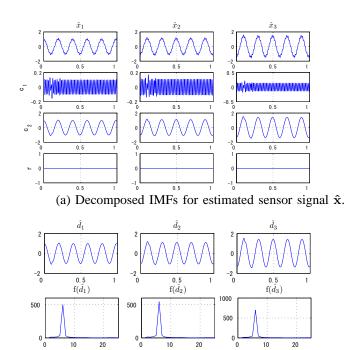
III. EXPERIMENTS AND RESULTS

In this section, we first conduct the computer simulation to compare the algorithm of MEMD with ICA pre-processing and MEMD. We then use the real-world EEG data collected from the quasi brain death patient to demonstrate the effectiveness of the proposed method.

A. Simulation Results for Artificial Data

Three artificial signals: a low frequency (5Hz) sine wave signal simulated a delta wave brain activity, an impulse signal simulated ECG and a higher frequency (50Hz) sine wave signal simulated the power supply interference are used to generate the sensor signals with random numerical matrix as Eq. (1). Moreover, an additive noise with a SNR (signal noise ratio) 16 dB (the power of noise is higher than the power of signal) was added to the sensor signal.

By applying FA and ICA algorithms described in Section II(A) to above 3 signals, we obtain the result shown in Fig. 1. As seen from Fig.1, the expected 5Hz sine wave component was decomposed well. Used the decomposed component



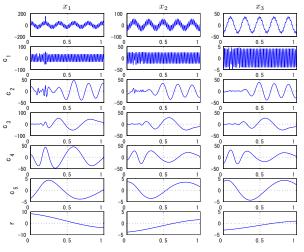
(b) Denoised components $\hat{\mathbf{d}}$ in time and frequency domains.

Fig. 3: The results for MEMD with ICA per-processing.

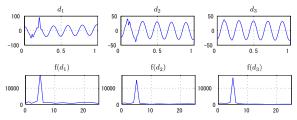
 y_1 , we can obtain the estimated sensor signals by $\hat{\mathbf{x}} = \mathbf{W}^{-1}\mathbf{Q}^{-1}y_1$, where Q and W are estimated matrix obtained from the factor analysis (FA) and independent component analysis (ICA), respectively. The result is shown in Fig. 2. As shown in Fig. 2, we found that the estimated sensor signals still have some unexpected high-frequency noisy component.

In order to obtain more accuracy estimated component, we further apply the MEMD algorithm described in Section II(B) to the estimated sensor signals $\hat{\mathbf{x}}$. The results were shown in Fig. 3. In Fig. 3(a), three estimated sensor signals \hat{x}_1, \hat{x}_2 and \hat{x}_3 were decomposed into two IMF components c_1 , c_2 and a monotonic residue component r from high frequency to low frequency simultaneously. Since the IMF components c_1 with a high frequency scales refer to noise and the residual component r is not the typical useful components considered, only the desired components c_2 is the denoised component, we named it as $\hat{\mathbf{d}} = [\hat{d}_1, \hat{d}_2, \hat{d}_3]$. By using fast Fourier transform (FFT), we can obtain $f(\mathbf{d})$ in frequency domain (see Fig. 3(b)). It should be noted that the above this result is obtained from two stages processing as ICA with FA decomposition and the multivariate empirical mode decomposition (MEMD). The result by using these stages processing was almost close to the ideal case.

Next, in order to comparison, we apply MEMD to the same set of artificial data \mathbf{x} without ICA pre-processing. The results were obtained in Fig. 4. In Fig. 4(a), many IMF components c_1 to c_5 and a residue component r were decomposed. By removing high frequency component c_1 and the residual component r, then combined the components c_2 to c_5 as a desired component named \mathbf{d} , we obtain the result shown in Fig. 4(b). Comparing the results shown in Fig. 4(b)



(a) Decomposed IMFs for sensor signal x.



(b) Denoised components d in time and frequency domains.

Fig. 4: The result for MEMD.

to Fig. 3(b), we found that the decomposed component in the time domain shown in Fig. 4(b) is not smoothly, this means that only used the MEMD is not sufficiently to remove noise completely. Moreover, the power spectrum of the decomposed component in the frequency domain shown in Fig. 4(b) is much higher than that of shown in Fig. 3(b).

Let's define the EEG energy using the power spectrum within the frequency band multiply by recorded EEG time. This definition can be also used to calculate the other signals energy generated by artificial data. Using the formula of energy, we can calculate and evaluate the energy of the sensor signal generated by an original source signal, the estimated sensor signal by used ICA pre-processing with the MEMD algorithms, and the estimated sensor signal used MEMD algorithms. The results were shown in Table I.

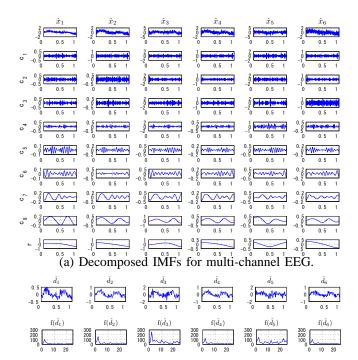
In Table I, the duration of signal is one second. As seen from Table I, we know that the accuracy of the estimated sensor signal used the MEMD with ICA pre-processing algorithm is much higher than that of used the MEMD algorithm. This illustrates the effectiveness of the proposed method.

B. Result for Patient's EEG Analysis

Based on the results in the previous sub-section, we now apply the proposed method to analysis the real-world EEG data recorded from quasi brain death. This patient's EEG data was directly recorded at the bedside of the patients in the intensive care unit (ICU) in a hospital of Shanghai. In the

TABLE I: The energy of sensor signal and its estimates

Energy of sensor signal and its estimates	Value
Sensor signal generated by a 5Hz original source signal	2.96×10^{3}
Estimated sensor signal	
used the ICA pre-processing and MEMD	2.45×10^{3}
Estimated sensor signal	
used the MEMD	8.07×10^4



(b) Denoised EEG signal in time and frequency domains.

Fig. 5: The result of the quasi brain death by MEMD with ICA per-processing.

EEG recording, only nine electrodes are chosen to apply to patients. Among these electrodes, six exploring electrodes as well as GND were placed on the forehead, and two electrodes (A1, A2) as the reference were placed on the earlobes. The sampling rate of EEG was 1000 Hz and the resistances of the electrodes were set to less than 10 k Ω .

To analyze the EEG energy of this patient, one second EEG data of patient is selected randomly. We firstly remove the unexpected high-frequency noisy component which is similar to the power supply interference by FA and ICA algorithms described in Section II(A) and obtain the estimated sensor signals $\hat{\mathbf{x}}$. In order to obtain more accuracy component, by using the MEMD algorithm described in Section II(B), the estimated sensor signals $\hat{\mathbf{x}}$ were decomposed into 10 IMF components $(c_1$ to c_9 and r)(Fig. 5(a)). Since IMF components c_1 to c_3 with high frequency scales refer to noise and the residual component r is not the typical useful components considered, the desired components from c_4 to c_9 are combined to form the denoised EEG signal $\hat{\mathbf{d}}$, and change into frequency domain by fast Fourier transform (FFT)(Fig. 5(b)). We calculate the

energy of EEG data and the energy of this patient is 3.34×10^3 . In this case, we know that this patient is in the coma state. The clinical doctor confirmed this result is correct.

IV. CONCLUSIONS

In this paper, we proposed a data analysis method based on MEMD with ICA pre-processing to calculate and evaluate the energy of EEG recorded from the quasi brain deaths. From the simulation result for artificial data, we demonstrate the proposed method is more effective than by using MEMD. In real-world EEG data, we can distinguish the coma patient from the quasi brain deaths correctly by proposed method. We have illustrated that the proposed method is useful and necessary in brain death determination.

ACKNOWLEDGMENT

This work was partly supported by KAKENHI (24360146, 22560425), and partly supported by JSPS and NSFC under the Japan-China Scientific Cooperation Program.

REFERENCES

- [1] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung, H. Liu,: "The Empirical Mode Decomposition and The Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis". Proceedings of the Royal Society of London, A 454, pp. 903-995, 1998
- [2] N. Rehman, D. Mandic: "Multivariate Empirical Mode Decomposition". Proceedings of the Royal Society A, vol. 466, No. 2117, pp. 1291-1302, 2010
- [3] Y. Yin, J. Cao, Q. Shi, D. Nabduc, T. Tanaka and R. Wang: "Comatose Patients and Brain Deaths' EEG Analysis Using MEMD". Processings of the 2011 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference.
- [4] J. F. Cardoso, and A. Souloumiac, "Jacobi angles for simultaneous diagonalization", SIAM Journal on Matrix Analysis and Applications, Vol. 17, No. 1, pp. 145–151, 1996.
- [5] J. Cao, N. Murata, S. Amari, A. Cichocki, and T. Takeda, "A robust approach to independent component analysis of signals with high-level noise measurements", *IEEE Transactions on Neural Networks*, Vol. 14, No. 3, pp. 631–645, May 2003.
- [6] R. Taylor, : "Reexamining the definition and of death", Semin Neurol, Vol. 17, pp. 265-270, 1997
- [7] J. Cao and Z. Chen: "Chapters 15: Advanced EEG Signal Processing in Brain Death Diagnosis" in Book [Signal Processing Techniques for Knowledge Extraction and Information Fusion] Editors: D.P. Mandic, M. Golz, A. Kuh, D. Obradovic, and T. Tanaka, pp.275-297, Springer. March 2008
- [8] T. Tanaka, D. Mandic: "Complex Empirical Mode Decomposition". IEEE Signal Processing Letters, vol. 14, No. 2, pp. 101-104, 2006
- [9] M. Altaf, T. Gautama, T. Tanaka, D. Mandic: "Rotation Invariant Complex Empirical Mode Decomposition". In Proc. IEEE Int. Conf. on Acoustics, Speech, Signal Processing, Honolulu, HI, pp. 1009-1012, 2007
- [10] G. Rilling, P. Flandrin, P. Goncalves, J. Lilly: "Bivariate Empirical Mode Decomposition". *IEEE Signal Process*. Lett. 14, pp. 936-939, 2007
- [11] N. Rehman, D. Mandic,: "Empirical Mode Decomposition for Trivariate Signals". *IEEE T. Signal Process*. (In print)
- [12] N. Huang, M. Wu, S. Long, S. Shen, W. Qu, P. Gloersen, K. Fan: "A Confidence Limit for The Empirical Mode Decomposition and Hilbert Spectral Analysis". Proc. R. Soc. Lond, A 459, pp. 2317-2345, 2003