Common Spatial Pattern Using Multivariate EMD for EEG Classification

Long Zhang*, Cheng Zhang*, Hiroshi Higashi* ‡, Jianting Cao† ‡, and Toshihisa Tanaka* ‡
* Tokyo University of Agriculture and Technology, Japan
* E-mail: {qiaoliu,chousei,higashi}@sip.tuat.ac.jp, tanakat@cc.tuat.ac.jp Tel/Fax: +81-42-388-7123
† Saitama Institute of Technology, Japan
† cao@sit.ac.jp Tel/Fax: +81-48-585-6854
‡ RIKEN Brain Science Institute, Japan

***Abstract*—Brain-computer interface (BCI) is a system to translate humans thoughts into commands. For electroencephalography (EEG) based BCI, motor imagery is considered as one of the most effective ways. This paper presents a method for classifying EEG during motor-imagery by the combination of well-known common spatial pattern (CSP) with so-called multivariate empirical mode decomposition (MEMD), which is effectively suitable for processing of multichannel signals of EEG. In the proposed method, the EEG signal is decomposed into intrinsic mode functions (IMF) using the MEMD. Different from EMD, the number of IMF is the same in each channel. Then by removing some of the IMFs, the reconstructed signal can carry more useful information than the original signal. Based on the MEMD, weights of CSP are found. By off-line simulation, the use of MEMD in CSP has shown to perform well in the application to the classification of EEG signals.**

I. Introduction

Brain-computer interface (BCI) can translate humans thoughts directly to the outside world. It is a new technology as a radically new communication option for those with neuromuscular impairments that prevent them from using conventional augmentative communication methods. It enables us to connect to the real world without peripheral nerves and muscles.

The brain normally produces tiny electrical signals that come from the brain cells and nerves which send messages to each other. These electrical signals can be detected and recorded by the electroencephalography (EEG) measurements. It has been chosen to capture brainwaves for BCI applications because of its simplicity, inexpensiveness and high temporal resolution.

Because of the volume conduction, EEG signals give a unclear image of brain activity [2]. Therefore, a spatial filtering preprocessing stage which performs source separation before feature extraction is often used to improve BCI's performance, so that it can give a clear response to the brain acticity. Recently, common spatial pattern (CSP) has been widely used in spatial filtering because of its efficiency [3] [4]. However, it can only reflect the separative ability of the mean power of two classes. Artifacts such as eye and muscle activities may dominate over the EEG signal, and thus they may give excessive power in some channels. Because of CSP simply pooling the covariance matrices of trials together, if an artifact happens to be unevenly distributed in different experiment conditions, CSP will capture it with a high eigenvalue. This will distort the following CSP spatial filter. Therefore, it must be removed as much as possible.

To remove such a non-linear artifact, the empirical mode decomposition (EMD) has been used [1] [2]. The EMD is a signal processing decomposition technique that decomposes the signal into waveforms modulated in both amplitude and frequency by extracting all of the oscillatory modes embedded in the signal. The key issue in EMD processing is the computation of the local mean of the original singal, a step which depends critically on finding the local extrema. In real-valued EMD, the local mean is computed by taking an average of upper and lower envelopes, which in turn are obtained by interpolating between the local maxima and minima. In the actual signal processing, multivariate signals are necessary to be dealt with. In general, for multivariate signals, the local maxima and minima may not be defined directly [5]. So it is only suitable for univariate (real valued) signals. Although artifacts will commonly appear over several channels, univariate EMD cannot consider inter-channel relationship in decomposition.

Recently, an n-variate EMD processing for signal processing decomposition, so-called multivariate empirical mode decomposition (MEMD) has been proposed [10], which is an extension of the basic EMD suitable for dealing with multivariate signals. MEMD generates multiple n-dimensional envelopes by taking signal projections along different directions in n-dimensional spaces, and then averaged to obtain the local mean.

In this paper, we propose to combine CSP with MEMD, to extract necessary components from multi-channel signals. By the application of MEMD processing, an n-variate signal is decomposed into a finite set of amplitude-and/or frequency-modulate components, where some artifact-related components are discarded and the reconstructed signal is obtained. This signal is used for designing CSP.

II. Common Spatial Pattern (CSP) – Review

CSP is an effective method for feature extraction and classification in two class motor imagery–BCI. This section reviews basic CSP processing [3] [4].

Let $\boldsymbol{X} \in \mathbb{R}^{M \times N}$ be an observed signal, where M is the number of channels and N is the number of samples. CSP finds a spatial weight vector, $\boldsymbol{w} \in \mathbb{R}^M$, in such a way that a variance of a signal extracted by linear combination of $\boldsymbol{X}$ is minimized in a class. Actually, we do not directly use $\boldsymbol{X}$, but use the filtered signal described as $\hat{\boldsymbol{X}} = \mathcal{H}(\boldsymbol{X})$ in CSP, where $\mathcal{H}$ is a bandpass filter which passes the frequency band related to brain activity of motor-imagery. Denote the components of $\hat{\boldsymbol{X}}$ by $\hat{\boldsymbol{X}} = [\hat{\boldsymbol{x}}_1, \ldots, \hat{\boldsymbol{x}}_N]$, where $\hat{\boldsymbol{x}}_n \in \mathbb{R}^M$, and the time mean of the observed signal is given by $\boldsymbol{\mu} = \frac{1}{N}\sum_{n=1}^N \hat{\boldsymbol{x}}_n$. Then, the variance of the extracted signal of $\hat{\boldsymbol{X}}$ is given by

$$\sigma^2(\boldsymbol{X}, \boldsymbol{w}) = \frac{1}{N}\sum_{n=1}^N |\boldsymbol{w}^T(\hat{\boldsymbol{x}}_n - \boldsymbol{\mu})|^2. \tag{1}$$

We assume that sets of the learning data, C_1 and C_2, where C_d contains the signals belonging to class d, $d \in \{1, 2\}$ is a class label, and $C_1 \cap C_2 = \emptyset$. CSP finds the weight vector that minimizes the intra-class variance in C_c, under the normalization of samples, where $c \in \{1, 2\}$. More specifically, for c fixed, CSP finds $\boldsymbol{w}_c$ by solving the following optimization problem [3] [4];

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & E_{X \in C_c}[\sigma^2(\boldsymbol{X}, \boldsymbol{w})] \\ \text{subject to} \quad & \sum_{d=1,2} E_{X \in C_d}[\sigma^2(\boldsymbol{X}, \boldsymbol{w})] = 1, \end{aligned} \tag{2}$$

where $E_{X \in C_d}[\cdot]$ denotes the expectation over C_d. Then, (2) can be rewritten as

$$\min_{\boldsymbol{w}} \boldsymbol{w}^T \boldsymbol{\Sigma}_c \boldsymbol{w}, \quad \text{subject to} \, \boldsymbol{w}^T(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\boldsymbol{w} = 1, \tag{3}$$

where $\boldsymbol{\Sigma}_d$, $d = 1, 2$, are defined as

$$\boldsymbol{\Sigma}_d = E_{X \in C_d}\left[\frac{1}{N}\sum_{n=1}^N (\hat{\boldsymbol{x}}_n - \boldsymbol{\mu})(\hat{\boldsymbol{x}}_n - \boldsymbol{\mu})^T\right]. \tag{4}$$

The solution of (3) is given by the generalized eigenvector corresponding to the minimum generalized eigenvalue of the generalized eigenvalue problem described as

$$\boldsymbol{\Sigma}_c \boldsymbol{w} = \lambda(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)\boldsymbol{w}_c. \tag{5}$$

For classification, the following method is applicable to CSP. Given an unlabeled data, $\boldsymbol{X}$, we classify $\boldsymbol{X}$ by the following rule:

$$u = \operatorname*{argmin}_c \sigma^2(\boldsymbol{X}, \boldsymbol{w}) \implies \boldsymbol{X} \in C_u. \tag{6}$$

III. Empirical Mode Decomposition

The EMD processing is a fully data-driven method designed for multi-scale decomposition and time-frequency analysis of real-world signals, whereby the original signal is modeled as a linear combination of intrinsic oscillatory modes, called intrinsic mode functions (IMFs), which are defined so as to exhibit locality in time and to represent a single oscillatory mode.

A. Univariate EMD

EMD decomposes the original signal into a finite set of amplitude-and/or frequency-modulated components, termed IMFs, which represent its inherent oscillatory modes. 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

$$x(k) = \sum_{i=1}^N c_i(k) + r(k) \tag{7}$$

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.

The first complex extension of EMD was proposed in [6]. An extension of EMD to analyze complex/bivariate data which operates fully in the complex domain was first proposed in [7], termed rotation invariant EMD (RI-EMD). An processing which gives more accurate values of the local mean is the bivariate EMD (BEMD) [8], 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 [9]; the estimation of the local mean and envelopes of a trivariate signal is performed by taking projections along multiple directions in three-dimensional spaces.

B. Multi-Variate EMD

For multivariate signals, the local maxima and minima may not be defined directly because the fields of complex numbers and quaternions are not ordered [9]. 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 was proposed [10]. The extrema of such projected signals are then interpolated componentwise to yield the desired multidimensional envelopes of the signal. In multivariate EMD, 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 $R = \sum_{j=1}^{n+1}(x_j - C_j)^2$ where a coordinate system in an n-dimensional Euclidean space can then be defined to serve as a pointset on an $(n-1)$ sphere. Let $\{\theta_1, \theta_2, \ldots, \theta_{n-1}\}$ be a set of 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

$$x_n = \sin(\theta_1) \times \cdots \times \sin(\theta_{n-2}) \times \sin(\theta_{n-1}) \tag{8}$$

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 n sphere. A convenient method for generating multidimensional ‘ low-discrepancy ’ sequences involves the family of Halton and Hammersley sequences. Let $x_1, x_2, \ldots, x_n$ be a set of the first n prime numbers, then the ith sample of a one-dimensional Halton sequence, denoted by r_i^x is given by

$$r_i^x = \frac{a_0}{x} + \frac{a_1^2}{x} + \frac{a_3^3}{x} + \cdots + \frac{a_s^{s+1}}{x} \tag{9}$$

where base$-x$ representation of i is given by

$$i = a_0 + a_1 \times x + a_2 \times x^2 + \cdots + a_s \times x^s \tag{10}$$

Starting from $i = 0$, the ith sample of the Halton sequence then becomes

$$(r_i^{x_1}, r_i^{x_2}, r_i^{x_3}, \ldots, r_i^{x_n}) \tag{11}$$

Consider a sequence of n-dimensional vectors $\{V(t)\}_{t=1}^T = \{v_1(t), v_2(t), \ldots, v_n(t)\}$ which represents a multivariate signal with n-components, and $X^{\theta_k} = \{x_1^k, x_2^k, \ldots, x_n^k\}$ denoting a set of direction vectors along the directions givenby angles $\theta_k = \{\theta_1^k, \theta_2^k, \ldots, \theta_{n-1}^k\}$ on an $(n-1)$ sphere. Then, the proposed multivariate extension of EMD suitable for operating on general nonlinear and non-stationary n-variate time series is summarized in the following.

1) Choose a suitable pointset for sampling on an $(n-1)$ sphere.
2) 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 $\mathbf{x}^{\theta_k}$, for all k (the whole set of direction vectors),giving $\{p^{\theta_k}(t)\}_{k=1}^K$ as the set of projections.
3) Find the time instants $\{t_i^{\theta_k}\}$ corresponding to the maxima of the set of projected signals $\{p^{\theta_k}(t)\}_{k=1}^K$.
4) Interpolate $\{t_i^{\theta_k}, \mathbf{v}(t_i^{\theta_k})\}$ to obtain multivariate envelope curves $\{\mathbf{e}^{\theta_k}(t)\}_{k=1}^K$.
5) For a set of K direction vectors, the mean $\mathbf{m}(t)$ of the envelope curves is calculated as

$$\mathbf{m}(t) = \frac{1}{K} \sum_{k=1}^{K} \mathbf{e}^{\theta_k}(t) \tag{12}$$

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 shifting processing. The optimal empirical value of S has been observed to be in the range of 2–3 [11]. In the multivariate EMD, we apply this criterion to all projections of the input signal and stop the shifting process once the stopping condition is met for all projections.

IV. Application of MEMD to CSP

We propose a new method by the combination of CSP with MEMD processing. In the first step MEMD application to the original signal $x(k)$, each channel has the same number of IMF index $c_i(k)$. Through some experiment, we remove some of IMFs, in this way a new signal $\hat{x}(k)$ is reconstructed by

$$\hat{x}(k) = \sum_{i \in \Omega} c_i(k) + r(k) \tag{13}$$

where Ω is IMF index, $r(k)$ is the residue, $c_i(k)$ reperents the IMFs corresponding to the original signal, and k is time. The reconstructed signal can carry more useful information than the original signal, and are used to design CSP.

V. Experimental Results

We experimented classification of EEG during imagined movement using the method we proposed. In the experiment, the proposed method (MEMD–CSP) is used to the classification of the EEG data. We also compare the performance of the proposed method to that of other methods (CSP and EMD–CSP).

A. Dataset

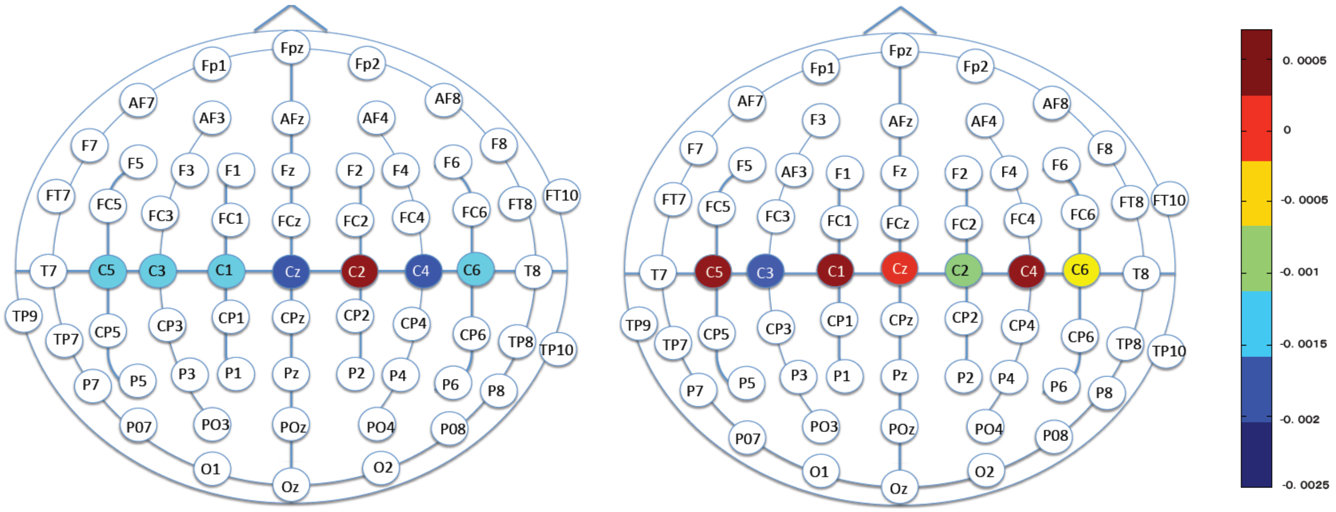
The EEG dataset which we test with is from BCI competitionIII 2005. This dataset comprises of 118 electrode channels out of an EEG amplifier sampled at 100Hz. These data are collected from five subjects *aa*, *al*, *av*, *aw*, *ay*. The EEG data used in this experiment consisted of two classes: right hand (R) and right foot (F) motor imageries. The visual cues at each trial last for 3.5 seconds and there are 280 trials each subject, and 140 trials in each class. [12].

In this experiment, we applied bandpass filter at first, which the passbands for each subject was 7–30 Hz. Because the selection of a suitable EEG reference can greatly influence the classification accuracy and sensitivity to artifacts [13] [14]. In this paper, we also performed another experiment as a comparison, using a small Laplacian reference [4] which is obtained by rereferencing an electrode to the mean of its four nearest neighboring electrodes. Before we apply MEMD or EMD, 100 of 140 trials in each class were randomly divided into 10 groups, all trials were completely disrupted the order, 9 of which were used for processing and learning, leaving 1 group as test data (leave-one-out).

B. Experiment Method

1) Channel selection: It is well-known that during the motor imageries, the parts of brain center are most active [15]. We selected 7 channels from the whole EEG: C1, C2, C3, C4, C5, C6, Cz, which obtained the biggest contrast between the weight spatial weights given by CSP. Here we call channel 51–57 for short.

In Fig. 1 the color level represents the coefficients of the spatial weights given by CSP. The figure shows that when the motor imagery occurred, the spatial weights of the active part of the brain was significantly different from the other parts.



(a) w_1 (b) w_2 (c) colormap

Fig. 1: The CSP weights of subject aa when the CSP was obtained with channel 51-47: (a) weights for right hand movement, and (b) weights for right foot movement. (c) The colormap represents the value of spatial weights.

2) MEMD1–CSP: In eq. (14), we need to determine Ω for reconstruction. To obtain the highest classification accuracy, we need to find a way to choose the proper IMFs $c_i(k)$ that obtained from MEMD processing, then add the $c_i(k)$ to reconstruct the new signal: $\hat{x}(k)$. We make 10 data groups from 100 trials each class, which mentioned in Dataset, because we used 9 of them for learning, we can obtain 90 $c_i^{tr}(k)$ each class ($i \in \Omega$), where tr is the number of trial. We make all combinations of IMF index Ω, reconstruct new signal $\hat{x}(k)$, and make test by CSP. The best result is showed in Table I.

3) MEMD2–CSP: We use the rule of IMF selection based on the combination of EEG with fractional Gaussian noise (fGn) in MEMD in the same way as [2]. It is well known that the EMD of fGn acts as a dyadic filter bank [16], hence the fGn could be considered as a reference signal to determine the IMFs representing the actual EEG signals. In each trial of data, we used the selected 7 channels of EEG and a channel of fGn, as the signal $x(k)$ being processed. After MEMD processing, we obtained IMF index $c_i(k)$ of each channel. To select the proper $c_i(k)$, we compare the average power of the IMF index of each channel (P_{ch}), to the average power of fGn (P_f). The part of the IMF index will be removed, when P_{ch} close to P_f, retain the $c_i(k)$ which P_{ch} is different to P_f apparently. To facilitate the observation, we demonstrate the most obvious and effective subject *aa* in Fig 2. Figure 2 illustrates that we determine the selection of the IMF index of EEG from the power of the signal. In this simulation, the number of IMF index we obtained is 8. The power of 5th, 6th and 7th IMF index of EEG ($c_i(k), i \in \{5,6,7\}$) close to the power of fGn. So we removed these 3 IMF indices of the signal $x(k)$, add the remaining IMF index together, to get a more clean EEG signal $\hat{x}(k)$, the new signal is in the following:

$$\hat{x}(k) = \sum_{i \in \Omega} c_i(k) + r(k), \quad \Omega = \{1,2,3,4,8\}. \tag{14}$$

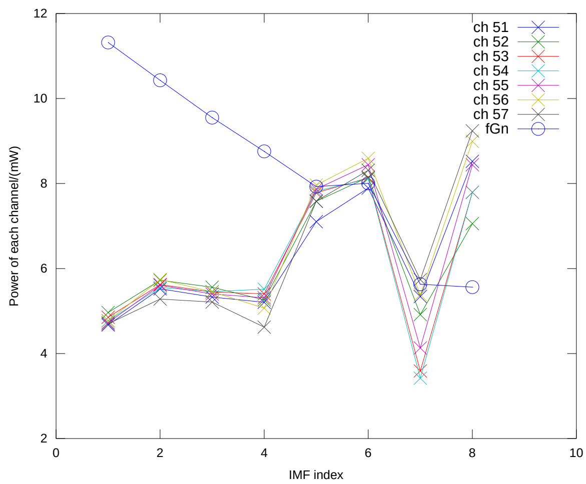


Fig. 2: The relationship of the average power between fGn and channel 51–57. The subject is *aa*. The horizontal axis represents IMF index, and vertical axis represents the power.

4) Laplacian derivative: As a contrast, the Laplacian derivative is applied before the process. The Laplacian reference is obtained by rereferencing an electrode to the mean of its four nearest neighboring electrodes. And then we do the experiment of CSP, EMD–CSP and MEMD–CSP again. Results are shown in Table II.

TABLE I: Classification accuracy

Method	subject				
	aa	al	av	aw	ay
CSP1	**71.3**	88.4	48.6	**89.9**	79.9
CSP2	65.3	**90.2**	63.7	80.3	87.3
EMD–CSP	68.4	89.6	64.1	82.5	86.9
MEMD1–CSP	68.8	90.0	**68.8**	76.3	**87.5**
MEMD2–CSP	60.3	82.9	55.3	60.7	74.0

TABLE II: Classification accuracy (using Laplacian derivative)

Method	subject				
	aa	al	av	aw	ay
CSP1	**70.7**	78.3	48.7	**88.5**	72.6
CSP2	63.7	87.7	64.0	79.1	86.3
EMD–CSP	62.8	89.5	**64.4**	79.6	85.9
MEMD1–CSP	63.8	**95.0**	63.8	71.2	**88.1**
MEMD2–CSP	55.3	79.7	54.7	70.0	75.6

C. Results

Table I shows classification accuracy obtained by each methods. Each method is applied to EEG signals as follows. In the application of the proposed method MEMD–CSP and the comparative method EMD–CSP, we used data of original signal as test data. Accuracy rate is given by 10×10 cross validation.

Table II shows classification accuracy obtained by each methods, using Laplacian derivative.

- CSP1 All of 118 EEG channels were classified by CSP.
- CSP2 For comparison, channel 51–57 of the EEG signals were selected for the classification by CSP.
- EMD–CSP We also classified the EEG signals using EMD processing. The channels we selected are 51–57. Because by definition EMD is suitable for univariate signals, we can obtain different number of IMFs one channel to another. However we removed the same number of IMF $c_i(k)$ from each channel carrying fewer feature value, then reconstruct the new signal $\hat{x}(k)$ and classified by CSP.
- MEMD1–CSP As the procedure shown in section V. B. 2), we classified the EEG signals by using the proposed method in the way of making test for each situation of IMF selection given by MEMD.
- MEMD2–CSP As the the procedure shown in section V. B. 3), to find a more effective method about IMF selection, we also used the method which on the basis of power comparison.

Table I shows that we obtained the best classification accuracy using the proposed method MEMD1–CSP in subjects *av* and *ay*.

In the method MEMD1–CSP for subject *aa*, by making test for IMF selection we obtained the highest classification accuracy is 68.8%, when $\Omega = \{2,5,6\}$. However, in the method MEMD2–CSP, according to Fig. 2, we removed 5th, 6th and 7th IMF index, recontructed the signal $\hat{x}(k)$ with $\Omega = \{1,2,3,4,8\}$, then obtained classification accuracy is 60.3%, it illustrates that the results are not satisfied.

VI. Conclusions

We have proposed a novel method where the EEG signal is decomposed into IMF using the MEMD, and are classified by CSP. By experiment, the use of MEMD–CSP has shown to perform well in the application to classification of EEG signals. However the current study are all based on No.51–57 channels. The experimental results may not be satisfactory. We also try other methods, such as seclecting channels by comparing the spatial weights or in different frequency.

Acknowledgment

This work was supported in part by KAKENHI (21360179).

References

[1] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. C. Yen, C. Tung, and H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc.of the Royal Society A,* vol. 454, no. 1971, pp. 903–995, 1998.

[2] Md. K. I. Molla, T. Tanaka, T. M. Rutkowski, and A. Cichocki, "Separation of EOG artifacts from EEG signals using bivariate EMD," in *Proc. 2010 IEEE International Conference on, Acoustics Speech and Signal Processing (ICASSP),* pp. 562–565, Dallas, TX, Mar. 14–19. 2010.

[3] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Designing optimal spatial filters for single-trial EEG classification in a movement task," *Clinical Neurophysiology*, vol. 110, pp. 787–798, 1999.

[4] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 4, pp. 441–446, 2000.

[5] J. Müller-Gerking, G. Pfurtscheller, and H. Flyvbjerg, "Classification of movement-related EEG in a memorized delay task experiment," *Clinical Neurophysiology*, vol. 111, no. 8, pp. 1353–1365, Aug. 2000.

[6] T. Tanaka and D. P. Mandic, "Complex empirical mode decomposition," *IEEE Signal Processing Letters,* vol. 14, pp. 101–104, 2007.

[7] M. U. Altaf, T. Gautama, T. Tanaka, and D. P. Mandic, "Rotation invariant complex empirical mode decomposition," in *Proc. 2007 IEEE International Conference on, Acoustics Speech and Signal Processing (ICASSP)*, Honolulu, HI, pp. 1009–1012, 2007.

[8] G. Rilling, P. Flandrin, P. Goncalves, and J. M. Lilly, "Bivariate empirical mode decomposition," *IEEE Signal Processing Letters,* vol. 14, pp. 936–939, 2007.

[9] N. Rehman and D. P. Mandic, "Empirical mode decomposition for trivariate signals," *IEEE Trans. Signal Process*. pp. 1059–1068, Mar. 2010.

[10] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proc.of the Royal Society A*, vol. 466, no. 2117, pp. 1291–1302, May. 2010.

[11] N. E. Huang, M. Wu, S. Shen, W. Qu, P. Gloersen, and K. Fan, "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis," *Proc. of the Royal Society. Lond. A*, vol. 459, pp. 2317–2345, 2003.

[12] B. Blankertz, K. R. Müller, D. J. Krusienski, G. Schalk, J. R. Wolpaw, A. Schlogl, G. Pfurtscheller, J. R. Millan, M. Schroder, and N. Birdaumer, "The BCI competition III: Validating alternative approaches to actual BCI problems," *IEEE Trans. Neural Syst. Rehabil. Eng.,* vol. 14, no. 2, pp. 153–159, 2006.

[13] D. J. McFarland, L. M. McCane, S. V. David, and J. R. Wolpaw, "Spatial filter selection for EEG-based communication," *Electroenc. Clin. Neurophys.*, vol. 103, no. 3, pp. 386–394, 1997.

[14] G. Pfurtscheller, "Mapping of event-related desynchronization and type of derivation," *Electroenc. Clin. Neurophys.*, vol. 70, pp. 190–193, 1988.

[15] C. Neuper, G. R. Müller-Putz, R. Scherer, and G. Pfurtscheller, "Motor imagery and EEG-based control of spelling devices and neuroprostheses," *Progress in Brain Research,* vol. 159, pp. 393–409, 2006.

[16] P. Flandrin, G. Rilling, and P. Goncalves, "Empirical mode decomposition as a filter bank," *IEEE Signal Processing Letters,* vol. 11, no. 2, pp. 112–114, Feb. 2004.