International Journal of Neural Systems, Vol. 21, No. 4 (2011) 335–350

DOI: 10.1142/S0129065711002870

World Scientific
www.worldscientific.com

# CONTINUOUS EEG SIGNAL ANALYSIS FOR ASYNCHRONOUS BCI APPLICATION

WEI-YEN HSU
*Graduate Institute of Biomedical Informatics*
*Taipei Medical University, 250 Wu-Xin Street*
*Taipei 110, Taiwan*
*shenswy@stat.sinica.edu.tw*



In this study, we propose a two-stage recognition system for continuous analysis of electroencephalogram (EEG) signals. An independent component analysis (ICA) and correlation coefficient are used to automatically eliminate the electrooculography (EOG) artifacts. Based on the continuous wavelet transform (CWT) and Student's two-sample t-statistics, active segment selection then detects the location of active segment in the time-frequency domain. Next, multiresolution fractal feature vectors (MFFVs) are extracted with the proposed modified fractal dimension from wavelet data. Finally, the support vector machine (SVM) is adopted for the robust classification of MFFVs. The EEG signals are continuously analyzed in 1-s segments, and every 0.5 second moves forward to simulate asynchronous BCI works in the two-stage recognition architecture. The segment is first recognized as lifted or not in the first stage, and then is classified as left or right finger lifting at stage two if the segment is recognized as lifting in the first stage. Several statistical analyses are used to evaluate the performance of the proposed system. The results indicate that it is a promising system in the applications of asynchronous BCI work.

*Keywords*: Asynchronous brain-computer interface (BCI); electroencephalogram (EEG); independent component analysis (ICA); wavelet transform; fractal dimension; support vector machine (SVM).

## 1. Introduction

Brain-computer interface (BCI) is a new communication system that provides an alternative channel to directly transmit messages from the human brain to computers by analyzing the brain's mental activities.[1–26] BCI systems based on the single-trial analysis of electroencephalographic (EEG) signals associated with finger movements have grown rapidly in the last decade.[3] EEG analysis focuses on discriminating left finger from right finger movement using event-related brain potentials (ERP), revealing that there are special characteristics of event-related desynchronization (ERD) and synchronization (ERS) in mu and beta rhythms over the sensorimotor cortices during mental tasks.[27,28]

Independent component analysis (ICA) is a computational method for separating a multivariate signal into additive subcomponents supposing the mutual statistical independence of the non-Gaussian source signals.[29] It has been applied extensively to remove artifacts for the analysis of EEG.[30–32] The ICA is used for source localization of EEG/MEG data,[30] and to remove ballistocardiogram artifacts from simultaneously recorded EEG and fMRI data.[31] The blind source separation could show neurophysiolgically and neuroanatomically meaningful neuronal components without the assumption of prior physic models.[32] In this study, we propose an automatic method to eliminate the electrooculography (EOG) artifacts. Firstly, the ICA is adopted for the decomposition of independent components. A natural measure of similarity is then used to detect the EOG artifact. Finally, the EEG signals without EOG artifacts are obtained from the recovery of remaining independent components.

Continuous wavelet transform (CWT) gives a complete representation of EEG signals in the time-scale domain, so it can be applied for the

precise localization of ERP components in the time-scale domain.[33,34] However, it is time-consuming in terms of computation time. We utilize the CWT and Student's two-sample t-statistics to obtain the location of active segments in the time-frequency domain at off-line. Moreover, feature extraction in BCI work is an important topic that greatly affects classification accuracy. Fractal dimension in fractal geometry[35,36] is an effective fractal feature which has been applied to various biomedical signal analyses,[37–40] such as seizure onset detection in epilepsy[39] and EEG analyses of sleeping newborns.[40] Discrete wavelet transform (DWT) is an efficient and structured approach[41–48] to achieving multiresolution representation of ERP components.[49–55] We use the DWT and fractal dimension to extract features for subsequent classification. More specifically, we use DWT to decompose raw EEG signals into multiscale ones, and then obtain multiresolution fractal feature vectors (MFFVs) from wavelet data by calculating their fractal dimensions. In addition to multiscale characteristics, MFFVs also contain important fractal information in the time-scale space. As MFFVs have been shown to achieve promising results in BCI applications,[56] we have opted to use MFFVs as features in this study.

The support vector machine (SVM)[57] recognizing the patterns into two categories from a set of data is usually used for the analyses of classification and regression.[58,59] For example, the min-max modular SVM is adopted for the prediction of multiple subcellular locations of proteins.[58] No-reference video quality measurement is performed with support vector regression.[59] Since it can balance accuracy and generalization simultaneously, it is used for classification in this study.[58–60]

In most previous works, the segment of EEG signals was usually too long to apply in real-time analysis. The EEG signals, however, are not restricted to a predefined segment in asynchronous BCI works.[61,62] That is, the mental tasks can be performed at any desired moment. Moreover, the duration of finger lifting is much less than that of finger resting, so the classification accuracy has to be evaluated for these two different conditions respectively. In this study, EEG signals are analyzed continuously in the 1-s segment with the overlapping of 0.5-s window for each test. The present system first automatically eliminates the EOG artifacts. It then selects the active segment in time-frequency domain from each trial. Finally, MFFVs are extracted from selected active segments and used for the training of SVM classifier. A two-stage recognition architecture is proposed for the asynchronous mode in this study. The segment is recognized to be lifted or not in the first stage, while it is classified as left or right finger lifting at stage 2. If the segment is recognized as lifting in the first stage, then it can be further processed at stage 2. Finally, the results are evaluated with the statistical analysis.

This paper is organized as follows: In Sec. 2, materials and proposed methods are presented. Sections 3 and 4 describes experimental results and discussion, respectively. Finally, a conclusion is given in Sec. 5.

## 2. Materials and Methods

A flowchart of the process of EEG signal analysis for single-trial finger-lifting classification is illustrated in Fig. 1. The flowchart describes the detailed procedures of both off-line training and on-line processing. The process consists of several important steps, including automatic EOG artifact elimination, active segment selection, feature extraction and classification. During the off-line training, the data are trained to obtain the essential parameters for the corresponding steps in on-line processing. First, ICA combined with correlation coefficient is used to automatically eliminate the EOG artifacts. Next, based on the CWT and Student's two-sample t-statistics, active segment selection detects the location of the active segment in the time-frequency domain. We then extract features with the proposed modified fractal dimensions from the DWT data. Finally, the fractal features are used to train the parameters of the SVM classifier. After the off-line training, the trained parameters are utilized for the on-line processing. First, the EOG artifacts are removed automatically from the test data. The detected locations of active segment are then directly adopted. Next, the same process is used to extract multiresolution fractal features. At last, the SVM classifier with trained parameters is used for the classification.

Moreover, an asynchronous BCI work is simulated by continuously analyzing EEG signals in the 1-s segment with 0.5-s window overlapped for each test. We propose a two-stage recognition architec-

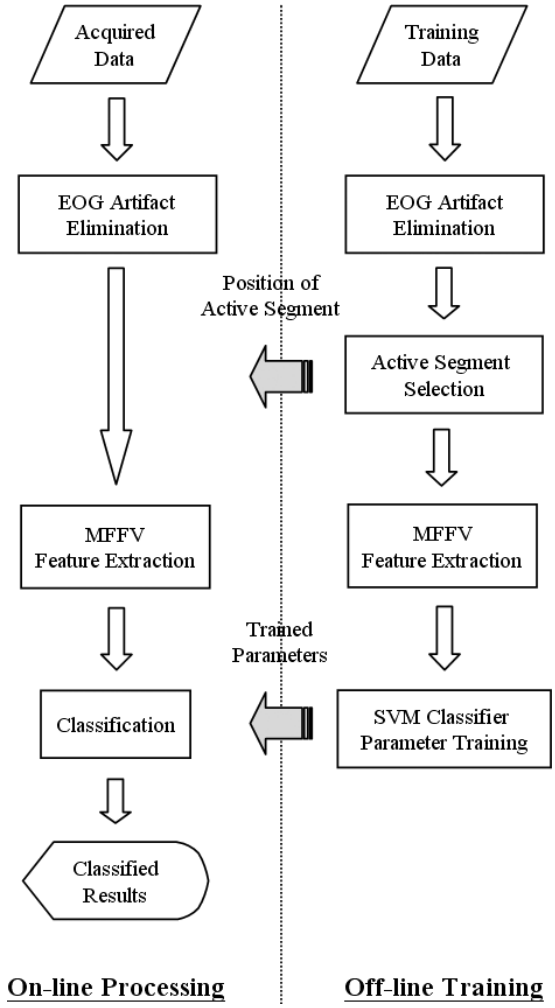



Fig. 1. Flowchart of proposed BCI system.

ture for the asynchronous mode. The segment is first recognized to be lifted or not in the first stage, and then it is classified as left or right finger lifting at stage 2, if the segment is recognized as lifting in the first stage.

### 2.1. *Data acquisition and mental task*

EEG signals were recorded from five untrained subjects (four males and one female, two left-handed and three right-handed) in a shielded room using 13 silver/silver chloride electrodes. As illustrated in Fig. 2, they included ten scalp EEG channels (C3, C5, FC3, C1, CP3, C4, C2, FC4, C6, and CP4), two EMG channels for monitoring left and right muscle activity, and one channel on the forehead to record possible EOG artifacts and eye blinks during the experiment.[63] All electrodes were referenced to the

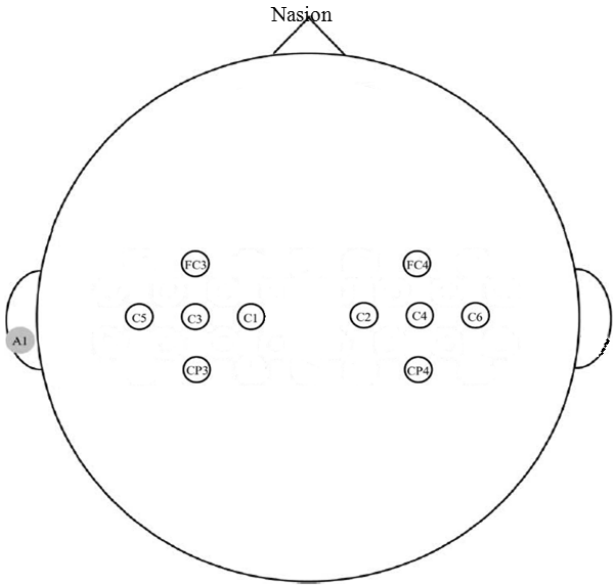



Fig. 2. EEG electrode positions with respect to the international 10–20 systems. Electrode positions marked with white circles are used to record EEG signals. All electrodes were referenced to the position A1 at the left earlobe.

A1 lead at the left earlobe. Before being sampled at a rate of 256 Hz, the EEG data were filtered by an analog band-pass filter with cutoff frequencies at 0.5 Hz and 100 Hz, and amplified by a multiple of ten thousand. During the experiments, each subject was asked to perform four trials, which included left finger lifting, resting, right finger lifting, and resting in sequence, in each test. Each trial was ten seconds in length; therefore, each test took forty seconds. For each lifting trial, the first 4 s was quiet and then an acoustic stimulus was given as a cue to signify the beginning of left or right finger lifting. At the same time, each subject was asked to execute finger lifting. We recorded sixty tests for each subject, and thus there were 240 trials for each subject. No trials were removed during the EEG data processing stage. Therefore, in total we used 1200 trials from five subjects in the finger movement experiments. Data segments representing three states were subsequently extracted from the experimental data set, that is, the above-mentioned left finger lifting, right finger lifting and finger resting. Data segments of the finger lifting were acquired from second −2 to second 2, where second 0 stands for the trigger of movement by detecting the peak EMG signal after linear envelope processing. (Only the data recorded between −2

and 2 s were considered to be event-related.) Data segments of finger resting were randomly acquired from 4-s windows within the finger resting sections of each trial. Data segments of all states were then used in the subsequent process.

### 2.2. *Automatic elimination of EOG artifacts*

ICA is a statistical method for transforming observed multidimensional mixed signals into components that are statistically as independent from others as possible. It is a method to resolve the blind source separation problem. That is, the source components are calculated with almost no advance knowledge of the nature of sources. In contrast to principal component analysis (PCA), which only ensures that output patterns are uncorrelated, ICA guarantees they are statistically independent. Statistical independence requires that all high-order correlations are zero, while decorrelation only minimizes the second-order statistics. The application of ICA for blind source separation of EEG signals is based on a reasonable assumption that EEG data acquired from multiple scalp electrodes are the linear combinations of temporally independent components.

Each test was arranged into an $m \times n$ matrix, where $m$ and $n$ represent channel number and sample points. That is, the $i$th row contains acquired signals from the $i$th channel, while the $j$th column means the sample at the $j$th time point across all channels. In this study, the FastICA algorithm[57] is adopted to eliminate the EOG artifacts due to the characteristic of its fast convergence. It removes the means of row vectors from the matrix and then uses a whitening procedure to transform the covariance matrix of zero-mean data into an identity matrix. Finally, it separates the whitened data into a set of components which are as mutually independent as possible.

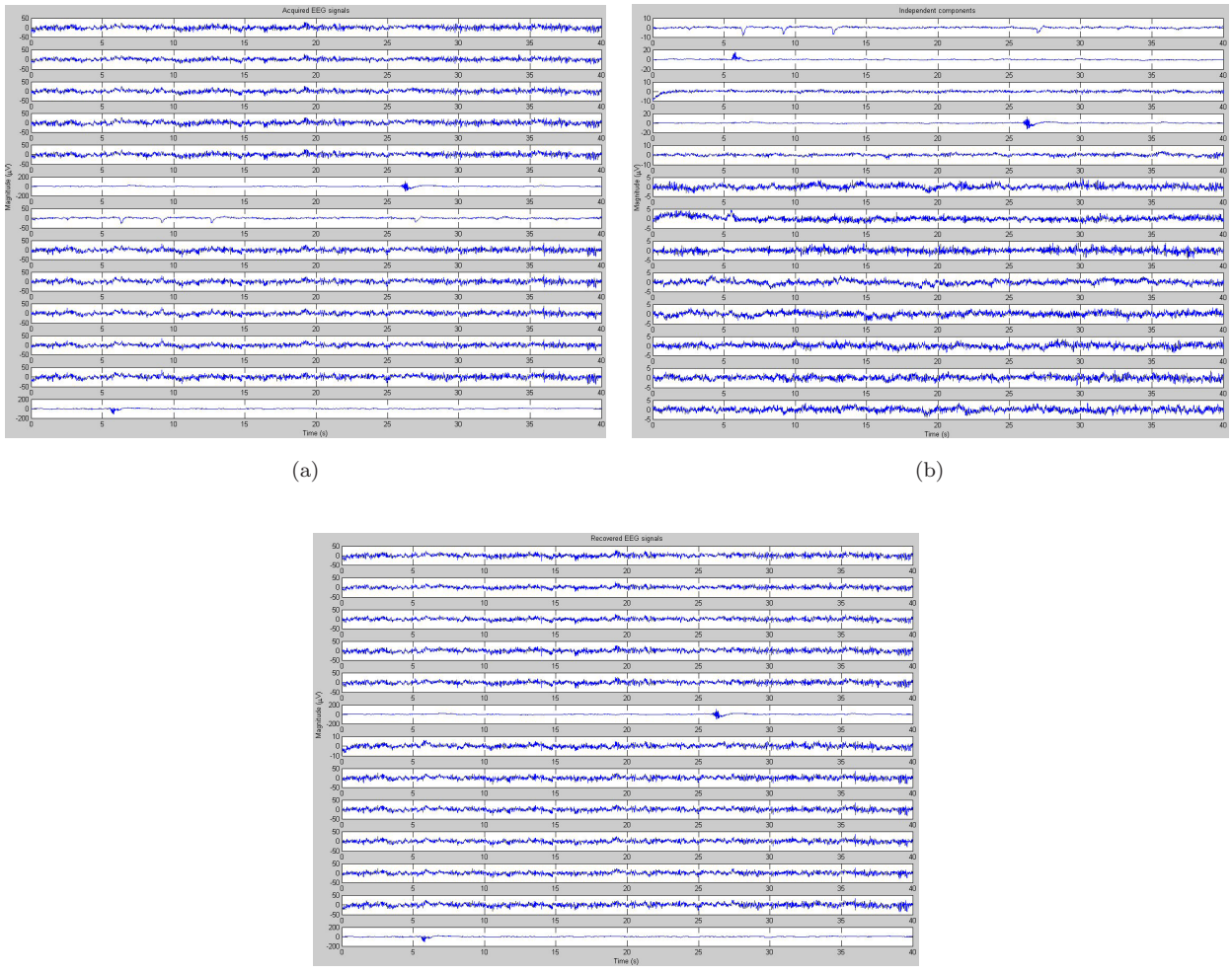
In this study, a natural measure of similarity — the absolute value of correlation coefficient between the EOG channel and estimated independent components — is proposed to automatically eliminate the EOG artifacts. The independent components with maximal similarity, which must be larger than a predefined threshold, are regarded as pure EOG artifacts. After the removal of EOG artifacts, the EEG signals without EOG artifacts are recovered from the remaining independent components. The procedure for automatic elimination of EOG artifacts is shown in Fig. 3. In Fig. 3(a), acquired (mixed) EEG signals are given. There are a total of 13 channels in the signals. The first five channels contain the signals of the C3 group (C3, C5, FC3, C1, and CP3), while the signals of the C4 group (C4, C2, FC4, C6, and CP4) are located in channels 8–12. The REMG and LEMG signals are situated in channels 6 and 13, respectively. The EOG artifact is acquired from channel 7. Figure 3(b) shows independent components after performing the FastICA algorithm. In this example, the maximum degree of similarity between the EOG channel and independent components is 0.8977, and this component is located at the first channel of independent components. The recovered EEG signals without EOG artifacts are shown in Fig. 3(c). We can observe that the EOG artifacts have clearly been eliminated and only EEG signals with low magnitude remain in the EOG channel. This is a reasonable result in the elimination of EOG artifacts.

### 2.3. *Data preprocessing*

Part of common non-EEG noise is different from EEG signals in both topographical and frequency characteristics. The mu and beta rhythms of the EEG are those components with frequencies distributed between 8–30 Hz and located over the sensorimotor cortex. EOG signals are maximal at low frequencies ($< 5$ Hz) and are prominently situated over the anterior head regions. Hence, an appropriate filtering method can increase the signal-to-noise ratio by enhancing EEG signals and reducing non-EEG noise at the same time. The Laplacian filter is a simple but effective filtering method.[64] It calculates the second derivative of the spatial voltage distribution for a selected electrode. It is a high-pass spatial filter that enhances localized activities and reduces background noise. This filter is achieved by subtracting the average potential of a set of surrounding electrodes from the electrode of interest,

$$V_i^{\mathrm{Lap}} = V_i - \frac{1}{N} \sum_{j \in S_i} V_j \tag{1}$$

where $V_i$ represents the potential between the $i$th electrode and the reference A1, and $S_i$ and $N$ stand for the set of electrodes surrounding the $i$th electrode and the number of surrounding electrodes, respectively. The distance between the selected electrode and its surrounding electrodes demonstrates




(a) (b)

(c)

Fig. 3. Procedure for automatic elimination of EOG artifacts. (a) Acquired EEG signals. (b) Independent components. (c) Recovered EEG signals without EOG artifacts.

the characteristic of Laplacian filtering which determines the greater the distance, the greater will be the insensitivity to highly localized potentials. The electrode of interest to each of its surrounding electrodes performed the Laplacian filtering, shown as follows,

$$V_{C3}^{\text{Lap}} = V_{C3} - \frac{1}{4}(V_{C5} + V_{FC3} + V_{C1} + V_{CP3})$$
$$V_{C4}^{\text{Lap}} = V_{C4} - \frac{1}{4}(V_{C2} + V_{FC4} + V_{C6} + V_{CP4}) \quad (2)$$

The 4-s window for (left and right) finger lifting was acquired from the finger-lifting trial with its center standing for the trigger of movement, while the 4-s window for finger resting was randomly acquired from the finger-resting trial.

### 2.4. *Active segment selection*

Before features were extracted, each trial included a redundant 4-s event-related window. In order to achieve efficient computation and high classification accuracy, it is important to reduce the length of this 4-s window by selecting only a 1-s active segment, which is defined as the most distinguishable in finger resting/lifting from the 4-s window. The better the selected active segments, the higher the classification accuracy.

The CWT in the time-frequency domain[65] is applied here for precise localization of the ERP components. In this study, active segment selection based on the CWT and Student's two-sample t-statistics is performed to obtain the optimal active segment in

the time-frequency domain. The CWTs of EEG signals for each trial, for finger resting or left or right finger lifting, are analyzed in the $C3$ and $C4$ channels, respectively,

$$W(j,k) = \int_R f(x) \frac{1}{\sqrt{j}} \psi\left(\frac{x-k}{j}\right) dx \tag{3}$$

where $\frac{1}{\sqrt{j}}\psi(\frac{x-k}{j})$ are the dilated and translated versions of the wavelet function $\psi(x)$ at scale $j$ and shift $k$, and $W(j,k)$ represents the CWT of the EEG signal $f(x)$. $W(j,k)$ is represented as a 2D time-scale plot that retains the scale separation of ERP components.[33,65–69]

However, the 2D time-scale plot $W(j,k)$ generated from the CWT is so noisy that it is difficult to accurately locate the active segment. The 2D anisotropic Gaussian filter is proposed to smooth the power spectrum and reduce the noise effects in the time-frequency domain,

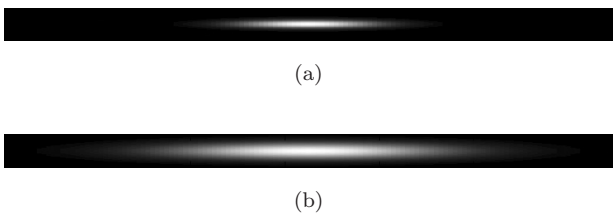
$$G_{ai}(x,y) = a \cdot e^{-\frac{1}{2}(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2})} \tag{4}$$

where $a = \frac{1}{\sum_x \sum_y G_{ai}(x,y)}$ is a normal coefficient. In addition, the best discriminant region is not situated at a single point but lies in a range of time and scale. We considered the time and scale range by convoluting the above-mentioned 2D anisotropic Gaussian filter (It is shown in Fig. 4 for different $\sigma$s), and found that by using a wider frequency range, a higher classification accuracy can be achieved from the acquired EEG signals than with a narrower one.[70] It is because the wider range contains all the mu and beta rhythmic components that are important for mental task classification. Therefore, we accumulate the power spectrum along the scale direction within the wide frequency range. The resulting profile represents the brain ERP responses for finger resting, left and right finger lifting.

Student's two-sample t-statistics between two of three distinct states are evaluated from the training data set in both $C3$ and $C4$ channels,

$$t(k) = \frac{|\mu_{s_1}(k) - \mu_{s_2}(k)|}{\sqrt{\left(\frac{(N_{s_1}-1)\cdot\sigma_{s_1}^2(k) + (N_{s_2}-1)\cdot\sigma_{s_2}^2(k)}{N_{s_1}+N_{s_2}-2}\right)\left(\frac{1}{N_{s_1}} + \frac{1}{N_{s_2}}\right)}} \tag{5}$$

where $s_1$ and $s_2$ belong to two different states, $N_s$ denotes the number of trials in state $s$, and $\mu_s(k)$ and $\sigma_s^2(k)$ stand for the mean and variance, respectively, of the profile in state $s$ from the training data set. The denominator in Eq. (5) indicates the pooled variance of two distinct states. $t(k)$ with different time $k$ can form a 1D profile with respect to time, but it contains different characteristics. The peaks in the profile $t(k)$ imply that there are local maximal differences between these two states in the time-scale domain, meaning they are best discriminated at a particular time and scale range. The t-statistics are used to select the optimal active segment. The active segment of length 1 s is then selected, with its center being the peak after the $C3$ and $C4$ channels are concatenated.



(a)

(b)

Fig. 4. 2D anisotropic Gaussian filter. (a) For small sigma. (b) For large sigma.

### 2.5. *Feature extraction*

Prior to classification, feature extraction is performed on the selected 1-s active segment rather than directly classifying the native EEG data without feature extraction. Feature extraction greatly affects the results of classification; accordingly, the better the extracted features, the higher the classification accuracy we can expect. In this study, we first band-pass filtered the active segments to the wide frequency range that contains all mu and beta rhythmic components using a Butterworth band-pass filter. We then performed the discrete wavelet transformed fractal dimension on the filtered active segments.

Multiresolution analysis decomposes a signal into numerous details at various resolutions, where each resolution represents a class of distinct physical characteristics within the signal. In other words, a signal is characterized with the formulation by decomposing it into sub-bands, and each sub-band can be treated individually based on its characteristics. Multiresolution representation of the filtered active segments is achieved by DWT. The active segment $A$ for each trial is represented in terms of the

DWT as

$$A(x) = \sum_{k=-\infty}^{\infty} S_J(k)2^{J/2}\phi(2^J x - k) + \sum_{j=1}^{J}\sum_{k=-\infty}^{\infty} D_j(k)2^{j/2}\psi(2^j x - k) \quad (6)$$

where the expansion coefficients are determined by

$$\begin{aligned} S_J(k) &= \langle A(x), 2^{J/2}\phi(2^J x - k)\rangle \\ D_j(k) &= \langle A(x), 2^{j/2}\psi(2^j x - k)\rangle \end{aligned} \quad (7)$$

where $S_J(k)$ and $D_j(k)$ represent the approximation and detail spaces of $A$, respectively, and $2^{J/2}\phi(2^J x - k)$ and $2^{j/2}\psi(2^j x - k)$ denote the dilated and translated versions of the scaling function $\phi(x)$ and wavelet function $\varphi(x)$, respectively. The active segment $A$ is then decomposed into individual subbands $S_J, D_J, \ldots,$ and $D_1$. Scale distribution of subbands for segment $A$ is listed in Table 1.

Fractal geometry provides a proper mathematical model to describe a complex shape that exists in nature with fractal features. The fractal dimension is one of the most popular fractal features, and we opted to use the fractal dimension due to the fact that it is relatively insensitive to signal scaling and shows a strong correlation with human judgment of surface roughness.[71] Several approaches for estimating the fractal dimension of signals and images have been presented.[71,72]

The differential box counting (DBC) method, which covers a wide dynamic range with a low computational complexity, is popular and frequently used. However, there are two significant faults in the original DBC method, which make the estimation of fractal dimension inaccurate. We have proposed a solution, namely the modified fractal dimension,[56] to resolve these issues. First, the original DBC method is based on the difference between the minimum and maximum rectangle numbers, and is easily disturbed by noise. As the standard deviation of the amplitude represents the dispersion of the signal, we use the standard deviation to replace the difference in rectangle numbers in the original DBC method. Second, DBC produces a stair-like function and may result in the underestimation of the fractal dimension. We overcome this fault by means of the floating calculation of the fractal dimension.

Table 1. Scale distribution of subbands for a segment.

| Scale Domain | | | |
|---|---|---|---|
| $A$ | | | |
| $S_1$ | | | $D_1$ |
| $S_2$ | | $D_2$ | |
| $S_3$ | $D_3$ | | |

The MFFVs are formed by concatenating various-scale fractal features, calculated from the active segment itself and all of its subbands using the modified fractal dimension. The MFFVs reflect the roughness of the EEG data at multiresolution.

### 2.6. *Classification*

Conventional neural networks can be difficult to establish because they usually need to choose an appropriate number of hidden layers and neurons to approximate the function in consideration of the desired accuracy. If the number in the network is too high, it may over-fit the training data and result in very poor generalization. The SVM, pioneered by Vapnik,[57] not only has a very steady theory of statistical learning, but guarantees to obtain the optimal decision function from a set of training data. The advantage of SVM is to balance the accuracy and generalization by maximizing the performance of network as well as minimizing the complexity of the learning machine at the same time. The main idea of SVM is to construct a hyperplane as the decision surface in such a way that the margin of separation between positive and negative examples is maximized. The SVM optimization problem is

$$\min_w \frac{1}{2}w^T w + C\sum_{i=1}^{N}\xi_i$$

subject to $\xi_i \geq 0, \forall\, i,$ and

$$d_i(w^T x_i + b) \geq 1 - \xi_i, \quad \forall\, i = 1, 2, \ldots, N \quad (8)$$

where $g(x) = w^T x + b$ represents the hyperplane, $w$ is the weighting vector, $b$ is the bias term, $x$ is the training vector with label $d, C$ is the weighting constant, and $\xi$ is the slack variable. It is then transformed into a convex quadratic dual problem. The discriminant function with optimal $w$ and $b, g(x) = w_o^T x + b_o$, posterior to the optimization form becomes

$$g(x) = \sum_{i=1}^{N}\alpha_i d_i K(x, x_i) + b_o \quad (9)$$

where $\alpha$ is a Lagrange multiplier and $K(x, x_i)$ is a kernel function. Generally, appropriate kernel functions[57] are the polynomial kernel function $K(x_i, x_j) = (x_i^T x_j + 1)^p$ and the radial basis function (RBF) kernel function $K(x_i, x_j) = \exp((-1/2\sigma^2)\|x_i - x_j\|^2)$.

### 2.7. *Simulation of an asynchronous BCI work*

An asynchronous BCI work is simulated by continuously analyzing EEG signals in 1-s segment each time with 0.5-s window overlapped forward for each test. A two-stage recognition architecture is proposed for these asynchronous applications. The segment is first recognized to be lifted or not in the first stage, and then it is classified as left or right finger lifting at stage 2 if the segment is recognized as lifting in the first stage.

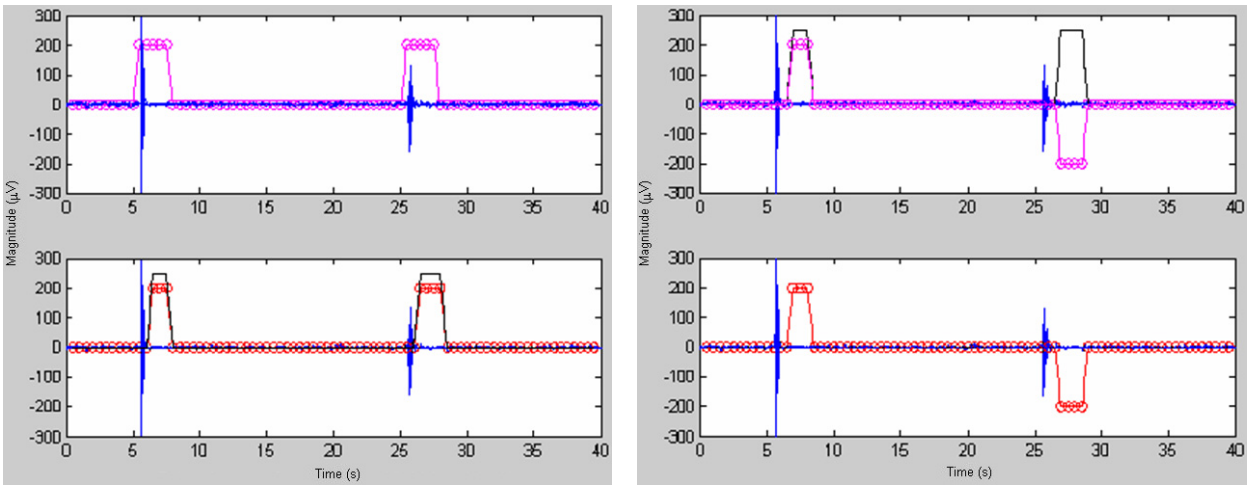
Prior to the recognition procedure, the parameters of SVM classifier have to be completely trained in advance. However, the count of segments for finger lifting (include left and right finger lifting) is much less than that for finger resting. In this study, the same count of segments is chosen from finger lifting and resting to desire to achieve the stability of parameter training of the SVM classifier for the first stage. After the classifier is trained, the procedure of analysis is performed by continuously recognizing 1-s segments every 0.5 seconds for each test, as shown in Fig. 5(a). In the second stage, the same count of segments acquired from left and right finger lifting is used to train classifier parameters. After the training, a similar analysis procedure is performed for each test, which is shown in Fig. 5(b). The procedure of continuous analysis for a test is shown in Fig. 5. Figure 5(a) shows the recognized results for finger lifting or resting in the first stage, while recognized results for left or right finger lifting in the second stage are illustrated in Fig. 5(b). For each sub-figure, the top figure shows the desired output based on the peak of EMG signals, whereas the recognized results of the proposed method are exhibited in the bottom figure.

## 3. Results

### 3.1. *Results of active segment selection*

The classification tests for EEG data are carried out using five-fold cross validation in this study. More specifically, the dataset for each subject is divided into five subsets, and the following procedure is repeated five times. One of five subsets is used as a test set and the other four are used as a training set each time. The average recognition rate is immediately evaluated across all five tests.

The results of active segment selection for each subject (subjects S1–S5) are shown in Fig. 6. The




(a) (b)

Fig. 5. Procedure of continuous analysis for a test. (a) Recognized results of finger lifting or resting in the first stage. (b) Recognized results of left or right finger lifting in the second stage. The top figure shows the desired output based on the peak of EMG signals, while the recognized results of proposed method are exhibited in the bottom figure.

original 4-s event-related window in each trial is too redundant to make the computation efficient. It is important to select most representative segment by reducing its original 4-s length to a 1-s segment. Moreover, we evaluate the performance of method selecting active segments from the 4-s event-related window. If the selected 1-s active segment contains essential information about ERP components, and can obtain better discrimination results than the original 1-s windows, it can greatly increase usability and achieve real-time applications. Moreover, the polynomial kernel function is used for the training of SVM classifiers to achieve stable results in the experiments. To verify the effectiveness of active segment selection, the 1-s data segments without and with active segment selection are compared for all three pairs of distinct states, including left finger lifting versus finger resting, right finger lifting versus finger

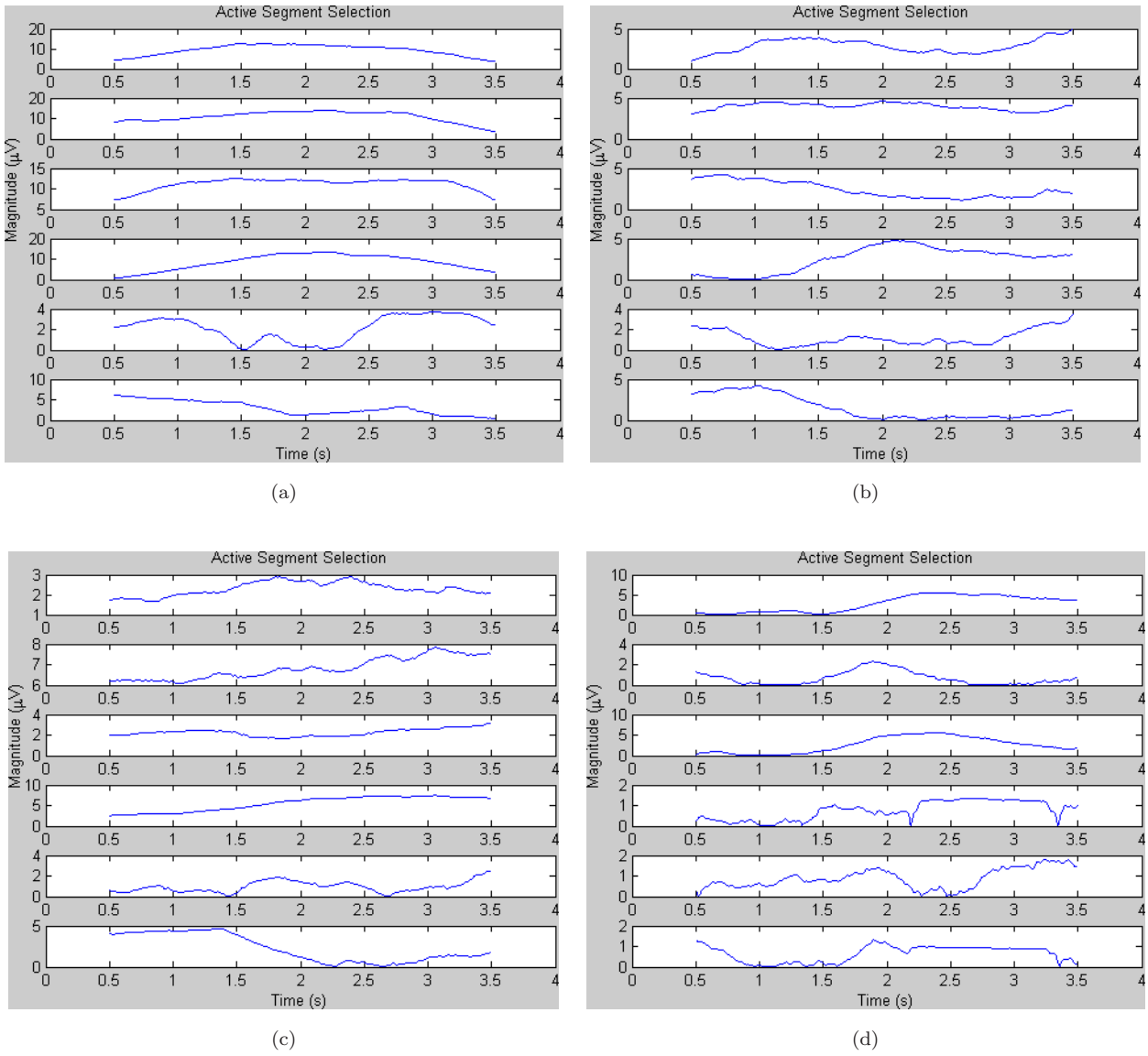
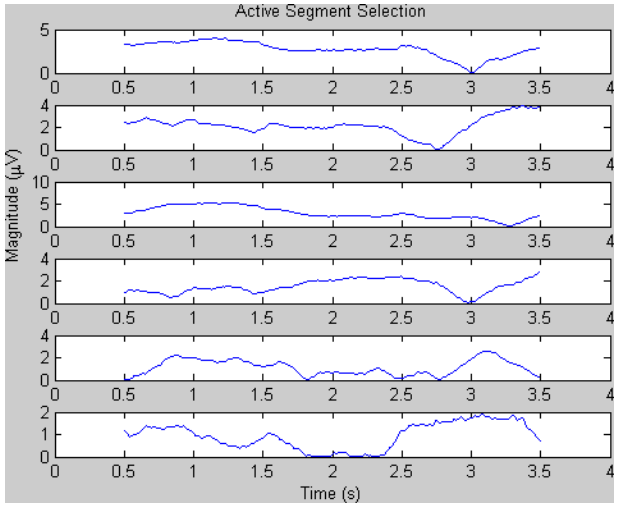



Fig. 6. Results of active segment selection. (a) For Subject S1. (b) For Subject S2. (c) For Subject S3. (d) For Subject S4. (e) For Subject S5. There are six plots within each figure. The first two plots represent channels $C3$ and $C4$ for the case of "left finger lifting versus finger resting", respectively; the center two plots represent channels $C3$ and $C4$ for the case of "right finger lifting versus finger resting", respectively; the last two plots represent channels $C3$ and $C4$ for the case of "left finger lifting versus right finger lifting", respectively.




(e)

Fig. 6. (*Continued*)

Table 2. Comparison of recognition rates between without and with active segment selection (ASS) for three pairs of distinct states.

| Subject | Left lifting *vs.* resting | | Right lifting *vs.* resting | | Left *vs.* right lifting | |
|---|---|---|---|---|---|---|
| | w/o ASS (%) | ASS (%) | w/o ASS (%) | ASS (%) | w/o ASS (%) | ASS (%) |
| S1 | 94.5 | 99.2 | 94.5 | 98.1 | 53.2 | 85.7 |
| S2 | 83.0 | 97.8 | 85.4 | 97.2 | 60.5 | 79.8 |
| S3 | 73.2 | 77.4 | 81.3 | 88.9 | 55.6 | 70.3 |
| S4 | 63.8 | 74.7 | 70.3 | 82.3 | 53.2 | 66.1 |
| S5 | 56.1 | 74.0 | 61.1 | 76.3 | 52.7 | 60.4 |
| Average | 74.1 | 84.6 | 78.5 | 88.6 | 55.0 | 72.5 |

resting, and left finger lifting versus right finger lifting. The results of comparison in recognition rates between them are listed in Table 2. In the case of "left finger lifting versus finger resting", the average recognition rates for without and with active segment selection are 74.1% and 84.6%, respectively. In the case of "right finger lifting versus finger resting", the average recognition rate is 78.5% without the active segment selection, while it increases to 88.6% after applying the selection of active segments. In the case of "left finger lifting versus right finger lifting", the average recognition rates for without and with active segment selection are 55.0% and 72.5%, respectively. The performance of the first two cases is improved by about 10%, whereas that of the last case is improved by 17.5%. It indicates that the approach of active segment selection can reserve the segment of essential information by discarding that of less essential information, including noise.

### 3.2. *Results of classifiers*

To validate the performance of the SVM, a common classifier, linear discriminant analysis (LDA), is used for the comparisons in all three pairs of distinct states. Table 3 lists the compared results in recognition rates between them. In the case of "left finger lifting versus finger resting", the average recognition rates for the LDA and SVM are 80.0% and 84.6%, respectively, whereas the average recognition rates for the LDA and SVM are 86.1% and 88.6%, respectively in the case of "right finger lifting versus finger resting". In the case of "left finger lifting versus right finger lifting", the average recognition rates for the

Table 3. Comparison of recognition rates between the LDA and SVM classifiers for three pairs of distinct states.

| Subject | LDA | | | SVM | | |
|---|---|---|---|---|---|---|
| | Left lifting *vs.* resting (%) | Right lifting *vs.* resting (%) | Left *vs.* right lifting (%) | Left lifting *vs.* resting (%) | Right lifting *vs.* resting (%) | Left *vs.* right lifting (%) |
| S1 | 99.0 | 97.3 | 83.5 | 99.2 | 98.1 | 85.7 |
| S2 | 85.6 | 92.2 | 79.1 | 97.8 | 97.2 | 79.8 |
| S3 | 72.4 | 87.1 | 65.4 | 77.4 | 88.9 | 70.3 |
| S4 | 71.8 | 80.5 | 64.2 | 74.7 | 82.3 | 66.1 |
| S5 | 71.1 | 73.5 | 57.8 | 74.0 | 76.3 | 60.4 |
| Average | 80.0 | 86.1 | 70.0 | 84.6 | 88.6 | 72.5 |



LDA and SVM are 70.0% and 72.5%, respectively. The performance of the first case is improved by 4.6%, whereas that of the last two cases is improved by about 2.5%. It denotes that the SVM has better performance in classification accuracy than the LDA.

### 3.3. *Performance evaluation of two-stage recognition architecture*

To evaluate the performance of proposed two-stage recognition architecture, a confusion matrix[73] is adopted in this study. The true positive (TP) is estimated by checking the number of points for desired output and recognition results, which are both marked as finger lifting, in the first stage. The positive represents the number of points that are located in active region of recognition results. The false positive (FP) is calculated by subtracting TP from the positive, while the false negative (FN) is calculated by subtracting TP from the number of points that are situated in the active region of real finger lifting. Finally, the true negative (TN) is obtained by calculating the point number of finger resting for both desired output and recognition results. We use sensitivity and specificity, which are statistical measures of the performance of a binary classification test, to estimate the recognition accuracy of finger lifting and resting in the first stage.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{10}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{11}$$

All the points that are recognized as lifting in stage 1 (including the part of false positive) are further classified as left or right finger lifting in the second stage. However, the part of false positive in stage 1 is generally regarded as incorrectly recognized results of the second stage, which would obviously increase the error rate of classification. The point number $N_p$ $(0 \leq N_p \leq \text{TP})$ of correct classification for the true positive is calculated from the desired output and recognition results, which obtain the same finger lifting, in the second stage. The recognized accuracy (RA) of stage 2 is defined as follows,

$$RA = \frac{N_p}{TP + FP} \tag{12}$$

In addition to RA, the receiver operating characteristics (ROC) curve is another popular approach to evaluate the performance of a method.[74] A ROC curve is obtained by plotting all sensitivity values (true positive fraction) on the $y$ axis against their equivalent specificity values (false positive fraction) on the $x$ axis for all available thresholds of classified results of test data. The AUC is usually taken to be an important index because it provides a single measure of overall accuracy that is not dependent upon a particular threshold. In this study, the performance of proposed recognition architecture in the second stage is evaluated by means of both the RA and AUC.

### 3.4. *Simulation results of an asynchronous BCI work*

All the acquired EEG data from these five subjects are used for the simulation of asynchronous BCI works to verify the practicability of proposed two-stage recognition architecture. In addition, each 40-s test for each subject takes the peak position of EMG signals into account to achieve a more accurate estimation of asynchronous BCI simulation.

Table 4. Simulation results of asynchronous BCI works for two-stage recognition architecture.

| Subject | 1st stage | | 2nd stage | |
|---|---|---|---|---|
| | Sensitivity (%) | Specificity (%) | RA (%) | AUC (%) |
| S1 | 98.6 | 99.6 | 88.9 | 0.87 |
| S2 | 97.3 | 99.1 | 83.7 | 0.79 |
| S3 | 82.7 | 93.4 | 72.1 | 0.73 |
| S4 | 77.0 | 88.3 | 64.3 | 0.62 |
| S5 | 75.9 | 89.0 | 61.5 | 0.65 |
| Average | 86.3 | 93.9 | 74.1 | 0.73 |



The simulation results of asynchronous BCI works for the two-stage recognition architecture are listed in Table 4. The average sensitivity and specificity in the first stage are 86.3% and 93.9%, respectively. In the second stage, the average RA and AUC are 74.1% and 0.73, respectively. It indicates that the two-stage recognition architecture can work well in the asynchronous BCI mode by effectively analyzing continuous EEG signals into three distinct states.

## 4. Discussion

### 4.1. *Automatic elimination of EOG artifacts*

Figure 3 illustrates the procedure for automatic elimination of EOG artifacts. Acquired EEG signals are shown in Fig. 3(a), the content of which contains C3 and C4 group, EOG artifacts, and LEMG and REMG signals. Figure 3(b) exhibits independent components after the ICA decomposition is performed. In this example, the independent component that is maximally similar to the EOG signal is located in the first channel, while the independent components with the maximal similarity to LEMG and RMEG signals are situated in channels 2 and 4, respectively. The remaining independent components are related to EEG signals and some other noise. After discarding the first independent component (including EOG artifacts), the recovered EEG signals are shown in Fig. 3(c). We can observe that EOG artifacts in channel 7 (EOG channel) have been completely eliminated. The remaining signals in this channel are EEG signals with low magnitude. This is reasonable since original EOG signals are mixed signals, which may contain some small EEG and EMG signals besides EOG artifacts. Moreover, the EOG artifacts that are mixed into the EEG signals of C3 and C4 group have also clearly been removed. This indicates that the method can increase the discriminant stability of EEG data by automatically eliminating the influence of EOG artifacts.

### 4.2. *Statistical analysis of active segment selection*

Active segment selection selects a 1-s active segment, which contains essential information about the ERP components, from a 4-s event-related window, hoping to obtain discerning results in a shorter segment. It can enhance the practicability of BCI work in real-time applications. Table 2 lists the comparison results of performance with and without active segment selection for three pairs of distinct states, including left finger lifting versus finger resting, right finger lifting versus finger resting, and left finger lifting versus right finger lifting. The performance in this experiment is evaluated by the recognition rate. In addition, two-way ANOVA and multiple comparison tests are performed to verify if the abovementioned results are significantly different or not. The results first indicate that the improvement of recognition rates for the case of "left finger lifting versus finger resting" is significant ($p$-value 0.0179), and is improved by 10.5% on average. The results then denote that there are also significant differences with and without active segment selection for the case of "right finger lifting versus finger resting" ($p$-value 0.0075). The average classification accuracy increases by 10.1%. Finally, the results show that the performance of case "left finger lifting versus right finger lifting" is significantly improved ($p$-value 0.0143) after active segment selection is performed. The recognition rate is enhanced by 17.5% on average. Accordingly, active segment selection can extract the segment of essential information from 4-s windows to further increase the computation efficiency.

### 4.3. *Explanation of performance difference for active segment selection*

Moreover, the results of selecting the location of active segments for each pair of different states are shown in Fig. 6. For the case of "left finger lifting versus finger resting," the recognition rates for subjects

S1 and S3 increase only by 4.7% and 4.2% respectively, while those for other subjects are improved by more than 10%. The performance difference among these five subjects will be explained from the results shown in Fig. 6. There are six plots within each figure for each subject. The first two plots represent channels $C3$ and $C4$ for the case of "left finger lifting versus finger resting," respectively; the middle two plots represent channels $C3$ and $C4$ for the case of "right finger lifting versus finger resting," respectively; and the last two plots represent channels $C3$ and $C4$ for the case of "left finger lifting versus right finger lifting" respectively. When observing the first two plots in Fig. 5, selected active segments for subjects S1 and S3 are close to the trigger position, which is situated at the center of the 4-s window, while those for other three subjects are considerably farther away from the trigger position. Accordingly, the use of active segment selection is more effective for subjects S2, S4 and S5 than for other two subjects in the case "left finger lifting versus finger resting." We can obtain a similar discussion for the case of "right finger lifting versus finger resting," which is illustrated in the middle two plots of Fig. 5. Finally, the last two plots in Fig. 5 are considered. The active segment for subject S5 is selected around the trigger position, but the four others are not. Thus, the recognition rate for subject S5 only slightly increases, while those for other subjects are improved substantially. Accordingly, active segment selection in this experiment is proven an effective and stable procedure to detect the location of a discerning segment.

### 4.4. *Two-stage recognition system for asynchronous BCI application*

Table 4 lists the results of simulation in asynchronous BCI works for the two-stage recognition system. All the EEG signals acquired from each subject are used for the simulation to validate the usability of proposed architecture. The sensitivity and specificity are defined and adopted to test the performance of "finger lifting or not/resting" in the first stage, while we use RA and AUC to judge the performance of "left or right finger lifting" in the second stage. The results denote that subjects S1 and S2 obtain higher sensitivity and specificity in the first stage and better RA and AUC in the second stage than others. That is, these two subjects have superior ability in the discriminant of "lifting or not" and "left or right lifting" after performing proposed methods. On the contrary, the subjects S4 and S5 obtain worse results in performance for the first and second stages. In other words, these two subjects are ordinary in recognizing the states of "lifting or not," whereas they are poor in discriminating between left and right finger lifting. Finally, the remaining subject, S3, in the first and second stages has ordinary results in the recognition of three different states. Accordingly, the two-stage recognition system is very suitable for the analyses of continuous EEG signals in asynchronous BCI work.

## 5. Conclusion

In this study, we have proposed a two-stage recognition system for continuous analysis of EEG signals. Firstly, the EOG artifacts are automatically eliminated by ICA and correlation coefficient. The active segment selection scheme is then used to reduce original 4-s event-related window to 1-s segment and increase the speed of feature extraction at the same time. Next, the MFFVs are obtained by modified fractal dimension from wavelet data. Finally, the SVM is used to classify the MFFVs. Asynchronous BCI works are simulated by continuously analyzing EEG signals in the two-stage recognition architecture. The signals are recognized to be lifted or not in the first stage, and are then classified as left or right finger lifting at stage two if they are recognized as lifting in the first stage. We also evaluate the performance of the proposed system with several statistical analyses. The results indicate that it is promising and has splendid potential in the applications of asynchronous BCI works.

In future studies, the proposed two-stage recognition system will be applied to continuously analyze motor imagery (MI) EEG/MEG data. Moreover, we will develop an incremental learning algorithm for real-world applications, where data varies over time.

## Acknowledgement

The author would like to express his sincere appreciation for the grant give him under shared facilities supported by the Program of Top 100 Universities Advancement, Ministry of Education, Taiwan.

## References


1. H. Lee, A. Cichocki and S. Choi, Nonnegative matrix factorization for motor imagery EEG classification, *International Journal of Neural Systems* **17**(4) (2007) 305–317.
2. B. Blankertz, F. Losch, M. Krauledat, G. Dornhege, G. Curio and K.-R Muller, The berlin brain-computer interface: Accurate performance from first-session in BCI-naïve subjects, *IEEE Trans. Biomedical Engineering* **55**(10) (2008) 2452–2462.
3. J. Lehtonen, P. Jylanki, L. Kauhanen and M. Sams, Online classification of single EEG trials during finger movements, *IEEE Trans. Biomedical Engineering* **55**(2) (2008) 713–720.
4. M. A. Kramer, F. L. Chang, M. E. Cohen, D. Hudson and A. J. Szeri, Synchronization measures of the scalp EEG can discriminate healthy from alzheimers subjects, *International Journal of Neural Systems* **17**(2) (2007) 61–69.
5. X. Wang, Y. Chen, S. Bressler and M. Ding, Evaluating causal relations among multiple neurobiological time series: Blockwise versus pairwise granger causality, *International Journal of Neural Systems* **17**(2) (2007) 71–78.
6. W. Y. Hsu, EEG-based motor imagery classification using neuro-fuzzy prediction and wavelet fractal features, *Journal of Neuroscience Methods* **189**(2) (2010) 295–302.
7. A. Montina, C. Mendoza and F. T. Arecchi, Role of refractory period in homoclinic models of neural synchronization, *International Journal of Neural Systems* **17**(2) (2007) 79–86.
8. M. Chiappalone, A. Vato, L. Berdondini, M. Koudelka and S. Martinoia, Network dynamics and synchronous activity in cultured cortical neurons, *International Journal of Neural Systems* **17**(2) (2007) 87–103.
9. D. E. Postnov, L. S. Ryazanova, R. A. Zhirin, E. Mosekilde and O. V. Sosnovtseva, Noise controlled synchronization in potassium coupled neural networks, *International Journal of Neural Systems* **17**(2) (2007) 105–113.
10. M. Chen, C. S. Jiang, Q. X. Wu and W. H. Chen, Synchronization in arrays of uncertain delay neural networks by decentralized feedback control, *International Journal of Neural Systems* **17**(2) (2007) 115–122.
11. N. Chakravarthy, S. Sabesan, K. Tsakalis and L. Iasemidis, Controling synchronization in a neural-level population model, *International Journal of Neural Systems* **17**(2) (2007) 123–138.
12. H. Osterhage, F. Mormann, T. Wagner and K. Lehnertz, Measuring the directionality of coupling: Phase versus state space dynamics and application to EEG time series, *International Journal of Neural Systems* **17**(3) (2007) 139–148.
13. I. Osorio and M. G. Frei, Seizure abatement with single DC pulses: Is phase resetting at play? *International Journal of Neural Systems* **19**(3) (2009) 149–156.
14. A. Shoeb, J. Guttag, T. Pang and S. Schachter, Non-invasive computerized system for automatically initiating vagus nerve stimulation following patient-specific detection of seizures or epileptiform discharges, *International Journal of Neural Systems* **19**(3) (2009) 157–172.
15. L. B. Good, S. Sabesan, S. T. Marsh, K. S. Tsakalis and L. D. Iasemidis, Control of synchronization of brain dynamics leads to control of epileptic seizures in rodents, *International Journal of Neural Systems* **19**(3) (2009) 173–196.
16. O. Faust, U. R. Acharya, L. C. Min and B. H. C. Sputh, Automatic identification of epileptic and background EEG signals using frequency domain parameters, *International Journal of Neural Systems* **20**(2) (2010) 159–176.
17. R. Acharya, E. C. P. Chua, K. C. Chua, L. C. Min and T. Tamura, Analysis and automatic identification of sleep stages using higher order spectra, *International Journal of Neural Systems* **20**(6) (2010).
18. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, Alzheimer's disease and models of computation: Imaging, classification, and neural models, *Journal of Alzheimer's Disease* **7**(3) (2005) 187–199.
19. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, Alzheimer's disease: Models of computation and analysis of EEGs, *Clinical EEG and Neuroscience* **36**(3) (2005) 131–140.
20. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Voxel-based morphometry in Alzheimer's patients, *Journal of Alzheimer's Disease* **10**(4) (2006) 445–447.
21. S. Ghosh-Dastidar and H. Adeli, Improved spiking neural networks for EEG classification and epilepsy and seizure detection, *Integrated Computer-Aided Engineering* **14**(3) (2007) 187–212.
22. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection, *IEEE Transactions on Biomedical Engineering* **55**(2) (2008) 512–518.
23. S. Ghosh-Dastidar and H. Adeli, Spiking neural networks, *International Journal of Neural Systems* **19**(4) (2009) 295–308.
24. S. Ghosh-Dastidar and H. Adeli, A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection, *Neural Networks* **22** (2009) 1419–1431.
25. M. Ahmadlou and H. Adeli, Fuzzy synchronization Likelihood with application to attention-deficit/hyperactivity disorder, *Clinical EEG and Neuroscience* **42**(1) (2011).

26. Z. Sankari, H. Adeli and A. Adeli, Intrahemispheric, Interhemispheric and distal EEG coherence in Alzheimer's disease, *Clinical Neurophysiology* (2011) doi:10.1016/j.clinph.2010.09.008.
27. G. R. Muller-Putz, D. Zimmermann, B. Graimann, K. Nestinger, G. Korisek and G. Pfurtscheller, Event-related beta EEG-changes during passive and attempted foot movements in paraplegic patients, *Brain Research* **1137** (2007) 84–91.
28. G. Pfurtscheller and dS. F. H. Lopes, Event-related EEG/MEG synchronization and desynchronization: basic principles, *Clin. Neurophysiol.* **110** (1999) 1842–1857.
29. A. Hyvarinen, J. Karhunen and E. Oja, *Independent Component Analysis*, John Wieley & Sons, Inc., New York (2001).
30. R. E. Banchs, H. Klie, A. Rodriguez, S. G. Thomas and M. F. Wheeler, A neural stochastic multiscale optimization framework for sensor-based parameter estimation, *Journal Integrated Computer-Aided Engineering* **14**(3) (2007) 213–223.
31. W. Nakamura, K. Anami, T. Mori, O. Saitoh, A. Cichocki and S. Amari, Removal of ballistocardiogram artifacts from simultaneously recorded EEG and fMRI data using independent component analysis, *IEEE Trans. Biomedical Engineering* **53**(7) (2006) 1294–1308.
32. A. C. Tang, B. A. Pearlmutter, M. Zibulevsky and S. A. Carter, Blind source separation of multichannel neuromagnetic responses, *Neurocomputing* **32–33** (2000) 1115–1120.
33. W. Y. Hsu and Y. N. Sun, EEG-based motor imagery analysis using weighted wavelet transform features, *Journal of Neuroscience Methods* **167**(2) (2009) 310–318.
34. Z. Zhou and H. Adeli, Time-frequency signal analysis of earthquake records using Mexican hat wavelets, *Computer-Aided Civil and Infrastructure Engineering* **18**(5) (2003) 379–389.
35. B. B. Mandelbrot, Fractal geometry of nature, *Freeman Press* San Francisco (1982).
36. Z. He, X. You, L. Zhou, Y. Cheung and Y. Y. Tang, Writer identification using fractal dimension of wavelet subbands in Gabor domain, *Integrated Computer-Aided Engineering* **17**(2) (2010) 157–165.
37. M. Ahmadlou, H. Adeli and A. Adeli, Fractality and a Wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder, *Journal of Clinical Neurophysiology* **27**(5) (2010) 328–333.
38. A. Ahmadlou, H. Adeli and A. Adeli, Fractality and a wavelet-Chao methodology for EEG-based diagnosis of Alzheimer's disease, *Alzheimer Disease and Associated Disorders* **24**(4) (2010).
39. B. R. Greene, S. Faul, W. P. Marnane, G. Lightbody, I. Korotchikova and G. B. Boylan, A comparison of quantitative EEG features for neonatal seizure detection, *Clinical Neurophysiology* **119**(6) (2008) 1248–1261.
40. P. D. Larsen, D. E. Elder, Y. C. Tzeng, A. J. Campbell and D. C. Galletly, Fractal characteristics of breath to breath timing in sleeping infants, *Respiratory Physiology and Neurobiology* **139**(3) (2004) 263–270.
41. A. Samant and H. Adeli, Feature extraction for traffic incident detection using wavelet transform and linear discriminant analysis, *Computer-Aided Civil and Infrastructure Engineering* **13**(4) (2000) 241–250.
42. A. Karim and H. Adeli, Fast automatic incident detection on urban and rural freeways using the wavelet energy algorithm, *Journal of Transportation Engineering* ASCE **129**(1) (2003) 57–68.
43. X. Jiang and H. Adeli, Wavelet packet-autocorrelation function method for traffic flow pattern analysis, *Computer-Aided Civil and Infrastructure Engineering* **19**(5) (2004) 324–337.
44. X. Jiang, S. Mahadevan and H. Adeli, Bayesian wavelet packet denoising for structural system identification, *Structural Control and Health Monitoring* **14**(2) (2007) 333–356.
45. H. Kim and H. Adeli, Wavelet hybrid feedback-LMS algorithm for robust control of cable-stayed bridges, *Journal of Bridge Engineering* ASCE **10**(2) (2005) 116–123.
46. H. Kim and H. Adeli, Hybrid control of smart structures using a novel wavelet-based algorithm, *Computer-Aided Civil and Infrastructure Engineering* **20**(1) (2005) 7–22.
47. S. Ghosh-Dastidar and H. Adeli, Neural network-wavelet micro-simulation model for delay and queue length estimation at freeway work zones, *Journal of Transportation Engineering* ASCE **132**(4) (2006) 331–341.
48. M. Rizzi, M. D'Aloia and B. Castagnolo, Computer aided detection of microcalcifications in digital mammograms adopting a wavelet decomposition, *Integrated Computer-Aided Engineering* **16**(2) (2009) 91–103.
49. W. Y. Hsu, W. F. P. Poon and Y. N. Sun, Automatic seamless mosaicing of microscopic images: Enhancing appearance with color degradation compensation and wavelet-based blending, *Journal of Microscopy-Oxford* **231**(3) (2008) 408–418.
50. T. Demiralp, J. Yordanova, V. Kolev, A. Ademoglu, M. Devrim and V. J. Samar, Time-frequency analysis of single-1sweep event-related potentials by means of fast wavelet transform, *Brain. Lang.* **66** (1999) 129–145.
51. H. Adeli, Z. Zhou and N. Dadmehr, Analysis of EEG records in an epileptic patient using wavelet transform, *Journal of Neuroscience Methods* **123**(1) (2003) 69–87.

52. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A wavelet-chaos methodology for analysis of EEGs and EEG sub-bands to detect seizure and epilepsy, *IEEE Transactions on Biomedical Engineering* **54**(2) (2007) 205–211.
53. S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection, *IEEE Transactions on Biomedical Engineering* **5**(9) (2007) 1545–1551.
54. H. Adeli, S. Ghosh-Dastidar and N. Dadmehr, A spatio-temporal wavelet-chaos methodology for EEG-based diagnosis of alzheimer's disease, *Neuroscience Letters* **444**(2) (2008) 190–194.
55. M. Ahmadlou and H. Adeli, Wavelet-synchronization methodology: A new approach for EEG-based diagnosis of ADHD, *Clinical EEG and Neuroscience* **41**(1) (2010) 1–10.
56. W. Y. Hsu, C. C. Lin, M. S. Ju and Y. N. Sun, Wavelet-based fractal features with active segment selection: Application to single-trial EEG data, *Journal of Neuroscience Methods* **163**(1) (2007) 145–160.
57. B. E. Boser, I. M. Guyon and V. Vapnik, A training algorithm for optimal margin classifiers, in *Proc. 5th Annu. Workshop on Computational Learning Theory* (1992) 144–152.
58. Y. Yang and B. L. Lu, Protein subcellular multi-localization prediction using a min-max modular support vector machine, *International Journal of Neural Systems* **20**(1) (2010) 13–28.
59. H. C. Lian, No-reference video quality measurement with support vector regression, *International Journal of Neural Systems* **19**(6) (2009) 457–464.
60. D. T. Lin and D. C. Pan, Integrating a mixed-feature model and multiclass support vector machine for facial expression recognition, *Integrated Computer-Aided Engineering* **16**(1) (2009) 61–74.
61. G. Townsend, B. Graimann and G. Pfurtscheller, Continuous EEG classification during motor imagery-Simulation of an asynchronous BCI, *IEEE Trans. Neural. Syst. and Rehab. Eng.* **12**(2) (2004) 258–265.
62. J. F. Borisoff, S. G. Mason, A. Bashashati and G. E. Brich, Brain-computer interface design for asynchronous control applications: Improvements to the LF-ASD asynchronous brain switch, *IEEE Trans. Biomed. Eng.* **51**(6) (2004) 985–992.
63. H. Jasper, The ten-twenty electrode system of the international federation, *Electroenceph. Clin. Neurophysiol.* **10** (1958) 371–375.
64. D. J. McFarland, L. M. McCane, S. V. David and J. R. Wolpaw, Spatial filter selection for EEG-based communication, *Electroencephalogr. Clin. Neurophysiol.* **103** (1997) 386–394.
65. W. Y. Hsu, Analytic differential approach for robust registration of rat brain histological images, *Microscopy Research and Technique* **74**(6) (2010) 523–530.
66. W. Y. Hsu, EEG-based motor imagery classification using enhanced active segment selection and adaptive classifier, *Computers in Biology and Medicine* (2011) DOI:10.1016/j.compbiomed.2011.05.014.
67. C. C. Chen, W. Y. Hsu, S. H. Chiu and Y. N. Sun, An efficient algorithm for point set registration using analytic differential approach, *IEICE Transactions on Information and Systems* **E93-D**(11) (2010) 3100–3107.
68. W. Y. Hsu, C. Y. Lin and W. F. Kuo, Unsupervised fuzzy C-means clustering for motor imagery EEG recognition, *International Journal of Innovative Computing, Information and Control* **7**(8) (2011).
69. W. F. Kuo, C. Y. Lin and W. Y. Hsu, Medical image segmentation using the combination of watershed and FCM clustering algorithms, *International Journal of Innovative Computing, Information and Control* **7**(9) (2011).
70. J. Müller-Gerking, G. Pfurtscheller and H. Flyvbjerg, Designing optimal spatial filters for single-trial EEG classification in a movement task, *Clin. Neurophysiol.* **110** (1999) 787–798.
71. A. P. Pentland, Fractal based description of natural scenes, *IEEE Trans. on Pattern Anal. Machine Intell.* **6** (1984) 661–674.
72. W. L. Lee, Y. C. Chen and K. S. Hsieh, Ultrasonic liver tissues classification by fractal feature vector based on M-band wavelet transform, *IEEE Trans. Medi. Imag.* **22** (2003) 382–392.
73. R. Kohavi and F. Provost, Glossary of terms, *Machine Learning* **30** (1998) 271–274.
74. J. A. Hanley and B. J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve, *Radiology* **143** (1982) 29–36.