UNSUPERVISED FUZZY C-MEANS CLUSTERING FOR MOTOR IMAGERY EEG RECOGNITION

Wei-Yen Hsu[1,4], Chi-Yuan Lin[2], Wen-Feng Kuo[3], Michelle Liou[1]
Yung-Nien Sun[4], Arthur Chih-Hsin Tsai[1], Hsien-Jen Hsu[5]
Po-Hsun Chen[6] and I-Ru Chen[7]

[1]Institute of Statistical Science
Academia Sinica
No. 128, Academia Road, Sec. 2, Nankang, Taipei 115, Taiwan
shenswy@stat.sinica.edu.tw; shen@csie.ncku.edu.tw

[2]Department of Computer Science and Information Engineering
National Chin-Yi University of Technology
No. 35, Lane 215, Sec. 1, Chung-Shan Road, Taiping City, Taichung 411, Taiwan

[3]Department of Medical Informatics
National Cheng Kung University Hospital
No. 138, Sheng Li Road, Tainan 701, Taiwan

[4]Department of Computer Science and Information Engineering
[5]Institute of Manufacturing Information and Systems
[6]Department of Electrical Engineering
National Cheng Kung University
No. 1, Ta-Hsueh Road, Tainan 701, Taiwan

[7]Department of Mathematics
Fu Jen Catholic High School
No. 270, Wu-Fon-Nan Road, Chia-I, Taiwan

Abstract. *In this study, an electroencephalogram (EEG) recognition system is proposed on single-trial motor imagery (MI) data. Fuzzy c-means (FCM) clustering is used for the unsupervised recognition of left and right MI data by combining with selected active segments and multiresolution fractal features. Active segment selection is used to detect active segments situated at most discriminable areas in the time-frequency domain. The multiresolution fractal features are then extracted by using modified fractal dimension from wavelet data. Finally, FCM clustering is used as the discriminant of MI features. The FCM clustering is an adaptive approach suitable for the clustering of non-stationary biomedical signals. Compared with several popular supervised classifiers, FCM clustering provides a potential for BCI application.*
Keywords: Brain-computer interface (BCI), Electroencephalogram (EEG), Motor imagery (MI), Fractal dimension (FD), Fuzzy c-means (FCM)

1. **Introduction.** The brain-computer interface (BCI) providing analternative channel to directly transmit messages to computers from the human brain by analyzing the brain's mental activities [1-6] is a new communication system. BCI systems based on the single-trial analysis of motor imagery (MI) electroencephalographic (EEG) signals have grown rapidly in the last decade [2]. It focuses on the recognition of left and right MIs using event-related brain potentials (ERP), which reveal that they possess special characteristics of event-related desynchronization (ERD) and synchronization (ERS) in mu and beta rhythms over the sensorimotor cortices during MI tasks [7-9].

The continuous wavelet transform (CWT) gives a highly redundant representation of EEG signals in the time-scale domain. It can be applied for the precise localization of ERP components [10], and however, it has heavy computation cost. The CWT and Student's two-sample t-statistics are adopted to off-line detect the location of active segments through the time-frequency domain. Besides, feature extraction is a substantial topic that greatly affects the classification success. That is, the better the extracted features are, the higher the classification accuracy we can expect. The DWT and fractal dimension are used to extract fractal features in multiscale for classification. Since the DWT is strong in selecting features from multiresolution, it is an efficient and structured approach to ERP representation [11].

Fractal dimension is one of the most common fractal features, which are used to express fractal geometry [12] that provides a proper mathematical model to describe complex and irregular shapes existing in nature. It has been applied to all kinds of biomedical signal analyses, such as seizure onset detection in epilepsy [13,14] and image coding [15,16]. In this study, the native EEG signals are decomposed into multi-scale bands by the DWT. The multiresolution fractal feature vectors (MFFVs) are extracted from multiscale EEG signals by calculating their fractal dimensions. That is, MFFVs simultaneously contain important multiscale characteristics and fractal information in the time-scale space. The MFFVs have been proven to achieve promising results in BCI classification [17], so they are opted to use as features again in this study.

To recognize MI EEG data by supervised classifiers has been applied in many BCI systems. Some of them, such as linear discriminant analysis (LDA) [18], multilayer perceptron (MLP) [19] and support vector machine (SVM) [20,21], are very popular and usually used for the discriminant. Moreover, the classifiers are supervised, whose parameters need to be trained in advance before on-line applications. Fuzzy c-means clustering that is an unsupervised approach partitions a collection of feature vectors into a number of subgroups based on minimizing the trace of a within-cluster scatter matrix [22,23]. EEG data are non-stationary and their characteristics vary as time, so the classification accuracy of EEG data with unsupervised FCM clustering may lead to better results than that with conventional supervised classifiers. Several properties of FCM clustering are described in the following. The FCM is not essential to label data in advance before training if it is used as a classifier, i.e., we can discriminate the data without *a priori* knowledge. The FCM is able to make flexible partitions of a finite data set. The FCM is an adaptive approach suitable to recognize non-stationary biomedical signals, such as MI EEG data [24]. It is the reason the MI EEG data are recognized by the FCM clustering in this study.

In addition, to evaluate the performance of FCM clustering, some supervised classifiers, such as LDA, MLP and SVM, are used for comparison. The experimental results show satisfactory results for the FCM. The results also indicate that it has great potential for unsupervised classifiers in BCI applications. This paper is organized as follows: in Section 2, MI EEG data description and analyses are presented; Section 3 describes experimental results and discussions on single-trial data; finally, conclusions are given in Section 4.

2. **Materials and Methods.** A flowchart of EEG signal analysis for single trial MI classification is illustrated in Figure 1. The process chiefly consists of three steps: active segment selection, feature extraction and fuzzy c-means clustering. Based on the CWT and Student's two-sample t-statistics, active segment selection detects the location of active segments in the time-frequency domain. We then extract feature vectors with proposed modified fractal dimension from the DWT data. Finally, we use fuzzy c-means

clustering to classify the fractal feature vectors into two classes, left and right motor imagery, without supervision.
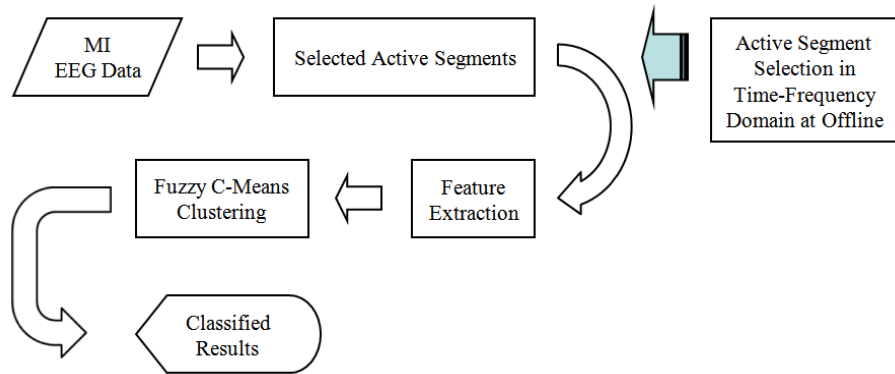

FIGURE 1. Flowchart of proposed BCI system

2.1. **Data description.** The EEG data was recorded by the Graz BCI group [25-27] from three subjects during a feedback experimental recording procedure. The task was to control a bar by means of imagining left or right hand movements. The order of left and right cues was random. The data was recorded on three subjects – the first subject, S1, performed 280 trials, while other two, S2 and S3, each performed 320 trials. The length of each trial was 8-9s. An example of a trial is given in Figure 2. The first 2s of the trial were quiet, then an acoustic stimulus indicated the beginning of the experiment at $t = 2$s, and a fixation cross + was displayed for 1s; then, at $t = 3$s, an arrow (left or right) was displayed as a cue. (Only the data recorded between 3 and 8s were therefore considered to be event-related.) At the same time, each subject was asked to move a bar by imagining left or right hand movements according to the direction of the cue. The recordings were made using a g.tec amplifier and Ag/AgCl electrodes; all signals were sampled at 128Hz and filtered between 0.5 and 30Hz. Two bipolar EEG channels were measured using two electrodes positioned at 2.5cm anterior and posterior to positions $C3$ and $C4$ according to the international standard 10/20 system [28], as shown in Figure 3.

2.2. **Active segment selection.** Before features were extracted, each trial included a redundant 5-s event-related window. In order to achieve efficient computation and high classification accuracy it is important to reduce the length of this 5-s window by selecting only a 1-s active segment, which is defined as the most representative in MI from the 5-s window. The better the selected active segments, the higher the classification accuracy.

The CWT in the time-frequency domain [29] 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, which was of either left or right MI, are analyzed in the $C3$ and $C4$ channels, respectively:

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

where $\frac{1}{\sqrt{j}} \psi\left(\frac{x-k}{j}\right)$ is 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 [29].
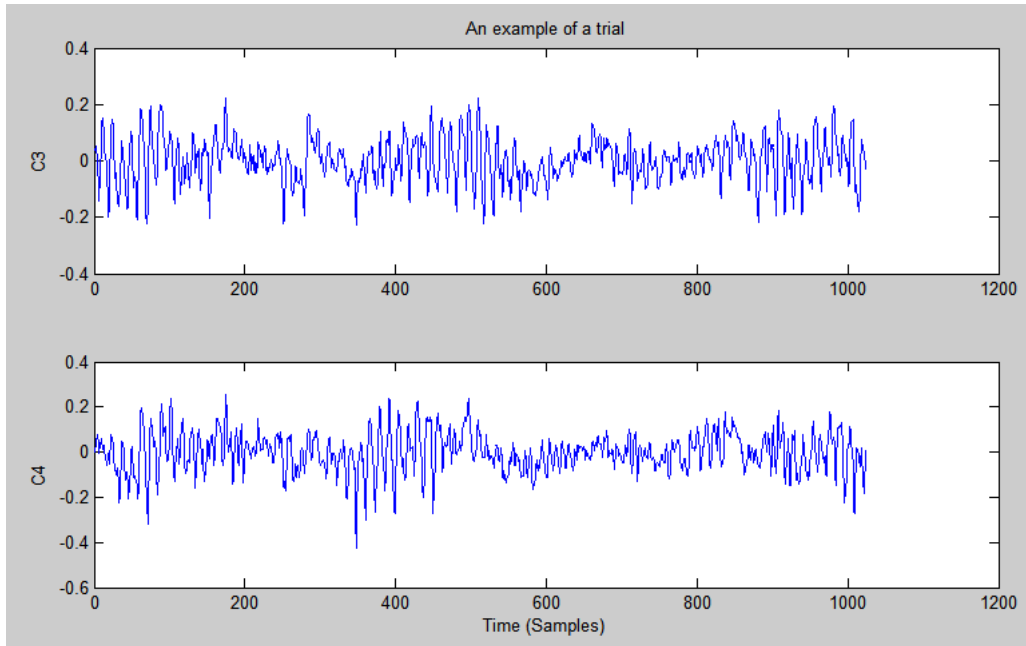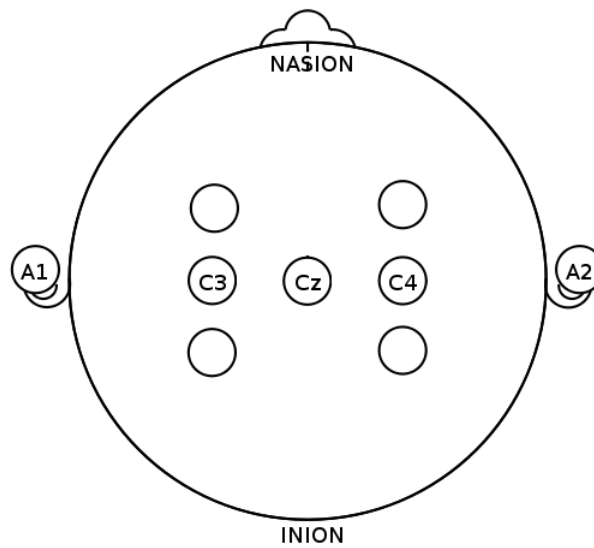
FIGURE 2. Example of a trial

FIGURE 3. Electrode position

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. We smooth the power spectrum and reduce the noise effects in the time-frequency domain. In addition, the best discriminant region is not situated at a single point but lies in a range of time and scale. Higher classification accuracy can achieve from the acquired EEG signals by using a wider frequency range than a narrower one [30]. It is because the wider range contains all the mu and beta rhythmic components that are important for mental task classification. Hence, the power spectrum is accumulated along the scale direction within the wide frequency range. The resulting profile represents the brain ERP responses for left and right MI. Student's two-sample t-statistics between the left and right MI are evaluated from the training data set in both $C3$ and $C4$ channels:

$$t\left(k\right) = \frac{\left|\mu_L(k) - \mu_R(k)\right|}{\sqrt{\left(\frac{(N_L-1)\cdot\sigma_L^2(k)+(N_R-1)\cdot\sigma_R^2(k)}{N_L+N_R-2}\right)\left(\frac{1}{N_L} + \frac{1}{N_R}\right)}} \tag{2}$$

where L and R represent left and right MI, respectively, 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 Equation (2) indicates the pooled variance of left and right MI. $t\left(k\right)$ with different time k can form a 1D profile with respect to time, but it contains different characteristics. The peaks in the profile $t\left(k\right)$ imply that there are local maximal differences between left and right MI in the time-scale domain, meaning the left and right MI 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 1s is then selected, with its center being the peak after the $C3$ and $C4$ channels are concatenated.

2.3. **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 transformand 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 discrete wavelet transform (DWT). The active segment A for each trial is represented in terms of the DWT as:

$$A\left(x\right) = \sum_{k=-\infty}^{\infty} S_J\left(k\right) 2^{J/2}\phi\left(2^J x - k\right) + \sum_{j=1}^{J}\sum_{k=-\infty}^{\infty} D_j(k)2^{j/2}\psi\left(2^j x - k\right) \tag{3}$$

where $S_J\left(k\right)$ and $D_j(k)$ represent the approximation and detail spaces of A, respectively, and $2^{J/2}\phi\left(2^J x - k\right)$ and $2^{j/2}\psi\left(2^j x - k\right)$ 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 $S_J, D_J, \ldots, D_1$.

Fractal geometry provides a proper mathematical model to describe a complex shape that exists in nature with fractal features. Fractal dimension (FD) is one of the most popular fractal features, and we opted to use the FD due to the fact that it is relatively insensitive to signal scaling and shows a strong correlation with human judgment of surface roughness [31]. Several approaches for estimating the fractal dimension of signals and images have been presented [31-33].

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 [17], 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 fractal dimension. We overcome this fault by means of the floating calculation of fractal dimension.

The multiresolution fractal feature vectors (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 MI data at multiresolution.

2.4. **Fuzzy c-means clustering.** Clustering is a process for classifying training samples in such a way that samples within a cluster are more similar to one another than samples belonging to different clusters. Similarity measures employed to classify samples depend on the object characteristics, e.g., distance, vector and entropy. Many clustering approaches have also been demonstrated such as the hard clustering algorithm [34,35] and the soft (fuzzy) clustering algorithm, each of which has its own special characteristics. There are many applications in clustering strategy [36,37]. The fuzzy clustering method assigns the sample with a number, m, between zero and one described as a membership function. In this paper, fuzzy clustering method is to classify the feature vectors extracted from the original EEG data, and recognize complicated brain mental tasks, such as left and right motor imagery.

The FCM clustering algorithm was first introduced by Dunn [38], the related formulations and algorithms were extended by Bezdek [39]. The FCM approach, like the conventional clustering techniques, minimizes an objective function in the least squared error sense. For class number c ($c \geq 2$), sample number n and fuzzification parameter m ($1 \leq m < \infty$), the algorithm chooses $u_i : X \to [0,1]$ so that $\sum_i u_i = 1$ and $w_j \in R^d$ for $j = 1, 2, \ldots, c$ to minimize the objective function.

$$J_{FCM} = \frac{1}{2} \sum_{j=1}^{c} \sum_{i=1}^{n} (u_{i,j})^m \|x_i - w_j\|^2, \tag{4}$$

where $u_{i,j}$ is the value of jth membership grade on ith sample x_i. The cluster centroids $w_1, \ldots, w_j, \ldots, w_c$ can be regarded as prototypes for the clusters represented by the membership grades. For the purpose of minimizing the objective function, the cluster centroids and membership grades are chosen so that a high degree of membership occurs for samples close to the corresponding cluster centroids. The steps of the FCM algorithm are listed below.

Step 1. Initialize the cluster centroids w_j ($2 \leq j \leq c$), fuzzification parameter m ($1 \leq m < \infty$) and the value $\varepsilon > 0$. Give a fuzzy c-partition $U^{(0)}$ and $t = 1$.

Step 2. Calculate the membership matrix $U^{(t)} = [u_{i,j}]$ by

$$u_{i,j} = \frac{\left(\frac{1}{(d_{i,j})^2}\right)^{1/(m-1)}}{\sum_{l=1}^{c} \left(\frac{1}{(d_{l,i})^2}\right)^{1/(m-1)}}, \tag{5}$$

where $d_{i,j}$ is the Euclidean distance between the training sample x_i and the class centroid w_j.

Step 3. Update the class centroids by

$$w_j = \frac{1}{\sum_{i=1}^{n} (u_{i,j})^m} \sum_{i=1}^{n} (u_{i,j})^m x_i \text{ for every } j \tag{6}$$

Step 4. Compute $\Delta = \max\left(\left|U^{(t+1)} - U^{(t)}\right|\right)$. If $\Delta > \varepsilon$, $t = t + 1$ and go to Step 2; otherwise, stop the process.

(a) Mean CWT for MI data

(b) Variance of CWT for MI data

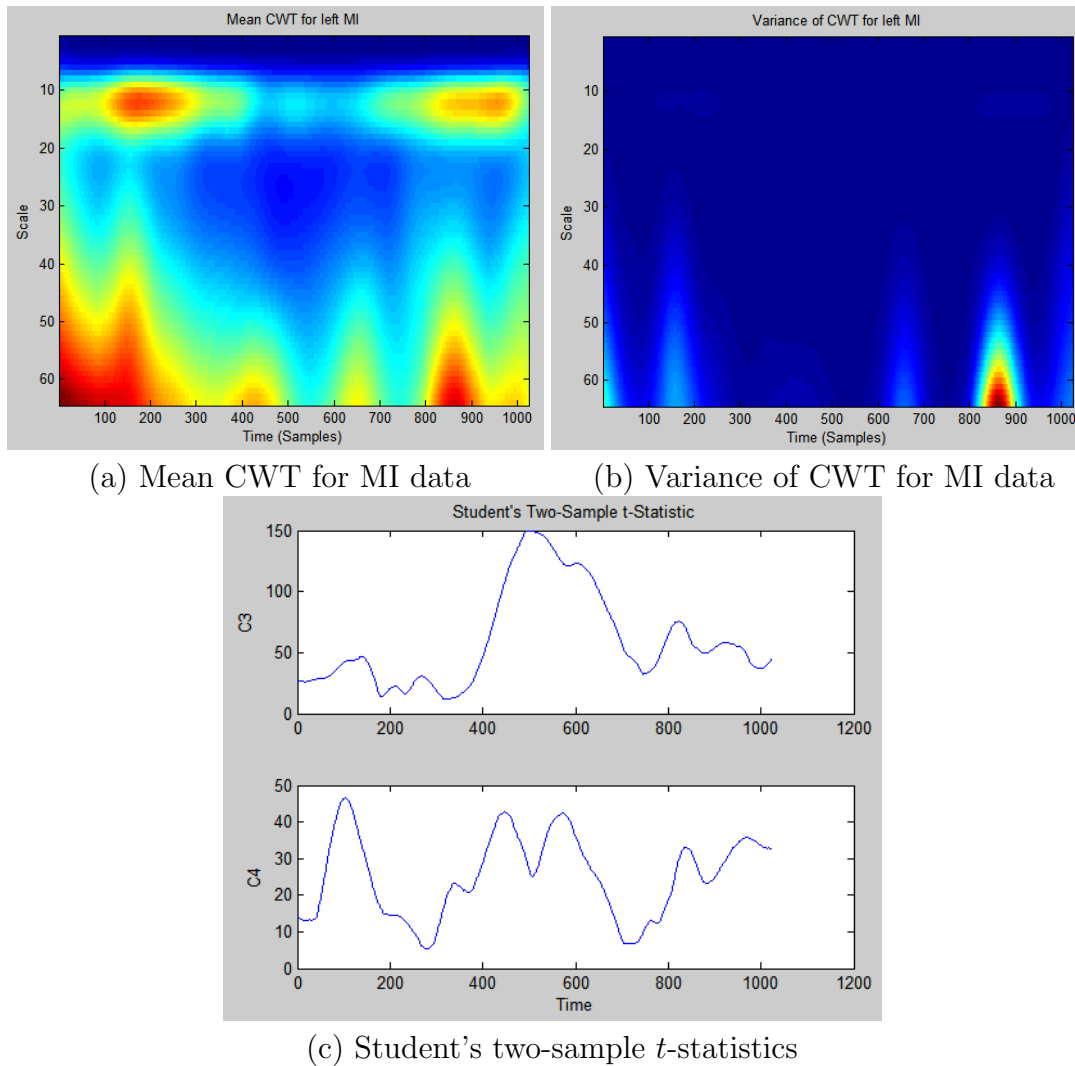(c) Student's two-sample t-statistics

FIGURE 4. Procedure for active segment selection

3. **Results and Discussion.** We demonstrate the process of active segment selection in Figure 4, which shows the detailed procedure for selecting the active segment in the time-frequency domain from the 5-s event-related window. The original 5-s event-related window in each trial is too redundant to make the computation efficient and the classification accurate. It is important to select most representative segment by reducing its original 5-s length to a 1-s window. In Figure 4(a), we show the mean CWT of the MI data from a subject; the variance of the CWT for the same subject is shown in Figure 4(b). After accumulating the power spectrum along the scale direction within the wide frequency range, we can then obtain Student's two-sample t-statistics, as illustrated in Figure 4(c).

The classification tests for MI data are in this study carried out using ten-fold cross validation. More specifically, the dataset for each subject is divided into ten subsets, and the following procedure is repeated ten times. Each time, one of the ten subsets is used as the test set and the other nine are used as training set. The average recognition rate is immediately evaluated across all ten folds. The classification accuracy of FCM clustering

for MFFV features is compared between without and with active segment selection to evaluate the efficacy of proposed active segment selection. In this experiment, in the case of "without active segment selection", the features are extracted from the 5-s event-related window. The comparisons of classification accuracy are listed in Table 1. The results show that the average accuracy for without active segment selection is 71.4%, while that for with active segment selection increases to 83.1%. Accordingly, it indicates that the proposed segment-selection method can significantly improve the overall performance in MI recognition. In other words, after using the proposed MFFV features together with active segment selection, classification accuracy is greatly improved for all subjects.

Furthermore, we perform two-way ANOVA to validate whether the use of active segment selection are significantly different or not. The results show that the classification improvement is significant (p-value 0.0511). In other words, active segment selection can extract essential information from event-related windows to further improve the efficiency and accuracy.

An unsupervised FCM method is proposed for left and right MI classification from single-trial EEG data in this study. The FCM algorithm is implemented to recognize left and right MI data without any training in advance. As FCM clustering is an unsupervised approach that partitions a collection of feature vectors into a number of subgroups based on minimizing the trace of a within-cluster scatter matrix, it is adopted to automatically classify MFFVs into two clusters, left and right MI. Moreover, EEG data are non-stationary and their inherent characteristics vary with time. The classification of MI EEG data using unsupervised FCM clustering may result in a better generalization performance than conventional supervised classifiers. Some attractive properties of FCM clustering are summarized as follows: first, FCM clustering is an unsupervised method, discriminating the data without labeling them before training; second, FCM clustering is capable of making flexible partitions of a finite data set; and third, FCM clustering is a robust approach suitable for the classification of non-stationary biomedical signals. Therefore, we use FCM clustering for the classification of MI EEG data in this study. We compare the performance of fuzzy c-means clustering with three supervised classifiers that are widely-used and have been shown to provide good results in BCI studies. The first classifier is linear discriminant analysis (LDA) [18], used to find the direction in the feature space along which the distance of the means relative to the within-class scatter reaches a maximum. The second classifier is a multilayer perceptron (MLP) with one hidden layer and sigmoid activation functions [19]. The third classifier is a nonlinear support vector machine (SVM) using a radial basis function (RBF) kernel [20], which is used to classify the feature patterns based on minimum error risk criterion. The optimal values for the parameters, weights and biases of all supervised classifiers were estimated using the aforementioned ten-fold cross validation methodology.

TABLE 1. Comparison of classification accuracy for FCM between without and with active segment selection

Classification Accuracy (FCM)	Subject 1	Subject 2	Subject 3	Average	Two-way ANOVA (p-value)
Without Active Segment Selection	74.3%	71.6%	68.4%	71.4%	0.0511
With Active Segment Selection	88.1%	77.8%	83.3%	83.1%	

TABLE 2. (a) Comparison of classification accuracy among LDA, MLP, SVM and FCM methods; (b) two-way ANOVA and multiple comparison tests for classification accuracyamong LDA, MLP, SVM and FCM methods

Classification Accuracy [%]	Subject 1	Subject 2	Subject 3	Average
LDA	82.9	74.3	81.7	79.6 ± 3.8
MLP	86.3	78.9	82.4	82.5 ± 3.0
SVM	86.7	79.6	84.6	83.6 ± 3.0
FCM	88.1	77.8	83.3	83.1 ± 4.1

(a)

$\mu_i - \mu_j$, p-value	μ_1	μ_2	μ_3
μ_2	0.1284		
μ_3	0.0293	0.1869	
μ_4	0.0808	0.5972	0.6261

(b)

An experiment was performed to assess the performance of the unsupervised fuzzy c-means clustering method. Table 2(a) shows the comparisons of classification accuracy among the LDA, MLP, SVM and FCM methods. All classifiers were used with the same selected active segments in this table; in other words, the listed values demonstrate only the deviations of performance among different classifiers. For Subject 1, the best recognition rate was achieved using FCM as the classifier, while the classification obtained using LDA had the worst recognition rate, the difference between the best and worst recognition rates being 5.2%. Moreover, using SVMas the classifier obtains the second highest recognition rate, which was only 1.4% less than the best one. For Subject 2, the best recognition rate was obtained using SVM as the classifier, while the classification obtained using LDA had the worst recognition rate. For Subject 3, the best and worst recognition rates were found to be the same as those of Subject 2. Overall, the best average recognition rate (83.6%) was obtained using SVM, while using LDA yielded the worst average recognition rate (79.6%), the difference being 4.0%. Moreover, FCM obtained the second highest recognition rate (83.1%), which is only slightly less (0.5%) than that of SVM.

In addition, two-way ANOVA and multiple comparison tests are performed again to test if the classification accuracy among classifiers is significantly different. The tested results are listed in Table 2(b). It denotes that there is significant difference between LDA and FCM (p-value 0.0808). Overall, FCM has a potential in BCI applications.

We observe that classification with SVM gives the best results for two subjects, and classification with FCM leads to the best results for one subject. This can be explained by the fact that FCM clustering is appropriate for the classification of non-stationary MI EEG signals because it is a robust discriminator. Traditional supervised classification methods, such as MLP, have suffered from difficulties in generalization, often obtaining models that over-fit the training data, which is why we chose to apply FCM clustering in the classification of MI EEG data in this study. In addition, FCM clustering is considered a good classifier because it is an unsupervised clustering-based method. For BCI work, FCM clustering does not need to train the parameters of the classifier in advance, and it is suitable for dealing with non-stationary biomedical data sets. In summary, FCM is

suitable for mental task classification. The difference between the classification results of FCM and SVM is small for all subjects; however, the computation time with SVM is larger than that with FCM. If we take the computation time into account, FCM is likely the most promising classifier among the four tested in this study in the classification of MI EEG data.

TABLE 3. (a) Comparison of AUC among LDA, MLP, SVM and FCM methods; (b) two-way ANOVA and multiple comparison tests for AUC among LDA, MLP, SVM and FCM methods

AUC	Subject 1	Subject 2	Subject 3	Average
LDA	0.79	0.70	0.75	0.75 ± 3.1
MLP	0.81	0.79	0.80	0.80 ± 2.7
SVM	0.83	0.78	0.82	0.81 ± 2.5
FCM	0.85	0.77	0.81	0.81 ± 2.7

(a)

$\mu_i - \mu_j$, p-value	μ_1	μ_2	μ_3
μ_2	0.1192		
μ_3	0.0342	0.4226	
μ_4	0.0028	0.6220	1.0000

(b)

Another popular approach, namely the receiver operating characteristics (ROC) curve, is also used to evaluate the performance of proposed method. The ROC curve is a plot of the true positive rate against the false positive rate for every possible cutoff of a diagnostic test. The performance is measured by the area under the ROC curve (AUC), which is used to describe the probability of random samples assigning to the correct class. The comparison of AUC among LDA, MLP, SVM and FCM methods is listed in Table 3(a). For Subject 1, using FCM clustering achieved the best AUC, while the worst AUC was obtained using LDA, the difference between them being 0.06. In addition, using SVM as the classifier obtained the second highest AUC, which was only 0.02 less than the best one. The best AUC was obtained using MLP as the classifier for Subject 2, while using SVM achieved the best AUC for Subject 3. The worst AUC was found when LDA was used as the classifier for all the three subjects. Overall, the best average AUC (0.81) was obtained using SVM and FCM, while using LDA yielded the worst average AUC (0.75), the difference being 0.06. The results also indicate that FCM clustering is a potential approach for MI classification.

Additionally, we perform two-way ANOVA and multiple comparison tests one more to verify if the AUC among classifiers is significantly different. The verified results are listed in Table 3(b). There are similar results to classification accuracy for the tests.

4. **Conclusions.** We have proposed an EEG analysis system for the single-trial recognition of MI EEG data in this study. Associated with active segment selection and multiresolution fractal features, the FCM is applied to discriminate left MI from right one. Active segment selection is an effective scheme that selects active segments in the time-frequency domain. It makes the length of original event-related window substantially

reduce to a 1-s segment and increases the speed of feature extraction at the same time. The multiresolution fractal features MFFVs are then obtained using proposed modified fractal dimension from DWT data. The features are so good and discriminative that they can enhance the classification of mental tasks. Finally, the FCM is used for the discriminant of MI EEG data. In addition, the FCM is not only a robust approach suitable for the classification of non-stationary MI EEG signals, but is also capable of making flexible partitions of a finite data set. The experimental results demonstrate that the recognition with the FCM possesses promising potential in the application of BCI works.

REFERENCES

[1] C. R. Hema, M. P. Paulraj, R. S. Y. Nagarajan and A. H. Adom, Brain machine interface: A comparison between fuzzy and neural classifiers, *International Journal of Innovative Computing, Information and Control*, vol.5, no.7, pp.1819-1827, 2009.

[2] J. R. Millán, F. Renkens, J. Mouriñ and W. Gerstner, Noninvasive brain-actuated control of a mobile robot by human EEG, *IEEE Trans. Biomed. Eng.*, vol.51, no.6, pp.1026-1033, 2004.

[3] D. Miao and S. Wang, A quantitative measurement of brain cognitive function based on human voice separation ability, *ICIC Express Letters*, vol.2, no.1, pp.15-21, 2008.

[4] W. Y. Hsu, EEG-based motor imagery classification using neuro-fuzzy prediction and wavelet fractal features, *Journal of Neuroscience Methods*, vol.189, no.2, pp.295-302, 2010.

[5] T. Yamaguchi, K. Nagata, P. Q. Truong, M. Fujio and K. Inoue, Pattern recognition of EEG signal during motor imagery by using SOM, *International Journal of Innovative Computing, Information and Control*, vol.4, no.10, pp.2617-2630, 2008.

[6] W. Y. Hsu, P. W. F. Poon and Y. N. Sun, Automatic seamless mosaicing of microscopic images: Enhancing appearance with colour degradation compensation and wavelet-based blending, *Journal of Microscopy*, vol.231, no.3, pp.408-418, 2008.

[7] 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*, vol.1137, pp.84-91, 2007.

[8] G. Pfurtscheller, G. R. Muller-Putz, A. Schlogl, B. Graimann, R. Scherer, R. Leeb, C. Brunner, C. Keinrath, F. Lee, G. Townsend, C. Vidaurre and C. Neuper, 15 years of BCI research at Graz university of technology: Current projects, *IEEE Trans. Neural Sys. Rehabil. Eng.*, vol.14, no.2, pp.205-210, 2006.

[9] G. Pfurtscheller and S. F. H. Lopes, Event-related EEG/MEG synchronization and desynchronization: Basic principles, *Clin. Neurophysiol.*, vol.110, pp.1842-1857, 1999.

[10] W. Y. Hsu and Y. N. Sun, EEG-based motor imagery analysis using weighted wavelet transform features, *Journal of Neuroscience Methods*, vol.176, no.2, pp.310-318, 2009.

[11] Z. Zhang, H. Ikeuchi, N. Saiki, T. Imamura, T. Miyake1, H. Toda and S. Horihata, Fast wavelet instantaneous correction and its application to abnormal signal detection, *International Journal of Innovative Computing, Information and Control*, vol.4, no.10, pp.2697-2710, 2008.

[12] B. B. Mandelbrot, *Fractal Geometry of Nature*, Freeman Press, San Francisco, 1982.

[13] M. E. Kirlangic, D. Perez, S. Kudryavtseva, G. Griessbach, G. Henning and G. Ivanova, Fractal dimension as a feature for adaptive electroencephalogram segmentation in epilepsy, *Proc. of the 23rd Annual Inte. Conf. IEEE on Engi. in Medi. and Biol. Soci.*, vol.2, pp.1573-1576, 2001.

[14] F. V. Nelwamondo, T. Marwala and U. Mahola, Early classifications of bearing faults using hidden markov models, Gaussian mixture models, mel-frequency cepstral coefficients and fractals, *International Journal of Innovative Computing, Information and Control*, vol.2, no.6, pp.1281-1299, 2006.

[15] Z. Zhang and Y. Zhao, Multiple description image coding based on fractal, *International Journal of Innovative Computing, Information and Control*, vol.3, no.6(B), pp.1616-1623, 2007.

[16] Z. Zhang and Y. Zhao, Improving the performance of fractal image coding, *International Journal of Innovative Computing, Information and Control*, vol.2, no.2, pp.387-398, 2006.

[17] 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*, vol.163, no.1, pp.145-160, 2007.

[18] N. Shimo, S. Pang, N. Kasabov and T. Yamakawa, Curiosity-driven multi-agent competitive and cooperative LDA learning, *International Journal of Innovative Computing, Information and Control*, vol.4, no.7, pp.1537-1552, 2008.

[19] H. B. Zhu, H. M. Kai, K. Eguchi, Z. Guo and J. Wang, Application of bpnn in classification of time intervals for intelligent intrusion detection decision response system, *International Journal of Innovative Computing, Information and Control*, vol.4, no.10, pp.2483-2491, 2008.

[20] B. H. Chen, L. P. Ma and J. L. Hu, An improved multi-label classification method based on SVM with delicate decision boundary, *International Journal of Innovative Computing, Information and Control*, vol.6, no.4, pp.1605-1614, 2010.

[21] X. Song, S. K. Halgamuge, D. Chen, S. Hu and B. Jiang, The optimized support vector machine with correlative features for classification of natural spearmint essence, *International Journal of Innovative Computing, Information and Control*, vol.6, no.3(A), pp.1089-1099, 2010.

[22] M. A. Ismail and S. Z. Selim, Fuzzy c-mean: Optimality of solutions and effective termination of the algorithm, *Pattern Recognition*, vol.19, no.6, pp.481-485, 1986.

[23] K. Shinkai, Decision analysis of fuzzy partition tree applying fuzzy theory, *International Journal of Innovative Computing, Information and Control*, vol.4, no.10, pp.2581-2594, 2008.

[24] K. Umayahara, S. Miyamoto and Y. Nakamori, Formulations of fuzzy clustering for categorical data, *International Journal of Innovative Computing, Information and Control*, vol.1, no.1, pp.83-94, 2005.

[25] C. Guger, A. Schlogl, C. Neuper, T. Strein, D. Walterspacher and G. Pfurtscheller, Rapid prototyping of an EEG based BCI, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol.9, no.1, pp.49-57, 2001.

[26] A. Schlogl, C. Keinrath, R. Scherer and G. Pfurtscheller, Estimating the mutual information of an EEG-based brain computer interface, *Biomedizinische Technik*, vol.47, pp.3-8, 2002.

[27] *The Graz Data Set and Description for the BCI 2003 Competition*, http://ida.first.fraunhofer.de/ projects/bci/competition/.

[28] B. J. Fisch, *Fisch and Spellmann's EEG Primer: Basic Principles of Digital and Analogue EEG*, Elsevier, New York, 1999.

[29] Y. Chen, J. Li, B. Feng and W. Guan, Hermite cubic spline multi-wavelet natural boundary element method, *ICIC Express Letters*, vol.3, no.2, pp.213-217, 2009.

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

[31] A. P. Pentland, Fractal based description of natural scenes, *IEEE Trans. on Pattern Anal. and Machine Intell.*, vol.6, pp.661-674, 1984.

[32] Z. Zhang, H. Ikeuchi, H. Toda, T. Miyake, T. Imamura, H. Ishii and S. Horihata, Wavelet and applying it in abnormal signal detection, *International Journal of Innovative Computing, Information and Control*, vol.4, no.4, pp.1009-1022, 2008.

[33] F.-X. Yu, J.-Y. Su, Z.-M. Lu, P.-H. Huang and J.-S. Pan, Multi-feature based fire detection in video, *International Journal of Innovative Computing, Information and Control*, vol.4, no.8, pp.1987-1993, 2008.

[34] K.-H. Huarng, T. H.-K. Yu and T.-T. Kao, Analyzing structural changes using clustering techniques, *International Journal of Innovative Computing, Information and Control*, vol.4, no.5, pp.1195-1201, 2008.

[35] J. Ren, C. Hu and R. Ma, HCLUWIN: An algorithm for clustering heterogeneous data streams over sliding windows, *International Journal of Innovative Computing, Information and Control*, vol.6, no.5, pp.2171-2179, 2010.

[36] H.-C. Liu, J.-M. Yih, W.-C. Lin and T.-S. Liu, Fuzzy C-means algorithm based on PSO and mahalanobis distance, *International Journal of Innovative Computing, Information and Control*, vol.5, no.12(B), pp.5033-5040, 2009.

[37] K. Zou, J. Hu, W. Li and L. Yu, FCM clustering based on ant algorithm and its application, *International Journal of Innovative Computing, Information and Control*, vol.5, no.12(B), pp.4819-4824, 2009.

[38] J. C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, *J. Cybern.*, vol.3, no.3, pp.32-57, 1974.

[39] J. C. Bezdek, *Fuzzy Mathematics in Pattern Classification*, Ph.D. Thesis, Applied Mathematics. Cornell University, Ithaca, New York, 1973.