International Journal of Medical, Medicine and Health Sciences
ISSN: 2517-9969
Vol:7, No:2, 2013

Active Segment Selection Method in EEG Classification Using Fractal Features

Samira Vafaye Eslahi

***Abstract*—**BCI (Brain Computer Interface) is a communication machine that translates brain massages to computer commands. These machines with the help of computer programs can recognize the tasks that are imagined. Feature extraction is an important stage of the process in EEG classification that can effect in accuracy and the computation time of processing the signals. In this study we process the signal in three steps of active segment selection, fractal feature extraction, and classification. One of the great challenges in BCI applications is to improve classification accuracy and computation time together. In this paper, we have used student's 2D sample t-statistics on continuous wavelet transforms for active segment selection to reduce the computation time. In the next level, the features are extracted from some famous fractal dimension estimation of the signal. These fractal features are Katz and Higuchi. In the classification stage we used ANFIS (Adaptive Neuro-Fuzzy Inference System) classifier, FKNN (Fuzzy K-Nearest Neighbors), LDA (Linear Discriminate Analysis), and SVM (Support Vector Machines). We resulted that active segment selection method would reduce the computation time and Fractal dimension features with ANFIS analysis on selected active segments is the best among investigated methods in EEG classification.

***Keywords*—**EEG, Student's t- statistics, BCI, Fractal Features, ANFIS, FKNN.

I. Introduction

BRAIN Computer Interface is a machine that translates brain activity into computer commands. This machine provides direct communication between brain and computer. These kinds of machines can help people with physical disability, does their daily tasks as well as healthy people. First of all the brain signals are recorded from the scalp and then will be processed. Sensorimotor rhythms (SMRs) are brain rhythmic waves that are among the frequency range of 8 to 12 Hz over the left and right sensorimotor cortices. Movement imagery in relaxation would, desynchronize these waves, and post-movement would synchronize SMRs [1].

The BCI is described that a person, has the ability to communicate with others without the prerequisite of brain's normal output pathways of peripheral nerves and muscles by controlling his EEG signals [2]. BCI systems based on MI EEG signals have become popular in the last decade [3]. Numerous methods have been presented such as linear regression, Kalman filtering [4], NN (Neural networks) [5], and FIS (fuzzy inference system) [6]. Linear regression is simple but it has low adaptation. NN can approximate any nonlinear functions but it needs a great deal of training data in feature space, In addition FIS that has a great capability of interpretation but its adaptability is low. ANIS [7], integrates the advantage of both NN and FIS and can be interpreted easily. Its training is fast and can converge on small data set too. Higuchi presented a method on time series fractal calculation that had good speed [8]. In EEG classification another approach was the usage of wavelet transformation. This method has used five ANFIS classifiers and five different kinds of EEG signals as inputs. For improving the accuracy they used one more classifier that its inputs are the outputs of those five classifiers [9]. After that for EEG classification, continuous wavelet features were extracted and using ANFIS in classification stage showed better performance than SVM (support vector machine). Another method with statistical features obtained from wavelet coefficients and FSVM (fuzzy support vector machine) classifier showed better accuracy than SVM [10]. FD was also good in separating various pairs of target EEG data. The general conclusion was that there is no particular discriminator uniformly suitable for all types of EEG data. Different discriminators perform differently on different data sets [11]. FD was used on EEG after feature extraction by a combination of continuous WT and student's t-statistics with the best classification accuracy in the 2003 BCI competition [12]. FD was used for random classification of EEG channels for BCI [13]. A comparison of FD and two of its variants with SVM and k nearest neighbor (KNN) algorithm on EEG data before onset of finger movements appears [14]. For a review of applications of linear discriminant analysis in BCI research see Reference [15]. In 2011 fractal dimension estimation features were classified with three classifiers, FKNN (fuzzy k- nearest neighbors), LDA (linear discriminate analysis) and SVM. It results that FKNN had the most accuracy among these three classifiers with Katz's fractal dimension method [16]. In this study the active segment selection is based on the CWT and Student's two-sample t-statistics and is used to obtain the location of optimal active segment in the time–frequency domain. Then we extract fractal dimension features as Katz and Higuchi fractal dimension estimation methods and we do binary classification with ANFIS, FKNN, LDA and SVM classifiers. Finally we compare them with these methods without active segment selection of the signal.

II. Materials and Methods

A. Dataset

The data set from BCI competition II (dataset III) have been provided by the Department of Medical Informatics, Institute for Biomedical Engineering, University of Technology Graz

Samira Vafaye Eslahi is with the Department of Biomedical Engineering, Islamic Azad University of Science and Research Branch. Tehran, Iran (e-mail: smr.vafa@gmail.com).

was analyzed. The data was obtained subject during imagery left and right hand movements over seven runs from a healthy 25 year old female. The signals were recorded with a sampling rate of 128 Hz from three electrodes placed at the standard positions of the 10-20 international system (C3, Cz and C4) and filtered between 0.5 and 30 Hz. Each run consisted of 40 trials and each trial was nine seconds long. During the first two or more seconds of each trial, neither a stimulus was presented nor did the subject perform any motor imagery task. After this period, an acoustic and a visual stimulus indicating the beginning of the motor imagery task were presented. Then, for six seconds, a cue (a left or right arrow) showing the required motor imagery task was presented (in a random order for each trial) and the subject did this task. During this period, a feedback bar was displayed. Each of the training and the testing are 140 samples.

B. Continuous Wavelet Transforms (CWT)

After Laplacian filtering in both channels C3 and C4, the CWTs of the three states of the data segments have been analyzed respectively:

$$W_S^{C,n}(j,k) = \int_R f_S^{C,n}(x)\frac{1}{\sqrt{j}}(x)\psi\left(\frac{x-k}{j}\right)dx \tag{1}$$

where $\frac{1}{\sqrt{j}}(x)\Psi\left(\frac{x-k}{j}\right)$ are the dilated and translated versions of the wavelet function $\Psi(\mathrm{x})$ at scale j and shift k, and $W_s^{C,n}(j,k)$ indicate the CWT of the data segment $f_s^{C,n}(x)$, in which state "s" belongs to one of three distinct states, data segment C belongs to either channel C3 or C4 after the Laplacian filtering, and "n" represents each single trial. Here, the values of $W_s^{C,n}(j,k)$ can be represented with a 2D time-scale plot, which holds the special properties and keep the optimal scale separation of ERP components [12]. The reason we choose the Daubechies wavelet as the CWT function in Equation1 is mainly due to the special characteristic that Daubechies family wavelets are compactly supported with extreme phase and highest number of vanishing moments for a given support width. In addition to the property that the associated scaling filters are minimum-phase, Daubechies wavelet can be excellently applied on DWT as well as CWT and is suitably used for the detection of ERP components and salient oscillations [17].

C. Student's 2D Sample T-Statistics

We can define it as means and variances for the three states, i.e. left hand and right hand moving, left hand and right hand resting, of data after the process of information accumulation in scale space in three channels C3 and C4 and Cz. They are calculated from the training data set as below [18].

$$\overline{\pi_s^C}(k) = \frac{1}{N_s}\sum_{n=1}^{N_s}\pi_s^{C,n}(k) \tag{2}$$

and:

$$\sigma^2(k) = \frac{1}{N_s-1}\sum_{n=1}^{N_s}(\pi_s^{C,n}(k) - \overline{\pi_s^C}(k))^2 \tag{3}$$

In (3), "N_s" shows the number of trials in state "S". Student's two-sample t-statistics evaluated between any two of the three distinct states are subsequently represented as (4):

$$t_{s1s2}^c(k) = \frac{\left|\overline{\pi_{s1}^C}(k) - \overline{\pi_{s2}^C}(k)\right|}{\sqrt{(((N_{s1}-1)\sigma^2{}_{s1}^C(k) + (N_{s2}-1)\sigma^2{}_{s2}^C(k))/(N_{s1}+N_{s2}-2))((1/N_{s1})+(1/N_{s2}))}} \tag{4}$$

In this equation, "s_1" and "s_2" belong to two different states. The denominator in (4) indicates the gathered variance of the two states for channel *C*. The values of $t_{s1s2}^c(k)$ with different time k for the two distinct states, s1 and s2, can form a 1D function with respect to time, but this compared to the original wavelet transformed signal contains different characteristics. Points that have local peak in $t_{s1s2}^c(k)$ represent that they are with local maximal difference between the two states in the time-scale domain. After applying information accumulation in scale space, we can then obtain the Student's two-sample t-statistics. Then for selecting the optimal active segment we will use t-statistics. The active 6 seconds segment for each hand movement states is selected with its center being the peak after the *C3* and *C4* and *CZ* channels that were concatenated together.

D. Fractal Features

Feature extraction is the fractal dimension estimation of the recorded EEG signals. Fractal dimension (FD) is a useful concept in describing natural objects, which gives their degree of complexity [19] in fractal geometry, the FD is a statistical quantity that gives an indication of how completely a fractal appears to fill the space, as one zooms down to finer and finer scales, accordingly there are many specific definitions of fractal dimension. The FD is a measure of how complicated a self-similar figure is. Hence the FD can be considered as a relative measure of number of basic building blocks that form a pattern [20]. We introduce two fractal dimension estimation methods as Katz's and Higuchi's methods.

E. Katz's Method

Katz's method [21], calculates the Euclidean distance between two samples as below:

$$D = \frac{\log(L)}{\log(d)} \tag{5}$$

"L" is the total length of the curve and "d" is the maximum distance between first sample and the farthest one. With normalizing the distance the fractal dimension becomes:

$$D_k = \frac{\log(n)}{\log(n) + \log\left(\frac{d}{L}\right)} \tag{6}$$

In this equation "n" is the number of steps and calculates as $n = L/a$, Where "a" is the average of the Euclidean distances between the successive points of the sample.

F. Higuchi's Method

Higuchi's method [8], calculates fractal dimension as follows: considering a time series consequence y(1),y(2), …,y(N) we can construct subsample sets y-m as below:

$$y_m^k = \{y(m), y(m+k), y(m+2k), \ldots, x(m+Mk)\}, m = 1,2,\ldots,k \quad (7)$$

where "k" is confined to [1, K-max] and m is confined to [1, k], and "M" is the sample size. The length of each y (m) is calculated as:

$$L_m(k) = \frac{1}{k}\left\{\frac{N-1}{Mk}\sum_{i=1}^{M}(|y(m+ik) - y(m+(i-1)k)|)\right\} \quad (8)$$

Finally with normalization factor *(N-1)/ MK*, the Higuchi fractal dimension can be obtained as:

$$D = \frac{ln\,(L(k))}{ln\,(\frac{1}{k})} \quad (9)$$

where *L (k)* is sum of sub sample sets as follow:

$$L(k) = \sum_{m=1}^{k} L_m\,(k) \quad (10)$$

In this method we considered k-max equals to 5 for ANFIS features and, 18 and 13 for others. Then we calculated fractal dimension of the active segments of signal.

G. Classification

In classification stage we use four classifiers as FKNN, LDA, SVM and ANFIS. We explain these methods briefly.

H. Fuzzy K-Nearest Neighbors

FKNN [22] searching is similar to simple KNN (k-nearest neighbors) search. In simple KNN, every data point can belong to only one class which is the majority class in the K-nearest neighbor search. Whereas in FKNN, a data point can belong to multiple classes with different membership functions associated to these classes.

I. LDA

Another way to classify data is to first create models of the probability density functions for data generated from each class. Then, a new data point is classified by determining the probability density function whose value is larger than the others. LDA is an example of such an algorithm. LDA assumes that each of the class probability density functions can be modeled as a normal density, and that the normal density functions for all classes have the same covariance.

J. SVM

A primary motivation behind SVM is to directly deal with the objective of good generalization by simultaneously maximizing the performance of the machine while minimizing the complexity of the learned model. Cover's theorem on the separability of patterns [23] essentially says that data cast nonlinearly into a high-dimensional feature space is more likely to be linearly separable there than in a lower-dimensional space. Even though the SVM still produces a linear decision function, the function is now linear in the feature space, rather than the input space. Because of the high dimensionality of the feature space, we can expect the linear decision function to perform well, in accordance with Cover's theorem. Viewed another way, because of the nonlinearity of the mapping to feature space, the SVM is capable of producing arbitrary decision functions in input space, depending on the kernel function.

K. Adaptive Neuro-Fuzzy Inference System

A specific approach in neuro-fuzzy development is the Adaptive Neuro-Fuzzy Inference System (ANFIS), which has shown significant results in modeling nonlinear functions. In ANFIS, the membership function parameters are extracted from a data set that describes the system behavior. The ANFIS learns features in the data set and adjusts the system parameters according to a given error criterion [7]. Successful implementations of ANFIS in biomedical engineering have been reported, for classification [24]-[27] and data analysis [28]. The ANFIS is a fuzzy Sugeno model put in the framework of adaptive systems to facilitate learning and adaptation [7]. Such framework makes the ANFIS modeling more systematic and less reliant on expert knowledge.

III. Results

In this paper after preprocessing the recorded signal we used student's 2D samples t-statistics to select the active segments of the signal to use only active parts then we used fractal dimension estimation methods to extract fractal features from the signal. These fractal dimension estimation methods are Katz and Higuchi methods. After this step we used four famous classifiers for binary classification of the features in BCI application consist of right or left motion imagery. We implemented this method with MATLAB 2007, and found the output accuracy which is shown in Table I and the computation time that is shown in Table II.

TABLE I
COMPARISON OF CLASSIFICATION ACCURACY ON FEATURES EXTRACTED FROM SELECTED ACTIVE SEGMENTS AND TOTAL SIGNAL

Accuracy (Percent)	Katz's Method	Higuchi's Method	Student's t-statistics on CWT (Katz)	Student's t-statistics on CWT(Higuchi)
FKNN	85% (k = 9)	79% (k = 7, k-max =18)	84%	80%
SVM	79%	80% (k-max=13)	72%	82%
LDA	77%	78% (k-max=13)	79%	76%
ANFIS	82%	88% (k-max = 5)	82%	89%

As we see in Table I, the classification accuracy of some methods as Higuchi, Katz feature extractions with FKNN, LDA, SMV and ANFIS the classifiers have been compared to these methods on selected active segments of the signal. In 0 I we see obtained accuracy. The calculated computation time, consist of both feature extraction and classification process, has shown in Table II on total signal and active segments.

TABLE II
CALCULATED COMPUTATION TIME OF SOME METHODS CONSISTS OF BOTH FRACTAL FEATURES AND CLASSIFICATION PROCESSES ON SELECTED ACTIVE SEGMENTS

Computation time	Katz's Method	Higuchi's Method	Student's t-statistics on CWT(Katz)	Student's t-statistics on CWT(Higuchi)
FKNN	0.16 (k = 9)	1.06 (k = 7, k-max = 18)	0.1(k = 9)	0.98(k = 7, k-max = 18)
SVM	0.32	1.05 (k-max= 13)	0.27	1.02 (k-max = 13)
LDA	0.14	0.8 (k-max =13)	0.14	0.55(k-max = 13)
ANFIS	0.36	0.33 (k-max=5)	0.17	0.15(k-max = 5)

IV. CONCLUSION

In this paper we filtered original EEG signals between 0.5 Hz to 30 Hz recorder by Graz laboratory from 3 channels C3, CZ, C4. We used student's 2D samples t-statistics on continuous wavelet transform (CWT) for active segment selection to reduce computation time. Then e extracted features by two fractal dimension estimation methods and classified inputs with four famous classifiers like Fuzzy k-Nearest Neighbors (FKNN), linear Discriminate Analysis (LDA), Support Vector Machine (SVM) and Adaptive Neuro –Fuzzy Inference System (ANFIS). As we see in Table I, ANFIS classifier with Higuchi fractal features have the most classification accuracy, and in comparison to FKNN with Katz fractal features it has been improved but the computation time of ANFIS classification with Higuchi fractal features is more than FKNN with Katz features. We applied student's t-statistics, on wavelet transform coefficients to find active segment selection and in this way we could achieve a good result in decreasing the computation time. Also some of the classification accuracy has been decreased but we could reach to a good improvement in computation times. As we see in Table I, ANFIS classification accuracy with Higuchi fractal features did not decreased and it had a good improvement in computation time too. So selection of active segment of the signal can cause a good speed in BCI applications. We see in Table II that ANFIS method with Higuchi fractal features do not have a good speed but we could decrease it and get more close to FKNN with Katz fractal features and other fast classification methods.

REFERENCES

[1] A. Kubler, and K. Muller, "an introduction to brain computer interfacing," in: G.dornhege, J.D.R. Millan, T. Hinterberger, D.J. McFarland and K. Muller (Eds). Toward Brain-Computer Interfacing. The MIT Press Massachusetts. 2007; 1-25.
[2] J. R. Wolpaw, N Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, and G. Schalk, "Brain–computer interface technology: a review of the first international meeting," IEEE Trans. Rehabil. Eng. 2000; 8(2):164–73.
[3] L. Parra, C. Alvino, A. C. Tang, B. A. Pearlmutter, N. Yeung, and A. Osman, "Linear spatial integration for single trial detection in encephalography. NeuroImage 2002; 7(1):223–30.
[4] X. Hu, D. V. Prokhorov, I. I. Wunsch, and C. Donald," Time series prediction with a weighted bidirectional multi-stream extended Kalman filter," Neurocomputing. 2007; 70: 2392–9.
[5] N. Stamatis, D. Parthimos, T. M. Griffith, "Forecasting chaotic cardiovascular time series with an adaptive slope multilayer perceptron neural network. IEEE Trans Biomed Eng 1999; 46(12):1441–53.
[6] H. J. Rong, N. Sundararajan, G. B. Huang, P. Saratchandran, "Sequential adaptive fuzzy inference system (SAFIS) for nonlinear system identification and prediction," Fuzzy Sets Syst 2006; 157 (9): 1260–75.
[7] J. S. Jang Roger, "ANFIS: adaptive-network-based fuzzy inference system," IEEE Trans SMC 1993; 23(3):665–85.
[8] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," Physica D Nonlinear Phenomena. 1988; vol.31; 277-283.
[9] İ. Güler, E. D. Übeyli, "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients," Elsevier. Journal of Neuroscience Methods. 2005; 148(2):113-21.
[10] Q. Xu, H. Zhou, Y. Wang, J. Huang, " Fuzzy support vector machine for classification of EEG signals using wavelet-based features," Elsevier, Medical Engineering & Physics. 2009; (31): 858- 865.
[11] K. Majumdar., R.M. Frank, "Three linear discriminators for separating scalp human EEG signals during RSVP tasks," unpublished work, 2006.
[12] V. Bostanov, "BCI competition 2003 – data sets Ib and IIb: feature extraction from evenet related brain potentials with the continuous wavelet transform and the t-value scalogram," IEEE Trans. Biomed. Eng. 2004, 1057–1061.
[13] S. Sun, C. Zhang, Y. Lu, "The random electrode selection ensemble for EEG signal classification," Pattern Recogn., 2008; 14(5): 1663–1675.
[14] B. Blankertz, G. Curio, K. Muller, "Classifying Single Trial EEG: Towards Brain Computer Interfacing," Advances in Neural Information Processing Systems (NIPS), MIT Press, Cambridge, MA, 2002, pp.157–164.
[15] F. Lotte, M. Congedo, A. Lecuyer, F. Lamarche, B. Arnaldi, "A review of classification algorithms for EEG-based brain-computer interface," J. Neural Eng. 2007; 4 (2): R1–R13.
[16] U. Güçlü, Y. Güçlütürk, C. Loo, "Evaluation of fractal dimension estimation methods for feature extraction in motor imagery based brain computer interface," Elsevier, Procedia Computer Science. 2011; 3: 589–594.
[17] I. Daubechies, "Ten lectures on wavelets," CBMS-NSFLecture Notes nr. 61, SIAM, 1992.
[18] W. Y. Hsu, C. C. Lin, M. S. Ju, Y. N. Sun, "Wavelet-based fractal features with active segment selection: Application to single-trial EEG data," Journal of Neuroscience Methods. 2007; 163(1): 145–160.
[19] M. F. Barnsley, "Fractals Everywhere," 2nd Ed. New York: Academic Press Professional; 1993.
[20] K. Falconer, "Fractal geometry: Mathematical foundations and applications," Wiley, New York. 1990.
[21] J. Katz, "Fractals and the analysis of waveform," Computers in Biology and Medicine. 1988; 18: 145-156.
[22] J. M. Keller, M. R. Gray, J. A. Givens,"A fuzzy k-nearest neighbor algorithm," IEEE Transactions on Systems, Man and Cybernetics. SMC-1985; 15(4), 580-585.
[23] D. Cohen, E. Halgren, "Magnetoencephalography (Neuromagnetism). Preprint from Encyclopedia of Neuroscience," Elsevier, 2003; 3rd Ed.
[24] J. Usher, D. Campbell, J. Vohra, J. Cameron, "A fuzzy logic-controlled classifier for use in implantable cardioverter defibrillators," Pace-Pacing Clin Electrophysiol. 1999; 22: 183–6.
[25] S. Y. Belal, A. F. G. Taktak, A. J. Nevill, S. A. Spencer, D. Roden, S. Bevan, "Automatic detection of distorted plethysmogram pulses in neonates and paediatric patients using an adaptive-network-based fuzzy inference system," Artif Intell Med 2002; 24:149–65.
[26] I. Guler, E. D. Ubeyli, "Detection of ophthalmic artery stenosis by leastmean squares backpropagation neural network," Comput Biol Med. 2003; 33(4):333–43.
[27] I. Guler, E. D Ubeyli, "Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with with partial epilepsy using feature extraction. Expert Syst Appl. 2004; 27(3):323–30.
[28] Virant-Klun I, Virant J. Fuzzy logic alternative for analysis in the biomedical sciences. Comput Biomed Res. 1999; 32:305- 21.