EEG classification using fractal features and Adaptive Neuro- Fuzzy Inference System analysis in BCI applications
Samira Vafaye Eslahi, Nader Jafarnia Dabanloo and Keivan Maghooli
Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

ABSTRACT
BCI (Brain Computer Interface) roles as a machine that 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. In these machines, the brain signals are recorded from the scalp and will be prepared for analyzing in three steps of preprocessing, feature selection and classification that what kinds of tasks have been imagined. In BCI applications a big challenge is to improve classification accuracy in parallel with the computation time. In this paper, in preprocessing level we filtered the samples of each electrode with band pass digital Butterworth filter with cutoff frequency of 0.5 to 30 HZ. 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 and compared it with three strong classifiers as FKNN (Fuzzy k-Nearest Neighbors), LDA (Linear Discriminate Analysis) and SVM (Support Vector Machine). We found ANFIS with Higuchi fractal features has the most classification accuracy (88%) among other investigated methods, but its speed is rather low among them.

© 2013 Elixir All rights reserved

Introduction
Brain Computer Interface is a machine that translates brain activity into computer commands. Sensory motor rhythms (SMRs) are brain rhythmic waves that are among the frequency range of 8 to 12 Hz over the left and right sensory motor cortices. Movement imagery in relaxation would desynchronize these waves, and post-movement would synchronize SMRs [16]. 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 [25].

BCI systems based on MI EEG signals have become popular in the last decade [18]. Numerous methods have been presented such as linear regression, Kalman filtering [12], NN (Neural networks) [21], and FIS (fuzzy inference system) [20]. 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 [13] 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 [11]. 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 [10]. After that for EEG classification, continuous wavelet features were extracted and the 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 [26]. 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 3 classifiers with Katz’s fractal dimension method [7]. The principal aim of this study is to extract fractal dimension features as Kats and Higuchi fractal dimension estimation methods and binary classification with ANFIS classifier to compare them with each other and with three strong classifiers like FKNN, LDA and SVM done. In tradeoff between accuracy and time, as the ratio of accuracy to computation time increases it means that the speed and accuracy of the system is more acceptable than the others.

Materials and methods:
Dataset:
The data set from BCI competition II (dataset III) provided by the Department of Medical Informatics, Institute for Biomedical Engineering, University of Technology Graz was analyzed. The data was acquired over seven runs from a healthy 25 year old female subject during imagery left and right hand movements. 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 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) indicating the required motor imagery task was presented (in a random order for each trial) and the subject performed this task. During this period, a feedback bar was displayed. Both the training and testing sets consisted of 140 samples.

Keywords
EEG; BCI, Fractal Features, ANFIS Classification, Motor imagery, FKNN.
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 [1], [16]. 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 [5]. We introduce two fractal dimension estimation methods as Katz and Higuchi.

Katz’s method:

Katz’s method [14], calculates the Euclidean distance between tow samples as below:

\[ D = \frac{\log(n)}{\log(n)+\log(d/L)} \]  (1)

“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:

\[ D_k = \log(n)/(\log(n)+\log(d/L)) \]  (2)

In this equation n is the number of steps and calculates as below:

\[ n = L/a \]

Where “a” is the average of the Euclidean distances between the successive points of the sample.

Higuchi’s method:

Higuchi’s method [11], calculate fractal dimension as follows: considering a time series consequence y(1), y(2), … , y(N) we can construct subsample sets ym as below:

\[ y = \{y(m), y(m+K), y(m+2k), \ldots, x(m+Mk)\}, m=1,2,\ldots,k \]  (3)

Where k is [1, Kmax] and m is [1, k], and M is the sample size. In the length of each ym(Lm) is calculated as:

\[ L_m(k) = 1/k \{\sum (y(m+ik) - y(m+(i-1)k))\} \]  (4)

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

\[ D = \ln(L(k))/\ln(1/k) \]  (5)

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

\[ L(k) = \sum L_m(k) \]  (6)

In this method we considered kmax equals to 5 and calculated fractal dimension of sub bands.

Classification:

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

Fuzzy K-Nearest Neighbors:

FKNN [15], search 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.

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.

SVM:

A primary motivation behind SVMs 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 [3], essentially says that data cast nonlinearly into a high-dimensional feature space is more likely to be linearly separable 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.

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 [13]. Successful implementations of ANFIS in biomedical engineering have been reported, for classification [23] and data analysis [24]. The ANFIS is a fuzzy Sugeno model put in the framework of adaptive systems to facilitate learning and adaptation [13]. Such framework makes the ANFIS modeling more systematic and less reliant on expert knowledge.

Results:

In this paper after preprocessing the recorded signal 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) software, and found the output accuracy (Table1) and the computation time (Table2). As we see in table 1, the classification accuracy of some methods like Higuchi, Katz feature extractions with FKNN, LDA, SVM and ANFIS the classifiers have been compared with each other. In table 1 we see obtained accuracy. The calculated computation time, consist of both feature extraction and classification processes has shown in table 2.

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 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 figure 1, ANFIS with Higuchi fractal features, has the most classification accuracy value, and in comparison with FKNN with Katz fractal features it is improved but as we see in fig 2, the computation time of ANFIS classification with Higuchi fractal features is more than FKNN with Katz features.
The goal of EEG classification in BCI applications is to increase classification accuracy and decrease computation time. In Fig 1 we can compare classification accuracy of these methods with each other. Obviously we can conclude ANFIS with Higuchi fractal features has the most classification accuracy, after that FKNN with Katz fractal features and ANFIS with Katz fractal features have acceptable accuracy values rather than others.

In contrary the computation time of SVM with Katz fractal features has the lowest value rather than others and also FKNN with Katz fractal features has a good speed in the process of feature extraction and classification but ANFIS with Higuchi fractal features with regard to the highest classification accuracy has more computation time than two methods of FKNN and SVM with Katz fractal features (Fig 2). Also ANFIS with Higuchi features had the most classification accuracy, but FKNN had the best ratio of accuracy to computation time.

In this article we conclude that ANFIS with Higuchi fractal had the most classification accuracy than other investigated methods and FKNN with fractal features had the most speed among them. Also classification accuracy is very important in BCI applications, but we cannot neglect the computation time.

In this paper we tried to classify EEG signals with ANFIS classifier and two fractal feature extraction methods as Higuchi and Katz and compared it to some other methods. Also we could improve the classification accuracy (88%) rather than FKNN with Katz fractal features (85%), but we could not reach to an acceptable speed, suggesting a method to improve both the accuracy and computation time together is a big achievement.

Table 1. Comparison of classification accuracy

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higuchi’s Method</td>
<td>79% (k = 7, kmax = 18)</td>
</tr>
<tr>
<td>Katz’s Method</td>
<td>80% (kmax = 13)</td>
</tr>
<tr>
<td></td>
<td>78% (kmax = 13)</td>
</tr>
<tr>
<td></td>
<td>88% (kmax = 5)</td>
</tr>
<tr>
<td>FKNN</td>
<td>85% (k = 9)</td>
</tr>
<tr>
<td>SVM</td>
<td>79%</td>
</tr>
<tr>
<td>LDA</td>
<td>77%</td>
</tr>
<tr>
<td>ANFIS</td>
<td>82%</td>
</tr>
</tbody>
</table>

Table 2. The results of calculated computation time of some methods consist of both fractal features and classification processes

<table>
<thead>
<tr>
<th>Method</th>
<th>Computation time (second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Higuchi’s Method</td>
<td>1.06 (k = 7, kmax = 18)</td>
</tr>
<tr>
<td>Katz’s Method</td>
<td>0.8 (kmax = 13)</td>
</tr>
<tr>
<td></td>
<td>0.33 (kmax=5)</td>
</tr>
<tr>
<td>FKNN</td>
<td>0.15 (k = 9)</td>
</tr>
<tr>
<td>SVM</td>
<td>0.32</td>
</tr>
<tr>
<td>LDA</td>
<td>0.14</td>
</tr>
<tr>
<td>ANFIS</td>
<td>0.36</td>
</tr>
</tbody>
</table>

References:


