Marcin KOŁODZIEJ, Andrzej MAJKOWSKI, Remigiusz RAK

Warsaw University of Technology, Institute of the Theory of Electrical Engineering, Measurement and Information Systems

Implementation of genetic algorithms to feature selection for the use of brain-computer interface

Abstract. The main goal of the article is to apply genetic algorithms to feature selection for the use of brain-computer interface (BCI). FFT coefficients of EEG signal were used as features. The best features for a BCI system depends on the person who uses the system as well as on the mental state of the person. Therefore, it is very important to apply efficient methods of feature selection. The genetic algorithm proposed by authors enables to choose the most representative features and electrodes.

Streszczenie. W artykule przedstawiono zastosowanie algorytmów genetycznych do selekcji cech na użytek interfejsów mózg-komputer (BCI). Najlepszy zestaw cech dla tego typu interfejsów jest zależny od osoby, która używa interfejsu, jak również od jej stanu psychicznego. Z tego powodu konieczne jest zastosowanie bardzo efektywnych metod selekcji cech. Jako cechy wykorzystane zostały współczynniki FFT sygnału EEG. Zaproponowany przez autorów algorytm genetyczny umożliwia wyznaczenie najbardziej reprezentatywnego zbioru cech, jak również elektrod, z których pobierany jest sygnał EEG. (Zastosowanie algorytmów genetycznych do selekcji cech na użytek interfejsów mózg-komputer)

Keywords: EEG, brain-computer interface, BCI, feature extraction, feature selection, ERD/ERS, genetic algorithms
Słowa kluczowe: EEG, interfejs mózg-komputer, ekstrakcja cech, selekcja cech, ERD/ERS, algorytmy genetyczne

Introduction

Implementing communication between man and machine by the use of EEG signals is one of the biggest challenges in signal theory. Such a communication system could improve the standard of living of people with severe motor disabilities. Disable persons cannot move, however they can think about moving their arms, legs and in this way produce stable motor-related EEG signals (so called event-related desynchronization/synchronization - ERD/ERS). The fundamental problem in all BCI systems is the proper interpretation of EEG signals [1,2,3]. EEG signals are read from the head surface. There are three main stages in EEG signal analysis: feature extraction, feature selection and classification (fig.1).
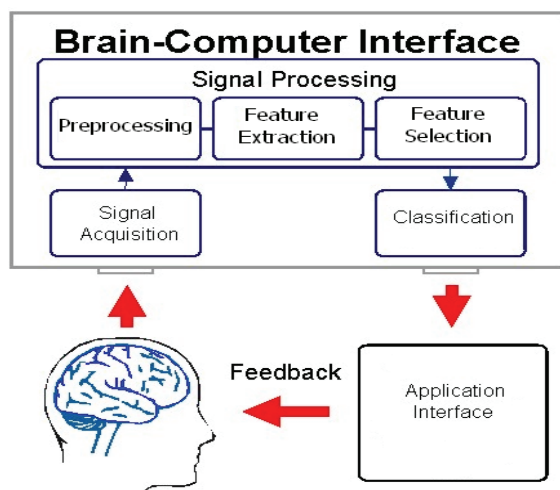

Fig.1. The block diagram of a BCI system

To enable proper interpretation of EEG signal first some features have to be extracted from the signal. In our case FFT coefficients of EEG signal were used as features. Because there is a really large number of so created features, to enable an effective classification only a small number of the most representative features must be selected. The process is farther complicated by the fact that the most representative features can depend on the person who uses the interface as well as on his mental state.

There are many methods of feature selection. To the best known belong: linear discriminant analysis (LDA), principal components analysis (PCA), sequential forward selection (SFS). The paper describes a novel method of feature selection based on the use of genetic algorithm (GA). As the final result the effects of classification of EEG signal will be presented.

Experiment description

The dataset used in the experiment is provided by IDIAP Research Institute (Silvia Chiappa, José del R. Millán). The set contains data from 3 normal subjects acquired during 3 non-feedback sessions. The subjects were relaxed, sat in a normal chairs with arms resting on their legs. Each subject had three tasks to execute:

1. Imagination of repetitive self-paced left hand movements, (*left*, class 2).
2. Imagination of repetitive self-paced right hand movements, (*right*, class 3).
3. Generation of words which starts with the same random letter, (*word*, class 7).

All 3 sessions were conducted on the same day. Each session lasted 4 minutes with 5-10 minutes breaks in between them. The subject performed a given task for about 15 seconds and then switched randomly to another task on the operator's request. EEG data was not split in trials. In the experiment we have focused on the first session of only one subject.

The EEG signal was divided into windows lasting one second. The aim of the experiment was to determine with the highest probability to which class the windowed EEG signal belongs. For every window a feature extraction method was applied. As features we used modules of Discrete Fourier Transform (DFT) coefficients. We considered 40 frequency points from the range: 1Hz - 40Hz. The EEG signal contained 32 channels. So that, for one-second window 40×32=1280 features were obtained. Such a large number of features doesn't allow for effective learning of the classifier. Therefore, in the next stage, selection of features was conducted. For this purpose genetic algorithm was used. It enabled to choose the most representative features. The features were next used as a learning set for the classifier. The procedure described above makes possible to determine which features represent most information.

As it is known which features are connected with the signal from which electrodes and which features are the

most representative, it allows also for drawing certain conclusions about location of the most important electrodes. That kind of information will contribute to construction of amplifiers with lower number of channels.

Genetic algorithms

A genetic algorithm (GA) is a method for solving optimization problems that are based on natural selection from the population members. The genetic algorithm repeatedly modifies a population of individual solutions. At each step the genetic algorithm tries to select the best individuals. From the current "parent" population genetic algorithm creates "children", who constitute next generation. Over successive generations the population evolves toward an optimal solution. The genetic algorithm uses three main rules at each step to create the next generation (fig.2):

- *Selection rules* select the individuals, called *parents*, that contribute to the population at the next generation.
- *Crossover rules* combine two parents to form children for the next generation.
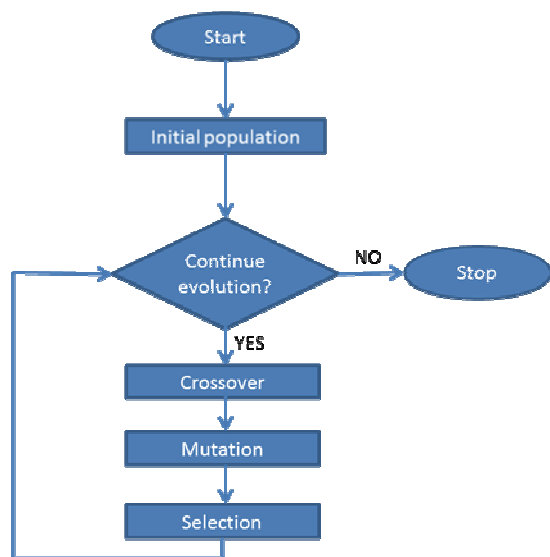- *Mutation rules* apply random changes to individual parents to form children.

Fig.2. Block diagram of genetic algorithm

Experiment procedure

All steps of our experiment, including feature extraction, feature selection and classification were implemented in Matlab. For the feature extraction the FE_toolbox was used [4]. Experiment began with the creation of feature set. As it has been already mentioned, for each, one second window of EEG signal, modules of Discrete Fourier Transform (DFT) coefficients were calculated. In this way 1280 features were received. While calculating DFT a rectangular time window was used. From the generated data three matrices were created, one for each class. Matrix named FC2, of dimension 1280x124, was related to the features of EEG signal for the case when users imagine the movement of their left hand by the time of 124 seconds. Respectively matrix named FC3 (imagining of right hand movement) and FC7 (generation of words) had dimensions 1280x152 and 1280x197. Relatively small number of vectors (time windows) with so many features did not allow for efficient classifier training. Moreover part of features were highly correlated (both: between DFT values and between channels). So efficient method of feature selection was

necessary. It was assumed that out of 1280 features only 30 features would be selected.

As it has been mentioned above, we used a genetic algorithm for feature selection [5,6]. The starting population consisted of 1000 individuals and each of them contained 30 randomly generated features. Next operations of mutation and crossover were performed (with some selected probability). In this way, exchange of genes was realized and what follows exchange of features. Only the best adapted individuals passed to the next step of the algorithm. In order to verify which individuals were the best adapted a special fitness function was used, which trained the classifier and next classified the data and returned the classification error.

For classification Linear Discriminant Analysis (Matlab function *classify* with the option *'quadratic'*) was used. The fitness function performed 10 times cross validation test and returned percentage of the classify error. Smaller error means more relevant features. If satisfactory level of error was achieved the algorithm was stopped. In this way 200 generations were calculated (fig. 3).
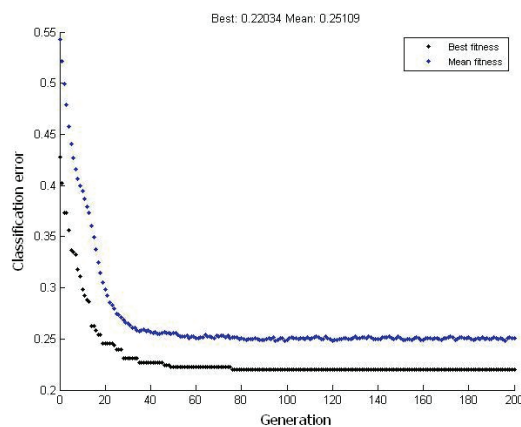

Fig.3. Classification error for 32 channels and 200 generations

Of course, each run of genetic algorithm may result in selection of different 30 features. So it is very important to determine which of these features (and the same time which channels) bring important information to the classification process.
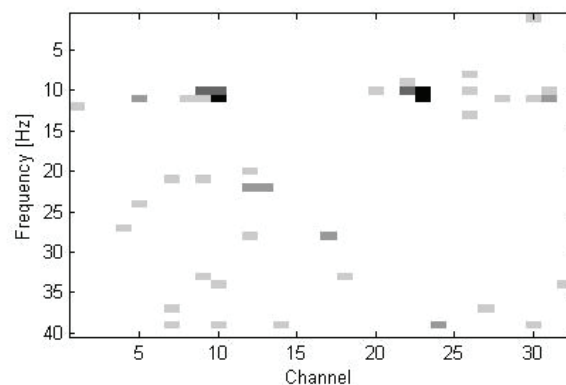

Fig.4. Features that repeated in 10 runs of genetic algorithm

To solve the problem genetic algorithm was launched ten times and next the selected features were compared. In this way we have chosen features which repeated in all runs. Results can be seen in figure 4. Darker shades mean that the features described by them occurred more frequently. Classification error for a single launch of genetic

algorithm ranged from 22% to 26%. It can be noticed that the most often selected features were taken from channels: 9, 10, 22 and 23 (that is from the electrodes: CP1, CP5, CP2 and C4). Such knowledge allows us to determine from which areas of brain signals should be collected. The results of our research showed that the most often selected features were attached to frequencies 10 Hz and 11 Hz. It can be observed in the fig.5 which presents the number of selected features per frequency for 10 runs of genetic algorithm.
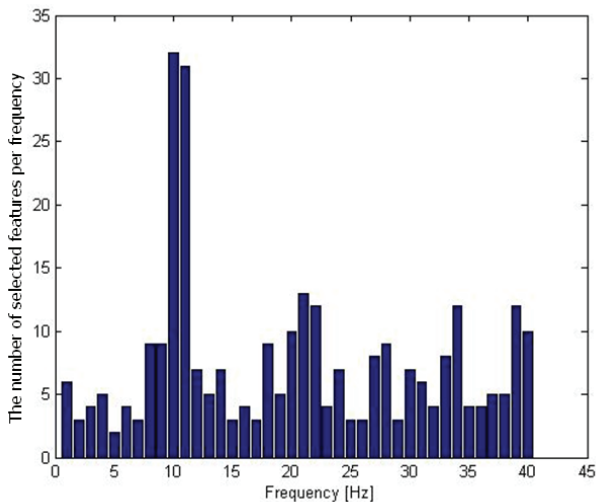

Fig.5. The number of selected features per frequency

Experiment for selected electrodes

As we knew which channels (electrodes) bring the most information, it would be interesting to check if information from only these channels (instead of 32 electrodes) would be sufficient for constructing of brain-computer interface. Our further studies went in this direction. We examined the classification error for only four electrodes CP1, CP5, CP2, C4 (selected earlier). Then the genetic algorithm was started to select the new best features and determine the classification error. It appeared that the classification error achieved in this case, for features from the range 1Hz - 40Hz, was only 23% (fig.6).
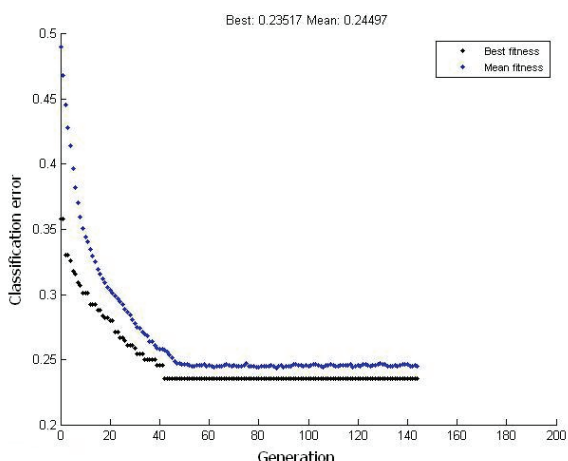

Fig.6. Classification error for 4 channels and 200 generations

The experiment showed truly that it is possible to reduce the number of electrodes without significant deterioration of the classification results.

Conclusions

There are many methods of feature selection. They can be divided into two groups: ranking and non-ranking. Studies showed that genetic algorithms, from the non-ranking group, can be successfully applied to feature selection. When using ranking methods the chosen feature vectors can contain features that are correlated with each other and at the same time do not bring in significant new information for the classifier. In turn to the main disadvantages of using GA belong: the long time the algorithm must run to produce results and the fact that each run of genetic algorithm creates different set of features. Next problem (as GA belongs to non-ranking methods) is that the selected features are not ranked, we do not know which of them brings more information for the classifier. So, in order to draw conclusions about "the importance of features", we must run the GA several times. It also enables us to determine where to place the electrodes and which frequencies of EEG spectrum are the most important.

In our experiment the classification error for the genetic algorithm stabilized after about 80 generations. Thus definitely, genetic algorithm can be stopped after reaching about 100 generations. It means that the time of its operation will be shortened. Additionally, the execution time of GA can be shortened by modifying the selected probabilities of mutation and crossover operations. However, the time required by GA for feature selection is fairly long in practice, so GA is rather suitable for the analysis of data in off-line mode.

The next step of the research should be testing GA for all users and all sessions and then practical verification of obtained results for the newly collected data. It is worth noting that in the experiment we did not take biofeedback into consideration. The biofeedback could significantly improve the results in practice.

REFERENCES
[1] Vidal, J.J., *Direct brain-computer communication*, Ann. Rev. Biophys Bioeng, 2, 1973.
[2] Molina G., *Direct Brain-Computer Communication through scalp recorded EEG signals*. PhD Thesis, École Polytechnique Fédérale de Lausane, 2004
[3] Wolpaw J.R., Birbaumer N., Heetderks W. J., Mcfarland D.J,Hunter Peckham P., Schalk G, Donchin E., Quatrano L.A., Robinson C.J, Vaughan T.M, *Brain–Computer Interface Technology: A Review ofthe First International Meeting*, IEEE Transactions on Rehabilitation Engineering, vol. 8, No. 2, June 2000.
[4] Kołodziej M., Majkowski A., Rak R. *Matlab FE_Toolbox - an universal utility for feature extraction of EEG signals for BCI realization*, Przegląd Elektrotechniczny 2010-1.
[5] Kantardzic M., " *Data Mining: Concepts, Models, Methods, and Algorithms* ", IEEE Press & John Wiley, November 2002.
[6] Documentation of Genetic Algorithm and Direct Search Toolbox™ - MATLAB.

Authors: *prof. dr hab. inż. Remigiusz J. Rak, e-mail: rakrem@iem.pw.edu.pl; dr inż. Andrzej Majkowski, e-mail: amajk@iem.pw.edu.pl; mgr inż. Marcin Kołodziej, e-mail: kolodzim@iem.pw.edu.pl Politechnika Warszawska, Instytut Elektrotechniki Teoretycznej i Systemów Informacyjno-Pomiarowych, ul. Koszykowa 75, 00-661 Warszawa.*