



# Event-Related Potentials Signal Feature Classification Algorithm Based on Genetic Algorithm

Bo Li<sup>1,2\*</sup>, Baoxing Bai<sup>1</sup>

## ABSTRACT

To solve the problem of insufficient feature information obtained by the single feature extraction method. The feature extraction was achieved by autoregressive model (AR) and wavelet transform. After the merging of feature sets, the genetic algorithm is used to select the optimal feature set. To test the validity of the proposed method, it is compared with the feature selection method based on SGA and the filter selection method based on Fisher distance. The classification accuracy of AGA is significantly higher than other methods, and the best pattern recognition performance is obtained. SGA-FS and AGA-FS take classification accuracy as the index to evaluate different feature subsets, so it is expected to obtain better classification results.

**Key Words:** Genetic Algorithm, Feature Classification Algorithm

**DOI Number:** 10.14704/nq.2018.16.6.1678

**NeuroQuantology 2018; 16(6):801-805**

## Introduction

The most natural way for humans to communicate with the outside world is through the muscular tissue of the body and the intentions formed in the brain are sent to specific parts of the body through the nervous system to make response (Folstein and Van, 2008). The human brain is a complex organ made up of hundreds of millions of cells called neurons. Electrical signals are used for communication among neurons. When a neuron discharges, it generates an electrical signal that can be detected on the scalp, which is called event-related potentials (ERP). It is a major problem in signal processing to realize the communication between humans and external devices through ERP signals. The communication system to achieve this purpose is called brain-computer interface, which will improve the living quality of humans, especially patients with physical disability but with normal thinking (Janelle, 2002). In current BCI systems, the non-invasive, ERP signals are generally used as signal sources.

The brain-computer interface (BCI) is a communication system of normal output channel that does not depend on the brain (peripheral nerves and muscles) and is a new human-machine interface method. Its essence is to infer human thoughts or objectives through the processing of brain signals so as to achieve human-computer communication. Among many brain signals, the electroencephalogram (ERP) has the advantages of large amount of information, high time resolution, portable acquisition equipment and low cost, thus becoming the major control signal in current brain-machine interface researches.

The basic principle of BCI is that people can generate specific pattern of ERP signals under certain conditions (such as imagining the movements of hands, feet, tongue, etc.) (as is shown in Figure 1). Through the feature extraction, selection and classification of ERP signals, different intentions of the subjects can be

**Corresponding author:** Bo Li

**Address:** <sup>1</sup>Changchun University of Science and Technology, Changchun 130000, China; <sup>2</sup>Changchun Institute of Technology, Changchun 130000, China

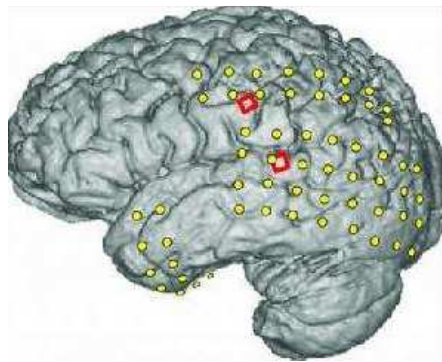
**e-mail** ✉ liboxstg1209@163.com

**Relevant conflicts of interest/financial disclosures:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Received:** 7 March 2018; **Accepted:** 5 May 2018



identified. Feature selection is one of the core issues in ERP pattern recognition and one of the key technologies in brain-computer interface researches. The task of feature selection is to select a set of optimal features with a quantity of  $d$  from a set of features with a quantity of  $D$  ( $D > d$ ), that is, to select some of the most effective features from all features for pattern recognition to reduce the dimensionality of the feature space and improve the classification performance (Orgs *et al.*, 2008).



**Figure 1.** Microelectrode brain signal transformation

At present, the methods used in the feature selection of ERP mainly include filter selection method (such as separability criteria and mutual information based on classification) and standard genetic algorithm (SGA). The former selects the feature subset with the largest amount of mutual information as the optimal subset, which is the feature election independent of the classifier, so it cannot guarantee to obtain the recognition rate with higher accuracy. The latter often uses the classification performance as its fitness function, so it can obtain better recognition results. Genetic algorithms (GA) is a global probability searching algorithm based on biological evolution mechanisms such as natural selection and genetic variation, which has many advantages and can maintain a good balance between the depth and breadth of the searching space. Through individual parameter coding, fitness function design, selection, cross and mutation, the fitness of the best individual in the group can improve continuously and finally the optimal solution to the problem can be found (Harle and Vickers, 2001). However, SGA is not very effective in many situations and it is prone to be disturbed by problems such as premature convergence and poor local optimization ability. In SGA, the crossover probability and the mutation probability are often estimated based on

experience, which will bring great blindness, thus affecting the global optimality and convergence of the algorithm. Based on this point of view, this paper proposes an adaptive genetic algorithm (AGA) to conduct adaptive selection of the above two probabilities. All features are used for optimization selection and finally the optimal feature subset can be obtained. This method is called as AGA-FS (Feature Selection).

### Methods

The AR model is an autoregressive model, whose meaning is that the current output of the model is the weighted sum of the current input and the previous  $P$  outputs. The formula is as follows:

$$y(n) = \sum_{j=1}^P a_j y(n-j) + x(n) \quad (1)$$

In the formula,  $P$  is the model order;  $a_j$  is each coefficient of the model. The commonly used methods for AR model parameter estimation are autocorrelation method, Burg method and least square method (Marple algorithm).

In this paper, AR model and wavelet transform are used to extract features and the mean, median and standard deviation of two types of features are calculated as additional features. The features extracted from these two are merged to form a feature set. Before the classification, the feature selection is performed on the original feature set according to the genetic algorithm based on k-Nearest Neighbor (KNN) algorithm classification error rate. And then, the selected feature is used as the input of classifiers such as support vector machine to verify the validity of the feature extraction method and genetic algorithm used for feature selection in this paper (Loze *et al.*, 2001). In this paper, the GA Toolbox provided by Matlab is used, in which the roulette algorithm is adopted to implement the selection operation. The single-point crossing algorithm and uniform mutation algorithm are selected for genetic operators (cross, mutation). The termination criteria of the algorithm is achieved by the combination of the generation number of the population and the Stall Genlimit parameter provided by the toolbox. Set Stall Genlimit to 50, that is and the genetic algorithm will terminate in the following two cases:

- (1) The population has reached the maximum number of generations;

(2) The maximum generation number was not reached, but the fitness function value does not change, within consecutive 50 generations of Stall Genlimit.

Then, the other two key issues are the coding scheme of the sample and the setting of the fitness function (evaluation function).

## Results and discussion

### Analysis of ERP signal model

The experimental data in this paper comes from the experimental database of "BCI Competition 2003". The structure of the original data is shown in Table 1.

**Table 1.** Meaning of variables in raw ERP data structure

variable	dimension	description
x-train	1152*3*140	Training set, channel C3, C4
y-train	140*1	Training set labels
x-test	1152*3*140	Test set, channel C3, C4

The data was collected in a 25-year-old normal female. The experimental task was to control the movement of the cursor on the screen by imagining the movement of the left or right hand. The instruction of the left and right arrows on the screen were given at random. All tests were conducted in the same day and the duration of each experiment lasted >T. The entire experiment consisted of 280 tests, of which there were 140 groups for left and right hand movements respectively. 70 for training and 70 for testing, containing 3 channel data of C3, C4, C2. The ERP sampling frequency is 128Hz.

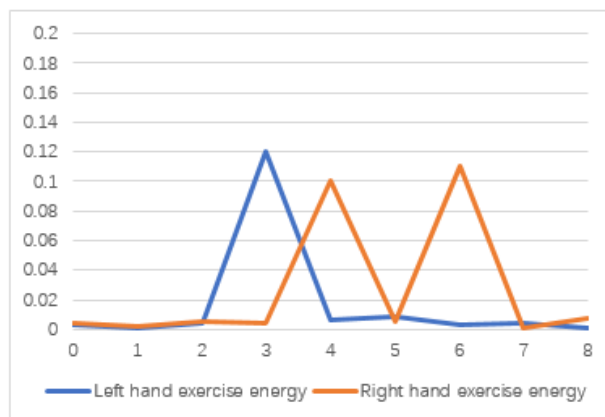
In this paper, the ERP data of C3 and C2 were selected because the features in these two channels of the imagination of left and right hand movements are the most obvious. In addition, in order to verify the validity of the extracted features and the classifiers in this paper, the test sets without class labels were discarded. The final extracted data is: 140\*2350, a total of 140 signals, and the meaning of each column was shown in Table 2.

**Table 2.** Structure of electrified ERP data

Column index	meaning
1	Class labels
2:1153	C3
1154:2350	C4

Firstly, the time domain interception was conducted on the extracted ERP data. On the one hand, the AR model is suitable for the processing of static stationary signals. On the other hand, the

event-related desynchronization (ERD) and the event-related synchronization (ERS) phenomena were relatively significant in the period of 3.5 to 6 seconds, as is shown in Figure 2. Therefore, the 3.5 to 6 s of the original data was intercepted as the subsequent processing signals.



**Figure 2.** Visualize the signal energy of C3 and C4 channels when imagining left and right hands

After many experiments, the order of the AR model was determined to be 10, and then the parameters matrix of channel C3 and C4 of the model obtained from the least square method were 140 matrix, recorded as coef3, coef4. The 140\*10 feature set was obtained by the subtraction of channel parameters. In addition, the mean, median and standard deviation of these 10 parameters were calculated as additional features. Finally, the 140\*13 feature set was obtained, recorded as FeaAR.

For the time domain signal after the interception, its energy was calculated. Then, the db10 wavelet function was used for the 3-layer wavelet decomposition of the obtained energy. The decomposition layers and the corresponding frequency were shown in Table 3.

**Table 3.** Decomposition layers and corresponding frequencies

Decomposition signal	Frequency range \ Hz	Number of layers
original	0-128	0
cA1	0-32 (32-64)	1
cA2	0-16 (16-32)	2
cA3	0-8 (8-16))	3

Among them, the approximation component cA3 contained the overall information of the original signal. From the perspective of the energy, the reconstructed signal A3 could be used as a feature to distinguish the two types of motion; cD3



contained the  $\alpha$  (8-12Hz) band; cD2 contained the  $\beta$  (18-24 Hz) band. The ERD/ERS of ERP signals mainly occurred in these two bands. Therefore, cD2 and cD3 were selected for the reconstruction to obtain D2 and D3. The reconstructed signals A3, D3 and D2 (the dimensions were both 280\*320; the first 140 behaviors were C3 and the last 140 behaviors were C2) were stored in the form of two channels respectively, as is shown in Table 4.

**Table 4.** Signals of C3 and C4 after reconstruction

Reconstruct signal (280*340)	C3 (140*320)	C4 (140*320)
A <sub>3</sub>	A <sub>3</sub> C <sub>3</sub>	A <sub>3</sub> C <sub>4</sub>
D <sub>3</sub>	D <sub>3</sub> C <sub>3</sub>	D <sub>3</sub> C <sub>4</sub>
D <sub>2</sub>	D <sub>2</sub> C <sub>3</sub>	D <sub>2</sub> C <sub>4</sub>

### Feature selection

The parameters adopted were population size P=100, evolutionary generation T=150, and selection pressure q=0.8. The selection of these three parameters was based on the general recommendation of the literature. The penalty coefficient w=0.2, the crossover probability P<sub>c</sub> and the mutation probability P<sub>m</sub> were calculated separately. The new parameters k1 and k2 followed the trial and error method, and finally k1=0.5 and k2=0.3.

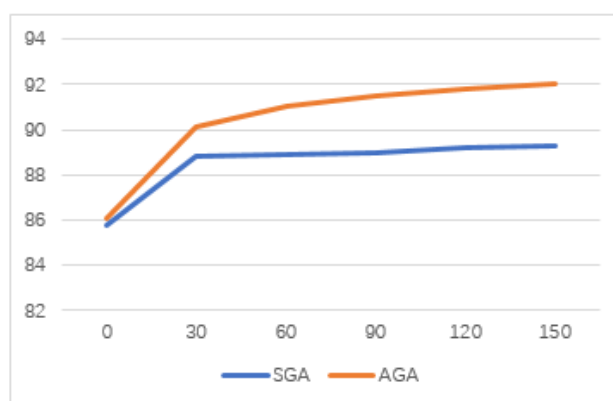
The specific implementation steps of SGA-FS were the same as AGA-FS. Its parameter selection is similar to AGA-FS, but the P<sub>c</sub> and P<sub>m</sub> were fixed values. Generally, the selection range of crossover probability P<sub>c</sub> was 0.5 to 1.0 and the selection range of mutation probability P<sub>m</sub> was 0.001 to 0.005. Finally, P<sub>c</sub>=0.8, P<sub>m</sub>=0.03.

Separate optimal feature combination, criterion value J(w) of the exclusive use of each feature was calculated and lined up. The corresponding first d features were selected as the selection result. This strategy was called F-FS(1); the sequential advancing method was the simplest top-down searching method, in which a feature was selected from the unselected features at a time, so that the J(w) was the maximum when it was combined with the selected features until the number of features increased to d. This strategy was called F-FS(2).

In the above several schemes, the feature subset size d=5,10,15,20,...,45,50 was adopted. The initial feature was based on adaptive personality feature extraction of wavelet packet optimal basis and the initial feature vector obtained by experimental data Dataset Ia was X. The dimension was 130 (namely, D=130). The

classifier used the probabilistic neural network (PNN) and the smoothing coefficient  $\delta$  took 1.

Figure 3 was the variation curve of the classification accuracy of SGA and AGA with genetic algebra when the feature vector dimension d=20 (Vine and Wilson, 2010). It could be seen that AGA has relatively high classification accuracy and fast convergence rate than SGA and other feature dimensions had similar results.



**Figure 3.** The curve of classification accuracy with genetic evolution

The feature selection not only allows us to obtain the optimal subset of features, but also allows the program to run multiple times. The selection frequency of each feature is calculated (the ratio of the number of features selected and the total number of feature selection), and then it is arranged in a descending order. The effectiveness of each feature is evaluated based on this, and then the effectiveness of each ERP channel can be evaluated.

### Conclusion and outlook

This paper uses the feature selection based on adaptive genetic algorithm based (AGA-FS) and apply it to BCI. To verify the validity of this method, it is compared with the feature selection based on standard genetic algorithm (SGA-FS) and the feature selection of filter algorithm based on Fisher distance (F-FS). Among these schemes, SGA-FS and AGA-FS use the classification accuracy as the index to evaluate different feature subsets, so it is expected to obtain relatively good classification results. SGA-FS uses the standard genetic algorithm for feature selection. Due to inherent defects of SGA, such as premature and local convergence, its performance is slightly worse than that of AGA. The experimental results show that this method obtains the best



classification recognition performance. For the processing of four types of motion imagination ERP signals, the event-related synchronization / desynchronization phenomenon is firstly analyzed and demonstrated, and then the CSP and wavelet analysis are used for feature extraction respectively. It is found that the classification effect is not satisfactory. Based on this, the paper proposes further screening of the ERP feature signals.

Although many scholars are currently carrying out researches on the classification of spontaneous ERP thinking mode and have achieved certain results, attempts to apply such research in the real-time communication of brain-computer interface are still confined to the laboratory research phase. There is still a big gap between the communication speed and the normal processing capacity of human beings (Lombardi *et al.*, 2015). The extraction of signal features is an important factor influencing the classification accuracy of thinking activities.

## References

- Charlton MH. Human brain and psychological processes. *Archives of Neurology* 1966; (6): 672.
- Folstein JR, Van PC. Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology* 2008; 45(1): 152-70.
- Harle SK, Vickers JN. Training quiet eye improves accuracy in the basketball free throw. *Sport Psychologist* 2001; 15(3): 289-305.
- Janelle CM. Anxiety, arousal and visual attention: a mechanistic account of performance variability. *Journal of Sports Sciences* 2002; 20(3): 237-51.
- Lombardi R, Lardo A, Manfredi S, Nappo F. Changing the university system of management: a study on the Italian scenario. *Journal of Innovation & Entrepreneurship* 2015; 4(1): 1-9.
- Loze GM, Collins D, Holmes PS. Pre-shot EEG alpha-power reactivity during expert air-pistol shooting: a comparison of best and worst shots. *Journal of Sports Sciences* 2001; 19(9): 727-33.
- Orgs G, Dombrowski JH, Heil M, Jansen-Osmann P. Expertise in dance modulates alpha/beta event-related desynchronization during action observation. *European Journal of Neuroscience* 2008; 27(12): 3380-84.
- Vine SJ, Wilson MR. Quiet Eye Training: Effects on Learning and Performance Under Pressure. *Journal of Applied Sport Psychology* 2010; 22(4): 361-76.