Automated Feature Selection Based on an Adaptive Genetic Algorithm for Brain-Computer Interfaces

Guo-zheng Yan, Ting Wu, and Bang-hua Yang

School of Electronic, Information and Electrical Engineering,
Shanghai Jiao Tong University, 200240 Shanghai, China
{gzhyan, wu_ting, ybh}@sjtu.edu.cn

Abstract. In brain-computer interfaces (BCIs), a feature selection approach using an adaptive genetic algorithm (AGA) is described in this paper. In the AGA, each individual among the population has its own crossover probability and mutation probability. The probabilities of crossover and mutation are varied depending on the fitness values of the individuals. The adaptive probabilities of crossover and mutation are propitious to maintain diversity in the population and sustain the convergence capacity of the genetic algorithms (GAs). The performance of the AGA is compared with those of the Standard GA (SGA) and the Filter method in selecting feature subset for BCIs. The results show that the classification accuracy obtained by the AGA is significantly higher than those obtained by other methods. Furthermore, the AGA has a higher convergence rate than the SGA.

1 Introduction

Brain-computer interfaces (BCIs) are devices intended to help disabled people communicate with a computer using the brains' electrical activity. The electrical activity can be measured by electroencephalogram (EEG) [1]. Most BCIs make use of spontaneous mental activities (e.g., thinking on moving a finger, the hand, or the whole arm, etc.) to produce distinguishable electroencephalogram (EEG) signals [2], [3]. The distinguishable EEG signals are then transformed into external actions. Over the past years a variety of evidences have evaluated the possibility to recognize a few mental tasks from EEG signals [4], [5]. However, how to improve the recognition performance of EEG signals in signal processing is still a key problem. This paper will focus on a feature selection approach.

Feature selection is the problem of selecting a subset of d features from a set of D ($D>d$) features based on some optimization criterion. An automated feature selection is crucial for classification because irrelevant features or redundant information are known to cause the classifier to have poor generalization, increase the computational complexity and require more training samples. Various kinds of possible features (autoregressive parameters, power spectral density, averages, wavelet packet energy, etc.) are used for classifying the EEG signals, but the most effective features remain unclear. So, the algorithms which can find a good approximation to the best subset need to be developed.

The most common algorithms for feature selection include Filter algorithms and Wrapper algorithms. The main disadvantage of the Filter algorithms is that it selects

T.-D. Wang et al. (Eds.): SEAL 2006, LNCS 4247, pp. 575 – 582, 2006.

feature subsets that are independent of classification algorithms and ignores the effects of the selected feature subset on the performance of the classification algorithm. As a kind of wrapper method, the genetic algorithm (GA) is often used to perform feature selection.

In the feature selection algorithms for BCIs, there are some reported applications [6], [7] which are based on SGA, but there are very few reported application based on AGA. However, in most cases, the AGA outperforms the SGA significantly [8]. We will explore an adaptive GA (AGA) method. It is compared with those of the Standard GA (SGA) and the Filter method in selecting feature subset for BCIs.

2 Dataset and Feature Extraction

2.1 Dataset

All data were acquired from six healthy subjects (three male and three female, 22-35 years old). The subjects were asked to move a cursor up and down (two mental activities) on a computer screen, while his slow cortical potentials (SCPs) were taken. Each trial lasted 6s and consisted of three phases: a 1s rest phase, a 1.5-s cue presentation phase, and a 3.5-s feedback phase. The cue presentation is a visual target appearing either at the top or bottom. Data were recorded during the 3.5-s feedback at a sampling rate 256Hz. The feedback is provided by a cursor whose vertical position indicated the current level of SCPs (Cz-Mastoids). The following six channels of EEG data were recorded (denotation follows the 10/20 system):

Ch1: A1-Cz (A1 = left mastoid)	Ch2: A2-Cz (A2 = right mastoid)
Ch3: (2 cm frontal of C3)-Cz	Ch4: (2 cm parietal of C3)-Cz
Ch5: (2 cm frontal of C4)-Cz	Ch6: (2 cm parietal of) C4-Cz

2.2 Feature Extraction

Three common types of feature extraction methods were used in this paper.

(1) Autoregressive model coefficients (AR): The autoregressive coefficients of 3 orders, obtained using the Yule-Walker method [9]. We can write the AR features of Ch1 as $\{f_1, f_2, f_3\}$, Ch2 as $\{f_4, f_5, f_6\}$, … , Ch6 as $\{f_{16}, f_{17}, f_{18}\}$.

(2) Average coefficients (AC): We select db4 wavelet functions to decompose the EEG signals up to sixth level giving 64 (2^6) sub-bands. The first 25 sub-bands whose frequencies lower than 50Hz are adopted and the other sub-bands are discarded as useless information. Average coefficients values of the 25 sub-bands are calculated. So, we can obtain 25-Dimensional AC features for each single channel. The AC features from Ch1 to Ch6 can be written as $\{f_{19}, \ldots, f_{43}; f_{44}, \ldots, f_{68}; f_{144}, \ldots, f_{168}\}$.

(3) Average energies (AE): The process of obtaining AE is similar with AC. We also can obtain 25-Dimensional AE features for each single channel. The AE features from Ch1 to Ch6 can be written as $\{f_{169}, \ldots, f_{193}; f_{194}, \ldots, f_{218}; f_{294}, \ldots, f_{318}\}$.

All the features can be written as $U=\{f_1, f_2, \ldots, f_{318}\}$.