Study on EEG Signal Analysis and Command Recognition Based on Deep Learning

Bowen Yao, Siyuan Gao, Shuo Zhang
Zhejiang University

Abstract

Electroencephalogram (EEG) analysis has wide prospects for brain science, rehabilitation engineering, automatic control, military and biomedical engineering in recent year. EEG signals recorded from sensorimotor areas during mental imagination of specific movements are classified by brain-computer interface (BCI) to recognize some subject-specific EEG patterns.

This paper evaluates various methods for EEG feature extraction and classification. Multi-variable Adaptive Autoregressive Model (AAR) model is applied to extract feature information from the EEG signals, with Kalman filtration method to estimate the coefficients. For classification, linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and artificial neural network (ANN) are mainly discussed in the paper. Results are compared and some remarks are given.

Key words: EEG, BCI, AAR, Kalman filtration, LDA, QDA, Artificial neural network

Ⅰ.Introduction

Brain-computer interface (BCI) is a direct way of communication between brain and an external device, based on recognition of subject-specific electroencephalogram (EEG) patterns. This technical system can be used for supporting communication possibilities for patients with severe neuromuscular disabilities, who are in particular demand of gaining reliable control via nonmuscular devices. As an advanced method of human-computer interaction, BCI even has broader applications. For people with a special working environment, such as astronauts and aquanauts, BCI provides them a brand new way of operating peripheral equipments. As for general public, this technology enables human mind to directly control devices and communicate with external environment.

The BCI researches mainly aim at improving the accuracy and robustness of classifying the input features, such as electrode positions and frequency bands. The pattern recognition procedure can be separated into two steps: (1) extraction of EEG features, and (2) classification of those extracted features into target output.

According to different signal detection methods, BCI can be divided into invasive BCIs and non-invasive BCIs. The former refers to those devices with electrodes implanted inside the skull.

It certainly provides better signals, but also facing the problem of infection and the complexity of operation. The EEG from the non-invasive BCIs contains more interference signals; however, due to the safety and simplicity of operation, it is more widely used than the invasive ones.

Because of the large amount of information produced in every trial and the existence of noisy interference signals, extracting EEG features is of great importance. Some methods are based on time and frequency domain analysis, such as chaos analysis, filter in spatial domain, spectrum analysis, fast Fourier transformation, and wavelet analysis, etc. Considering the intertemporal correlation of the information, we focus on methods derived from autoregressive model (AR model), among which the adaptive autoregressive model (AAR model) is finally chosen.

An important step toward real-time processing and feedback presentation is the setup of a subject-specific classifier. Several machine learning algorithms can be applied to this task, for example, linear discriminant analysis (LDA), neural network, support vector machine (SVM), discriminant analysis based on Mahalanobis distance (MDA), and leave-one-out method, etc. The different EEG classification methods have been compared, and the neural network is shown to be the most suitable algorithms for the dataset used in our research.

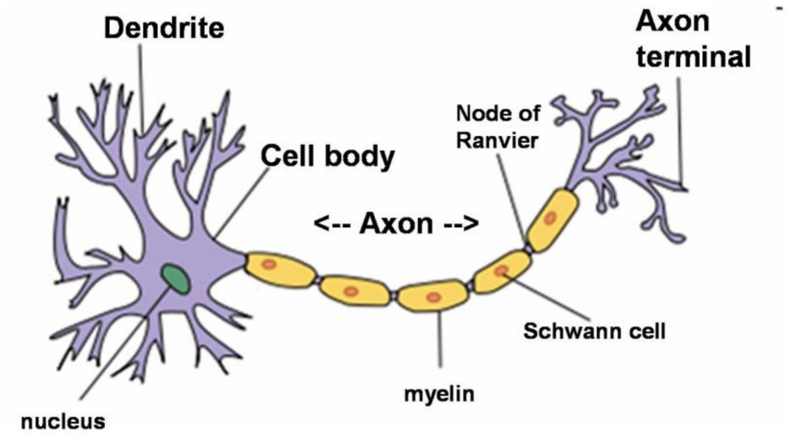
Ⅱ. Backgrounds

A. The generation of EEG signals

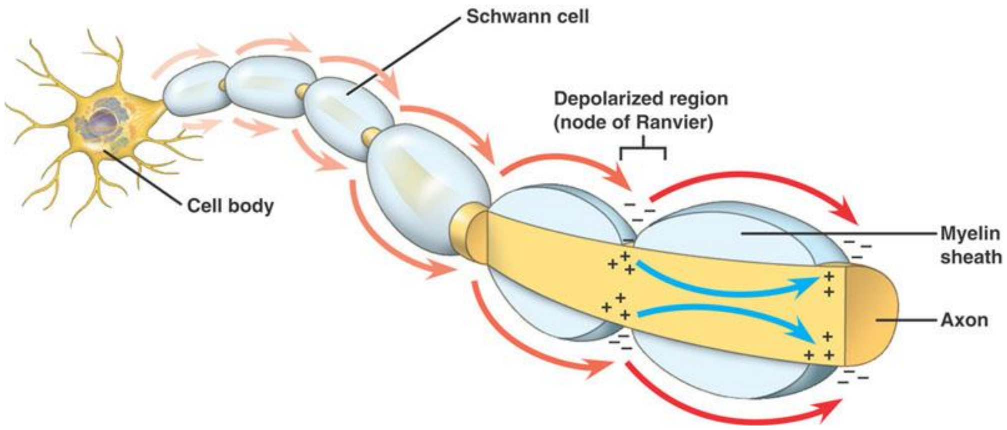
The nervous system is the part of an animal's body that coordinates its voluntary and involuntary actions and transmits signals between different parts of the body. The human nervous system contains of two parts, the central nervous system (CNS) and the peripheral nervous system (PNS). The CNS includes the brain and spinal cord, while the PNS consists of the nerves and ganglia outside pf the brain and spinal cord [1].

The brain, an organ serving as the center of the human nervous system, exert centralized control over the other organs of the body, issuing instruction to the body to complete voluntary actions, or allowing rapid and coordinated responses to changes in the environment. The outer layer of the brain is called the cerebral cortex, which is estimated to contain 15-33 billion neurons and has functional roles in motor and sensation [2]. Different parts of the cerebral cortex, defined as Brodmann areas, are correlated to different cortical functions. By the chance of surgeries, neurophysiologists are able to stimulate parts of the cerebral cortex, in order to find the correlation between Brodmann areas and physiological functions.

As the core components of the nervous system, neuron, also known as neurone or nerve cell, is an electrically excitable cell that processes stimulations and transmits information through electrical and chemical signals. A typical neuron possesses a cell body (soma), dendrites, and an axon. The cell body of a neuron often gives rise to multiple dendrites, but only one axon, which could branch hundreds of times before it terminates [3]. Neurons are connected by their dendrites and axons, with the several junctions named synapses.



Neurons are electrically excitable cells that use electrical signals as a way of communication. All animal cells are surrounded by a membrane acting as both an insulator and a diffusion barrier to the movement of ions. Because of the differences of the type and density of the ions inside and outside the membrane, there is a difference in electric potential between the interior and the exterior of a biological cell, stated as membrane potential [4]. When a cell is stimulated, the electrical membrane potential of a cell rapidly rises and falls across the cell membrane, producing an electric current along the dendrites and axons, further spreading explosively to all the cells that are connected, a greater electric current generated.



Electroencephalogram (EEG) measures voltage fluctuations resulting from ionic current flows within the neurons of the brain. From each electrodes attached to the human scalp, a record of the oscillations of electric potentials over time can be obtained. EEG is such a diagram with time as the horizontal axis, and electrical potential the vertical axis.

B. The features of EEG signals

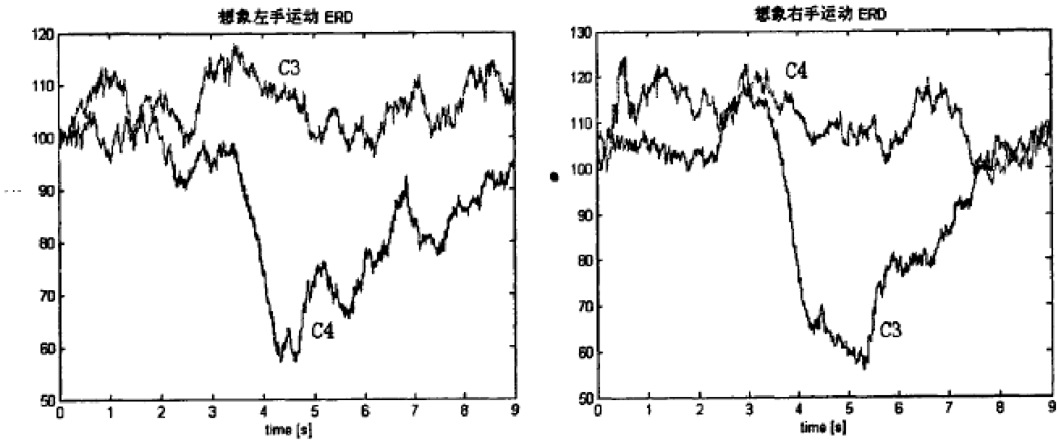
The EEG signals have a broad frequency bandwidth, but what are usually focused on clinically are those between 0.5-35Hz. Empirically, they are sorted as follows:

a) δ rhythm, with a frequency between 0-4Hz, is usually associated with the deep stage of sleep, also known as slow-wave sleep [5].

b) θ rhythm, has a frequency between 4-7Hz. Cortical θ rhythm is observed frequently in young children. In older children and adults, it tends to appear during meditative, drowsy, or sleeping states, but not during the deepest stages of sleep. Several types of brain pathology can give rise to abnormally strong or persistent cortical θ rhythm [6].
c) α rhythm, is neural oscillation in the frequency range of 7.5-12.5Hz. It predominantly originates from the occipital lobe during wakeful relaxation with closed eyes, and is reduced with open eyes, drowsiness and sleep [7].
d) μ rhythm, at a frequency of 7.5-12.5Hz, is most prominent when the body is physically at rest. It is found over the motor cortex, in a band approximately from ear to ear. It is suppressed when the body performs a motor action or, with practice, when the subject visualizes performing a motor action [8].
e) β rhythm, with a frequency range of 12.4-30Hz, is associated with waking consciousness, especially with active, busy, or anxious thinking and active concentration [9].

The EEG signals have following features:
a) The EEG signals are faint compared to the interference signals, such as the signals of blinking eyes, mental or muscular tension.
b) The signals have complicated components. Besides EEG, there are a large number of other signals like electrocardio (ECG). Some of the interference signals have been explained, while some are still remained unrecognized.
c) The EEG signals have a clear spectrum feature, under the condition of which, methods based on spectrum analysis are well applied.
d) The EEG signals are unstationary stochastic process. Its randomness and unstationarity derive from not only the irregular activities of the related physiological structures, but also the reaction to the unexpected changes in the environment. Thus the statistical methods, such as time series analysis should be considered.

The human brain is separated into two distinct cerebral hemispheres, and observes a principle of cross-control, which is, the left-hemisphere dominates the motor of the right-body, and vise versa. Therefore, when the subject performs, even imagine left hand or right hand movement, the EEG signals from the two hemispheres could be of significance difference. Researches have shown that, during physical and motor imagery of left or right hand movements, α rhythm and β rhythm of the contralateral hemisphere attenuates obviously. This phenomenon is called event-related desynchronization (ERD). Conversely, after the movement and the visualization finish, the amplitude of α rhythm and β rhythm will increase to normal. This is event-related synchronization (ERS). ERD/ERS behaves different between voluntary and involuntary motors, also among the movements from different part of the body, for instance, the characteristic frequency band on which the ERD/ERS changes can be different, as well as the corresponding part of the cerebral cortex. Thus, ERD/ERS is able to take an important role as a characteristic for pattern recognition.

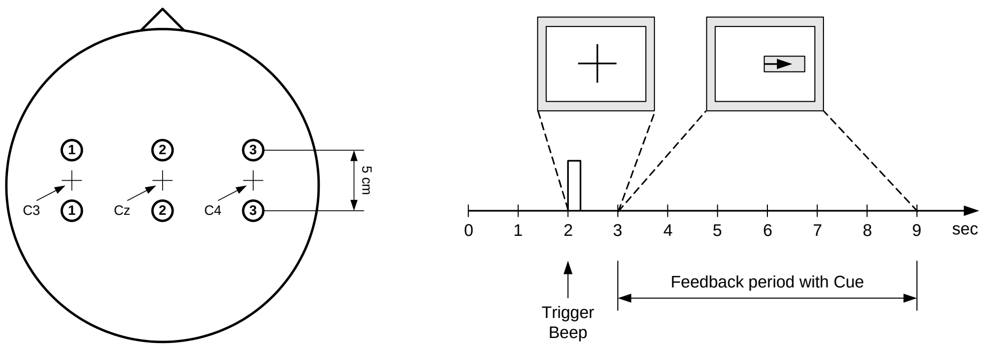


C. Description of the data set

In order to measure the quality of our algorithm and results, we choose a widely used database, the Data SetⅢ of the BCI Competition 2003, which is provided by the Department of medical Informatics, Institute for Biomedical Engineering, University if Technology, Graz, Austria.

The subject is a 25-year-old health female. The task was to control a feedback bar in one dimension by imagination of left or right hand movements. The order of left and right cues was random.

The experiment included 280 trials with the time of 9 seconds each. The first 2s were quiet. At t=2s, an acoustic stimulus indicated the beginning of the trial, and a cross "+" was displayed for 1s. Then, ant t=3s, and arrow (left or right) was displayed as a cue stimulus. The subject was asked to use imagination as described above to move the feedback bar into the direction of the cue. The feedback was based on AAR parameters calculated from channels C3 and C4. The AAR parameters were combined with a discriminant analysis into one output parameter. The recording was mede using g.tec amplifier and Ag-AgCl electrodes. Three bipolar EEG channels were measured over C3, Cz, and C4. EEG was sampled with 128Hz and was filtered between 0.5-30Hz. The trials for training and testing were randomly selected to prevent any systematic effect due to the feedback. [10]



Ⅲ. Extraction of EEG features

A. The general model

Considering the intertemporal correlation of the information, we focus on time series models. In real practice, AR model has a wider application than MA model and ARMA model, because the parameters of AR model are rather simply estimated, which provides a better real-time reaction. However, AR model can only be used to describe stationary time series. For event-related EEG, which is not stationary most of the time, an Adaptive Autoregressive Model (AAR) is applied to characterize it, in the following form:

$$u_n = \sum_{i=1}^{p} w_{ni} u_{n-i} + \varepsilon_n$$

$$\varepsilon_n \sim W\left(\mu, \sigma^2\right),\ n = p+1, p+2, \ldots, N$$

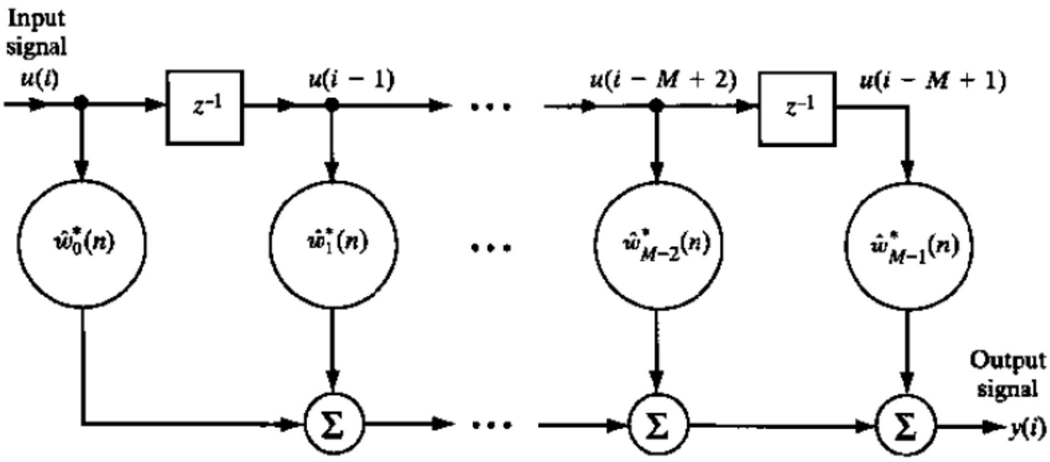
Or it can be written as

$$u_n = w(n)^T u(n) + \varepsilon_n$$

where $w(n) = \left(w_{n1}, w_{n2}, \ldots, w_{np}\right)^T$ is the tap-weight vector at time n

$u(n) = \left(u_{n-1}, u_{n-2}, \ldots, u_{n-p}\right)^T$ is the tap-input vector at time n

The AAR model differs from an AR model in that the parameters w_{ni} can vary with time, and they represent the signal characteristics only of a short period of time. (Even though the EEG signal turns out to be stationary, it is still in the scope of AAR model, as a special case.)



A rough estimate of the model order can be derived from the number of peaks in the spectrum. For each peak, two parameters are needed. Examination of EEG spectra shows a prevalent existence of 2-3 peaks, giving an approximate model order of 6. Therefore, we apply an AAR model with p=4 in our recognition task.

B. Estimation of ARR parameters: Kalman filtration or Recursive Least Squares Algorithm (RSL)

The ε_n term in the model, also called the error process, represents the part that cannot be explained by the model. In general, the EEG signal is described more accurately by a model if the error process is smaller. We defined the Mean Square Error (MSE) as

$$MSE = \frac{\sum_{n=1}^{N} \varepsilon_n^2}{N}$$

And the total power of the EEG is defined as

$$MSY = \frac{\sum_{n=1}^{N} u_n^2}{N}$$

Furthermore, we define the Relative Error Variance (REV) as

$$REV = \frac{MSE}{MSY}$$

which is a measure of how well the AAR model describes the EEG signal.

The Kalman filtration method is mainly about minimize the following function:

$$\underset{w(n)}{\arg\min}\, \xi_n = \sum_{i=1}^{n} \lambda^{n-i} e_i^2 + \delta\lambda^n \left\| w(n) \right\|^2,\ e_i = u_i - w(n)^T u(i) \quad \text{, for each n}$$

$\lambda \in (0,1]$ and λ^{n-i} is a weighted factor that will become smaller when i getting smaller. It is used to ensure that data in the distant past are “forgotten” in order to afford the possibility of following the statistical variations of the observable data when the filter operates in a nonstationary environment. In other words, the closer to time n, the more important the error term is. Similarly, for the term $\delta\lambda^n \left\| w(n) \right\|^2$, an approximate form of regularization, the exponential weighting factor λ means that the beneficial effect of adding $\delta\lambda^n \left\| w(n) \right\|^2$ to the cost function is forgotten with time.

The above target function is equivalent to a system of linear equations written in the form of matrix:

$$\Phi(n)\hat{w}(n) = z(n)$$

where $\Phi(n) = \sum_{i=1}^{n} \lambda^{n-i} u(i) u(i)^T + \delta\lambda^n I,$

$$z(n) = \sum_{i=1}^{n} \lambda^{n-i} u(i) u_i$$

To solve this, we use the method of iteration, so that we can avoid matrix inversion. Noted that

$$\Phi(n) = \lambda\Phi(n-1) + u(n)u(n)^T$$
$$z(n) = \lambda z(n-1) + u(n)u_n$$

We therefore have

$$\pi(n) = P(n-1)u(n)$$
$$k(n) = \frac{\pi(n)}{1-UC + u(n)^T \pi(n)}$$
$$e_n^* = u_n - \hat{w}(n-1)^T u(n)$$
$$\hat{w}(n) = \hat{w}(n-1) + k(n)e_n^*$$
$$P(n) = (1-UC)^{-1} P(n-1) - (1-UC)^{-1} k(n)u(n)^T P(n-1)$$
$$UC = 1-\lambda$$

UC is called the update coefficient. $(1-UC)^{-1} = \lambda^{-1}$, a value slightly larger than 1, controls the adaptation ratio. The selection of an appropriate UC is essential for AAR models. We update the UC based on the final accuracy rate after classification. The detailed selection procedure will be illustrate in 'Practice and results'.

For each sub-AAR model (with u_n at the left-hand side of the equation), the estimated parameters can form a vector as the feature vector for the time point n. Thus, for each time point n, we can obtain 140 feature vectors from 140 randomly selected training trials. Also, for each time point n, we use the group of 140 training feature vector to do the classification following.

Ⅳ. **Classification of extracted feature vectors**

Method 1: Linear Discriminant Analysis (LDA)

The extracted feature vectors from the training data set are denoted as $\vec{x}$, each with known class y (moving towards left or right). LDA aims to find a good predictor for the class y of any observation $\vec{x}$ from the same distribution. LDA approaches the problem by assuming that the conditional probability density functions $p(\vec{x} \mid y=0)$ and $p(\vec{x} \mid y=1)$ are both normally distributed:

$$p(\vec{x} \mid y=0) \sim N(\overrightarrow{\mu_0}, \Sigma_0)$$

$$p(\vec{x} \mid y=1) \sim N(\overrightarrow{\mu_1}, \Sigma_1)$$

also with the homoscedasticity assumption $\Sigma_0 = \Sigma_1 = \Sigma$, with full rank.

Under this assumption, the Bayes optimal solution is to predict points as being from the second class if the log of the likelihood ratios is below some threshold T, i.e.

$$(\vec{x}-\overrightarrow{\mu_0})^T \Sigma_0^{-1}(\vec{x}-\overrightarrow{\mu_0}) + \ln|\Sigma_0| - (\vec{x}-\overrightarrow{\mu_1})^T \Sigma_1^{-1}(\vec{x}-\overrightarrow{\mu_1}) - \ln|\Sigma_1| < T$$

further under the homoscedasticity assumption, the criterion becomes

$$\vec{\omega}\cdot\vec{x} > c$$

where $\vec{\omega} \propto \Sigma^{-1}(\overrightarrow{\mu_1}-\overrightarrow{\mu_0})$

$$c = \frac{1}{2}(T - \overrightarrow{\mu_0}^T \Sigma_0^{-1}\overrightarrow{\mu_0} + \overrightarrow{\mu_1}^T \Sigma_1^{-1}\overrightarrow{\mu_1})$$

Method 2: Quadratic Discriminant Analysis (QDA)

QDA is an extension of LDA, releasing the homoscedasiticity assumption, i.e. the two covariance matrixes are not necessarily identical. Then the discriminant statistic and criterion is given by

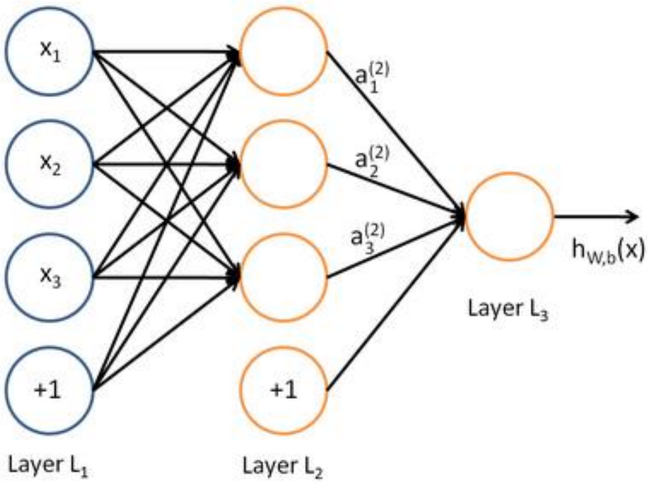
The likelihood ratio= $$\frac{\sqrt{2\pi|\Sigma_1|}^{-1}\exp(-\frac{1}{2}(x-\overrightarrow{\mu_1})^T \Sigma_1^{-1}(x-\overrightarrow{\mu_1}))}{\sqrt{2\pi|\Sigma_0|}^{-1}\exp(-\frac{1}{2}(x-\overrightarrow{\mu_0})^T \Sigma_0^{-1}(x-\overrightarrow{\mu_0}))} < t$$

for some threshold t.

Method 3: Artificial Neural Network (ANN)

The idea of ANN is derived from biological neural network. In a human brain, neural network has an estimated 1011 highly interconnected neurons acting in parallel. After receiving electrochemical signals from other neurons, the neuron is activated and sends signals to other neurons.

The word *network* in the term 'artificial neural network' refers to the inter–connections between the neurons in the different layers of each system. An example system has three layers. The first layer has input neurons which send data via synapses to the second layer of neurons, and then via more synapses to the third layer of output neurons. More complex systems will have more layers of neurons with some having increased layers of input neurons and output neurons. The synapses store parameters called "weights" that manipulate the data in the calculations.

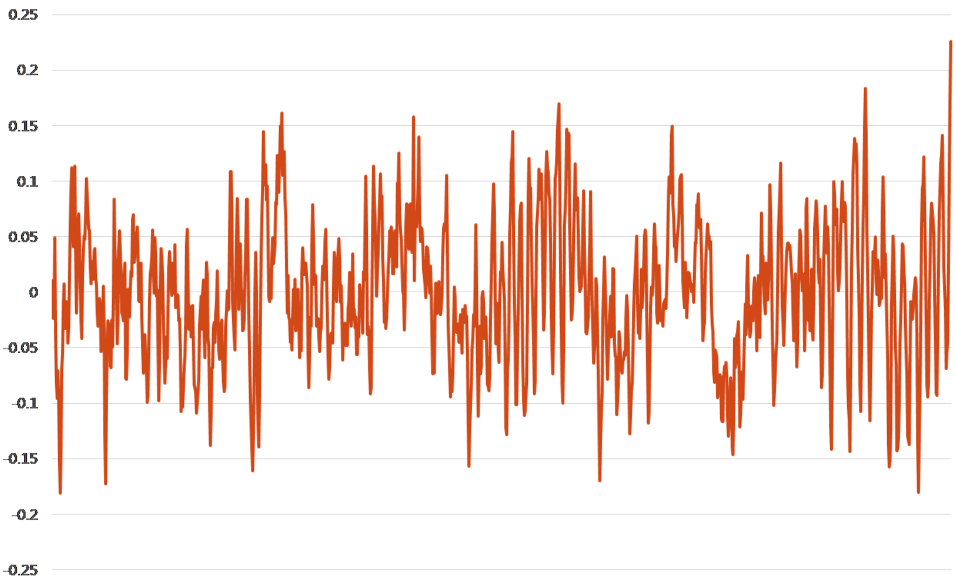


An ANN is typically defined by three types of parameters:

1. The interconnection pattern between the different layers of neurons
2. The learning process for updating the weights of the interconnections
3. The activation function that converts a neuron's weighted input to its output activation.

Ⅴ. Practice and results

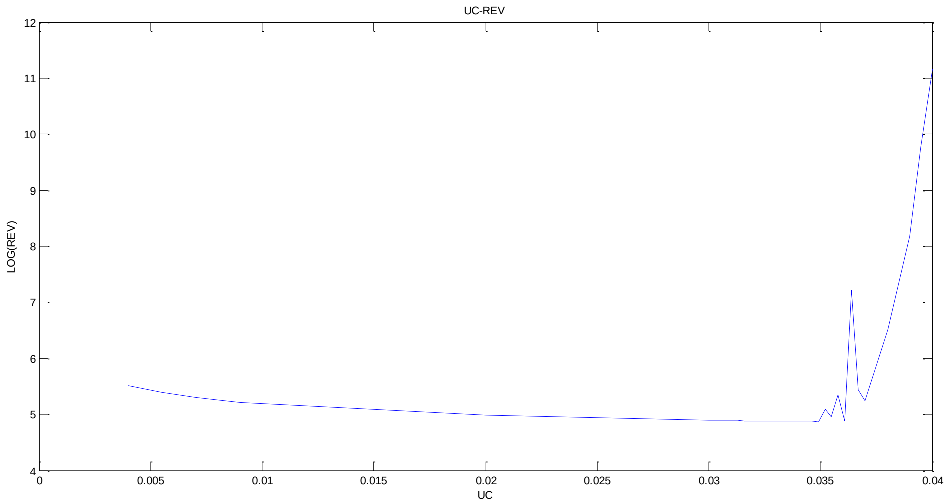
We plot the EEG signals of the 140 training trials from Data SetⅢ of the BCI Competition 2003. The time series are as following:



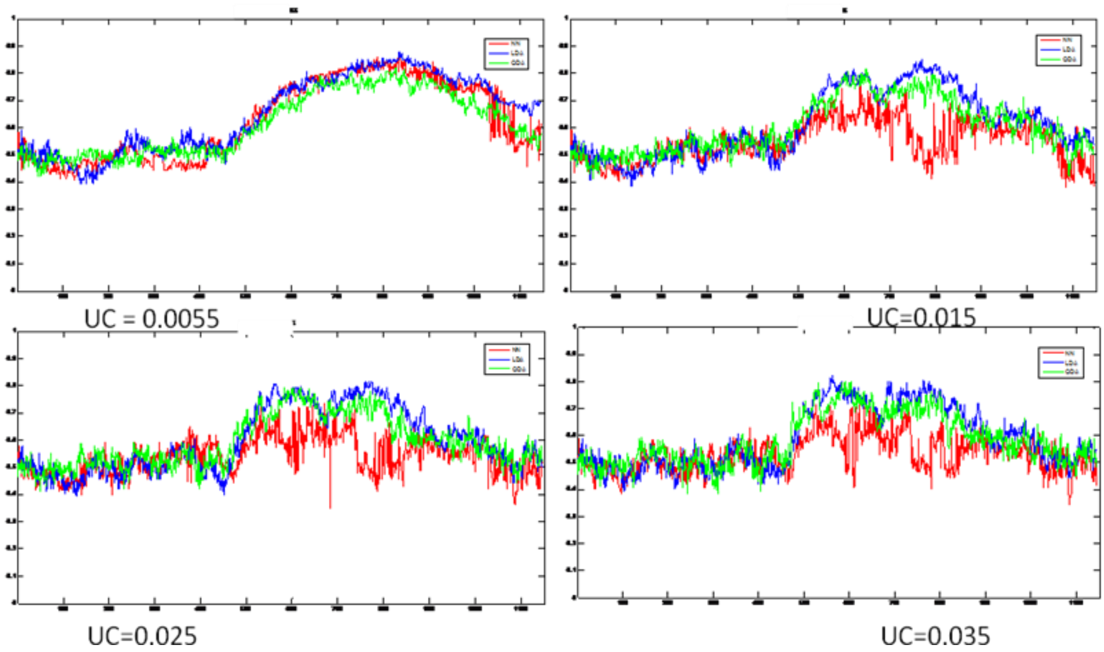
A. Selection of Update Coefficient (UC)

The selection of an appropriate UC is essential for AAR models. We update the UC based on the final accuracy rate after classification.

Firstly, using different value of UC to form AAR model, we get the relationship between the Relative Error Variance (REV) and the UC. A diagram of LOG(REV) against UC is shown below:



From the diagram, we can see that UC=0.035 gives the smallest REV, after which the LOG(REV) begins to fluctuate upwards sharply. However, when UC<0.035, the LOG(REV)s are rather small. Therefore, we select UC=0.0055, 0.015, 0.025, 0.035 to do the classification respectively. And the accuracy rates are given below, with the red line representing the classification method of neural network, blue line the linear discriminant analysis, and green line the quadratic discriminant analysis. (For each combination of AAR and classification method, we can get a time series of accuracy rates. This is because for each time point n, a group of feature vectors are extracted from all trials, then a classification method is applied to the training vectors, thus different accuracy rates are generalized for every time point n, forming a time series of accuracy rate.)

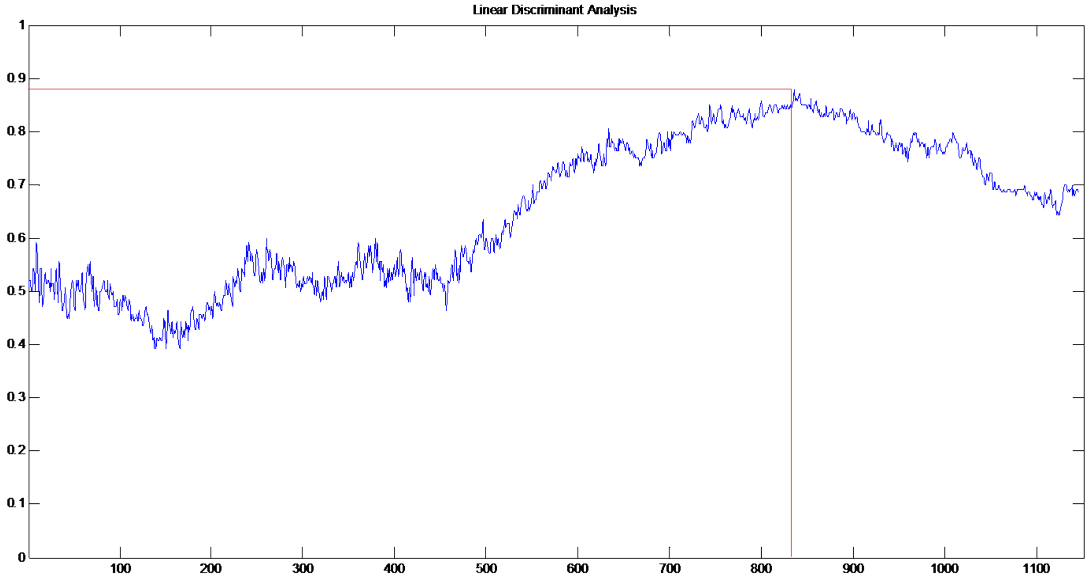


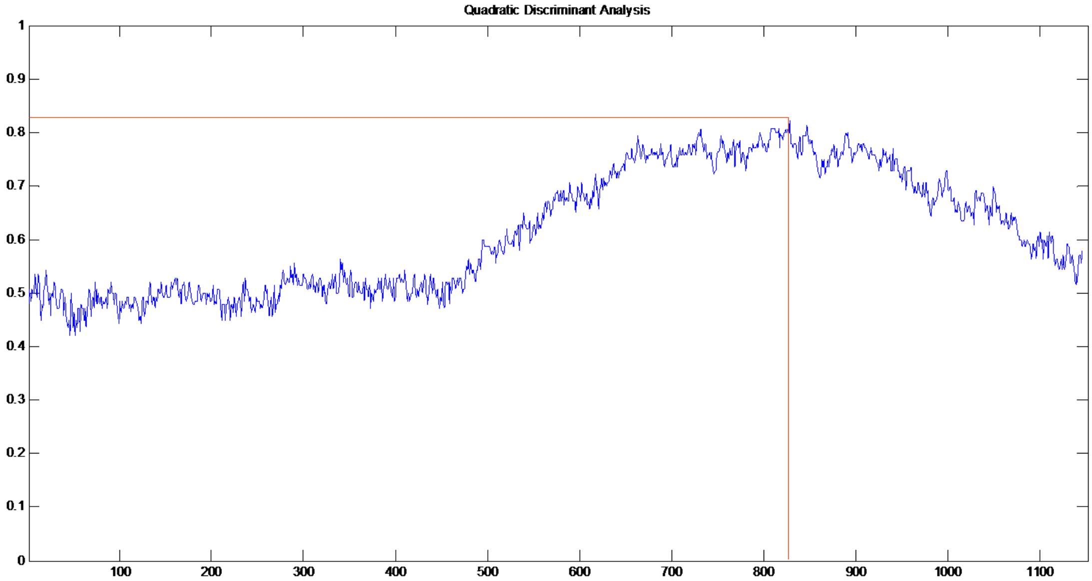
Secondly, from the results we can see that, although UC=0.0035 gives the smallest REV of AAR model, the final classification accuracy is not as stable as UC=0.0055, especially for the ANN

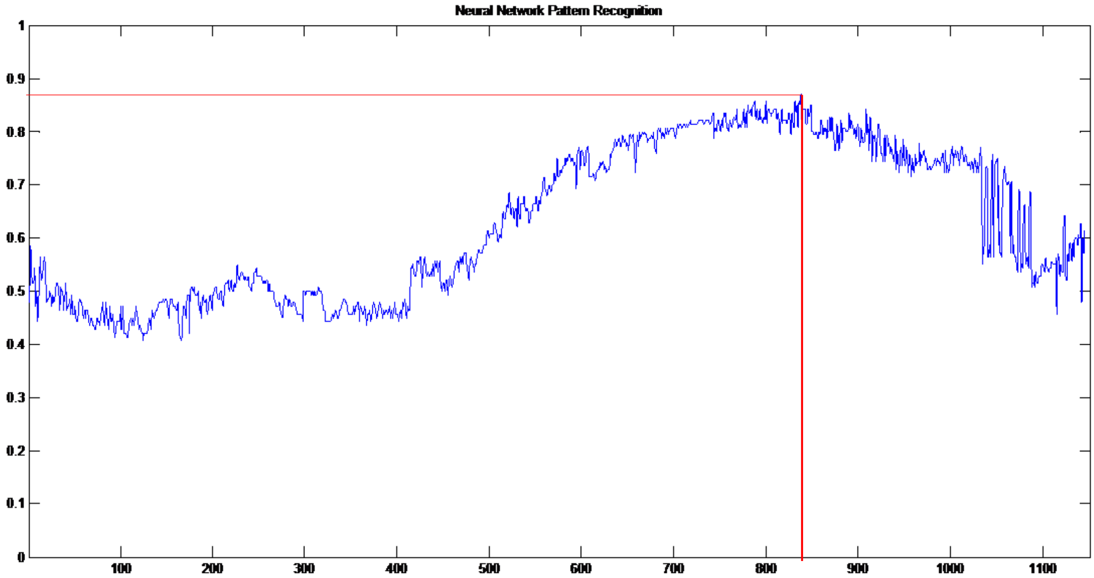
classification method. Under UC=0.0055, the accuracy rate is not only the most stable, but also the highest for all three classification methods in a global aspect.
Therefore, UC=0.0055 is selected.

B. Application of classification methods

The accuracy rates respect to different classification method are shown below:







	LDA	QDA	ANN
Mean	0.6507	0.6136	0.6282
Max	0.8786(6.53s)	0.8214(6.47s)	0.8714(6.55s)
Std	0.1377	0.1144	0.1434

C. Remarks

a) The maximum of accuracy rate for LDA, QDA, and ANN are all greater than 80%, which indicates that the extraction of signal features based on AAR model and Kalman filtration is efficient and robust.

b) From the final accuracy rate, LDA (87.86%) and ANN (87.14%) is better than QDA (82.14%). However, ANN is more time-consuming and less stable than the other two methods. Looking at the factors above, LDA is the optimum classification method for this specific recognition task.

c) Theoretically, QDA is a more generalized method than LDA, thus should have a better outcome than LDA. But this is not the case in our research. Maybe there is some underlying problem in our method needed to find out and the model needs to be further improved.

d) The optimal time for classification is different within three classification methods, and among different people as well, but the maximum of accuracy rates are all appear between 4-6s. This time period should be attached more significance in practical application.

e) Selecting the value of UC so that to reach a relative small REV is an effective way. However, simply chosing an UC with the minimum REV might lead to overfitting. Better to consider the whole recognition process as a whole to decide the value of UC.

References:

[1] http://en.wikipedia.org/wiki/Membrane_potential

[2] http://en.wikipedia.org/wiki/Delta_wave

[3] http://en.wikipedia.org/wiki/Theta_rhythm

[4] http://en.wikipedia.org/wiki/Alpha_wave

[5] http://en.wikipedia.org/wiki/Mu_wave

[6] http://en.wikipedia.org/wiki/Beta_wave

[7] The BCI Competition 2003: Progress and Perspectives in Detection and Discrimination of EEG Single Trials.

[8] http://en.wikipedia.org/wiki/Linear_discriminant_analysis

[9] http://en.wikipedia.org/wiki/Quadratic_classifier

[10] http://en.wikipedia.org/wiki/Artificial_neural_network

[11] Schlögl A, Flotzinger D, Pfurtscheller G. Adaptive Autoregressive Modeling used for Single-trial EEG Classification-Verwendung eines Adaptiven Autoregressiven Modells für die Klassifikation von Einzeltrial-EEG-Daten[J]. Biomedizinische Technik/Biomedical Engineering, 1997, 42(6): 162-167.

[12] Haykin S S. Adaptive filter theory[M]. Pearson Education India, 2007.

[13] Jansen B H, Bourne J R, Ward J W. Autoregressive estimation of short segment spectra for computerized EEG analysis[J]. Biomedical Engineering, IEEE Transactions on, 1981 (9): 630-638.

[14] Pfurtscheller G, Neuper C, Guger C, et al. Current trends in Graz brain-computer interface (BCI) research[J]. IEEE Transactions on Rehabilitation Engineering, 2000, 8(2): 216-219.

[15] Pfurtscheller G, Neuper C. Motor imagery and direct brain-computer communication[J]. Proceedings of the IEEE, 2001, 89(7): 1123-1134.

[16] Puskorius G V, Feldkamp L A. Neurocontrol of nonlinear dynamical systems with Kalman filter trained recurrent networks[J]. Neural Networks, IEEE Transactions on, 1994, 5(2): 279-297.

[17] Pfurtscheller G, Neuper C, Schlogl A, et al. Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters[J]. Rehabilitation Engineering, IEEE Transactions on, 1998, 6(3): 316-325.

[18] 邹清. 基于 AR 模型的脑电信号特征提取与识别[D]. 中南大学, 2008.