TELKOMNIKA, Vol. 11, No. 5, May 2013, pp. 2381~ 2386
ISSN: 2302-4046

Study on Fisher Analysis of Electroencephalograph Data

Yuan Shi*, Qi Wei, Ruijie Liu, Yuli Ge
Dalian Institute of Science and Technology
Dalian Lvshun Economic Development Zone
*Corresponding author, e-mail: 20088041@qq.com

Abstract

Objective in this paper, we have done Fisher discriminant analysis to Electroencephalogram (EEG) data of experiment objects which are recorded impersonally, come up with a relatively accurate method used in feature extraction and classification decisions. The present study is the groundwork analysis for other analysis in EEG study. Methods In accordance with the strength of α wave, the head electrodes are divided into four species. In use of part of 21 electrodes EEG data of 63 people, we have done Fisher discriminant analysis to EEG data of six objects. EEG data processing and statistic analysis adopted independently designed EEG analysis toolbox and the program of correlation analysis. Results In use of part of EEG data of 63 people, we have done Fisher discriminant analysis, the electrode classification accuracy rates is 82.3%. Conclusions Fisher discriminant has higher prediction accuracy, EEG features (mainly α wave) extract more accurate. Fisher discriminant would be better applied to the feature extraction and classification decisions of EEG data.

Keywords: *EEG, Electroencephalogram,· Fisher discriminant, α rhythm*

1. Introduction

The ordinary EEG (electroencephalo-graph) examination, also known as brainwave discriminant analysis, aims to find out whether brain waves are normal, and to assist in the diagnosis of brain lesions. Traditional electroencephalogram discriminant analysis is conducted through interpreting multi-channel EEG in the recording paper by EEG experts, that is, visual inspection is adopted to understand and evaluate the EEG. The essence of experts' experience-based method is to uses experts' experience to remove the signal interference and artifact, extract features from the EEG signals according to signals' frequency, amplitude, and phase information, and make category descriptions for extracted features by using accepted experience, thus analysis and evaluation of EEG can be completed [1]. So far, this method is still widely applied in clinics. Visual methods can capture pathological waveform to a certain extent, and even identify brain lesion locations. However, EEG is characterized by being non-stationary and non-linear, and the visual method depends largely on the EEG analysts' knowledge and experience [2]. These two aspects require new methods to make breakthroughs in EEG research. Fisher discriminant analysis is adopted in EEG research, which will greatly facilitate the extraction and classification of EEG signal information, thus it will help in EEG examination and quantitative analysis, providing an effective analytical tool for EEG checks [3].

2. Research Method

2.1. Subjects

The subjects are 28 men and 35 women, aged from 20 to 60 years old, with an average age of 36.7 years. All subjects are healthy, with no record of serious nervous system disease and psychiatric drug use, which means these people are individuals chosen from the normal population.

2.2. Establishing Mathematical Models for EEG Data Selection

The sampling frequency for lab records of EEG is 100Hz. According to the international 10-20 system, data of 21 electrodes are recorded: C3, CZ, C4, FP1, FPZ, FP2, F7, F8, FZ, F3,

Received January 10, 2013; Revised March 5, 2013; Accepted March 15, 2013

F4, O1, OZ, O2, P3, PZ, P4, T5, T6, T3, T4. One block (representing a small period of time) of EEG data is retrieved at a time, with 512 sampling points for each block and a record time of 5.12 seconds. The EEG of normal people will mainly show α rhythm, with α wave appearing in the back of the head and weakening gradually from back to front. According to different intensities of α wave in various parts of head, the 21 conductive electrodes can be divided into four categories, namely the front head electrodes, the side head electrodes, the central area electrodes, occipital electrodes. The specific classification is as follows:
(1) The first category: the central area electrodes (C3, CZ, C4).
(2) The second category: the front head electrode (FZ, F3, F4, FP1, FPZ, FP2, F7, F8).
(3) The third category: the occipital electrode (P3, PZ, P4, O1, OZ, O2, T5, T6).
(4) The fourth category: the side head electrode (T3, T4).

2.3. Computer Processing of the EEG Data

In order to facilitate the analysis of the raw data of the EEG, MATLAB programming is used to design a dedicated EEG Toolbox. In EEG Toolbox when raw data is accessed and saved in matrix, the horizontal lines represent the time points of lab records (namely sampling points) and the vertical columns stand for electrodes. Before analysis, all data of every subject should be accessed in EEG Toolbox and there should be visual EEGs, with one page including a block of EEG data.

In accordance with the electrode classification method described above, 4 overall Fisher discriminant analyses divide sample data into four categories. First, according to the classification we will put the 21 electrode EEG data which has mathematical models into 4 matrixes $\overline{X}_i (i=1,2...4)$. Then we put EEG data of the current block into the matrix X, whose measurement is 512 × 21. Now we will use Fisher discrimination analysis to predict the electrode classification results, and the results should be displayed in the form of the vector.

Fisher discrimination analysis will build a linear classifier surface in the feature space to project the space sample points along this direction, and identify categories according to values of these projections in this direction. Fisher discrimination analysis procedures in this research program are based on the multi-channel EEG data. First we must establish mathematical models, namely the discrimination functions, then to predict EEG data categories according to the discrimination rules. Fisher discrimination analysis can be explained by the following mathematical formula:

$$y = u'X \quad (1)$$

In the formula, u is the feature vector corresponding to the maximum characteristic root of each $E^{-1}A$ .Multi-dimensional discrimination need to use many feature vectors to form multiple discriminant functions. We put unclassified EEG data X into four Fisher discriminant functions, and get the minimum distance and match it with the corresponding whole. We can use Fisher discriminant functions to predict classifications for EEG data of each block, and show the predicted classification results visually combined with the actual classification in Fisher discriminant analysis prediction map.

2.4. The algorithm of Fisher Discriminant Analysis

The algorithm[4]:
Input: EEG data of the current block.
Output: Predicted classification results and accuracy of Fisher discriminant analyses.
Step 1 : Calculate mean vectors.

$$\overline{\mu}_j = \frac{1}{n_j}\sum_{i=1}^{n_j} X_i^{(j)} = \overline{X}^{(j)}, (j=1,2\cdots l) \quad (2)$$

Step 2 : Calculate overall mean vector.

$$\overline{X} = \frac{1}{n}\sum_{k=1}^{l} n_k \overline{X}^{(j)}, (j=1,2,\cdots,l) \quad (3)$$

Step 3 : Use estimate value of sample to calculate the deviation of groups, as mean vectoers or overall mean vector are unknown.

$$E_l = n_1(\bar{X}^{(1)} - \bar{X})(\bar{X}^{(1)} - \bar{X})' + \cdots + n_l(\bar{X}^{(l)} - \bar{X})(\bar{X}^{(l)} - \bar{X})' \tag{4}$$

Step 4 : Calculate deviation of groups.

$$A = \sum_{i=1}^{n_1}(X_i^{(1)} - \bar{X}^{(1)})(X_i^{(1)} - \bar{X}^{(1)})' + \cdots + \sum_{i=1}^{n_l}(X_i^{(l)} - \bar{X}^{(l)})(X_i^{(l)} - \bar{X}^{(l)})' \tag{5}$$

Step 5 : Calculate u is the feature vector corresponding to the maximum characteristic root of each $E^{-1}A$.

Step 6 : Build Fisher discriminantfunctions $y_l = u_l'X$.

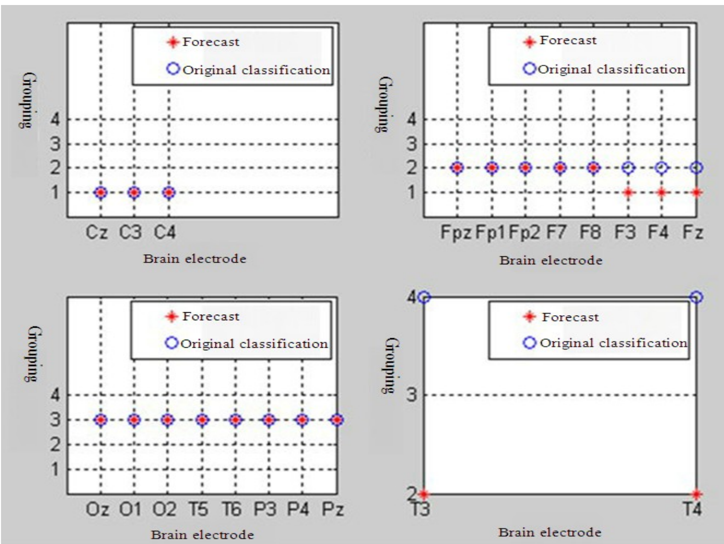
Step 7 : Put unclassified EEG data X into four Fisher discriminantfunctions, and get the minimum distance and match it with the corresponding whole.

$$\left| y(X) - \overline{y}^{(i)} \right| = \min_{1 \le j \le k} \left| y(X) - \overline{y}^{(j)} \right| \tag{6}$$

Step 8 : Predict classification results and calculate the classification accuracy.

2.5. The Analysis Results of Fisher Discriminant Analysis

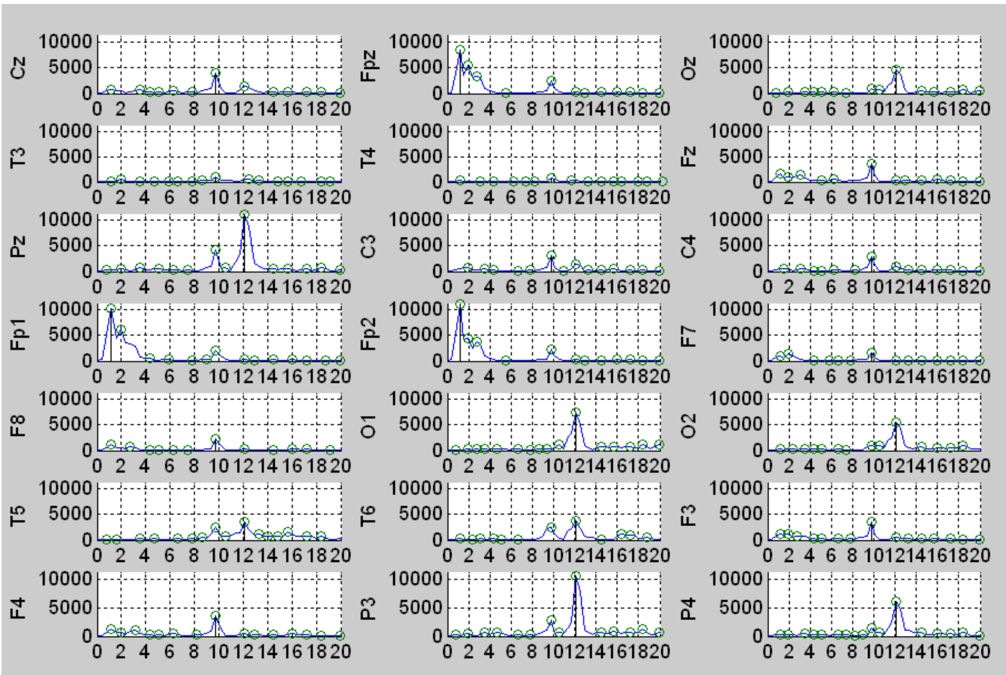
Fisher discrimination analysis procedure can predict classifications of all blocks of EEG data belonging to different subjects. Because space is limited, we will give classification results of a block of 21 electrodes for one subject in detail, and forecast classification results in other blocks are similar here. Fisher discriminant analysis will be conducted in the twenty-second block of EEG data for Subject 1.



Note 1): forecast results in the upper left corner are for first-category electrodes, results in the upper left corner are for second-category electrodes, results in the lower left corner are for third-category electrodes, and results in the lower right corner are for fourth-category electrodes. 2): x-coordinate is for electrode names of all types, and y-coordinate is for categories. 3): red * indicates forecast classification, and blue O represents the actual classification. When forecast classification and actual classification are consistent, * will coincide with O.

Figure 1. Fisher forecast results for the twenty-second block Subject 1

Figure 1 shows Fisher predicted results for the twenty-second block of EEG data of Subject 1. In the current block, we have accurate analysis for electrodes in the first category in the central area and the third category in occipital electrodes, but second-category electrodes FZ, F3, F4 in the front head is classified into the first category, and fourth-category electrodes T3 and T4 in the side head is classified into the second category. The EEG of normal people will show α rhythm, with wave α weakening in a row from back of the head, the central area, head side, front head. In predicted results, FZ, F3, and F4 are classified into the central area, indicating wave α of FZ, F3 and F4 have stronger electrodes than normal front head; T3 and T4 are classified into the front head, indicating wave α of T3 and T4 have weaker electrode wave than normal temporal electrode wave. The estimated spectral density of EEG data for the current block (shown in Figure 2) is retrieved by using psd function. By doing this, we found that wave α of Subject 1 for the current block conforms with that of electrodes in rear head, and tends to weaken gradually along the direction of front head. Wave α of front head electrodes are stronger, and are not weakened compared with electrodes in the central area. As a result, these electrodes are wrongly classified into the first category. Wave α of temporal electrodes with the same intensity of wave α in front head electrodes, and this is why T3 and T4 are wrongly put into the second category. The intensities of wave α in electrodes showed spectral density estimates is basically in agreement with those in predicted results. Fisher discriminant analysis will build a linear classification surface in the feature space to project the space sample points along this direction, and combining data to be analyzed in the whole with lowest projection value along this direction. We will take F3 as an example to illustrate the process of the discrimination analysis. We put F3 into four Fisher discrimination functions. Thus the distance from projected F3 to four wholes is 2.5061e +017, 4.4660 e +018, 5.7533 the e +018, 3 .8302 e + 018 respectively. In these values, the distance from projected F3 to the first whole is the smallest, thus it's put into the first category. Through projection Fisher discrimination analysis can more accurately distinguish superiority levels of wave α with better predicted results. The right predicted number for the current block is 16, and the wrong number is 5, with 76.1% as the accuracy rate.



Note 1): psd function is used to obtain the power spectral density estimates in the 22nd block for Subject 1. The x-coordinate is for the frequency, and the ordinate-coordinate is for the power estimate. 2): the frequency of wave α is 8 ~ 13Hz; the larger the ratio of the square of wave α to the overall area is, the stronger α wave is. We can see intensities of α wave in electrodes.

Figure 2 Power spectral density of the 22nd block for Subject 1

Fisher discrimination analysis is conducted for 63 subjects. Predicted classifications of 10 blocks of EEG data are randomly selected from each subject, with an average accuracy rate of 82.3% (shown in Table 1). Compared with the Mahalanobis distance discrimination analysis, Fisher discrimination analysis has higher prediction accuracy rate, and more accurate EEG feature (mainly wave α) extraction. Predicted results can reflect intensities of wave α in each electrode, but these can not remove amplitude modulation, approximation and participants' individual differences as well as the emergence of other band waves' impact on classification results.

Table 1. Average accuracy rate of using Fisher discriminant analysis to predict EEG classifications for 63 subjects

Accuracy rate	70%-75%	75%-80%	80%-85%	85%-90%	Average accuracy rate
Subject(s)	4	25	33	1	82.3%

3. Results and Analysis

According to different intensities of α wave in various parts, the 21 conductive electrodes can be divided into four categories. Mathematical models for Fisher discriminant analysis can be established by using data of 21 electrodes. Fisher discriminant analysis is conducted for 63 normal subjects. Forecast classifications of 10 blocks of EEG data are randomly selected from each subject, with an average accuracy rate of 82.3%. It has higher prediction accuracy rate, and more accurate EEG feature (mainly wave α) extraction, and makes better distinctions in intensities of wave α in electrodes. The tests show that the Fisher discriminant analysis method can extract EEG features of normal people in a more favorable manner, and can be applied in EEG data classifications[5], [6].

Fisher discriminant analysis can not remove amplitude modulation, approximation and participants' individual differences as well as the emergence of other band waves' impact on classification results, which leads to wrong classifications. By using theories of multivariate statistics, we find out reasons for electrode misjudgment as follows:

(1) Fisher discriminant analysis will build a linear classification surface in the feature space to project the space sample points along this direction, and identify categories according to values of these projections in this direction. Since it is impossible to get an infinite number of sample data, the overall projection areas can only estimated by using limited EEG data within the centralized training set. This may affect the prediction accuracy.

(2) Misjudgment exists in Fisher discrimination analysis when it's used to analyze EEG data for normal people. In fact, Fisher discriminant reduces dimensions, namely it projects the data of the 512-dimensional EEG vector onto a straight line, and then classifies patterns according to the obtained one-dimensional characteristic (also called scalar). It views the Fisher discriminant analysis as a feature extraction algorithm, we research the way of obtaining the best projection direction from the perspective of feature extraction so that one-dimensional projection characteristic can distinguish four types of electrodes in most favorably manner[7-8]. If four types of electrodes are overlapping when projected, and data to be analyzed is just in the overlapping area, it is probably to misclassify the sample. We take the first-category and second category electrodes (Figure 3) as examples to illustrate that one-dimensional characteristic obtained from various projected directions, may affect greatly in classified performances. For instance, first-category and second category electrodes have some overlapping area in space. If it is projected to the x-axis direction, Fisher discriminant analysis can easily distinguish electrodes projected in the region of R4 and R6. However, it cannot accurately distinguish electrodes projected in the region of R5 where many electrodes overlap, and thus misclassification occurred. We hope to find a line L, so that projected electrodes can be separated as far as possible. As it can be seen in Figure 4.9, the effect is better to project along the L direction than the x direction. Although there will still be some overlapped electrodes, the electrodes in the overlapped area R2 have been significantly reduced. In fact, four catogeries of electrodes partially overlap in the 512-dimensional space, and cannot be accurately distinguished after they are projected, and then it may lead to misclassification.

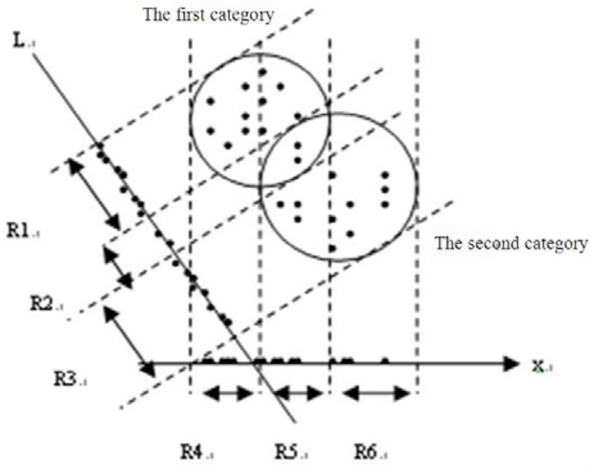


Figure 3. Fisher projection maps of the collectivity of both

(3) There are individual differences in EEG[9]. Different subjects have different EEG amplitudes, frequencies, and waveforms; each block of EEG signals differs as well. Fisher discriminant analysis method can not eliminate individual differences of the normal EEG data when it is used to analyze each block of EEG data, all of which result in misclassification in electrodes categories [10].

4. Conclusion

Fisher discriminant analysis is conducted for 63 normal subjects. Forecast classifications of 10 blocks of EEG data are randomly selected from each subject, with an average accuracy rate of 82.3%. It has higher prediction accuracy rate, and more accurate EEG feature (mainly wave α) extraction, and makes better distinctions in intensities of wave α in electrodes. The tests show that the Fisher discriminant analysis method can extract EEG features of normal people in a more favorable manner, and can be applied in EEG data classifications.

References

[1] Gabor AJ, Leach RR, Dowla FU. Automated seizure detection using a self-organizing neural network. *Electroencephalography and Clinical Neurophysiology*. 1995; 99(3): 257-266.
[2] S Blanco, et al. Applying time-frequency analysis to seizure EEG activity. *IEEE Engineering in medicine and biology magazine*. 1997; 16: 65-71.
[3] Williams WJ. Time-Frequency Analysis of Biological Signals. *IEEE Electrical Computer Science*. 1993; 12(1): 83-86.
[4] Sebastian M, Gunnar R, Jason M, et al. Fisher discriminant analysis with kernels. *Proceedings of IEEE International Work shop on Neural Networks for Signal Processing*. Madison, Wisconsin. 1999; 41-48.
[5] John Trinder, John A van Beveren, Philip Smith, et al. Correlation between ventilation and EEG arousal during sleep onset in young subjects. *Journal of Applied Physiology*. 2001; 83: 2005-2011.
[6] P Comon. Independent component analysis—a new concept. *Signal Processing*. 2000; 36: 287-314.
[7] W Zhao, R Chellappa, and A Krishnaswamy. Discirminant analysis of principle components for face recognition. *Proc. of Inter. Conf. on Automatic Face and Gesture Recognition*. Nara, Japan. 1998; 336-341.
[8] Hendra Kusuma, NFN Wirawan, Adi Soeprijanto. Gaborbased Face Recognition With Illumination Variation Using Subspace Linear Discriminant. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(1): 119-128.
[9] Zhang Shao-bai, Huang Dan-dan. Electroencephalography Feature Extraction Using High Time Frequency Resolution Analysis. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(6): 1415-1421.
[10] Kasper K, Schuster HG. Easily calculable measure for complexity of spatial temporal pattern. *Physical Review Online Archive*. 2010; 36 (2): 842-848.