Computer Modelling and New Technologies, 2013, Vol.17, No. 4, 17-20
Transport and Telecommunication Institute, Lomonosov 1, LV-1019, Riga, Latvia

STUDY ON MAHALANOBIS DISCRIMINANT ANALYSIS OF EEG DATA

Yuan Shi, Linlin Yu, Li Zou

[1]Dalian Institute of Science and Technology
Dalian Jiao Tong University
BinGang Str.999-26, DaLian, China
Phone: (+0086)-13664247491. E-mail: 20088041@qq.com

Objective In this paper, we have done Mahalanobis Discriminant analysis to EEG data of experiment objects that are recorded impersonally come up with a relatively accurate method used in feature extraction and classification decisions. 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 Mahalanobis Discriminant analysis to EEG data of six objects. Results In use of part of EEG data of 63 people, we have done Mahalanobis Discriminant analysis, the electrode classification accuracy rates is 64.4%. Conclusions Mahalanobis Discriminant has higher prediction accuracy, EEG features (mainly α wave) extract more accurate. Mahalanobis Discriminant would be better applied to the feature extraction and classification decisions of EEG data.

Keywords: Electroencephalogram, Mahalanobis Discriminant, α hythm

1. Introduction

The purpose of routine brain wave inspection is to evaluate whether the brain wave is normal or not and provide help to diagnose the brain disorders which is also known as brain wave interpretation. The traditional brain wave interpretation is realized through reading the multi-channel electroencephalogram on the recording paper by experts, which is to understand and evaluate electroencephalogram (EEG) with the method of visual inspection. The essence of this method based on expertise is that experts utilize experience to wipe out the disturbance and artefact of signals, conduct feature extraction to the EEG according to the frequency, range, phase position and other information, and carry out the category description for the extracted features with the recognized experience to analyse and evaluate the EEG [1]. Up to now, this method is widely applied to the clinic. The visual inspection, to some extent, can catch the pathological waveform or even confirm the position of the brain focus. However, due to the strong non-stationary and nonlinear characteristics of EEG, with the addition of the great dependence of visual inspection on knowledge-level and experience of EEG analysis personnel, the new method must be explored to realize the breakthrough of EEG research [2].

Mahalanobis Discriminant analysis has been introduced into the research of EEG, which will actively promote the extraction and classification of EEG data to assist the inspection and quantitative analysis of EEG and provide the effective analysis means for the EEG examination.

2. Objects and Methods

2.1. Object of Study

We take 28 men and 35 women as the research objects, whose age is ranging from 20 to 60, and the average age is 36.7. All the subjects are enjoying good health without serious nerve system diseases and history of taking psychotropic drugs, and they are selected from the normal population.

2.2. Build the Selection of Mathematical Modelling EEG Data

The sampling frequency of experiment recording of EEG is 100Hz which is recording 21 electrode data according to the lead location in international 10-20 system: C3, CZ, C4, FP1, FPZ, FP2, F7, F8, FZ, F3, F4, O1, OZ, O2, P3, PZ, P4, T5, T6, T3, T4. A block (indicating a short time period) of EEG data is acquired at every turn, and the number of sampling points for each block is 512 with the recording time of 5.12s. The electroencephalogram of normal people is mainly in α rhythm, the strengths of α wave appear in the occiput, and then weakening gradually from back to front. Classify the 21 conducting electrodes into 4 categories in accordance with the intensity differences of the α rhythm in various parts of the head, which is, former head electrode, side head electrode, central electrode, occiput electrode. The specific classification situation is as follows:

(1) The first category: central electrode (C3, CZ, C4)
(2) The second category: former head electrode (FP1, FPZ, FP2, F7, F8, FZ, F3, F4)
(3) The third category: occiput electrode (O1, OZ, O2, P3, PZ, P4, T5, T6)
(4) The forth category: side head electrode (T3, T4).

2.3. The Computer Processing of EEG Data

The electroencephalogram dedicated toolbox EEG Toolbox is designed with the MATLAB programming language in order to facilitate analysing the original data of the electroencephalogram. In EEG Toolbox, after the original data was introduced, it was saved in the matrix, and line represents the timing of experiment recording (that is sampling point) while column indicates the electrode. All the data of every subject were introduced before analysis and the electroencephalogram should be shown intuitively, and a block of EEG data should be displayed on each page [3].

The 4-population Mahalanobis distance discriminance classifies the sample data into four categories based on the electrode classification method introduced above. Firstly, put the 21 electrode EEG data, which will build the mathematical models into the four matrixes $\overline{X}_i(i=1,2...4)$ on the basis of classification. Put the current block of EEG data into the Matrix X, which is 512×21 matrix. The electrode classification results are predicted with Mahalanobis distance discriminance and expressed by putting them into vector.

When the Mahalanobis distance discriminance is adopted, the discriminating data will be classified into the nearest category based on the distance length of population centre from the distinguishing EEG data. The Mahalanobis distance analysis procedure in the research is on the basis of multi-channel EEG data design. Firstly, the mathematical model, namely, discrimination function, should be built, then predict the category of EEG data according to the discrimination rules. The Mahalanobis distance discriminance could be explained by the following mathematical formula:

$$d_j^2(X) = A * B, \tag{1}$$

$$A = (X - \frac{1}{n_j}\sum_{i=1}^{n_j} X_i^{(j)})', \tag{2}$$

$$B = [\frac{1}{n-n_j}\sum_{j=1}^{n_j}(X_i^{(j)} - \overline{X}^{(j)})(X_i^{(j)} - \overline{X}^{(j)})']' A \cdot \tag{3}$$

Formula (1-3) is the Mahalanobis distance discriminance function, and 4-population Mahalanobis distance discriminance will build four discriminance functions. Substitute an unknown classified EEG data X into the four Mahalanobis distance discrimination functions, and acquire the minimum Mahalanobis distance and discriminate it to the corresponding totality.

Each block of EEG data can be predicted and classified by the Mahalanobis distance discrimination function and the predicated classification results as well as actual classifications can be intuitively displayed in the Mahalanobis distance discriminance predicating result chart.

3. EEG Data Analysis Results from Mahalanobis Distance Discriminance

Classify the 21 conducting electrodes into four categories according to the intensity differences of α wave on various parts of the head, and forecast 21 electrodes classification conditions of six subjects in the current block with Mahalanobis distance discriminance: C3, CZ, C4, FP1, FPZ, FP2, F7, F8, FZ, F3, F4, O1, OZ, O2, P3, PZ, P4, T5, T6, T3, T4. Draw the Mahalanobis distance discriminance predicating result chart and 2D pie chart of Mahalanobis distance discrimination accuracy rate with Mahalanobis distance discriminance procedure [4].

Mahalanobis distance discriminant analysis procedure can forecast and classify the EEG data of all blocks for various subjects. Due to the space constraints, only the forecast results of the EEG data in six blocks for three subjects are given in detail, the whole situation can be reflected by showing only classification results of two blocks for each subject, and the forecast classification results in other blocks are similar to these. Analyse the predicated results of one subject in detail. We number the six subjects for the sake of convenient description: 1, 2, 3, 4, 5, 6. First conduct the Mahalanobis distance discriminant analysis for the 12th block of EEG data of Subject 1. The figure of EEG is shown in Fig1.

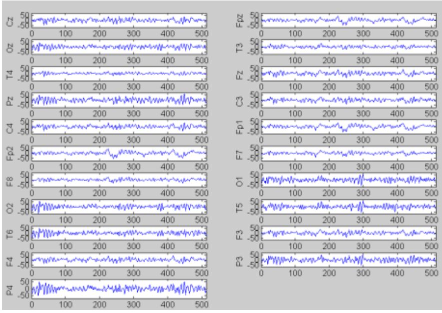

Figure 1. EEG data of the twelfth block of in subject 1

Note 1): The horizontal ordinate is frequency (Unit: Hz), while the vertical coordinate is voltage (Unit: µV), and the electrode parts have been marked on the left side of the data.

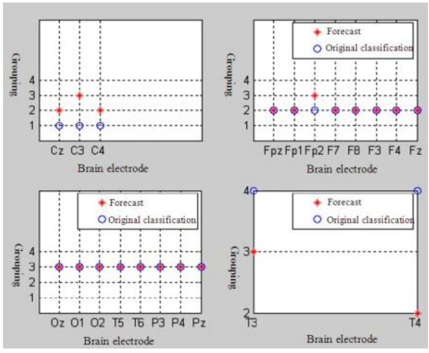

Figure 2. Mahalanobis distance predicating results of the twelfth block in Subject 1

Note 1) The top left corner is the predicating results of the first–category electrode; the top right corner is the predicating results of the second–category electrode; the left bottom is the predicating results of the third–category electrode; and the bottom right corner is the predicating results of the forth–category electrode. 2) The horizontal ordinate is the names of various kinds of electrodes and the category is listed on the vertical coordinate. 3) red* represents the predicated classification situation, blue O indicates the actual category situation. * And O will coincide when the predicated classification is consistent with the actual classification. 4) first–category electrode (C3, CZ, C4), second–category electrode (FZ, F3, F4, FP1, FPZ, FP2, F7, F8), third–category (P3, PZ, P4, O1, OZ, O2, T5, T6), forth–category electrode (T3, T4)

Carry out the Mahalanobis distance discriminant analysis for the six subjects, and randomly extract 10 blocks of EEG data for every subject to predict classification. The Mahalanobis distance predicating results is shown by Fig2. The average accuracy rate is 64.4% (shown in Table 1). Overall, the forecast results of Mahalanobis distance discriminance are better; the extraction of the EEG characteristics (mainly α wave) is relatively accurate. The predicating results can reflect the intensity differences of wave α in various parts of the head. To some extent, the occurrence rate or amount of wave α, namely, the quantity of wave α recorded in the EEG within a certain period of time, has huge differences in individuals. The forecast results are influenced by such situations where wave α constantly appears in some people and sporadically in other people, or when other frequency waves appear. The mis-discriminance can also be caused by the amplitude modulation and right-and-left difference.

Table 1. The average accuracy rate of Mahalanobis distance discriminance predicating EEG classification

Subject	1	2	3	4	5	6	Average Accuracy Rate
Accuracy Rate	66.7%	63.2%	64.2%	63.5%	65.4%	63.1%	64.4%

4. Conclusions

Because Euclidean distance is oversimplified, and the absolute distance and Chebyshevy distance cannot completely express the characteristic differences of the multidimensional data in the high-dimensional space, therefore, we usually analyse the EEG data with Mahalanobis distance discriminance in experiments. Classify 21 conducting electrode into four categories according to the intensity differences of the wave α in every part, build Mahalanobis distance discriminance mathematical model with 21 brain electrode data, and conduct the Mahalanobis distance discriminant analysis to the EEG data of 6 normal subjects. The predicating classification accuracy rate is 64.4%. On the whole, the predicating results of Mahalanobis distance discriminance is better, the extraction of EEG characteristics (mainly wave α) is more accurate, and the predicating results can reflect the intensity differences of wave α on the various parts of the head. The experiment indicates that Mahalanobis distance discrimination can preferably extract the EEG characteristics of normal people and can be applied to the classification decision of EEG data.

The EEG of normal people presents α rhythm and wave α is the major EEG characteristic of normal people. The predicating classification results of different blocks are not completely equivalent, which reflects that the EEG is a non-stationary random signal and wave α is constantly changing. The amplitude modulation phenomenon, left-and-right difference and individual differences in subjects will exert an influence on predicating classification of EEG data and cause the mis-discriminance. We analyse the reasons for mis-discriminance of the electrodes as follows:

(1) When the Mahalanobis distance classification discriminance is adopted, the discriminating data will be classified into the nearest category based on the distance length of each population centre from the distinguishing EEG data [5]. It is impossible to get infinite EEG data samples, so the limited samples with centralized training are used to estimate every population centre. The waveforms, amplitudes and phase positions of the normal people collected in the experiment are different in the EEG data, which may lead to the inaccuracy of predicating classification.

(2) When Mahalanobis distance discrimination is adopted to analyse the EEG data of normal people, where mis-discrimination exists. The principle of Mahalanobis distance discriminance is to classify the discriminating samples to the nearest category [6]. If the two kinds of the samples are overlapped and the discriminating samples are just in the overlapped area, the samples will probably be classified wrongly. We take the two populations for example in order to give a better explanation. Suppose that the sample x is one-dimension variable, G1 and G2 are two populations, and Figure 3 is the distribution situation of the two populations. For instance, if x belongs to Population G1, but it falls on the left side of $\bar{\mu}$, in accordance to the rules, x is classified to G2; similarly, the point in G2 can be classified to G1 wrongly with Mahalanobis distance discriminance. The error probability is shown in the shadow area of the diagram. If the two populations are near to each other, the rate of mis-discrimination must score high [7].

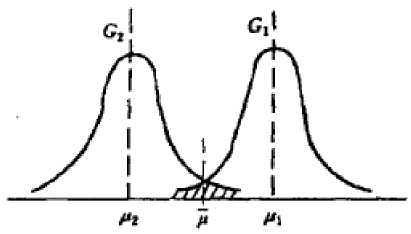

Figure 3. Distribution diagram of two populations

Sponsored by Hi-Tech Res earch and Development Program of China (863) (2012AA040912).

References

1. Trinder, J., A. van Beveren, J., Smith, P., et.al. (2005) Correlation between ventilation and EEG arousal during sleep onset in young subjects [J], *Journal of Applied Physiology*, 83, 2005-2011.
2. Kasper, K., Schuster, H. G. (2001) Easily calculable measure for complexity of spatial temporal pattern. Physical Review Online Archive, 36 (2), 842-848.
3. Viglione, S. S., Ordon V. A., Risch, F. (1999) A methodology for detecting ongoing changes in the EEG prior to clinical seizures. In: Western Institute on Epilepsy, West Huntington Beach, 27-28.
4. Williams, W. J. (1999) Time-Frequency Analysis of Biological Signals, *IEEE Electrical Computer Science*, 12(1), 83-86.
5. Blanco, S., et.al. (2007) Applying time-frequency analysis to seizure EEG activity, *IEEE Engineering in medicine and biology magazine*, 16, 65-71.
6. Williams, W. J., Hitten, P., Zaveri, J. (1995) Time-frequency Analysis of Electrophysiology Signals in Epilesy, *IEEE Transaction on Biolmedical Engineering*, 3(2), 133-142.
7. Richman, J. S., Moorman, J. R. (2000) Physiological time-series analysis using approximate entropy and sample entropy, *AJP Heart and Circulatory Physiology*, 278(6), 2039-2049.

Received on the 17th of June 2013