A New Improved Feature Extraction Method in Memory EEG Data

Luo Jing[1, a] ,Yun Li[2,b] and Hong Zhang[3,c]

[1]Sichuan police college , luzhou,Sichuan, P. R. China.

[2]Southwest University No. 2, Tiansheng Road Beibei, Chongqing, P. R. China.

[3]University of Oregon1250 Siskiyou Boulevard Ashland,Oregon,USA

[a]Email:209528872@qq.com, [b]Email:ly110710@swu.edu.cn,[c]Email:zhangh@swu.edu.cn

Keywords: extraction; EEG; LDA; classification tree

Abstract. various papers and conferences about EEG data can be found at present. There are various feature extraction methods reviewed oversimplified in section one, such as zerocrossing, low zero-crossing rates, coherence analysis, subspace methods, the mean absolute amplitude, standard variance, kurtosis and so on. The feature extraction methods such as self-produced mother wavelet feature extraction method, best basis-based wavelet packet entropy feature extraction, empirical mode decomposition and non-linear feature extraction using correlation dimension and Hurst exponent are detailed introduced in section two. Those feature extraction methods are complex and limited, which often used in some specific fields. In this paper, a new feature extraction is proposed named incremental value, which considers the changes in brain waves. Next LDA and classification tree are used to analyze the results of feature extraction and to predict with unequal memory error compared with the feature extraction methods, such as mean absolute amplitude, standard variance and kurtosis. The method that we proposed is concise and accurate than other methods.

I. Introduction

The electroencephalogram signals known as EEG is proposed for assessing the functional state of the brain. In memory processes, the brain activity changed with various persons and memory degrees. The study of the brain electrical activity, through the electroencephalographic records, is one of the most important tools for determining whether a person has been remembered.

As we know, since EEG data is high-dimensional and complex, the electroencephalogram signals pre-processing, feature extraction and feature classification are difficult. At present, there are so many fields of feature extractions and classification algorithms used in study the feature extraction and classification of EEG data.

Features can be defined as parameters that provide information about the underlying structure of a signal [1]. There have been numerous techniques applied to EEG signals for the purpose of feature extraction [2]. There are major methods to extract EEG feature, such as temporal features, spectral features and nonlinear features (complexity measures). We will introduce a few of these methods. The simplest are temporal feature including the mean absolute amplitude, standard variance and kurtosis, as well as measures relating to the probability density function of the waveform, such as mode, median or the entropy [1]. The next is zero-crossing, low zero-crossing rates are associated with low frequency processes [3]. The third method is hjorth parameters used in expressing EEG waveforms [4]. The fourth method is coherence analysis which reflects the degree of synchrony between frequency components of two signals [5]. The fifth is subspace methods which is a form of parametric spectral estimation [6]. The sixth is Correlation dimension that provides a bound on the fractal dimension of the attractor of the underlying dynamical system [7]. Furthermore, there are so many method to extract feature, including detrended fluctuation analysis (DFA)[8][9][10], non-parametric spectral estimation methods, higher-order spectral analysis

(HOSA)[11], short time fourier transform (STFT)[12], the wavelet transform, matching pursuits (MP)[13], entropy measures and Lyapunov exponents and so forth[14][15]. After introducing the feature extraction, we will introduce some commonly used classification methods. Once feature extraction finished, we will use classification to evaluate the effect of feature extraction [21]. We usually use the following method to classify. The first is neural network (NN) classification or artificial neural networks (ANN) is proposed as the result of attempts to understand how the human brain learns and solves problems [1]. Using several neurons in conjunction, ANNs becoming capable of modeling very complex nonlinear systems are suitable for EEG data [16]. The second is self-organising maps (SOM) or Kohonen maps used in analysis and visualizing high dimensional data [17]. The third is linear discriminant analysis (LDA) known as fishers Linear Discriminant analysis, which is simple to implement, to interpret and generally works well in classifying linearly separable data [18]. The fourth is hidden Markov model (HMM) used in analysis non-stationary signals [19]. The fifth is fuzzy classification proposed by Zadeh in 1965 which can be effectively integrated with other classification systems for better performance [20]. In addition to the above methods, there are combined classifiers and classification tree. In this paper, linear discriminant analysis (LDA) and classification tree are used for analysis the results of feature extraction and prediction whether the subjects remember. The following is a flow chart of EEG feature extraction.

This paper is organized as follow: in section two, the state of the art about the advanced feature extraction method were introduced. In section three, a new feature extraction method will be proposed. Then an experiment is described detailed. Next is the conclusion and future work. Finally the results are depicted.

Figure 1.1.   the flow chart of EEG feature extraction architecture

II. State of the art

The former section introduced the various foundation feature extraction methods oversimplified. In this section some major and novel feature extraction methods are exhaustively introduced, including self-produced mother wavelet feature extraction method, best basis-based wavelet packet entropy feature extraction, empirical mode decomposition and nonlinear feature extraction using correlation dimension and hurst exponent. The following is the summary of the four feature extraction methods.

Table 1   the summary of the four feature extraction methods

Feature extraction methods	Application field in paper	Advantages and disadvantages
self-produced mother wavelet	Motor Imagery Brain-Computer Interface	The performance of this work derived more accurate results than variance and FFT methods.
best basis-based wavelet packet entropy	epileptic detection	The proposed method was successfully applied to EEG signals for the epileptic

		detection. But the method is too complex.
Empirical mode decomposition	Epilepsy EEG data	The extracted features of seizure-free and ictal EEGs are effective and satisfying sensitivity and specificity. But the Algorithm complexity is high.
non-linear feature extraction using correlation dimension and Hurst exponent	Epileptic EEG data	using CD and Hurst exponent to extract non-linear features of EEG signals. High precision and high complexity

A. self-produced mother wavelet

First of all, we will introduce the self-produced mother wavelet feature extraction method proposed by W.-L. Yeh, Y.-C. Huang, J.-H. Chiou, J.-R. Duann and J.-C. Chiou in 35th Annual International Conference of the IEEE EMBS, Osaka, Japan, July 2013[24]. The method based on the concept of Continuous Wavelet Transform compute the correlation coefficients between the wavelet and the EEG signal, and the correlation coefficients are employed as the features. The following is the definition:

$$C = \sum_{j=1}^{n-m} a(j) * \frac{\sum_{i=1}^{m} w(i) * E(i+j)}{\sum_{i=1}^{m} E(i+j)^2} \quad (2.1) \qquad C(j) = 1, if \sum_{i=1}^{m} w(i) * E(i+j) > C_{th}, else\ C(j) = 0 \quad (2.2)$$

Where c is the correlation coefficient, C_{th} is the threshold we set, w is the wavelet; E is the EEG signal segment we want to classify. Moreover, n and m are the number of samples of EEG segment and the wavelet respectively. According to above parameters, we can get a feature vector C. This method often used in motor imagery brain-computer interface and it is better than variance and FFT methods.

B. best basis-based wavelet packet entropy

Secondly, the paper named Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection was written by Deng Wang, Duoqian Miao, Chen Xie, which was published in Expert Systems with Applications in 2001. He proposed the feature extraction method best basis-based wavelet packet entropy feature extraction. The procedure of feature extraction is introduced in following detailed presented in [25].

(1) Select a wavelet function W and specify the decomposition level L.

(2) Calculate the sample mean SM.

(3) Decompose SM at the specified level with the selected wavelet function, and return a wavelet packet tree T. Let $B_{i,k}$ be the set of WPT basis vector, $0 \leq l < L, 1 \leq k \leq 2^l - 1$.

(4) Calculate energies $E_{i,k}$ for all subbands using Eq.

(5) Se tthe initial basis $B = \{Bl-1,1, Bl-1,2,..., Bl-1,k,..., Bl-1,2l-1\}$ related to the subbands at the bottom level.

(6) Compare the entropy of a parent node $E_{i,k}$ with the sum of the entropy of two child nodes ($E_{l+1,2k-1} + E_{l+1,2k}$).

If $El,k \leq (El+1,2k-1 + El+1,2k)$, then replace $Bl+1,2k-1$ and $Bl+1,2k$ by Bl,k in B; else set $Ei,k = (El+1,2k-1 + El+1,2k)$, i.e., assign the sum of the childrens entropy to the parent node.

(7) Repeat (6) for the next higher level until the root is reached.

(8) Select a sample from training set.

(9) Decompose the sample to L using W.

(10) Calculate wavelet coefficients in the corresponding best basis B.

(11) Calculate Max, Min, Mean and Standard deviation of the wavelet coefficients to form a

4-dimension feature.

(12) Repeat steps (8) to (11) for all samples.

Where this method is was successfully applied to EEG signals for the epileptic detection. When used in this memory EEG data, we should revised the method.

C. empirical mode decomposition

Thirdly, Feature extraction and recognition of ictal EEG using EMD and SVM published in Computers in Biology and Medicine in 2013, which was written by Shufang Li, Weidong Zhou who are the person in school of information science and engineering, Shandong University, Jinan, China. They presented a novel method for feature extraction and recognition of ictal EEG using EMD and SVM. Empirical mode decomposition (EMD) technique is to decompose a signal automatically into a set of the band-limited functions named Intrinsic Mode Functions, which in the whole data set, the number of extreme points and the number of zero crossings must either equal or differ at most by one, and at every point, the mean value of the envelopes defined by local maxima and local minima is zero [26]. EMD can separate a segment of EEG signal x(t) into n IMFs and r, which is reconstructed as a linear combination:

$$x(t) = \sum_{n=1}^{N} imf_n + r \qquad (2.3)$$

In the paper, they said that the main features of the ictal EEG are closely related to the first five IMFs, and IMF1CIMF5 of each EEG segment are used to extract the EEG features. And they defined the coefficient of variation for one IMF:

$$v_c = \sigma^2 / u^2 \qquad (2.4)$$

where is the variation and is the mean. So the v_c can present the variations of the signal amplitude. The definition of fluctuation index is proposed:

$$F_i = \frac{1}{l} \sum_{j=1}^{l} |IMF(j+1) - IMF(j)| \qquad (2.5)$$

Where IMF represents one of the IMFs with length l. The fluctuation indices of IMF1C5 are computed as the features. As we known the ictal EEG has more fluctuation than the normal EEG, so the method should amend according to the certain circumstances.

D. non-linear feature extraction using correlation dimension and hurst exponent

Last but not the least, EEG non-linear feature extraction using correlation dimension and Hurst exponent written by Shujuan Geng and Weidong Zhou, published in Neurological Research which is a journal of progress in Neurosurgery, Neurology and Neurosciences. They proposed non-linear feature extraction using correlation dimension and Hurst exponent in the paper. The correlation dimension (CD) was delimited as follow [27]:

$$X(i) = x(i), x(i+), .., x[i + (m-1)] \qquad (2.6)$$

Where i=1,2,,N-(m-1), X(i) is the phase space vectors. Presume x to be a data sequence containing N data points, where x = [x(1),x(2),,x(N)]. is the time lag and m is the embedding dimension. Then they computed the correlation integral C(r):

$$C(r) = \frac{2}{M(M-1)} \sum_{i \neq 1}^{M} \theta[r - |X(i) - X(j)|] \qquad (2.7)$$

Where the r is the radial distance around each reference point X(i), M is the number of data points in phase space, X(i) and X(j) are phase space vectors, (x) the Heaviside step function. According to the above statements, the CD is proposed:

$$CD = \lim_{x \to 0} [\frac{\ln(C(r))}{\ln(r)}] \qquad (2.8)$$

This paper proposed that The Hurst exponent (H) was used to evaluate the presence or absence of long-range dependence and its extent in a time series Where C(r) is the correlation integral. The

Hurst exponent was shown as following:

$$w_j = \sum_{i=1}^{j} [X_i - \bar{X}_i] \qquad (2.9)$$

$$R(n) = max(O, w_1, w_{21}, \ldots w_n) - min(O, w_1, w_{21}, \ldots w_n) \qquad (2.10)$$

$$S(n) = \{\frac{1}{n}\sum_{i=1}^{n}[x_i - \bar{X}_i]^2\}^{1/2} \qquad (2.11)$$

$$\frac{R}{S} = \frac{R(n)}{[S(n)^2]^{1/2}} \qquad (2.12)$$

$$\frac{R(n)}{S(n)} = an^H \qquad (2.13)$$

where $\bar{x_n}$ the mean of x_i, $R(n)$ is the range, $S(n)$ is the sample standard deviation of n points, $S(n)2$ is the variance of n points, H is called the Hurst exponent.

In this paper the authors calculated CD and Hurst exponent to extract non-linear features of EEG signals. Then they compared epileptic EEG and interictal EEG. They said that the CD could quantify the complexity levels of the epileptic EEG and interictal EEG signals. And the Hurst exponent is used to evaluate the presence or absence of long range dependence and its degree in a time series. So the nonlinear parameters such as CD and Hurst exponent can help interpret epileptic EEG and interictal EEG. The former are just a brief overview about feature extraction in EEG data. There are so many EEG feature extraction methods have not been covered in this paper.

III. A new feature extraction method

Wherever Times is specified, Times Roman or Times New Roman may be used. If neither is available on your system, please use the font closest in appearance to Times. Avoid using bit-mapped fonts if possible. True-Type 1 or Open Type fonts are preferred. Please embed symbol fonts, as well, for math, etc. Before propose the new feature extraction method, we will define this problem. In this paper, we take the most commonly used standard electrode placement method, which is 10 to 20 system electrode. The following are issues related to the definition:

Definition 1: electrode placement of the electrode cap is stored in a vector, X =

[PO3,PO4,F3,F4,C3,C4,P3,P4,O1,O2,OL,OR,T3, T4,T5,T6,POz,Cz,Fz,Pz]. Xi is the i-th electrode, where Xi = X[i]. For example, X1 = X[1] = PO3.

Definition 2: The EEG data that we collected will be stored in a three-dimensional array Am∗n∗k. where m represents the number of times each participated subjects in the experiment, n stands for the position of the electrode cap which is the value i in definition 1, k represents the time axis.

Definition 3: The i-th electrode collected signal in the t-minute is expressed as Xit.where i = 1,2,...19,20. Where t = 1,2,...,k − 1,k.

Definition 4: The subject will get a test data in each experiment. The difference between test data and the true value is called memory error presented by ξ, where ξ is a real number.

The method that we proposed is named incremental value method. In mathematical, incremental value presents the variety of a set data, and stands for the manner and extent of changes in a data set. Therefore, the incremental value can better reflect the manner and extent of EEG Changes. The detailed defined as follows.

$$X_{it} = 1, where \ X_{it+1} \geq X_{it}, t = 1, 2, ..., k - 1, k \quad (3.1)$$

$$X_{it} = 0, where \ X_{it+1} < X_i t, t = 1, 2, ..., k - 1, k \quad (3.2)$$

$$x_i = \sum_{i=1}^{k} X_{it}, where \ t = 1, 2..., k \quad (3.3)$$

Where xi is presented the feature extraction of the i-th electrode cap, X(it) the i-th electrode

collected signal in the t-minute, t represents the time axis, i stands for the position of the electrode cap which is the value i in definition 1.

After extracting EEG feature, we dropped data Am∗n∗k from the three-dimensional to the two-dimensional. That is, we removed the time dimension. We use Am∗(n−1) presented, which defined as follows:

$$A = \left\{ \begin{array}{cccc} x_{11} & x_{12} & \cdots & x_{1(n-1)} \\ x_{21} & x_{22} & \cdots & x_{2(n-1)} \\ \cdots & \cdots & \cdots & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{m(n-1)} \end{array} \right\} \quad (3.4)$$

Where x_{ji} stands for the ith electrode cap obtained value through incremental value feature extraction methods in the jth experiment. Where m represents the number of times each participated subjects in the experiment, n stands for the position of the electrode cap which is the value i in definition 1. The following is the main idea of the algorithm with R language.

Algorithm 2-1 The new feature extraction algorithm

Input ： A_{m*n*k} the data set which is acquired ； *i is the start time of the selected data in the time axis k, j is the ender time of the selected data in the time axis k*； E is the correct data and F is the value of experiments.

Output ： the result of feature extraction ；

1. Selected the useful data set in A_{m*n*k} according the input value of i and j. The selected data set is $A_{m*n*[i:j]}$, denoted as B_{m*n*z}
2. For (x in 1:n){
3. For (y in 1:m){
4. For (z in 1:z){
5. If (B[y,x,z+1]<B[x, y ,z])
6. C[y, z]=1;
7. Else   C[y, z]=1;
8. C_m is the summation of the matrix by columns in matrix C_{m*Z}}}
9. Then each column will be composed of a matrix named D_{m*n} by columns
10. ϵ is the subjects experimental error, where ϵ =E-F.
11. set a threshold δ
12. if ϵ<=δ ,the objects momeried
13. else the objects did not memory
14. Divide the data set to training data and testing data.
15. Use the classification method LDA and classification tree to classify in training data.
16. And then predict with the testing data

When the feature extraction is finished, linear discriminant analysis (LDA) and classification tree are used for analysis the results of feature extraction and prediction whether the subjects have been remembered. Fishers classical linear discriminant analysis (LDA) is still one of the most widely used techniques for data classification. For two normal distributions with common covariance matrix γ and different means 1 and mean 2, LDA classifier achieves the minimum classification error rate [22]. Classification trees have become increasingly important due to their conceptual simplicity and their computational efficiency [23].

IV. Experiment and results

A. experiment

The experiment was designed by department of psychology university of Oregon United State.

We will introduce this experiment in detailed. Figure 4.1 depicts the course of the experiment in general.
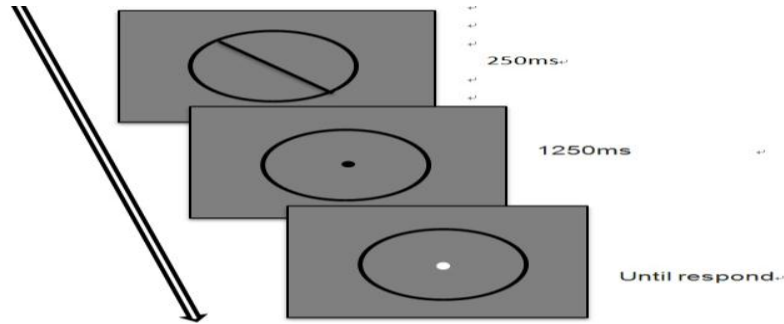

Figure 4.1. Experimental flowchart each experiment

Where the left figure in figure 3.1 is the memory screen sustained 250 milliseconds. During this time the subjects should keep in mind the direction of the black line inside the black circle. In next 1250 milliseconds the center figure is emerged in which the slash is disappeared and the black circle is sustained. When the center black point becomes white, the subject should draw the slash just like the slash in left figure according the memory. The above is a detailed process of the experiment. A subject should complete the experiment 960 times. And there is a 1500 milliseconds blank screen between the neighboring two trails.

During the experiment we should acquire the signal every 4 milliseconds. We storage the collected signals in a three-dimensional array $A_{960*20*751}$. 960 represents the times of experiments each subject. 20 is the electrode cap. 751 is the time axis. For example $A_{,2*1*1}$ is the signal value of PO3 at the first timestamp in the second experiment.

B. results

In this paper we have not discussed the pre-proposing about the collected signals. First we should remove the null values. After removing the null values we get the signals $A_{960*20*751}$. In older to analyze whether the subject has been remembered the direction of the line, we selected the EEG data in stages of memory to analyze. In general the effective data is $A_{960*20*[375:438]}$. The memory stages signals of OP3 are visualized in Figure 4.2. The x axis is the time in stages of memory. The y axis is the times of experiments. The z axis is the value of signals. The Figure4.3 is the memory stages signals of all electrode cap in one experiment.
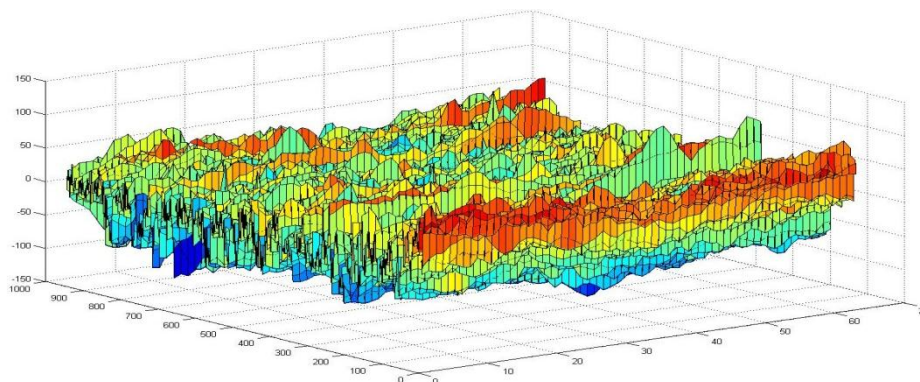

Figure 4.2. the memory stages signals of OP3

First of all, we use the method of incremental value which proposed in section 2. We use the formula of 3.1-3.3 to extract the EEG data of memory stages. Then we get the feature extraction value matrix A which is depicted in figure 4.4. A is a matrix of 905 rows and 20 columns. Rows are the number of trials and columns are the 20 electrodes.
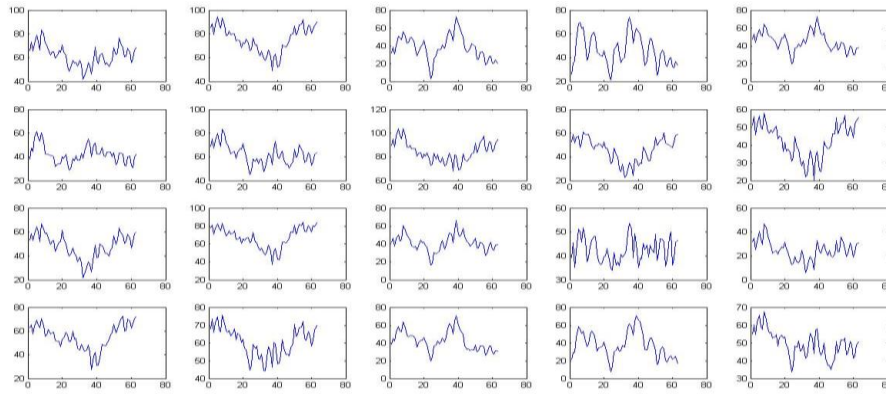
Figure 4.3. the signals of 20 electrodes in one experiment

	PO3	PO4	F3	F4	C3	C4	P3	P4	O1	O2	OL	OR	T3	T4	T5	T6	POz	Cz	Fz	Pz
1	31	32	29	32	29	29	29	32	32	32	33	32	29	30	29	32	29	29	28	28
2	30	34	33	37	29	34	30	33	32	30	29	34	34	34	32	37	33	31	34	31
3	31	30	27	29	32	32	31	31	31	30	35	30	33	37	33	27	33	28	26	32
4	31	32	32	35	34	35	33	33	32	32	33	31	36	35	31	35	36	35	36	34
5	32	33	33	35	31	35	31	33	32	32	35	35	30	34	33	35	33	31	32	34
6	37	32	32	27	30	30	33	30	36	32	33	31	30	33	34	32	32	27	30	31
7	35	31	31	32	33	32	33	33	31	29	37	31	33	29	32	33	29	32	32	30
8	34	33	35	32	35	31	34	33	38	39	34	38	36	37	35	37	37	32	32	35
9	35	38	33	34	35	35	35	34	31	35	36	38	35	35	34	37	35	36	33	35
10	33	37	30	27	33	33	31	37	36	35	31	35	37	38	37	35	33	28	30	35
11	31	32	33	34	35	34	35	33	30	34	31	35	37	31	32	35	34	32	33	35
12	33	32	31	33	30	30	35	29	33	32	34	33	36	34	37	34	33	31	32	30
13	31	30	36	34	35	32	31	32	34	33	33	31	33	30	35	33	33	36	31	34
14	38	35	33	29	34	35	34	34	34	29	34	30	31	34	37	33	32	31	31	32
15	33	32	29	27	30	32	32	31	37	32	38	33	33	36	33	32	32	31	30	32
16	34	34	34	33	31	39	31	34	32	34	35	32	28	34	36	35	33	32	32	32

Figure 4.4. a part of the feature extraction results with incremental value

After extracting feature, classification is raised. Figure 4.5 is the accuracy rate of using LDA classification to predict with unequal memory error . Figure 4.6 is the accuracy rate of using classification tree to predict with unequal memory error ε. The results of feature extraction with incremental value are compared with the mean absolute amplitude, standard deviation/variance and kurtosis in figure 4.5 and figure 4.6. Figure 4.7 is the result of accuracy rate with LDA classification to classification tree. In figure 4.5 we find that the extraction feature with incremental value method is generally better than the mean absolute amplitude, standard deviation/variance and kurtosis. The figure 4.6 also illustrates this point. In figure 4.7 we observed the result of LDA classification is better than classification tree. But the differences of two classification methods are not so considerable.
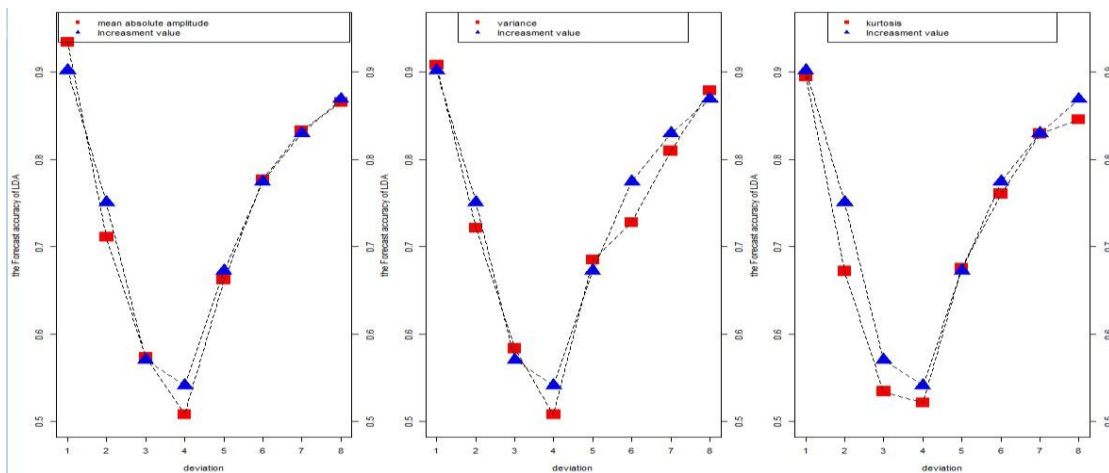

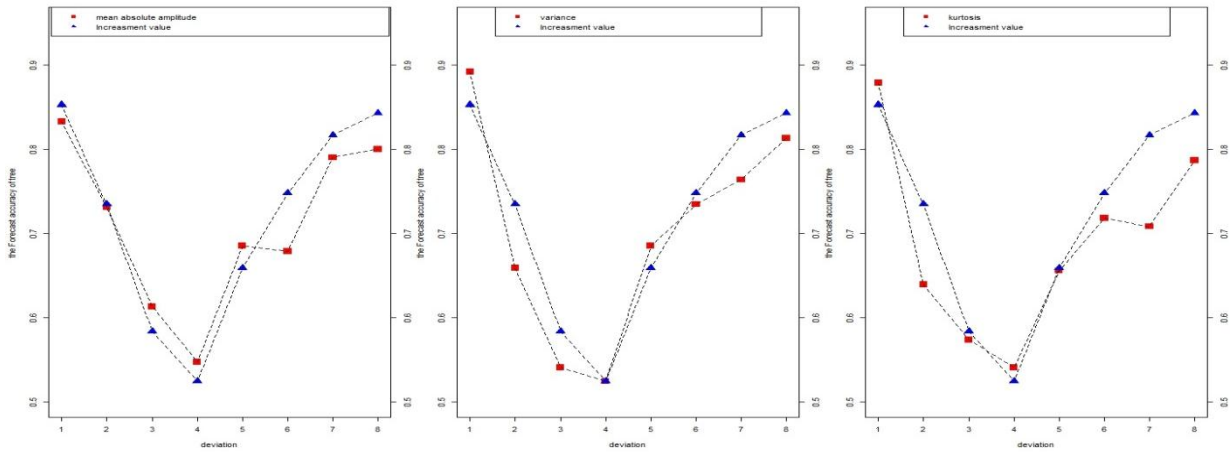Figure 4.5. the accuracy of using LDA to predict with unequal ε

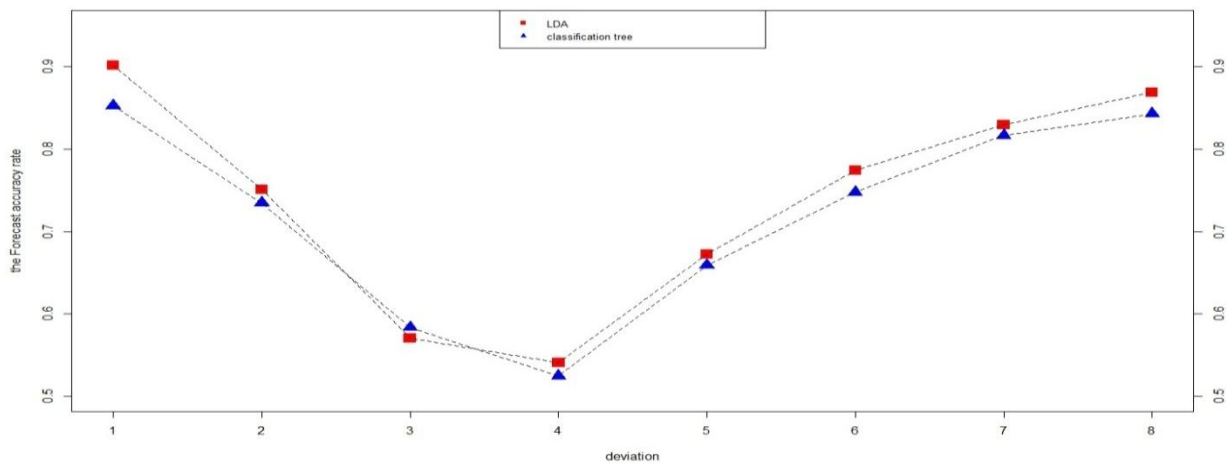Figure 4.6. the accuracy of using classification tree to predict with unequalε

Figure 4.7. the accuracy of using classification tree to LDA

V. Conclusion and future work

In this paper, we firstly reviewed the various feature extraction methods. Then we proposed a new feature extraction method named incremental value method. Next we used the classification tree and LDA to classify and predict the objects memory. We found that the method we proposed has a higher accuracy rate in testing dataset. In future work, other more exotic classifiers such as support vector machines, neural networks or other hybrid pattern classifiers will be considered.

VI. References

[1] Shayan Motamedi-Fakhra, Mohamed Moshrefi-Torbati et. Signal processing techniques applied to human sleep EEG signalsłA review. Biomedical Signal Processing and Control.10(2014)21-33.

[2] N.V. Thakor, S.B. Tong, Advances in quantitative electroencephalogram analysis methods, Annu. Rev. Biomed. Eng. 6 (2004) 453C495.

[3] M. Carrozzi, A. Accardo, F. Bouquet, Analysis of sleep-stage characteristics in full-term newborns by means of spectral and fractal parameters, Sleep 27(2004) 1384C1393. [4] B. Hjorth, EEG analysis based on time domain properties, Electroencephalogr. Clin. Neurophysiol. 29 (1970) 306C310.

[5] M. Kaminski, K. Blinowska, W. Szelenberger, Topographic analysis of coherence and propagation of EEG activity during sleep and wakefulness, Electroencephalogr. Clin. Neurophysiol. 102 (1997) 216C227

[6] S.V. Vaseghi, Advanced Signal Processing and Digital Noise Reduction, WileyTeubner, 1996

[7] P.S. Addison, Fractals and Chaos: An Illustrated Course, IOP publishing Ltd.,1997 Vienna, Austria, 2005, p. 6.

[8] C.K. Peng, S.V. Buldyrev, S. Havlin, M. Simons, H.E. Stanley, A.L. Goldberger, Mosaic organization of DNA nucleotides, Phys. Rev. E 49 (1994) 1685C1689.

[9] L. Jong-Min, K. Dae-Jin, K. In-Young, S.I. Kim, Analysis of scaling exponents of waken and sleeping stage in EEG, in: J. Mira, A. Prieto (Eds.), Proceedings, Part 1, 6th International Work-Conference on Artificial and Natural Neural Networks, IWANN 2001 Granada, Connectionist Models of Neurons, Learning Processes, and Artificial Intelligence, Springer-Verlag, Granada, 2001, pp. 450C456, Spain, June 13C15.

[10] Y. Shen, E. Olbrich, P. Achermann, P.F. Meier, Dimensional complexity and spectral properties of the human sleep EEG, Clin. Neurophysiol. 114 (2003) 199C209.

[11] C.L. Nikias, J.M. Mendel, Signal processing with higher-order spectra, IEEE Signal Process. Mag. 10 (1993) 10C37

[12] L. Cohen, TimeCFrequency Analysis, Prentice-Hall, New Jersey, 1995

[13] J. Zygierewicz, K.J. Blinowska, P.J. Durka, W. Szelenberger, S. Niemcewicz, W. Androsiuk, High resolution study of sleep spindles, Clin. Neurophysiol. 110 (1999) 2136C2147

[14] S.M. Pincus, I.M. Gladstone, R.A. Ehrenkranz, A regularity statistic for medical data-analysis, J. Clin. Monit. 7 (1991) 335C345.

[15] Roschke, J. Fell, P. Beckmann, The calculation of the 1st positive Lyapunov exponent in sleep EEG data, Electroencephalogr. Clin. Neurophysiol. 86 (1993) 348C352. [16] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control Signals Syst. 2 (1989)303C314

[17] T. Kohonen, T. Honkela, Kohonen Network, 2007

[18] L. Devroye, L. Gy?rfi, G. Lugosi, A Probabilistic Theory of Pattern Recognition (Stochastic Modelling and Applied Probability), Springer, 1996.

[19] A. Flexer, G. Dorffner, P. Sykacek, I. Rezek, An automatic, continuous and probabilistic sleep stager based on a hidden Markov model, Appl. Artif. Intell. 16(2002) 199C207.

[20] J.C. Bezdek, S.K. Pal, Fuzzy models for pattern recognition: background, significance, and key points, in:Fuzzy Models for Pattern Recognition, IEEE Press, 1992

[21] F Lotte, M Congedo et. A review of classification algorithms for EEG-based brainCcomputer interfaces. 31 January 2007.

[22] Nguyen Hoang Huy, Stefan Frenzel et. Two-Step Linear Discriminant Analysis for Classification of EEG Data. 2014

[23] Saul B. Gelfand, C. S. Ravishankar et. AN ITERATrvE GROWING AND PRUNING ALGORITHM FOR CLASSIFICATION TREE DESIGN.1989

[24] W.-L. Yeh, Y.-C. Huang et. A Self-Produced Mother Wavelet Feature Extraction Method for Motor Imagery Brain-Computer Interface. 35th Annual International Conference of the IEEE EMBS,Osaka, Japan, 3 - 7 July, 2013. [25] Deng Wang, Duoqian Miao, Chen Xie.Best basisbased wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection,Expert Systems with Applications, 2001.

[26] Shufang Li, Weidong Zhou et. Feature extraction and recognition of ictal EEG using EMD and SVM. Computers in Biology and Medicine.2013

[27] Shujuan Geng, Weidong Zhou et. EEG non-linear feature extraction using correlation dimension and Hurst exponent, Neurological Research, 2014.