

Pattern Identification of Subthalamic Local Field Potentials in Parkinson's Disease

Kun Zhang^{1,2}, Bin Feng^{1,2}, Yun-Peng Zhang¹, Yong-Zhi Huang^{2,3} and Shou-Yan Wang^{1,4,*}

¹Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, Jiangsu, China

²School of Communication and Information Engineering, Shanghai University, Shanghai, China

³The Functional Neurosurgery Group, Department of Surgery, John Radcliffe Hospital, Oxford, UK

⁴Institute of Science and Technology for Brain-Inspired Intelligenc, Fudan University, Shanghai, China

*Shouyan WANG: shouyan@fudan.edu.cn

Abstract. The local field potentials (LFP) in Parkinson's disease (PD), which contain abundant information related to disease and symptoms, are important for clinical treatment. The amplitudes of oscillations in LFP and the balance between them were found to be involved in brain functional state of PD patients. The LFP recorded from subthalamic nucleus before and after medication treatment were selected in this study. The power spectral ratio between frequencies and the percentage of energy corresponding to the wavelet packet nodes in the overall signal energy based on wavelet packet analysis related to symptoms were extracted as features. The brain states related to medication treatment in patients were classified using machine learning. The Naive Bayesian classifier and support vector machine (SVM) classifier were used to classify the states of off and on medication conditions. The results showed that the Naive Bayesian classifier was better than SVM classifier with higher accuracy. The specificity of Naive Bayesian classifier reached to 82.4%. The method proposed in this paper can accurately identify the brain functional state of PD patients.

1. Introduction

Parkinson's disease (PD) is a common degenerative disease of the central nervous system. At present the PD is estimated to affect around 180 out of 1,00,000 people.

The accurate identification of PD is of great significance to clinical diagnosis and treatment. At the present stage, machine learning method has been widely used in clinical medical science research. The use of machine learning to diagnose PD had achieved a good diagnostic effect [1]. PD patients with wearable sensor gait data and patients voice signals data also were used to diagnose PD [2].

The subthalamic nucleus (STN) is closely related to symptoms of PD patients [3]. The local field potential (LFP) signals were recorded by implanted electrode into the patients STN which contains abundant information related to disease and symptoms [4,5]. Therefore, this paper used the machine learning procedure to study the patients LFP before and after medication treatment, achieved the identification and judgment of brain function in PD patients.

In this paper, the amplitudes and dynamic regularity of oscillations in STN LFP as PD features were analysed. The balance in LFP different frequency bands and the power spectral ratio between frequencies and the percentage of energy corresponding to the wavelet packet nodes in the overall signal energy based on wavelet

packet analysis related to symptoms were extracted as features. In this paper, established the model of brain function judgment in PD patients and provided a method that identified brain functional state of PD patients in order to achieve intelligent, adaptive and closed-loop deep brain stimulation (DBS).

2. Methods

2.1 LFP Recording and Preprocessing

17 patients participated in this study as they were undergoing surgical DBS treatment at John Radcliffe Hospital (JRH), Oxford, UK. Postoperative magnetic resonance imaging confirmed the successful implantation of DBS electrode within the STN. The bipolar electrode recordings were made 3-5 days post-operatively from adjacent pairs of the externalized electrode. DBS electrode is 3389 (Medtronic, Minneapolis, MN, USA), it has four ring Pt-Ir cylinder surface contact. The distance between every contacts is 0.5 mm, the diameter of contact is 1.27 mm, length is 1.5 mm. The four contacts defined as 0-1-2-3, using the adjacent contact differential record when recording. The LFP signals were then amplified using isolated CED 1902 amplifier (Cambridge Electronic Design, UK), magnification of 10000 times,

digitized using CED1401 II and acquired using Spike2. The signals LFP sampling frequency is 2000 Hz.

17 patients underwent bilateral implantation of quadripolar DBS electrodes in STN. Five sides recording LFP signals had problems, so a total of 29 sides data. The recordings took place after 1.5 h of taking the usual pre-surgery dose of medication for the on medication condition. The recordings took place after withdrawal of medication 24h and medication efficacy had diminished for the off medication condition. Patients were instructed to relax for the recordings, which lasted around 5 min. Interception of stable and non-tremor 25 s data length for analysis. LFP were low-pass filtered (<90Hz, Chebyshev I). Interference around 50Hz is removed by an adaptive filter based on the signal in 40-60Hz band. Then LFP were high-pass filtered (>3Hz, Chebyshev I)[6].

2.2 Classification Model Flow of Brain Function State

Firstly, the model is used to analyze the PD patients LFP before and after medication treatment, and then extracts features that are closely related to the PD. So as to we can classify the PD patients LFP before and after medication treatment, and accurately identify the brain functional state of PD patients. The brain functional state of PD patients identification model flow chart is shown in Figure 1:

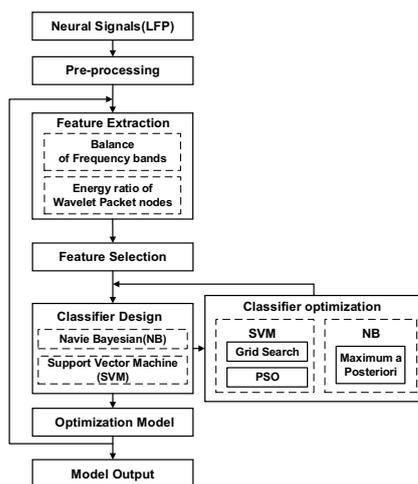


Figure 1. Flow chart of identify the brain functional state identification

3. Feature Extraction and Classifier Design

3.1 Balance Feature

Medication have different effects on the different frequency bands of LFP oscillations. In this paper, we used the power spectrum to analyze the data of 17 PD patients with a total of 29 sides. Quantify the degree of synchronization of the different frequency bands of patients LFP before and after medication treatment

(Figure.3). From the Figure 3, we can see that the power spectral density (PSD) amplitudes of LFP oscillations increase in theta bands (4-8 Hz), and high beta bands (20-30 Hz), and decrease in the range of low beta bands (12-18 Hz). And only in the low beta bands Dopamine drugs can significantly suppress the STN oscillations of the band ($P < 0.05$, t-test, Figure.3). We can also find that the balance of PSD between different frequency bands in LFP before and after the medication are also different.

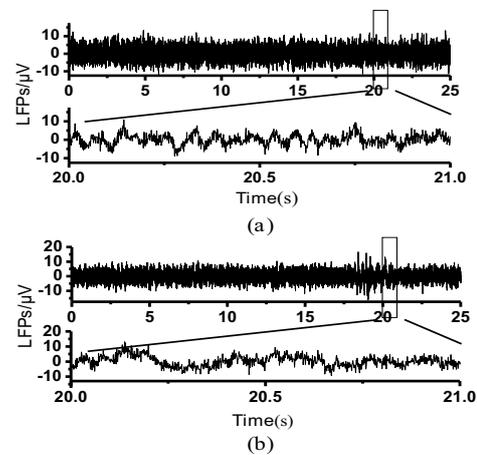


Figure 2. Local field potentials recorded off (a) and on (b) medication

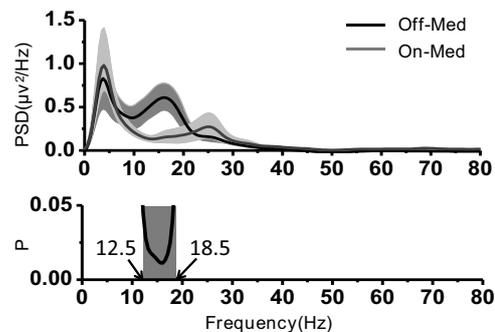


Figure 3. Power spectral density of LFP and the t-test significant of on and off medication

In this paper, the low frequency components of LFP signals were divided into four bands: theta (4-8 Hz), alpha (8-12 Hz), low beta (12-18 Hz) and high beta (20-30 Hz). Respectively calculated the power spectral ratio between each frequency band, used paired sample t-test analysis of STN LFP before and after medication to extract characteristics of different frequency bands energy ratio closely related to PD, for quantifying balance between different frequency bands.

Table 1 shows the balance between the low frequency different components of LFP signals for PD patients

Before and after medication treatment. The energy ratio between the different bands is calculated by the ratio of the parameters in the longitudinal direction of the table to the parameters in the transverse direction. As can be seen in Table 1 that the ratio of 20-30 Hz to 12-18 Hz inpatients STN LFP significantly increase after medication treatment (t-test, $P < 0.01$), while the ratio of 12-18 Hz and 8-12 Hz, 12-18 Hz and 4-8 Hz, 8-12 Hz

and 4-8 Hz, the significantly decrease after medication treatment (t -test, $P < 0.05$). Therefore, this paper selected these four parameters as the characteristic biomarkers for PD patients LFP, and was used to accurately identify the STN brain functional state of PD patients.

Table 1. The balance of different LFP frequency bands on and off medication

Balance	Theta	Alpha	L-beta	H-beta
Theta	1			
Alpha	1.39±1.19 **	1		
L-beta	5.89±9.39 *	2.81±2.79 **	1	
H-beta	3.29±5.54 1.43±2.27	2.06±2.54 2.12±2.51	1.04±0.85 1.78±1.31 **	1

Note :Theta:(4-8 Hz), Alpha:(8-12 Hz), L-beta:(12-18 Hz), H-beta:(20-30 Hz). * represent $P < 0.05$, ** is $P < 0.01$, t -test.

3.2 Wavelet Packet Feature

3.2.1 Wavelet Packet Transform

Wavelet packet transformation (WPT) is widely used in signal analysis, it has excellent time-frequency local characteristics and can describe the characteristics of a non-stationary signal within a local time range. WPT provides a more accurate method than others to divide the frequency bands and can make further decomposition in the high frequency bands [7]. The i level p point node (i, p) can be calculated using fast algorithm for orthogonal WPT:

$$\psi_{i+1}^{2p}(k) = \sum_{n=-\infty}^{\infty} l[n] \psi_i^p(k - 2^i n) \quad (1)$$

$$\psi_{i+1}^{2p+1}(k) = \sum_{n=-\infty}^{\infty} h[n] \psi_i^p(k - 2^i n)$$

where $l[n]$ is a low pass wavelet filter and $h[n]$ is a high pass wavelet filter. Wavelet Packet Decomposition is a kind of high and low pass filtering using a pair of conjugated high-pass and low-pass filter to iterate the signals. Different wavelet basis correspond to different wavelet filters, we used Meyer wavelet basis analyzing LFP signals[8]. The coefficients of decomposition of the i th layer can be calculated according to the coefficients of $i-1$, by this way all other coefficients can be obtained. The frequency range of each node can also be calculated through: $[0, f_s / 2^{i+1}]$, $[f_s / 2^{i+1}, 2f_s / 2^{i+1}]$, $[2f_s / 2^{i+1}, 3f_s / 2^{i+1}]$, ..., $[(2^{i-1}-1)f_s / 2^{i+1}, f_s / 2]$. f_s is the sampling frequency.

3.2.2 Energy of LFP Wavelet Packet node

WPT of PD patients STN LFP before and after medication can decompose accurately. LFP signals In this paper, sampling frequency of the LFP signals down to 384 Hz and the down-sampling LFP signals was decomposed by 6-layers WPT. The decomposition results are showed in the Table 2. Then we reconstructed each wavelet packet node and calculated the percentage of energy of wavelet packet nodes reconstructed signals and total energy of signals to achieve features wavelet packet nodes. Assuming that (i, j) is the j th node of the i th layer. The energy ratio of reconstructed signals is:

$$p_{i,j} = \frac{E_{i,j}}{\sum_{j=0}^{2^N-1} E_{i,j}} \quad (2)$$

$p_{i,j}$ is the energy ratio of energy of reconstructed signals and total energy of signals and N is the layer number of decomposition.

Table 2 shows Wavelet Packet nodes of PD patients before and after medication and their energy ratio. As seen from Table 2, we can find that there are some differences between energy ratio before and after medication. The energy ratio of the Wavelet Packet nodes [6.2], [6.6] and [6.7] reconstruct signals significantly decrease after the medication. while energy ratio of nodes [6.0] and [6.1] significantly increase. Then we calculate every node using Fisher feature extraction method, then choose the subset of features that have the greatest distinguishing feature and obvious features that has the bigger Fisher score. Fisher score of nodes in Table 2 can be arranged from large to small as [6.7], [6.6], [6.1], [6.2] and [6.0]. We choose the first three nodes [6.7], [6.6] and [6.1] to analysis (Figure.4). From Figure.4 we find that these Wavelet Packet feature nodes can distinguish PD patients before and after taking medicine accurately.

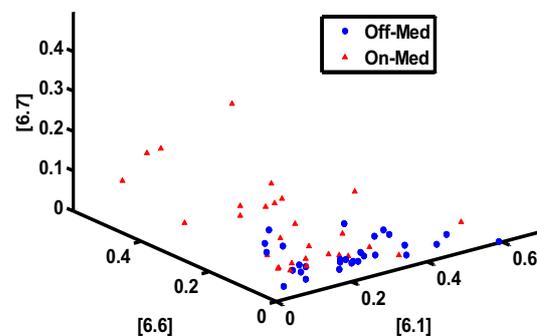


Figure 4. Scatter diagram of feature distribution. Energy ratio of Wavelet Packet nodes [6.1], [6.6] and [6.7] on and off medication

Table 2. Wavelet Packet nodes and energy ratio of Wavelet Packet nodes on and off medication

Wavelet Packet nodes	[6.0]	[6.1]	[6.3]	[6.2]	[6.6]	[6.7]
Frequency band/Hz	0-3	3-6	6-9	9-12	12-15	15-18
Off-Med energy ratio	0.037± 0.040	0.14± 0.12	0.095± 0.065	0.088± 0.053	0.13± 0.13	0.11± 0.097
On-Med energy ratio	0.065± 0.081	0.23± 0.14	0.12± 0.072	0.062± 0.027	0.048± 0.024	0.045± 0.031
<i>P</i> (t-test)	0.02*	0.0008**	0.1432	0.0171*	0.0016**	0.0009**
Fisher	0.10	0.2	0.04	0.19	0.372	0.457

Note: * represent $p < 0.05$, ** is $p < 0.01$, t-test

Bring the results of the calculation to the original function and simplify it to get function dual form:

3.3 Brain Functional State Classification

3.3.1 Navie Bayesian Classifier

Navie Bayesian(NB) Classifier can flexibly process and classify data, it has quite fast calculation rate of classification. Navie Bayesian Classifier is a kind of inverse probability derivation method that infers the probability of the unknown according to probability of something happened[9]. In this paper, firstly, we analyzed the nerve oscillation characteristics of PD patients in known states, calculated the probability distribution of its characteristics and probability of occurrence of LFP signals nerve oscillation of unknown state patients. Finally, we inferred the state of brain function of unknown state patients using Maximum a Posteriori(MAP) Navie Bayesian Classifier. After all, we can build a detection and identification model of PD patients brain function state to achieve fast patients brain function state determination.

$$P(\omega_i|X) = P(\omega_i) \prod_{k=1}^n P(X_k|\omega_i) \quad (3)$$

The sample is defined as $X=(X_1, X_2, X_3, \dots, X_k)$. Where X_k means the k th characteristic. $P(\omega_i)$ is prior probability of category ω_i .

3.3.2 Support Vector Machine Classifier

Support Vector Machine(SVM) Classifier is based on Structural Risk Minimization and statistics theory to deal with classification problem and is widely used in Biomedical field[10,11]. SVM classifier is a maximum interval function classifier containing hard interval

$$\min P = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \xi_i \quad (4)$$

support vector machine and soft margin support vector machine. SVM classifier aims to:

C is regularization parameter that penalizes the errors made by the model and ξ_i is slack vector.

$$\max_a = \sum_{i=1}^n a_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_i a_j y_i y_j \langle \phi(X_i), \phi(X_j) \rangle \quad (5)$$

The constraint : $0 \leq a_i \leq C, i=1, \dots, n, \sum_{i=1}^n a_i y_i = 0$.

When proceeding PD patients brain function state detection, we used LibSVM software package developed by team of professor Zhiren Lin of Taiwan University to analyze patients LFP. We analysed following four steps: Firstly, normalize PD patients characteristic LFP signals. Then used Radial basis function to transform signals dimension. And then used cross validation to train data and used Particle Swarm Optimization(PSO) algorithm to choose SVM Classifier optimal parameters, regularization parameter C and gamma in Kernel Function. Finally, implement an overall training to form a SVM classifier having stronger generalization ability and more precision as well as to establish the brain functional state of PD patients identification model.

4. Experimental Results

The balance of energy ratio between 4 bands in the low frequency components of PD patients STN LFP and the percentage of energy corresponding to the wavelet packet nodes [6.7], [6.6] and [6.1] in the overall signal energy based on wavelet packet, that can effectively reflect the brain functional state of PD patients. In this paper, we selected these seven parameters as the features of the classification of brain function state of PD patients, used the naive Bayesian classifier and SVM classifier for classification and used 5-fold cross validation, the results shown in Table 3.

Table 3. Classifier comparison

Algorithm	NB	SVM
Accuracy(%)	71.90±2.31	68.10±2.19
Sensitivity(%)	61.03±4.00	59.31±3.91
Specificity(%)	82.41±3.79	77.93±5.91

As shown in Table 3, the Naive Bayesian classifier is more accurate classification than the SVM classifier in PD patients, and naive Bayesian classifier classification detection the highest specificity is 82.7±13.79% for the PD and SVM classification the highest specificity is only 77.93±5.91. For the detection sensitivity is also naive Bayesian classification accuracy higher than SVM classifier.

5. Discussion

In the diagnosis of PD, the behavioral characteristics of the patients were diagnosed in the past and were not analyzed from the intrinsic patterning characteristics of the patient. With the continuous development of DBS at present, it has been more important to analyze the patterning characteristics of LFP signals that contains a wealth of physiological information. In this paper, specific nerve oscillations were closely related to the pathogenesis of PD: the energy balance between the low frequency bands and the energy ratio of each nodes to reconstruct the signals after wavelet packet transformation. The detection model was established using the maximum a posteriori Naive Bayesian and SVM classifier to achieve the classification of the brain functional state of PD patients. The results show that the accuracy, specificity and sensitivity of Naive Bayesian classifier identification algorithm are higher than SVM classifier. In addition, the specificity of the identification algorithm is high, but the sensitivity still has room for improvement. Less sample data, prone to over-fitting, is one of the reasons for the lower accuracy of classification, so to increase the experimental sample of PD will be the focus of the next step. The study of the characteristics of PD patientd LFP can improve the accuracy of detection classification in the further, which will be the important direction of the next step.

References

1. S Patel, K Lorincz, R Hughes, IEEE T INF TECHNOL B, 13, 864 (2009)
2. MA Little, PE Mcsharry, EJ Hunter, IEEE Trans Biomed Eng, 56, 1015 (2009)
3. JS Brittain, P Brown, Neuroimage, 85, 637, (2014)
4. G Foffani, A Priori, M Egidio, Brain, 126, 2153 (2003)
5. N Jenkinson, P Brown, TRENDS NEUROSCI, 34, 611 (2011)

6. YN Wang, XY Geng, YZ Huang, J Biomedical Eng, 33, 49 (2016)
7. MA Li, JY Ma, JF Yang, Chin J Sci Instru, 33, 1721 (2012)
8. KA Mamun, M Mace, ME Lutman, J NEURAL ENG, 12, 056011 (2015)
9. M Peker, B Sen, D Delen, J HEALTHC ENG, 6, 281 (2015)
10. E Byvatov, U Fechner, J Sadowski, J Chem Inf Comput Sci, 43, 1882 (2003)
11. S Pan, S Iplikci, K Warwick, EXPERT SYST APPL, 39, 10764 (2012)