DOI: 10.1051/itmconf/20171203020

An Automatic Parkinson 's Disease Recognition System Based on Multi - Feature Selection of Motion Signals

Tian-Yu Shen[1,2], Ji-Ping Wang[2], Jing Chen[2], Da-Xi Xiong[2], and Li-Quan Guo[2,a]

[1]*School of communication and information engineering, Shanghai University, 200444 Shanghai, China*
[2]*Suzhou Institute of Biomedical Engineering, Chinese Academy of Sciences, 215163 Suzhou, China*
[a] *Corresponding author: guolq@sibet.ac.cn*

Abstract. Parkinson's disease (PD) seriously affects human health so it has wide application value for its automatic diagnosis. In this study, 5 wearable inertial sensors were used in acquiring the acceleration and angular velocity signals under 4 paradigm actions. Total of 27 features were extracted from the signals, including amplitude, frequency, fatigue degree, self similarity, cross correlation and approximate entropy of an action. Genetic algorithm and BP neural network was used for feature selection and data classification. The experiment data were acquired from 10 PD patients and 10 healthy subjects. The results showed that the classification efficiency was improved after feature selection, and the average sensitivity, specificity and accuracy of the classification were 87%, 100% and 93% respectively. It may have certain application value in computer aided diagnosis of Parkinson's disease.

1. Introduction

Parkinson's disease (PD) is a neurological degenerative disease whose main clinical symptoms include motor symptoms such as tremor, muscle rigidity, dyskinesia, postural balance disorders, and non-motor symptoms such as loss of mood, sleep disorders and depression[1-2].The prevalence rate of people over 65 years old in China is 1700 per 10 million, and increases with age. So It is of great significance to quickly and accurately diagnose Parkinson's disease [3]. Currently, the diagnosis of PD is mainly based on clinical scales such as UPDRS, which contains medical history, clinical symptoms and anti-PD drug response [4]. Clinical diagnosis has high sensitivity, but the specificity is less than 75%. Identification of subclinical or early stage of Parkinson's disease provides the possibility of early neuroprotective therapy, delaying the development of the disease [5].

Quantitative assessment of Parkinson's disease based on motion signals is an objective assessment method that can improve the accuracy of PD diagnosis, improve the diagnostic criteria for PD, and establish a large-scale data analysis of PD. It uses wearable inertial sensors on the body to collect acceleration, angular velocity and other motion signals to be analyzed to obtain quantitative parameters to assess the motor function of patients[6][7]. Körver R J P et al. [8] used the asymmetry parameters to evaluate the shoulder motion function; Yoneyama M et al. [9] used periodic, cross-correlation and other parameters for gait analysis of PD patients. But most of the current studys only focus on a single symptom or one body part, and an automatic identification model of Parkinson's disease with high accuracy is needed.

Therefore, five wearable inertial sensors were used in our study to obtain the three-axis acceleration and angular velocity signals of the patients when doing the paradigm motion tasks. Twenty-seven quantitative parameters describing the motor function of different body parts were extracted. Finally feature selection algorithm and BP Neural network was used to establish a Parkinson 's disease automatic identification model.

2. Materials and Methods

2.1 Wearable Sensors

A total of five sensor nodes were used in this study, which can collect three axis acceleration, angular velocity and magnetic signals at the same time. The size of each node is 39×33×16mm, the average weight is 18g, and the sampling rate is 100HZ. The hardware circuit of the sensor node mainly includes the sensor module, the power module and the control module. MPU9250 chip is used as the motion sensor module, CC3200 chip is used as the control module, and lithium battery is used for power supply. When working each node acquires and transfers the motion data to the host computer wirelessly through the wifi communication network. The physical map of the sensor node and the positions to wear in body are shown in Figure 1:

DOI: 10.1051/itmconf/20171203020

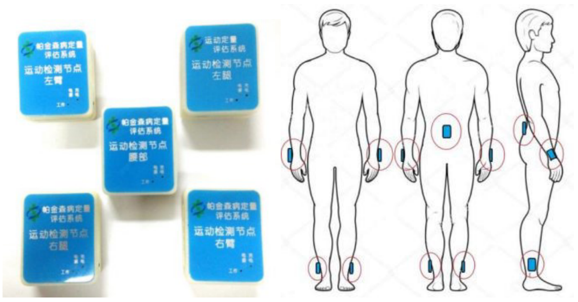


Figure 1. Wearable sensor nodes and the wearing positions.

2.2 Experiment & Data Acquisition

The PD group in the experiment included 10 patients with mild to moderate Parkinson's disease, age 58.7±8.5 years, height 165.8 ± 7.8 cm, weight 69.8 ± 11.5 kg, Hoehn-Yahr stage I-III. The control group included 10 young healthy subjects without any motion related disease, age 25.2 ± 2.2 years, height 168.1 ± 7.1cm, weight 62.8 ± 10.4kg. The 4 paradigm action tasks included 10m time up and go, hand rotation, leg raising and standing still, time duration of each action was 20s, and between every two tasks there was a 1~2min rest. The 5 wearable inertial sensors were worn on the left hand, right hand, left leg, right leg and waist of the subject's body using an elastic bandage. When completing the action tasks, the sensor nodes send the motion signals to the host computer to be displayed and saved. The experiment was conducted at the Second Affiliated Hospital of Soochow University and the informed consent of the subjects and their families was obtained.

2.3 Data Preprocessing

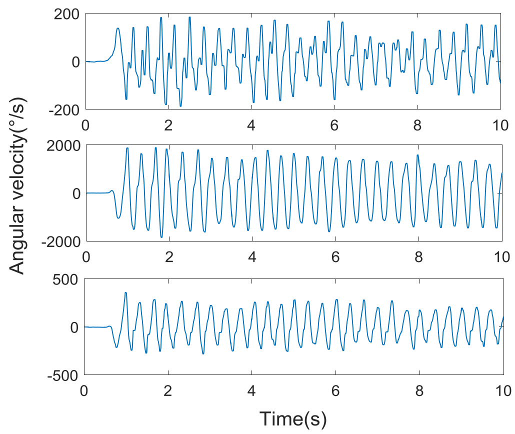
The x, y, z triaxial acceleration and angular velocity signals were filtered using a median filter of 5 points to remove isolated noise points. In order to eliminate the baseline offset, all the signals were detrended. For 10m time up and go, hand rotation and leg raising task, the angular velocity with maximum signal to noise ratio from the corresponding three-axis signals were selected through the calculation of variance to do further analysis. For standing still task, the x, y two-axis acceleration signals from the waist node were used. Due to abnormal shaking during standing, there were some abnormal signals. So the standard deviation σx(i) of the data from a 100-point sliding window and the standard deviation σx of the residual data were calculated for comparison. If σx(i) ≥ 2σx, the outliers x(i) from the window were removed.

2.4 Motion Function Features Extraction

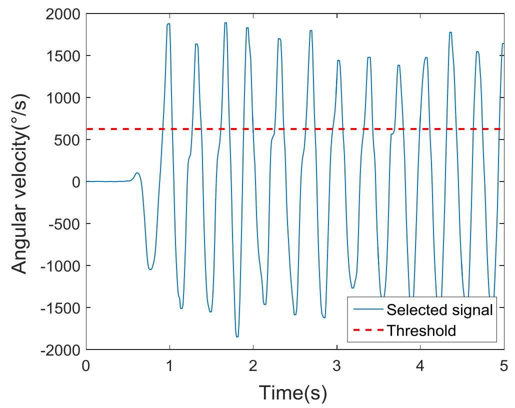
For the 10m time up and go, hand rotation and leg raising task, angular velocity signal v(t) after preprocessed was used. Mean value E(v(t)) and the standard deviation σ(v(t)) of the signal were calculated, and the judgment threshold value T0 was obtained according to the formula (1). Figure 2 shows the triaxial angular velocity signals of left hand rotation and the judgment threshold T0 in the selected signal.

$$T_0 = E(v(t)) + \sqrt{2} * \sigma(v(t))/2 \qquad (1)$$

The value of v(t) was compared with the threshold value T0 each time to obtain the number and frequency of accomplished actions. Then the angle signal A(t) was obtained by integrating the velocity signal v(t), and the amplitude information of the action was calculated. Due to the turning action in walking, the time required for the angle signal A(t) of the waist node to change from 40° to 140° was defined as the turning time. In addition, the fatigue level was defined as the ratio of action frequency for the first five actions and the last five actions. For the acceleration signal of the waist node when standing still, the root mean square RMSA was calculated, and then the velocity signal v(t) was obtained by integrating the acceleration signal to get the average velocity MV.



(a)



(b)

Figure 2. (a) The triaxial angular velocity signals of left hand rotation. (b) The judgment threshold T0 in the selected signal.

The parameters above were the results of intuitive analysis of the signal. In order to obtain the randomness characteristic, the approximate entropy (ApEn) was calculated. First, the acceleration or angular velocity signal X(i) was processed according to the formula (2) in order to reconstruct the m-dimensional phase space, in which $1 \le i \le N - m + 1$.

$$\mathbf{X}(i)=\left[x(i),x(i+1),\cdots \quad m-1)\right] \tag{2}$$

The distance between vectors **X(i)** and **X(j)** was calculated using formula (3), in which $0 \le k \le m-1$, $1 \le i, j \le N-m+1$.

$$d_{ij} = \max\left|x(i+k)-x(j+k)\right| \tag{3}$$

The similarity tolerance $r > 0$ was selected , for each **X(i)**, the number of $d_{ij} \le r$ were counted, and the ratio of the number to the total number of vectors $C_i^m(r)$ was calculated:

$$C_i^m(r) = \frac{1}{N-m+1}\{number\ d_{ij} \le r\} \tag{4}$$

Then take the logarithm of the ratio $C_i^m(r)$ and find the mean value $\Phi^m(r)$:

$$\Phi^m(r) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\ln C_i^m(r) \tag{5}$$

Finally increase the dimension m by 1, repeat the above steps to obtain $C_i^m(r)$ and $\Phi^{m+1}(r)$. Then the approximate entropy (ApEn) was calculated:

$$ApEn(m,r,N) = \Phi^m(r) - \Phi^{m+1}(r) \tag{6}$$

In order to obtain more features of the gait, the angular velocity signals of the left leg and the right leg were used to do the detrend fluctuation analysis (DFA) to obtain their self-similar parameters, and the detrend cross correlation analysis (DCCA) was used to obtain the cross-correlation parameters. First, the mean value $E(x)$ for the signal was calculated, i = 1, 2 ... N, and the cumulative sequence $y(k)$ was obtained by the formula (7):

$$\mathrm{y}(k) = \sum_{i=i}^{k}(X_i - E(x)) \tag{7}$$

The sequence $y(k)$ was divided into [N/n] intervals of length n, and the fitting function $yn(k)$ of each interval was obtained by linear least squares method. Then, the fluctuation mean square $F(n)$ of $y(k)$ was calculated :

$$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k)-y_n(k)]^2} \tag{8}$$

By repeating the above operations on all time scales, the relation curves of $F(n)$ and fragment size can be obtained. The slope determine scaling exponent alpha of the *log* ($F(n)$) to *log* (n) was calculated as self similar parameter. The cross-correlation parameters were obtained from the two time series covariates in the same way.

According to the methods mentioned above, a total of 27 features related to motor function were extracted, as shown in Table 1:

Table 1. Features extracted from motion signals.

Paradigm action tasks	Body parts	Quantitative parameters
10m time up and go	bilateral lower limbs waist	amplitude(°), frequency(count/min), ApEn, DFA, DCCA turning time(s)
hand rotation	bilateral upper limbs	amplitude(°), frequency(count/s), fatigue, ApEn
leg raising	bilateral lower limbs	frequency(count/s), fatigue, ApEn
standing still	waist	*RMSA*(m/s^2), *MV*(m/s), ApEn

2.5 Feature selection & Classification

The 27 extracted features were normalized to construct a feature vector as input to the network, 14 from the 20 subjects were randomly selected as the training set and 6 were used as the test set to do the BP neural network classification. Since the feature vector dimension was too large, the feature selection algorithm based on genetic algorithm was used to optimize the BP neural network. The chromosome length was set to 27, the population size was set to 20, and the maximum evolutionary algebra was set to 100. The fitness function was chosen as the reciprocal of the square sum of the classify error, as shown in formula (9), where Tp was the predicted value and T was the true value.

$$f(X) = \frac{1}{SE} = \frac{1}{sse(T_p - T)} \tag{9}$$

A total of five times classifications were carried out to improve the reliability of the results. The classification efficiency was compared before and after using the genetic algorithm each time. Accuracy, sensitivity, specificity and modeling time were used as indicators to evaluate the classification efficiency.

3. Results and Discussion

The results of the five times classifications were shown in Table 2 and Figure 3, where BP represented the BP neural network classification result and GA represented the result after genetic algorithm feature selection. It can be seen that the sensitivity, specificity and accuracy of the classification after feature selection were improved. Finally, the overall classification sensitivity, specificity and accuracy rate reached 87%, 100% and 93% respectively. The feature vector dimension required for each modeling after the feature selection was reduced to about half of the original dimension, which also shortened the time required for modeling. These results showed that the classification efficiency of the BP neural network was improved after the feature selection algorithm.

Table 2. The classification results of each time.

Num		Sen	Spc	Acc	Modeling time(s)
1	BP	0.50	1.00	0.67	0.16
	GA	1.00	1.00	1.00	0.12
2	BP	0.50	1.00	0.83	0.14
	GA	1.00	1.00	1.00	0.06
3	BP	0.50	1.00	0.67	0.06
	GA	0.75	1.00	0.83	0.06
4	BP	1.00	0.75	0.83	0.14
	GA	0.50	1.00	0.83	0.06
5	BP	1.00	0.33	0.67	0.09
	GA	1.00	1.00	1.00	0.06
Mean	BP	0.67	0.80	0.73	0.12
	GA	0.87	1.00	0.93	0.07

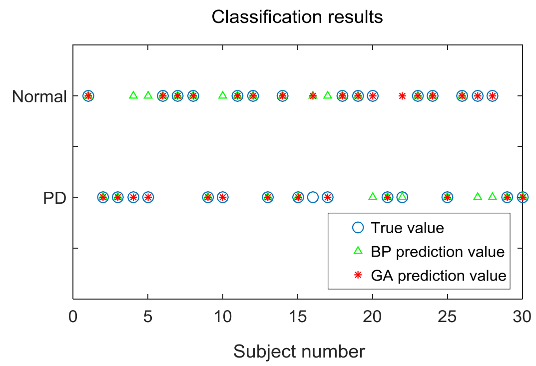


Figure 3. The overall classification results.

The evolution curve of the fitness function of the genetic algorithm was shown in Figure 4. It can be seen that as the evolutionary algebra increased to 100, the average fitness and the optimal fitness rised and eventually stabilize. Due to the number of experimental subjects was limited, the classification performance of this method need further verification, and the future work can also study the influence of frequency domain features on classification results.

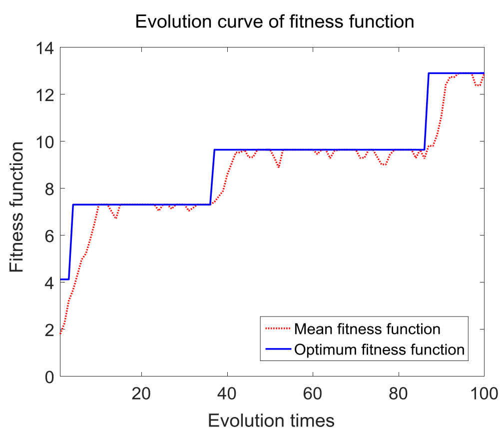


Figure 4. The evolution curve of fitness function.

4. Conclusion

In order to quickly diagnose Parkinson's disease, this paper proposed a classification algorithm based on multi-feature of motion signals. 5 wearable inertial sensors were used to obtain the acceleration and angular velocity signals of motion, multi-scale motor function parameters of different body parts were extracted from the signals, and genetic algorithm was used to select the parameters to participate in the BP neural network classification. The results showed that the classification sensitivity, specificity and accuracy of the method were 87%, 100% and 93% respectively. This may have certain application value in the early diagnosis and computer-aided diagnosis of Parkinson's disease.

References

1. Jankovic, J. Parkinson's disease: clinical features and diagnosis. Journal of Neurology Neurosurgery & Psychiatry ,79.4(2008)
2. Moustafa A A, Chakravarthy S, Phillips J R, et al. Motor symptoms in Parkinson’s disease: A unified framework. Neuroscience & Biobehavioral Reviews, 68(2016)
3. Zhang ZX RG, Hong Z, et al. Parkinson's disease in China: prevalence in Beijing,Xian, and Shanghai. Lancet ,365.3(2005)
4. Goetz, C. G. The Unified Parkinson's Disease Rating Scale (UPDRS): Status and recommendations. Movement Disorders, 18.7(2003)
5. Postuma, Ronald B, et al. MDS clinical diagnostic criteria for Parkinson's disease. Movement Disorders Official Journal of the Movement Disorder Society ,30.12(2015)
6. Yang, K., et al. Objective and quantitative assessment of motor function in Parkinson's disease-from the perspective of practical applications. Annals of Translational Medicine ,4.5(2016)
7. Dai H, Lin H, Lueth T C. Quantitative assessment of parkinsonian bradykinesia based on an inertial measurement unit. BioMedical Engineering OnLine, 14.1(2015)
8. Körver R J P, Heyligers I C, Samijo S K, et al. Inertia based functional scoring of the shoulder in clinical practice[J]. Physiological measurement, 35.2(2014)
9. Yoneyama, M, et al. Accelerometry-based gait analysis and its application to Parkinson's disease assessment--part 1: detection of stride event. IEEE Trans Neural Syst Rehabil Eng, 22.3(2014)