

## Research of P300 Feature Extraction Algorithm Based on Wavelet Transform and Fisher Distance

WANG Pan<sup>a,\*1</sup>, SHEN Ji-Zhong<sup>a,\*2</sup>, SHEN Jin-He<sup>a</sup>

<sup>a</sup> *Institute of Electronic Circuit and Information System, Zhejiang University, Hangzhou 310027, China*

---

### Abstract

With the aim to improve the divisibility of the features extracted by wavelet transform in P300 detection, we research the P300 frequency domain of event related potentials and the influence of mother wavelet selection towards the divisibility of extracted features, and then a new P300 feature extraction method based on wavelet transform and Fisher distance is proposed, which overcomes the drawbacks of no systematic feature selection method during traditional P300 feature extraction based on wavelet transform. In this paper, both the BCI Competition 2003 and the BCI Competition 2005 data sets of P300 were used for validation, the results showed that the proposed method can increase the divisibility by 121.8% of the features extracted by wavelet transform, and contribute to the followed classification.

**Index Terms:** BCI; Wavelet Transform; P300; Fisher Distance; Feature Extraction

© 2011 Published by MECS Publisher. Selection and/or peer review under responsibility of the International Conference on E-Business System and Education Technology

---

### 1. Introduction

A brain-computer interface (BCI) is a device that uses brain signals to provide a non-muscular communication channel between human beings and computers or peripheral electronic equipments[1]. At present, many kinds of EEG (electroencephalography) signals are utilized successfully in the field of BCI, such as motor imagery brain signals and P300[2]. Among these different kinds of EEG signals, P300 is drawing more and more attentions of the researchers because of its excellent characteristics, such as needless of training[3].

The P300 is a psycho-physiological correlates of neurocognitive functioning that reflect the response of the brain to events in the external or internal environment of the organism, which is classified as event related potentials of brain, and it is named so due to the positive deflection of the EEG at the central electrodes around 300ms post stimulus[4].

Recently, the researchers have proposed lots of feature extraction algorithms in order to extract the characteristics precisely, such as wavelet transform[5], independent component analysis[6] and adaptive filtering[7]. Wavelet transform is perfect for processing non-stationary random signals like EEG due to its

Corresponding author:

E-mail address: <sup>\*1</sup> [kel\\_1\\_y@126.com](mailto:kel_1_y@126.com); <sup>\*2</sup> [jzshen@zju.edu.cn](mailto:jzshen@zju.edu.cn)

multi-resolution[8], but traditionally, EEG signals are decomposed by wavelet, and then the time-frequency features between 0 Hz and 30 Hz are extracted and classified simply, refer to literature[9]. These similar algorithms are fast, but cause low divisibility at the same time[9].

With the aim to improve the divisibility of the features extracted by wavelet transform in P300 detection, we research the P300 frequency domain of event related potentials, and then a new P300 feature extraction method based on wavelet transform and Fisher distance was proposed, which overcome the drawbacks of no systematic feature selection method during traditional P300 feature extraction based on wavelet transform. We also research the influence of mother wavelet selection towards the divisibility of extracted features, and then we improve the P300 detection algorithm of literature [9] through proper electrode selection and wavelet decomposition layer.

## 2. Experiment Design and Data Source

The datasets we used are from P300 Speller Paradigm supplied by BCI Competition II Data set IIB and BCI Competition III data set II. The P300 Speller was proposed by Farwel and Donchin in 1988, which was based on oddball experiment paradigm, see Fig.1[10].

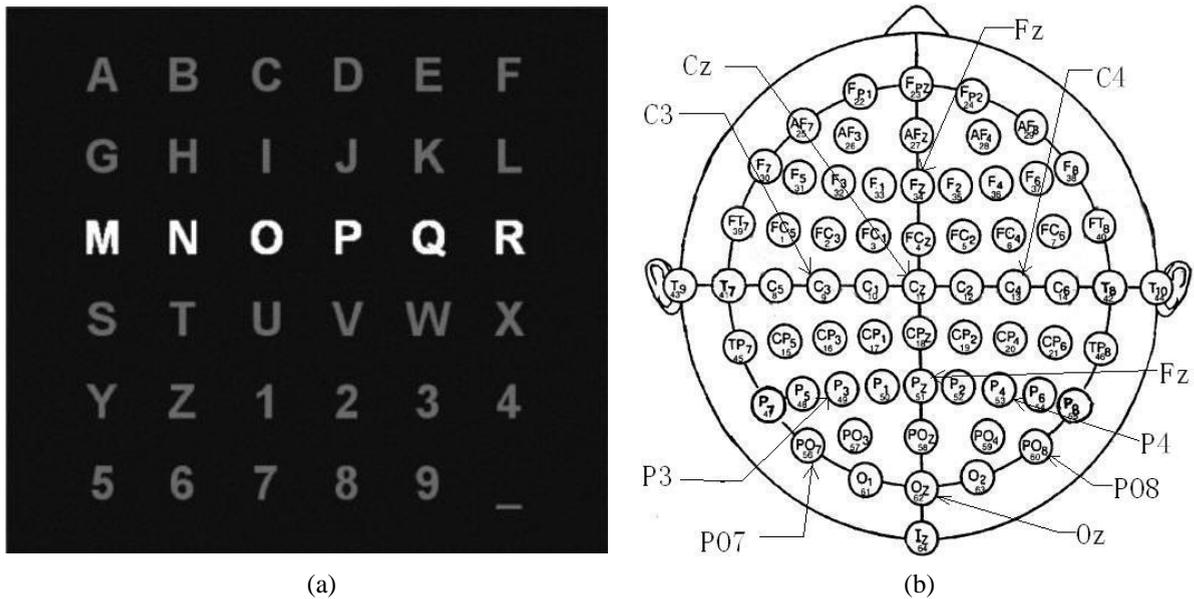


Fig. 1. Experimental set-up.  
 (a) Example of a 6 × 6 user display in P300 Speller.;  
 (b) Channel location and assignment numbers used for EEG acquisition

A 6 × 6 matrix that includes all the alphabet letters as well as other useful symbols is presented to the user on a computer screen, see Fig. 1(a). The rows and columns of the matrix are intensified successively in a random order. At any given moment the user selects one of the letters or symbols that he wishes to communicate, and maintain a mental count of the intensified chosen symbol. In response to this counting, a potential is elicited in the brain. This procedure is called one trail, which is repeated 15 times for each selected symbol. The EEG signal is acquired from 64 electrodes, whose locations are showed in Fig.1(b), and the sampling rate is 240Hz. Related literatures demonstrate that electrode Fz, Cz, Pz, Oz, C3, C4, P3, P4, PO7 and PO8 have much more

influences to P300 detection[11], so we choose the data from this 10 electrodes in our experiment, and extract 600ms' data after 12 rows/columns being intensified for each trail[4,10].

### 3. Feature Extraction and Divisibility Judging

In order to make the experiment results more reasonable, two classes of EEG signals should be balanced before feature extraction. We randomly choose 1000 groups of EEG signals with P300 and 1000 groups without, and call them target dataset and non-target dataset respectively.

First, wavelet transform is used to decompose original EEG signals, and extract approximate coefficients as features. Second, we use Fisher distance to analyze the divisibility of the extracted features.

#### 3.1. Wavelet transform

Wavelet transform is potentially one of the most powerful signal processing techniques because of its ability to adjust to signal components and its multi-resolution which is broadly used to analyze EEG signals in recent years[11].

The wavelet transform and inverse wavelet transform are defined below:

$$d_{j,k} = 2^{-j/2} \sum_{n=-\infty}^{\infty} f(x) \bar{\phi}(2^{-j}x - k), j, k \in Z \tag{1}$$

$$f(x) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_{j,k} \phi_{j,k}(x), j, k \in Z \tag{2}$$

In (1)(2),  $f(x)$  is the EEG signal, and  $d_{j,k}$  is wavelet coefficient, and  $\phi_{j,k}(x) = 2^{-j/2} \phi(2^{-j}x - k)$  is the mother wavelet. Integer  $j$  and  $k$  means decomposition scale and time shifts respectively. Mallat algorithm is usually applied to calculate the coefficients of wavelet transform in practice[11].

For example,  $S$  means a signal, and then the whole Mallat algorithm is finished within  $\log_2 N$  steps, see Fig.2[11].

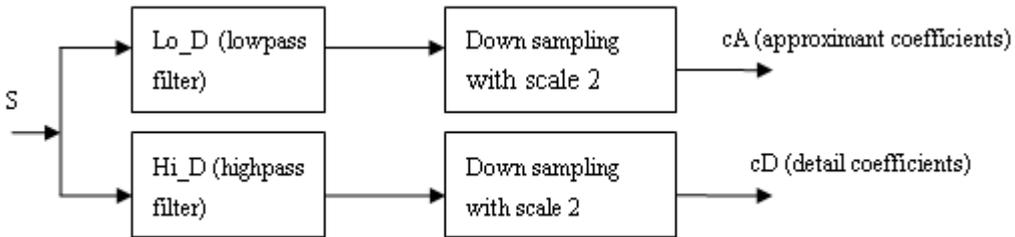


Fig. 2. One-step of the Mallat algorithm

If  $f$  is the sampling rate of signal  $S$ , then  $cA$  and  $cD$  represent  $S$ 's characteristic within band  $0$  to  $f/(2 \times 2)$ , and band  $f/(2 \times 2)$  to  $2/f$  respectively. Mallat is to repeat the step in Fig.2 until the satisfied resolution, which results multi-resolution of wavelet transform.

### 3.2. Fisher distance

Fisher distance is an efficient criterion of divisibility between two classes, which is broadly used in pattern recognition, and it computes the ratio of between-class scatter degree and within-class scatter degree between two classes. Larger ratio means larger divisibility of the two classes. So Fisher distance can be used as the criterion to choose optimal features. Fisher distance is defined as below[12]:

$$FisherDis = \frac{S_B}{S_W} \quad (3)$$

In (3),  $S_B$  and  $S_W$  represents between-class matrix and within-class matrix respectively. Assume we have  $n$  samples,  $x_1, \dots, x_n$ , divided into  $c$  classes, then between-class matrix and within-class matrix are defined as below:

$$S_B = \sum_{i=1}^c n_i (m_i - m)(m_i - m)^T \quad (4)$$

$$m = \frac{1}{n} \sum_{k=1}^n x_k \quad (5)$$

$$m_i = \frac{1}{n_i} \sum_{i=1}^{n_i} x_i \quad (6)$$

$$S_W = \sum_{i=1}^c S_i \quad (7)$$

$$S_i = \sum_{j=1}^{n_i} (x_j - m_i)(x_j - m_i)^T \quad (8)$$

Among (4) to (8),  $m$  means the average value of all the samples,  $m_i$ ,  $n_i$  and  $S_i$  represents the average value, sample number and within-class matrix of class  $i$  respectively.

## 4. Experiment Procedures and Results Analysis

Firstly, we extract the data from the 10 electrodes mentioned in section 1 from dataset target and non-target, and in order to compare with literature [9], we choose haar<sup>[9]</sup> as mother wavelet, and then decompose the EEG signals from each electrode from layer 3 to layer 6, and extract the approximate coefficients as features after decomposition. After that, we compute the Fisher distance of EEG signals with P300 and EEG signals without

P300, the results are showed in table 1. In table 1, “Total” column means the Fisher distance of the combination of the features from 10 electrodes mentioned in section 1.

Literature [9] used 3 layer wavelet transform, and extracted approximate coefficients as the features. From table 1, we can see that the feature divisibility increases obviously after the decomposition layer of wavelet is increased to 4 and 5.

Table 1. The Fisher distance of two kinds of EEG signals under different layers

<b>Eloc layer</b>	<b>Fz</b>	<b>Cz</b>	<b>Pz</b>	<b>Oz</b>	<b>C3</b>	<b>C4</b>	<b>P3</b>	<b>P4</b>	<b>PO7</b>	<b>PO8</b>	<b>Total</b>
3	0.0156	0.0171	0.0098	0.0025	0.0162	0.0134	0.0070	0.0037	0.0073	0.0096	0.0101
4	0.0185	0.0195	0.0114	0.0027	0.0200	0.0156	0.0081	0.0041	0.0083	0.0102	0.0115
5	0.0200	0.0204	0.0114	0.0029	0.0224	0.0159	0.0086	0.0041	0.0085	0.0114	0.0117
6	0.0081	0.0113	0.0091	0.0015	0.0124	0.0113	0.0080	0.0034	0.0031	0.0010	0.0066

The definition of wavelet transform in section 2 shows that the number of the extracted features will decrease as the decomposition layer increases, and the number of features is a key factor to the running time of classification system, so reducing the feature number will definitely speed up the classification system, and contribute to the online request of BCI.

The corresponding band of approximate coefficients from different wavelet decomposition layer is computed according to the definition of wavelet transform, see table 2.

Table 2. The corresponding frequency of approximate coefficients under different wavelet decomposition layer

<b>Wavelet decomposition layer</b>	<b>Band/Hz</b>
3	0-30
4	0-15
5	0-7.5
6	0-3.75

Some existing researches show that P300 appears between 0 Hz to 10 Hz, even lower band[13]. Table 1 and table 2 demonstrate that feature divisibility is largest between 0 Hz to 7.5 Hz as approximate coefficients being chosen to be the features, and feature divisibility declines when the feature band decreases, so the phenomenon of P300 is the most obvious in band 0-7.5Hz. The two tables also show that choosing the appropriate frequency band can improve the divisibility of the extracted features.

The common method to select EEG signal features is to put the candidates into classifier, and select the optimal features according to the classification accuracy. In this paper, we propose a feature selection algorithm based on wavelet transform and Fisher distance, which provides a systematic feature selection method before final classification. Fisher distance algorithm is simple and fast, while classification has two main parts, training and testing, which will take a long time, so the designing lifecycle of EEG signal processing algorithm will be dramatically decreased.

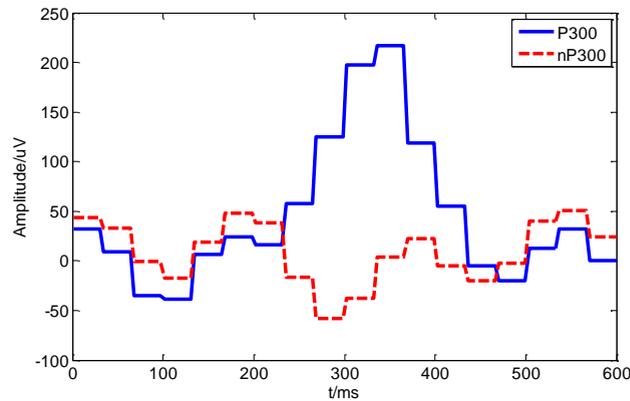
The last column from table 1 shows that feature divisibility isn't improved when combining features from 10 electrodes as one feature vector. Although some researches indicate that these 10 electrodes have larger influence to P300 detection, but P300 has low Signal Noise Ratio and it is random and non-stationary, which is different from one person to another [14]. So it is unreasonable to extract features only from these 10 electrodes for every subjects, features should be selected from different electrodes according to different subjects, and choosing electrodes can be taken as a kind of feature selection. Based on above analysis, a feature

selection method based on wavelet transform and Fisher distance is proposed, which can solve this issue to some extent:

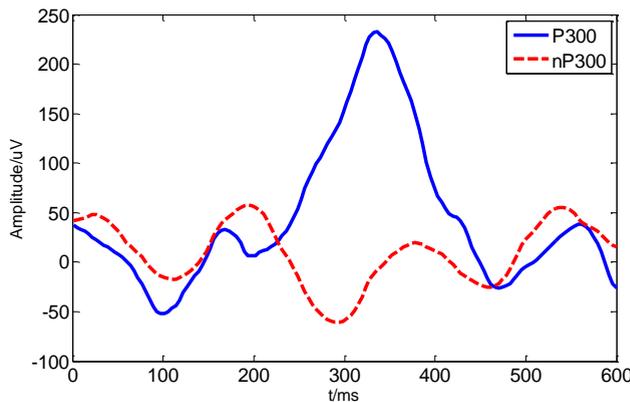
- (1) Signals from all electrodes are decomposed by wavelet, and then extract features from all electrodes and compute their Fisher distance, and select the features from one electrode which has the biggest Fisher distance as the initial feature vector;
- (2) Select features from one of the other electrode, and then combine them with the initial feature vector and compute the Fisher distance. If the new distance is smaller than the initial one, discard this new selected electrode; if the new distance is bigger than the initial one, take this new feature vector as initial feature vector, and turn to step (2) until all the electrodes have been selected. The final “initial feature vector” is the optimal features related to specific subjects. Different optimal features will be selected related to different subjects.

Using the proposed method to select features from the 10 electrodes, and the results show that feature divisibility reaches its maximum value when only C3 is selected. The maximum value is 0.0224, which is increased by 121.8% compared with 0.0101, obtained from [9].

In order to research on the influence of mother wavelet selection towards feature extraction with wavelet transform and compare the results with existing literatures, we select EEG signals from electrode Cz, which has a higher divisibility, from dataset target and non-target, and decompose the signals with different layer by wavelet transform and reconstruct them with their approximate coefficients, the results are showed in Fig.3.



(a)



(b)

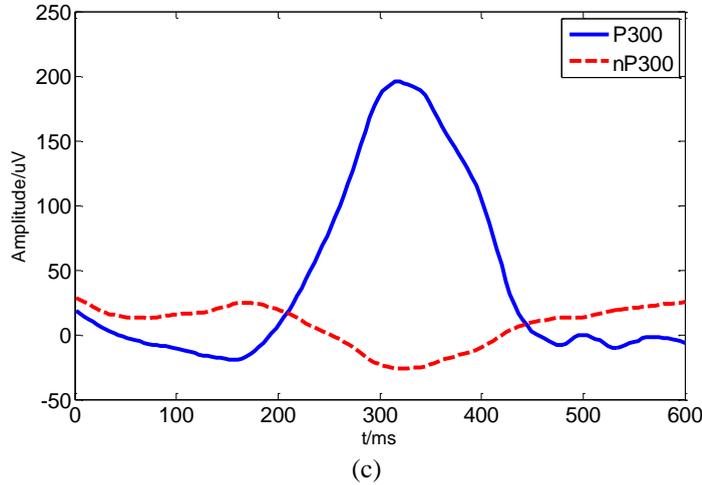


Fig. 3. The wavelet reconstructed signal.

- (a) Three layers wavelet reconstructed signal, haar;
- (b) Three layers wavelet reconstructed signal, db4;
- (c) Five layers wavelet reconstructed signal, db4.

In Fig.3(a), the EEG signals are reconstructed by the method in literature [9], in order to comparing, we select 3 layer and 5 layer reconstructed signals by Daubechies4[5] mother wavelet, see Fig.3(b)(c). The comparison of Fig.3(a) and Fig.3(b) demonstrate that Choosing haar as the mother wavelet is less efficient than choosing Daubechies4 as the mother wavelet. So the selection of mother wavelet is very important when deploy wavelet transform into the field of EEG feature extraction.

The comparison of Fig.3(b) and Fig.3(c) indicates that 5 layer wavelet decomposition has more obvious P300 phenomenon compared with 3 layer decomposition, and this result also verifies the conclusion that P300 phenomenon of band 0 -7.5Hz is the most obvious.

## 5. Conclusion

The P300 is the response of the brain to events in the external or internal environment of the organism, and it is drawing more and more attentions of the researchers because of its excellent characteristics, such as needless of training. But P300 has low Signal Noise Ratio and it is random and non-stationary, which is different from one person to another, so traditional signal processing algorithm cant not extract its features precisely. Wavelet transform is perfect for processing non-stationary random signals like EEG due to his multi-resolution, which is broadly used in the field of P300 detection. But the features extracted by wavelet transform have a low divisibility.

With the aim to solve the above issue in P300 detection, we research the P300 frequency domain of event related potentials, and then a new P300 feature extraction method based on wavelet transform and Fisher distance is proposed, which overcomes the drawbacks of no systematic feature selection method during traditional P300 feature extraction based on wavelet transform. Firstly, the original EEG signals are decomposed by wavelet transform, and extract wavelet coefficients from related band; secondly, Fisher distance of different electrode combinations is deployed to select optimal features; finally, the optimal features are sent to the classifier. We also research the influence of mother wavelet selection towards the divisibility of extracted features, and then we improve the P300 detection algorithm of literature [9] through proper electrode selection and wavelet decomposition layer. The experiment results show that the proposed method can increase

the divisibility by 121.8% of the features extracted by wavelet transform, and contribute to the followed classification.

## References

- [1] D.J. Krusienski, E.W. Sellers, T.M. Vaughan. Common Spatio-Temporal Patterns for the P300 Speller[A]. In: Proceedings of the 3rd international IEEE EMBS Conference on Neural Engineering[C]. Hawaii, USA. 2007. 421-424.
- [2] Alexander Lenhardt, Matthias Kaper, Helge J. Ritter. An Adaptive P300-Based Online Brain-Computer Interface[J]. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2008, 16(2): 121-130.
- [3] Kana Omori, Tomonari Yamaguchi, Katsuhiko Inoue. Feature Extraction from EEG Signals in P300 Spelling System[A]. In: ICROS-SICE International Joint Conference 2009[C]. Fukuoka International Congress Center, Japan. 2009. 849-852.
- [4] Mathew Salvaris, Francisco Sepulveda. Wavelet and Ensemble of FLDs for P300 Classification[A]. In: Proceedings of the 4rd international IEEE EMBS Conference on Neural Engineering[C]. Antalya, Turkey. 2009. 339-342.
- [5] YANG Bang Hua, YAN Guo Zheng, YAN Bo. Extracting EEG Feature in Brain-computer Interface Based on Discrete Wavelet Transform [J]. Chinese Journal of Biomedical Engineering, 2006, 25(5): 518-522(in Chinese)
- [6] WU Xiao Pei, YE Zhong Fu, GUO Xiao Jing, *et al.* Dynamic Independent Component Analysis and its Application to EEG of Motor Imagination[J]. Chinese Journal of Biomedical Engineering, 2007, 26(6): 818-824(in Chinese).
- [7] Chan F, Lam F, POoh P, *et al.* Detection of brainstem auditory evoked potential by adaptive filtering[J]. Med Biol Eng Comput, 1995, 33(2): 69-75.
- [8] Seyedehmina Ayoubian Markazi, Lampros.S.Stergioulas, Anusha Ramchurn, *et al.* Wavelet Filtering of the P300 Component in Event-Related Potentials[A]. In: Proceeding of the 28<sup>th</sup> IEEE EMBS Annual International Conference[C]. New York, USA. 2006. 1719-1722.
- [9] YANG Li Cai, LI Jin Liang, YAO Yu Cui, *et al.* P300 Detection Algorithm Based on Wavelet Decomposition and Support Vector Machines[J]. Chinese Journal of Biomedical Engineering, 2007,26(6): 804-809(in Chinese).
- [10] Hilit Serby, Elad Yom-Tov, Gideon F. Inbar. An Improved P300-Based Brain-Computer Interface[J]. IEEE Trans on Neural System and Rehabilitation Engineering, 2005,13(1): 89-98.
- [11] WANG Da Kui, PENG Jin Ye. Wavelet Analysis and Application in Signal Processing [M]. Beijing: Publishing House of Electronics Industry, 2006.1(in Chinese).
- [12] D. Huang, C. Xiang, S.S. Ge. Recursive Fisher Linear Discriminant for BCI Applications [A]. In: ISSNIP 2007 3rd International Conference [C]. Melbourne, Australia: 2007. 383-388.
- [13] Ozair Idris Khan, Sang-Hyuk Kim, Adil Khan, *et al.* Extraction of P300 Using Constrained Independent Component Analysis[A]. In: 31st Annual International Conference of the IEEE EMBS[C]. Minnesota, USA. 2009. 4031-4034.
- [14] Abeer E. Selim, Manal Abdel Wahed, Yasser M. Kadah. Machine Learning Methodologies in P300 Speller Brain-Computer Interface Systems[A]. In: 26rd National Radio Science Conference[C]. Future Univ, Egypt. 2009. 1-8.