Spectral Power Estimation for Unevenly Spaced Motor Imagery Data

Junhua Li[1,2], Zbigniew Struzik[2], Liqing Zhang[1], and Andrzej Cichocki[2]

[1] MOE-Microsoft Key Laboratory for Intelligent Information and Intelligent Systems, Shanghai Jiao Tong University, Shanghai, 200240, China
junhua.li@riken.jp, zhang-lq@cs.sjtu.edu.cn
[2] Laboratory for Advanced Brain Signal Processing, Brain Science Institute, RIKEN, Saitama, 351-0198, Japan
{zbigniew.struzik,a.cichocki}@riken.jp

Abstract. The human brain can send a command to external devices or communicate with the outside environment by the means of a brain computer interface (BCI) system. The effectiveness depends on how precisely specific brain activities can be identified from EEG. Noise is usually mixed into the EEG signal, and cannot be separated or filtered out in some cases. In a practical BCI system, the whole segment of EEG is discarded when a portion of that segment is contaminated by extreme noise or artifacts. This leads to the weakness that the BCI system cannot output decoding results during the period of that discarded segment. In order to solve this problem, we employed a Lomb-Scargle periodogram to estimate the spectral power based on an unevenly spaced segment, a portion of which has been removed due to noise contamination. According to the classification results of motor imagery data, the accuracy is not dramatically decreased along with increased proportion of data removal.

Keywords: Spectral Power Estimation, Brain Computer Interface, Motor Imagery, Unevenly Spaced Data, Classification.

1 Introduction

Brain computer interface (BCI) has attracted increasing attention of researchers coming from diverse research fields, and is one of interdisciplinary research hotspots. With a BCI system, healthy people can obtain fantastic manipulation experience contrary to their familiar perception [1], and disabled people can restore their abilities of communication [2] and degenerated motor function [3-4]. In the practical application of a BCI system, the intention of the user needs to be translated into a control command continuously in order to give the user an experience of smooth manipulation. This requires all EEG segments to be included for decoding. If some of EEG segments are discarded due to noise contamination, there is no output of commands during the periods of those discarded segments. Hence, it would be good to utilize the remaining portion of EEG segment after removing the portion of noise contamination. The power features are commonly used to distinguish different motor

M. Lee et al. (Eds.): ICONIP 2013, Part I, LNCS 8226, pp. 168–175, 2013.

imageries (e.g., left-hand and right-hand motor imageries) [5-7], because power features are robust in represent of information underlying motor imagery. For the complete EEG segment, Fourier transform can be used to transform temporal data points into spectral domain, but it is impossible to process unevenly spaced data like the EEG segment after removing a portion of noise contamination. In order to utilize the segments with unevenly spaced data points to let user feel smooth manipulation, we employed Lomb-Scargle periodogram to estimate the spectral power [8-9] and support vector machine (SVM) [10-11] to predict the class of motor imagery. Two categories of data are used to prove the feasibility of the method. One is the simulated data and the other is two-class motor imagery data. We used simulated data to illustrate that spectral power can be correctly estimated when data come to be unevenly spaced after removing some data points of them. And, we used real motor imagery data to demonstrate that classification accuracy does not dramatically decreased when different proportional portions of segments have been removed. Hence, the method of combination of Lomb-Scargle periodogram and SVM is suitable for using in the BCI system when a portion of segment has be removed.

2 Data Acquisition

The simulated data were generated by mixing two sinusoidal signals, which are 3Hz and 6Hz, respectively. The maximal amplitude of 3Hz sinusoid signal was 1.5 times of that of the 6Hz sinusoid signal. The motor imagery data came from three subjects. Fourteen electrodes were used to record the EEG signal on the sensorimotor cortex while the subject was conducting motor imagery at sampling rate of 250 Hz. Those electrodes were referenced at the mastoids behind ears and grounded at AFz. Each subject participated in four sessions. Each session consists of 15 trials, each of which was four-second length. Subject conducted either left hand motor imagery or right hand motor imagery according to the cue shown on the computer monitor.

3 Method

We first divided a four-second trial into 25 segments of one-second length with an overlap of 87.5%. A segment is denoted by X, which is N by T matrix. Where N is the number of channels, and T is the number of sampling points. Spectral power of each channel time series $y(t_i)$ is estimated by the Lomb-Scargle periodogram [8-9]. The estimated spectral power at frequency Ω_f can be obtained through minimizing the following sum of difference squares:

$$\min_{\substack{a>0 \\ \phi\in[0,2\pi]}} \sum_{i=1}^{T} (y(t_i) - \alpha\cos(\Omega_f t_i + \phi))^2 \ . \qquad (1)$$

Let

$$a = \alpha \cos \phi \tag{2}$$

and

$$b = -\alpha \sin \phi , \tag{3}$$

we can rewrite equation (1) as:

$$\min_{a,b} \sum_{i=1}^{T} (y(t_i) - a\cos(\Omega_f t_i) - b\sin(\Omega_f t_i))^2. \tag{4}$$

The optimal parameters $\hat{a}, \hat{b}$ can be obtained through minimizing equation (4)

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = R^{-1} r \tag{5}$$

where

$$R = \sum_{i=1}^{T} \begin{bmatrix} \cos(\Omega_f t_i) \\ \sin(\Omega_f t_i) \end{bmatrix} \begin{bmatrix} \cos(\Omega_f t_i) & \sin(\Omega_f t_i) \end{bmatrix} \tag{6}$$

and

$$r = \sum_{i=1}^{T} \begin{bmatrix} \cos(\Omega_f t_i) \\ \sin(\Omega_f t_i) \end{bmatrix} y(t_i) . \tag{7}$$

The power of specific frequency Ω_f is then estimated with respect to optimal parameters $\hat{a}, \hat{b}$ as follows:

$$\begin{aligned} &\frac{1}{T} \sum_{i=1}^{T} \left(\begin{bmatrix} \hat{a} & \hat{b} \end{bmatrix} \begin{bmatrix} \cos(\Omega_f t_i) \\ \sin(\Omega_f t_i) \end{bmatrix} \right)^2 \\ &= \frac{1}{T} \begin{bmatrix} \hat{a} & \hat{b} \end{bmatrix} R \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} \\ &= \frac{1}{T} r^{\mathrm{T}}(\Omega_f) R^{-1}(\Omega_f) r(\Omega_f) . \end{aligned} \tag{8}$$

Similarly, the minimization of squares mentioned above is used to estimate spectral powers at all frequencies. After that, spectral estimation for one channel is finished. Those steps are repeated for all channels and all segments to get the spectral powers. Because the frequency range of 8-30 Hz is mostly related to motor imagery task [7], we divided that band into four subbands with a bandwidth of 5 Hz (i.e., 8-12 Hz, 13-17 Hz, 18-22 Hz, and 23-27 Hz). Subband powers were obtained by averaging spectral powers within the corresponding frequency band range for each channel. Then, subband powers (four features for each channel) for all channels are concatenated into a feature vector:

$$F = [f_{11}, f_{12}, f_{13}, f_{14}, f_{21}, f_{22}, f_{23}, f_{24}, \cdots, f_{N1}, f_{N2}, f_{N3}, f_{N4}]^{\mathrm{T}}, \tag{9}$$

where N is the number of channels. Subsequently, features are normalized as:

$$f_{qp} = \log\left(\frac{f_{qp}}{\sum_{i=1}^{N} \sum_{j=1}^{4} f_{ij}} \right). \tag{10}$$

The normalized features were fed into a linear SVM classifier to distinguish which class it belongs to.

4 Results

4.1 Simulated Data

Figure 1 shows the spectral power estimation from a mixed signal, which mingles two sinusoidal signals with 3 Hz and 6 Hz respectively. From top left to bottom right, spectral power estimations for the complete signal, proportional data point removals from 10% to 80% are shown, respectively. The data points removed are chosen randomly. The powers shown in figure 1 were normalized by dividing by a proportional factor (1-p, p is the removed percentage) in order to keep the same scale between cases of different proportional data removal. For example, the estimated power is divided by the proportional factor of 0.7 when 30% of data points are removed from the signal. From figure 1, we can see the components at 3 Hz and 6 Hz can be better estimated even up to 80% of data point removal.

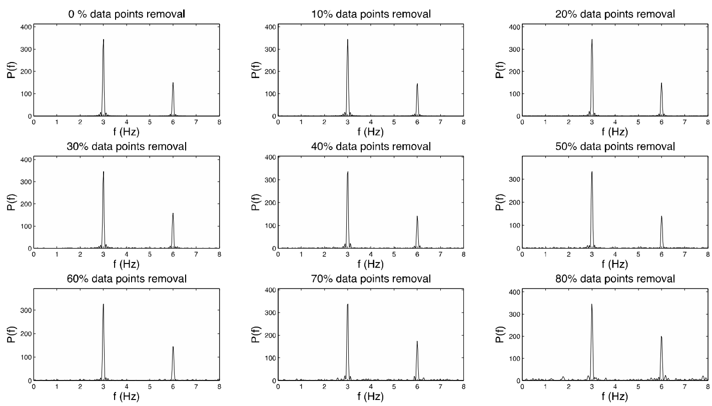

Fig. 1. Spectral power estimations for the complete signal and signals after data point removal

4.2 Real Motor Imagery Data

In this section, we showed results tested on real motor imagery data. The proposed method can solve the problem that the whole segment has to be discarded due to partial noise contamination on that segment, if the classification accuracy for segments with data removal can remain the same or slightly decrease. Here, we used two ways to randomly remove data points. One is that data points are randomly removed (see figure 2 for an example). The other is data blocks are randomly removed (see figure 3 for an example). The width of removed blocks is generated according to a normal distribution with a mean of 20 and standard deviation of 10.

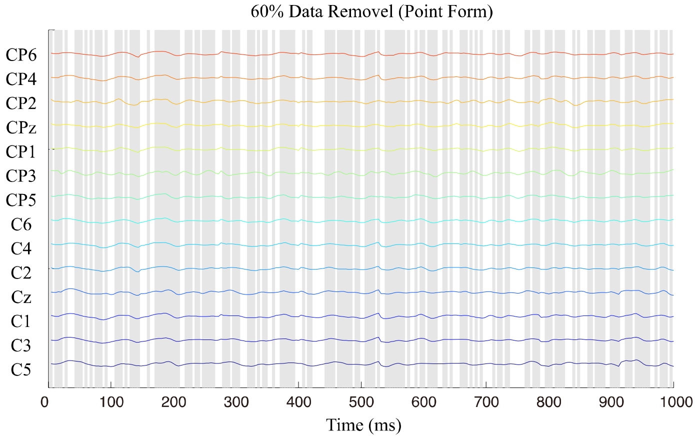

Fig. 2. An example of data point removal. The data points shown with gray background are removed while data points shown with white background are retained.

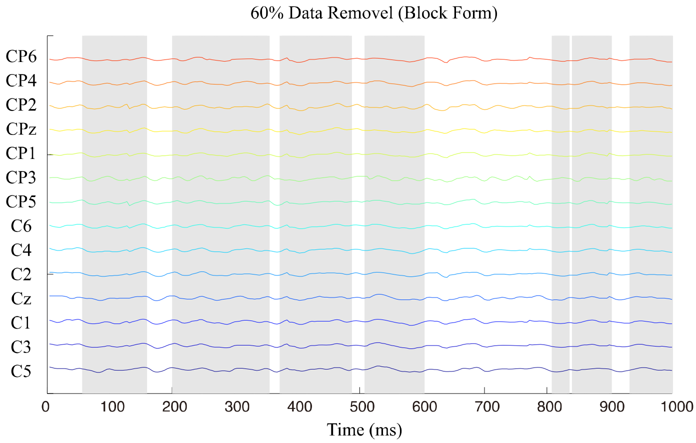

Fig. 3. An example of block point removal. The data points shown with gray background are removed while data points shown with white background are retained.

The data from the preceding session were used for training and the data from the following session were used for testing. Sliding time window accuracies were calculated through the number of correct classification segments divided by the number of all segments. A trial was classified to the class that most of sliding time windows within that trial belonged to. Then, trial accuracies were obtained by the ratio of correct classification trials. Figure 4 and figure 5 show testing accuracies for the conditions of data point removal and data block removal, respectively. In general, the accuracies for all sessions of all subjects are not dramatically decreased. Trial accuracies varied more than that of sliding time window across different proportional portions of data removal. The reason is that trial was counted as correct classification trial even if the number of correct classification sliding time windows is one more than that of wrong classification sliding time windows, and vice versa. Therefore, in some cases, trial accuracy changed greatly while sliding time widow accuracy did not change too much. A comparable classification accuracy can be achieved even when 80% of data were removed. The high accuracies can be kept no matter how many data points were removed from 10% to 80% for subject 1, especially for session 2 and session 3. The accuracies for 80% data removal are largely worse than that for 70% data removal for subject 1 in the condition of block data removal. It seems that our method is relatively sensitive to the form of block data removal.

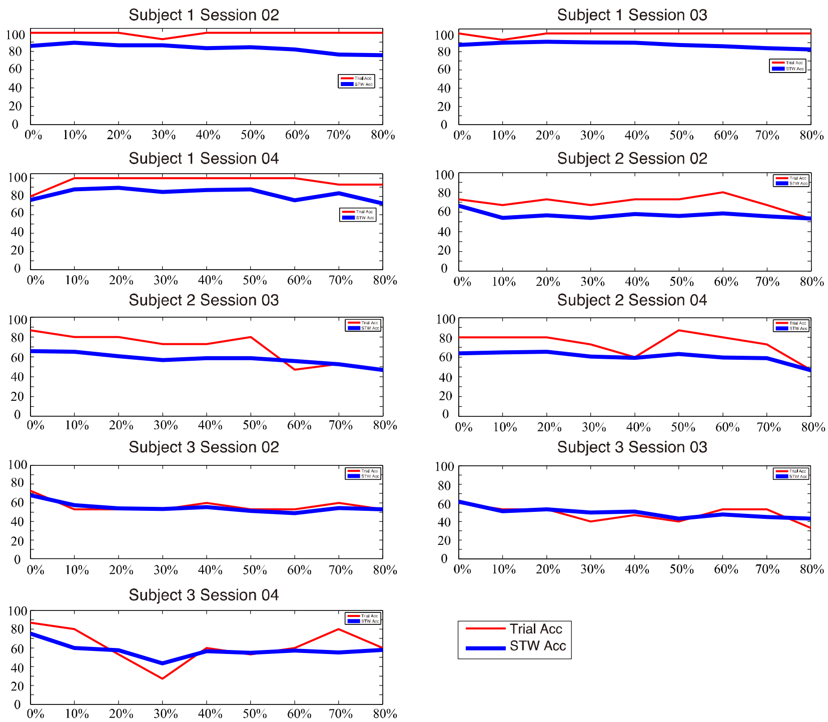

Fig. 4. Classification accuracies for the form of data point removal. The thin red lines represent trial accuracies, and the bold blue lines represent sliding time window accuracies.

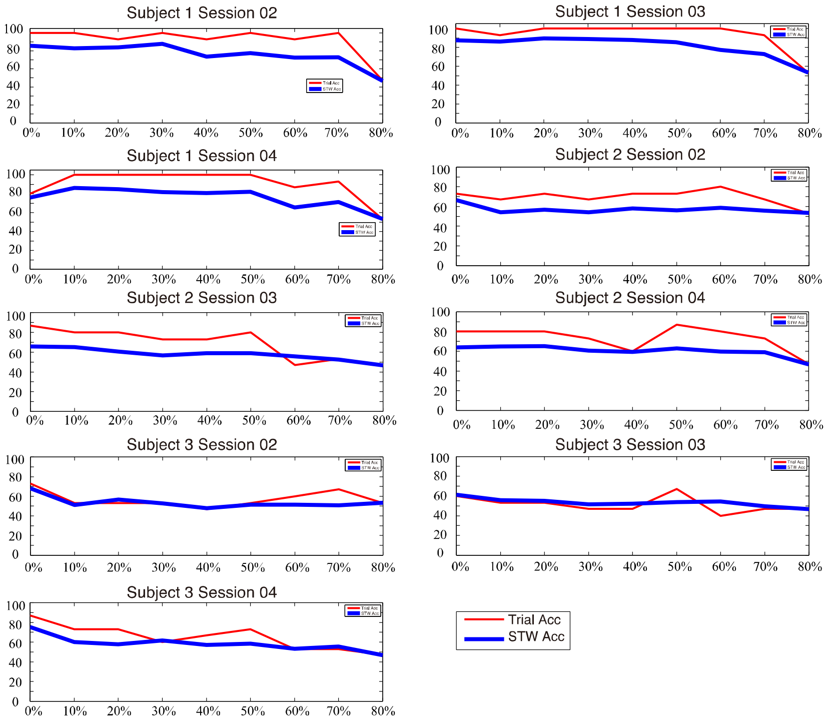

Fig. 5. Classification accuracies for the form of block point removal. The thin red lines represent trial accuracies, and the bold blue lines represent sliding time window accuracies.

5 Conclusion

We proposed the combination of Lomb-Scargle periodogram and SVM classifier to distinguish the EEG segment with a portion of data removal due to noise contamination. The results indicated that classification accuracy was not dramatically decreased when different percentages of data were removed. Therefore, the classification performance using the proposed method for segments with data removal is acceptable for a BCI application system. This means that the segment with noise contamination can still be utilized to output commands after only removing the noisy portion, rather than discarding the whole segment, which is conventionally taken by the BCI system. In brief, the proposed method can achieve comparable classification performance even when most of data points of a segment have been removed. It avoids the problem that there is no output of commands when a segment is discarded, because Fourier transform cannot be used to estimate spectral power after a portion of data has been removed due to noise contamination.

Acknowledgments. The work was in part supported by the National Natural Science Foundation of China (Grant No. 61272251, 91120305), and in part supported by the NSFC-JSPS International Cooperation Program (Grant No. 61111140019).

References

1. Li, J., Liu, Y., Lu, Z., Zhang, L.: A Competitive Brain Computer Interface: Multi-person CarRacing System. In: 35th Annual International Conference of the IEEE EMBS, Osaka, Japan, July 3-7, pp. 2200–2203 (2013)
2. Muller, K.R., Tangermann, M., Dornhege, G., Krauledat, M., Curio, G., Blankertz, B.: Machine Learning for Real-time Single Trial EEG Analysis: from Brain-computer Interfacing to Mental State Monitoring. J. Neurosci. Methods. 167(1), 82–90 (2008)
3. Li, J., Liang, J., Zhao, Q., Li, J., Hong, K., Zhang, L.: Design of Assistive Wheelchair System Directly Steered by Human Thoughts. Int. J. Neural Syst. 23(3), 1350013 (2013)
4. Pfurtscheller, G., Muller, G.R., Pfurtscheller, J., Gerner, H.J., Rupp, R.: 'Thought'-Control of Functional Electrical Stimulation to Restore Hand Grasp in A Patient withTetraplegia. Neurosci. Lett. 351(1), 33–36 (2003)
5. Pfurtscheller, G., Guger, C., Muller, G., Krausz, G., Neuper, C.: Brain Oscillations Control Hand Orthosis in A Tetraplegic. Neurosci. Lett. 292(3), 211–214 (2000)
6. Palaniappan, R.: Utilizing Gamma Band to Improve Mental Task Based Brain-computer Interface Design. IEEE Trans. Neural Syst. Rehabil. Eng. 14(3), 299–303 (2006)
7. Li, J., Zhang, L.: Active Training Paradigm for Motor Imagery BCI. Exp. Brain Res. 219(2), 245–254 (2012)
8. Lomb, N.R.: Least-squares Frequency Analysis of Unequally Spaced Data. Astrophys. Space Sci. 39(2), 447–462 (1976)
9. Stoica, P., Li, J., He, H.: Spectral Analysis of NonuniformlySampled Data: ANew Approach Versus the Periodogram. IEEE Trans. Signal Process. 57(3), 843–858 (2009)
10. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, New York (1995)
11. Hearst, M.A., Dumais, S.T., Osman, E., Platt, J., Scholkopf, B.: Support Vector Machines. IEEE Intell. Syst. 13(4), 18–28 (1998)