Journal of Neuroscience Methods 193 (2010) 373–379

Contents lists available at ScienceDirect

Journal of Neuroscience Methods

journal homepage: www.elsevier.com/locate/jneumeth

Bilateral adaptation and neurofeedback for brain computer interface system

Junhua Li*, Liqing Zhang

MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems, Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

ARTICLE INFO

Article history:
Received 27 March 2010
Received in revised form
20 September 2010
Accepted 21 September 2010

Keywords:
Bilateral adaptation
Neurofeedback
Brain computer interface

ABSTRACT

Brain computer interface (BCI) provides an alternative communication pathway between human brain and external devices without the participation of peripheral nerves and muscles. Although the BCI techniques have been developing quickly in recent decades, there still exist a number of unsolved problems, such as instability, unreliability and low transmission rate in real time applications of BCI. In the present study, we design a bilateral training framework for both human and the BCI system to improve recognition accuracy and to reduce the impact caused by non-stationary EEG signal. The statistical analysis is used to test whether there is an obvious improvement in recognition performance after using the bilateral adaptation strategy. The statistical analysis indicates that our algorithm is significantly different from the existing method in both conditions of trials ($p = 0.0073$) and sliding time windows ($p = 0.00077$). The results of statistical analysis reconfirm that performance using our algorithm is distinctly improved. The online experiments also demonstrate that the proposed algorithm achieves higher prediction accuracy and reliability compared with the existing method. The objective of our research is to transfer this strategy to some practical applications (e.g., electrical wheelchair control) for the better performance.

1. Introduction

The cutting-edge research fields between bioinformatics and computer science have been developing dramatically recently. Brain computer interface (BCI) is now becoming one of hot research topics due to the following three reasons. First, it provides a new approach to understand neurophysiologic mechanism of how brain is executing specific task (Pfurtscheller and Neuper, 1997). Second, BCI is of practical significance in real applications. The technique can be used for helping people with severe motor disabilities (Müller-Putz et al., 2005; McFarland et al., 1997; Kübler et al., 1999) and applied to clinical rehabilitation of motor functions (Duffau, 2006; Daly and Wolpaw, 2008; Cooper et al., 2008). Third, it also provides a possibility to combine brain intelligence and machine (e.g., computer) intelligence, and people could enhance their ability of manipulating objects in the world by combining themselves' intelligence and machine's intelligence.

Generally speaking, there are a number of ways of measuring brain activity, such as electrical signal and blood oxygen level dependent (BOLD) signal. Near-Infrared Reflectance Spectroscopy (NIRS) (Xu et al., 2007) and Functional Magnetic Resonance Imaging (fMRI) (Ciuciu et al., 2008) are based on the measure of BOLD signal. Electrical signal is acquired from electrodes mounted on the surface of scalp or implanted into the tissue of brain, namely, non-invasive or invasive manners. For the invasive manner, electrodes should be implanted into the brain or laid on the cortex surface of the brain. For example, Ikeda and Shibbasaki (1992) recorded movement-related cortical potentials through an invasive system. Although such an invasive system has the higher spatial resolution, it is easy to cause the damage of users' brain tissue. In this paper, we rather concentrate on noninvasive BCI based on Electroencephalogram (EEG) measurements (Keirn and Aunon, 1990).

EEG signal has an advantage in temporal resolution compared with BOLD signal. But it also suffers a number of drawbacks for BCI based on motor imagery such as temporal variation of EEG. A model trained several days ago (even a few hours ago) is no longer suitable to classify the current EEG data due to EEG variation over time. Therefore, a practical BCI system should be continually adapted to track user's current EEG pattern in order to achieve a good performance. A number of adaptation strategies have been proposed for adaptive BCI systems in recent years. Usually adaptive technique can be applied to feature extractor, classifier, or both of them. Parameters in these models, such as mean and variance, are continually updated for the purpose of tracking the EEG state of the subject. Vidaurre et al. (2006) used an adaptive autoregressive model for extracting features and an adaptive Quadratic Discriminant Analysis (QDA) for pattern classification. In addition,

* Corresponding author at: MOE-Microsoft Key Laboratory for Intelligent Computing and Intelligent Systems, Department of Computer Science and Engineering, Room 3-508, SEIEE Buildings, Shanghai Jiao Tong University, Shanghai 200240, China. Tel.: +86 02154748390.

E-mail addresses: juhalee@sjtu.edu.cn (J. Li), zhang-lq@cs.sjtu.edu.cn (L. Zhang).

0165-0270/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2010.09.010

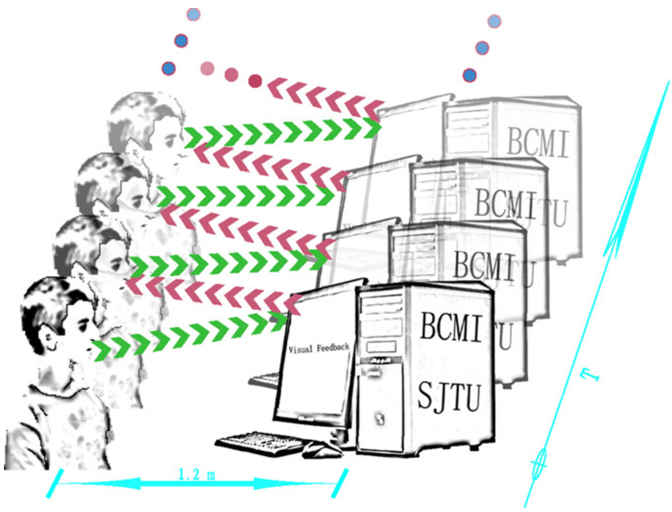

Fig. 1. Sketch map of bilateral training process. Green arrow represents BCI system training. Red arrow represents humans training. Dots mean to continually repeat training over the whole experiment time. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

the adaptive Linear Discriminant Analysis (LDA) and Bayesian classifier are usually adopted as adaptation of classifier (Shenoy et al., 2006; Sykacek et al., 2004). It should be highlighted that the literatures mentioned above just used single feature extractor or classifier, thus any mistakes caused by feature extractor or classifier cannot be corrected.

In order to improve the robustness of BCI system, we suggest that EEG data is projected into several different subspaces and is classified by several different hyperplanes (classifiers) in our bilateral training framework. The voting strategy is employed to improve accuracy of BCI system. First, each hyperplane separating EEG data is trained as a voting machine, then these voting machines are constructed a strong classifier by using committee voting rule so as to achieve the better classification performance. Previous studies mainly emphasize on improving BCI system by using adaptation of BCI models while ignoring human neuro-feedback adaptive adjustment, which is considered as another dimension to improve performance of BCI. In our bilateral training framework, we take human factors into consideration during training a practical BCI system. Subjects are requested to adjust their brain activities according to neuro-feedback reflecting ongoing neural activities and to try their best to conduct correct tasks of motor imagery. We call this course as human training. Fig. 1 shows the process of the bilateral training framework. BCI model training and human training are implemented alternately. Human and BCI system are both trained in order to adapt to each other and gradually attain a dynamic equilibrium. Hence, the reliability of system is improved due to the mutual adaptation of human subject and BCI model.

2. Bilateral training framework

Bilateral training framework is divided into two parts: BCI system training and human training. The objective of BCI system training is to train a BCI model by using EEG data recorded during subject training so as to achieve a good generalization performance. The subject modulates his/her own brain activities to make BCI system classify his/her EEG signals more correctly based on neuro-feedback principle during human training. Classification accuracy will be gradually improved by interaction between subject modulation and BCI system adaptation.

2.1. BCI system training

The BCI system training consists of the following three steps: EEG signal preprocessing, feature extraction and pattern learning. In the signal preprocessing stage, we mainly dealt with the artifact removal, baseline drift and so forth. Artifacts, such as EOG and EMG, should be removed from EEG recordings. At the same time, a visual feedback was given to subject as an indication of artifacts. Subject should try to keep any portion of body motionless when the indication of artifacts was displayed on the screen. The EEG signals after removing artifacts were further processed by using a band-pass filter with bandwidth between 8 and 30 Hz. At the feature extraction stage, we employed common spatial patterns (CSP) (Anthony and Zoltan, 1995; Ramoser et al., 2000) to extract features from EEG signals. Those features obtained based on the optimal component separation property of CSP are optimal for discriminating two populations of EEG.

A segment of EEG signal is represented as an N by T matrix E, where N is the number of recording electrodes and T is the number of samples per electrode in a segment. A segment E is projected into space of common special patterns as orthogonal components by the projection matrix V. So a segment of EEG is decomposed as

$$Z = VE. \tag{1}$$

We will obtain $2m$ time series if m largest eigenvectors of each group are chosen. Then the $2m$ features are calculated for a segment (EEG data) by following equation:

$$feature_i = \sum_{t=1}^{T} (z_i(t))^2 \quad i = 1, 2, \ldots, 2m, \tag{2}$$

where z_i represents row of Z, t is sampling time of a segment. In order to normalize the distribution of elements, features are reevaluated by the following equation:

$$feature_i = \log\left(\frac{feature_i}{\sum_{j=1}^{2m} feature_j}\right). \tag{3}$$

Those features are used to train a support vector machine (SVM) classifier (Vapnik, 1995; Hearst et al., 1998).

2.2. Human training

It usually is not sufficient to improve performance of BCI system without considering the quality of signal source. It is because some signals measured from a few subjects are mingled with each other and is unclassifiable essentially, even the best classified algorithm used in BCI system cannot classify the signals with a satisfactory performance. Hence, a signal source of high quality is critical for the BCI system's performance. For the purpose of getting the signals with high quality, a subject has to generate EEG patterns which are easily recognized by the BCI system. It becomes easier for a human subject to enhance own brain activities corresponding to the cues if a neuro-feedback is given. Under the circumstances, the subject can immediately adjust own brain activities according to visual feedback on the computer screen. To this end, in our training framework, we design two rectangular bars on the screen as feedback to indicate which task the subject is imagining now. During the experiments, there are two rectangular bars displayed on the screen, one is near the left side of the screen and the other is near the right side. Left bar is increased while subject is imagining left movement. The opposite bar is increased while subject is imagining right movement (Fig. 2). The subject should strive to maintain the imaginary movement of given direction (direction of the arrow, see Fig. 2) and make corresponding bar increase.

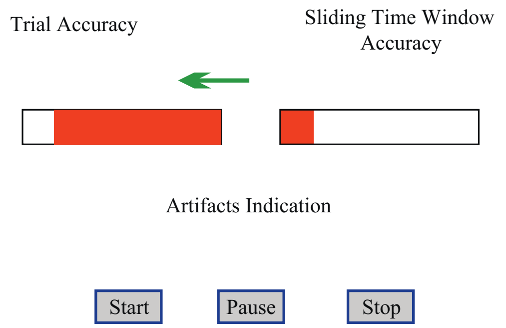

Fig. 2. The feedback displayed on the screen. Accuracy is displayed on top of the screen. The rectangular bars show neurofeedback in real time. Artifacts indication is given to subject when artifacts are found by the BCI system.

2.3. *Interaction between human training and BCI system training*

In bilateral training framework, EEG signal adjustment and BCI system updating are executed alternately. With mutual adjustment goes on, the classification accuracy will be raised gradually. During the experiments, subject tries to make BCI system improve the classification accuracy by modulating his/her brain activities. At the same time, the BCI system is also updated across sessions. We train the first model using EEG data of the first session. The second model is trained by not only EEG data from the second session but also the first session (see Fig. 3). EEG data used for training is chosen randomly from each of last five sessions according to a specific proportion. We set the highest proportion for the latest session. The proportion is declined by degrees from the latest session to the previous sessions up to the fifth prior session. The proportion of training data from each session is defined as

$$P_{sn} = \frac{I_n}{\sum_{i=1}^{n_{\max}} I_i} \quad \left(n \in \left\{1, 2, 3, 4, 5\right\}\right), \tag{4}$$

where $I = [I_1\ I_2\ I_3\ I_4\ I_5] = [1\ 3\ 5\ 7\ 9]$, $n_{\max}$ is the maximum value in the parameter n. n equals the number of sessions when the number of sessions is not more than five, and the value of n will always be the same five when the session started from the sixth. The proportion of training data from the latest session decrease from the first session to the fifth session respectively 1, 3/4, 5/9, 7/16, 9/25. But the proportion of the latest session is higher than that of all previous sessions. The reason is that the importance of the latest session reduces with sessions go, but the data from the latest session is still the most important.

The output of each model is weighted to produce the result of classification for EEG of every sliding time window. The classification result of a trial is obtained by averaging within all probabilistic outputs of sliding time windows. The weight of each model a_i is calculated by

$$a_i = \frac{a_i}{\sum_{j=1}^{5} a_j}, \tag{5}$$

$$a_j = -\frac{1}{2}\log\left(\frac{1-TA}{TA}\right), \tag{6}$$

where TA is the training accuracy corresponding to each model. A model which has the higher training accuracy is the more reliable. Hence, we give it a higher weight. The classification result of each sliding time window is obtained by weighting all outputs of models.

$$Pr\ o = \sum_{i=1}^{5} a_i \times o_i, \tag{7}$$

where o is probabilistic output of each model corresponding to given class. a_i is obtained by Eq. (5). The EEG data of a trial is stored if J is greater than threshold given (we set threshold as 0.55, except subject 4 for 0.75). J is defined as follows:

$$J = \frac{1}{M}\sum_{k=1}^{M} Pr\ o_k, \tag{8}$$

where M is the number of sliding time windows.

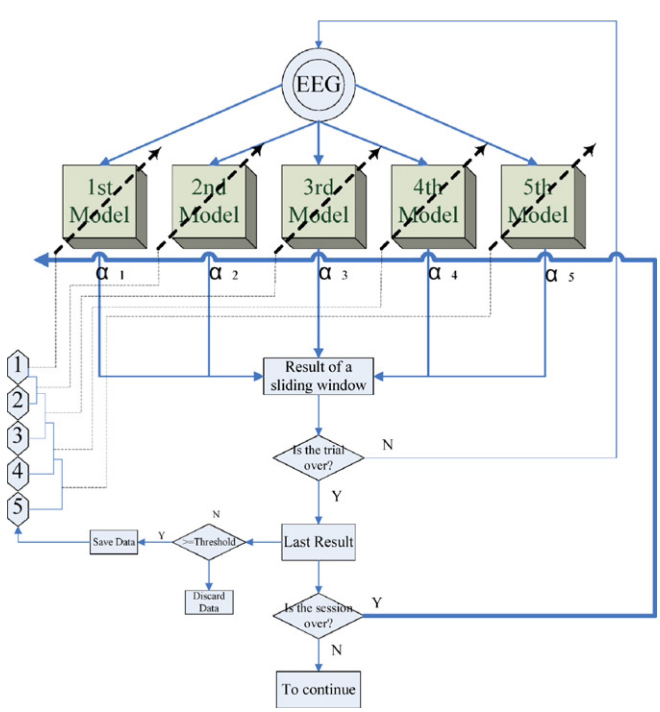

Fig. 3. The flow chart of interaction between humans training and BCI system training. A segment EEG is classified respectively by each model. The result is given by weighted outputs of each model. The EEG data of a trial is saved if J (see Eq. (8)) is greater than threshold given, otherwise discarded. The most previous model is replaced by latest model when the number of models is more than five. According to this rule, models are constantly updated.

3. Experimental setup

3.1. *Subjects*

The subjects participated in this study are six adults without any sensory-motor diseases or history of psychological disorders. All subjects had not attended related BCI experiments previously and were given introduction before experiments. Their age ranges from 23 to 26 years old and two of them are female. All subjects have given their written informed consent for the study.

3.2. *EEG signals recording*

We recorded EEG signals from fourteen channels with a digital DC EEG amplifier (Neuro Scan System). All electrodes were mounted in a standard EEG cap according to the 10-20 international system. The distribution of electrodes used in our experiments is shown in Fig. 4. The EEG recordings were referenced to the blue electrodes and grounded at the electrode of GND. Fourteen electrodes used for recording were placed on the region related to sensorimotor cortex. Electrodes' impedances were kept below 5 kΩ. All electrodes were digitized at 250 Hz.

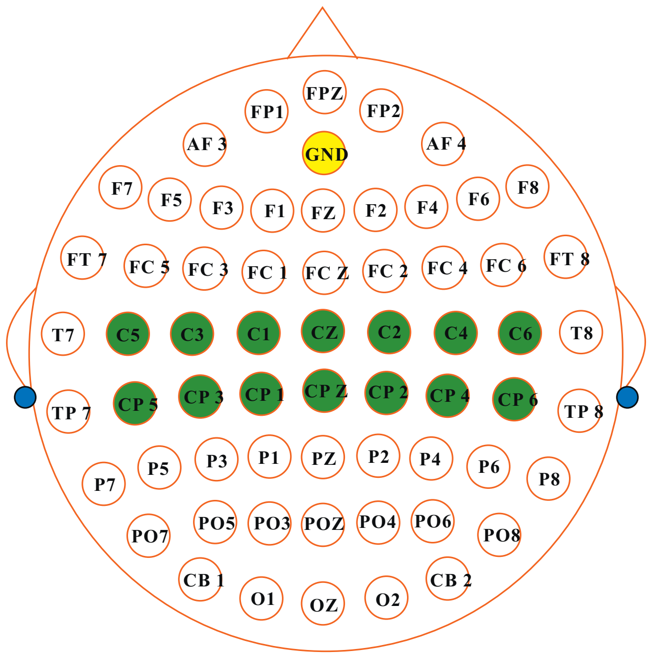

Fig. 4. The distribution of electrodes using in experiments. Green electrodes were used for collecting EEG data. Blue electrodes were as reference. Yellow electrode is ground. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

3.3. Experimental paradigm

The subjects sat still in a comfortable chair facing a 22-in. wide-screen monitor 1.2 m away (see Fig. 1), and were asked to remain motionless during sessions. Each subject participated in two experimental paradigms. One is a normal training paradigm corresponding to the existing method. Subjects took part in different number of sessions ranged from seven to ten (see Table 1 for each subject in detail). There were 2 min for rest between successive sessions. Each session consisted of twelve trials separated by intervals of 2 s. The word 'Attention' was displayed on the screen for 3 s at the beginning of each session. A trial was 4 s long and consisted of 25 sliding time windows. The width of sliding time window was 1000 ms. The time window slid forward every 125 ms till the end of a trial. After a session, a trial was divided into segments of 1 s length with an overlap of 87.5%, resulting in 25 segments per trial. So a total of 300 samples were used for training in the normal training paradigm.

With regard to bilateral training paradigm, the configuration of trials for subject adaptation is the same as normal training paradigm. The difference is combination between human training and BCI system training as described in the section bilateral training framework. In bilateral training paradigm, subjects completed a adaptive procedure in each session after 12 pure trials EEG data were successfully collected, where the 'pure trial' means the mean of probability corresponding to correct class (J, see Eq. (8)) is greater than threshold given (see Table 1). The session also stopped when the amount of trials reached to maximal number (we set it as 24). Hence, each session would stop within brown block, as shown in Fig. 5. The numbers of trials consisted of a session varied among different sessions of each subject. It is noticed that the number of pure trials collected in bilateral training paradigm in one session is not larger than 12. And the total number of pure trails chosen from last all five sessions also is not larger than 12 due to the proportion normalization (see Eq. (4)). Hence, the number of samples (EEG segments) for training is not larger than 300 after dividing trials.

4. Results

4.1. On-line performance comparison

We tested two paradigms online to assess performance. In the aspect of accuracy, the bilateral training paradigm (BTP) outperformed the normal training paradigm (NTP). The average classification accuracy in trials across all sessions of all subjects was 86% and 76.99% for bilateral training paradigm and normal training paradigm respectively. We carried out performance comparison from the sixth session of bilateral training paradigm because five models used in bilateral training paradigm had already been trained at that point. For some subjects, the number of sessions in which they participated is different in two different paradigms. So we compared performance from the sixth session to the last session of one of the two paradigms. From Fig. 6, we can see that all subjects achieved better performance when bilateral training paradigm was employed in experiment. In particular, the average accuracy in trials using BTP was significantly higher than that of NTP for subject 2. The recognition accuracy would be improved when averaging probabilistic outputs of all sliding time windows during a trial (which is equivalent to voting). We, furthermore, compared BTP with NTP in terms of the accuracy of sliding time windows so as to eliminate the effect of that mentioned above. The overall mean of recognition accuracy in sliding time windows was 79.2% and 70.29% for BTP and NTP respectively. The results showed that improvements on recognition rates can be again obtained by using BTP for all subjects. Table 1 shows detailed information to each subject with different paradigms.

For the variation of classification accuracy across sessions, BTP was smoother than NTP. As seen from Fig. 6, the variation of classification accuracy between sessions was larger corresponding to NTP. In other words, the curve of NTP was more fluctuant than that of BTP. This phenomenon was more obvious for subject 4. Therefore, BTP has a better reliability for EEG signal classification.

4.2. Statistical analysis

The statistical analysis was employed to test whether there was an obvious improvement in recognition performance after using the bilateral adaptation strategy. A paired t-test was used to compare the overall performance of BTP with NTP across all subjects. Results of the paired t-test indicated that recognition accuracy of BTP was significantly higher than recognition accuracy of NTP in both conditions of trials ($p = 0.0073$) and sliding time windows ($p = 0.00077$). Furthermore, we verified the performances across sessions for each subject and listed the results in Table 2. The results showed recognition accuracy using BTP was significantly improved for most of subjects (except for subject 3).

4.3. Off-line analysis

We performed an off-line analysis to investigate which of the aspects of algorithm used in the BTP were actually effective for performance improvement. The data used for off-line analysis was collected from the NTP of subject 4. In the first analysis, we addressed whether the 'pure' trial selection was necessary. To this end, we evaluated performance of models which were trained in 'pure' trials in comparison to all trials of a session. The second analysis was to explore whether the method of training data randomly chosen from previous sessions according to a specific proportion was useful for performance. We compared the two conditions. One is training data proportionally and randomly come from previous five sessions. The other is training data come from the latest session. For instance, the data proportionally and randomly chosen from the second session to the sixth session was used for

Table 1
The parameter setting and recognition accuracy for each subject. The thresholds are listed in the second column. NS is the number of sessions in normal training paradigm. AT represents average recognition accuracy in trials across all sessions by percentage. AW represents average recognition accuracy in sliding time windows across all sessions by percentage. *B* represents bilateral training paradigm. *N* represents normal training paradigm. The overall mean averaging each column is shown in the last row.

Subject	*T*	NS	AT(*B*)	AT(*N*)	AW(*B*)	AW(*N*)
1	0.55	10	89	79.1	82.31	70.83
2	0.55	10	93.4	76.7	79.86	67.8
3	0.55	7	70.73	65.57	62.77	57.99
4	0.75	10	97	95.1	94.2	87.22
5	0.55	8	97	86.38	92.69	81.62
6	0.55	10	68.9	59.1	63.38	56.26
Overall mean			86	76.99	79.2	70.29

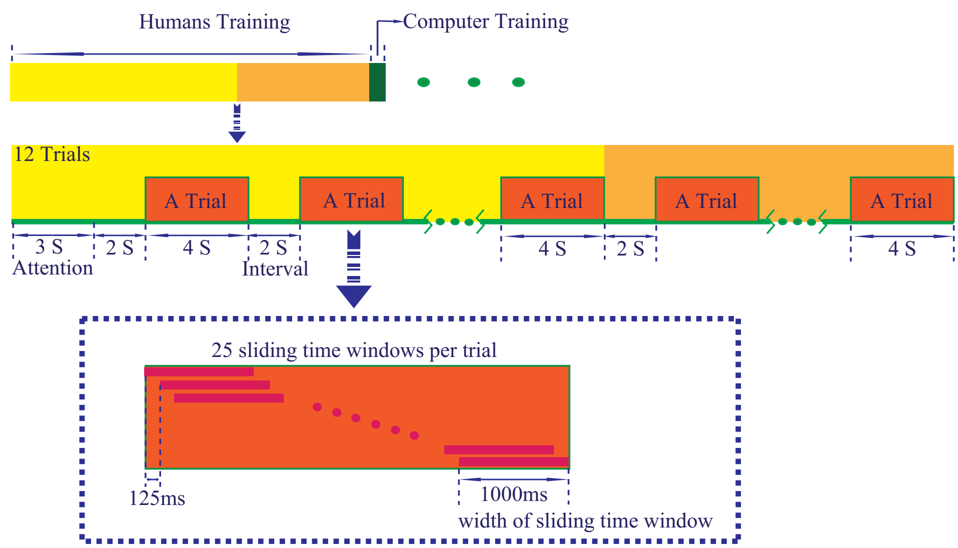

Fig. 5. The time arrangement of bilateral training. The word "Attention" was displayed on the screen for 3 s at the beginning of each session. The interval between trials was 2 s. A trial was 4 s long and consisted of 25 sliding time windows. Structure of a trial was enlarged inside the dashed border.

Table 2
The significant differences for recognition accuracy in condition of sliding time windows. The *p* is the paired *t*-test used in condition of sliding time windows.

Subject	1	2	3	4	5	6
p	0.018	0.0001	0.27	0.048	0.029	0.039

training model A. The total number of chosen data used for training model A is equivalent to the number of a session. Only the data of the sixth session was used for training model B. Then, we compared the recognition accuracy of model A with the model B. The rest can be done in the same manner. In the third analysis, we compared multi-models with single model. The data of one session was used for training a model and five models combined to form the multi-models. The single model was trained using all data of five sessions. The results of off-line analysis are listed in Table 3. The comparison indicated that 'pure' trial selection hardly made a contribution to performance improvement. The recognition accuracy was almost same after utilizing the strategy of 'pure' trial selection. However, training data proportionally and randomly chosen from previous five sessions was useful for performance improvement. We found it's dominant in all sessions (from session 7 to session 10). With regard to multi-models, off-line analysis showed the multi-models, overall, was better than the single model.

Table 3
The off-line analysis results. PTS represents recognition accuracy when 'pure' trial selection is used. Contrarily, NPTS represents recognition accuracy without 'pure' trial selection. *W* refers to recognition accuracy of model which is trained using the data proportionally and randomly chosen from previous five sessions. NW refers to recognition accuracy of model trained in the latest session data. MM denotes recognition accuracy of multi-models. SM denotes recognition accuracy of single model trained in all data of five sessions.

Session	PTS	NPTS	*W*	NW	MM	SM
3	98.26	97.57				
4	93.75	93.75				
5	94.44	95.14				
6	86.81	86.81				
7	89.58	86.46	93.06	86.46	96.53	95.49
8	77.08	81.25	87.5	81.25	85.42	88.19
9	97.92	90.63	95.14	90.63	99.31	96.88
10	96.18	95.83	96.53	95.83	97.92	97.22

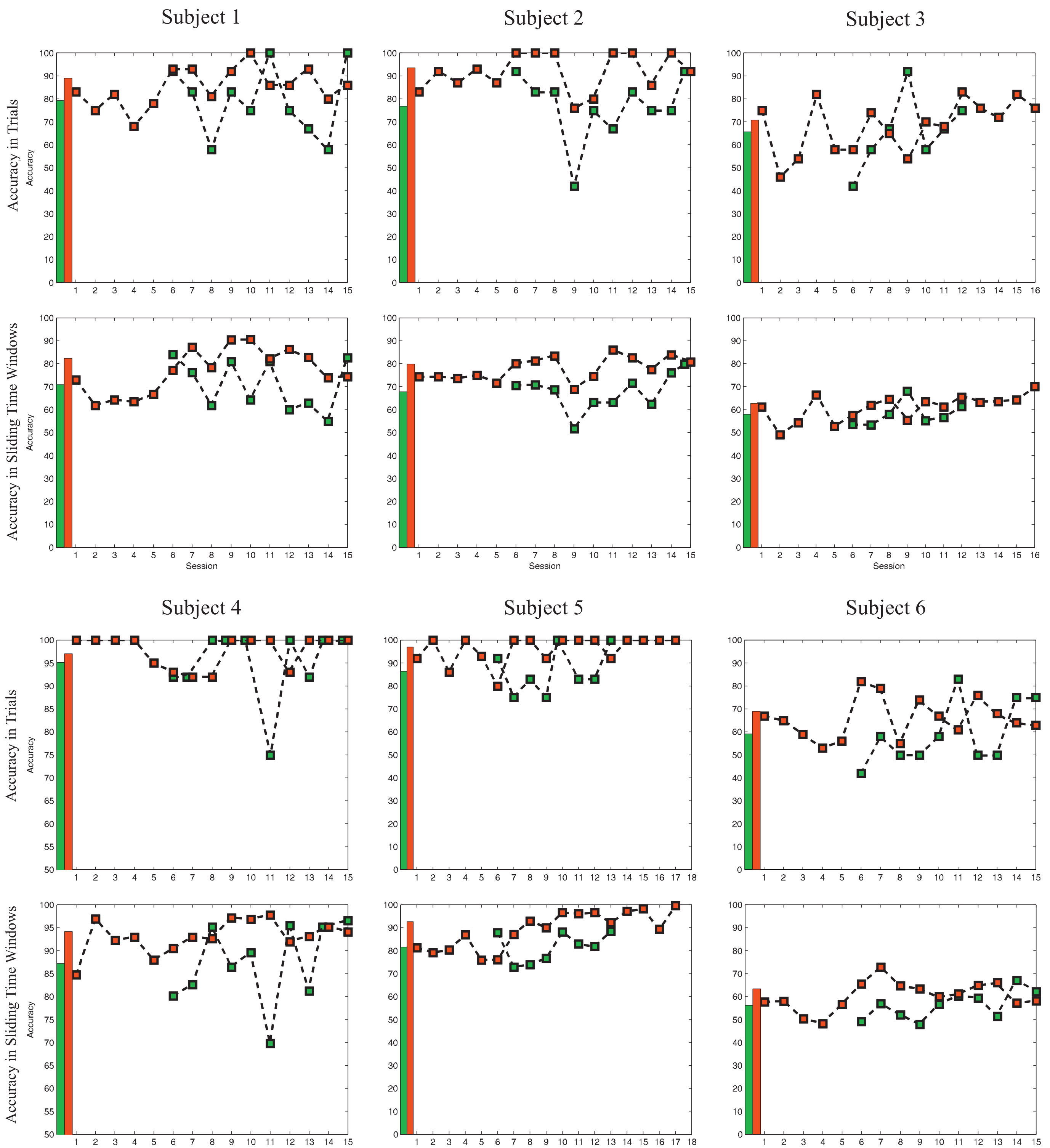

Fig. 6. A comparison in classification accuracy of the bilateral training paradigm (BTP) and the normal training paradigm (NTP). Red squares represent classification accuracy of BTP. In contrast, green squares represent classification accuracy of NTP. The bars drawn on the left of each graph indicates the mean of classification accuracy over all sessions starting with the sixth session (red bar for BTP and green bar for NTP). The first row and third row show accuracy in trials for each subject. The second row and fourth row show accuracy in sliding time windows for each subject. The squares are staggered for clear viewing when red square and green square are overlapped. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

5. Discussion and conclusion

This study combines human training and BCI system training so as to reduce the impact caused by non-stationary EEG signal and improving the recognition accuracy of BCI. Compared with normal training paradigm, the proposed paradigm using bilateral adaptation strategy achieves higher recognition accuracy. In addition, it is also observed that the bilateral training paradigm has better reliability. The reasons why bilateral training paradigm is more suitable for EEG signal classification can be explained as follows.

- The model used in bilateral training paradigm is updated in real-time by the EEG data from not only the current session but also the previous sessions. Hence, the model learns the latest information in time, meanwhile the information of the previous sessions is not lost. In contrast, the model used in the normal training paradigm just learns the latest information.
- All EEG data signals are used to train the model in normal training paradigm. In this way, the model simultaneously learns useless features for classification, even some features disturbed correct discrimination, while model is learning helpful features. Our method discards interferential EEG signals and only reserves EEG signals which are correctly classified in previous sessions for training model. Moreover, the model not only learns from the data of one previous session, but also from the data of a specific previous period. In addition, the system runs online and new data can be acquired over time for online system adaptation.
- There are five models in bilateral training paradigm. Each model is equivalent to a voting machine. The classification performance is enhanced by a combination of several voting machines. The outputs of five models are weighed to produce classification result. Weights are obtained according to training accuracy. Each weight indicates the reliability of corresponding model.
- Last but not least important, subjects are able to adjust immediately their own brain activities according to neuro-feedback. This is a good way for subjects to train how to control their own brain activities.

We also employed statistical analysis to reconfirm the comparison of two paradigms. The results of paired *t*-test illustrated bilateral training paradigm using bilateral adaptation strategy significantly outperformed normal training paradigm corresponding to the existing method in both conditions of trials and sliding time windows. Further more, we performed an off-line analysis to investigate which of the aspects of algorithm used in the BTP were actually effective for performance. As seen from results, training data proportionally and randomly chosen from previous sessions and multi-models are useful for performance. For practical or clinical applications, high recognition accuracy and reliability are necessary. Therefore, improving performance of BCI systems is important for transferring BCI technique to practical applications.

Acknowledgments

This paper was supported by the Science and Technology Commission of Shanghai Municipality (Grant No. 08511501701), National Basic Research Program of China (Grant No. 2005CB724301) and the National Natural Science Foundation of China (Grant No. 60775007).

References

Anthony CS, Zoltan JK. Principal-component localization of the source of the background EEG. IEEE Trans Biomed Eng 1995;42(1):59–67.
Ciuciu P, Abry P, Rabrait C, Wendt H. Log wavelet leaders cumulant based multifractal analysis of EVI fMRI time series: evidence of scaling in ongoing and evoked brain activity. IEEE J Sel Top Signal Process 2008;2(6):929–43.
Cooper RA, Dicianno BE, Brewer B, LoPresti E, Ding D, Simpson R, et al. A perspective on intelligent devices and environments in medical rehabilitation. Med Eng Phys 2008;30(10):1387–98.
Daly JJ, Wolpaw JR. Brain–computer interfaces in neurological rehabilitation. Lancet Neurol 2008;7(11):1032–43.
Duffau H. Brain plasticity: from pathophysiological mechanisms to therapeutic applications. J Clin Neurosci 2006;13(9):885–97.
Hearst MA, Dumais ST, Osman E, Platt J, Scholkopf B. Support vector machines. IEEE Intell Syst 1998;13(4):18–28.
Ikeda A, Shibbasaki H. Invasive recording of movement-related cortical potentials in humans. J Clin Neurophysiol 1992;9(4):409–520.
Keirn ZA, Aunon JI. A new mode of communication between man and his surroundings. IEEE Trans Biomed Eng 1990;37(12):1209–14.
Kübler A, Kotchoubey B, Hinterberger T, Ghanayim N, Perelmouter J, Schauer M, et al. The thought translation device: a neurophysiological approach to communication in total motor paralysis. Exp Brain Res 1999;124:223–332.
McFarland DJ, Lefkowicz AT, Wolpaw JR. Design and operation of an EEG-based brain–computer interface with digital signal processing technology. Behav Res Methods Instrum Comput 1997;29:337–45.
Müller-Putz GR, Scherer R, Pfurtscheller G, Rupp R. EEG-based neuroprosthesis control: a step towards clinical practice. Neurosci Lett 2005;382:169–74.
Pfurtscheller G, Neuper C. Motor imagery activates primary sensorimotor area in man. Neurosci Lett 1997;239:65–8.
Ramoser H, Muller-Gerking J, Pfurtscheller G. Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans Rehabil Eng 2000;8(4):441–6.
Shenoy P, Krauledat M, Blankertz B, Rao R, Müller K. Towards adaptive classification for BCI. J Neural Eng 2006;3:R13–23.
Sykacek P, Roberts S, Stokes M. Adaptive BCI based on variational Bayesian kalman filtering: an empirical evaluation. IEEE Trans Biomed Eng 2004;51(5):719–27.
Vapnik V. The nature of statistical learning theory. New York: Springer-Verlag; 1995.
Vidaurre C, Schloogl A, Cabeza R, Scherer R, Pfurtscheller G. A fully on-line adaptive BCI. IEEE Trans Biomed Eng 2006;53(6):1214–9.
Xu M, Takata H, Ge S, Hayami T, Yamasaki T, Tobimatsu S, et al. NIRS measurement of hemodynamic evoked responses in the primary somatosensory cortex by finger stimulation. In: IEEE/ICME international conference on complex medical engineering; 2007. p. 1425–9.