35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Extending Motor Imagery by Speech Imagery for Brain-Computer Interface

Li Wang, Xiong Zhang, and Yu Zhang

Abstract—**An electroencephalogram (EEG)-based brain computer interface (BCI) is a novel tool that translates brain intentions into control signals. As the operational dimensions of motor imagery are limited, we describe in this paper an extension of its capability by including speech imagery. Our new system was tested with the help of subjects, whose native language is Chinese. The tests were divided into two steps. The first step was speech imagery; consequently motor imagery and speech imagery were merged in the second step. Feature vectors of EEG signals were extracted from both common spatial patterns (CSP) and cross-correlation functions; then these vectors were classified by a support vector machine (SVM). The distinguishing accuracies of two intentions were found to be between 79.33% and 88.26%. This result shows that the capability of BCI for motor imagery can be extended by combining motor imagery and speech imagery.**

I. Introduction

Recently, brain-computer interface (BCI) technology based on classifying single trial electroencephalography (EEG) signals has been paid more attention by the researchers. Due to the function of exchanging information with the external world feasibly, this technology has been introduced in the state-of-the-art medical devices for motor dysfunction [1], such as amyotrophic lateral sclerosis (ALS) and traumatic brain injury. Many BCI systems, which are based on neuronal activities, are implemented by classifying EEG signals during motor imagery, e.g. imagining the movement of hand, foot and tongue [2]. These different imagined movements indicate the different spatial distribution of EEG (mainly are mu and beta rhythms) appearing on the contralateral hemisphere. These rhythms can be used to explained specific energy change of event-related desynchronization/synchronization (ERD/ERS) [3].

The maximum dimensions of BCI based on motor imagery is four[4]. However, with the increase of dimensions, training time will be extended and the classification accuracy will be reduced. In order to extend the dimensions of BCI based on motor imagery, other imagine paradigm ,e.g. speech imagery is proposed. Eric C Leuthardt had used electrocorticography (ECoG) speech network to control a BCI [5]. In his study, the letters of the alphabet, including OO, EE, AH and EH, were selected as materials of the speech imagery. Charles S. DaSalla proposed /a/ and /u/ as vowel speech imagery for EEG-based BCI [6]. In Data set V of BCI Competition III [7], mental imagery is classified into 3 categories: left hand motor imagery, right hand motor imagery and word association. The first letter of the words should be kept same in the word association. Most of English words are polysyllabic, which are more complicate in the application of BCI comparing with letters of the alphabet and vowels.

Based on the above reasons, the speech imagery according to Chinese characters is proposed to extend motor imagery BCI systems in this paper. Different from English words, the correlation between the shape and the meaning of Chinese characters (a kind of ideographic language) is closer than that between the shape and sound. Furthermore, Chinese characters are monosyllabic pronunciation, e.g. "左" is pronounced as "zuo" in third tone, which means "left" in English; "壹" is pronounced as "yi" in first tone, which means "one"; "移" is pronounced as "yi" in second tone, which means "move".

A specific rhythm , such as mu and beta rhythms linking to ERD/ERS of motor imagery, can be captured as a strong or attenuated activity using a bandpass-filter. The method of common spatial patterns (CSP) is an advanced feature extraction algorithm for extracting discriminant spatial features of EEG. It constructs spatial filters to maximize the variance for one class of imagery and to minimize the variance for the other one simultaneously. Based on rhythm modulation, the result of EEG synchronization in different cerebral cortex can be regarded as the feature value in the BCI systems. Cross-correlation function is an appropriate algorithm of EEG synchronization. Because the energy characteristics and synchronization of EEG are different and independent in the physiological mechanism, both CSP and the cross-correlation function are used to extract the eigenvalues of EEG respectively. The extracted eigenvalues are classified by support vector machine (SVM).

II. Methods

A. Data Acquisition

Six Chinese, right handed students (four males and two females) of Southeast University participated in the no feedback experiment. Their ages are from 22 to 27 with the average of 23.33. All of the students didn't attend similar experiment before. They were explained the purpose the procedure of the experiment and signed the Informed Consent. The experiment was under the guidance of the

This work was supported by the National Key Basic Research Program of China (No. 2010CB327705), the National High Technology Research and Development Program of China (No. 2012AA03A302), College Graduate Research and Innovation Foundation of Jiangsu Province under Grant CXLC12_0095, and this work was also supported by Shenzhen China Star Optoelectronics Technology Co. LTD.

L. Wang is with School of Electronic Science and Engineering, Southeast University, Nanjing 210096, China (e-mail: wangly1986@gmail.com).

X. Zhang and Y. Zhang are with School of Electronic Science and Engineering, Southeast University, Nanjing 210096, China.

Academic Ethics Committee. The experiment data were recorded by a SynAmps 2 system (Neuroscan Co., Ltd.).

This experiment was divided into two steps. The subject performed speech imagery according to the training paradigm in Fig. 1(a) firstly. Then he/she performed both motor imagery and speech imagery following the training paradigm in Fig. 1(b). In the first step, 15 EEG channels (the left of Fig. 2) are used to cover the positions involving the Broca's area and Wernicke's area. In the second step, 35 EEG channels (the right of Fig. 2) cover the cerebral cortex including the Broca's area, Wernicke's area, primary motor area (M1) and supplementary motor area (SMA). To measure the influence of ocular artifacts, the electrooculogram (EOG) (horizontal and vertical pairs) was also recorded. Channel-level preprocessing was performed before applying the EOG correction, referencing the signals to left mastoid and grounding the signals to forehead. Signals were sampled at 250Hz, and preprocessed by a 0.1-100Hz band-pass filter.

B. Experimental Paradigm

The training paradigms of two steps are almost identical except the “Cue” in Fig. 1. The “Cue” is a Chinese character, such as “左” or “壹” in Fig. 1(a) and is a Chinese character “移” or an arrow in Fig. 1(b). Timing of training paradigms is shown in Fig. 1. To start with each trial, a fixed cross is displayed with black background, which is ready period of 1 s. After the ready period, a “Cue” appears for 1 s on the screen. In the next 4s, the subject is required to keep reading the Chinese character in mind or imagining his/her hand movement according to the content of “Cue”. The subject can't move lip or make a sound when the “Cue” is a Chinese character and he/she also can't move his/her hands when the “Cue” is an arrow. In order to prevent the subject adapting this experiment, a fixed asterisk with black background is displayed for 2~3 s to suggest the subject have a rest after the imagery period is finished. Each step of the experiment includes 5 runs and each of cues is randomly displayed 15 times in each run. Between each run, the subject may have a break of 5 minutes.

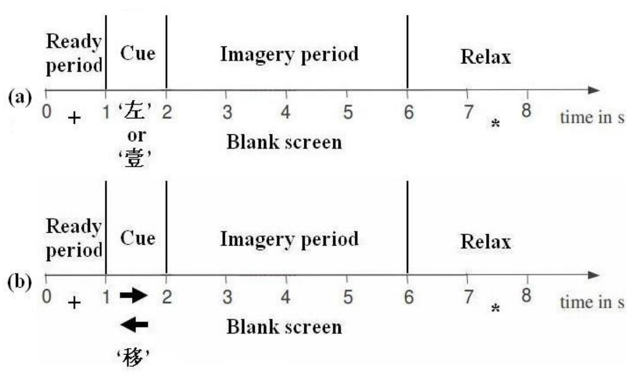


Figure. 1 (a) Timing of a trial of the first step training paradigm. 3~5 s of every trial, and 2 s of relax period before imagery period of each Chinese character is regarded as Rest. These two fragments of EEG signals are calculated by CSP spatial filter and cross-correlation function. (b) Timing of a trial of the second step training paradigm. EEG signals from 3~5 s of every trial are calculated by CSP spatial filter and cross-correlation function.

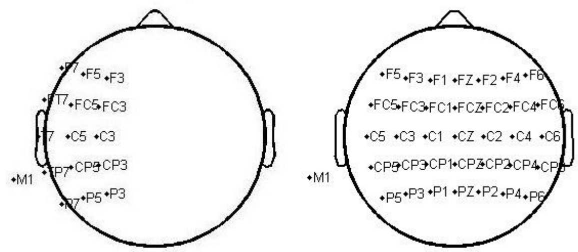


Figure. 2 Position of EEG electrodes. Left electrode setup has 15 channels for the first step and the right one has 35 channels for the second step. These electrode positions of both setups correspond to the international 10–20 system.

C. Feature Extraction

As EEG signals mainly distribute in the alpha and beta wave, signals are filtered by band-pass zero-phase filter at a range of 6-30 Hz before analysis and classification of EEG.

1). Common Spatial Patterns (CSP)

CSP is a supervised method to extract the task related components. The signal-to-noise ratio of EEG can be improved effectively by this method. As the result of simultaneous diagonalization of the two corresponding covariance matrices, two different imagery categories of EEG signals are projected into low-dimensional spatial subspace by CSP spatial filters. The variance of two types of EEG signals matrices can be maximized by this transformation. According to the two tasks shown in Fig. 1, EEG signals can be modeled as the combination of specific components and common components. More detail of CSP is referenced to [8]. The best projection matrix W is calculated by CSP, and W^{1} is the inverse matrix of W. The columns of W^{1} are common spatial patterns.

Fig. 3 shows the four most significant common spatial patterns from the result of “壹 (one)” versus Rest. As the weight values of the common spatial patterns of “壹 (one)” versus Rest are larger than those of “左 (left)” versus “壹 (one)”, it can be speculated that it is harder to distinguish which Chinese character is read in mind than to distinguishing whether subjects are reading one character. This speculation is coincident with subsequent results.

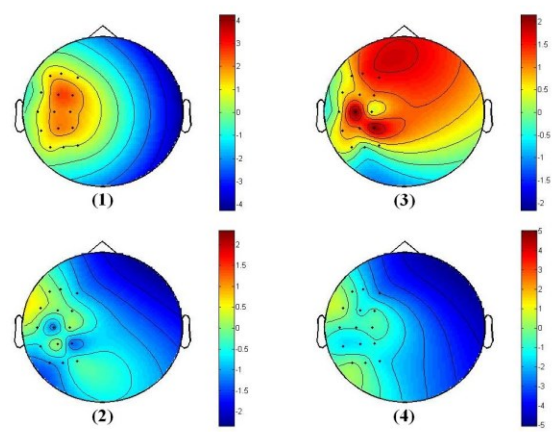


Figure. 3 The “壹 (one)” versus Rest topographic maps of four most meaningful common spatial patterns are obtained by CSP for subject S2. (1), (2), (3) and (4) of each subgraph are the four spatial patterns.

The mapping of each trial can be converted into $Z=WX$ (X is the EEG signal). First m columns and last m columns of the Z matrix are used for the construction of the classifier. Then the feature vectors f_p can be calculated from signals Z_p ($p = 1\ldots 2m$) as (1). The typical number of m is 2.

$$f_p = \log\left(\frac{\mathrm{var}(Z_p)}{\sum_{i=1}^{2m}\mathrm{var}(Z_i)}\right) \tag{1}$$

2). Cross-correlation Function

EEG synchronization phenomenon is the important performance [9], which considered as functional collaboration and integration of different brain regions to complete the cognitive behavior. To improve the accuracy of classification, it's appropriate to extract feature values of the EEG synchronization in different cerebral cortex for BCI systems.

The methods of the analyzing EEG synchronization include cross-correlation function, coherence function and phase synchronization [10] et.al. The cross-correlation function is the simplest method and suitable for real-time BCI systems. EEG synchronization is analyzed by this method in this paper. The cross-correlation function is defined as follow:

$$\mathrm{c}_{xy}(\tau) = \left\langle \left(\frac{x_i - \bar{x}}{\sigma_x}\right)\left(\frac{y_{i+\tau} - \bar{y}}{\sigma_y}\right)\right\rangle_i = \frac{1}{N-\tau}\sum_{i=1}^{N-\tau}\left(\frac{x_i - \bar{x}}{\sigma_x}\right)\left(\frac{y_{i+\tau} - \bar{y}}{\sigma_y}\right) \tag{2}$$

Two EEG signals are set as x_i and y_i, i=1,…, N. $\bar{x}$ and $\bar{y}$ are the mean of x and y. σ_x and σ_y are the corresponding standard deviation. τ represents a time-shift. The measure of the linear synchronous between x and y is given by cross-correlation function c_{xy}. The range of the cross-correlation function is [0 1], in which 0 indicates that there is no synchronization between the x and y, and 1 represents the greatest degree of synchronization. Cross-correlation are calculated for the channels F3, F5, F7, FC3, CP3, CP5, CP7 and P7 pair to pair. These channels are near the Broca's area and Wernicke's area, which have close relationship with speech imagery. The result of one subject (S3) is shown in Fig. 4. The cross-correlations of F5-CP5 and F3-CP3 increase significantly when the subject reads a Chinese character.

D. Feature Classification

The feature vectors of EEG are classified by the support vector machine (SVM), which is based on statistical learning theory. It solves the problem of searching a hyperplane to separate the training data X with labels Y. LIBSVM [11] with radial basis kernel is selected in this paper because of its wide application.

III. RESULTS

As shown in Fig. 5, for EEG signals of speech imagery in the first step, two fragments of signals will be analyzed. One fragment data has 2 s time period, from 3 s to 5 s of every trial in Fig.1(a), and another one is 2 s of relax period in Fig.1(a). After filtered by CSP spatial filters, four feature values are extracted by (1). At the same time, the channel pairs of F3-CP3, F5-CP5, F3-F5, and CP3-CP5, near the Broca's area and Wernicke's area, are selected to calculate EEG synchronization phenomenon between these two areas. Four feature values are extracted by (2). As EEG signals between two channels are linearly related, τ is set to 0. Table I presents the accuracy that calculated by 10×10 cross-validation. This method randomly splits the data into ten parts, and nine of which are used to train the classifier and the remaining one is used to test it. This process is repeated ten times.

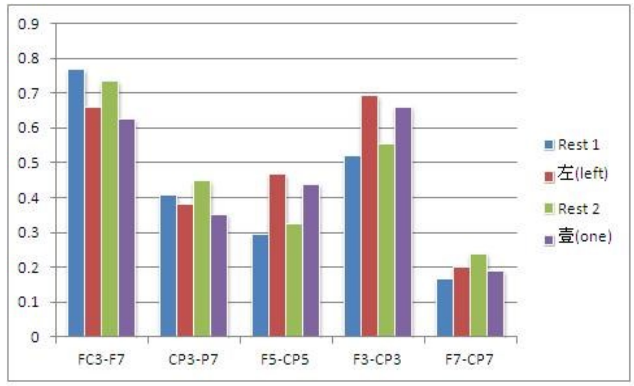
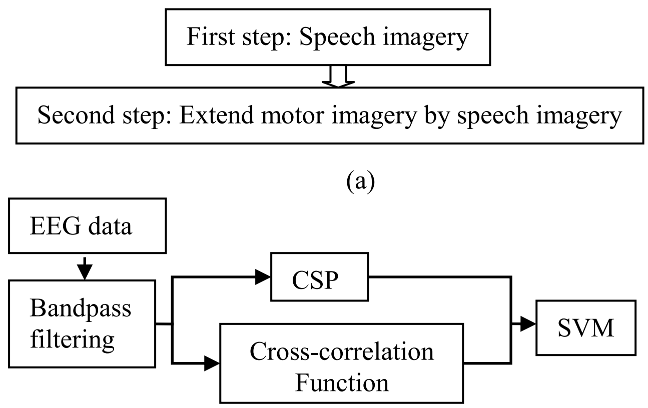


Figure. 4 The values of cross-correlation are calculated from average of 75 times for subject S3. τ is set to 0. "Rest 1" is corresponded to "左(left)" and "Rest 2" is corresponded to "壹(one)".



(a)

(b)

Figure. 5 (a) Flowchart of experimental paradigm. (b) Flowchart of EEG data processing

As shown in Table I, it is hard to separate which Chinese characters is read, but it is still effective to distinguish whether subjects are reading a character in mind. In order to compare with the combined algorithm, feature vectors are extracted by CSP and the cross-correlation function respectively. Three average results of the combined algorithm are 85.39, 85.51 and 66.08. They are better than the average results of CSP, which are 83.92, 83.51, and 66.22. They are also better than the average results of the cross-correlation function, which are 79.54, 77.46 and 59.4. As language and motor are processed by different cerebral cortex, speech imagery and motor imagery are classified in the next part of paper.

TABLE I. THE CLASSIFICATION RESULT OF SPEECH IMAGERY

Subject	Accuracy±std (%)		
	"左"(left) vs Rest	*"壹"(one) vs Rest*	*"左"(left) vs"壹"(one)*
S1	78.67±1.46	79.33±0.89	64.67±1.29
S2	89.46±1.91	91.86±0.73	68.74±1.26
S3	80.06±0.79	79.2±1.12	66.81±2.77
S4	93.53±0.72	94.67±0.78	61.27±1.61
S5	83.12±1.42	81.83±1.61	69.87±1.73
S6	87.33±0.54	86.19±1.21	65.13±2.83
Mean	85.39	85.51	66.08

In order to compare disparity of different channels between speech imagery and motor imagery, event-related spectral perturbation (ERSP) is plotted by EEGLAB [12]. ERSP is superposition of single trial energy spectrum distribution. As channel C3 and F3 are close to the primary motor area of left brain and the Broca's area separately, ERSP of these two channels from subject S2 are plotted in Fig. 6 respectively.

As shown in Fig. 6, 5-12Hz EEG of channel C3 is displaying ERD when subject S2 is imagining right hand movement and the energy of 10-15Hz EEG from channel F3 increased when he is reading "移 (move)" in mind. EEG data processing is similar with the first step, and it is also better to combine CSP and cross-correlation function than single algorithm. The results of combination are shown in Table II. Besides channels F3, CP3, F5 and CP5, the channel pair of C3-FCz and C4-FCz are selected to calculate synchronization phenomenon between M1 and SMA [13]. Two average accuracies of speech imagery vs motor imagery are better than left vs right (see Table II). The result indicates that it is necessary to improve the accuracy of motor imagery by long-term training.

IV. CONCLUSION

In day-to-day life, it's common to utilize both real and imagined speech by an individual. So a more easily operable paradigm can be offered by using speech imagery for BCI systems. Motor imagery can be extended by speech imagery using one Chinese character without long-term training. In our future work, the separation of reading two Chinese characters in mind will be improved by utilizing more appropriate algorithm. The comparison between speech imagery, motor imagery, and speech+motor imagery will also be studied.

TABLE II. THE CLASSIFICATION RESULT OF EXTENDING MOTOR IMAGERY BY SPEECH IMAGERY

Subject	Accuracy±std (%)		
	Left vs Right	*Left vs"移"(move)*	*Right vs"移"(move)*
S1	81.13±0.89	82.52±1.03	80.86±2.22
S2	74.64±2.94	79.33±1.86	80.53±2.26
S3	77.33±1.54	84.76±2.03	85.33±2.13
S4	76.6±2.46	86.67±1.67	87.47±1.63
S5	71.53±0.73	88.26±0.92	82.67±1.05
S6	74.4±1.22	83.28±0.89	86.13±1.32
Mean	75.94	84.14	83.83

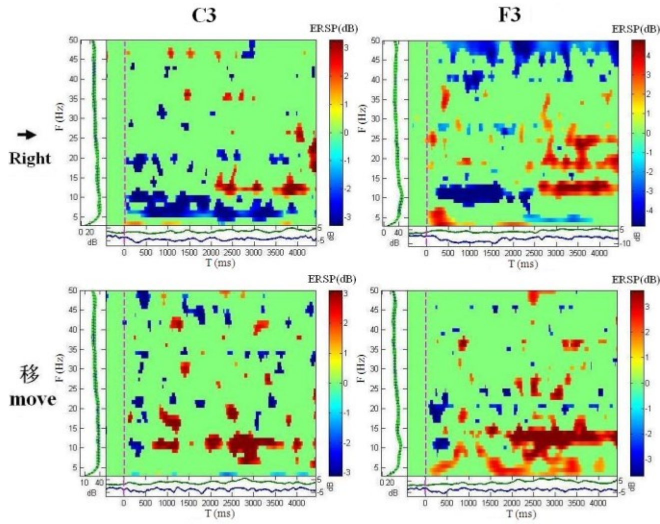


Figure. 6 ERSP of two types of imagery from channels C3 and F3 for subject S2. Bootstrap significance level is 0.01, and t=0 s is corresponded to t=1 s in Fig. 2, when the cue appears. The horizontal axis represents time and the vertical axis represents frequency.

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurt-scheller, and T.M. Vaughan, "Brain-computer interfaces for communication and control," *Clin. Neurophysiol.*, vol. 113, pp. 767–791, 2002.

[2] V. Morash, O. Bai, S. Furlani, P. Lin and M. Hallett, "Classifying EEG signals preceding right hand, left hand, tongue, and right foot movements and motor imageries," *Clin. Neurophysiol.,* vol. 119, pp. 2570–2578, 2008.

[3] G. Pfurtscheller, and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles," *Clin. Neurophysiol.,* vol.110, pp.1842-1857, 1999.

[4] M. Naeem, C. Brunner, R. Leeb, B. Graimann and G. Pfurtscheller, "Seperability of four-class motor imagery data using independent components analysis," *J. Neural Eng.*, vol. 3, pp. 208-216, 2006.

[5] E. C Leuthardt, C. Gaona, M. Sharma, N. Szrama, J. Roland, Z. Freudenberg, J. Solis, J. Breshears and G. Schalk, "Using the electrocorticographic speech network to control a brain–computer interface in humans," *J. Neural Eng*., vol. 8, pp. 1-11, 2011.

[6] C. S. DaSalla, H. Kambara, M. Sato and Y. Koike, "Single-trial classification of vowel speech imagery using common spatial patterns," *Neural Networks*, vol. 22, pp. 1334-1339, 2009.

[7] S. Chiappa, and J. del R. Millán, "Data set V of BCI Competition III," http://www.bbci.de/competition/iii/desc_V.html, 2005.

[8] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Trans. Rehabil. Eng*., vol. 8, no. 4, pp. 441-446, Dec. 2000.

[9] F. Varela, J. P. Lachaux, E. Rodriguez and J. Martinerie, "The brainweb: Phase synchronization and large-scale integration," *Nat. Rev. Neurosci.*, vol. 2, pp. 229-239, 2001.

[10] R. Quian Quiroga, A. Kraskov1, T. Kreuz1, and P. Grassberger, "Performance of different synchronization measures in real data: A case study on electroencephalographic signals," *Phys. Rev. E.*, vol. 65:041903, 2002.

[11] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," http://www.csie.ntu.edu.tw/cjlin/libsvm, 2011.

[12] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics," *Journal of Neuroscience Methods*, vol. 134, pp. 9-21, 2004.

[13] Y. Wang, B. Hong, X. Gao and S. Gao, "Design of electrode layout for motor imagery based brain-computer interface," *Electron. Lett.*, vol. 43, pp. 557-558, 2007.