J. Neural Eng. **4** (2007) 120–129 doi:10.1088/1741-2560/4/2/012

Amplitude and phase coupling measures for feature extraction in an EEG-based brain–computer interface

Qingguo Wei[1,2], Yijun Wang[2], Xiaorong Gao[2] and Shangkai Gao[2]

[1] Department of Electronic Engineering, School of Information Engineering, Nanchang University, Nanchang 330031, People's Republic of China
[2] Department of Biomedical Engineering, School of Medicine, Tsinghua University, Beijing 100084, People's Republic of China

E-mail: gsk-dea@tsinghua.edu.cn (Shangkai Gao)

Received 18 July 2006
Accepted for publication 20 February 2007
Published 28 March 2007
Online at stacks.iop.org/JNE/4/120

Abstract
Most of the feature extraction methods in existing brain–computer interfaces (BCIs) are based on the dynamic behavior of separate signals, without using the coupling information between different brain regions. In this paper, amplitude and phase coupling measures, quantified by a nonlinear regressive coefficient and phase locking value respectively, were used for feature extraction. The two measures were based on three different coupling methods determined by neurophysiological *a priori* knowledge, and applied to a small number of electrodes of interest, leading to six feature vectors for classification. Five subjects participated in an online BCI experiment during which they were asked to imagine a movement of either the left or right hand. The electroencephalographic (EEG) recordings from all subjects were analyzed offline. The averaged classification accuracies of the five subjects ranged from 87.4% to 92.9% for the six feature vectors and the best classification accuracies of the six feature vectors ranged between 84.4% and 99.6% for the five subjects. The performance of coupling features was compared with that of the autoregressive (AR) feature. Results indicated that coupling measures are appropriate methods for feature extraction in BCIs. Furthermore, the combination of coupling and AR feature can effectively improve the classification accuracy due to their complementarities.

1. Introduction

A brain–computer interface (BCI) is a communication system that does not depend on the normal output pathways consisting of periphery nerves and muscles. A BCI transforms mental decision and reaction into control commands by analyzing the bioelectrical brain activity (Wolpaw *et al* 2002). The technique can help patients, having totally lost motor ability but having intact cognition (e.g., people with amyotrophic lateral sclerosis, cerebral palsy or locked-in syndrome), realize the control of external facilities, motor neuroprostheses, wheelchairs and so on, to improve their living quality.

Most of the existing BCIs utilize electroencephalograms (EEGs) recorded from the scalp as input signals because they have the appeal of being both easily available and free from procedural risks. Significant technological progress has been achieved in the past decade toward developing an EEG-based BCI system; its practical use, however, is still limited by a low information transfer rate (ITR). The key to increase ITR is to boost the classification accuracy of single-trial data. For a chosen experiment paradigm, the classification performance mainly depends on whether an appropriate feature vector is extracted from the multi-channel EEG recordings that reflect brain intention.

The commonly used features are extracted from band power (Pfurtscheller *et al* 2000, Pfurtscheller and Neuper 2001), autoregressive (AR) model (Anderson *et al* 1998, Huan and Palaniappan 2004, Dornhege *et al* 2004) or common

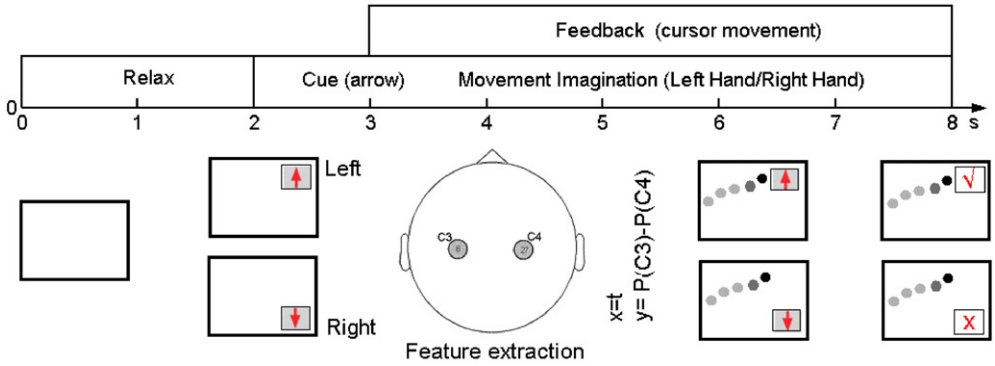

Figure 1. Online feedback experiment paradigm of the motor imagery-based brain–computer interface.

spatial pattern (CSP) (Muller-Gerking *et al* 1999, Ramoser *et al* 2000). All these methods are based on the dynamic behavior of single signals; none of them make use of the coupling information between two EEG signals. However, brain signals also exhibit coupling phenomena between different brain regions because cognitive acts require the integration of different functional areas widely distributed over the brain and the communication among them (Varela 1995, Friston *et al* 1997, Tononi and Eelman 1998, Rodriguez *et al* 1999, Varela *et al* 2001). Therefore, coupling measures might be a new method for feature extraction, and are drawing increasing attention.

Signal coupling or interdependence can be described by linear and nonlinear methods. Linear interdependence is denoted by cross-correlation in a time domain and by coherence in a frequency domain (Gersetein *et al* 1978, Gotman 1987, Zaveri *et al* 1999), while nonlinear coupling is represented by nonlinear regressive (NLR) coefficients and phase locking value (PLV) (Le Van Quyen *et a*l 1998, Fernandes de Lima *et al* 1990, Chavez *et al* 2003, Bartolomei *et al* 2001, Lopes da Silva *et al* 1989, Tass *et al* 1998, Lachaux *et al* 1999, Mormann *et al* 2000, Spiegler *et al* 2004, Brunner *et al* 2005). However, the dynamical behavior of brain electrical activity is essentially nonlinear during a cognitive task and can be better characterized by nonlinear interdependences. Furthermore, coherence does not separate the effects of amplitude and phase in the interrelation. In contrast, NLR coefficients and PLV characterize the nonlinear couplings of amplitude and phase respectively and, consequently, could be discriminatory classification features.

Amplitude and phase coupling measures have already been employed for the recognition of different mental tasks in BCIs. Gysels and Celka (2004) demonstrated that phase coupling measure provided relevant information for the classification of mental tasks in the framework of BCIs. They successfully recognized three cognitive tasks by combining PLVs and power spectral density estimation. Wei *et al* (2006) presented amplitude coupling measurements to classify single-trial electrocorticogram (ECoG) data. NLR coefficients between signals on ten leads are extracted in two frequency bands, 0–3 Hz and 8–30 Hz, as classification features. A classification accuracy of 93% was achieved by selecting an optimal feature subset from a high-dimensional feature space. Also, Wang *et al* (2006) have applied the phase synchrony measurement in motor cortex for classifying single-trial EEG during motor imagery.

The purpose of this paper is to investigate whether amplitude and phase couplings are appropriate methods for feature extraction in an EEG-based BCI, and assess the performance of coupling features by comparing classification accuracies of the NLR coefficient and PLV with those of the AR coefficients.

2. Method

2.1. Experimental paradigm and data acquisition

Five healthy right-handed volunteers (S1–S5), four males and one female, aged between 22 and 25, took part in the experiment. They were selected randomly from the students in the department. Most of them had previously participated in BCI studies and were familiar with the experimental environment.

Figure 1 shows the experiment paradigm of online BCI control with visual feedback, which is same as that utilized by Wang *et al* (2006). During the experiment, the subject was sitting in a comfortable armchair facing a computer screen approximately 1 m in front of him/her. Each trial was 8 s long and started with a blank screen during which the subject could relax. At second 2, a visual cue (arrow) appeared on the screen, indicating the cognitive task to be performed. Depending on the direction of the arrow, the subject was instructed to imagine a movement of either the left or the right hand. The 'left hand' and 'right hand' movements' imagination were designated to control the one-dimensional cursor movement, up and down, respectively. Starting from second 3, feedback information was provided at each sampling point by a cursor movement from left to right of the screen. The cursor moves at a steady rate, with its vertical movement controlled by the power difference of μ rhythm between the signals on electrodes C3 and C4, which was caused by contralateral event-related desynchronization (ERD) during unilateral hand motor imagery (Pfurtscheller and Lopes da Silva 1999). The power difference was calculated using all samples before each feedback point. At second 8, a true or false mark appeared to display the final result of the trial, and then another trial began. For the task 'left', if the cursor finally arrived at the upper part of the designated right border, the mark is 'true',

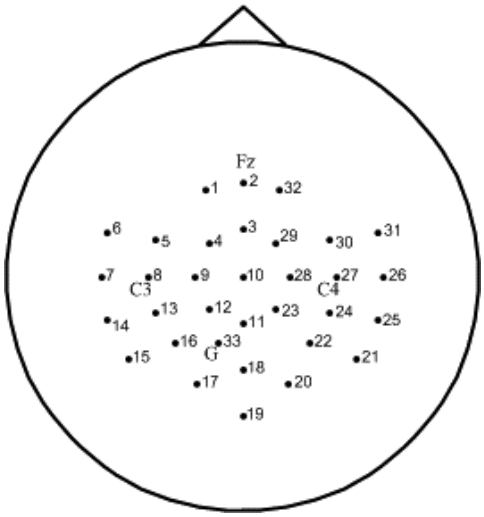

Figure 2. Electrode locations for EEG recordings. The 32 monopolar silver–silver chloride electrodes overlie the primary sensorimotor area (SM1) and the supplementary motor area (SMA). The ground electrode is positioned at G.

otherwise the mark is 'false'. For the task 'right', marks are just the contrary.

EEG was recorded with 32 monopolar silver–silver chloride scalp electrodes, spaced with an approximately 2.5 cm distance, over the primary sensorimotor area (SM1) and the supplementary motor area (SMA) as illustrated in figure 2. The ground electrode is positioned at G. The EEG signals were amplified and sampled at 1024 Hz with 24 bit resolution by a BioSemi ActiveTwo system, subsampled to 256 Hz by Labview software and then stored.

The experiment was divided into four runs, consisting of 60 trials each. The interval between two runs was about 5 min. The sequence of 'left' and 'right' trials was randomized throughout each run, and the number of two tasks was equally partitioned. No trials were discarded. Thus, the dataset for each subject contained a total of 240 trials (120 trials per class) and all trials were used for the following offline analysis.

2.2. Data preprocessing

Considering the reaction time of a subject to visual stimulation, we extracted a window of length 5 s from the data of each electrode starting from 0.5 s after the appearance of the visual cue (arrow). To save the computational load and reduce the requirement for memory, all trials were downsampled by keeping every other sample starting with the first. Thus for each trial and each electrode, we obtained an EEG time series consisting of 640 samples. To enhance the difference between these two tasks and reduce the effect of artifacts, common average reference (CAR) was used to re-reference the resampled data.

Prior to the extraction of coupling features, all EEG channels were filtered between 8 and 30 Hz by a band pass digital filter of Chebysheve type I. The filtering was performed forwardly and reversely to avoid amplitude and phase distortion. The 8–30 Hz frequency range was chosen because it contains all μ and β frequency components of the EEG which are important for the discrimination task (Ramoser *et al* 2000). Figure 3 shows the averaged amplitude spectra over all trials of each task on electrodes C3 and C4 from subject S1.

To extract features from AR coefficients, all EEG channels were high pass filtered at 4 Hz. The 4 Hz high pass filtering was performed to prevent the AR model from being distorted by EEG baseline drifts and, meanwhile, to retain the oscillating components in EEG signals as much as possible. In order to sharpen the spectral information to focal brain sources, spatial Laplacian filters were applied.

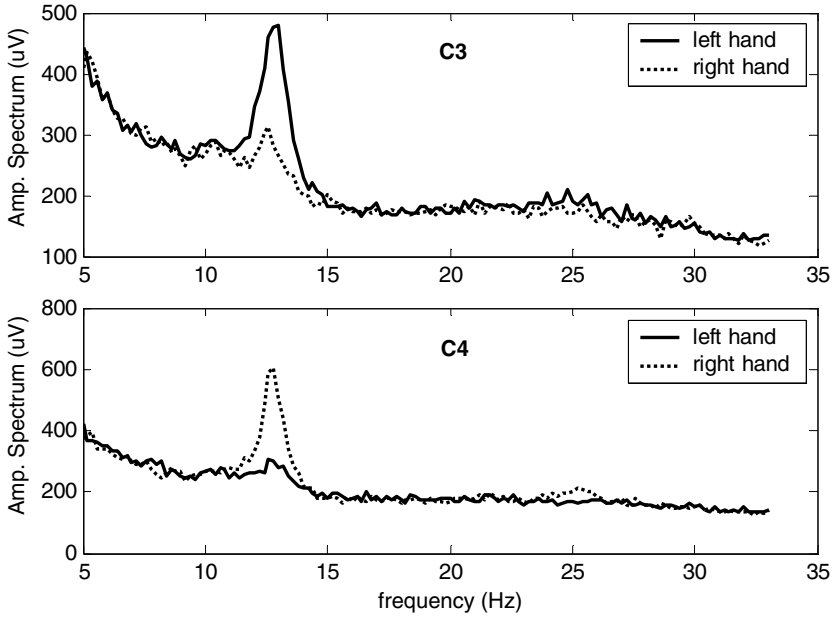

Figure 3. The averaged amplitude spectra over all trials of each task on electrodes C3 and C4 from subject S1.

2.3. Amplitude and phase coupling measures

The nature of the interdependence between EEG signals is characterized both by the amplitude association and by the corresponding phase relationship.

2.3.1. Amplitude coupling measure: NLR coefficient. Assume x and y are the signals on two leads, respectively. Given signal x, the expectation of signal y, denoted as $\mu_{y|x}$, is the regression curve of y on x:

$$\mu_{y|x}(x) = \int_{-\infty}^{+\infty} y p(y|x)\,\mathrm{d}y, \tag{1}$$

where $p(y|x)$ is the conditional probability of y given signal x.

The reduction of variance of y that can be obtained by predicting the y values from x according to the regression curve is the coupling measure given by

$$\eta^2_{y|x} = \frac{\mathrm{var}(y) - E[(y - \mu_{y|x}(x))^2]}{\mathrm{var}(y)}. \tag{2}$$

$E[(y - \mu_{y|x}(x))^2]$ estimated from the regression curve is called the explained variance, i.e. it is explained or predicted on the basis of x. By subtracting the explained variance from the total variance $\mathrm{var}(y)$, one obtains the unexplained variance. The basic idea is that if the amplitude of signal y is thought of as a function of the amplitude of signal x, given a certain value of x, the value of y can be predicted according to a NLR curve.

The estimation of this measure is called the NLR coefficient h^2. To obtain an approximation of the regression curve, the x amplitude values are subdivided into M bins (M is determined experimentally and taken as 20 in this study.) For each bin, the x value in the midpoint (p_i) and the average of y values (q_i) are calculated, and the resulting points (p_i, q_i) are connected by segments of straight lines. Consequently, the NLR coefficient h^2 can be calculated according to the following expression:

$$h^2 = \frac{\sum_{n=1}^{N}(y_n - \langle y\rangle)^2 - \sum_{n=1}^{N}(y_n - \hat{\mu}_{y|x}(x_n))^2}{\sum_{n=1}^{N}(y_n - \langle y\rangle)^2}, \tag{3}$$

where $\hat{\mu}_{y|x}(x_n)$ is the linear piecewise approximation of the regression curve and $\langle y\rangle$ denotes the average of y values over the N points of the time series. The estimator h^2 represents the strength of the association between the two signals and can take values between zero (y is totally independent of x) and one (y is totally dependent on x). The NLR coefficient h^2 can denote both linear and nonlinear relationships between two signals. If the relationship is linear, then $h^2_{y|x} = h^2_{x|y}$; otherwise $h^2_{y|x} \neq h^2_{x|y}$.

2.3.2. Phase coupling measure: PLV. Phase coupling between two signals is quantified by PLV. This measure requires first the estimation of the instantaneous phase of each signal that can be obtained by means of either the analytic signal or the wavelet analysis. Le Van Quyen *et al* (2001) showed that these two methods are equivalent for the analysis of EEG signals. Here, we follow the analytic signal method to calculate the instantaneous phase of an arbitrary signal. Given a signal $x(t)$, the analytic signal $\eta_x(t)$ is defined as

$$\eta_x(t) = x(t) + \mathrm{i}\tilde{x}(t) = A_x(t)\,\mathrm{e}^{\mathrm{i}\theta_x(t)}, \tag{4}$$

where $\tilde{x}(t)$ is the Hilbert transform of $x(t)$:

$$\tilde{x}(t) = \frac{1}{\pi}\mathrm{p.v.}\int_{-\infty}^{+\infty}\frac{x(t)}{t-\tau}\,\mathrm{d}\tau. \tag{5}$$

p.v. denotes that the integral is taken in the sense of Cauchy principal value, and the instantaneous phase is calculated as follows:

$$\theta_x(t) = \arctan\left(\frac{\tilde{x}(t)}{x(t)}\right). \tag{6}$$

The instantaneous phase $\theta_y(t)$ of signal $y(t)$ can be determined according to the same procedure as above.

Signals x and y are said to be $n : m$ synchronized if the difference of their instantaneous phases $\Delta\theta(t) = n\theta_x(t) - m\theta_y(t)$, for $\{n, m\} \in \mathbb{Z}$, remains bounded for all t. Since our goal is to measure the phase coupling between signals derived from the same physiological system (i.e. the brain), it is very likely that the phase locking ratio $n : m$ equals 1:1. Thus we adopt the case of $n = m = 1$, and subsequently phase locking is quantified as

$$\mathrm{PLV}_t = |\langle \mathrm{e}^{\mathrm{i}\Delta\theta(t)}\rangle_t|, \tag{7}$$

where $\langle\cdot\rangle_t$ is the operator of averaging over time. For discrete signals, the phase locking value is calculated as follows:

$$\mathrm{PLV} = \left|\frac{1}{N}\sum_{n=1}^{N}\mathrm{e}^{\mathrm{i}\Delta\theta(n)}\right|. \tag{8}$$

A PLV is equal to the average length of all unit vector $\mathrm{e}^{\mathrm{i}\Delta\theta(n)}$ in one window. PLVs are normalized and vary between zero and one. A PLV of zero means that the phase of the two signals is not coupled at all, while a PLV of one means that the signals are perfectly coupled because there is a constant phase difference over all time points.

Although the instantaneous phase can be obtained for arbitrary broad band signals, this has a clear physical meaning only for narrow band signals (Boashash 1992). As a result, filtering is required to separate the frequency band of interest from the wide band activity.

2.4. Feature extraction

To correctly translate brain intention into a control signal, the most discriminative information must be extracted from the EEG recordings. Feature extraction is the transformation of the original data to a data set with a reduced number of variables so that in the low-dimensional signal space, the difference of brain state between two cognitive tasks is more distinguished. It is here where neurophysiological *a priori* knowledge is very beneficial.

2.4.1. Coupling-based features. Prior to coupling measures, an important issue that must be solved is how to determine meaningful associate areas for a given cognitive task. Gerloff *et al* (1998) demonstrated that for both externally and internally paced finger extensions, functional coupling occurred between the primary sensorimotor cortex (SM1) of both hemispheres and between SM1 and the mesial premotor

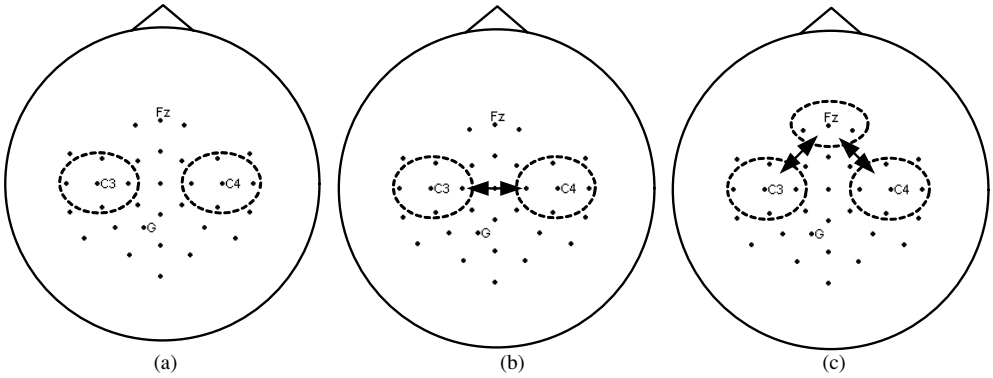

Figure 4. Three coupling methods and corresponding electrodes of interest: (a) coupling between any two electrodes within each of the two ellipses around C3 and C4 (CW), (b) coupling between each electrode within the ellipse around C3 and each electrode within the ellipse around C4 (CB1), (c) coupling between each electrode within the ellipse around Fz and each electrode within the two ellipses around C3 and C4 (CB2).

Table 1. The corresponding relationship of each feature (dimension) with electrode locations for the three types of couplings: CW, CB1 and CB2.

Feature	1	2	3	4	5	6	7	8	9	10
Electrode										
CW	5–7	5–8	5–9	5–13	7–8	7–9	7–13	8–9	8–13	9–13
CB1	5–30	5–28	5–27	5–26	5–24	7–30	7–28	7–27	7–26	7–24
CB2	5–1	5–2	5–32	7–1	7–2	7–32	8–1	8–2	8–32	9–1
Feature	11	12	13	14	15	16	17	18	19	20
Electrode										
CW	30–28	30–27	30–26	30–24	28–27	28–26	28–24	27–26	27–24	26–24
CB1	8–30	8–28	8–27	8–26	8–24	9–30	9–28	9–27	9–26	9–24
CB2	9–2	9–32	13–1	13–2	13–32	30–1	30–2	30–32	28–1	28–2
Feature	21	22	23	24	25	26	27	28	29	30
Electrode										
CW										
CB1	13–30	13–28	13–27	13–26	13–24					
CB2	28–32	27–1	27–2	27–32	26–1	26–2	26–32	24–1	24–2	24–32

areas (PM), probably including the supplementary motor area (SMA). The study of event-related coherence showed that synchronization between μ rhythms occurred in the precentral area and SM1 (Andrew and Pfurtscheller 1996). Spiegler *et al* (2004) investigated phase coupling between different motor areas during tongue-movement imagery and found that phase-coupled 10 Hz oscillations were induced in SM1 and SMA. These studies give convincing evidences that coupling within hemispheres, coupling between hemispheres and coupling between SM1 and SMA exist during different cognitive and motor tasks.

Although every electrode recording in these three areas provides relevant information for classification, the use of fewer electrodes is desired for a practical BCI. Thus we selected five electrodes surrounding C3 or C4 in each hemisphere and three electrodes in SMA, as illustrated in figure 4. The coupling measures were based on these electrodes of interest and, accordingly, three coupling methods were defined: (a) coupling between any two electrodes within each of the two ellipses around C3 and C4 (CW), (b) coupling between each electrode within the ellipse around C3 and each electrode within the ellipse around C4 (CB1), (c) coupling between each electrode within the ellipse around Fz and each electrode within the two ellipses around C3 and C4 (CB2). These coupling methods were separately used for calculating the NLR coefficient and PLV, and thus six different feature vectors were obtained.

Each feature vector was constructed by concatenating NLR coefficients or PLVs from all electrode pairs. The dimensions of feature vectors derived from CW, CB1 and CB2 are 20, 25 and 30 respectively for phase coupling measure (quantified by PLV) and double respectively for amplitude coupling measure (quantified by NLR). Although the strengths of amplitude couplings in two directions are not equivalent, their difference is small. For a fair comparison of classification performance of amplitude coupling with phase coupling, we measured unidirectional amplitude coupling to make the dimensions of NLR and PLV feature vectors equal.

Coupling measures could be used for feature extraction because the coupling strength is different for different mental tasks. Figure 5 illustrates the averaged NLR coefficients and PLVs over all trials of each task for subject S3 based on the three coupling methods. We can see from this figure that both NLR vectors and PLV vectors of the two tasks evidently differ in most dimensions for each coupling method. The corresponding relationship of each feature (dimension) with electrode locations for the three types of couplings is listed in table 1.

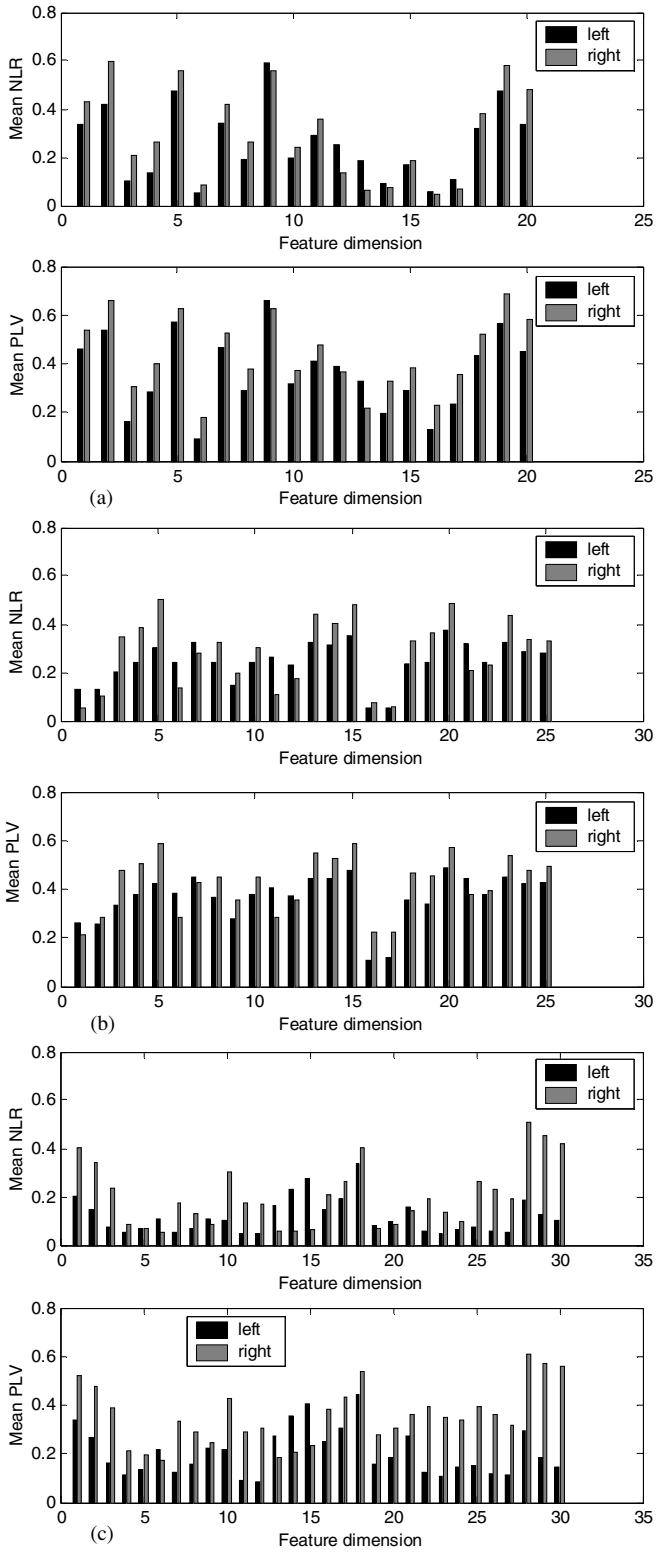

Figure 5. Averaged feature vectors over all trials of each task derived from the NLR coefficient and PLV for subject S3 based on three coupling methods: (a) coupling method CW, (b) coupling method CB1, (c) coupling method CB2.

2.4.2. Features extracted from the AR model. When used as classification features in BCIs, AR coefficients reflect the spectral distribution of the most prominent brain rhythms. In an AR model of order p, each data point of a time series is denoted as a linear combination of the last p data point. The Burg algorithm was utilized to estimate AR coefficients. Compared to other methods such as Levinson–Durbin, the Burg algorithm is more accurate because it minimizes both forward and backward errors by using more data points. The model order was chosen as six because the AR process of such an order is proved most suitable for mental task classification (Anderson *et al* 1998, Keirn and Aunon 1990).

Unlike the NLR coefficients and PLVs, AR coefficients are computed from single EEG signals. The most useful electrodes for differentiating between left and right motor imageries are located in SM1. To compare the performance of AR features with that of coupling features, we selected the same ten electrodes as used for coupling measures in SM1. The AR coefficients from all ten electrodes were concatenated to form a 60-dimensional feature vector.

Table 2. The value of r^2 for each subject and for each feature derived by reducing each coupling feature vector to one dimension. r^2 is the proportion of the total variance in the signal feature that is accounted for by the user's intent, and thus its values reflect the correlation between features or feature vectors and left and right motor imagery tasks.

	CW		CB1		CB2	
Subject	NLR	PLV	NLR	PLV	NLR	PLV
S1	0.815	0.772	0.793	0.742	0.863	0.821
S2	0.543	0.539	0.533	0.544	0.656	0.677
S3	0.829	0.792	0.779	0.766	0.853	0.791
S4	0.562	0.474	0.341	0.279	0.582	0.609
S5	0.955	0.913	0.891	0.840	0.767	0.833
All	0.741	0.698	0.667	0.634	0.744	0.746

The values show a significance at the $P < 0.001$ level.

2.5. Classification

The classifier used for single feature vectors is the Fisher discriminant analysis (FDA) (Webb 2002). FDA is a linear discriminant analysis whose purpose is to project data from high dimension onto a line so that the new data in one-dimensional space are more manageable. To obtain good separation of the projected data, FDA maximizes the difference of the sample means between two classes and minimizes the total within-class variance of the projected samples. When two feature vectors are combined for classification, they are first reduced to one dimension by FDA, and then a linear support vector machine (SVM) is employed as the classifier. In contrast to FDA, SVMs separate data by trying to find such an optimal hyperplane that maximizes the margin between the nearest samples of two categories (Vapnik 1995).

Linear classifiers are less prone to overfitting than their nonlinear counterparts when limited samples are available (Muller *et al* 2003). Linear classifiers are preferred also because of their low computational complexity that is desired especially in the case of online application.

To assess classification performance, the generalization accuracy is estimated by a 10 × 10-fold cross validation. Specifically, the original training set is randomly permuted

Table 3. The classification accuracy and standard deviation (%, rounded) of 10 × 10-fold cross validation derived from NLR-, PLV- and AR-based feature vectors for five subjects. The NLR- and PLV-based feature vectors are obtained from three coupling methods. The best result of each subject and their average for the six coupling feature vectors are set in bold face.

	Subject					
Feature vector	S1	S2	S3	S4	S5	Average
NLR-CW	96.1 ± 3.6	84.7 ± 7.9	**98.7 ± 2.0**	84.0 ± 6.9	**99.6 ± 1.3**	92.6 ± 4.3
NLR-CB1	94.3 ± 4.4	82.2 ± 6.4	94.3 ± 4.9	71.7 ± 8.6	99.5 ± 1.4	88.4 ± 5.1
NLR-CB2	**97.3 ± 3.1**	89.3 ± 6.8	97.4 ± 3.2	**84.4 ± 7.2**	94.5 ± 4.2	92.8 ± 4.9
PLV-CW	92.3 ± 5.7	84.7 ± 5.8	95.2 ± 3.8	79.2 ± 7.7	**99.6 ± 1.3**	90.2 ± 4.9
PLV-CB1	92.5 ± 5.2	85.8 ± 6.5	93.9 ± 4.3	67.0 ± 8.9	97.7 ± 2.8	87.4 ± 5.5
PLV-CB2	95.7 ± 4.1	**89.6 ± 6.2**	98.0 ± 2.9	84.2 ± 8.2	97.2 ± 3.2	**92.9 ± 4.9**
AR	95.8 ± 4.0	85.4 ± 7.9	98.7 ± 2.3	87.1 ± 5.5	99.6 ± 1.3	93.3 ± 4.2

for ten times and each time the randomly permuted data are split into ten parts: each part is used for testing and the remaining parts are used for training the classifier. This cross validation procedure leads to 100 classification tests, and the generalization accuracy is decided by averaging the 100 test accuracies.

3. Results

To see the coupling correlation between coupling features and the task from all users, each feature vector was reduced to one dimension using FDA, and a statistical measure r^2 (Winer *et al* 1991), the proportion of the total variance in the signal feature that is accounted for by the user's intent, was conducted on the basis of these dimension-reduced features. The values of r^2 are listed in table 2. An analysis of variance (ANOVA) revealed that the significance level is all at $P < 0.001$. Among the three coupling methods, the CB2 achieves the best results (largest values of r^2), while the CB1 is the poorest. As to the two measures, there is no distinct difference.

Table 3 shows the classification accuracy and standard deviation (%, rounded) of 10 × 10-fold cross validation derived from NLR-, PLV- and AR-based feature vectors for five subjects. The NLR- and PLV-based feature vectors are obtained from three coupling methods. The averaged classification accuracies of the five subjects range from 87.4% to 92.9% for the six feature vectors and the best classification accuracies of the six feature vectors range between 84.4% and 99.6% for the five subjects. These results indicate that coupling measures are appropriate feature extraction methods for differentiating two brain tasks in BCIs.

Among the three coupling methods, the best averaged classification accuracies are achieved by the coupling method CB2 for amplitude and phase coupling measures (92.8% and 92.9% respectively). Note that CB2 utilizes three more electrodes than the other two coupling methods. With the same ten electrodes, the averaged classification accuracy of the coupling method CW is higher than that of CB1 for amplitude and phase coupling measures (i.e. NLR and PLV). This means that the coupling measures obtained from CW can separate these two classes of imagined movements more easily than CB1.

As for the two methods for coupling measurement, the averaged classification accuracies of NLR coefficients over five subjects are slightly better than those of PLV for CW and CB1, but they do not have statistical distinction for CB2.

Table 4. The classification accuracy and standard deviation (%, rounded) of 10 × 10-fold cross validation derived from the combinations of AR-based feature vector with NLR- or PLV-based feature vectors for three subjects. The NLR- and PLV-based feature vectors are obtained from three coupling methods.

	Subject		
Feature combination	S1	S2	S4
AR+NLR-CW	99.0 ± 1.8	89.1 ± 6.0	94.5 ± 5.1
AR+NLR-CB1	99.1 ± 1.7	89.0 ± 5.4	90.1 ± 5.9
AR+NLR-CB2	99.3 ± 1.4	90.7 ± 6.0	94.2 ± 4.6
AR+PLV-CW	98.6 ± 2.7	89.0 ± 5.9	92.3 ± 5.5
AR+PLV-CB1	98.0 ± 2.6	89.9 ± 5.7	89.8 ± 6.1
AR+PLV-CB2	99.2 ± 1.7	91.9 ± 5.2	93.5 ± 5.0

When only coupling methods CW and CB1 are considered (i.e. the same ten electrodes in SM1 as the estimation of AR coefficients are exploited), the best averaged classification accuracy of coupling measures (92.6%) is a little lower than that of AR coefficients (93.3%). In terms of individual subjects, however, the best classification accuracy of coupling measures is equal to or higher than that of the AR coefficient for three subjects (96.1% for S1, 98.7% for S3 and 99.6% for S5). This suggests that the classification performances of coupling features and AR features are comparable; they do not exhibit a clear distinction.

Since coupling measures and AR coefficient reflect different electrical physiological processes, the features extracted from them are at least partially independent. Their combination can provide complementary information and consequently increase classification accuracy. The classification accuracy and standard deviation (%, rounded) of 10 × 10-fold cross validation derived from the combinations of AR-based feature vector with NLR- or PLV-based feature vectors for three subjects (S1, S2 and S4) are listed in table 4. The NLR- and PLV-based feature vectors are obtained from three coupling methods. For each measuring method, each coupling method and each subject, the classification accuracy of feature combination is clearly higher than that of the best single feature vector. Especially for subject S4, the maximal gain of 7.4% was achieved when the feature vectors AR and NLR-CW were combined.

Figure 6 illustrates one feature vector distribution of training and test set derived from 10 × 10-fold cross validation

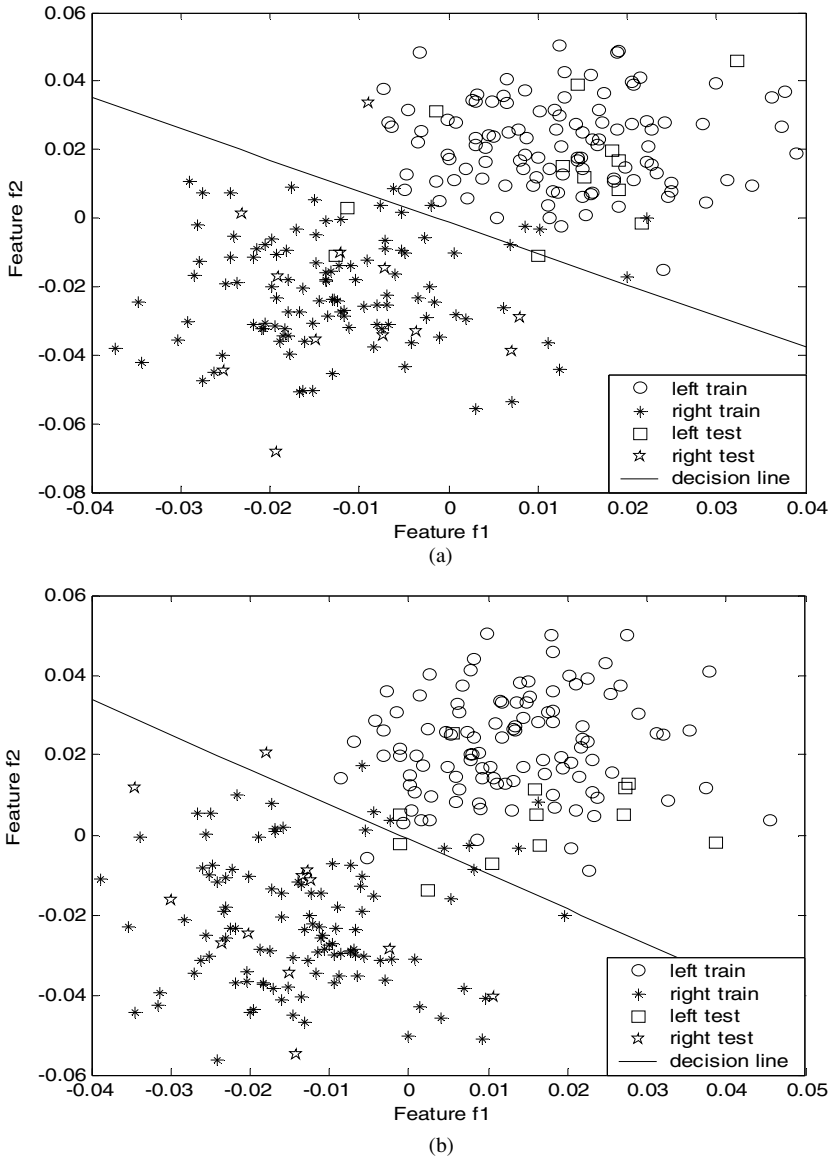

Figure 6. One feature vector distribution of training and test sets derived from 10 × 10-fold cross validation for subject S2 and the corresponding decision line for two methods of feature combination: (a) features f1 and f2 are obtained from AR and NLR coefficients respectively, (b) features f1 and f2 are obtained from AR coefficients and PLV respectively. Each decision line is determined by an SVM classifier derived from training data. The NLR coefficients and PLV are calculated by the coupling method CB2.

for subject S2 and the corresponding decision line for two methods of feature combination: (a) features f1 and f2 are obtained from AR and NLR coefficients respectively and (b) features f1 and f2 are obtained from AR coefficients and PLV respectively. Each decision line is determined by an SVM classifier derived from the training data. The NLR coefficients and PLV are calculated by the coupling method CB2. Obviously, if only one feature vector (f1 or f2) is used for classification, the decision line will be perpendicular or parallel to the abscissa axis. No matter what position it is located at, greater classification errors will result. The combination of two features f1 and f2, however, makes the decision line deflect toward the best position for separating two classes of samples and, subsequently, the classification accuracy will be improved.

4. Discussion

The method of feature extraction is of great importance in the recognition of mental tasks of a BCI system. The classification performance mainly depends on the quality of features or feature vectors used. This paper presents a study on the use of coupling measures as classification features in an EEG-based BCI during motor imagery. The performance of NLR coefficients and PLVs for classifying two mental tasks of motor imagery is evaluated and compared with that of AR coefficients.

The averaged classification accuracies of the five subjects for six different coupling vectors are between 87.4% and 92.9%. This was achieved in the case that only a small number of recording electrodes and training samples were exploited.

The averaged classification accuracy of five subjects for the coupling feature NLR-CW is 92.6%, only 0.7% less than that for the AR feature. There is no statistical difference between them. Hence, the classification performance of coupling measures could be said to be satisfactory. Although the classification performance of PLV-based features is slightly poorer than that of NLR-based features for the two coupling methods CW and CB1, the computation of PLV is very fast. This is an important asset in the framework of BCIs, especially for online application.

The classification results gave further evidence that short-range and long-range couplings exist, and subsequent measures can be utilized for classification features in an EEG-based BCI during motor imagery. Short-range coupling occurs between electrodes within regions, whereas long-range coupling exists between regions which are a bit widely separated. Some authors argued that whereas long-range couplings do reflect cognitive processing, short-range couplings might be caused by the volume conduction of neighboring electrodes (Lachaux *et al* 1999). Since the electrode interval in this study is larger than 2 cm, the electrode couplings within regions are basically due to cognitive tasks.

The good classification performance of feature vectors NLR-CB2 and PLV-CB2 (see table 3) demonstrated that the long-range coupling between SMA and SM1 provided discriminatory information for different tasks. SMA has long been recognized to play an important role in the planning or preparation of a movement behavior (Dechent *et al* 2004, Roland *et al* 1980), whereas its value in motor imagery-based BCIs was rarely explored. SMA activation and its coupling with SM1 related to two motor imagery tasks were validated by the classification results in this study.

When amplitude and phase coupling measures are used for feature extraction, the neurophysiological *a priori* knowledge could help us to determine the functional coupling areas associated with specific mental tasks. Without considering the coupling regions, high-dimensional feature vectors would result and subsequently, an intricate observation would arise. In contrast, the method in this study for the selection of coupling regions and electrodes of interest has the benefits of convenient electrode preparation, fast data processing and high feasibility for application.

In general, two steps cannot be absent from coupling measures, i.e. data intercepting and statistical significance test. For the analysis of a large amount of continuous data, a moving-window technique is needed to intercept the data into multiple segments in order to ensure that each segment of data is approximately stationary. In a BCI, the data of single trials are a few seconds in length (5 s in this study) and can be considered as quasi-stationary. As a result, it is not necessary to window the single-trial data. The purpose of the statistical significance test is to differentiate significant couplings against background fluctuations. The results from all subjects indicate that the coupling correlation between features and the two imagined hand movements exists and the significance level is high.

The number of subjects is relatively limited, and the validity of the methods should be tested on data sets from more extensive subjects. However, the subjects used in this study were randomly selected from the students in the department. Thus, this study, to some extent, is reasonable.

Future work is to carefully pick out the electrode pairs with a significant difference of coupling strength between different mental tasks on the basis of individuals and test the performance of current classification algorithms in online experiments.

5. Conclusion

Although the technology of BCIs developed rapidly, they still have a long way to go before practical application. Coupling measures between two EEG signals subserve a further understanding of the neuro physiological mechanisms underlying cognitive activities, whereas the finding of most discriminatory features is the pivotal work of BCI research. This study demonstrated that coupling measures are appropriate methods for feature extraction, and the combination of coupling features and AR feature could effectively improve classification performance in an EEG-based BCI during motor imagery.

Acknowledgments

The project is supported by National Natural Science Foundation of China (no. 60318001), Beijing Natural Science Foundation (no. 3051001).

References

Anderson C W, Stolz E A and Shamsunder S 1998 Multivariate autoregressive models for classification of spontaneous electroencephalogram during mental tasks *IEEE Trans. Biomed. Eng.* **45** 277–86

Andrew C and Pfurtscheller G 1996 Event-related coherence as a tool for studying dynamic integration of brain regions *Electroencephalogr. Clin. Neurophysiol.* **98** 144–8

Bartolomei F, Wendling F, Bellanger J J, Regis J and Chauvel P 2001 Neural networks involving the medial temporal structures in temporal lobe epilepsy *Clin. Neurophysiol.* **112** 1746–60

Boashash B 1992 Estimating and interpreting the instantaneous frequency of a signal: I. Fundamentals *Proc. IEEE* **80** 520–38

Brunner C, Graimann B, Huggins J E, Levine S P and Pfurtscheller G 2005 Phase relationships between different subdural electrode recordings in man *Neurosci. Lett.* **375** 69–74

Chavez M, Le Van Quyen M, Navarro V, Baulac M and Martinerie J 2003 Spatio-temporal dynamics prior to neocortical seizures: amplitude versus phase couplings *IEEE Trans. Biomed. Eng.* **50** 571–83

Dechent P, Merbodt K D and Frahm J 2004 Is the human primary motor cortex involved in motor imagery? *Cogn. Brain Res.* **19** 138–44

Dornhege G, Blankertz B, Curio G and Muller K-R 2004 Boosting bit rates in noninvasive EEG single-trial classification by feature combination and multiclass paradigms *IEEE Trans. Biomed. Eng.* **51** 993–1002

Fernandes de Lima V M, Pijn J P, Nunes C and Lopes da Silva F 1990 The role of hippocampal commisures in the interhemispheric transfer of epileptiform after discharges in the rat: a study using linear and nonlinear regression analysis *Electroencephalogr. Clin. Neurophysiol.* **76** 520–39

Friston K J, Stephan K M and Fraukowiak R S J 1997 Transient phase locking and dynamic correlations: are they the same thing? *Human Brain Mapp*. **5** 48–57

Gerloff C, Richard J, Hadley J, Schulman A E, Honda M and Hallett M 1998 Functional coupling and regional activation of human cortical motor areas during simple, internally paced and externally paced finger movements *Brain* **121** 1513–31

Gersetein G L, Perkel D H and Subramanian K N 1978 Identification of functional related neural assemblies *Brain Res*. **140** 43–62

Gotman J 1987 Interhemispheric interactions in seizures of focal onset: data from human intracranial recordings *Electroencephalogr. Clin. Neurophysiol*. **67** 120–33

Gysels E and Celka P 2004 Phase synchronization for the recognition of mental tasks in a brain-computer interface *IEEE Trans. Neural Syst. Rehabil. Eng*. **12** 406–15

Huan N J and Palaniappan R 2004 Neural network classification of autoregressive features from electroencephalogram signals for brain-computer interface design *J. Neural Eng*. **1** 142–50

Keirn Z A and Aunon J I 1990 A new mode of communication between man and his surroundings *IEEE Trans. Biomed. Eng*. **37** 1209–14

Lachaux J P, Rodriguez E, Martinerie J and Varela F J 1999 Measuring phase synchrony in brain signals *Human Brain Mapp*. **8** 194–208

Le Van Quyen M, Adam C, Baulac M, Martinerie J and Varela F 1998 Nonlinear interdependencies of EEG signals in human intracranially recorded temporal lobe seizures *Brain Res*. **792** 24–40

Le Van Quyen M, Foucher J, Lauchaux J P, Rodriguez E, Lutz A, Martinerie J and Varela F 2001 Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony *J. Neurosci. Methods* **111** 83–98

Lopes da Silva F, Pijn J P and Boeininga P 1989 Interdependence of EEG signals: linear vs. nonlinear associations and the significance of time delays and phase shifts *Brain Topogr*. **2** 9–18

Mormann F, Lehnertz K, David P and Elger C E 2000 Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients *Physica* D **144** 358–69

Muller-Gerking J, Pfurtschller G and Flyvbjerg H 1999 Designing optimal filters for single-trial EEG classification in a movement task *Clin. Neurophysiol*. **110** 787–98

Muller K R, Anderson C W and Birch G E 2003 Linear and nonlinear methods for brain-computer interfaces *IEEE Trans. Neural Syst. Rehabil. Eng*. **11** 165–9

Pfurtscheller G and Lopes da Silva F H 1999 Event-related EEG/EMG synchronization and desynchronization: basic principles *Clin. Neurophysiol*. **110** 1842–57

Pfurtscheller G and Neuper C 2001 Motor imagery and direct brain-computer communication *Proc. IEEE* **89** 1123–34

Pfurtscheller G, Neuper C, Guger C, Harkam W, Ramoser H, Schlogl A, Obermaier B and Pregenzer M 2000 Current trends in Graz brain-computer interface (BCI) research *IEEE Trans. Rehabil. Eng*. **8** 216–9

Ramoser H, Muller-Gerking J and Pfurtschller G 2000 Optimal spatial filtering of single-trial EEG during imagined hand movement *IEEE Trans. Rehabil. Eng*. **8** 441–6

Rodriguez E, George N, Lachaus J, Marttinerie J, Renault B and Varela F 1999 Perception's shadow: long-distance synchronization of human brain activity *Nature* **397** 430–3

Roland P E, Larsen B, Lassen N A and Skinhoj E 1980 Supplementary motor area and other cortical areas in organization of voluntary movements in man *J. Neurophysiol*. **43** 118–36

Spiegler A, Graimann B and Pfurtscheller G 2004 Phase coupling between different motor areas during tongue-movement imagery *Neurosci. Lett*. **369** 50–4

Tass P, Rosenblum M G, Weule J, Kurths J, Pikovsky A, Volkmann J, Schnitzler A and Fruend H J 1998 Detection of *n*:*m* phase locking from noisy data: application to magnetoencephalography *Phys. Rev. Lett*. **18** 3291–94

Tononi G and Eelman G M 1998 Consciousness and complexity *Science* **282** 1846–51

Vapnik V N 1995 *The Nature of Statistical Learning Theory* (New York: Springer)

Varela F J 1995 Resonant cell assemblies: a new approach to cognitive functions and neuronal synchrony *Biol. Res*. **28** 81–95

Varela F J, Lachaux J P, Rodriguez E and Martinerie J 2001 The brainweb: phase synchronization and large-scale integration *Nat. Rev. Neurosci*. **2** 229–39

Wang Y, Hong B, Gao X and Gao S 2006 Phase synchrony measurement in motor cortex for classifying single-trial EEG during motor imagery *Proc. 28th IEEE EMBS Annual Int. Conf. (New York)* pp 75–8

Webb A R 2002 *Statistical Pattern Recognition* (New York: Wiley)

Wei Q, Gao X and Gao S 2006 Feature extraction and subset selection for classifying single-trial ECoG during motor imagery *Proc. 28th IEEE EMBS Annual Int. Conf. (New York)* pp 1589–92

Winer B J, Brown D R and Michels K M 1991 *Statistical Principles in Experiment Design* (New York: McGraw-Hill)

Wolpaw J R, Birbaumer N, McFarland D J, Pfurtscheller G and Vaughan T M 2002 Brain-computer interface for communication and control *Clin. Neurophysiol*. **113** 767–91

Zaveri H P, Williams W J, Sackellares J C, Beydoun A, Duckrow R B and Spencer S S 1999 Measuring the coherence of intracranial electroencephalograms *Clin. Neurophysiol*. **110** 1717–25