Abstract

Classification of multichannel EEG recordings during motor imagination has been exploited successfully for brain-computer interfaces (BCI). In this paper, we consider EEG signals as the outputs of a networked dynamical system (the cortex), and exploit novel features from the collective dynamics of the system for classification. Herein, we also propose a new framework for learning optimal filters automatically from the data, by employing a Fisher ratio criterion. Experimental evaluations comparing the proposed dynamical system features with the CSP and the AR features reveal their competitive performance during classification. Results also show the benefits of employing the spatial and the temporal filters optimized using the proposed learning approach.

1. Introduction

A brain-computer interface (BCI) is a communication system that relies on the brain rather than the body for control and feedback (Wolpaw et al., 2002). Ideally, it should run in a servo mode, allowing the subjects to initiate the communication anytime and anywhere without resorting to external stimuli or triggers. Such an interface not only offers a promising prosthetic device for those severely paralyzed but also signifies a radically new technology for the general public. Current BCI research is still in its early stage and the emphasis is placed on the design of algorithms to decode a pre-specified set of brain states. This involves three main aspects:

Brain states. Only brain states consciously controllable by the subjects are suitable for BCI. Besides, these states should generate distinct, repeatable and measurable patterns whenever accessed. Among the most commonly used brain states are imaginations of body movements (motor imaginations). Motor imaginations can reliably change the neural activities over sensorimotor cortices. Depending on the part of the body imagined moving, these changes exhibit distinct spatial distributions (Pfurtscheller & Lopes da Silva, 1999). Recognition of these patterns can then be translated into control signals, as is the case in this study.

Recording devices. Motor imaginations can be recorded by both electroencephalography (EEG) and magnetoencephalography (MEG). EEG remains the most popular way to record BCI signals, and will be the focus of this study. It measures scalp electrical activities diffused from the cortex. Compared to MEG, it is portable and inexpensive. However, EEG can only measure blurred cortical activities due to the diffusion of the skull and the skin. Thus EEG is normally used for studying cortical patches in the centimeter scale. Furthermore, EEG signals are contaminated by noise from various sources, such as muscle activities and power line interference. Spatial and temporal filters are commonly applied before any further analysis (Perrin et al., 1989; Dornhege et al., 2006).

Decoding algorithms. Pre-filtered EEG signals still contain considerable noise, which poses a challenge for its decoding. Statistical machine learning (ML) techniques have been introduced into BCI to combat these variations. Techniques like Artificial Neural Networks, Support Vector Machine (SVM) (Müller et al., 2003) and Linear Discriminant Analysis (Dornhege et al., 2006) have been employed to learn patterns from training EEG signals and then classify new EEG signals. This strategy often results in increased decoding success and significant shortening of subject training time.
(from several months down to several days). The most prominent examples include the Berlin BCI (Dornhege et al., 2006), the MPI BCI (Lal et al., 2005) and the Graz BCI (Ramoser et al., 2000).

Apart from the classifiers, these ML-based BCIs also differ in the features they extract from EEG signals. The most successfully used features include autoregressive (AR) coefficients (Dornhege et al., 2004; Lal et al., 2005) and common spatial patterns (CSP) (Ramoser et al., 2000; Dornhege et al., 2006). In this paper, we will employ a novel type of feature based explicitly on the neurophysiology of EEG signals instead. Basically, we consider EEG signals as the outputs of a networked dynamical system. The nodes of this system consist of cortical patches, while the links correspond to neural fibers. A large and complex system like this often generates interesting collective dynamics, such as synchronization in the activities of the nodes, and they result in the change of EEG patterns measured on the scalp. These features from the collective dynamics of the system can be employed for classification (Gysels & Celka, 2004; Song et al., 2006). This will be elaborated in section 2.

To recover the cortical dynamics from the EEG signals, subject-specific spatial and temporal filtering is usually needed (Dornhege et al., 2006; Song & Epps, 2006). Instead of manually tuning these filters, we propose a common framework in section 3 to learn them from the data. Our basic idea is to optimize the filters so that the separability of the two classes is improved. Experimental results show that the learned filters not only reduce the classification errors of the DS features but also extract physically meaningful information from the EEG signals. Comparisons are also made between the DS features with the learned filters and the CSP and the AR features with manually tuned filters. These comparisons together with further comparisons to other filter learning methods, such as the CSSP (Lemm et al., 2005) and CSSSP (Dornhege et al., 2006) method, demonstrate the competitive performance of our method (section 4). Finally, the conclusion is given in section 5.

2. Dynamical System Features

The cortex is a highly folded sheet of neurons (≈100 billion neurons) and they self-organize into clusters. These neuronal clusters not only tightly connect with their neighbors but also communicate with distal clusters through neural fibers. Each cluster is often associated with certain aspect of information processing. The collaboration of these clusters achieves the normal functioning of the brain. In this section, we will first describe a simple mathematical model of the cortex, and then show how it leads to dynamical system features related to motor imaginations.

2.1. Mathematical Model of the Cortex

Typically, a neuronal cluster will generate electrical oscillations. It has been modeled as an oscillator with phase θ and output s. Its dynamics are governed by a simple phase model (Pikovsky et al., 2003):

\[
\begin{align*}
\dot{s} &= f(\theta) \\
\dot{\theta} &= \omega + g(t),
\end{align*}
\]

where ω is the intrinsic frequency of the oscillation and f is a function 2π-periodic in θ. g(t) is the input to the oscillator. g(t) will accelerate the oscillation if it assumes positive values, and slow it down if negative.

The whole cortex can then be modeled as a networked dynamical system \(\mathbb{D}\), as shown in Figure 1. Each node in the system represents a neuronal cluster and each link a neural interaction. The input, \(g(t)\), to each neuronal cluster now consists of two parts: influence from other clusters and modulation by subcortical structures (Pfurtscheller & Lopes da Silva, 1999). Suppose the links of the network are represented as an adjacency matrix \(G\) (\(G_{ij} = 1\) if node \(i\) and \(j\) are connected; \(G_{ij} = 0\) otherwise). Then the dynamics of a node \(i\) take a more specific form:

\[
\dot{\theta}_i = \omega_i + \sum_j \epsilon_{ij} G_{ij}(s_j - s_i) + h_i(t),
\]

where \(\sum_j \epsilon_{ij} G_{ij}(s_j - s_i)\) represents the influence from other nodes, and \(h_i(t)\) is the subcortical input. Note that there is an added parameter \(\epsilon_{ij}\) in (2), which controls the strength of the influence from node \(j\) to \(i\).

2.2. Desynchronization of Neuronal Clusters

Two properties of the network of oscillators in (2) are of particular interest to BCI (Pikovsky et al., 2003): (i) Without the input \(h(t)\), all nodes will settle down into an oscillation of the same frequency \(\omega_0\), if the network is connected and the influence \(\epsilon\) is sufficiently strong (mutual synchronization). (ii) If the input \(h_i(t)\) to node \(i\) is sufficiently strong and oscillates at a fre-

Figure 1: Networked dynamical system model of the cortex.
frequency $\omega_0$, node $i$ will then be forced to oscillate in the same frequency $\omega_0$ (forced synchronization).

These two properties explain well the spatial distribution of the EEG signals during motor imaginations (Pfurtscheller & Lopes da Silva, 1999): (i) If no imagination is carried out, the neuronal clusters in the idle sensorimotor cortex tend to synchronize with each other and oscillate in the frequency range of 8-26 Hz (EEG $\alpha$ and $\beta$ rhythm). The spatial summation of this unison is a strong $\alpha$ (and/or $\beta$) rhythm in EEG signals. (ii) If the subject is actively engaged in motor imaginations, the associated neuronal clusters will be strongly modulated by the subcortical structures. The dynamics of these clusters will then stray away from their former synchronous state. This results in a decrease of $\alpha$ (and/or $\beta$) power in EEG signals.

This phenomenon is called Event-Related Desynchronization (ERD) in the neuroscience literature. Depending on the part of the body imagined moving, neuronal clusters at different locations will be active. These clusters desynchronize with other clusters, and the spatial distribution of the desynchronization will be different as the imagination contents change. ERD suggests that the strength of the synchronization between neuronal clusters can be used as features for classification (Gysels & Celka, 2004; Song et al., 2006).

### 2.3. Features for Motor Imaginations

An EEG electrode measures mostly the activities of the neuronal cluster directly underneath it (we will qualify this in section 3). Suppose the pairwise synchronization of the measured neuronal clusters can be computed from EEG signals and organized into a matrix $S$ ($S$ is symmetric with entry $S_{ij}$ for cluster $i$ and $j$). Each entry in $S$ is a dynamical system feature and the similarity between two EEG signals can then be quantified in terms of these features as:

$$K(S, \hat{S}) = \text{Tr}((S \circ A)^T (\hat{S} \circ A)), \quad (3)$$

where $A$ is a weighting matrix, $\text{Tr}(\cdot)$ computes the trace of a matrix, and $\circ$ represents element wise matrix product. Essentially, this measure transforms EEG trials into vectors in a new space and then computes the inner product in that space. The dimensions of this space consist of the synchronization values for each pair of EEG electrodes. Since we will use a SVM classifier for our later experiments, this measure can be interpreted as a kernel.

Preliminary analysis of our motor imagination data set (this data set are further explained in section 4) indicates that the synchronization in our data appears to be either in-phase ($\theta_i - \theta_j = 0$) or anti-phase ($\theta_i - \theta_j = \pi$). These two types of synchronization can be well detected simply using the covariance. Therefore, classifying EEG signals using the DS features consists of three steps: (i) filter EEG signals; (ii) compute the entries of $S$ using the covariance; and (iii) compute the kernel matrix according to (3) and pass it to SVM.

## 3. Learning Optimal Filters

Filtering EEG signals is important for later classifications. Due to the diffusion of the skull and skin, an EEG electrode actually measures a mixture of signals from several neuronal clusters. Spatial filters, such as a Laplacian filter, are usually applied to concentrate the signals to a single neuronal cluster. Furthermore, EEG signals are contaminated by various noise, such as electrical signals from muscle movements. Our interest lies in oscillation in the frequency range of 8-26 Hz ($\alpha$ and $\beta$ rhythm). Bandpass filtering is usually needed to suppress other signals.

As previous BCI researchers have experienced (Dornhege et al., 2006), the optimal filters for each subject are very different, and it is quite inconvenient to manually choose these filters. Attempts have been made to learn these filters from the training EEG data. Pioneering works has been reported in Lemm et al. (2005) and Dornhege et al. (2006), where FIR (temporal) filters are learned for the CSP features to improve the separability of the two classes. Our work is inspired by their ideas, but our approach are different in two aspects. First, our approach is directed to the dynamical system features. Second, we have proposed a common framework for the learning of both the spatial and the temporal filters. In the following subsections, the common framework is described before it is specialized into the spatial and the temporal filter learning.

### 3.1. Common Framework

Our filter learning framework involves three steps: (i) quantify the quality of a feature using the Fisher ratio; (ii) express the Fisher ratio using the filter parameters; (iii) and then maximize the Fisher ratio with respect to the filter parameters. Given the data and the filter parameter $a$, our framework can be formulated mathematically as:

$$\max_a Q(a) = \frac{(\mu_+ (a) - \mu_- (a))^2}{\sigma_+ (a)^2 + \sigma_- (a)^2}, \quad (4)$$

where $Q$ is the fisher ratio, $\mu$ the mean value of a feature and $\sigma$ its standard deviation (the subscripts $+$ and $-$ restrict computation to positive and negative class respectively). Higher values of $Q$ usually indicate better separation of the two classes. This learning
framework can be applied to various problems. However, only local optimum can be guaranteed for the solution, since \( Q \) is in general not convex in terms of \( a \). This is also the case in learning the filters for the DS features. To find an optimal solution efficiently, we will employ the subspace optimization technique.

The filter learning is performed on each pair of EEG electrodes separately. For a pair, two filters are learned, one for each electrode. Suppose the parameters of the two filters are \( a \) and \( b \) respectively. It turns out that for both the spatial and the temporal filtering, \( Q \) assumes a form bi-quadratic in \( a \) and \( b \).

More specifically, if \( b \) is fixed (for example), \( Q \) becomes the quotient between \( a^T V(b) a \) and \( a^T W(b) a \), where \( V(b) \) and \( W(b) \) are matrices quadratic in \( b \). The optimal \( a \) can then be obtained by solving the following constrained optimization problem:

\[
\max_a \quad a^T V(b) a \\
\text{s.t.} \quad a^T W(b) a + \gamma b^T b a^T a = c, \tag{5}
\]

Note that the additional term \( \gamma b^T b a^T a \) does not originate from \( Q \). It is a regularized product of the norms of \( a \) and \( b \), and the strength of this regularization is controlled by \( \gamma \).

Using the Lagrange multiplier method (let the multiplier be \( \lambda \)), the optimal \( a \) can be derived from the following generalized eigenvector problem:

\[
\tilde{V}(b) a = \lambda \tilde{W}(b) a, \tag{6}
\]

where

\[
\tilde{V}(b) = V(b) + V(b)^T, \\
\tilde{W}(b) = W(b) + W(b)^T + 2\gamma b^T b I. \tag{7}
\]

The optimal \( a \) is then the generalized eigenvector corresponding to the largest eigenvalue. Similarly, \( b \) can be optimized by fixing \( a \). Local maxima can then be found by optimizing \( a \) and \( b \) alternately (Algorithm 1). In our experiments, the solution usually changes very little after two iterations, and henceforth only two iterations are used. To specialize this algorithm into the learning of the spatial and the temporal filters, we only need to derive the exact forms of \( V(a), W(a), V(b) \) and \( W(b) \) for these two cases respectively.

### 3.2. Learning Spatial Filters

Studies show that the spherical spline Laplacian filter is useful for the study of cortical dynamics (Srinivasan et al., 1998). This method models the shape of the head as a unit sphere and uses orthogonal bases on the sphere to spatially interpolate EEG signals (Perrin et al., 1989). The filtering is then achieved by computing the analytical Laplacian of the interpolation function. This filter only high-passes EEG signals, and is unable to emphasize interesting signals in the middle frequency range (Song & Epps, 2006). This section will start with a reformulation of the spherical spline Laplacian filter, which leads to a class of spatial filters. The exact forms of \( V(a), W(a), V(b) \) and \( W(b) \) are then derived.

For square integrable functions on a sphere, the Legendre polynomials \( p_n(\cos \theta) \) evaluated at \( \cos \theta \) constitute a set of orthogonal bases. The parameter \( n \) is the degree of the polynomial and it controls the spatial frequency of a basis. A \( p_n \) with larger \( n \) will generally represent higher spatial frequency. \( \theta \) is the latitudinal (zonal) angle. In this study, a maximum of \( n = 20 \) is used for the interpolation of EEG signals (due to the low spatial variation of EEG signals).

Suppose a position on the unit sphere is \( e \), and the position of the \( i \)th EEG electrode is \( e_i \). Let \( \cos(e,e_i) \) denote the cosine of the angle between \( e \) and \( e_i \), we can construct a matrix \( P(e) \) with entries:

\[
(P(e))_{i,j} = \frac{1}{4\pi(n(n+1))^{2}} p_n(\cos(e,e_i)), \tag{8}
\]

where \( i \) ranges through the index of the electrodes, and \( n = 1 \ldots 20 \). Then EEG signals at position \( e \) can be interpolated as:

\[
u(e) = c_0 + C^T P(e) 1,
\]

where \( 1 \) is a vector of all ones. \( c_0 \) (a vector) and \( C \) (a matrix) are the interpolation coefficients estimated from actual EEG signals. The solution of these coefficients can be found using two constraints: (i) The interpolated function has to pass the actual EEG measurements; (ii) \( C^T . 1 = 0 \). Our formulation in equation (9) is equivalent to equation (1) in Perrin’s original formulation (Perrin et al., 1989). The difference is that equation (9) describes the interpolation of a time series rather than that of a single time point.
Spatial filtering of EEG signals can then be achieved by simply removing the DC component $c_0$ and re-weighting other frequency components (the bases). Suppose the filter (weighting) is $a$. Thus spatial filtering can be computed as:

$$\hat{u}(e_i) = C^TP(e_i)(1 \circ a) = C^TP(e_i)a. \quad (10)$$

The spherical spline Laplacian filter can be obtained by simply setting entries of $a$ to $-n(n+1)$ (equivalent to equation (5) in Perrin et al. (1989)). With formula (10), other types of filtering can also be implemented by varying $a$. For example, a bell-shaped bandpass filter can be obtained by setting the entries of $a$ to $-\exp(-\kappa n(n+1))n(n+1)$ ($\kappa$ is a parameter controlling the width and the peak).

Suppose filter $a$ and $b$ are applied to electrode $e_i$ and $e_j$ respectively, the covariance between the two filtered EEG signals can then be computed as:

$$cov_{ij} = \frac{1}{l} \hat{u}(e_i)^T\hat{u}(e_j) = \frac{1}{l} a^TP(e_i)CC^TP(e_j)b,$$

where $l$ is the number of time points. Further denote $\tilde{C}_{ij} = P(e_i)CC^TP(e_j)$ (Since the following derivations are the same for each pair of electrodes, the superscripts $ij$ are dropped henceforth for convenience.) Then $\mu_+$ in (4) can be computed as:

$$\mu_+ = \frac{1}{m} \sum_{k \in +} cov_k = a^T \left( \frac{1}{ml} \sum_{k \in +} \tilde{C}_k \right) b = a^T D_+ b, \quad (12)$$

where $k \in +$ means that the index ranges through all $m$ trials in the positive class (Suppose the negative class also has $m$ trials). The variance $\sigma_+$ can be computed as:

$$\sigma_+^2 = \frac{1}{m} \sum_{k \in +} (cov_k - \mu_+)^2 = a^T E_+ (b)a, \quad (13)$$

where

$$E_+(b) = \frac{1}{ml^2} \left( \sum_{k \in +} \tilde{C}_k \right) b^2 - \frac{1}{ml^2} \sum_{k \in +} \left( \tilde{C}_k b \right)^2. \quad (14)$$

Similarly $\mu_- = a^T D_- b$ and $\sigma_- = a^T E_- (b)a$. $V(b)$ and $W(b)$ can then be derived as:

$$\begin{align*}
(\mu_+ - \mu_-)^2 &= a^T (D_+ b - D_- b)^2 a = a^T V(b)a \\
(\sigma_+^2 + (\sigma_-)^2) &= a^T (E_+(b) + E_-(b))a = a^T W(b)a \quad (15)
\end{align*}$$

Since $a$ and $b$ are symmetric, $V(a)$ and $W(a)$ can be derived analogously by exchanging the positions of $a$ and $b$ and transposing $\tilde{C}_k$ in (12)-(15). Substituting $V(a)$, $W(a)$, $V(b)$, $W(b)$ into Algorithm 1 will then produce the optimal filters.

### 3.3. Learning Temporal Filters

Unlike Dornhege et al. (2006) who formulated the learning of the temporal filters in the time domain (FIR filter), our formulation works directly in the frequency domain. The basic idea of our approach is to place weighting directly on the complex coefficients of the discrete Fourier transformation (DFT). The computation of the covariance in our approach needs two forward and inverse DFT. However, it can be made more efficient using the Correlation Theorem.

Weighting the frequency components of an EEG signal $u(e_i)$ will transform it to:

$$\hat{u}(e_i) = F^{-1}(F(u(e_i)) \circ a), \quad (16)$$

where $a$ is the filter (weighting), and $F$ represents forward DFT ($F^{-1}$, the inverse DFT). Suppose filters $a$ and $b$ are applied to EEG electrodes $e_i$ and $e_j$ respectively. The covariance of the filtered signals can then be computed as:

$$cov = \frac{1}{l} a^T \hat{u}(e_i)^T \hat{u}(e_j)$$

$$= \frac{1}{l} \left( F^{-1}(F(u(e_i)) \circ a) \right)^T \left( F^{-1}(F(u(e_j)) \circ b) \right)$$

(17)

(Note that the superscripts are dropped for convenience.) Computation (17) is inefficient, since two forward and inverse DFT are needed. The computation, however, can be reduced using the Correlation Theorem. This theorem states that the covariance between two signals $u(e_i)$ and $u(e_j)$ is equal to $(F(u(e_i)))^* F(u(e_j))$ ($^*$ denotes conjugate transpose). Thus (17) can be simplified to:

$$cov = \frac{1}{l} a^T \left( (F(u(e_i)))^* \circ F(u(e_j)) \right) b, \quad (18)$$

where $Diag(\cdot)$ transforms its vector argument into a diagonal matrix. Formula (18) requires only two DFT computations and hence it is more efficient.

The derivations for $V(a)$, $W(a)$, $V(b)$ and $W(b)$ become straightforward, if we compare equation (18) with equation (11). By setting $\tilde{C}_{ij} = Diag ((F(u(e_i)))^* \circ F(u(e_j)))$, they can be obtained from (12)-(15). Substituting these matrices into Algorithm 1 produces the optimal filters.

### 4. Results and Comparison

The dynamical system (DS) features and the filter learning approach are evaluated using data set IVa from the Berlin BCI group (Dornhege et al., 2004). This data set contains EEG signals (118 channels, sampled at 100 Hz) for five healthy subjects (labeled
Classifying EEG for Brain-Computer Interfaces

‘aa’, ‘al’, ‘av’, ‘aw’ and ‘ay’ respectively). During the recordings, they were prompted by visual cues to imagine for 3.5 s either right hand (the positive class) or right foot movements (the negative class). Our analysis will be confined to the data between 0.5 s and 2.5 s (i.e. 200 time points each channel), since in an online BCI setting a sliding window seldom exceeds 2 s (Dornhege et al., 2006). Each type of imagination was carried out 140 times. Thus there are 280 labeled trials in total for each subject (note: each trial is a multivariate time series of size 118 × 200). The task is to classify the type of the imagination for each trial in an offline fashion.

Five experiments were conducted in our evaluations. We will first describe procedures common to these experiments. All our classifications are carried out using SVM and the errors are obtained from 10 × 10 fold cross-validation. An identical temporal pre-filtering (bandpass between 8-40 Hz) is applied to all subjects. In the case of the DS features, an identical spatial pre-filtering (a bell-shaped bandpass filter − exp(−κ(n + 1)n(n + 1)) with κ = 0.01) is also applied for all subjects. Furthermore, only the top 20 DS features (in terms of their Fisher ratios) are used for classification. This is equivalent to setting only 20 entries of $A$ in equation (3) to 1 and all others to 0.

4.1. Learned Filters Improve Classification

For each pair of EEG electrodes (equivalent to a DS feature), the optimal spatial and temporal filters were learned sequentially. In Table 1, we present the classification errors using: (i) the DS features without the spatial and the temporal filter optimization (DS column); (ii) the DS features only with the spatial filter optimization (DS+S column); (iii) the DS features only with the temporal filter optimization (DS+T column); (iv) the DS features with both the spatial and the temporal filter optimization (DS+S+T column).

The results demonstrate that both the learned spatial and temporal filters improve the classification (DS+S and DS+T columns). Although there is no absolute winner in the two types of filters, when applied separately, the temporal filters outperform the spatial filters in general (The winning filter for each subject is highlighted in bold). Especially for subject ‘aa’ and ‘aw’, the temporal filters reduce about 5% more errors than the spatial filters.

The combined application of the learned filters almost always further reduces the errors (only subject ‘av’ slightly violates this rule). The maximum reductions is around 7% (for subject ‘aa’ and ‘aw’). The errors obtained (DS+S+T column) are now lower than 10% for 4 of the 5 subjects (except ‘av’). It seems that the learned filters help less for some subjects (‘al’ and ‘ay’). The reason can be that the pre-filtering is already near the optimal solution.

The classification for subject ‘av’ has the largest error. Our preliminary studies indicate that the most responsive frequency range of this subject shifts above 26 Hz (contrary to the usual 8-26 Hz). While most energy in the EEG signals concentrates below 26 Hz, this makes it difficult to extract good features for the subject.

| Sb  | DS  | DS+S | DS+T | DS+S+T |
|-----|-----|------|------|--------|
| aa  | 16.7±7.2 | 14.6±7.0 | 9.7±5.7 | 9.5±5.7 |
| al  | 3.7±3.3  | 3.2±3.2   | 3.6±3.4  | 2.7±3.1  |
| av  | 27.3±7.9 | 25.1±8.0  | 21.4±7.9  | 21.5±7.6  |
| aw  | 13.1±6.0 | 12.1±5.7  | 7.5±4.4  | 6.2±4.5  |
| ay  | 11.0±5.3 | 9.6±5.0   | 9.7±5.1  | 8.5±5.0  |

4.2. Learned Filters Extract Meaningful Information

Several details related to section 4.1 are clarified here. The spatial and the temporal filters can be interpreted as weighting in the corresponding frequency domain. We have further restricted them to be polynomial models in our experiments. The results in Table 1 are obtained with polynomial functions of degree 6 (for both the spatial and the temporal filter learning). The regularization parameters $\gamma$ for the spatial and the temporal filters are $10^{-7}$ and $10^{-13}$ respectively. For the case of the temporal filter, a bell-shaped pre-filter is also applied (− exp(−κ(n + 1)n(n + 1)) with κ = 0.001 for all subjects). Since a DS feature is bilinear in the filters applied to the two channels, our optimization in Algorithm 1 only has exact control over their multiplicative effect. It is expected that the multiplication of the two learned filters will be more meaningful than any individual one.

Figure 2(a) shows a learned spatial filter (thin line, bow-shaped) and the pre-filter (thin line, bell-shaped) for one channel. Although both filters are simple, their multiplicative effect creates a double-peak characteristics (dotted line). This is equivalent to emphasizing the frequency contributions under these two peaks. The overall effect of the learned filters from two channels (dotted lines in Figure 2(b)) is also double-peaked (thick line in Figure 2(b)). We believe that these peaks are somehow related to the electrode spacing on the scalp. It is likely that the learned filters weight the
information from the actual electrodes more heavily than that from the interpolated positions.

For the temporal filters, we will interpret the learned filters in terms of their effects on the power spectrum. Hence only the absolute values of the weighting are displayed. The final filter for an example channel (dotted line in Figure 3(a); it is the multiplication of a pre-filter and a learned filter, both in thin lines) does not appear to emphasize the motor imagination signals (i.e. ERD in 8-26 Hz). The meaning, however, becomes clearer when we examine the filters from two channels together. In Figure 3(b), the filters from two channels are shown in dotted lines and their multiplication in thick line. The multiplication creates the strongest peak within 10-18 Hz, and a second strongest peak within 18-28 Hz. This corresponds well to the most responsive frequency range of the motor imaginations.

4.3. Dynamical System Features are Competitive

The DS features obtained with learned filters were compared to the CSP and the AR features obtained with manually chosen parameters. The parameters for the CSP features (filtering frequency, selected channels and the number of projection subspaces) and the AR features (filtering frequency, selected channels and the order of the AR model) were tuned according to the winning entry of BCI competition III (Wang et al., 2005). The results are shown in Table 2.

Overall, the CSP features perform the best, the DS features follow, and the AR features produce lower accuracy. Furthermore, the DS features always obtain either the best (highlighted in bold) or the second best place (highlighted in italic). Especially for subject ‘av’, the DS features outperform the CSP features by 6%. It is important to note that the parameters for the CSP and AR features have been tuned manually and intensively, while the results for the DS features are obtained with exactly the same starting parameters. This shows the usefulness the DS features and our filter learning approach.

Table 2: Classification errors (%) of the CSP, the AR and the DS features with optimized filters.

| Sb  | CSP+AR | CSP+DS | AR+DS | ALL |
|-----|--------|--------|-------|-----|
| aa  | 7.6±5.0| 7.3±5.1| 7.7±4.7| 7.3±4.9|
| al  | 1.6±2.3| 0.9±1.9| 1.6±2.5| 1.5±2.2|
| av  | 22.3±7.4| 22.5±7.8| 21.4±7.4| 21.6±7.1|
| aw  | 3.5±3.2| 2.8±3.1| 5.2±3.8| 3.4±3.2|
| ay  | 8.9±4.6| 5.5±4.3| 9.1±4.6| 8.7±4.5|

4.4. Dynamical System Features Extract Complementary Information

The CSP, AR and DS features are computed differently from the EEG signals. An interesting question is whether they complement each other during classification. To investigate this, we combine more than two types of features (CSP+AR, CSP+DS, AR+DS and ALL three) using the META scheme described by Dornhege et al. (2004). The classifications of the combined features are presented in Table 3. The combination with the smallest error for each subject is highlighted in bold and the second place in italic. Furthermore, we surround an error with a box, if it is the smallest ever (in Table 2 and 3) for a subject.

Table 3: Classification errors (%) of the combinations of the CSP, the AR and the DS features.

| Sb  | CSP+AR | CSP+DS | AR+DS | ALL |
|-----|--------|--------|-------|-----|
| aa  | 7.6±5.0| 7.3±5.1| 7.7±4.7| 7.3±4.9|
| al  | 1.6±2.3| 0.9±1.9| 1.6±2.5| 1.5±2.2|
| av  | 22.3±7.4| 22.5±7.8| 21.4±7.4| 21.6±7.1|
| aw  | 3.5±3.2| 2.8±3.1| 5.2±3.8| 3.4±3.2|
| ay  | 8.9±4.6| 5.5±4.3| 9.1±4.6| 8.7±4.5|

The DS features indeed complement the CSP and the AR features, as is evidenced by the further reduction of errors in subject ‘aa’, ‘av’ and ‘aw’. The reduction, however, is not large (the largest being around 1% for subject ‘aa’). Furthermore, the combination of all three types of features does not necessarily further reduce the errors. This happens when the best features have already extracted almost all information about
the separability of the two classes. Additional features may only provide redundant or even conflicting information for the classification. This is very likely in our case since we have optimized each type of feature intensively. Finally, our results suggest that the combination of the CSP and the DS features performs the best, and the DS features complement the CSP features better than the AR features.

4.5. Our Learning Framework is Competitive

The DS features obtained with the learned filters were compared to the CSP features produced by the CSSP (Lemm et al., 2005) and the CSSSP (Dornhege et al., 2006) method. These two methods are also designed to remove the manual filter tuning, and they have incorporated the filter learning into the original CSP method. The comparisons are presented in Table 4.

It can be seen that the three methods are quite competitive. Each method has its best performance in certain subjects. Notably, our method does the best in subject ‘av’, outperforming the other two methods by about 10%. As mentioned earlier, the most responsive frequency range of ‘av’ shifts above the normal α and β band (8–26 Hz). This seems to suggest that, for such BCI “abnormal”, the DS features may be a better choice for the classification task.

Table 4: Classification errors (%) of the CSSP, the CSSSP and the DS+S+T method.

|       | CSSP | CSSSP | DS+S+T |
|-------|------|-------|--------|
| aa    | 14.6±6.2 | 11.6±6.3 | 9.5±2.1 |
| al    | 2.3±3.0  | 2.1±2.7  | 2.7±3.1 |
| av    | 32.6±7.6 | 31.8±7.7 | 21.5±7.6 |
| aw    | 3.5±3.3  | 6.5±4.3  | 6.5±4.5 |
| ay    | 6.0±3.9  | 10.5±5.7 | 8.5±5.0 |

5. Conclusion

In this paper, we exploited the collective dynamics of the cortex as features for BCI. We also proposed a framework for learning the optimal spatial and temporal filters during the extraction of these features. For 4 of the 5 subjects tested, our automated approach reduces classification errors to less than 10%. This performance is comparable to that of the CSP features obtained with manually tuned parameters. Further comparisons with other filter learning approaches also show the competitive performance of our method. Our results suggest that the dynamical system features combined with filter learning approach are very promising for BCI. More investigation is needed to fully demonstrate its advantage.

Acknowledgments

We thank Ms. Yingxin Wu and the reviewers for useful comments. National ICT Australia is funded through the Australian Governments Backing Australian Ability initiative, in part through the Australian Research Council.

References

Dornhege, G., Blankertz, B., Curio, G., & Müller, K. (2004). Boosting bit rates in non-invasive EEG single-trial classifications by feature combination and multi-class paradigms. IEEE Trans. Biomed. Eng., 51, 992–1002.

Dornhege, G., Blankertz, B., Krauledat, M., Losch, F., Curio, G., & Müller, K.-R. (2006). Optimizing spatio-temporal filters for improving BCI. In Advances in Neural Info. Proc. Sys. 18.

Gysels, E., & Celka, P. (2004). Phase synchronization for the recognition of mental tasks in a BCI. IEEE Trans. Neural Sys. Rehab. Eng., 12, 406–415.

Lal, T. N., Schröder, M., Hill, N. J., Preissl, H., Hinterberger, T., Melling, J., Bogdan, M., Hofmann, T., Birbaumer, N., & Schölkopf, B. (2005). A brain computer interface with online feedback based on MEG. ICML.

Lemm, S., Blankertz, B., Curio, G., & Müller, K.-R. (2005). Spatio-spectral filters for improving the classification of single trial EEG. IEEE Trans. Biomed. Eng., 52, 1541–1548.

Müller, K.-R., Anderson, C. W., & Birch, G. E. (2003). Linear and non-linear methods for BCIs. IEEE Trans. Neural Sys. Rehab. Eng., 11, 165–169.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. (1989). Spherical splines for scalp potential and current density mapping. EEG and Clin. Neurophys., 72, 184–187.

Pfurtscheller, G., & Lopes da Silva, F. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. Clin. Neurophys., 110, 1842–1857.

Pikovsky, A., Rosenblum, M., Kurths, J., & Chirikov, B. (2003). Synchronization : A universal concept in nonlinear sciences. Cambridge University Press.

Ramoser, H., Müller-Gerking, J., & Pfurtscheller, G. (2000). Optimal spatial filtering of single trial EEG during imagined hand movement. IEEE Trans. Rehab. Eng., 8, 441–446.

Song, L., & Epps, J. (2006). Improving separability of EEG signals during motor imagery with an efficient circular Laplacian. ICASSP.

Song, L., Gysels, E., & Gordon, E. (2006). Phase synchrony rate for the recognition of motor imagery in BCI. In Advances in Neural Info. Proc. Sys. 18.

Srinivasan, R., Nunez, P. L., & Silberstein, R. B. (1998). Spatial filtering and neocortical dynamics: Estimates of EEG coherence. IEEE Trans. on Biomed. Eng., 45, 814–826.

Wang, Y., Gao, X., Zhang, Z., Hong, B., & Gao, S. (2005). BCI competition III—data set IVa: classifying single-trial EEG during motor imagery with a small training set. IEEE Trans. Neural Sys. Rehab. Eng. (submitted).

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain-computer interfaces for communication and control. Clin. Neurophys., 113, 767–791.