Classifying EEG-based motor imagery tasks by means of time–frequency synthesized spatial patterns

Tao Wanga, Jie Denga, Bin Hea,b,*

*a University of Illinois at Chicago, Chicago, IL, USA
b Department of Biomedical Engineering, University of Minnesota, 7-105 BSBE, 312 Church Street, Minneapolis, MN 55455, USA

Accepted 17 June 2004
Available online 21 August 2004

Abstract

Objective: To develop a single trial motor imagery (MI) classification strategy for the brain–computer interface (BCI) applications by using time–frequency synthesis approach to accommodate the individual difference, and using the spatial patterns derived from electroencephalogram (EEG) rhythmic components as the feature description.

Methods: The EEGs are decomposed into a series of frequency bands, and the instantaneous power is represented by the envelop of oscillatory activity, which forms the spatial patterns for a given electrode montage at a time–frequency grid. Time–frequency weights determined by training process are used to synthesize the contributions from the time–frequency domains.

Results: The present method was tested in nine human subjects performing left or right hand movement imagery tasks. The overall classification accuracies for nine human subjects were about 80% in the 10-fold cross-validation, without rejecting any trials from the dataset. The loci of MI activity were shown in the spatial topography of differential-mode patterns over the sensorimotor area.

Conclusions: The present method does not contain a priori subject-dependent parameters, and is computationally efficient. The testing results are promising considering the fact that no trials are excluded due to noise or artifact.

Significance: The present method promises to provide a useful alternative as a general purpose classification procedure for MI classification.

Keywords: Brain–computer interface (BCI); Electroencephalography (EEG); Motor imagery; Event-related desynchronization (ERD); Spatial correlation; Time–frequency weighting

1. Introduction

There is an increasing interest in finding new communication channels between human and environment instead of using nerves and muscles (termed as brain–computer interface, BCI) (Wolpaw et al., 2000; Vaughan et al., 2003). The main applications of BCI technology shall be helping those people suffering severe physical disabilities (locked-in) but cognitively intact, and providing ways of interacting with environment through decoding the mental states measured and discriminated by computer (for review see Wolpaw et al., 2002).

Brain states can be measured by invasive means, by which electrocorticographic signals, local-field potentials or even spikes can be recorded (Kennedy et al., 2000; Levine et al., 2000). However, scalp recorded electroencephalograms (EEGs) have the unique feature of providing a noninvasive convenient way eavesdropping brain signals from cortical areas. By now, various components from EEG have been found useful in this purpose, e.g. the localized changes in spectral power of spontaneous EEG related to sensorimotor processes (Wolpaw et al., 1991; Pfurtscheller et al., 1997); slow cortical potentials (Birbaumer et al., 2000); and various types of event-related potentials (Donchin et al., 2000; Middendorf et al., 2000). BCI operation could be either in synchronous (or cue-driving) mode where mental states alter according to predefined external stimulus or cue, or in asynchronous
(or intermittent) mode where the subjects initiate the intent of control (Mason and Birch, 2000; Millán and Mourino, 2003). For the widely adopted synchronous mode, brain signals within a cue-triggered time windows are processed to extract the patterns sensitive to mental states.

For motor imagery (MI) tasks the subjects are instructed to imagine themselves performing a specific motor action without overt motor output. Positron emission tomography and functional magnetic resonance imaging studies suggest that, cortical sensorimotor systems are activated during MI (Jeannerod, 1994). Other experiments have demonstrated that, the supplementary motor area (SMA), prefrontal area, premotor cortex, cerebellum, and basal ganglia are activated during both movement execution and imagery (Ersland et al., 1996; Decety et al., 1990). Several EEG studies also further confirm the notion that MI can activate primary sensorimotor areas (Beisteiner et al., 1995; Lang et al., 1996). It has been found that planning and execution of movement leads to a short-lasting and circumscribed attenuation in the mu (8–12 Hz) and the central beta (13–28 Hz) rhythm known as event-related (de-)synchronization (ERD/ERS) which has played an important role in BCI study (Pfurtscheller and Neuper, 2001). Such kind of simultaneously attenuated and enhanced EEG rhythms can be used to classify brain states related to the planning or even imagination of different types of limb movements (Pfurtscheller et al., 1997).

In most circumstance, it is easy to record EEGs from multiple electrodes covering a large part of the scalp. Although, much evidence shows the active area of the cortex should be well localized inside the brain, such localized brain activities would be spatial smeared when volume conducted through the skull and the scalp. Müller-Gerking et al. (1999) introduced a method called common spatial pattern (CSP), in an attempt to increase the signal-to-noise ratios by taking advantage of the inherent correlation between neighboring channels. They decomposed the signal into spatial patterns that are extracted from the data of two populations of EEGs in a manner that maximizes their differences. These spatial patterns provide a weighting of the electrodes, which is derived directly from the data. An analysis of multi-channel EEG data has shown that classification accuracy can be increased by training an artificial neural network for each EEG channel and combining all networks to a committee, i.e. a set of channels (Peters et al., 1997). A distinction-sensitive learning vector quantization classification algorithm uses a weighted distance function and adjusts the influence of different input features through a learning algorithm (Pfurtscheller et al., 1997). Furthermore, spatio-temporal principle component analysis has also been used in classify MI from a multi-channel EEGs (Vallabhaneni and He, 2004).

EEG rhythmic components related to MI are time-locked to the preparation and execution events and localized asymmetrically over bilateral hemisphere. The ERD/ERS time course is revealed by computing as the percentage changes of inter-trial variance related to the reference interval with moving average over time to smooth the ripple frequency (Kalcher and Pfurtscheller, 1995). In MI-based BCI system, it is crucial to optimize these sensitive parameters in terms of time, frequency and spatial area to accommodate subject variability. Usually, subject specific parameters are set a priori without objective criterion. In addition to explicitly taking advantage of event-related activities as the feature description, modeling a segment of EEG with autoregressive parameters has been investigated on MI cases, using fewer channels to replace traditional band power approach, which required a reactive frequency bands to be specified a priori (Pfurtscheller et al., 1998). However, directly modeling raw EEG segments is found to be sensitive to the artifact, which is inevitable in such single trial based scenario. The power spectrum of EEG trials (estimated by classical Welch periodogram method) with respect to that of baseline also proved to be practical by combining low-resolution surface Laplacian and linear classifiers (Cincotti et al., 2003).

A mental states classification algorithm universally applying to human subjects should be able to set parameters adaptively through training. Positive contributions could come from a large scope of EEG signals with different extent in time, frequency and spatial domain; however, noise artifacts would also deteriorate the performance if without a proper selecting mechanism. The previous results based on this philosophy have shown the promise only in cases with a small number of channels due to the drawback of the poor generalization capability for involving more channels (Wang and He, 2004). From these considerations, we present a method to recognize mental states of subjects who are performing synchronous MI tasks (imagination of left or right hand movement) by taking advantage of combining rhythmic components in temporal, frequency and spatial domains. The working hypotheses are: (i) envelopes of the oscillatory EEG components can be used as time–frequency features to extract ERD/ERS properties; (ii) spatial patterns derived from the envelopes at time–frequency grids are discriminable between two MI states using EEG; (iii) weighted synthesizing of spatial patterns over time and frequency bins would benefit the performance of classification statistically.

2. Methods

2.1. Data description

The EEG data files consisting of the synchronized imaginary movement experiments were made available by Dr Allen Osman of the University of Pennsylvania (Osman and Robert, 2001) for a data analysis competition during the Neural Information Processing Systems (NIPS 2001)
Brain-Computer Interface Workshop (Whistler, Canada, December 2001) (Sajda et al., 2003).

Scalp EEGs were recorded from 59 channels (placed according to the international 10/20 system) with a 100 Hz sampling rate. Nine subjects were asked to imagine either right or left hand movement indicated by a highly predictable timed cue, which is a six-second-long epoch consisting of several pieces of timing arrangement. The subjects were well trained until their responses were consistently within 100 ms of the synchronization signal. No feedback was furnished to the subjects during the execution of mental tasks. Two important timing cues that should be mentioned here are that at 3.75 s of a trial epoch, a cue (preparation cue, lasting 250 ms) of one letter (‘L’ or ‘R’) appeared on the screen indicating which hand movement should be imagined; and at 5.0 s another cue (execution cue, lasting 50 ms) appeared, indicating that it was time to make the requested response. For each subject, 180 trials were recorded, 90 for imaginary right hand movement and the other 90 for left.

2.2. Surface Laplacian filtering

The raw scalp EEGs are spatially smeared due to the head volume conductor effect. Surface Laplacian filter has been extensively studied to correct the volume conduction effect in the body surface potentials (Hjorth, 1975; Perrin et al., 1987; Le et al., 1992; He and Cohen, 1992; Nunez et al., 1994; Babiloni et al., 1996; He and Wu, 1997; He, 1998, 1999; He and Wu, 1999; He et al., 2001; Zhao and He, 2001). The surface Laplacian can be considered as a spatial high pass filter, which is necessary before further signal processing and feature extraction. McFarland et al. (1997) have compared the EEG classification results using different spatial filters and concluded that the common average and high pass filter, which is necessary before further signal processing and feature extraction. McFarland et al. (1997) have compared the EEG classification results using different spatial filters and concluded that the common average and surface Laplacian yielded good performance.

In the present study, the finite difference implementation of the surface Laplacian (Hjorth, 1975) was used, with the assumption that the distances from the channel of interest to its neighboring channels are approximately equal,

\[ V_{j}^{Lap} = V_j - \frac{1}{4} \sum_{k \in S_j} V_k \]  

where \( V_j \) is the scalp potential EEG of the \( j \)-th channel, and \( S_j \) is a index set of the four neighboring channels.

2.3. Frequency decomposition

Frequency decomposition could be either constant bandwidth or constant-Q (also called proportional-bandwidth) of a data stream. We adopted the latter scheme due to the fact that it could give about a slightly better classification accuracy, meanwhile with less frequency bins. We thus constructed a set of fifth order Butterworth band-pass filters, which span the indicated bands where the ratio (Q) of center frequency over bandwidth is the same for all filters, and adjacent filters overlap according to a criterion: the currently central frequency is the start frequency point of next bins (Fig. 1). If Q is chosen to be large, the bandwidth of the filters is relatively narrow, leading to a large numbers of bins in frequency axis. For small Q, the bandwidth is large with fewer bins, which might result in narrowband signal components contaminated with substantial noise. The overlapping criteria guarantee a proper redundacy of signal in neighboring bins. The total range of frequency bands was chosen from around 6–30 Hz, which covers both mu and beta band.

2.4. Rhythmic component description

We delineated the ERD/ERS properties of the decomposed EEG rhythms of each frequency bin with their envelopes, which can be regarded as a smooth outline of the instantaneous power of the decomposed EEG rhythm for the frequency bin under consideration. The envelop of a narrow band signal can be extracted by using Hilbert transform technique (Papoulis, 1977), and the ERD/ERS phenomenon thus can be evidently observed through the grand average approach over the envelopes. Fig. 2 illustrates the concept of this procedure (see Wang and He, 2004 for details).

Event-related activities are time-locked to the onset of cue, and for MI experiments the ERD may last for around 2 s. Two methods could be used to reduce sampling rate of the envelopes, while retaining the original information. One is to downsample the signal envelopes. Our study showed that the sampling rate could be reduced by a factor of 10 because the frequency band for most envelopes is below 5 Hz. The other one that we finally adopted is to integrate the envelop signal over several consecutive values, that is actually a moving average filter in the time domain that may allow the time smoothing of the envelopes. Specifically, we selected a 2250 ms time range starting from 'preparation cue' to the end of imagination. After moving average filtering, a new time serial data equivalent to 6.7 Hz
2.5. Classification by correlation of spatial patterns

Once the subjects initiated a MI state, the spatial patterns of the piecewise integrated envelopes of the rhythmic component would be a good candidate as a feature vector. In the training stage, at each time–frequency grid point, we obtained a spatial feature pattern denoted by \( p(t,f) \), the correlation coefficients, \( C(t,f)=C(p(t,f), P(t,f)) \), which is defined below,

\[
C = \frac{(p - \bar{p})(P - \bar{P})}{\|P - \bar{P}\| \cdot \|P - \bar{P}\|}
\]

(\( \bar{p} \) and \( \bar{P} \) denote the corresponding mean values of \( p \) and \( P \) between \( p \) and \( P \) was used to measure which kind of mental state the current trial would be more likely to be. We thus define an assignment function of \( p \)

\[
h_{t,f}(p) = \text{sgn}(C(p, P_L) - C(p, P_R))
\]

i.e. \( h_{t,f}(p)=1 \) means that \( p \) is assigned to left hand movement, and for \( h_{t,f}=-1 \) for right hand. Go through all trial in training set in this way, the correct classification accuracy, denoted by \( a_{t,f} \) was obtained.

2.6. Weighted synthesis in time and frequency

We hypothesized that these accuracy values reflect the adaptation or sensitivity of the feature patterns at every frequency bin and time segment. Thus, we defined a normalized time–frequency weight as

\[
w_{t,f} = \begin{cases} 
[(a_{t,f} - E)/(1 - E)]^m, & a_{t,f} > E \\
0, & a_{t,f} \leq E 
\end{cases}
\]

so as to allow \( w_{t,f} \in [0, 1] \), where \( m \) is a control parameter used to further emphasize those bands with larger accuracy values. \( E \) is a threshold, which can be set manually. In the present study \( E=0.6, m=2 \).

According to the time–frequency weights, synthetic decision for the whole time segments over all frequency bins was made by the following synthetic assignment function

\[
g(p) = \text{sgn} \left[ \sum_{i=1}^{T} \sum_{f=1}^{F} w_{i,f} h_{i,f}(p) \right]
\]

where \( T \) and \( F \) denote the sampling numbers in time and frequency domains, respectively. In the testing stage, the classification was accomplished by Eq. (5). The value of \( g(p) \) is either 1 or \( -1 \), indicating left and right, respectively. Since \( w_{t,f} \) functioned as an indicator of the classification capability for a time segment at a frequency bin, the final classification decision was made by synthesizing all frequency bins and time segments according to their contributions as rated by training.

3. Results

A period of 2250 ms starting from the preparation cue was selected (i.e. from 3.75 to 6.0 s of a trial epoch). All times mentioned in the context are relative to the raw
EEG data. In order to compensate for the edge effect of the temporal filtering, a short piece of extra EEG ahead was included in the preprocessing stage. Time coordinate was evenly divided into 14 time intervals with 300 ms involved and 50% overlapping. There are 180 trials available for each subject with equal number of two MI tasks. A usual testing procedure for small dataset is 10-fold cross-validation. We set the number of two kinds of EEG patterns to be even in each fold so as to obtain unbiased training performance.

3.1. Classification accuracy rate

After Laplacian filtering of the EEG signals at 59 electrodes, the 10 electrodes at the periphery were excluded from the present analysis. We selected six montages covering different scope of the area of interests for comparing the performance (see Fig. 3). These electrodes are all selected on the base of the neighboring electrodes of C3 and C4, two most important sites that are very close to the sensorimotor area of the cortex. Table 1 lists the accuracy rates of 9 subjects obtained under six selected montage. The high scores were achieved in montage II, III, IV. Large montage consisting of many electrodes (montage V and VI) might deteriorate the scores (see Table 1).

Fig. 3. Six electrode-montages were selected for comparing classification performance by adopting different electrodes’ spatial patterns. The electrodes inside the different frame (solid, dashed and dot) are selected, and the montages are labeled with Roman number (I–VI). High classification accuracy was achieved in proper localized montage II, III, IV. Large montage consisting of many electrodes (montage V and VI) might deteriorate the scores (see Table 1).

<table>
<thead>
<tr>
<th>Subject</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>IV*</th>
<th>IV#</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>82</td>
<td>81</td>
<td>81</td>
<td>83</td>
<td>76</td>
<td>72</td>
<td>75</td>
<td>91</td>
</tr>
<tr>
<td>S2</td>
<td>92</td>
<td>91</td>
<td>91</td>
<td>88</td>
<td>90</td>
<td>84</td>
<td>95</td>
<td></td>
</tr>
<tr>
<td>S3</td>
<td>78</td>
<td>77</td>
<td>74</td>
<td>75</td>
<td>72</td>
<td>69</td>
<td>65</td>
<td>81</td>
</tr>
<tr>
<td>S4</td>
<td>74</td>
<td>66</td>
<td>71</td>
<td>78</td>
<td>74</td>
<td>65</td>
<td>71</td>
<td>89</td>
</tr>
<tr>
<td>S5</td>
<td>73</td>
<td>77</td>
<td>77</td>
<td>76</td>
<td>74</td>
<td>73</td>
<td>68</td>
<td>83</td>
</tr>
<tr>
<td>S6</td>
<td>77</td>
<td>84</td>
<td>80</td>
<td>77</td>
<td>76</td>
<td>66</td>
<td>71</td>
<td>93</td>
</tr>
<tr>
<td>S7</td>
<td>85</td>
<td>89</td>
<td>88</td>
<td>91</td>
<td>90</td>
<td>89</td>
<td>78</td>
<td>96</td>
</tr>
<tr>
<td>S8</td>
<td>72</td>
<td>76</td>
<td>75</td>
<td>71</td>
<td>55</td>
<td>54</td>
<td>53</td>
<td>90</td>
</tr>
<tr>
<td>S9</td>
<td>66</td>
<td>76</td>
<td>74</td>
<td>74</td>
<td>63</td>
<td>61</td>
<td>61</td>
<td>94</td>
</tr>
<tr>
<td>Mean</td>
<td>78</td>
<td>80</td>
<td>79</td>
<td>80</td>
<td>74</td>
<td>71</td>
<td>70</td>
<td>90</td>
</tr>
</tbody>
</table>

*This column is calculated from the best time segment and frequency bin (see Section 3.1 for details). # This column displays the rate obtained in training set for montage IV.
within a proper scope. Therefore, the small value of $Q$ is preferred for less number of frequency bins.

In the present study, we assume that the subjects, through specific training sessions, would have made required mental states which, however, can not be guaranteed in terms of practical EEG patterns. Artifacts and the possible absence or weakness of subjects intention might occur which makes the perfect classification (100% score) virtually impossible. The strategy dealing with the uncertainty issue might be introducing a third category that contains the rejecting trials with high artifact and ‘doing nothing’ trials, as well as balance mechanism on error rate and communication transfer rate. This could be justified by making a simple calculation on the values inside the square brackets of Eq. 5 whose absolute value may be considered as a likelihood evaluation for the two classless. If we would set those trials with less likelihood (i.e. close to zero) to be the third class as rejected trials in the testing session, the overall scores would be expected to be improved. Computation along this line suggests that if we would set a threshold allowing 10% rejected trials, the mean classification accuracy of 9 subjects would become 82.3% (from 80%) which is close to the ideal improvement at 83.3% provided that the rejected trials are randomly classified in the previous two-classes scheme. For 20% third-class setting, the classification accuracy would become 85% (vs. ideal classification accuracy at 87.5% in this case). In other words, about 40% trials in the third-class were falsely classified previously.

3.2. Time–frequency weight distributions

Time–frequency weight distributions for 9 subjects, as shown in Fig. 4, illustrate the performance contribution of characteristic patterns at time–frequency grids. The figure was obtained under a setting of $Q=4$ on montage IV. Time coordinates were evenly divided into 14 time intervals; zero point indicates the ‘preparation cue’ and 1250 ms point indicates the ‘execution cue’. Frequency coordinates represent 13 center frequencies between 6 and 30 Hz. For most subjects, we could observe that there are peaks located at two ($\mu$ and $\beta$) frequency bands along the time axes. However, the magnitudes vary across subjects. For subjects 2 and 7, for instance, the peaks at beta band appeared quite later in time axis. For subjects 3 and 4, the feature patterns at

![Fig. 4. Time–frequency weight distributions of 9 subjects for the case montage IV and $Q=4$. Time coordinates were evenly divided into 14 time intervals; 0 ms point indicates the ‘preparation cue’ and 1250 ms point indicates the ‘execution cue’. Frequency coordinates represent 13 center frequencies between 6 and 30 Hz (cf. Fig. 1).](image)
beta-rhythm are practically negligible, while for subjects 5 and 8 to the feature patterns at mu-rhythm appeared negligible.

3.3. Topography of synthesized characteristic feature patterns

The weighted sum of characteristic patterns \( P(t,f) \) at time–frequency grids yields a synthesized characteristic pattern (denoted by \( P \)), reflecting the features of the event-related potential in time, frequency and spatial domains. Let \( P_L \) and \( P_R \) denote this pattern from left and right MI-EEG, then we shall be able to derive, \( P_L = P_C + P_D \) and \( P_R = P_C - P_D \), where \( P_C \) is called common-mode pattern (CMP) and \( P_D \) is called differential-mode pattern (DMP). By simple mathematical manipulation we obtain \( P_C = (P_L + P_R)/2 \), and \( P_D = (P_L - P_R)/2 \).

The ERDs associated with movement planning (or preparation) and execution have been demonstrated in Fig. 5. CMPs usually show symmetrical patterns indicating there are common components occurring for both left and right hand MI tasks, which might reflect the bilateral identity for the hand movement events; and many subjects appeared strong activities in the mid-central area as well. DMPs, on the other hand, although relatively weak with regard to CMPs (see Table 2) but explicitly (through the temporal-frequency synthetic manipulation) show contralateral attribution of hand movement. Most of sources concentrate in the vicinity of C3 and C4, while others also appeared on their neighboring areas, supporting the notion that contralateral sensorimotor area is mainly in charge of hand movement operations (Pfurtscheller and Neuper, 2001). However, there is exception in subject 9, which shows ipsilateral patterns to the imagined moving hand.

Fig. 5. Topographical mapping of the CMP (left panels) and DMP (right panels) for the 9 subjects (Montage V was used in order to display clearly the overall spatial patterns of DMP; \( Q = 4 \)). A clear and consistent contralateral patterns can be observed in DMP maps.
A DMP to CMP ratio (DCR) was thus defined to measure proportion of these two components in terms of the ratio of dynamic ranges (i.e. the peak-to-peak value). For all subjects the ratios are around 0.4 (i.e. −8 dB), which indicates that dynamic range of CMP is about 2.5 times larger than that of DMP (Table 2). Suppose the DMP represents the underlying ‘signal’ component which is useful for classification, and given the fact that the single trial ‘signal variance’ is small, the mean ‘noise’ level could thus be estimated by directly averaging the decomposed spatial patterns over all trials and time–frequency grids without weighting procedure. We could thus define an equivalent signal-to-noise ratio for spatial patterns as follows

\[ SNR = 10 \log \left( \frac{1}{N} \sum_{n=1}^{N} \sum_{f=1}^{F} \text{var}(P_{n}(t,f)) \right) \]

where \( p_{n}(t,f) \) denotes the \( n \)-th spatial pattern in a data set containing \( N \) trials, \( T \) is the total number of downsampled time intervals, and \( F \) is the total number of decomposed frequency bins. The results are listed in Table 3. Compare the \( SNR \) of all the 9 subjects with the values of accuracy in Table 1, we can see they are well correlated, i.e. the subject with higher \( SNR \) generally lead to higher classification accuracy.

### 4. Discussion

MI tasks would elicit the non-phase-locked EEG components, and can be displayed with different intensity at various frequency bins by frequency decomposition methods. These frequency bins normally lie in the sensorimotor mu and beta bands, but it is difficult to exactly demarcate these ranges for a given subject. The instantaneous power of the rhythmic components estimated effectively by enveloping is distributed over the scalp with distinguishable and unstationary patterns. The present method attempts to employ useful information in time, frequency and spatial domains by accumulating positive contributions to the classification. The human–machine adaptation was implemented in terms of the time–frequency weights evaluated in the training process. The basic classification component was based on similarity of spatial feature patterns corresponding to each time segment and frequency bin. In addition, the present classification algorithm is computationally efficient in both training and testing stages, since the major computational procedures are linear weighted summations.

As Wolpaw et al. (2002) pointed out, it is important to consider the adaptation processes of BCI by adaptively adjusting parameters of a BCI algorithm through on-line BCI experiments. The availability of such on-line testing would also enable rejecting of bad trials which are judged to be artifact based upon other simultaneous measurements. The present algorithm, however, provides a general purpose classification method to classify MI tasks. Further research can be done to explore on-line adaptive process of further improving the classification accuracy. Given the lower \( SNR \) nature of single trial EEG (Table 3) and on-line processing provision, underlying messages can only be augmented through efforts on every measurable domain. For synchronized MI tasks, techniques of extracting event-related activities of both phase-locked and non-phase-locked components of scalp EEG could serve this purpose.

As shown in Table 1, the mean accuracy scores of all 9 subjects is 80% which is promising considering the small number of trials and the fact that no rejection is considered in the present study (all trials in the dataset were used). Most results reported in BCI literature for classifying overt or imagined movement of limbs (including hands and feet) ranged from less than 65 to more than 95% for individuals (see e.g. Millán et al., 2002; Babiloni et al., 2000, 2001; Peters et al., 2001). Among these methods, CSP has been well studied and found favorable for binary discrimination also using multi-channel setting and could obtain discriminated spatial patterns similar to the Fig. 5 (see e.g. Müller-Gerking et al., 1999; Ramoser et al., 2000). Their results have achieved over 90% accuracy rate for three human subjects with MI experiments, under the condition of artifact free EEG trials.

The main idea of CSP algorithm is to use a linear transform with a projection matrix after decorrelation processing so as to maximize the variance of two-class features in the direction of a few projected vectors. The covariance matrix estimation, which is crucial for CSP is quite sensitive to occasional deviate EEG trials and channels. This might lead to the first-important spatial

<table>
<thead>
<tr>
<th>Subjects</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
</table>

### Table 2

Ratio of dynamic range of DMP over CMP (in dB)

<table>
<thead>
<tr>
<th>Subjects</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCR</td>
<td>−7.39</td>
<td>−8.06</td>
<td>−8.16</td>
<td>−8.35</td>
<td>−8.50</td>
<td>−8.56</td>
<td>−11.30</td>
<td>−11.30</td>
<td>−11.30</td>
</tr>
</tbody>
</table>

### Table 3

Equivalent signal-to-noise ratio (in dB)
patterns (i.e. the first and the last eigenvectors of the inverse projection matrix) that explain the largest variance of one task and the smallest variance of another task fail to provide the best classifying partition in corresponding axis of the transformed space. Therefore, the CSPs derived from the selected vectors of the inverse projection matrix fail to provide the most discriminative source distributions. Fig. 6 shows examples of the spatial patterns for subject 2 obtained from two cases of 20 channels (montage IV) and 28 channels (montage V) with other computational conditions remaining the same. The discriminative spatial patterns over sensorimotor areas are displayed for the first case (Fig. 6a), while for that with 28 channels these patterns apparently stride across the first-important and the second-important spatial patterns (Fig. 6b), specifically the last vector and the second (rather than the first) vector of the inverse projection matrix. The corresponding accuracy rates reduced from 81 to 53% if only the first-important eigenvectors are taken into account. Another drawback of CSP methods is that since there is no substantial difference for the two eigenvectors with opposite direction during eigen-decomposition, the polarity of the spatial pattern can be opposite which could make the inconsistent polarity for the same type of task. So far the overall accuracy rate for the 9 subjects we currently achieved applying the CSP method is 76% if appropriate time–frequency and spatial parameters can be set.

As to the feature characterization, autoregressive model as a well-established methods has been found promising as applied to classify the binary MI tasks (e.g. Pfurtscheller et al., 1998), in which a comparison of adaptive autoregressive parameter with their DSLVQ method was investigated. The best results for the 4 subjects and artifact free trials were reported as between 5.8 and 32.8% in terms of error rate. For the results applied to the same dataset as ours during the competition session (Sajda et al., 2003), method based on three combination features known from BCI literature, i.e. ERP, AAR model and CSP, has achieved the best of 76% in average. It can be seen, however, the scores of individual subjects are quite different with the presented one which might indicates that individuals probably have their own sensitive feature description. It has been acknowledged that the score strongly depends on experimental settings, training of subjects, overt or imagined movements, and the number of MI tasks, which make the comparison of the capacities of different methods less convincible. However, these methods contribute to this important issue with various perspectives, which would definitely benefit the practical solution in the future.

In summary, we have developed a spatio-time–frequency approach for mental state classification based on the oscillatory event-related activity during MI tasks from multi-channel scalp EEG recordings. The classification principle is based on the spatial correlation between feature patterns as obtained from multi-channel recordings and time–frequency decomposition. The time–frequency weighting determined through training process has been shown to provide design of subject-tailored classifiers. The present method does not contain subject-dependent parameters, and is computationally efficient. The testing results in nine human subjects are promising despite the nature of off-line testing and the fact that no trials are excluded due to noise or artifact. The present method promises to provide a useful alternative as a general purpose classification procedure for MI classification.

Acknowledgements

The authors are grateful to Dr Allen Osman of University of Pennsylvania for making his data available, and Lei Ding for useful discussions in the data analysis. This work was supported in part by NSF BES-0218736, NSF CAREER Award BES-9875344, and NIH R01EB00178.

References


