HUMAN EXPERT SUPERVISED SELECTION OF TIME-FREQUENCY INTERVALS IN EEG SIGNALS FOR BRAIN–COMPUTER INTERFACING

Alban Duprès, François Cabestaing, José Rouillard
Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISTAL
Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France
{firstname.lastname@univ-lille1.fr}

Abstract—In the context of brain–computer interfacing based on motor imagery, we propose a method allowing a human expert to supervise the selection of user-specific time-frequency features computed from EEG signals. Indeed, in the current state of BCI research, there is always at least one expert involved in the first stages of any experimentation. On one hand, such experts really appreciate keeping a certain level of control on the tuning of user-specific parameters. On the other hand, we will show that their knowledge is extremely valuable for selecting a sparse set of significant time-frequency features. The expert selects these features through a visual analysis of curves highlighting differences between electroencephalographic activities recorded during the execution of various motor imagery tasks. We compare our method to the basic common spatial patterns approach and to two fully-automatic feature extraction methods, using dataset 2A of BCI competition IV. Our method (mean accuracy $m = 83.71 \pm 14.6$ std) outperforms the best competing method ($m = 79.48 \pm 12.41$ std) for 6 of the 9 subjects.

Index Terms—brain–computer interface, EEG signal processing, sparse feature set, feature selection, human expertise.

I. INTRODUCTION

Brain-computer interfaces (BCI) are devices that enable users to control effectors using only their cerebral activity. For now, non-invasive BCI have mainly been used to restore a communication channel between a severely disabled user and a computer, letting him/her recover a partial autonomy. In the so-called active BCI paradigms, users consciously control their mental activity at their own pace and independently from external events [1]. Motor imagery (MI), i.e. imagination of a specific motor action, is the most commonly used mental task in active BCIs. In this context, it is crucial to define appropriate features, computed from EEG signals, allowing the BCI to distinguish between different MI tasks performed by the user.

During MI, very specific neurophysiological patterns are elicited in electroencephalographic (EEG) signals, such as event-related desynchronisations (ERD) and event-related synchronisations (ERS). ERD and ERS are characterized respectively by a decrease of EEG power during MI and an increase of EEG power at the end of MI [2]. ERD and ERS are known to appear mainly in EEG signals recorded over the motor cortex with a spatial distribution that depends on the MI task, following the cortical motor homunculus [3]. EEG frequency bands in which ERD and ERS can be detected during MI correspond to $\mu$ (7–13 Hz) and $\beta$ (13–25 Hz) rhythms. Nevertheless, it is well known that the most relevant frequency bands and spatial locations are variable over subjects and MI tasks [4].

In order to facilitate the discrimination of MI tasks, the EEG signal processing pipeline typically includes a spatial filtering stage. CSP (Common Spatial Patterns) is the most widely used linear spatial filtering approach involving user-specific parameters [5]. A user-tuned CSP filter increases the variance of filtered EEG signals for one specific MI task while minimising their variance for other MI tasks or for non MI-related mental states. The actual discriminative performance of CSP filters depends on the frequency bands in which the signals are processed and on the time interval during which signal power is determined. For instance, CSP spatial filters computed on raw EEG signals or on EEG signals filtered in inappropriate frequency bands yield poor classification performance.

To solve this problem, several approaches have been described in the literature. Some researchers have proposed to keep a wide frequency range, i.e. encompassing $\mu$ and $\beta$ rhythms and therefore valid for any user, and to improve the spatial filtering stage. For example [6] has compared several techniques for determining regularized versions of CSP, showing that spatial filtering can be significantly improved compared to basic CSP. Other researchers have proposed to select user-specific frequency bands in which ERD/ERS detection is more effective. For example, [4] implements multiple band-pass frequency filters and computes a specific CSP for each frequency band. Then, a feature selection algorithm keeps the most relevant frequency/CSP features for a given user.

However, all these techniques include a fully automatic feature selection stage, which implies defining empirically several meta-parameters, such as the number of features. But they do not specify the sparsity of the feature space by taking into account some a priori neurophysiological knowledge during this feature selection stage. Studies have shown that sparsity of the feature space allows for good classification performance because the BCI is less sensi-
tive to covariate shifts in EEG signals. For instance, Raza et al. have proposed to increase the sparsity using either forward-addition or backward-elimination of features in the space [7]. However, no detailed neurophysiological-based analysis of the ERD/ERS patterns is performed in order to select the most appropriate frequency intervals, spatial locations and time intervals in which the EEG signals should be analysed.

In this paper, we describe a technique in which a human expert selects a small number of time-frequency features during a first stage. Obviously, the expert does not examine the raw EEG signals to make his/her decision, which would require a very high level of expertise, but a set of curves plotted using samples of time-frequency processed EEG signals. Then, in a second stage, a CSP filter is determined for each retained time-frequency feature. This approach yields a sparse feature space, whose sparsity is controlled by a human expert and not by blindly setting the values of a few meta-parameters. We will also see that this approach enables us to gather detailed informations about the neurophysiological patterns of a given user during MI, and therefore better understand his/her performance.

II. Method

Since our method is supervised, the signal processing pipelines are slightly different for the training and the online processing modes. In the training mode, which aims at selecting features and adapting parameters, a set of labelled EEG trials is analysed. The BCI paradigm used for building this training set must be cue-based in the sense that the user is told when to start and when to stop imagining two or more different motions. In the online processing mode, the paradigm is not necessarily cue-based and the user can freely perform any of these MI tasks when he/she wants to.

Figure 1 illustrates the processing pipeline for the training mode. It is composed of four successive stages: spatial filtering, band-pass frequency filtering, power estimation, and aggregation over trials. In the training mode, we wanted to use the same processing techniques and parameters for all the users. Thus, the surface Laplacian was selected for spatially filtering EEG signals since it allows spatial noise removal and source identification without requiring user-tuning [8]. Then, for highlighting user specificities in the frequency domain, like other authors we use a bank of band-pass filters encompassing the frequency bands of $\mu$ and $\beta$ rhythms [7], [4]. For highlighting ERD/ERS in the signals, we compute their log-variance — equivalent to their power — in a sliding window of fixed duration. Finally, for getting a data representation easily understandable by the human expert, we compute the average and standard deviation, over all trials of each MI task, of signal power at every instant.

Indeed, the aim of the proposed method is to help the human expert to analyse neurophysiological time-frequency patterns related to MI tasks. For each MI task, he/she is asked to review a set of curves, one for each frequency band, showing the time-course of a specific signal known to be correlated to this MI. In order to clearly exhibit differences between the analysed MI task and other MI tasks, an additional baseline curve is shown to the expert for each frequency band. This baseline curve is computed

Figure 1: Processing pipeline for the training mode.

Figure 2: Set of curves for expert review of left hand motor imagery, electrode C4.
by averaging values of the signal under consideration for all
the trials corresponding to other MI tasks, i.e. following a
one-versus-rest strategy.

For example, figure 2 shows the first set of six curves,
one for each frequency band, displayed for review by the
expert in order to tune a BCI based on left hand vs. right
hand MI. The EEG signal recorded at location C4 over the
right sensorimotor cortex is known to exhibit clear ERD
when the user performs MI of the left hand. For this MI
task, the curve of interest in each frequency band (solid red
lines in figure 2) is the time-course of average power of
this signal over all corresponding trials, after spatial noise
removal by a Laplacian filter, in a time window starting one
second before MI onset and ending one second and half
after MI offset. Baseline curve in each frequency band (blue
dotted lines in figure 2) is the average power of the same
signal computed over all trials that do not correspond to the
MI task under review. To display the statistical significance
of signal variations, two additional curves are plotted around
each curve at plus and minus half standard deviation.

A visual analysis of these curves enables the expert to
select several time-frequency intervals that best discriminate
this particular MI task from others. The expert pays atten-
tion to band-passed signals that highlight neurophysiological
patterns related to a MI task, such as ERD or ERS patterns.
Neurophysiological knowledge of a human expert is useful
to visually identify such patterns, which have different fre-
quency and temporal distribution over subjects. For example,
two time-frequency intervals [0 – 2.5 s, 8 – 12 Hz] and
[0 – 2.5 s, 20 – 24 Hz] can be selected as the best candidates
to discriminate the two MI. After this, CSP filters are
computed using epochs of all EEG signals in each of these
time-frequency intervals, in order to optimise spatial filtering
for this user compared to the non-adaptive Laplacian. Three
pairs of CSP filters are kept for each time-frequency interval.

For the online mode, the processing pipeline includes
four successive stages, as illustrated in figure 3. EEG sig-
als are filtered in the spatial and frequency domains by
CSP/band-pass pairs determined during the training mode.
Then, the log-variances of filtered signals are computed over
time intervals that were considered as most discriminant for
each frequency by the expert, yielding a small number of
features. Finally, the signal epoch is analysed by a LDA
(Linear Discriminant Analysis) classifier. This processing
pipeline can be performed over sliding overlapped time
windows when the paradigm is not cue-based.

Since the features are determined considering a one-versus-
rest strategy, we follow the same approach for classifying
signal epochs. Each LDA classifier outputs both a class and
a confidence score, for instance the distance between the
feature vector and the separating hyperplane. A standard
voting procedure is used to determine the most appropriate
class according to all outputs.

We also allow the expert to compare different sets of
features, by excluding time-frequency intervals that were
initially selected or including other intervals. To compare
the efficiency of these various sets, the system evaluates
the correct classification rate that each of them yields when
applied on the training set of EEG signals. However, the
expert can decide to keep a smaller set of features even
with a slightly lower classification rate, if he/she considers
that it is probably more robust to deal with covariate shifts
in signals.

III. RESULTS AND COMPARISON

In this paper, our method is evaluated on EEG signals
freely available in data set 2A of BCI competition IV, which
has been widely used for comparison purposes [9]. This data
set comprises raw EEG data recorded by 22 electrodes from
9 subjects. Subjects were asked to perform left hand, right
hand, feet, and tongue MI. All MI tasks were performed
during four seconds just after presentation of a cue. Each
user performed two sessions on different days in order to
obtain a training data set and an evaluation data set.

Only data recorded during left hand and right hand MI
were used in the evaluation, in order to compare our method
with those described in [6] and [7]. EEG signals recorded at
locations C3 and C4 were spatially filtered by a Laplacian,
yielding two signals of interest for further processing, one
for each MI: the signal derived from spatial filtering of C3
(resp. C4) is known to exhibit ERD when the user performs
right (resp. left) hand motor imagery. These signals were
filtered by two banks of six band-pass frequency filters (5th
order Butterworth), yielding twelve signals of interest. Fi-
nally, their log-variance was computed, time-averaged over
a sliding window, and aggregated to determine time-courses
of averages and standard deviations over trials. A sliding
window of one second length allows to keep a good temporal
resolution and to highlight ERD/ERS patterns by smoothing
the signal power.

Then, the expert reviewed the curves plotted for each MI
and each frequency band in order to retain time-frequency
intervals that he considered as the most discriminant be-
tween left and right hand MI. For instance, curves displayed
in figure 2 correspond to the signals of the training set
for subject 9, that were reviewed by the expert. CSP fil-
ters were then determined for each time-frequency interval,
LDA classifiers trained, and correct classifications scores
computed for data of the training set. These scores, although
obtained on the training data, could be used by the expert to
add/remove time-frequency intervals by considering a trade-
off between performance and sparsity.
For each subject, figure 4 shows the first time-frequency interval that our expert considered as the most discriminant hand MI. The vertical black dashed lines indicate boundaries of the selected time interval in each frequency band. One can observe that differences between the two MI tasks are more visible in the processed EEG for subjects 1, 3, 7, 8, and 9 than for other subjects. Table 1 indicates the sets of time-frequency intervals that were finally retained by our expert.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Time-frequency intervals $[t_1 - t_2 \text{ (s)}; f_1 - f_2 \text{ (Hz)}]$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[0.5 – 3.3; 8 – 12], [0.5 – 4; 12 – 16], [0.5 – 1.5; 20 – 24], [0.5 – 2.5; 24 – 28]</td>
</tr>
<tr>
<td>2</td>
<td>[0.5 – 1.5; 12 – 16], [0.5 – 1.5; 16 – 20]</td>
</tr>
<tr>
<td>3</td>
<td>[0.5 – 4; 8 – 12]</td>
</tr>
<tr>
<td>4</td>
<td>[0 – 4; 8 – 12], [1 – 2.5; 16 – 20], [0.5 – 2; 24 – 28]</td>
</tr>
<tr>
<td>5</td>
<td>[0 – 3; 4 – 8], [0 – 1.7; 20 – 24], [0 – 3; 24 – 28]</td>
</tr>
<tr>
<td>6</td>
<td>[0 – 1.5; 8 – 12], [0 – 2.5; 12 – 16], [0.5 – 2; 24 – 28]</td>
</tr>
<tr>
<td>7</td>
<td>[0 – 4; 3 – 8], [0.5 – 4; 8 – 12], [0.5 – 3.5; 16 – 20]</td>
</tr>
<tr>
<td>8</td>
<td>[0 – 1.5; 8 – 12], [1 – 3; 12 – 16], [0.5 – 3; 16 – 20], [0.5 – 2.5; 20 – 24]</td>
</tr>
<tr>
<td>9</td>
<td>[0.5 – 2.5; 8 – 12]</td>
</tr>
</tbody>
</table>

TABLE 1: Time-frequency features for each subject

Using this set of time-frequency features, and the online processing pipeline of figure 3, our method was compared to: 1) a standard CSP; 2) the weighted Tikhonov regularized CSP (WTRCSP) presented in [6]; 3) the backward-elimination (BE) method described in [7]. The basic CSP and the WTRCSP, which is the best regularized CSP algorithm according to [6], are computed on a wide frequency range from 8 to 30 Hz and a time interval between 0.5 and 2.5 s after the cue. However the BE method computes CSP filters on user-specific frequency bands and a time interval between 0 and 3 s after the cue.

Table 2 shows the classification accuracies (in %) for each subject and each processing method. Methods are assessed on the evaluation data set, only mean and standard deviation are used for quantified and detailed analysis of results. The highest classification score for each subject is indicated in bold font. Our method (mean accuracy $m = 83.71 \pm 14.6 \text{ std}$) outperforms the other methods for 5 of the 9 subjects. The performance of our method is worse for subjects 2, 5, and 9 but equal for subject 3. Our method is mainly profitable for subjects 4 and 7 for whom the accuracy is increased respectively by 9 and 15 percentage points compared to the best competing method. One can verify that methods which select user-specific frequency intervals, such as our method (mean $m = 83.71$) and the BE ($m = 79.43$) method, obtained better results than a basic CSP ($m = 78.01$) and a WTRCSP ($m = 78.47$) both computed on a wide frequency range.

IV. DISCUSSION

In Figure 4, according to the spatial location and the MI under review, we can observe for all subjects a decrease, even weak, of the EEG power recorded over the contralateral motor cortex. ERD patterns are present at location C4...
Table 2: Classification scores for each subject.

<table>
<thead>
<tr>
<th>Subject</th>
<th>basic CSP</th>
<th>WTRCSP</th>
<th>BE</th>
<th>Our method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>88.89</td>
<td>88.89</td>
<td>90.28</td>
<td>92.36</td>
</tr>
<tr>
<td>2</td>
<td>51.39</td>
<td>54.86</td>
<td>63.19</td>
<td>61.11</td>
</tr>
<tr>
<td>3</td>
<td>96.53</td>
<td>96.53</td>
<td>93.75</td>
<td>96.53</td>
</tr>
<tr>
<td>4</td>
<td>70.14</td>
<td>70.14</td>
<td>70.14</td>
<td>79.17</td>
</tr>
<tr>
<td>5</td>
<td>54.86</td>
<td>65.97</td>
<td>72.92</td>
<td>62.5</td>
</tr>
<tr>
<td>6</td>
<td>73.53</td>
<td>61.81</td>
<td>65.97</td>
<td>75.69</td>
</tr>
<tr>
<td>7</td>
<td>81.25</td>
<td>81.25</td>
<td>75</td>
<td>96.53</td>
</tr>
<tr>
<td>8</td>
<td>93.75</td>
<td>93.75</td>
<td>91.67</td>
<td>97.22</td>
</tr>
<tr>
<td>9</td>
<td>93.75</td>
<td>90.97</td>
<td>92.36</td>
<td>92.36</td>
</tr>
<tr>
<td>Mean</td>
<td>78.01</td>
<td>78.47</td>
<td>79.48</td>
<td>83.71</td>
</tr>
<tr>
<td>Std</td>
<td>14.6</td>
<td>15.65</td>
<td>12.41</td>
<td>17.01</td>
</tr>
</tbody>
</table>

We have proposed an easy to implement method in order to select the time-frequency intervals that best discriminate different classes in the context of MI-based BCI. The selection of time-frequency intervals is specific to each subject and is performed offline, using a set of pre-recorded signals. It is supervised by a human expert who reviews a set of curves determined for each MI task through a time-frequency analysis of the recorded EEG signals. The results of our study confirm the fact that a correct selection of time-frequency intervals impacts the performance of CSP, as mentioned in the literature. Moreover our method gathers detailed informations about the specific neurophysiological patterns appearing in EEG signals when the users performs MI. It enables a better understanding of the difference in classification accuracy between the subjects. We are currently developing a user-friendly software interface that will allow an easier selection of time-frequency intervals by the expert, as well as a visual validation of spatial patterns determined using the training set.

References