MSNE Python Workshop

Offline Analysis of the BCICIV Motor Imagery Data Set

Course Instructor: Karahan Yilmazer `karahan.yilmazer@tum.de`

Winter Semester 23/24

Introduction

Now we can use a publicly available motor imagery data set that is assumed to be of high quality to test different preprocessing and analysis methods. Our goal in this tutorial will be to classify between different classes of motor imagery.

Info: These tutorials are designed keeping in mind that the class comprises of people from different backgrounds. Do not be overwhelmed, the topics will be covered in as much detail as required to enable you to start using the technologies presented in the tutorials.

As mentioned, we will be analyzing a motor imagery (MI) data set from the BCI Competition IV (BCICIV). It contains EEG data from 7 participants performing different motor imagery tasks (two out of the three classes: left hand, right hand, and both feet). The data was recorded from 59 channels with a sampling frequency of 1000 Hz. It was then band-pass filtered between 0.05 Hz and 200 Hz. We will be using the downsampled version of this data set at 100 Hz, but you can install and use the original 1000 Hz version of the data set if you want to. Further information on the data set can be found here.

You will find the data set in the folder **`T4-motor-imagery-analysis\data`**. Each participant has their own MATLAB construct which we can read in Python using `scipy.io.loadmat()`.

The Info object contains the necessary information for the recording and is the base Setting Up the Software As you can see, since the learning outcome is low we already read in the contents of a file that contains EEG data. The data is saved in .mat format that can be easily read by MATLAB but in Python, you have to deal with nested arrays to access the content.

1 MNE

For the analysis of EEG data, we will be using the MNE library. Keep in mind that you can implement every MNE functionality we will be using for this course yourself. But for now it is more straightforward to use MNE for most of our analysis. We will be touching on the basics of using the library. **However, the MNE website offers an amazing tutorial section for nearly all the use cases. Do check it out if you are interested.**

The two base classes we will be using are **`Raw`** and **`Epochs`**. Raw is the object that represents the continuous EEG recording. Epochs is the epoched version of this data. This means that we cut out the so-called *trials* or *epochs* from the continuous EEG data. For that, we use the time markers recorded simultaneously with the EEG.

Time markers are sent every time an important event occurs during the experiment. They signal the important time period to focus on during the data analysis part. The most important markers for us are the cue onsets where the imagery period starts. We call the time period before the cue was shown the *pre-stimulus period* and the one after the cue onset *post-stimulus period*. The post-stimulus period is where participant is imagining moving their limb and the pre-stimulus period will be our baseline.

By epoching the continuous data, we get samples of how EEG data looks like for a given event, like left hand imagery. By analyzing these segments of data, we can train classifiers to separate different types of motor imagery.

Most of the explanation for the offline analysis will be done in the tutorial session. You will also see that the analysis script is full of comments and small explanations for each processing step. However, in the following you will find a section about the common spatial pattern (CSP). It is the method that we will be using to increase the separability of different classes and the explanation is intended for those students who are interested in the mathematics behind it.

Later, you will find two sections about motor imagery and ERD/ERS curves. Especially the ERD/ERS curves section is meant for the ones who are seeking a nice challenge. In the analysis script `berlin_mi_analysis.py` you can find a function called **`calc_erds()`**. This function calculates and plots the ERD/ERS curves. However, in the script called **`exercise_berlin_mi_analysis.py`** it is left empty for you to fill. Arguably, the most fun way to fill it is to go through this paper and try to implement it yourself by minimizing the help you receive from this tutorial. But of course, you can always refer to the explanations below.

2 Common Spatial Pattern (CSP)

The EEG frequency band powers play a very important role when it comes to designing BCIs. Bearing in mind that a convenient way to get an estimate of band powers of filtered EEG signals is to compute their variance [1], it becomes clear that an efficient algorithm that can discriminate different types of brain activity shall rely on the variances of filtered data. Given that it is known which class a data segment belongs, CSP can be used for this purpose.

CSP is a supervised spatial filtering method that transforms EEG data into a new space where the variance of one class is maximized whereas the variance of the other class is minimized [2]. Consequently two classes can be discriminated more easily. However, it was shown that it can also be extended to be used in a multi-class setting [3].

CSP was introduced in [4]. It was first used in EEG analysis to extract abnormal components from the clinical EEG [5]. CSP was later adapted to EEG data classification of movement-related patterns [6]. Now, it is a widely used method to classify EEG data, especially motor imagery.

In most studies a wide filter band is used. This band can vary from 8-30 Hz [1] to 4-40 Hz [7] depending on the application. The argumentation behind this choice is that this broad frequency band contains both the mu and beta frequency bands which were shown to be important for movement classification [8]. In one study it was shown that the 8-30 Hz broad band outperformed narrow bands in classification [6].

Although the implementation can vary slightly based on the study, the general outline will be explained below.

The first step is to band-pass filter the data in the range of interest, in accordance with the previous discussion about the frequency band. The filtered continuous data is then epoched. This means that it is cut into individual trials or epochs. Each epoch is usually time-locked to a cue, for example the start of a motor imagery period during the recording.

Let it be assumed that raw EEG data for a single epoch is represented as an (N x T) matrix X, where N is the number of channels and T is the number of samples per channel in each epoch. Trivially, T is dependent on the sampling rate of the EEG system and the epoching window.

Due to the band-pass filtering, the constant part of the EEG data has been removed and the mean of the distribution is thus zero. That is why the first place to look for characteristic information is in its second moments, or the covariance matrix [6]. The normalized spatial covariance matrix of the EEG can then be estimated as:

$$R = \frac{XX^T}{trace(XX^T)} \tag{1}$$

The matrix products correspond to averaging over time [6] and the trace normalization is done to eliminate magnitude variations in the EEG among different individuals [5]. It has to be noted that this normalization can be omitted. In this case the covariance matrix is estimated as XX^T.

The covariance matrix is then computed for both classes separately:

$$R_1 = \frac{X_1 X_1^T}{trace(X_1 X_1^T)}$$

$$R_2 = \frac{X_2 X_2^T}{trace(X_2 X_2^T)}$$

where the subscripts denote different classes (e.g. left and right hand imagery).

The next step is the eigenvalue decomposition of the composite covariance matrix which is the sum of the covariance matrices computed earlier:

$$R_c = R_1 + R_2$$

$$R_c = U \lambda U^T \tag{2}$$

Since R_c is a real symmetric matrix, U is an NxN orthogonal matrix whose columns are orthonormal eigenvectors of R_c and λ is a diagonal matrix whose entries are the corresponding eigenvalues.

The whitening transformation is then defined as:

$$P = \lambda^{-1/2} U^T \tag{3}$$

This transformation equalizes the variances in the space spanned by the eigenvectors in U [6]. If the class-specific covariance matrices are transformed using this whitening matrix by

$$S_1 = P R_1 P^T \tag{4}$$

$$S_2 = P R_2 P^T \tag{5}$$

then S_1 and S_2 share common eigenvectors (basic spatial patterns), since

$$S_1 + S_2 = P R_c P^T = I$$

If the eigenvalue decomposition of S_1 is given as

$$S_1 = B \psi_1 B^T \tag{6}$$

then the S_2 can be factored as:

$$S_2 = B \psi_2 B^T \tag{7}$$

where B is orthonormal.

In this case the diagonal eigenvalue matrices ψ_1 and ψ_2 always add up to the identity matrix:

$$\psi_1 + \psi_2 = I \tag{8}$$

Thus the variance accounted for by the first m eigenvectors which correspond to the m largest eigenvalues in ψ_1 will be maximal for class 1. Due to the constraint on ψ_2 in equation 8, the variances accounted for by these eigenvectors must be minimal for class 2. The reverse will hold true for the last m eigenvectors [5].

The CSP projection matrix is calculated as:

$$W_{CSP} = B^T P \tag{9}$$

Finally, the CSP projection matrix can be computed as:

$$Z = W_{CSP}^T X \tag{10}$$

The columns of of W_{CSP}^{-1} are the so called common spatial patterns and can be seen as time-invariant EEG source distribution vectors.

The feature vector can be calculated as:

$$g_p = log(\frac{var(Z_p)}{\sum_{p=1}^{2m} var(Z_p)}) \tag{11}$$

To calculate the feature vector, the first and last m rows of Z are used. a common value for m is 2 [6]. Conversely g_p has the length $2m$. This vector can be computed for each trial and then concatenated into a feature matrix F:

$$G = [g_1, g_2, ..., g_{2m}]^T \tag{12}$$

3 Motor Imagery

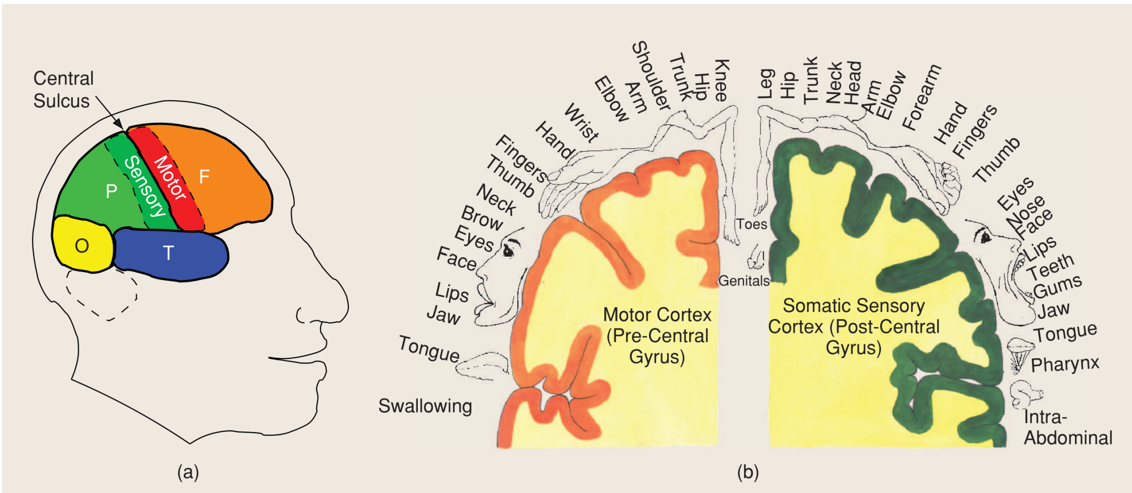
EEG recordings are well-suited to capturing oscillatory brain activity or "brain waves" [9]. These brain waves are generated by the synchronous activity of large neuron populations and have characteristic frequency ranges and spatial distributions. One of such brain waves is called "alpha waves" or "alpha rhythm" and falls in the range of 8-13 Hz. It is a very dominant brain wave that can even be seen in EEG recordings without the need for extensive filtering.

A brain wave of interest for this tutorial is the mu rhythm that oscillates with a frequency of 8-12 Hz. As it can be seen, the alpha and mu rhythm frequency bands overlap. The difference is, however, their spatial distributions. While the alpha rhythm is mostly observed in the occipital region, the mu rhythm is seen over the sensorimotor region (see Figure 1). Thus it is also referred to as the sensorimotor rhythm. Actually sensorimotor rhythms fall into two more major frequency bands: beta (18-30) and gamma (30-200+ Hz) [10]. But due to the Nyquist Theorem, to be able to reliably capture the whole Gamma rhythm, sampling rates over 500 have to be used. So, most researchers focus on the mu and the beta bands. For the purpose of this tutorial, the main focus will be on the mu band.

It was repeatedly shown that there is a decrease in the mu rhythm during limb movement [11] [12] [13]. This decrease is called the event-related desynchronization (ERD) [14]. If however the SMRs are increased in association with sensorimotor rhythms, it is called a event-related synchronization (ERS) [15]. Interestingly the ERDs are observed also when a movement is imagined. This is called a motor imagery (MI). It is the act of imagining a motion but not executing it.

It was shown that MI produced neural activity that is spatiotemporally similar to the motor execution case but smaller in magnitude [16]. The main focus of this tutorial is the MI of the hands. The MI imagery of the hands cause an ERD for the contralateral and an ERS for the ipsilateral hemisphere [9]. In simpler terms, this means that right hand MI causes an ERD on the left hemisphere while a left hand MI causes an ERD of the right hemisphere. Looking at the electrode positions on a typical EEG setup, it can be concluded that during a left hand MI an ERD has to be seen at the C4 electrode, whereas a right hand MI should cause an ERD at the C3 electrode. The corresponding plots can be seen in 3a.

This phenomenon is also captured by the so called motor homunculus seen in Figure 1 (b). The homunculus is a geometric mapping between body parts and motor/somatosensory cortex and it again shows that the hands are contralaterally arranged.



Blankertz - Optimizing Spatial filters for Robust EEG Single-Trial Analysis [17]

Figure 1: (a) Lobes of the brain: frontal, parietal, occipital, and temporal (named after the bones of the skull beneath which they are located). The central sulcus separates the frontal and parietal lobe. (b) Geometric mapping between body parts and motor/somatosensory cortex. The motor cortex and the somatosensory cortex are shown at the left and right part of the figure, respectively. Note that in each hemisphere there is one motor area (frontal to the central sulcus) and one sensori area (posterior to the central sulcus). The part which is not shown can be obtained by mirroring the figure folded at the center. [17]

4 ERD/ERS Curves

The classical ERD/ERS method calculates the instantaneous power as:

$$\overline{P}_j = \frac{1}{N}\sum_{i=1}^{N} x_{f_{i,j}}^2 \tag{13}$$

where $\overline{P}_j$ is the averaged power estimation of band-pass filtered data (averaged over all trials) and $x_{f_{i,j}}$ is the j-th sample of the i-th trial of the band-pass filtered data [18].

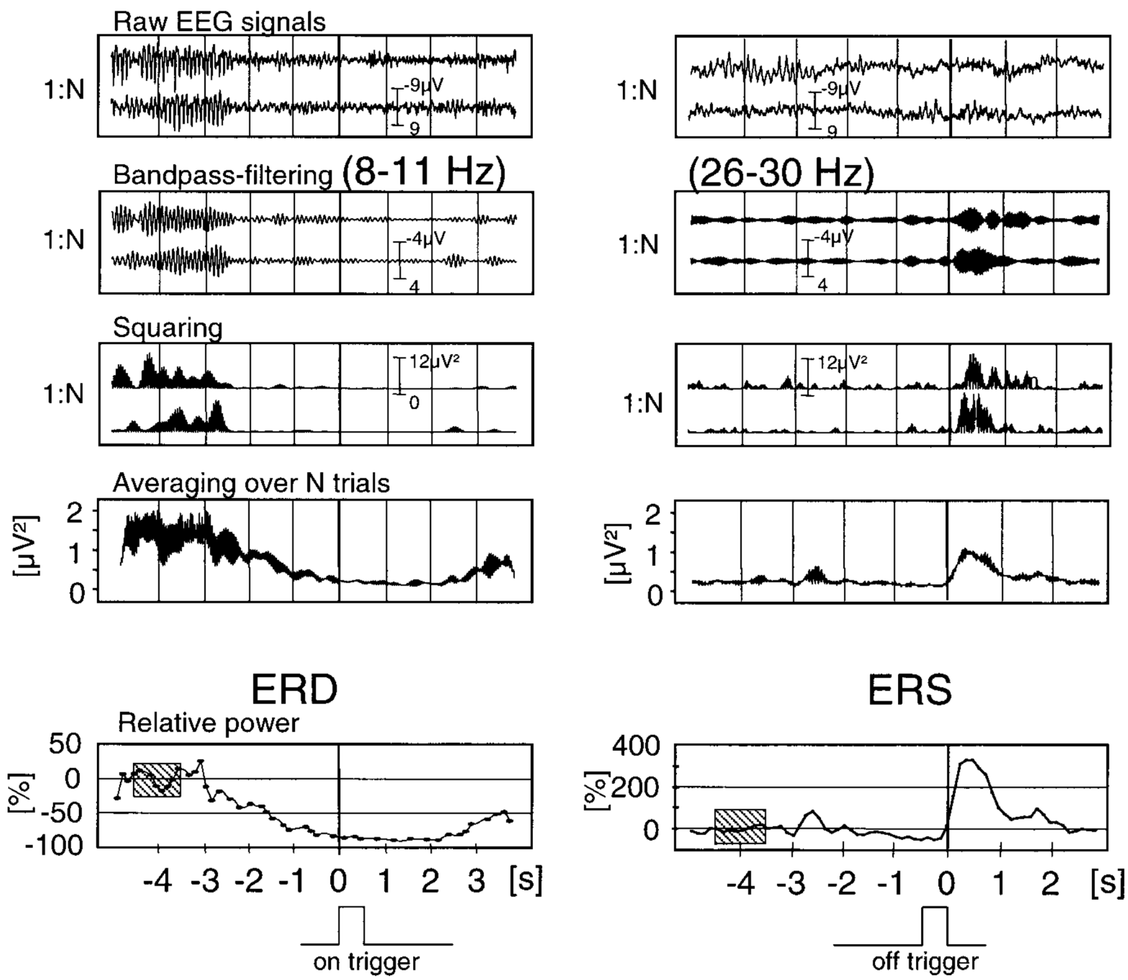
This estimation of the band power is then divided into the period of interest A and the baseline or reference period R. Consecutively, the ERD or ERS is defined as [19]:

$$ERD_j = \frac{A_j - R}{R} \cdot 100\% \tag{14}$$

with

$$R = \frac{1}{k} \sum_{j=n_0}^{n_0+k} A_j \tag{15}$$

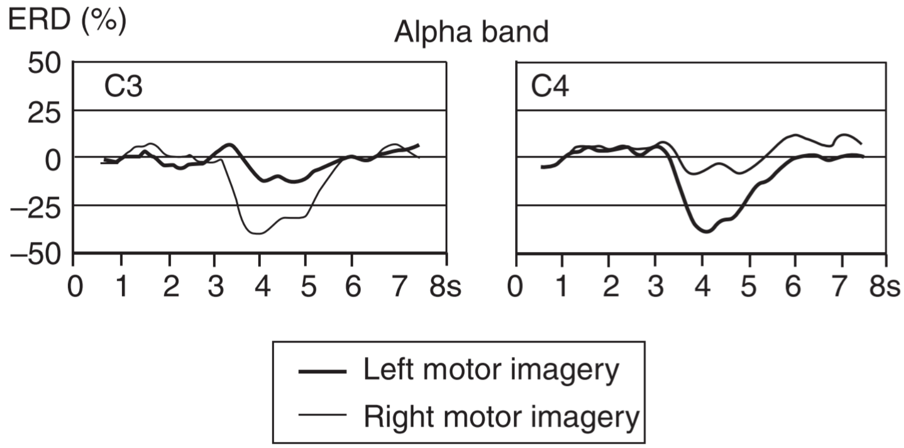
where A_j is the power at the j-th sample.
A schematic showing the steps of this method can be found in Figure 2.



Pfurtscheller and Lopes da Silva - Event-related EEG/MEG synchronization and desynchronization: Basic principles [19]

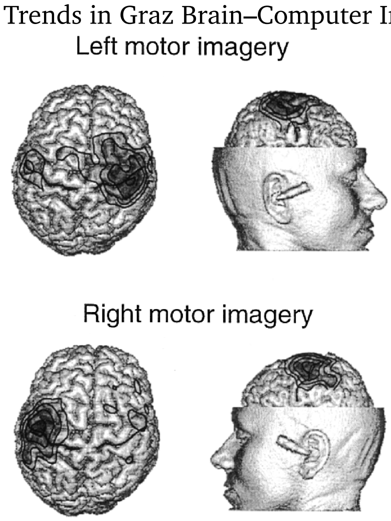
Figure 2: Schematic diagram showing the computation steps of ERD/ERS. The raw EEG data is first band-pass filtered in the frequency range of interest. Then the filtered signal is squared and averaged over each trial. On the left plot an ERD can be seen, that is recognizable by the decrease in the relative power. The right plot shows an increase in the band power, which is referred to as an ERS.

Pfurtscheller et al. - Current Trends in Graz Brain–Computer Interface (BCI) Research [20]



(a) Average power in the alpha band (here, 9–13 Hz; called the mu band over motor areas) during motor imagery based on EEG signals from the left (C3) and right sensorimotor cortex (C4). Positive and negative deflections, with respect to baseline (0.5 to 2.5 seconds), represent a band power increase (ERS) and decrease (ERD) respectively. The cue was presented at 3s for 1.25 seconds. [9]

Pfurtscheller et al. - Current Trends in Graz Brain–Computer Interface (BCI) Research [20]



(b) Distribution of ERD on the cortical surface calculated from a realistic head model, shown 625 ms after presentation of the cue [9]

Figure 3: Oscillatory EEG activity used in the Graz BCI.

References

[1] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," vol. 8, no. 4, pp. 441–446.

[2] Natasha Padfield, Jaime Zabalza, Huimin Zhao, Valentin Masero, and Jinchang Ren, "EEG-Based Brain-Computer Interfaces Using Motor-Imagery: Techniques and Challenges," vol. 19, no. 6, pp. 1423.

[3] Moritz Grosse-Wentrup and Martin Buss, "Multiclass Common Spatial Patterns and Information Theoretic Feature Extraction," vol. 55, no. 8, pp. 1991–2000.

[4] Keinosuke Fukunaga, *Introduction to Statistical Pattern Recognition*, Electrical Science. Academic Press.

[5] Z.J. Koles, "The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG," vol. 79, no. 6, pp. 440–447.

[6] Johannes Mueller-Gerking, Gert Pfurtscheller, and Henrik Flyvbjerg, "Designing optimal spatial filters for single-trial EEG classification in a movement task," p. 12.

[7] Yu Zhang, Guoxu Zhou, Jing Jin, Xingyu Wang, and Andrzej Cichocki, "Optimizing spatial patterns with sparse filter bands for motor-imagery based brain–computer interface," vol. 255, pp. 85–91.

[8] G. Pfurtscheller, Ch. Neuper, D. Flotzinger, and M. Pregenzer, "EEG-based discrimination between imagination of right and left hand movement," vol. 103, no. 6, pp. 642–651.

[9] Rajesh P. N. Rao, *Brain-Computer Interfacing: An Introduction*, Cambridge University Press.

[10] Jonathan R. Wolpaw and Elizabeth Winter Wolpaw, Eds., *Brain-Computer Interfaces: Principles and Practice*, Oxford University Press.

[11] "Berger: Über das elektroenkephalogramm des menschen - Google Scholar," .

[12] H. H. Jasper and H. L. Andrews, "Electroencephalography. III. Normal differentiation of occipital and precentral regions in man," vol. 39, pp. 96–115.

[13] Herbert Jasper and Wilder Penfield, "Electrocorticograms in man: Effect of voluntary movement upon the electrical activity of the precentral gyrus," vol. 183, no. 1-2, pp. 163–174.

[14] G. Pfurtscheller and A. Aranibar, "Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement," vol. 46, no. 2, pp. 138–146.

[15] G. Pfurtscheller, "Event-related synchronization (ERS): An electrophysiological correlate of cortical areas at rest," vol. 83, no. 1, pp. 62–69.

[16] Kai J. Miller, Gerwin Schalk, Eberhard E. Fetz, Marcel den Nijs, Jeffrey G. Ojemann, and Rajesh P. N. Rao, "Cortical activity during motor execution, motor imagery, and imagery-based online feedback," vol. 107, no. 9, pp. 4430–4435.

[17] Benjamin Blankertz, Ryota Tomioka, Steven Lemm, Motoaki Kawanabe, and Klaus-robert Muller, "Optimizing Spatial filters for Robust EEG Single-Trial Analysis," vol. 25, no. 1, pp. 41–56.

[18] J. Kalcher and G. Pfurtscheller, "Discrimination between phase-locked and non-phase-locked event-related EEG activity," vol. 94, no. 5, pp. 381–384.

[19] G. Pfurtscheller and F.H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: Basic principles," vol. 110, no. 11, pp. 1842–1857.

[20] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlogl, B. Obermaier, and M. Pregenzer, "Current trends in Graz brain-computer interface (BCI) research," vol. 8, no. 2, pp. 216–219.