Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 957–964
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Spatial filtering and selection of optimized components in four class motor imagery EEG data using independent components analysis

Clemens Brunner *, Muhammad Naeem, Robert Leeb, Bernhard Graimann, Gert Pfurtscheller

Laboratory of Brain–Computer Interfaces, Institute for Knowledge Discovery, Graz University of Technology, Krenngasse 37, 8010 Graz, Austria

Received 10 May 2006; received in revised form 15 December 2006
Available online 17 January 2007

Communicated by G. Borgefors

Abstract

Three independent components analysis (ICA) algorithms (Infomax, FastICA and SOBI) have been compared with other preprocessing methods in order to find out whether and to which extent spatial filtering of EEG data can improve single trial classification accuracy. As reference methods, common spatial patterns (CSP) (a supervised method, whereas all ICA algorithms are unsupervised), bipolar derivations and the original raw monopolar data were used. In addition to only performing ICA, the number of components was reduced with PCA before calculating a spatial filter for Infomax and FastICA.

The multichannel data (22 channels) of eight subjects, consisting of two sessions recorded on different days, was analyzed. The task was to perform motor imagery of the left hand, right hand, foot or tongue, respectively, during predefined time slices (cued paradigm). For a measure of fitness, classification accuracies for both cross-validated results using data from just one session as well as simulated online results (representing the session-to-session transfer) were calculated. In the latter case, the spatial filters and classifiers were computed for one session and applied to the completely unseen second session.

For the data analyzed in this study, Infomax outperformed the other two ICA variants by far, both in the cross-validated as well as in the simulated online case. CSP, on the other hand, yielded significantly lower classification accuracies than Infomax for the cross-validated results, whereas there is no statistically significant difference when it comes to simulated online data. Performing PCA before ICA improved the results in the case of FastICA, whereas the classification accuracies dropped significantly for Infomax.

Keywords: Spatial filtering; Independent components analysis (ICA); Common spatial patterns (CSP); Principal components analysis (PCA); Electroencephalogram (EEG); Brain–computer interface (BCI); Motor imagery

1. Introduction

Independent components analysis (ICA) is an unsupervised statistical method used for decomposing a complex mixture of signals into independent sources (Vigário et al., 2000). It is especially suitable for preprocessing multichannel electroencephalographic (EEG) data in brain–computer interface (BCI) research because it can remove a number of different artifacts such as electromyogram (EMG) or electrooculogram (EOG) signals (Jung et al., 2000a,b). It can also be used to separate different rhythmic EEG components, such as right- and left-hemispheric mu rhythms, from ongoing EEG (Makeig et al., 2004).

In this study, a feature selection algorithm automatically selected a small number of ICA components that are optimally suitable to differentiate between different brain states associated with four motor imagery tasks in a BCI experiment. The main goal behind this strategy was to improve the single trial EEG classification accuracy by using ICA

* Corresponding author. Tel.: +43 316 873 5315; fax: +43 316 873 5349.
E-mail address: clemens.brunner@tugraz.at (C. Brunner).

0167-8655/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2007.01.002

for spatial preprocessing and a subsequent feature selection algorithm for selecting the most relevant components. Here, bandpower features in a number of different bands between 8 Hz and 30 Hz were calculated from the preprocessed data. For comparison, the well-known spatial filtering method called common spatial patterns (CSP) (Koles, 1991; Müller-Gerking et al., 2000) was applied to the same data. In contrast to ICA, CSP is a supervised method that requires additional a priori information about the class of the data. As another reference method, bipolar derivations (which are simply differences between two monopolar EEG channels) were calculated. All preprocessing methods were compared with the results obtained from the original (monopolar) raw EEG data.

2. Subjects and experimental paradigm

In this study, the EEG data of eight subjects (three females and five males with a mean age of 23.8 years and a standard deviation of 2.5 years), recorded during a cue-based four class motor imagery task, was analyzed. Two sessions on different days were recorded for each subject, each session consisting of six runs separated by short (a couple of minutes) breaks. One run consisted of 48 trials (12 for each of the four possible classes), yielding a total of 288 trials per session.

The subjects were sitting in a comfortable armchair in front of a computer screen. As mentioned above, the paradigm consisted of four different tasks, namely the imagination of movement (motor imagery) of the left hand, right hand, foot, and tongue, respectively. At the beginning of each trial ($t = 0$ s), a fixation cross-appeared on the black screen. In addition, a short acoustic warning tone was presented at this time instant. After two seconds (at $t = 2$ s), a cue in the form of an arrow pointing either to the left, right, down or up (corresponding to one of the four classes left hand, right hand, foot or tongue) appeared for 1.25 s, prompting the subjects to perform the desired motor imagery task. No feedback (neither visual nor acoustic) was provided. The subjects were asked to carry out the mental imagination until the fixation cross-disappeared from the screen at $t = 6$ s. A short break followed, lasting at least 1.5 s. After that, the next trial started. The paradigm is illustrated in Fig. 1 (left).

Twenty two Ag/AgCl electrodes (with inter-electrode distances of 3.5 cm) were used to record the EEG, the setup is depicted in Fig. 1 (right). Monopolar derivations were used throughout all recordings, where the left mastoid served as reference and the right mastoid as ground. The signals were sampled at 250 Hz and bandpass-filtered between 0.5 Hz and 100 Hz. An additional 50 Hz notch filter was enabled to suppress line noise.

Although a visual inspection of the raw EEG data was performed by an expert, no trials were removed from the subsequent analysis in this study in order to evaluate the robustness and sensitivity to outliers and artifacts of each method. The fraction of artifacteous trials over all subjects was rather low anyway, namely 7.5% on average (median value of 6.1%).

3. Methods

3.1. Preprocessing

3.1.1. Spatial filters

The contamination of EEG signals with a variety of different artifacts such as EOG or EMG is an important issue in EEG data analysis. Appropriate precautions have to be taken in order to deal with this problem. Furthermore, the spatial resolution of EEG signals is compromised due to volume conduction through the scalp, skull and other layers of the brain. In the field of BCI research, these factors influence the classification accuracy of task-related activity. To address these problems, various spatial filtering techniques, for example common average reference (CAR), orthogonal source derivations, common spatial patterns (CSP), principle components analysis (PCA) and independent components analysis (ICA), can be utilized.

All these spatial filtering methods seek to solve the problems mentioned above by creating new components from the original data channels. In general, a spatial filter tries to estimate a so-called unmixing matrix $\boldsymbol{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n]$ such that the obtained components $\boldsymbol{y}(t) = [y_1(t), \ldots, y_n(t)]$ are as representative of the underlying sources as possible.

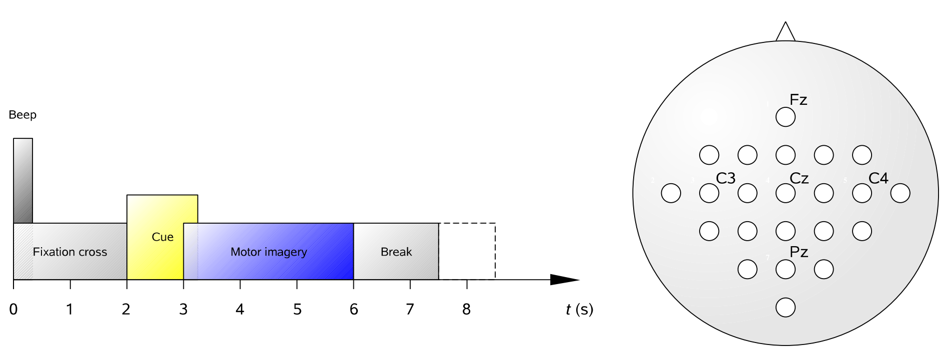

Fig. 1. Timing scheme of the BCI paradigm (left) and electrode setup of the 22 channels with inter-electrode distances of 3.5 cm. Some locations corresponding to the international 10–20 system are labeled (right).