Visuo-Spatial Attention Frame Recognition for Brain-Computer Interfaces *

Ferran Galán[1,2], Julie Palix[3], Ricardo Chavarriaga[1], Pierre W. Ferrez[1,4], Eileen Lew[1,4], Claude-Alain Hauert[3], and José del R. Millán[1,4]

[1] IDIAP Research Institute, Martigny, Switzerland
[2] University of Barcelona, Barcelona, Spain
[3] Lab. du Développement et des Apprentissages Moteurs, UniGe, Geneva, Switzerland
[4] Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Ferran.Galan@idiap.ch

Abstract. *Objective:* To assess the feasibility of recognizing visual spatial attention frames for Brain-computer interfaces (BCI) applications. *Methods:* EEG data was recorded with 64 electrodes from 2 subjects executing a visuo-spatial attention task indicating 2 target locations. Continuous Morlet wavelet coefficients were estimated on 18 frequency components and 16 preselected electrodes in trials of 600 ms. The spatial patterns of the 16 frequency components frames were simultaneously detected and classified (between the two targets). The classification accuracy was assessed using 20-fold cross-validation. *Results:* The maximum frames average classification accuracies are 80.64% and 87.31% for subject 1 and 2 respectively, both utilizing frequency components located in gamma band.

1 Introduction

Asynchronous EEG-based brain-computer interfaces (BCI) [1] allow subjects to control devices spontaneously and at their own pace, contrarily to synchronous BCI systems [2], and without requiring external cues such as in the case of relying on evoked potentials [3]. To this end, people learn how to voluntary modulate different oscillatory EEG rhythms by the execution of different mental tasks. A limitation of using mental tasks as control commands (e.g., imagining movements or doing arithmetics) is that subjects need to keep performing those mental tasks during the whole interaction, what can be exhausting, especially for novel users. An alternative is to exploit conscious behaviors that do not require sustained attention. Recent studies have demonstrated the possibility to modulate EEG alpha band by orienting visuo-spatial attention [4]. In an ideal case, BCI users could make a wheelchair *turn left* just by orienting their attention (without any eye movement) to some location in the left visual field, what is more natural than, for instance, imagining a left hand movement. Moreover, once the wheelchair just turn

* We thank Drs. S. Gonzalez and R. Grave, Geneva Univ. Hospital, for suggesting the use of visuo-spatial attention in BCI. This work was supported in part by the Swiss National Science Foundation through the NCCR 'IM2' and by the European IST Programme, Projects FP6-003758 and FP6-IST-027140. This paper only reflects the author's views and funding agencies are not liable for any use that may be made of the information contained herein.

left, users will simply stop attending to any particular spot of their visual field and the wheelchair, endowed with an intelligent controller [1], will move forward.

In this paper we assess the feasibility of recognizing user's voluntary modulation of EEG rhythms associated to visuo-spatial attention in an experimental setup close to the ecological conditions of asynchronous EEG-based BCIs. To this end, we compare both, a *traditional BCI approach* and a *frames approach*. These frames, as described by Freeman [5], correspond to active intermittent induced spatial patterns of amplitude modulation of beta-gama oscillations in response to conditioned stimuli. Based on those findings we address the following questions: *(i)* Does this discontinuous mode of function (i.e., frames) also appear in response to voluntary modulation of EEG rhythms? *(ii)* In this case, is it possible to classify these frames with respect to the attended location? *(iii)* Which frequency ranges yields better classification accuracy? *(iv)* Can this approach improve BCI performance? We hypothesize that *traditional approaches* (assuming sustained modulation of EEG rhythms over time) would face methodological problems: they will label (for training purposes) and classify samples extracted from periods of time where the underlying brain phenomena is either not present or is not salient enough. Then, a *frames approach* (which only classifies those samples where the induced episodic frames are detected) would be more appropriate. This paper addresses these questions and presents some hints for future work.

2 Methods

Data were recorded from 2 subjects with a portable Biosemi acquisition system using 64 channels sampled at 512Hz and high-pass filtered at 1Hz. The sampling rate was fixed at 512Hz to ensure a good estimation of the highest frequency component under analysis. The subjects were sitting in a chair looking at a fixation cross placed at the center of a monitor. The subjects were instructed to covertly attend to one of two possible target locations (lower-left and lower-right monitor's corners). The target location was specified by the operator in a pseudo-random balanced order. The subjects specified when they started to shift their attention. Each subject participated in 10 sessions composed by 4 trials each, 2 trials for each target. The duration of each trial was 7 seconds but only the first 600ms were utilized in this study.

The signal was spatially filtered using common average reference (CAR) previous to estimate the continuous Morlet wavelet coefficients on 18 frequency components (7, 8, 9, 10, 11, 12, 28, 32, 36, 40, 44, 48, 56, 64, 72, 80, 88, and 96 Hz) and 16 electrodes (F5, FC5, C5, CP5, P5, AFz, Fz, FCz, Cz, PCz, Pz, F6, FC6, C6, CP6, P6). The selection of electrodes was based on preliminary analysis of continuous Morlet wavelet coefficients scalp topography. Thus, each trial is composed by 512×0.6 samples and 18×16 features. The analysis carried on aims to compare the recognition rates over the different frequency components using two different approaches, namely the *traditional BCI approach* and the *frames approach*. The process was structured in two steps:

1. One canonical space was built per each frequency component (18 canonical spaces) [6] using 16-dimensional vectors (estimated wavelet coefficients at 16 electrodes). Since it is a 2-class problem, canonical spaces are defined by 1 canonical function.
2. A linear discriminant classifier (LDA) was built following two different approaches:

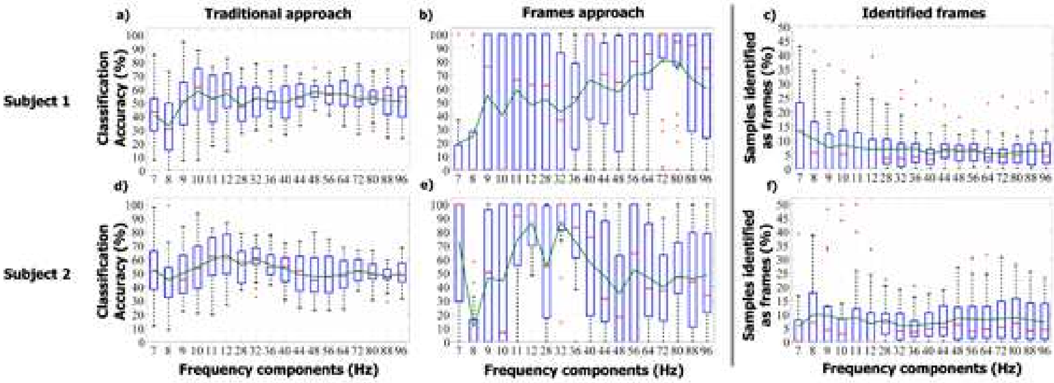

Fig. 1. Classification results using 20-fold crossvalidation over the 18 frequency components. Solid line represents the mean. LDA classification accuracy distributions utilizing traditional approach (*left*) and frames approach (*center*). *Right*, percentage of the total trial samples identified as frames.

(a) *Traditional BCI approach*: using all the training projected samples on the canonical space and classifying all the test projected samples.

(b) *Frames approach*: only a subset of the projected samples (i.e. frames) are used for training and classification. A sample was considered as a frame if its projection on the training canonical space was located on the opposite tails of each class distribution. Eight percentiles were utilized as thresholds: P_{40}, P_{35}, P_{30}, P_{25}, P_{20}, P_{15}, P_{10} and P_5. Thus, a sample was identified as a frame either its projection was below a given percentile (i.e: P_5) of class 1 or above the opposite percentile (i.e: P_{95}) of class 2. From now, the reference to one percentile also includes its opposite.

Both approaches were assessed using k-fold crossvalidation, $k = 20$. Each fold was integrated per one trial of each condition respecting the timing when they were recorded.

3 Results and Conclusions

The average LDA classification accuracy is higher utilizing the frames approach. For both subjects, the maximum classification accuracy is reached utilizing P_5. We report on detail the results obtained on this percentile. The maximum average classification accuracy classifying all the samples (i.e. traditional approach) is 58.41% at 10Hz and 63.08% at 12Hz for subject 1 and 2 respectively (see Figure 1 *left*), both in the alpha range. Utilizing frames approach, the maximum average classification accuracies are 80.64% at 72Hz, and 87.31% at 32Hz for subject 1 and 2 respectively (see Figure 1 *center*), both in the gamma range. It represents an absolute increase of 22.23% and 27.13% for subject 1 and 2 respectively. Notice that these classification accuracies are computed only on those samples identified as frames. The average percentage of samples identified as frames out of the total of samples of a trial is 5.85% for subject 1 and

5.92% for subject 2 (see Figure 1 *right*) at 72Hz and 32Hz respectively. In case of subject 1, only in 1 fold out of 20 it was not possible to identify any frame. In case of subject 2, it was not possible in 4 out of 20 folds. To understand the implication of these results in a real BCI application, each trial has been labelled according to the class maximum recognized by the classifier, using all the samples in case of traditional approach, and using only frames in case of frames approach. In the first case, the trial classification accuracies are 60.00% and 57.50% for subjects 1 and 2 respectively, what implies that channels capacities are .05 bits/second and .03 b/s (using estimator proposed in [1]). Using frames approach, the trial classification accuracies are 60.00% and 47.50%, but with only 12.50% of error recognition in both cases, what implies that channels capacities are .55 b/s and .46 b/s. Using frames approach the BCI theoretical channel capacity is boosted by 10.

This preliminary study on visuo-spatial attention frame recognition for BCI provides relevant hints for further research. First, it is possible to voluntary modulate EEG rhythms by orienting visuo-spatial attention in order to use asynchronous noninvasive EEG-based BCI's. Second, the intensity of this modulation is not sustained over time. This fact can be related to the active intermittent induced spatial patterns of amplitude modulation (frames) in response to conditioned stimuli described by Freeman [5]. In this case these patterns are voluntary driven by the subject. Third, it is possible to classify the frames generated by orienting the attention to different visual locations with high classification accuracies (above 80%). Fourth, these classification accuracies are maximum in gamma band ($>$ 30Hz), corresponding to endogenous shifts of attention effects [7]. Fifth, classification accuracies utilizing a traditional approach, i.e. assuming modulations sustained over time, are around the chance level. It suggests that this approach is not optimal to recognize induced EEG phenomena, what is confirmed comparing the BCI theoretical channel capacity achieved using both approaches. Using frames approach the BCI theoretical channel capacity is drastically increased.

References

[1] Millán, J. del R., Renkens, F., Mouriño, J., Gerstner, W.: Noninvasive brain-actuated control of a mobile robot by human EEG. IEEE Trans. Biomed. Eng. **51** (2004) 1026–1033.

[2] Birbaumer, N., Ghanayim, N., Hinterberger, T., Iversen, I., Kotchoubey, B., Kübler, A., Perelmouter, J., Taub, E., Flor, H.: A spelling device for the paralysed. Nature **398** (1999) 297–298.

[3] Kelly, S. P., Lalor, E. C., Reilly, R. B., Foxe, J. J.: Visual spatial attention tracking using high-density SSVEP data for independent brain-computer communication. IEEE Trans. Neural Sys. Rehab. Eng. **13** (2005) 172–178.

[4] Thut, G., Nietzel, A., Brandt, S., Pascual-Leone, A.: Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. J. Neurosci. **26** (2006) 9494–9502.

[5] Freeman, W. J.: Origin, structure, and role of background EEG activity. Part 3. Neural frame classification. Clin. Neurophysiol. **116** (2005) 1118–1129.

[6] Galán, F., Ferrez, P. W., Oliva, F., Guàrdia, J., Millán, J. del R.: Feature extraction for multiclass BCI using canonical variates analysis. IEEE Int. Symp. Intell. Signal Process. (2007).

[7] Palix, J., Hauert, C. A., Leonards, U.: Brain oscillations: Indicators for serial processing in inefficient visual search? Perception **35** (2006) 234.